

对立、特征和发音动作

端木三

提要 本文讨论现有特征理论存在的问题，并介绍以发音动作为基础的特征系统。本文提出，现有特征的数量可以减少。本文还讨论有关的理论问题，如复合音的结构，如何表示细微的发音区别，以及特征是内在的还是约定俗成的。

关键词：对立，区别特征，主动发音器官，发音动作，复合音

1. “对立” (contrast) 和 “最小对立组” (minimal pair)

“对立”可以泛指两个听上去不同的词，即两个语音形式不同的词，例见(1)。

(1) 对立组

- a. [mau]-[xei] 猫-黑
- b. [mau]-[kau] 猫-高

虽然两组词都不同，但是可以看出，(1a)的语音区别比(1b)的多。(1a)里三个音都不一样，而(1b)里只有第一个音不一样。有时候“对立”专指“最小对立组”(minimal pair)，即只有一个音不同的一对词。因此，(1a)不是最小对立组，而(1b)是最小对立组。最小对立组也可以叫做“单音对立组”。(2)是两个其它例子。

(2) 最小对立组(单音对立组)

- a. [mau]-[nau] 帽-闹
- b. [mau]-[xau] 帽-号

最小对立组一般指两个词。可是因为其对立是由其中的一对音决定的，比如(2a)的对立是由[m]-[n]决定的，(2b)的对立是由[m]-[x]决定的，所以我们也可以说这一对音有对立功能。

虽然(2a)和(2b)分别只有一对音不同，因此都是最小对立组，可是可以看出，(2a)里[m]-[n]的区别比(2b)里[m]-[x]的小。可见“最小”这个概念，还不够十分清楚，这点我们下面还会讨论。

声调的处理，带来一个新的问题。比如“毛”和“帽”听上去不一样，应该是对立组。可是它们是不是最小对立组呢？如果我们用1、2、3、4表示普通话的四声，那么声调好像是一个基本单位，“毛”和“帽”好像应该算是一个最小对立组，见(3)。

(3) 声调对立：升降调看成基本单位

- [mau2]-[mau4] 毛-帽

可是如果把升降调看成平调的组合，即把升调(如普通话的二声)分析成“低”加“高”，把降调(如普通话的四声)分析成“高”加“低”，情况就不一样了，见例(4)，其中[´]表示高调，[ˋ]表示低调。

- (4) 声调对立：升降调看成平调的组合
[màú]-[máù] 毛-帽

如果双元音算两个元音，而每个元音各取一半声调，那么“毛”和“帽”就不应该算最小对立组，因为它们的韵腹韵尾都不一样，有两对不同的音。当然，如果把双元音看成一个音，或者把韵母看成一个音(游汝杰等, 1980)，或者把声调看成一个单位，那么，声调对立还是可以算最小对立组。

2. “本语有对立”、“他语有对立”、“无语有对立”

一个语言(或方言)所用的对立不一定和其它的语言一样。比如，[l]-[n]在普通话有对立功能，“兰”-“南”不同。而[l]-[n]在成都话无对立功能，“兰”-“南”不分，读[lan]或[nan]都可以。所以，“对立”这个概念跟具体语言(或方言)有关。我们可以区别三种情况，见(5)-(7)。

- (5) 本语有对立：

在一个语言(或方言)里有对立，如[l]-[n]在普通话里有对立功能，可以区别“兰”-“南”。

- (6) 他语有对立：

在一个语言(或方言)里无对立，可是在别的语言(或方言)里有对立。比如，[l]-[n]在在成都话里无对立功能，“兰”-“南”不分，而在普通话里“兰”-“南”不同，有对立。所以，[l]-[n]对成都话来说，是他语有对立。

- (7) 无语有对立：

一种存在的(可以听得出或者测量得出的)语音区别，但是没有任何语言(或方言)用它来作对立(即没有任何语言用它来区别词)。

在描写一个语言(或方言)时，特别是在分析其音位系统或为其设计拼写系统时，一般只包括“本语有对立”的情况。在描写音位变体(allophonic variation)时，可以以“他语有对立”作为基础，即如果是“他语有对立”的变体，应该包括在描写里，而“无语有对立”的变体，可以不用描写。

