

普通话塞音爆破段谱的特性分析

阳晶^{1,2} 陈肖霞²

¹中国社会科学院研究生院

²中国社会科学院语言研究所语音室

摘要

人们对塞音发音部位在声学上的判定提出过很多理论, 其中 Stevens 和 Blumstein 提出的声学不变量是基于塞音爆破段自身的声学特性来研究其发音部位的, 他们的研究证明塞音自身的声学表现与生理发音之间存在对应关系。本文对汉语普通话双唇、齿龈和软腭三类塞音在不同元音环境下, 在送气不送气两种条件下爆破段自身谱的特性进行分析, 并运用 SPSS 统计检验对样本进行验证, 得出三类塞音不同声学模式及相互之间的关系。

1. 前言

通常来说, 在连续的语流当中音素与音素是线性连接在一起的, 相邻两者之间会互相影响, 不管是辅音与辅音相接, 元音与元音相接还是辅音与元音相接。把这些语流当中的音素单切出来与这些音素单独发音的情况进行比较我们会发现, 两者差别很大, 但是这并不影响我们对语流中音素的正确判断, 除了语境对我们的提示之外, 还有很重要的一点是每个音素总是包含着表示自身特性的线索。元音的特征体现在共振峰的模式上, 而共振峰的模式与元音发音时舌位的高低、前后、平展与否有着对应的关系, 即声学特征与生理发音上有对应关系。辅音也同样如此, 辅音的不同发音方法在声学上表现为不同的语图模式, 塞音是冲直条, 擦音是摩擦乱纹, 塞擦音是冲直条加乱纹, 鼻音和边音声带振动, 都有浊音横杠。

判定发音部位时, 塞音冲直条时长很短, 因此很难确定。早在 1957 年, M.Halle, G.W.Hughes 和 Redley 就提出, 当塞音与元音相邻时, 通常有三种判定线索: 静音段, 爆破段和过渡段。在对塞音的爆破段进行谱的特性分析时, 他们通过实验, 取爆破段开始的 20ms 进行谱分析, 得出如下结论: 谱本身包含辨别塞音的信息, b 在低频部分有能量集中区域, d 的谱趋向平滑或者在高频部分能量突出, 在 500Hz 区域也有能量集中, g 在 1500—4000Hz 有强的能量集中。

另外, 塞辅音与元音相接时, 两者之间的过渡也是

判定不同塞音的重要因素。Delattre, Liberman, Borst 和 Gerstman 1952 年提出相邻元音的 F2 过渡可以作为判定浊塞音部位的线索, 后来 Delattre, Liberman 和 Cooper 1955 通过实验发现相邻元音的 F2 过渡还可以判定清塞音和与塞音对应部位的鼻音。元音的 F2 过渡很直接的表示了从塞音发音部位到后接元音目标位置的发音动程。同时他们通过实验还发现在第二共振峰的特殊位置可以清楚的辨别不同塞音, 这一特殊位置就是通常所说的音轨, 邻接元音 F2 指向的音轨不受不同元音色彩的影响, 而 F1 对塞音辨别作用不大。这就是音轨理论, 即塞音与不同元音相接时元音过渡段指向表示塞音特性的一点, 即音轨。不同部位的塞音有不同的音轨频率, 双唇塞音的音轨频率是 700Hz, 齿龈塞音的音轨频率是 1800Hz, 软腭塞音的音轨频率一般有两个, 1700Hz 和 3200Hz。

Stevens 和 Blumstein 在 1978 年将前人的关于根据爆破段谱的特性区分不同部位塞音的研究更进一步, 提出声学不变量理论, 从塞音爆破段本身寻找线索来判定其发音部位的特性。他们认为塞音的发音部位和声学表现上有一定的对应关系, 即塞音的爆破段存在一定的声学不变性。声学不变性理论的提出进一步证明了声学模式与生理上发音部位的对应关系。

