

普通话单元音的共振峰数据测量方法研究*

李爽¹, 熊子瑜²

(1. 中国社会科学院研究生院 语言学系, 北京 102488; 2. 中国社会科学院语言研究所, 北京 100732)

文 摘: 本研究基于大规模语音数据库, 通过设计 Praat 脚本程序自动提取并分析普通话单元音的共振峰数据。研究发现, 设置不同的共振峰分析上限频率, 即 Maximum Frequency (下文中简称 MF) 值, 会严重影响到第二共振峰(F2)的数据分析结果。基于这一发现, 本研究通过设置不同的 MF 值, 提取出相应的共振峰数据, 再通过数据统计分析和声学元音图来考察如何合理设置 MF 值, 以提取出更为准确有效的共振峰数据。研究结果表明, 针对不同的单元音, 可以设置不同的 MF 值以提取更为准确有效的共振峰数据。

关键词: 普通话; 单元音; 共振峰数据;

中图分类号: H1; H116.2;

元音的声学性质主要依赖于它的共振峰模式^[1]。随着语音实验和数字信号处理技术的发展, 对元音共振峰数据的提取、分析和研究成为研究元音的主要切入点。1969 年 Fant 等人根据瑞典语长元音的共振峰模式特征, 提出通过改变长元音共振峰的取样时段来提取出相对准确的共振峰数据^[2]。1976 年 Sundbergd 等人针对俄语元音和瑞典语元音研究了发音时喉头的升降对单元音共振峰数据产生的影响^[3]。尽管第一共振峰(F1)和第二共振峰(F2)及其模式是一个元音最重要的声学特征, 它们基本上可以决定一个元音的音色^[4], 但许多实验证明, 在共振峰模式中受其他因素影响较大、需要进一步研究的是 F2 的绝对频率。如吴宗济先生认为有些音的 F3 对 F2 影响很大, 如单元音/i/, 测算/i/的 F2 时需用一种公式来补偿 F2 的频率值^[5]; Bladon 和 Fant 根据元音合成—听辨实验总结出 F2 经过权重调节之后, 合成出来的元音才最接近原始的元音^[6]。

20 世纪 50.60 年代以来, 普通话单元音共振峰数据的分析和提取等方面取得了显著成就。比较有代表性的测量普通话单元音共振峰数据的方法有 1964 年吴宗济先生提出的偏峰值算法^[7]、1965 年徐焕章先生提出的消声室宽带分析法、混响室(集体发声)法、歌声分析法三种方法^[8]、1979 年张家驂先生提出的准动态元音分析法^[9]、2006 年孟子厚先生提出的共振峰统计特性测量方法^[10]等。对比这些研究的结果, 可以看出部分单元音的共振峰数据相差还很大, 如吴宗济先生测量出/a/的 F1 为 1000Hz, 而在徐焕章先生的数据中却为 770Hz; 徐焕章先生测量出/i/的 F2 为 2100Hz, 而在孟子厚先生的数据中却为 2871Hz。这些差异说明普通话单元音的共振峰数值还需要进一步分析和研究。