“无语有对立”的情况很多。比如，男女的音色不一样，可是没有任何语言用这种音色来作对立。一个人累了和不累的时候说话的声音也不一样，可是也没有任何语言用这种音色来作对立。说话的快慢，也很容易听出来，可是也没有任何语言用它来对立词义(元音或辅音的相对长短除外)。还有，两个语言(或方言)之间，看上去是相同的一个音，也可能有细微的不同。比如，根据 Ladefoged (1992: 173-174) 的观察，美国加利福尼亚的

人发 [θ] 这个音时，舌尖一般是伸出来的，而英国南部的人发 [θ] 时，舌尖一般不伸出来。不过，这种细微区别，好像也没有任何语言用来作对立。

3. 区别性特征 (distinctive features)

上面说了，词与词的最小区别，即“最小对立组”，是以音为单位的。现在我们讨论音与音之间的区别。音与音之间的最小区别，一般叫做“区别性特征”或“区别特征”(distinctive features)，有时也叫“语音特征”或简称“特征”。比如，我们说 [m]-[n] 的区别比 [m]-[x] 的小，这是因为 [m]-[n] 之间相同的特征值多，而 [m]-[x] 之间相同的特征值少。具体说，我们可以把 [m]、[n]、[x] 分析成 (8)，其中加号表示符合此特征，减号表示不符合此特征。比如 [+鼻音] 表示鼻音，[-鼻音] 表示非鼻音。

(8)	部位	擦音	塞音	鼻音	有声
[m]	双唇	-	+	+	+
[n]	舌尖	-	+	+	+
[x]	舌根	+	-	-	-

根据这个分析，[m]-[n] 之间只有一个特征有区别(部位的区别)，而 [m]-[x] 之间有五个特征有区别(每项特征都有区别)。

因为特征是区别语音的基础，而语音是区别词的基础，所以特征实际上也是区别词的基础。也就是说，“最小对立组”实际上是以特征为基础的。那么，区别性特征为什么不叫做“对立特征”呢？这可能是因为“最小对立组”这个概念流行得早，区别性特征这个概念当时还没有流行。否则，把区别性特征叫做“对立特征”也未尝不可。还有一个原因，就是音与音的对立比较容易判断，而特征与特征的对立不大容易判断。比如 [mau]-[nau] 是一个音的对立，[mau]-[xau] 也是。可是 [m]-[x] 有多少特征的对立呢？这就没那么容易判断了。在(8)里，[m]-[x] 好像有五个特征不同。可是在普通话里 [m] 并不需要五个特征，只要两个就够了，即 [双唇、鼻音]。同样，[x] 也只需要两个特征就够了，即 [舌根、擦音]。这样，[m]-[x] 之间的特征区别好像就只有两个，即发音“部位”(place)和发音“方式”(manner)的区别。

语言学的历史上有两个重要的里程碑。一个是对音位的发现，即任何语言的任何词都可以用一组元音辅音来代表。这个发现使我们可以为任何语言编写一套拼写系统。第二个里程碑恐怕要算对区别性特征的发现，即任何一个音都可以用一组区别性特征来代表。Ladefoged and Halle (1988:577) 认为，区别性特征是语言学过去的一个世纪里最重要的发现。语音学的专家 (Ladefoged) 和音系学的专家 (Halle) 能有这样的共识是很难得的，可见区别性特征的重要性有多大。

然而，虽然区别性特征这个概念已经提出几十年了，一些根本性的问题仍然缺乏共识。比如，区别性特征的数量应该怎么确定？区别性特征是以什么为基础的(比如以发音为基础，听觉为基础，还是两者兼顾)？区别性特征是先天的(内在的)还是后天的(约定俗成的)？下面我们分别讨论这些问题。

4. 区别性特征的数量

区别性特征的数量到底有多少？这个问题的答案，取决于我们对区别性特征的目的是怎么看待的。一种看法是，区别性特征的目的仅仅是区别语音的“对立”，非对立方面的变化，比如说话的快慢，虽然也很重要，但不属于区别性特征的描写范围，应该分别处理。另一种看法是，区别性特征应该能够描写语音的所有变化，包括非对立性的变化。

