

普通话单字调的归一化模型(进展报告)

杨 顺 安

一、前 言

为了合成出自然的普通话音节以及词和句,就必须先考察普通话单音节的音高变化规律,即字调模式或调型。现在通行的五度值表示法,能简明地表征一个音节的相对调型,对一般语音的描写以及教学来说,是有实用意义的,但对语音合成来说,似乎粗略了些。因此,还要进一步分析单音节的调型规律,总结出一组便于合成应用的调型曲线和表达式。

很早以来,国内外许多研究者用较近代化的测试工具已对普通话字调进行过一些研究。60年代初,林茂灿^①用音高显出器测量了294个单字的音高变化,提出了有区别意义的调型段概念;70年代,蒋秋广等^②分析了三千00个音节的基频数据,得出了四声的平均基频值等参数,并以平均基频值为基准进行归一处理,得出了四声的调型曲线。

79年,宋知用等^③在计算声门波时,用了一种非归一的四声多项式表达式;最近,李子殷^④用了一组归一化的四声基频值作为合成参数,提高了合成自然度。这些研究对字调处理都有益的。

我们现在从合成应用出发，对四种声调的基频数据进行了分析，以同一^准基频率和调域为基准，进行归一化，得出调型模式及多项式表达式，并能将这种调型表示法很方便地转换到五度值表示法上去。

二、测量和数据处理

我们用语图测算了一位女发音人发的阴、阳、上、去各10个音节的基频数据，然后进行时间域和频率域的归一化处理。关于归一化的处理方法在另外一篇文章^⑤中，有较详细的说明，这里只是频域归一化的公式和基准值与前文有些不同。

图1是经过时间归一后的四种声调的基频 F_0 的平均值 \bar{F}_0 ($i = 1, 2, 3, 4$)，其中频率轴(左纵轴)用对数刻度。可以看到，每一种声调的起始段或收尾段都有一定的“弯头”和“降尾”^①。

图1中，最大频率是297.8赫，位于去声的起始段，最小频率是109.6赫，位于去声的终止点。在计算调域时，一般不取去声终点值为下限值，而取上声凹点频率(这种处理法是文献〔1〕用过的，也是吴宗济先生提议的)。在此，上声凹点频率为146.0赫，即调域上限值 $F_H = 297.8$ 赫，下限值 $F_L = 146.0$ ，调域 $F_d = F_H - F_L = 151.8$ 赫。

