

电磁发音仪与宁波方言的元音研究

胡方

中国社会科学院语言研究所

提要 本文用电磁发音仪研究宁波话元音的发音，并试图解释元音发音的生理机制。结果发现元音高低与实验所得的舌位材料并没有直接、清楚的关联。但对舌面发音点所作的平行因子分析表明：表面上纷繁复杂的舌位材料可以成功分解出两个底层舌移动机制，“回缩及后举”（retraction and back raising）和“前举”（front raising）。语言学上元音高低、前后之对立均可以用这两个机制来解释。

关键词 电磁发音仪 元音高低、前后 平行因子分析

一 引论

自从 Bell (1867) 以来，学者们一般都用高低（height）和前后（backness）等发音特征来描写元音。也就是说，每一个元音在口腔内都有一特定的舌位形状。此一理论成为语音学内元音描写（如国际语音学会的元音图）以及音系学上元音区别性特征的基础（如 Chomsky and Halle 1968）。在英国的语音学传统上，元音的高低、前后更明确地以舌位最高点来定义（Jones 1909）。然而这个传统的元音描写模式并未得到经验材料的验证，相反，学者们在发音研究中得到不少反证。早在 1910 年，Meryer 就发现 [ɪ] 的舌位比 [e] 还低；Russel (1928) 更以大量的 X-ray 的材料否定了舌位高低、前后的传统说法：[ɪ] 舌位比 [e] 低，[u] 可能是前元音，而 [ɑ ɔ ɔ] 等元音的发音点是在咽腔内，与口腔内的舌位绝对高低并无直接关系。20 世纪中叶以来，Ladefoged 及其同事对这一问题进行了大量的生理、声学、以及感知方面的研究（Ladefoged 1967, 1971, 1975, 1976; Ladefoged 等 1972）。结果表明：传统元音描写模型描写的是元音的听觉印象，但在表述的时候却把它们翻译成了生理术语，因为元音高低、前后并没有发音生理上的事实基础，然而却得到元音的声学事实的支持。例如，Ladefoged 等 (1972) 发现，一些美国英语的发音人 [ɪ e e] 的舌位几乎是一样的；

Ladefoged (1975, 1976) 发现 Stephen Jones 所发的基本元音 (Cardinal Vowels) 的舌位完全不是所定义的那一套，“甚至连前、后元音都没有很好的区分”，“相对高低更是错得离谱” (1976: 10-12)。而另一方面，Ladefoged (1967) 的心理声学实验发现，受过训练的语音学者，能够很好地分别元音的高低、前后，即使他们听到的只是录音材料。这使他相信元音高低、前后与舌位无关，只是受过训练的语音学家能够“标注他们所听到的元音共振峰结构并模仿出来” (Fromkin 1985: 5)。Ladefoged (1975, 1976) 更进一步指出：传统的基本元音图与以听觉基础为刻度的声学元音图有本质关联。

虽然现在语音学界的许多学者一般都同意元音的高低、前后应该从声学、或者心理声学的角度来阐释，但对于究竟是怎样的生理上的发音机制形成了这种声学或听觉上的语音对立，并不是完全清楚。本文应用电磁发音仪，分析宁波方言元音的发音特征，并试图对上述的理论问题做出回答。宁波方言有丰富的单元音，其中在 (C) V 位置上有音位对立的就有十个：[i y ɣ e ø e a ɔ o u]。另外，还有出现在音节尾喉塞音前面的两个短元音 [a ɔ]，以及两个舌尖元音 [ɿ ʮ]。限于篇幅，本文只讨论十个一般单元音。

二 实验方法

七位宁波发音人（四男三女，年龄介乎 20 至 30 岁）参与了此项研究。发音材料是十个包含目标元音的单音节词，除了元音 [ø] 必须与齿龈音（alveolar）同现外，所有测试词都是零声母，高平调。