根据第一种看法，区别性特征的数量应该有两个要求，即“足够多”和“尽量少”，见(9)。

(9) 以“对立”来确定区别性特征的数量

- a. 足够多：区别性特征应该区别所有的对立，包括“本语有对立”和“他语有对立”。
- b. 尽量少：“无语有对立”的情况不属于区别性特征的描写范围。

“足够多”可以用“本语有对立”和“他语有对立”来定义，即任何一个语言里有的对立都应该区分。“尽量少”是为了理论的精简，可以用“无语有对立”来定义，即任何语言都没有对立的语音区别，不必用特征来表示。

比如，假如我们从普通话开始，我们无疑需要区别 [m]-[n]、[m]-[x]、[n]-[l]，等等。普通话的 [e] 和 [o] 要不要区别呢？它们在普通话里是没有对立功能的，是由介音决定的，如 [je]-[wo] “叶” - “我” 不是最小对立组。可是在成都话里 [e] 和 [o] 是有对立功能的，不由介音决定，如 [pe]-[po] “白” - “剥” 是个最小对立组。所以，[e] 和 [o] 需要有特征来区别。同样，北京话的 [sʰ] 和 [ç] 没有对立功能，比如“香”有人读 [sʰaŋ] (“女国音”)，有人读 [çaŋ]。可是在苏州话里，[sʰ] 和 [ç] 是有对立功能的，[sʰaŋ]-[çaŋ] “箱” - “香”不同，所以，[sʰ] 和 [ç] 需要有特征来区别。

下面考虑另外一种情况。汉语普通话有两个用舌尖的非颚化的清擦音 [s、s̺]，英语也有两个 [θ、s]。如果分别分析这两个语言，那么一个特征[舌尖向前] ([anterior]) 就足以区别它们了，见(10)。

(10)		[+舌尖向前]	[-舌尖向前]
	汉语	s	s̺
	英语	θ	s

根据这种分析，普通话的 [s] 是 [+舌尖向前]，[s̺] 是 [-舌尖向前]。同样，英语的 [θ] 是 [+舌尖向前]，[s] 是 [-舌尖向前]。这也等于说，汉语的 [s] 相当于英语的 [θ]，而汉语的 [s̺] 相当于英语的 [s]。

可是如果我们让人们用语感来鉴别，大家可能会认为汉语的 [s̺] 更像英语的 [s]，不像英语的 [θ]。这种感觉如果很清楚，一点不含糊，那么也可以在特征理论上有所反映。这样的话，要区别汉语的 [s、s̺] 和英语的 [θ、s]，一个特征就不够了，还得加一个。为了行文方便，我们暂且把这个新的特征叫做[舌尖向后]。如果我们排除[+舌尖向前、+舌尖向后]，即舌尖不能同时又向前又向后，那么可以区别三种舌尖部位，见(11)。

- (11) 考虑语言间的对应关系：汉语的 [s] 等于英语的 [s]
 [+舌尖向前] [-舌尖向前]
 [-舌尖向后] [+舌尖向后]
- | | | | | |
|----|---|---|--|---|
| 汉语 | | s | | ʃ |
| 英语 | θ | s | | |

也许有人会说，汉语的 [s] 和英语的 [s]，国际音标都是 [s]，当然应该对应了。可是应当记住，国际音标是有一定的灵活性的，在本语无对立的情况下，对符号的选择有时不止一种。比如，英语的双唇塞音一般用的音标是 [p、b]，可是它们和法语的 [p、b] 很不一样，所以 Iverson and Salmons (1995) 认为英语应该用 [p^h、p] 才对。当然，传统形成了以后，要修改是很不容易的。

还有人可能会说，语言之间的对应关系，恐怕是个程度问题。两个听上去相似的音，并不一定有什么内在的关系。比如，试考虑 (12) 中的三个假想语言(甲、乙、丙)，每个都有两个擦音(A-F)，按发音部位(简称 1、2、3、4 等)从左到右排列。

- (12) 发音部位 → 1 2 3 4 5 6 7 ...
- | | | | | | | | |
|---|--|---|---|---|---|---|---|
| 甲 | | A | | | B | | |
| 乙 | | | C | | | D | |
| 丙 | | | | E | | | F |