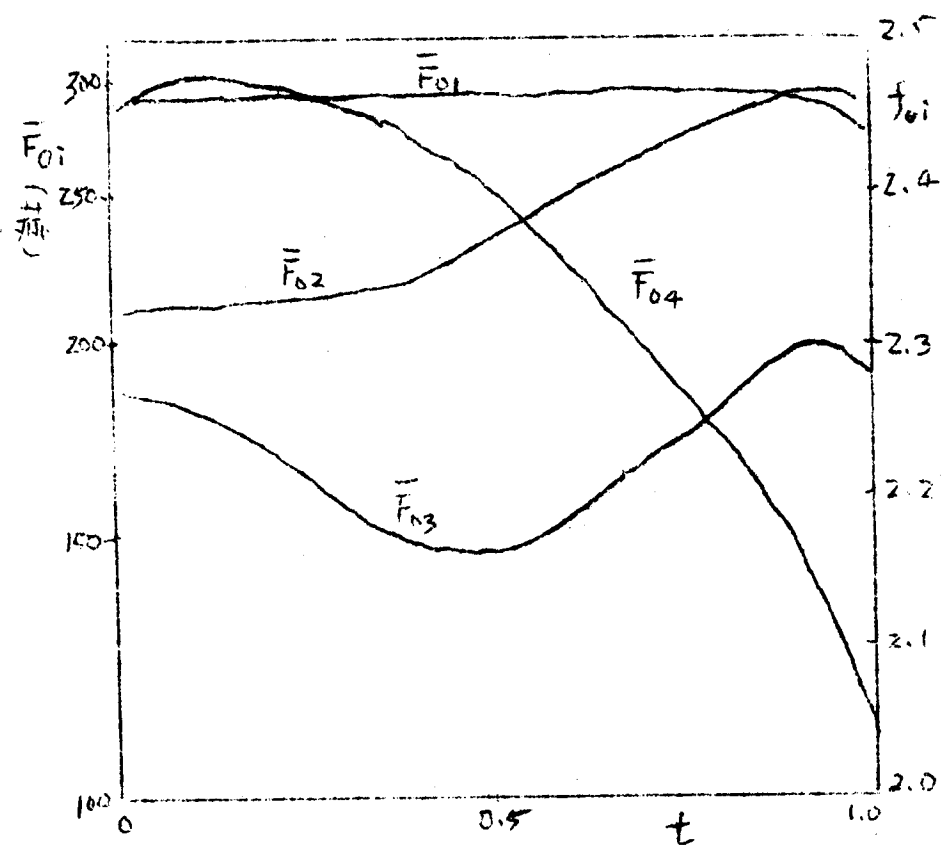


图1 四声的时归—平均曲线

上、下限值之比为 2.04，近似为 2，即二者相差约为一个倍频程。

如果以 f 表示 F 的十进对数值， f_d 、 f_H 和 f_L 分别相应于调域及其上、下限值，并取 f_H 和 f_L 的中间值为中值 f_c ，则对四种声调归一化的公式为

$$\left. \begin{aligned} f_i(t) &= (\bar{F}_{0i}(t) - f_c) / f_d \\ f_d &= f_H - f_L \quad (i=1, 2, 3, 4) \\ f_c &= f_L + 0.5f_d \end{aligned} \right\} (1)$$

这样一来，四种声调都统一地用两个基准值，中值 f_c 和调域 f_d 来归一。

通过简单的函数变换，可将上述归一调值 $f_i(t)$ 转换为五度制调值 $f'_i(t)$

$$f'_i(t) = 4f_i(t) + 3 \quad (2)$$

图 2 是归一化处理后的调型曲线，左纵轴是归一调值，右纵轴是五度制调值。

利用最小二乘法，对 $f_i(t)$ 的数据进行多项式拟合，当阶次为 4 时，可得如下表达式：

$$\left. \begin{aligned} f_1(t) &= 0.453 + 0.295t - 1.456t^2 + 2.574t^3 - 1.468t^4 \\ f_2(t) &= 0.011 + 0.160t - 0.913t^2 + 1.375t^3 - 2.560t^4 \\ f_3(t) &= -0.115 + 0.246t - 1.785t^2 + 16.36t^3 - 8.73t^4 \\ f_4(t) &= 0.44 + 1.205t - 5.58t^2 + 6.44t^3 - 3.39t^4 \end{aligned} \right\} (3)$$