近来, 本研究在采用 Praat 程序大批量分析和

提取普通话单元音共振峰数据时发现: 设置不同的 MF 参数, 所分析出来的单元音共振峰数据会发生显著变化, 如下图 1 和图 2 所示。

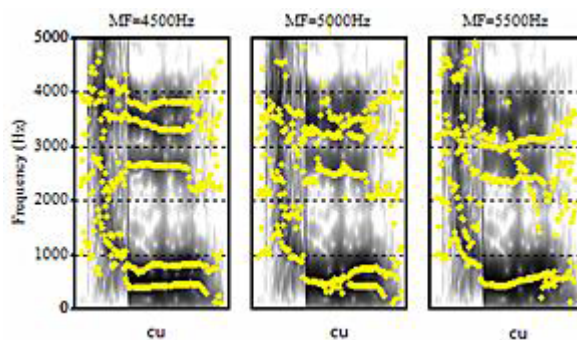


图 1 女发音人发“cu”音的共振峰语图

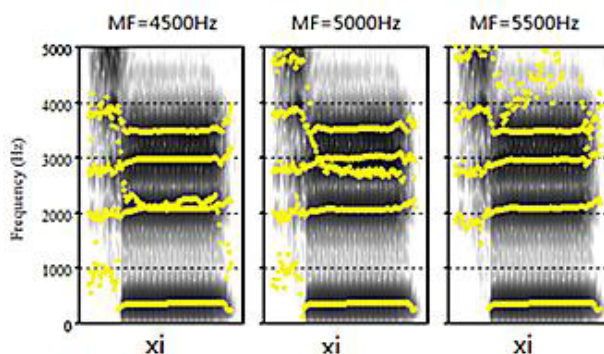


图 2 男发音人发“xi”音的共振峰语图

在图 1 中, 当 MF 值设置为 5500Hz 时, Praat 程序对单元音/u/分析出来的 F1 和 F2 数据存在明显错误, 当 MF 值设置为 4500Hz 时, F1 和 F2 的数据清晰可辨, 看起来更为准确有效。在图 2 中, 当 MF 值设置为 4500Hz 时, Praat 程序对单元音/i/分析出来的 F2 和 F3 数据纠缠在一起, 这显然不利于共振峰数据的正确提取, 当 MF 值设置为 5500Hz

时,分析出来的 F1、F2 和 F3 数据清晰可辨,显得更为准确有效。

基于这一观察结果,本研究认为通过合理地设置 MF 值,有可能提取出更为准确有效的共振峰数据。本文接下来将通过设置不同的 MF 值提取出相应的共振峰数据,然后在此基础上对比分析 MF 值的不同设置是否会对共振峰数据的分析结果产生显著影响,以探寻 MF 值的合理设置方式。

1 数据

本研究所使用的语音材料全部取自“普通话基础语音数据库”。该语料库包含了 71 位北京地区成年发音人(35 男,36 女)的录音材料,内容覆盖字、词、句和篇章等材料,所有录音过程都在录音室环境下完成。本文仅使用其中单元音韵母的单音节材料,一共 8 千多个。首先对这些单音节材料进行声韵母切分和标注,然后利用 Praat 脚本自动提取出全部单元音韵母的 F1 和 F2 的均值数据。提取共振峰数据时,本研究分别将 MF 值设置为 3500Hz、4000Hz、4500Hz、5000Hz、5500Hz、6000Hz 和 6500 Hz,从而针对每个单元音,可以分析并提取出了七组不同的共振峰数据。

2 共振峰数据分析

为了综合评价在设置哪种 MF 值时所提取到的共振峰数据更为准确有效,本研究利用 SPSS 统计软件对这些共振峰数据进行了决策树判别分析和均值分析,并通过声学元音图对提取出来的共振峰数据进行观察和评价,以考察这些共振峰数据的效度和信度,从而确定最佳的 MF 参数和最佳的共振峰数据。

2.1 判别分析

本文根据普通话单元音的 F1 和 F2 数据,采用 SPSS 软件所提供的决策树分析工具对全部单元音进行判别分析,以考察基于共振峰数据来判别普通话单元音的准确率。一般而言,判别的准确率越高,表明所采用的共振峰数据越有效度,越能反映出不同元音之间的区别。

考虑到普通话的两个舌尖元音 /ɿ/ 和 /ʅ/ 具有互补分布、发音相近的特点,纯粹基于 F1 和 F2 数据很难有效区分,因此在进行判别分析时,本文先将这两个舌尖元音合并处理,并标记为 I,以区别于舌面元音 /i/。下面的表 1 和表 2 分别给出了男女发音人的单元音判别准确率,粗体字表示各单元音判别的最高准确率。

表 1 男发音人的单元音判别准确率(%)