如果我们把语言甲和语言乙相比，我们可能会认为 A 对应于 C。如果我们再把语言乙和语言丙相比，我们可能会认为 C 对应于 E。这样我们得出的对应关系是 A-C-E，即 A 对应于 E。可是如果我们直接把语言甲和语言丙相比，我们会认为 B 对应于 E，而不是 A 对应于 E。这就是说，我们从不同的对应方法得到了互相矛盾的结果。这个问题怎样解决呢？

上面的推理有一个假设，即发音部位产生的效果是渐变的，均匀的，就像一把测量长度的尺。可是实际上，发音部位产生的效果，无论从发音的角度，还是从声学效果的角度来看，都不是渐变的或均匀的。比如，从发音的角度来说，“唇”和“唇齿”是两个相邻的部位，可是唇音不用牙齿，而唇齿音要用牙齿，所以发音动作很不一样。同样，“齿间”音 [θ] 和“齿”音 [s] 是两个相邻的部位，可是它们产生的声学效果很不一样：[θ] 的噪音很弱，[s] 的噪音很强。根据 Stevens (1972) 的“突变理论”(quantal theory)，发音部位上有些突变点(quantal points)。如果发音器官越过突变点，就会产生强烈的效果。而在没有突变点的区域，发音器官即使有些前后移动，产生的声学效果也不大。所以，(12) 提出的问题，是一个没有根据的假想，实际上是不会出现的。

上面讨论的是如何决定“足够多”。下面简单讨论一下“尽量少”的情况，举两个例子。国际音标有一个符号 [ɿ]，叫做“无除阻”(unreleased)，即一个塞音只有闭口(成阻和持阻)，没有开口(除阻)，汉语方言入声韵尾的 [p、t、k] 都是这种情况。比如，广东话的“叶”，最后的 [p] 不开口，国际音标应该是 [jɿp]。而英语的 [ʃɪp] “船”，最后的 [p] 开不开口都可以，可以读 [ʃɪp]，也可以读 [ʃɪpɿ]。既然发音上有区别，那么我们需不需要一个特征[无除阻]呢？可是光看发音上的区别还不够，还要看 [p] 和 [pɿ] 有没有对立功能。比如，有没有一个语言可以用 [ʃɪp]-[ʃɪpɿ] 这样的最小对立组来区别两个词呢？如果没

有，那么 [p]-[p̚] 的区别属于“无语有对立”，没有必要用特征来区别。同样，英语的 [k]，在 [ki] “钥匙”里比较靠前，在 [kar] “车”里比较靠后。这样的前后区别是很明显的，需不需要用特征来表示呢？这也要看有没有对立。如果没有任何语言用它来区别词义，那么不必用特征来表示。如果有语言用它来区别词义，那么应该用特征来表示。

还有一个例子，就是韩语的所谓三种清塞音。塞音一般可以分清浊 (voice)，如 [p]-[b]，还可以分是否送气 (aspirated)，如 [p]-[p^h]。所以，印度语有四种塞音，两清两浊，如 [p、p^h、b、b^h]，这样的语言并不奇怪。韩语如果有三种清塞音，这是十分奇怪的，因为现有的特征不够，还得再加一个。这个新的特征，一般叫做[紧] (tense)，音标用 [*] 表示。三种清塞音例见(13)，特征分析见(14)。

(13) 传统分析：区别全在声母

/t^hal/ “面具”
/t*al/ “女儿”
/tal/ “月亮”

(14)

	[t ^h]	[t*]	[t]
送气	+	-	
紧		+	-

Chin-Wu Kim 在 (1965) 报导了韩语的三种清塞音，在语言学界引起了很大的兴趣。可是至今在别的语言里还没有发现类似的情况，所以韩语的现象仍然是一个独例 (Cho 等人 2002)。这使得我们对韩语的语料不得不产生一些疑问，因此最好能重新检查一下前人的推理。Kim and Duanmu (2004) 指出，(13)并不是真正的最小对立组，因为这三个词有声调的区别：“面具”和“女儿”是高调(或降调)，“月亮”是低调(或升调)。也就是说，“女儿”和“月亮”的区别也许不在声母上，而是在韵母的调上，起码我们没有看到足够的理由证明 [t*]-[t] 本身具有对立功能，因此也没有足够的理由增加[紧]这个新特征。没有[紧]这个特征，我们仍然可以区别 (13) 中的三个词，分析见 (15)，其中 [á] 是高调，[à] 是低调。