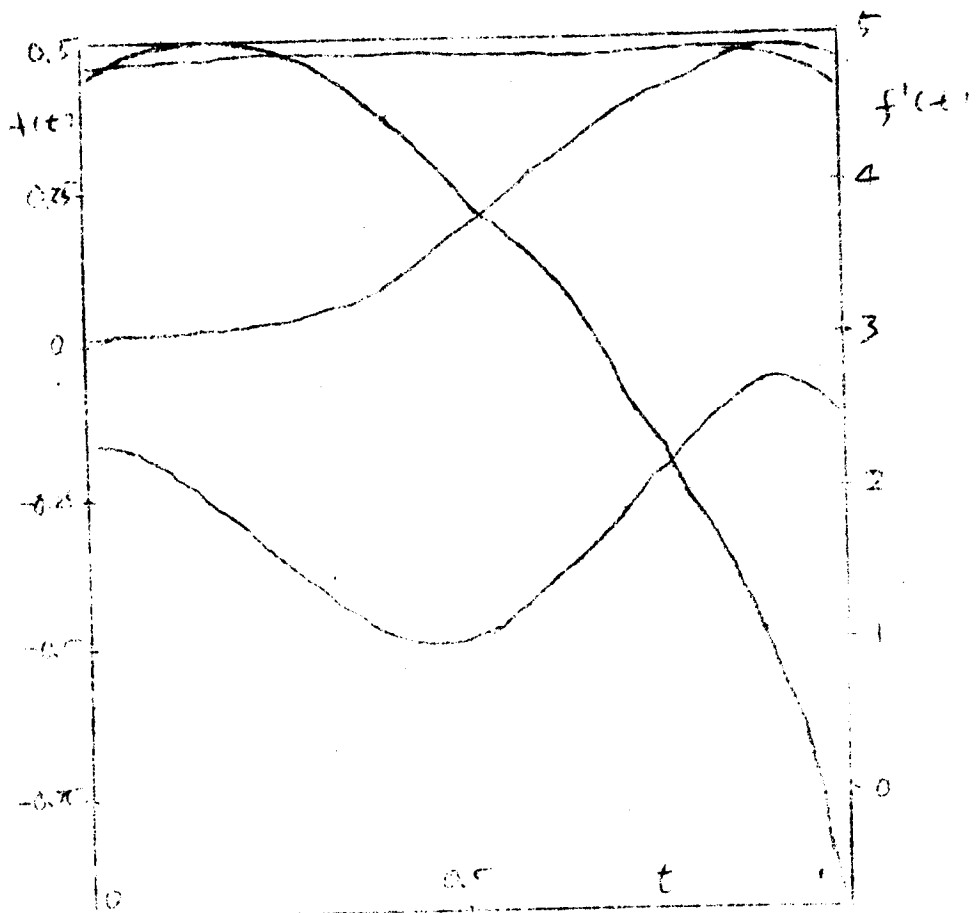


图2 四种声调的归一调型曲线

图3中圆点代表原 $f_k(t)$ 的函数数值。细线是根据式(3)画出的拟合值。由图可见，此拟合已足够准确了。

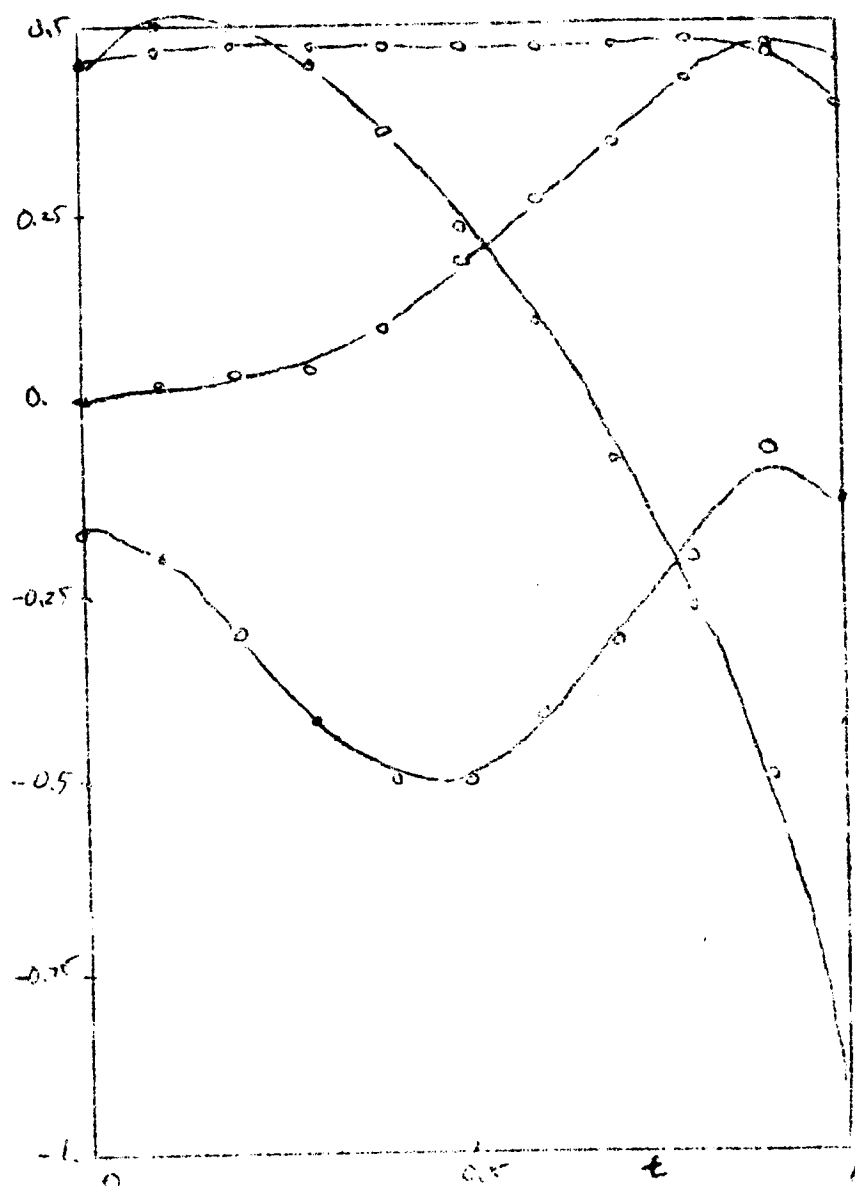


图3 四声调型的多项式拟合曲线

三、 讨 论

根据上述归一化公式(1)和(8)，我们可以得到字调的归一化表示或归一模型。利用(1)式，可得基频值 F_{0i}

$$F_{0i}(t) = \log^{-1} [f_c + f_d \bar{f}_{0i}(t)] \quad (4)$$

式中 \log^{-1} 表示十进对数的逆运算，即以 10 为底的幂。参照藤崎等^⑤提出的日语 F_0 产生模型的形式，我们得到如图 4 所示的普通话字调产生模型。虽然，字调归一表达式(8)是根据一个人的数据得到的，但是图 4 的产生模型却有普遍性。因为，字调的变化是相对的^⑦，因此，模型中值频率 f_c 、调域 f_d 和归一模式 $\bar{f}_{0i}(t)$ 都可在一定范围内因人而异。改变 f_c ，可以产生不同性别、不同年龄的 F_0 。改变 f_d ，可以模拟不同讲话速率的 F_0 变化：一般情况下，调域 $f_d = \log 2 = 0.301$ ，即在一个倍频程内变化。如果是强调语气或慢速情况， f_d 要增大一些，反之，在快速发音时， f_d 可小一些。就归一模式 $\bar{f}_{0i}(t)$ 而言，变化肯定会有，但

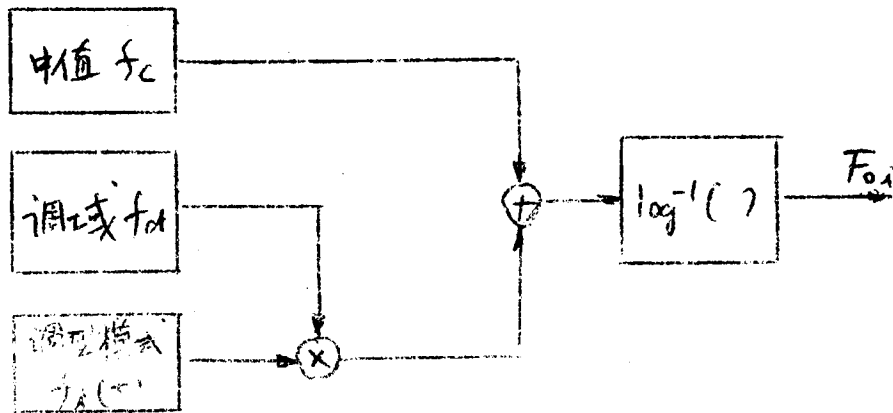