	3500	4000	4500	5000	5500	6000	6500
/a/	94.9	96.5	99.7	99.7	98.7	96.8	97.3
/ə/	86.8	95.1	97.6	97.8	95	92.9	82.8
/i/	82.7	70.2	79.1	92.9	94.1	88	80.5
I	69.3	79	79	74.9	76.8	81.8	79.6
/o/	72.9	86.9	80.7	77.6	81	68.2	65.1
/u/	89.4	90.5	88.7	86.7	79.1	83	85.8
/y/	9.4	70.6	68.5	72.7	64.7	69.6	61.5
all	80.9	87.1	88.5	89.9	88	86.4	83.4

注:表中 3500、4000.....6500 这些数据的单位是“Hz”,即分析共振峰数据时所设置的 MF 值,下文表格同此。

表 2 女发音人的单元音判别准确率(%)

	3500	4000	4500	5000	5500	6000	6500
/a/	76.8	87.3	87.4	93.2	99.1	97.6	96.1
/ə/	69.0	91.0	88.4	91.3	88.0	90.0	93.4
/i/	60.6	70.7	73.6	51.6	66.4	83.6	91.0
I	54.2	56.1	72.6	85.0	87.5	83.8	78.5
/o/	69.6	73.6	76.1	70.6	66.0	72.0	72.7
/u/	82.7	85.3	91.1	84.7	82.7	85.0	83.5
/y/	12.9	0.0	12.5	85.0	88.9	90.9	84.3
all	67.9	75.7	79.4	81.7	84.0	88.0	88.1

注:由于女发音人的特殊性,本实验进一步提取了 MF=7000Hz 时的共振峰数据,并测出其综合准确率为 87.1%,低于 MF=6500Hz 时的准确率。

对比分析表 1 和表 2 中的判别准确率数据可以得出:(1)当 MF 值设置为 5000 Hz 时,男性发音人的单元音判别准确率达到最高;当 MF 值设置为 6500 Hz 时,女性发音人的单元音判别准确率会达到最高。这表明,在分析和提取男性发音人的单元音共振峰数据时,把 MF 值设置为 5000 Hz 是合理的,而在分析和提取女性发音人的单元音共振峰数据时,把 MF 值设置为 6500 Hz 会比 5500Hz 更为有效,分析和提取出来的共振峰数据会有更好的区分度。(2)针对不同的单元音,可以设置不同的 MF 值,以获取更好地判别准确率。

根据表 1 和表 2 的判别分析结果,本文得到各个单元音所对应的最佳 MF 值,如下表 3 所示。再根据各个单元音所对应的最佳 MF 值,分析和提取出各个单元音的共振峰数据。然后再次进行单元音判别分析。结果表明,男女发音人的单元音综合判别准确率都有一定程度的提高,分别达到了 93% 和 89.9%。因此本研究认为,针对不同单元音分别设置不同的 MF 值,有利于分析和提取出更有效度的共振峰数据,从而能在一定程度上提高对普通话单元音的判别准确率。

表 3 普通话各个单元音所对应的最佳 MF 参数

单元音	最佳 MF (女)	最佳 MF (男)
/a/	5500Hz	4500 Hz 或 5000 Hz
/ə/	6500 Hz	5000 Hz
/i/	6500 Hz	5500 Hz
/ɿ/	5500 Hz	6000 Hz
/ʅ/	5500 Hz	6000 Hz
/o/	4500 Hz	4000 Hz
/u/	4500 Hz	4000 Hz
/y/	6000 Hz	5000 Hz

2.2 均值分析

本文采用 SPSS 软件对普通话单元音的 F1 和 F2 数据进行均值分析, 并通过对标准差的大小来考察设置不同的 MF 值会对共振峰数据的分析结果产生哪些影响。一般而言, 标准差越小, 表明所提取出来的共振峰数据越有信度, 其内部一致性越高, 数据越稳定。

统计结果表明, 在分析普通话单元音的共振峰数据时, 若设置不同的 MF 参数, 对元音的 F1 数据并不会产生较大影响, 而对 F2 数据产生的影响较大。下面表 4 和表 5 按男女发音人分类给出了各个单元音的 F2 均值、样本数和标准差等数据。