材料收集使用的是 Carstens 公司的电磁发音仪 (Electromagnetic Articulograph, 简称为 EMA)。其中二位男发音人录音用的是 AG100 系统，实验在香港城市大学中文、翻译及语言学系的语音实验室进

行；其他发音人用的是 AG200 系统。电磁发音仪最大的优点是可以实时监测发音器官的动作，尤其是以前比较难测量的口腔内的发音器官如舌头，另外，由于其对人体几乎无损害，因此没有以前基于 X 光技术仪器的局限，使得大量采集发音材料成为可能，甚至可以应用于采集儿童的语言。该系统的基本工作原理是应用通电后所产生的交互电磁场追踪在发音人中分矢形面（mid-sagittal）上的发音器官的移动。简单地说，有三个发射环（transmitter coil）被互相等距离地安装在一个固定头盔上，这样就可以三个不同的频率产生一个径向均衡的交互电磁场。当发音人戴上头盔，位于发音人中分矢形面上的发音器官（如舌头、下颚、唇等）上的小型传感器（transducer coil）所产生的感应电压（induced voltages）就会以不同的高频率被交互电磁场所采样。根据电磁场力的衰减与其与发射点的距离之立方约成正比这一物理定律，利用采样电压就可以测量出每个采样点与每个发射点之间的距离，以及其相对于发射轴的斜度。然后，根据一定的计算方式，每一个采样点（也即是发音点或参照点）就可以在一个笛卡尔坐标系上标示出来（关于 EMA 完整的技术背景报告，可以参见：Perkell 等 1992，也可以参见 Hoole 1996，以便对 EMA 技术在方法论上的详细考虑有所了解）。

在实验中，我们将接收传感器沿矢状平面中线粘贴在发音人的上唇（Upper Lip, UL）、下唇（Lower Lip, LL）、下齿龈脊（the gum ridge at the lower teeth, Jaw）、舌尖（Tongue Tip, TT）、舌中（Tongue Mid, TM）、舌背（Tongue Dorsum, TD）。其中，TT 粘在离舌尖不到 1 厘米的位置，TM 距 TT 约 3 厘米，TD 距舌尖约 6 厘米。另外，还有两个传感器分别粘贴在发音人的鼻梁（the bridge of the nose, Ref. 1——未在图中显示）和上齿龈脊（the gum ridge at the upper teeth, Ref. 2），作为数据校正用的参照点¹。在本项研究中，我们主要关注粘在舌头上的三个接收传感器。具体如图 1 所示：

正式录音前，有一定的时间给予发音人练习，

以便使其习惯粘着传感器说话。发音人感觉可以自然说话了之后，录音开始。发音人阅读以随机顺序写在纸上的测试词，每个测试词都放在一个引导句中：[ɲo io do? __ pa? nau tʰiŋ] “我要读__给你听”。语音录音与发音器官动作录音同时进行，用的是 EMA 系统自带的语音录音频道。每个测试词重复五遍。

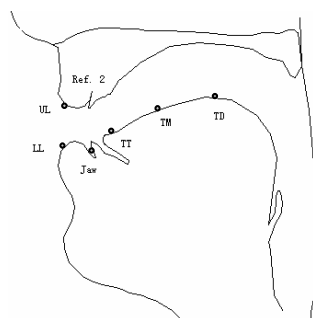


图 1 接收传感器粘贴示意图

当语料录音结束后，我们录制了每个发音人的咬合面（occlusal plane）方位，以便进行必要的坐标旋转。方法是：在一信用卡大小的硬纸片上粘上两个传感器，插入发音人口中，让发音人轻轻咬住尾部，然后根据两个传感器提供的数据确定录音坐标与发音人咬合面之间的角度。这样，所有录制的材料就可以根据这一角度调整，使得新坐标的横轴与发音人的咬合面相平行。根据发音人咬合面旋转坐标使得不同发音人之间的发音数据比较成为可能，因为旋转之后发音人之间的发音数据有了相同的参照系。