图 4 字调产生模型

基本趋势是一致的。

这种字调的产生模型还可以推广应用到词组和句中。只是要按一定的规律改变 f_c 和 f_d ，并加上音节间的过度规律。例如，在二字组中，一般后字是正常重音，音长较长而调型完整^③，此时，可为后字设定一个比前字大的 f_d ；在问句或强调句中，句尾或主要词句中基频升高，可设定一种渐升的 f_c ；在加速语句中，基频变化范围缩小^⑨，可设定一较小的 f_d 。

当然，这种归一化字调产生模型还有待合成试验来验证和进一步完善。

本文原始数据是由颜景助同志提供的，在此表示感谢。

参考文献

- ① 林茂灿，“音高显示器与普通话声调音高特征”，《声学学报》，2，1，8—15（1965）。
- ② 萩秋广等，“标准中国语の单音节语の四声の音响特征”，《日本音响学会志》，31，6，369（1975）。
- ③ 宋知用，张家禄，“声门波平均频谱的计算”，《声学学报》，2，134—149（1979）。
- ④ 李子殷，“关于按规则合成汉语连续语言及提高其自然度的方法”，《语言与通信信号处理学术会议论文集》，p. 8（1982）。
- ⑤ 杨顺安，曹剑芬，“普通话二合元音的动态特性”，《语音实验年报》，1982。
- ⑥ 藤崎博也等，“日本語单语アクセントの基本周波数パターンの生成机构のモデル”《日本音响学会志》，27，9，445—453（1971）。
- ⑦ 赵元任，《语言问题》，p. 66，（商务1980版）。
- ⑧ 林茂灿，颜景助，孙国华，“二字组正常重音声学特性”，《语音实验年报》，1982。
- ⑨ 吴宗济，“普通话语句中的声调变化”，《中国语文》，82年，6期，439—449。