表 4 男发音人的单元音 F2 的均值分析结果 (按 MF 分类)

	3500	4000	4500	5000	5500	6000	6500
/a/	1143	1229	1273	1296	1299	1300	1307
	939	939	939	939	939	939	939
	123	107	92	94	97	100	104
/ə/	1025	1073	1082	1082	1097	1120	1144
	721	721	721	721	721	721	721
	119	85	87	88	97	125	169
/i/	1475	1902	2050	2193	2234	2242	2261
	631	631	631	631	631	631	631
	577	428	345	204	169	161	155
/ɿ/	1275	1281	1329	1334	1345	1359	1393
	138	138	138	138	138	138	138
	144	142	112	110	112	111	172
/ʅ/	1373	1468	1466	1497	1546	1571	1587
	181	181	181	181	181	181	181
	198	131	139	136	125	122	130
/o/	761	785	794	800	834	926	1036
	321	321	321	321	321	321	321
	78	75	83	86	191	421	603
/u/	674	693	714	753	962	1160	1265
	980	980	980	980	980	980	980
	86	81	122	214	531	718	791
/y/	1483	1770	1834	1843	1894	1936	1977
	286	286	286	286	286	286	286
	430	224	140	129	130	135	151

注: 表中各单元音对应的三行数据中, 第一行数据表示的是单元音在相应 MF 值下的均值, 单位是“Hz”; 第二行表示的是单元音的样本数目,

单位是“个”; 第三行表示的是单元音在相应 MF 值下共振峰数据的标准差。各组的粗体数据表示的是标准差最小的一系列数据。下表 5 同此。

表 5 女发音人的单元音 F2 的均值分析结果 (按 MF 分类)

	3500	4000	4500	5000	5500	6000	6500
/a/	1189	1327	1408	1490	1579	1583	1577
	951	951	951	951	951	951	951
	123	139	162	142	110	111	109
/ə/	1042	1179	1206	1250	1255	1250	1255
	743	743	743	743	743	743	743
	165	106	101	94	97	96	97
/i/	773	1351	1890	2176	2538	2676	2716
	639	639	639	639	639	639	639
	354	607	615	628	410	278	226
/ɿ/	1164	1422	1524	1536	1590	1590	1591
	139	139	139	139	139	139	139
	328	194	163	140	136	143	142
/ʅ/	1363	1562	1735	1801	1823	1842	1869
	182	182	182	182	182	182	182
	344	259	167	160	163	164	179
/o/	832	882	905	946	960	956	960
	326	326	326	326	326	326	326
	86	93	106	114	137	139	137
/u/	692	715	750	844	866	831	870
	1001	1001	1001	1001	1001	1001	1001
	77	96	146	209	234	224	370
/y/	1137	1662	1728	2184	2237	2216	2219
	287	287	287	287	287	287	287
	506	475	460	183	161	157	148

为了进一步说明 MF 参数对单元音 F2 数据的影响, 本文采用标准差作为度量 F2 数据稳定性的测算指标, 制作出相应的变化曲线, 如下图 3 和图 4 所示, 变化曲线能较好地反映出在设置不同的 MF 参数时所提取出来的共振峰数据的稳定性。其中, 变化幅度越大, 说明 F2 的分析结果受到 MF 参数的影响越大。

在普通话的几个单元音中, 不同单元音的 F2 数据受 MF 值的影响是不同的, 从图 3 和图 4 中可以看出某些单元音的变化幅度相对较大, 如男发音人的/u/、/o/、/i/、/y/等单元音, 女发音人的/i/、/y/、/u/、/ɿ/、/ʅ/等单元音, 说明这些单元音的 F2 数据很容易受到 MF 值的影响, 若 MF 值设置不当会严重影响 F2 数据的分析结果。而另外几个单元音的变化幅度却相对平缓, 其中男发音人的/a/和女发音人的/a/、/ə/是变化幅度最平缓的。