数据分析主要牵涉到提取目标元音的发音器官的位置信息。在本项研究中，元音的发音目标点（vowel target）的确定既用发音上的标准，也用声学上的标准。具体操作上，当主发音点的切线速度在元音中间附近位置达到最小值时，我们就定义这里是元音的发音目标点。在发音研究上，切线速度最小值标准（the tangential velocity minima criterion）是常用的确定发音目标值的方法（参见：Löfqvist 等 1993, Löfqvist 1999）。在本项研究中，这一标准并不严格应用到确定每个发音点上，而是只应用在确定主发音点上，然后根据主发音点同时决定其它发音点的位置信息。这样，元音的发音目标点就在同一时间点上被确定。在操作上，距离元音的窄缩位置（constriction location）最近的采样点就被认为

¹采样前两个男发音人所使用的 AG100 系统是个只有 5 个传感器频道的系统，实验中粘贴在下发音器官上（LL、Jaw、TT、TM、TD）。

是该元音的主发音点，具体讲，非低前元音 [i y ɤ ø ε] 是舌中(TM)点，低元音及后元音 [a ɔ o u] 是舌背(TD)点(关于元音发音上的窄缩位置，可以参见: Wood 1979)。

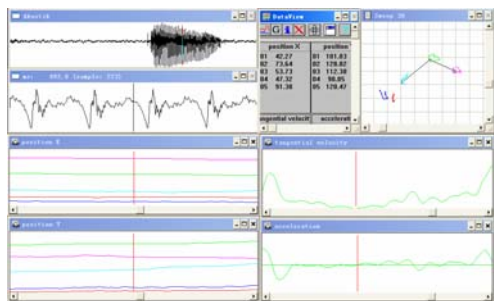


图 2 数据分析窗口例(元音 [e], 男发音人一)。左侧子窗口(自上而下): 声波信号、光标附近放大声波信号、发音位置横轴、发音位置纵轴; 右侧子窗口(自上而下): 坐标点数据信息(左上)、数据 X/Y 坐标显示(右上)、切线速度曲线、加速度曲线(后两者只显示了 TM)

图 2 是 Emalyse 软件的一个分析窗口例子，声波信号窗口显示所录元音的声波信息，放大声波信号窗口显示光标附近元音的声波信息，发音位置横轴窗口显示元音 5 个采样点在时间轴上的横轴位置信息，发音位置纵轴窗口显示元音 5 个采样点在时间轴上的纵轴位置信息，数据信息窗口显示在光标点上的各项数值，X/Y 坐标窗口将发音位置横轴和纵轴上的数据信息标示在一个笛卡尔坐标系上，切线速度窗口显示各采样点在时间轴上的切线速度信息，加速度窗口显示各采样点在时间轴上的加速度信息(后两个窗口，在图中只显示了 TM 采样点的信息)。所有非数值窗口中的光标都是同标的(synchronized)。为了清楚显示有用信息，图中发音位置横、纵轴窗口，X/Y 坐标窗口，切线速度窗口，以及加速度窗口都有适度放大。图中切线速度窗口和加速度窗口只显示了 TM 采样点的信息，这是因为对元音 [e] 来讲，TM 采样点就是所定义之主发音点。如图所示，在光标所在位置，在元音中部附近，主发音点的切线速度达到最小值(请注意此时加速度接近零)，根据上文所述原则，我们认为此处就是该元音的发音目标点，因此，所有采样发音点的位置信息都在这一点上确认，其数据就被记录下来作进一步分析之用。

三 研究结果

图 3 显示了宁波声学元音图(具体参见: 胡方 2005)。图中横轴和纵轴的标度是赫兹值，但刻度已经转化为 Bark，而且，纵轴刻度是横轴的两倍，以加强第一共振峰在元音感知上的重要性，这样，元音的分布更符合听觉结果。图中的每个元音都来自 50 个采样值的均值(10 个发音人×每人每个元音重复 5 遍)。