1977—1981年刊出稿和学术报告

任宏谟，“北京话塞音的研究”，中国社会科学院研究生院

1981届研究生毕业论文简介，p. 285（1982版）。

吴宗济，“汉语普通话辅音声学特性的初步分析”，数字通信终端技术交流会，77，1月，武汉。

吴宗济，杨力立，曹剑芬，“实验语音学知识讲话”，中国语文，79年，1，2，4，5，6期。

吴宗济，“A preliminary study to distinctive features and their correlations in standard Chinese”，Proc. 9th. Int. Cong. Phon. Sci., Copenhagen (1979)。

吴宗济，曹剑芬，“普通话辅音声学特征的几个问题”，第二届声学学术会议，79年，北京。

吴宗济，“试论普通话语音的‘区别特征’及其相互关系”，中国语文，80年，5期。

吴宗济，“什么叫‘区别特征’”，国外语言学，80年，1期。

吴宗济，“实验语音学十讲”，北京大学讲义，80—81年，（整理中。）

吴宗济，杨力立，“普通话‘语调’规则初探”，语言通讯和语言信号处理学术交流会，81，1月，北京。

吴宗济, “实验语音学与语言学”, 语文研究, 81年, 1期。

吴宗济, “普通话语调的实验研究”, 第一届中国语言学会年会,

81, 10月, 成都。

周殿福, 《艺术语言发声基础》, 中国社会科学出版社, 80,

5月第一版。

周殿福, “关于广播语言的吐字发声的问题”, 全国播音工作经

验交流会, 81年, 8月。

周殿福, “语音实验和听觉”, 语文教学与研究, 81年, 2期。

林茂灿, “现代合成技术与语音特性”, 数字通信终端技术交流

会, 77年1月, 武汉。

林茂灿, 林联合, 夏光荣, 曹雨生, “A study of tone-

sandhi in standard Chinese with computer”,

Proc. 9th. Int. Cong. Phon. Sci., Copenhagen,

1979.

林茂灿, 林联合, 夏光荣, 曹雨生, “普通话二字词变调的实验

研究”, 中国语文, 80年, 1期。

林茂灿, 颜景助, “北京话轻声的声学性质”, 方言, 80年,

2期。

林茂灿, “语音研究与现代科学技术”, 语文现代化, 80年,

4期。

林茂灿，颜景助，“北京话轻声的声学性质”，语言通讯和语言信号处理学术交流会”，81，1月，北京。

杨顺安，“声带波的提取方法及其应用（综述）”，语言通讯和语言信号处理学术交流会，81，1月，北京。

杨顺安，“用数字信号处理技术提取声带波”，中国社会科学院研究生院1981届毕业论文简介，p. 286（1982年版）。

曹剑芬，“怎样发普通话的zh、ch、sh、r”，浙江师范学院学报，81年，3期。

曹剑芬，“沙洲县常阴沙话数词‘十’的读音”，中国语文，81年，4期。

鲍怀翘，“普通话主要元音的声学特征”，数字通信终端技术交流会，77年1月，武汉。

鲍怀翘，林祥荣，“骨导语言清晰度测量”，第二届声学学术会议，79年，北京；电声技术，79年，3期。

鲍怀翘，林祥荣，“骨导语言频谱分析”，第二届声学学术会议，79年，北京。

鲍怀翘，“利用FFT进行元音合成”，语言通讯和语言信号处理学术交流会，81年1月，北京。

鲍怀翘，“语音研究的新发展”，吉林大学学报（社），80年，1期。

鲍怀翘：“利用FFT进行语言频谱分析”，北京仪器仪表学会处
理学术年会。81年，1月。