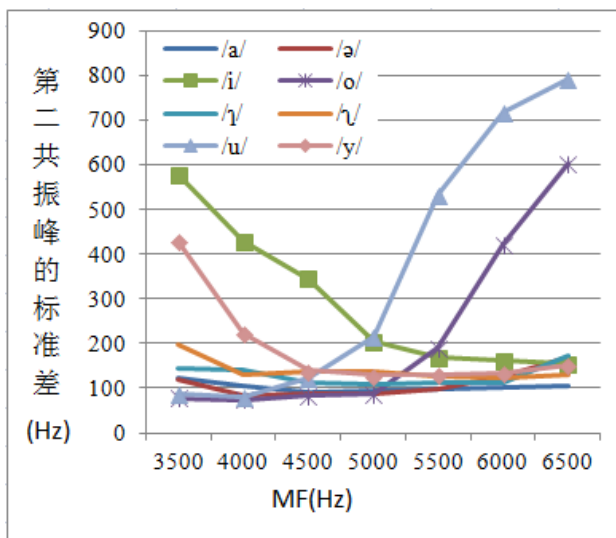


图3 各单元音 F2 的“标准差”变化曲线图（男）

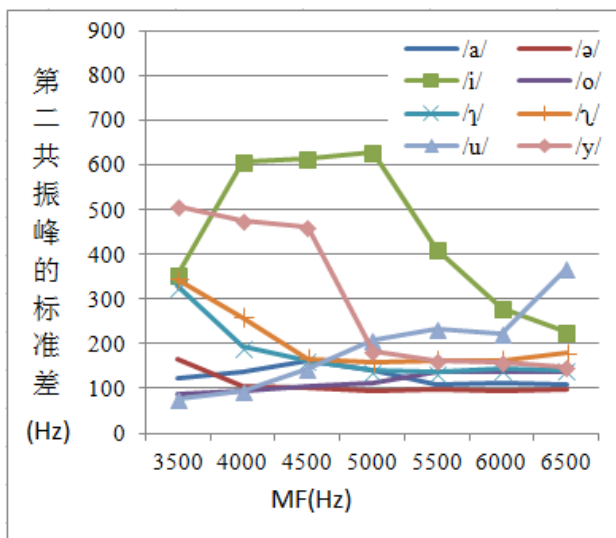


图4 各单元音 F2 的“标准差”变化曲线图（女）

通过设置不同的 MF 值并提取出相应的共振峰数据，再根据对 F1 和 F2 的均值分析结果，选取其中标准差最小的均值数据，本文得到男女发音人各单元音 F1、F2 的最佳 MF 值（见表 6）以及最佳的共振峰均值数据（见表 7）。

表 6 分析普通话单元音共振峰数据时的最佳 MF 参数(Hz)

单元音	男 F1	男 F2	女 F1	女 F2
/a/	5000	4500	5500	6500
/ə/	3500	4000	4000	5000
/i/	4000	6500	6500	6500
/ɨ/	3500	5000	4000	5500
/ɥ/	3500	6000	4000	5000
/o/	4000	4000	3500	3500
/u/	4000	4000	3500	3500
/y/	4000	5000	4500	6500

表 7 普通话单元音的 F1 和 F2 数据 (Hz)

单元音	男 F1	男 F2	女 F1	女 F2
/a/	833	1273	1061	1577
/ə/	470	1073	542	1250
/i/	315	2261	367	2716
/ɨ/	396	1334	458	1590
/ɥ/	391	1571	456	1801
/o/	485	785	528	832
/u/	357	693	390	692
/y/	322	1843	379	2219

2.3 声学元音图

根据元音 F1 和 F2 的频率值画出来的元音分布图称为元音频率图或声学元音图^[11]。其纵轴表示 F1，横轴表示 F2。作图时一般用声学刻度，单位是赫兹 (Hz)，也有用听感标度的，单位是巴耳克 (Bark)，赫兹与巴耳克之间可以进行换算。声学元音图反映了各个元音的声腔共鸣特性，采用的是客观测量出来的声学语音学数据，因此会比较客观。本研究制作的声学元音图同时给出了这两种刻度，以便于观察和比较。