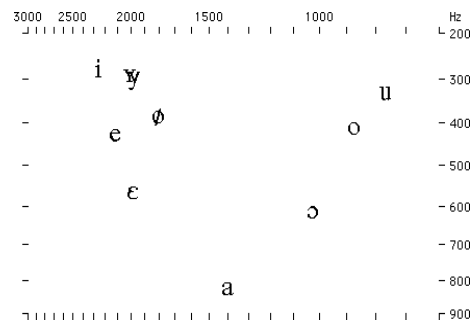


图 3 宁波元音在声学语音图上的分布

如引论中所述，根据听觉-声学的结果，元音的高低和前后是非常清楚的。从图 3 我们可以看到，宁波话的 [i y ɤ u] 是高元音，[e ø o] 是半高元音，[ε ɔ] 是半低元音，[a] 是低元音；[i y ɤ e ø ε] 是前元音，[u o ɔ] 是后元音，[a] 是央元音，即低元音不区分前后。

这里的问题是：元音的高低、前后在发音上是怎样实现的？传统的观点认为元音的高低对应于舌位的高低，元音的前后对应于舌位的前后。这种对应关系，在我们的发音材料中是否可以找到证据？

图 4 显示了发音人一在发宁波元音时根据三个舌面采样点所确立的舌位示意图，图中每个国际音标代表该元音每个采样点的平均值。图中元音上方的曲线是所录制的发音人的上颚曲线。

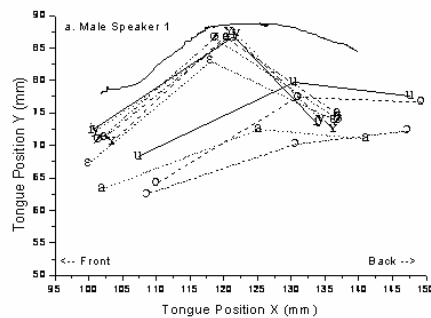


图 4 宁波元音舌位图: 男发音人一(单位: 毫米)

如图中所示，就采样的三点所反映的元音舌位图而言，前、后元音的区别还是很明显的。粗略地讲，前元音的整个的舌头位置是相对前置的，三个采样点中 TM 是最高点，而且不同程度地靠近硬腭，显示了在发前元音时，舌中间部位是主要发音位置，其主动接近硬腭以形成窄缩点；而后元音的整体舌头位置则是相对地后缩的，TD 一般是最高点，显示了舌背部分才是主要发音点，舌背拉动使得整个舌头向后，因而其窄缩点是在软腭或者是更后的位置。就舌位前后而言，低元音似乎是处在前、后元音的一个中间状态，因而在声学图上，它也是在央元音的位置上。从所得的材料看，如果说舌位前后与元音前后还有较好的对应关系，那么，元音高低与舌位高低就显得杂乱无章了。五个高、半高前元音 [i y e ø] 都有类似的舌位，高和半高后元音 [u o] 的 TD 点也在差不多的高度上。而且，我们很难在同高度的前、后元音之间以舌位高低建立起关系。因此，正如 Wood (1979: 25) 所指出的，传统上用舌位高低、前后来描写元音，是使用了错误的描写维度。也就是说，言语自动控制系统 (speech motor system) 在调控元音发音时，并不是通过调控舌位高低、前后来实现的。那么，这里的问题就变为：是什么样的舌移动机制在调控元音的发音？为了回答这个问题，我们对测量所得的舌位资料进行了平行因子分析 (parallel factor analysis, 简称为 PARAFAC)。

平行因子分析是 Harshman (1970) 发展出来的三众数 (three-mode) 因子分析模型，不同于一般的二众数 (two-mode) 模型的是它可以给出唯一解，也就是没有了二众数因子分析中的因子轴旋转问题。Harshman 等 (1977) 首次将它应用到了元音的发音研究，结果他们从五个美国英语发音人的 X 光片材料中发现了两个舌移动机制，前举 (front raising) 和后举 (back raising)。Hoole (1999) 将它应用到分析以电磁发音仪所得到的发音材料上，根据七个德语发音人的发音材料，结果也是发现了两个类似的舌移动机制。在本次研究中，我们发现也是二因子的模型最适合把握元音发音的舌移动特征，解释了约 90% 的方差 (variance)。舌发音点的因子装载作用 (factor loading effect) 可以直观地以图形表示为相对于所有发音人所有元音的平均舌位