2. 实验方法

汉语普通话塞音的发音部位也有三类: 双唇、龈、腭, 与英语不同的是, 汉语普通话塞音有送气不送气的区别, 没有清浊的区别, 我们希望研究汉语普通话塞音声母在不同元音环境下谱的特性, 因此, 对汉语单音节语料库的 m01、m02、m08、m10、m12 的在单元音 a、i、u 环境下谱的特性进行分析。我们对这五个人的单元音条件下塞音声母的音节进行音素标音, 对声母部分标注出所包含的声学段, 即爆破段、擦段、气段, 在标注的基础上选取爆破段, 进行 10000Hz 的重新采样, 然后进行 FFT 分析, 得到二维谱, 再进行 14 个极点的 LPC 平滑, 每个选取的爆破段得到一条曲线。我们运行 Praat 脚本将数据提取出来, 因为个人不同、音的不同和声调的不同, 每个音节塞音声母的爆破段时长不同, 提取出的数

据有的按 78.125Hz 来提取点数,有的按 39.0625Hz 取点。因此,我们按照 5000Hz 取 65 个点即以 78.125Hz 为间隔取点的标准,进行初步归一。由于汉语音节的拼合规则,g、k 不能与 i 相拼,有的音节并不是四声俱备,同时,在分析过程中,还有些音爆破段每出现,或者发音不好的,我们予以排除,最后得到每个人 b、d、g、p、t、k 6 个塞音声母在 a、i、u 三个元音环境下四个声调的约 400 个音节的塞音爆破段的二维谱。x 轴为频率 Hz 值, y 轴为能量 dB 值。

3. 实验结果与分析

3.1. 塞音爆破段谱的分析

3.1.1. 不同塞音之间谱的模式分析

根据 Stevens 等人的实验,塞音的三个发音部位对应三个不同的二维谱的模式,在汉语普通话中,m01 发音人 ba、da、ga、的谱图:

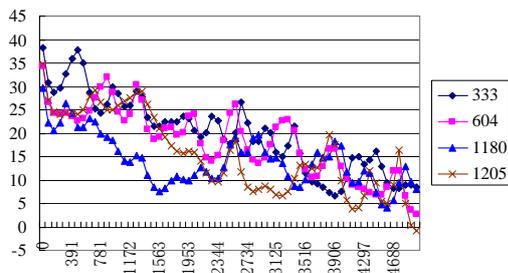


图 1 ba

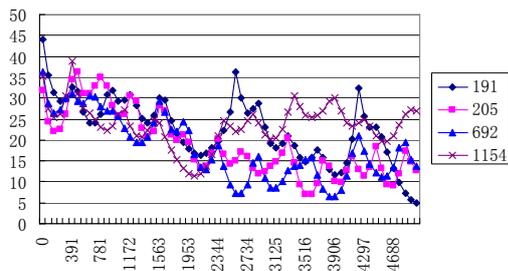


图 2 da

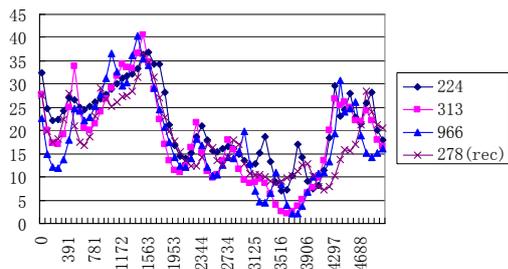


图 3 ga

以上谱图中,每个图包含 4 条曲线,分别表示音节的四个声调,示例的数字比如“191”、“205”表示该音节在语料库字表中的位置,后有“rec”字样的表示经过重新取点初步归一的。由于每个音节每个声调只有一个音,样本太少,无法分声调进行统计分析,同时考虑到声调主要负载在韵母部分,因此对塞音声母部分爆破段的分析没有考虑声调的区别。从以上谱图可以看出,双唇、齿龈、软腭塞音各有不同模式,b 是下降的,斜率很大,d 在 2500—4000Hz 范围内会有峰的突出,有的音突出不明显,中高频部分呈现渐升的趋势,总体呈下降趋势,但下降斜率没有 b 大,g 在频率范围的低频和高频部分有两个很明显的突出的峰,而在中频部分的峰并不突出。送气音与对应的不送气音谱的模式也是极为相似的,排除峰值位置的细节不看,p 与 b 都是下降的模式,d 与 t 都是中部略微突出,下降斜率小于 b、p, g 与 k 与 b、d 两类明显不同,在中频和高频部分有明显突起的峰且前一个峰比后一个峰更为突出,其余峰则很低。从谱上我们就可以将 g 类与其他两类区分开来。