(15) Kim and Duanmu (2004) 的分析：有的区别在声调

/t^hál/ “面具”
/tál/ “女儿”
/tàl/ “月亮”

一种少有的语言现象，到底不可靠，需不需要增加一个新特征，我们无法事先知道。不过，有了“足够多”和“尽量少”这两个条件，加上对“最小对立组”的严格定义，在多数情况下还是可以判断。

现在我们讨论对区别性特征的第二种看法，即区别性特征不仅应该描写有对立的区别，而且应该描写语音的其它变化，包括“无语有对立”的区别。提倡这个观点的人不少，早先有 Ladefoged (1972)，近来的有 Port and Leary (2006)。比如，Ladefoged (1972)

说, Jakobson 等人 (1952) 提出的区别性特征远远不够用, 因为有很多细微的区别无法表示。下面我们具体讨论一下 Ladefoged (1972) 用来批评 Jakobson 等人的三个例子。

第一个例子见(16)。这两个句子看上去完全由同样的音组成, 可是读起来不一样, 特别是 [pr] 这两个音的读法 (Ladefoged, 1972:279)。

- (16) a. I'm going to get my lamb prepared. “我要做羊肉”
 b. I'm going to get my lamp repaired. “我要修电灯”

既然两句句子有对立, 我们就应该加以区别。可是如果读起来词与词之间没有停顿, 如何区别这两种 [pr] 的读法呢? 如果 [p] 和 [r] 分别用特征来表示, 的确很难分出两种 [pr]。可是如果我们考虑音节结构, 那解决的办法就多了。比如, 根据 Duanmu (2009) 的分析, [pr] 在 prepared 里是个复合音 (complex sound), 占一个时间单位, 而在 lamp repaired 里是两个音, 占两个时间单位。关于复合音, 我们在下面还会讨论。毫无疑问, 区别性特征不是描写语音变化的唯一因素。既然有别的因素, 比如重音和音节结构, 我们就不必指望用区别性特征来解释所有的语音变化。

Ladefoged (1972) 的第二个例子是英语的 [p、b] 和法语的 [p、b] 之间的区别。英语的 [p] 在词首有很强的送气, 而法语的 [p] 没有。英语的 [b] 在词首经常是清音, 而法语的 [b] 一般是浊音。Ladefoged 说得很对, 如果用同样的特征来表示英语和法语的 [p], 或英语和法语的 [b], 那么两个语言之间的区别就无法得到体现。

不过, 这个问题也有解决的办法。关键不在于 [p] 在英语和法语里有不同的体现, 而是英语的 [p] 和法语的 [p] 根本就不是同一个音。根据 Iverson and Salmons (1995) 的分析, 英语的 [p] 应该是 [p^h], 而法语的 [b] 应该是 [p], 具体见(17)。

(17)	清送气	清不送气	浊
英语	[p ^h]	[p]	
	pat 拍	bat 蝙蝠	
法语		[p]	[b]
		peau 皮	beau 美

这样的分析有很多好处。首先, 它修改了前人分析英语的一个错误。第二, 它在语音上更精确。第三, 它解释了语言之间的差异。

Ladefoged 的第三个例子, 是一个特征在不同的音里有不同程度的声学体现。比如, [+有声]([+voice]) 这个特征, 应该指声带的振动。在元音里这样的振动很明显, 在擦音里这样的振动要少得多, 而在塞音里这样的振动就更少了。Ladefoged 指出, 要精确地描写语音的声学或物理性质, 只靠区别性特征是不够的, 还应该具体的规则, 把特征转换成物理数据。Ladefoged 的这个说法当然是对的。比如, 要区别一个人和另外一个人的发音, 只靠区别性特征肯定是不够的。

Ladefoged 的结论是什么呢? 很显然, 要想解决这么多的问题, 光靠区别性特征是不够的。实际上, Ladefoged 在 (1972: 282) 和 (2001: 170) 都明确指出, 连音段 (segment) 这个概念都应该抛弃, 取而代之的是以音节为基本单位。既然区别性特征是描写音段的, 一旦抛弃了音段, 区别性特征当然也没有存在的价值了, 这就是 Ladefoged 的结论。