的移动，见图 5。其中，七个发音人的平均上颚曲线也标示在了图中，以作参照。从图中，我们可以看到，因子一主要模拟了舌体的前后位置移动，以及舌体后部的上下位置移动，因子二则是模拟了舌面相对于硬腭的上下位置移动。这里的因子二与 Harshman 等 (1977) 和 Hoole (1999) 因子一完全一致，因子一也与他们的因子二类似。因此，所提取的两个舌移动机制可以分别称之为“回缩及后举” (retraction and back raising) 和“前举” (front raising)。

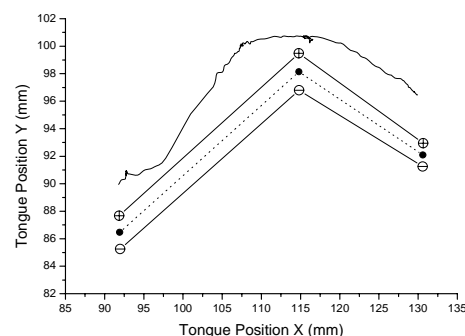
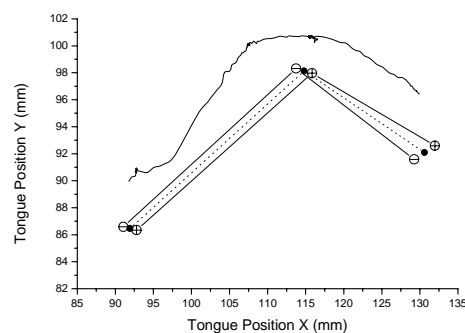


图 5 舌发音点二因子模型的图形化示例 (上: 因子一; 下: 因子二)。虚线所连实点是两个发音人所有元音的舌位均值, 实线所连的正、负号填塞的小圈分别表示相对于平均舌位的 ± 2 个标准差之位移。

每个宁波元音的发音可以视作是这两个舌移动机制的共同作用的结果。图 6 是元音的因子装载作用图, 数值指明构建每个元音所需之因子权重。图 6 与传统的“元音舌位图”或者声学元音图有一定的类似性, 这并不奇怪。不同的是, 声学元音图 (不管是纯声学的或者是基于听觉的) 所根据的是所有发音器官, 如舌、下颚高低、唇形等, 共同作用的结果, 而图 6 则是完全基于舌位的一个概括。如图,

坐标横轴（因子一）标示了“回缩及后举”的权重，而纵轴则标示了“前举”的权重。正值表明了该元音相对于平均舌位需要正位移，负值表明了需要负位移。比如：从图中我们可以看到，所有前元音都需要较大的因子一负值，这说明发前元音时舌头是趋前的；同理可见，发后元音时，舌头是后缩的。