在 Stevens 等人的实验中,认为不同部位一—双唇、齿龈、软腭的塞音爆破段的谱有其固有的特性和模式,他们通过在谱图上作两条参考线来确定各个模式的具体特性,同时不考虑 700Hz 以下的情况。双唇塞音是分散下降或略平的模式,在 1200—3500Hz 的范围内,至少有一个突出的峰比其他峰都要高,2400Hz 以下以及 2400—3600Hz 范围内至少有两个峰落在参考线区域内。齿龈塞音是分散上升的模式,参考线内至少有两个峰存在,而且在 2200Hz 以上必须有一个峰高于 2200Hz 以下的峰,他们同时提到,在齿龈塞音谱图中,可能在 800—1600Hz 之间会出现符合模式之外的峰,这是因为声门下共鸣,另外,在 1800Hz 左右会出现另外的峰是因为音轨所致。所谓分散就是谱中峰值相隔至少 500Hz,由此双唇与齿龈塞音与软腭塞音区别开来,因为软腭塞音是集中的模式。在 1200—3500Hz 频率范围内有一系列的峰重叠在一起,并且中频部分必须有一个峰是最突出的,高于其他频率位置的峰。

从我们分析所作的谱图来看,它们与 Stevens 文中所描述的模式是基本一致的,但是 g 类跟他们的描述有不同,g、k 与 u 元音相接时,最高的峰不是在 1500—3500Hz 范围内,而是在 700Hz 左右。3.1.2 同一塞音在不同元音条件下谱的特性

由于协同发音的影响,音节中相邻音素之间会互相

影响，普通话音节中塞音声母在不同元音之前也会受到元音音素的影响，下面是 m01 的 da、di、du 三组音的谱图（见图 7—图 8，da 见上图）：

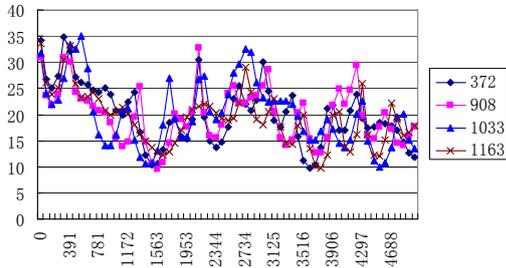


图 4 di

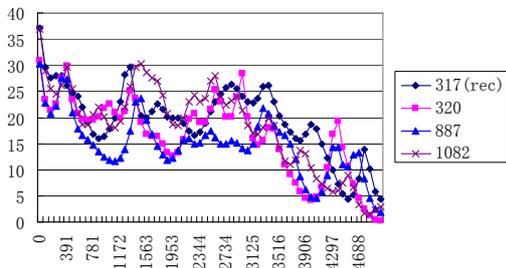


图 5 du

从图中可以看出，这三幅图细节上各有不同，但是从总体上看，都是在 2500—4000Hz 之间有突出的峰，强度比低频范围的峰还要大，而总体上有下降的趋势。

3.2. 爆破段二维谱峰值的分析

3.2.1 谱的特征峰点

为了更进一步的得到塞音爆破段谱的特性，我们从上一步得到的五个人的塞音爆破段谱图中，对每张图手工选取 7 个峰值，取值的标准是看每个人每个音四个声调之间的一致性，将得到的 7 个峰值数据进行散点作图，得到 16 张峰值分布图。由于是在谱图的基础上取点，我们对 g、k 两类音根据谱图只对突出的两个峰进行取点。