声学元音图比判别分析、均值分析更直观。通过图中各单元音的相对位置、椭圆形状和大小、不同单元音的重合度以及同一单元音的集中度就可以对比出各组数据的优劣。其中，各单元音在声学元音图中的相对位置是否与发音时的舌位基本一致最为重要。其次，椭圆越小则表明同一单元音的数据一致性就越高。再次，不同单元音的数据重合度越小，则说明各单元音之间的区分度越高。

根据表 3 和表 6 中给出的 MF 值，本研究提取出相应设置下的共振峰数据并制作出其声学元音图，进而把它们与设置为常规 MF 值(即男 5000Hz，女 5500Hz)时的声学元音图加以比较，详见图 6-图 10 (由于 Praat 软件不识别舌尖元音 /ɨ/ 和 /ɥ/，本实验在用 Praat 画声学元音图时用 ii 代替 /ɨ/，用 iii 代替 /ɥ/，以便和单元音 /i/ 相区别，如图中所示)。根据观察不难发现，对于男发音人的单元音共振峰数据而言，图 7 中的数据比图 6 中的更为准确有效，图 6 中的数据比图 5 中的更为准确有效。对于女发音人的单元音共振峰数据而言，图 10 中的数据比图 9 中的更为准确有效，图 9 中的数据比图 8 中的更为准确有效。这表明，在分析单元音的共振峰数据时，若能针对不同的元音设置不同的 MF 值（见表 3），其数据质量会比设置同一个 MF 值(即男 5000Hz，女 5500Hz)的更好，如果还能进一步针对不同元音的 F1 和 F2 来分别设置不同的 MF 值（见表 6），则能取得更加准确有效的共振峰数据。

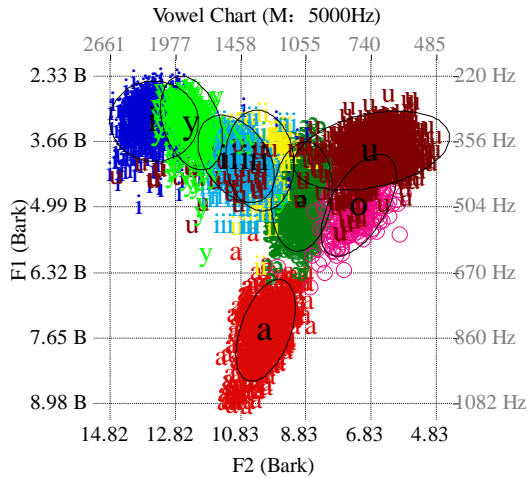


图5 单元音 MF 为 5000Hz 时的声学元音图 (男)

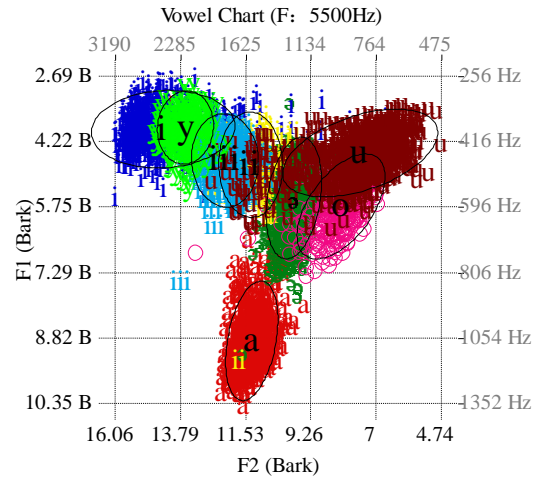


图8 单元音 MF 为 5500Hz 时的声学元音图 (女)