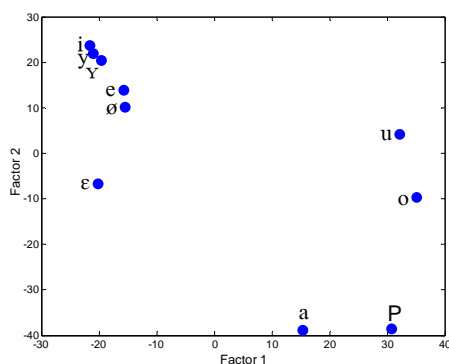


图 6：因子装载——元音。

四 结论

本文用电磁发音仪采集了宁波方言发音人的元音发音材料，从直接测量的元音舌发音点的位置信息看，元音的高低与舌位的高低没有明确的对应关系，这进一步证明了传统的元音描写术语，如元音高低、前后，并没有在生理上和发音上的实证基础，而只是有语言学上的分类意义，在语音学上，与之相关联的，正如前人所指出的，应该是声学上或者是心理声学上的证据。而在发音上，对材料所进行的平行因子分析结果显示，表面上纷繁复杂的舌位信息可以成功地分解出两个底层的舌移动机制，“前举”和“回缩及后举”，所有元音的发音都可以用这两个机制来模拟，而传统的元音描写维度如高低、前后，也可以用这两个机制来作重新的阐释。与前人对英语及德语等西方语言研究结果的一致性也表明，这两个舌移动机制在元音发音上是具有普遍性的²。

参考文献

² Jackson (1988) 在分析冰岛语时认为需要三个舌移动机制才能模拟冰岛元音的发音，不过，Nix等 (1996) 已经推翻了他的分析，证明也是两个类似的底层机制在起作用。

胡方 2005 《宁波方言元音的声学语音学研究》，《吴语研究》，341-348 页，上海教育出版社，上海。

Bell, A. M. 1867 *Visible speech or self-interpreting physiological letters for the writing of all languages in one alphabet*, Simpkin & Marshall, London.

Cho, T. 2001 *Effects of Prosody on Articulation in English*, Ph.D. dissertation, UCLA.

Chomsky, N. and Halle, M. 1968 *The sound pattern of English*, Harper & Row, New York.

Harshman, R. 1970 Foundations of the PARAFAC procedure: Models and procedures for an ‘explanatory’ multi-modal factor analysis. *UCLA Working Papers in Phonetics* 16.

Harshman, R., Ladefoged, P. and Goldstein, L. 1977 Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*, 62: 693-707.

Hoole, P. 1996 Issues in the acquisition, processing, reduction and parameterization of articulographic data. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation München (FIPKM)*, 34: 158-173.

Hoole, P. 1999 On the lingual organization of the German vowel system. *Journal of the Acoustical Society of America*, 106: 1020-1032.

Jackson, M. T. T. 1988 Analysis of tongue positions: Language-specific and cross-linguistic models. *Journal of the Acoustical Society of America*, 84: 124-143.

Jones, D. 1909 *The pronunciation of English*, Cambridge University Press, Cambridge.

Ladefoged, P. 1967 *Three Areas of Experimental Phonetics*. Oxford University Press, London.

Ladefoged, P. 1971 *Preliminaries to Linguistic Phonetics*, University of Chicago Press, Chicago.

Ladefoged, P. 1975 *A Course in Phonetics*, Harcourt Brace Jovanovich, New York

- Ladefoged, P. 1976 The phonetic specification of the languages of the world. *UCLA Working Papers in Phonetics*, 31: 3-21.
- Ladefoged, P., DeClerk, J., Lindau, M. and Papçun, G. 1972 An auditory-motor theory of speech production, *UCLA Working Papers in Phonetics*, 22: 48-75.
- Löfqvist, A. 1999 Interarticulator phasing, locus equations, and degree of coarticulation. *Journal of the Acoustical Society of America*, 106: 2022-2030.
- Löfqvist, A., Gracco, V. L. and Nye, P. W. 1993 Recording speech movements using magnetometry: One laboratory's experience. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation München (FIPKM)*, 31: 143-162.
- Meyer, E. A. 1910 Untersuchungen über Lautbildung. *Die neueren Sprachen 18 (Festschrift Wilhelm Vietor)*, 166-248.
- Nix, D. A., Papçun, G., Hogden, J. and Zlokarnik, I. 1996 Two cross-linguistic factors underlying tongue shapes for vowels. *Journal of the Acoustical Society of America*, 99: 3707-3718.
- Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I. and Jackson, M. 1992. Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America*, 92: 3078-3096.
- Russell, G. O. 1928. *The vowel*, Ohio State University Press, Columbus.
- Wood, S. 1979. A radiographic analysis of constriction locations for vowels. *Journal of Phonetics* 7: 25-43.