下面是 ba、da、ga 的峰值分布图：

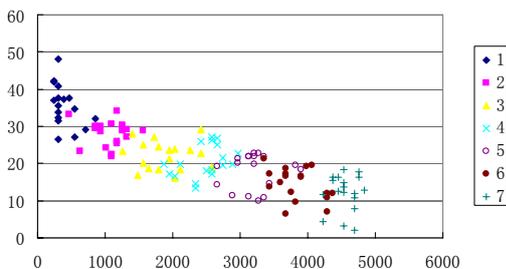


图 6 ba

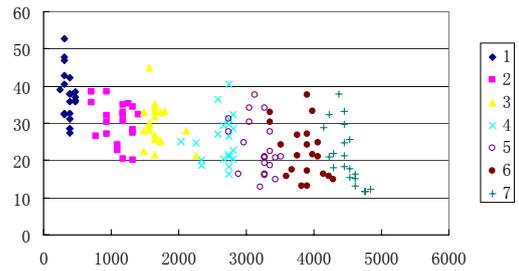


图 7 da

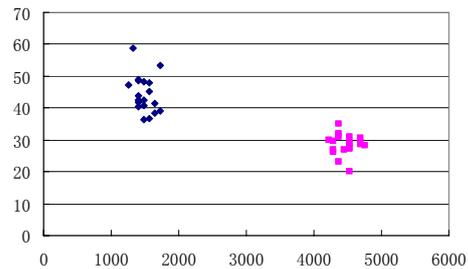


图 8 ga

我们在谱的分析一步中认为 g 与 b、d 两类有明显的不同，因此取点的时候从谱的特性出发只取两点。对于 b 与 d，我们用 SPSS 对 ba、da 两组音的 7 个峰点分别进行频率与能量值的非参数检验，得到的频率值结果为（见表 1）：

variable2 到 variable8 表示 7 个峰点，从频率值的统计结果看，这两组音在频率值上 7 个点都大于 0.05 说明不存在显著性差异。也就是说在频率域上，b、d 两类没有显著性差异。

能量值的统计结果为（见表 2）：从能量值的统计结果看，有四个点（第 3 点，第 4 点，第 6 点，第 7 点）值都明显小于 0.05，存在明显差异，这些有差异性的点处在中频高频位置。而第一点对与所有音都存在一致性，这有可能是受基频的影响，这说明 ba、da 两个音的差别是在体现在峰点的能量上，而不是频率值上。

3.2.2 后接元音对谱的特性的影响

我们知道相邻音素之间会互相影响，塞音声母后的元音音素对塞音爆破段的影响是不是会改变塞音发音部位的声学模式？下面是 ga、gu 的谱图和峰值图：

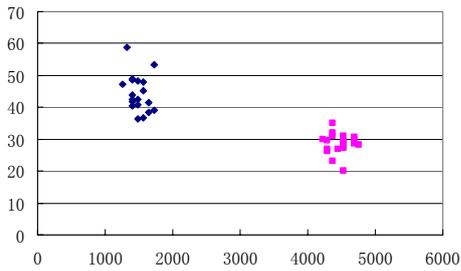


图 9 ga

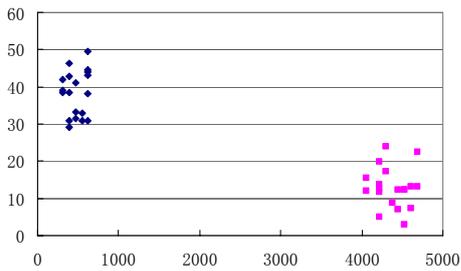


图 10 gu

可以看到，在 ga 音中，前一个突出的峰分布在 1400Hz 左右，而 gu 音中，前一个突出的峰分布在 500 左右，显然，这是受了元音的影响，a 音的共振峰 F2 就在 1300 左右，u 的 F1 一般在 400 左右，F2 在 800 左右，因此，即使只是爆破段，后接元音也会对它产生一定影响，这是不可避免的，因为人们在发音过程中，虽然是一个一个音素线性发出来的，但是前后相邻的音素发音时存在耦合作用，塞音声母发音时部位往往向后接元音靠拢。我们对 ga, gu 两组峰值的频率值和 dB 值进行非参数检验，频率值结果为（见表 3）：

从结果可以看出，第一个峰值点统计结果为 0.000，说明两者之间有明显的差异，而在第二点上不存在差异，这与我们观察到的结果是一样的。对能量值的非参数检验结果显示在能量值上，ga 与 gu 也存在明显差异，这两个结果说明后接元音对前面声母的爆破段是有影响的。虽然受元音影响，但 g、k 的谱的声学模式并没有改变。都是两个突出的峰，而且前一个中低频范围的峰更突出。并没有因为后接不同元音而多出现或减少突出的峰值。对于双唇和齿龈塞音，我们在提取出的峰值数据基础上计算每一点的平均值和标准差，得到每一个音七个峰点的平均值，并将每个点平均值做出折线图，得到峰点的平均值曲线图，下面是 ba、bi、bu 的峰点平均值曲线：