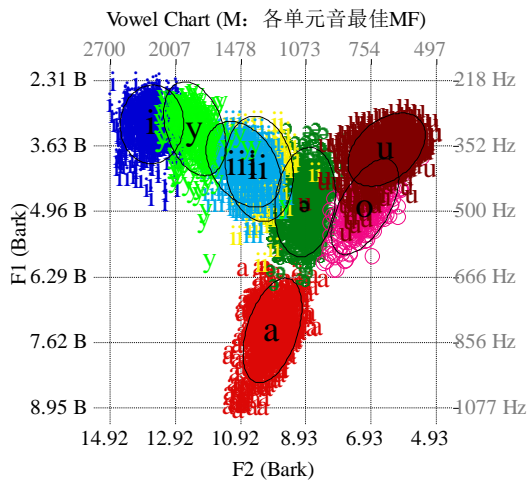


图6 各单元音最佳 MF 参数时的声学元音图 (男)

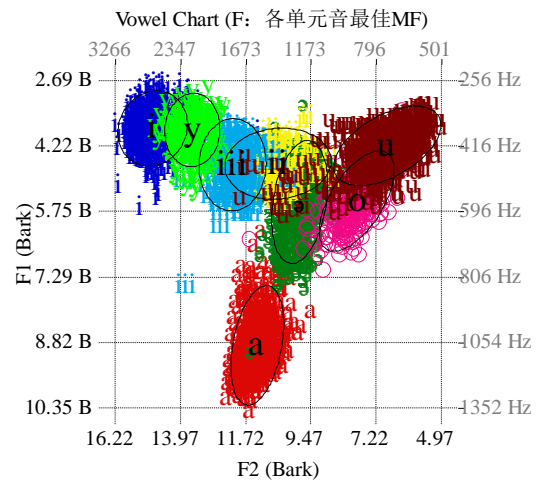


图9 各单元音最佳 MF 参数时的声学元音图 (女)

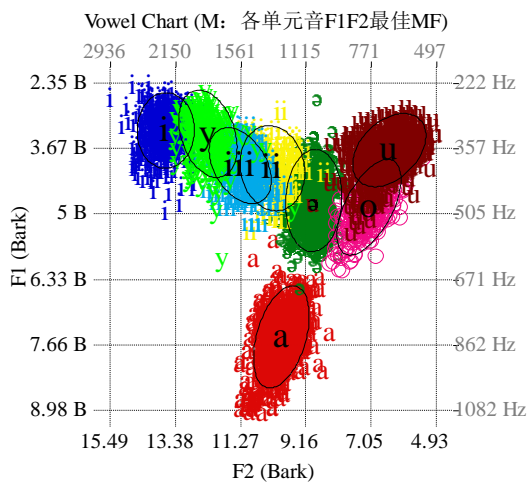


图7 各单元音 F1、F2 最佳 MF 参数时的声学元音图 (男)

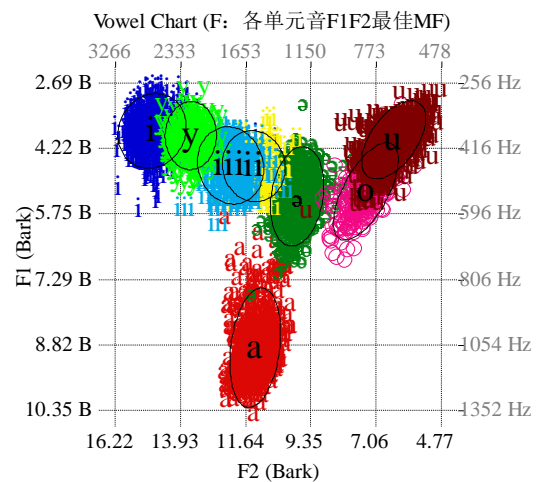


图10 各单元音 F1、F2 最佳 MF 参数时的声学元音图 (女)