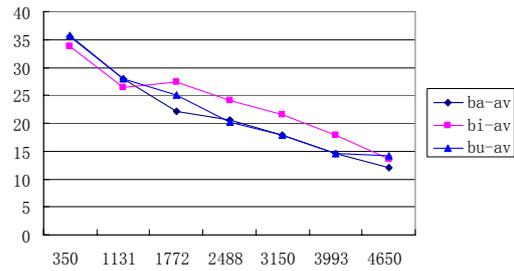


图 11 ba、bi、bu

与双唇塞音 b 相对的齿龈塞音的峰点平均值曲线为：

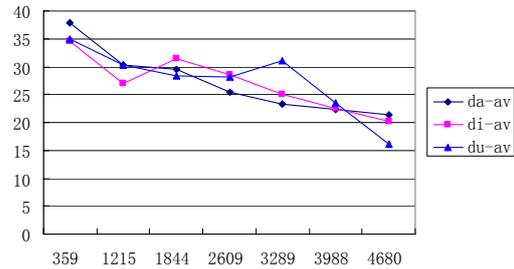


图 12 da、di、du

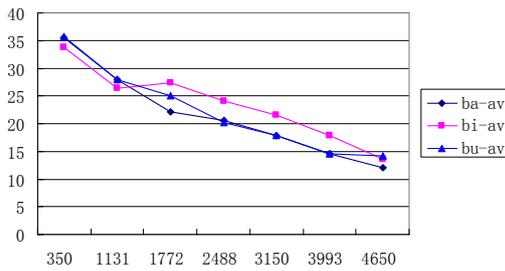
从这两张峰点平均值曲线图可以看出，后接元音的不同会对前面的塞音声母产生影响，但是，这种影响不足以改变不同发音部位的塞音之间谱的模式的区别，d 组的曲线下降斜率要明显小于 b 组，而且在中频部分都会有峰突出，高于其他频率位置的峰。这种与发音部位对应的谱的模式的区别是不受元音音素的影响的。

我们同样对 ba、bi、bu 的 7 个峰点分别进行频率值与能量值的非参数检验，得到的频率值结果见表 4（略），从统计结果看，ba、bi、bu 在频率值上不存在显著差异。

得到的能量值的结果为（见表 5 略），从统计结果看，在能量值上，这三个音也不存在明显差异，这说明，b 在三个元音之后存在明显的一致性，换句话说，不同元音对塞音声母爆破段没有明显影响。但是不同塞音的具体情况有所不同，从上面 ga、gu 的统计结果看，g 类音受元音的影响就比 b、d 两类要明显。

3.2.3 塞音送气不送气的对立对谱的特性的影响

我们知道。汉语普通话没有清浊的对立，有送气与不送气的对立，送气段一般表现为爆破段之后有很长一段送气过程，在语图上显示为乱纹，但是与摩擦乱纹不同的是，送气段有明显的能量集中区，而且受后接元音影响，有明显的过渡，普通话送气与不送气塞音的爆破段特性是怎样的呢？下面是 ba、bi、bu 的峰值平均值曲线与 pa、pi、pu 的峰值平均值曲线比较。



13 ba、bi、bu

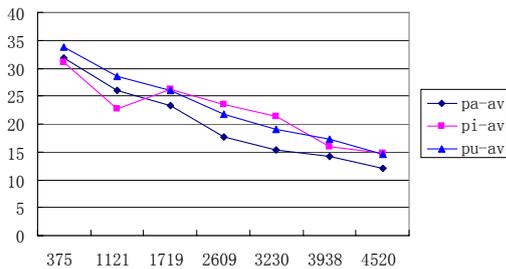


图 14 pa、pi、pu

从这五个发音人峰值平均数据的曲线可以看出，不送气的 ba、bi、bu 与送气的 pa、pi、pu 峰值平均值曲线的一致性很强。

同样我们对 ba、pa、da、ta、ga、ka 峰值的频率值和能量值进行非参数检验，得到的结果显示，送气与不送气两组在频率值和能量值上都没有显著差异。说明送气与否对不同部位塞音的谱的特性没有明显影响。