3 结论

本研究基于大规模语音数据库，通过设计 Praat 脚本程序自动提取并分析普通话单元音的共振峰 F1 和 F2 的数据。研究发现，设置不同的

MF 值, 会严重影响到 F2 数据的分析结果。基于这一发现, 本文通过设置不同的 MF 值, 提取出相应的共振峰数据, 再通过声学元音图以及数据统计分析来考察如何合理设置 MF 值, 以提取更为准确有效的共振峰数据。研究表明, 针对不同的单元音, 可以设置不同的 MF 值以提取出更为准确有效的共振峰数据。

参考文献

- [1] Ladefoged, P. Vowels and Consonants[M]. Los Angeles: Blackwell Publishers, 2001
- [2] Fant, G. and Henningsson, G. and St º alhammar, U. Formant frequencies of Swedish vowels[J]. STL-QPSR, 1969, 10(4):026-031
- [3] Sundberg, J. and Nordstr ºm, P-E. Raised and lowered larynx -the effect on vowel formant frequencies[J]. STL-QPSR, 1976, 17(2-3) : 035-039.
- [4] 林焘, 王理嘉. 语音学教程[M]. 北京: 北京大学出版社, 1992. LIN Tao, WANG Lijia, . Course of Phonetics [M]. Beijing: Peking University Press , 1992. (in Chinese).
- [5] 吴宗济等. 汉语普通话单音节语图册[M]. 北京: 中国社会科学出版社, 1986. WU Zongji, et al. The Spectrographic Album of Mono-Syllables of Standard Chinese [M]. Beijing: Chinese Social Sciences Press, 1986. (in Chinese).
- [6] 吴宗济, 林茂灿等. 实验语音学概要[M]. 北京: 高等教育出版社, 1989. WU Zongji, LIN Maocan, et al. The Summary of Experimental Phonetics [M]. Beijing: Higher Education Press , 1989. (in Chinese).
- [7] 吴宗济. 普通话元音和辅音的频谱分析及共振峰的测算[J]. 声学学报, 1964, 1 (1): 33-39 WU Zongji. The spectrographic analysis of the vowels and consonants in standard colloquial Chinese [J]. ACTA ACUSTICA, 1964, 1 (1): 33-39. (in Chinese).
- [8] 徐焕章. 普通话元音的平均声谱[J]. 声学学报, 1965, 2 (1): 1-7. XU Huanzhang. The mean spectral characteristics of the vowels of the standard Chinese [J]. ACTA ACUSTICA , 1965, 2 (1): 1-7. (in Chinese).
- [9] 张家驩等. 准动态元音分析法[J]. 声学学报, 1979, 1 (1): 23-29. ZHANG Jialu, et al. A semi-dynamic method for spectrographic analysis of vowels [J]. ACTA ACUSTICA , 1979, 1 (1): 23-29. (in Chinese).
- [10] 孟子厚. 普通话单元音女声共振峰统计特性测量[J]. 声学学报, 2006, 31 (1): 199-202. MENG Zihou. Statistical survey of female pure vowel formant in Chinese [J]. ACTA ACUSTICA , 2006, 31 (1): 199-202. (in Chinese).
- [11] 林焘, 王理嘉. 语音学教程[M]. 北京: 北京大学出版社, 1992. LIN Tao, WANG Lijia, . Course of Phonetics [M]. Beijing: Peking University Press , 1992. (in Chinese).

The Research into the Measurement Method of Formant Data of Mandarin Monophthong

Abstract: Based on a large speech database, several praat scripts are designed in our experiment in order to extract and analyze the formant data of Mandarin monophthong automatically. what the experiment find is, different Maximum Frequency ("MF" for short) parameters can affect the result of the second formant ("F2" for short) significantly. In view of this fact, the research set different MF parameters to extract the corresponding formant data, then examined those formant data in which parameters is more accurate by comparing the statistical analysis of data and painting vowel chart. The research result shows that we should set different MF parameters for each Mandarin monophthong in order to extract more accurate formant data.

Key words: Mandarin; monophthong; formant data;

(本文原载 NCMMS2013 中国贵阳 2013 年 8 月)