4. 结论与讨论

通过本文所做的实验研究，我们发现普通话不同发音部位塞音的爆破段存在相对稳定的声学模式，这些模式不受后接元音音色与送气不送气的影响，呈现出自身的一致性。双唇塞音谱的峰值分布很广，能量渐降，斜率很大；齿龈塞音峰值与双唇塞音在峰点的频率值上不存在显著差异，两者的不同主要体现在能量值上，齿龈

The characteristic analysis for stop burst spectrum in Standard Chinese

Yang Jing & Chen Xiaoxia

Department of Linguistics, the Graduate School of CASS Institute of Linguistics, CASS

ABSTRACT: There are many theories about the acoustic determination for stop's articulation place. Among these

塞音中高频部分的峰能量比低频范围的峰要强，因此在中高频部分，齿龈塞音表现为上升的模式，虽然总体还是下降的。软腭塞音峰值在整个频率范围内只有两个突出的峰，其中前一个最突出，虽然随着后接元音的不同，最突出峰的峰值有不同。同时这些模式得到了统计上的支持说明它们是客观存在的。

参考文献

1. Delattre, P.C., Liberman, A.M., Cooper F.S.,(1955), "Acoustic Loci and Transitional Cues for Consonants" J. Acoust. Soc. Am, 27,769-773
2. Halle, M., Hughes, G.W., Redley, J.-P.A., (1957), "Acoustic properties of stop consonants" J. Acoust. Soc. Am, 29,107-116
3. Stevens, K.S., Blumstein, S.E., (1979), "Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants" J. Acoust. Soc. Am, 66, 1001-1017
4. Blumstein, S.E. (1986), "On Acoustic Invariance in Speech" Lawrence Erlbaum Associates, Inc., Publishers, 178-201
5. wu zongji & Lin Maocan(1988), outline of experimental phonetics, High Education Press
6. Yan, Jingzhu & Sun, Guohua, (1991), "Data bank of the acoustic parameters in monosyllables of Standard Chinese", RPR(IL,CASS), pp. 1-16

characteristics and its pronunciation in Physiology. This paper gave an analysis in different vowel environment for three types with bilabial, alveolar, velar articulation places and aspirated and unaspirated articulation manners. Using

SPSS software to process different examples and got the relationship of different acoustic patterns among three stop types.

附录:

表 1 ba、da 峰点频率值的非参数检验

	VAR2	VAR3	VAR4	VAR5	VAR6	VAR7	VAR8
Mann-Whitney U	147.500	164.500	117.500	113.500	162.500	169.500	142.000
Wilcoxon W	300.500	335.500	327.500	284.500	333.500	340.500	352.000
Z	-.717	-.458	-1.842	-1.969	-.515	-.309	-1.122
Asymp. Sig. (2-tailed)	.474	.647	.065	.049	.606	.758	.262
Exact Sig. [2*(1-tailed Sig.)]	.497(a)	.654(a)	.067(a)	.051(a)	.613(a)	.762(a)	.276(a)

a Not corrected for ties. b Grouping Variable: VAR1

表 2 ba、da 峰点能量值的非参数检验

	VAR2	VAR3	VAR4	VAR5	VAR6	VAR7	VAR8
Mann-Whitney U	134.000	126.000	49.000	91.000	113.000	68.000	62.000
Wilcoxon W	287.000	297.000	220.000	262.000	284.000	239.000	233.000
Z	-1.097	-1.579	-3.830	-2.602	-1.959	-3.274	-3.450
Asymp. Sig. (2-tailed)	.273	.114	.000	.009	.050	.001	.001
Exact Sig. [2*(1-tailed Sig.)]	.283(a)	.119(a)	.000(a)	.009(a)	.051(a)	.001(a)	.000(a)

a Not corrected for ties. b Grouping Variable: VAR1

表 3 ga、gu 峰点频率值的非参数检验

	VAR2	VAR3
Mann-Whitney U	.000	129.500
Wilcoxon W	190.000	319.500
Z	-5.303	-1.507
Asymp. Sig. (2-tailed)	.000	.132
Exact Sig. [2*(1-tailed Sig.)]	.000(a)	.138(a)

a Not corrected for ties. b Grouping Variable: VAR1