这个结论我们认为有点仓促。比如，上面的三个例子，有两个都有解决办法。而且，第三个例子的要求似乎过高。我们对“有对立”的现象还不能说已经完全理解了，就要进一步去处理一切“无对立”的现象，恐怕操之过急。还有，Ladefoged 本人并不专门研究音节，可能不完全了解音节理论里同样存在很多问题。比如，对有人音节研究了多年，最后却提出要完全抛弃音节这个概念(如 Steriade, 1999, Blevins 2003)。所以，Ladefoged 关于抛弃音段的看法只能说是一种设想。

以上两种对特征的看法，还可以用另外一个例子来阐述。假如我们研究人的移动方式，我们可以区别走和跑(也许还有爬)。这两个动作是很不一样的，是人人都能做的，而且人人都能区分的，即要么是走，要么是跑。可是走和跑这两个概念，能不能足够表达人与人之间的不同的移动速度呢？答案很清楚，这两个概念肯定是不够的。我们最多可以说，走路一般比跑步要慢，可是人与人之间的区别，或者男女之间的区别，我们无法表达，还需要更精确的描写手段，比如引进速度这个概念。有了速度这个概念，我们可以区别人与人之间的快慢，男女之间的快慢，同一个人上坡时和下坡时的快慢，精力充沛时和疲倦时的快慢，等等。那么，既然走和跑这两个概念有不足，而速度这个概念又那么精确，我们是不是应该抛弃走和跑这两个概念，只用速度这个概念呢？答案应该是否定的：走和跑这两个概念的目的不是表达速度，而是表达两种不同的行为方式，而速度不能直接表达这两种行为方式。走和跑虽然和速度有一定关系，但是影响速度的有很多因素，我们应该把这些因素分别处理。

回到区别性特征这个问题上，“对立”是一种行为上的区别。如果一个人在发音上区别 [θ] 和 [s]，那么肯定是有意识地产生了两个不同的行为，所以我们应该有特征来表达这些不同行为。一个行为的物理效果如何，是另外一回事，要受到很多因素影响，行为只是其中的一个因素。比如一个人累的时候说的 [s]，和不累的时候说的 [s]，应该在行为上有个共同之处，这正是特征所要表达的。这两个 [s] 的不同，可能和其它因素有关系。如果我们对行为的物理效果感兴趣，我们当然应该考虑各种因素，包括行为以外的因素。但是，研究行为的种类(比如走和跑)和研究行为的物理效果(比如速度)，不完全是一回事，应该有所区别。

当然，Ladefoged 有一个看法是合理的，即区别性特征不应该太抽象，而应该尽量有发音上的，或物理上的性质。这个看法我们应该尽量采用。

以上的讨论主要是从“对立”的角度出发来看我们需要多少特征。还有一种角度，是考虑“自然音类”(natural classes)的种类。一个自然音类一般指受到一条语音规则影响的一组音，或者是能够触发一条语音规则的一组音。比如，德语有一条浊音清化规则，所有受它影响而产生变化的音就是一个自然音类。又比如英语的名词复数后缀有三个形式，根据它前面是什么音而定：前面是 [s、z、ʃ、ʒ、tʃ、dʒ] 时用 [ɪz]，前面是 [f、θ、p、t、k] 时用 [s]，其它情况用 [z]。人们一般认为，[s、z、ʃ、ʒ、tʃ、dʒ] 应该是个自然音类，因为它们可以触发 [ɪz]。[f、θ、p、t、k] 也好像是个自然音类，因为它们可以触发 [s]。剩下的其它音也许是个自然音类，也许不是。一个自然音类里的音，起码应该有一个共同的特征，有时候有两个或更多的共同特征。比如，英语 [s、z、ʃ、ʒ、tʃ、dʒ] 的共同特征好像是[非齿，舌尖，擦音](具体特征名称见下面的讨论)。[f、θ、p、t、k] 的共同特征是[清音]。不过，因为自然音类牵涉到语音规则，而语音规则有时候相当复杂，我们这里无法进行更多的讨论。所以，下面我们主要从对立的角度考虑特征的选择。

5. 特征和发音动作

区别性特征是以什么为基础的？是以发音为基础，还是以声学性质为基础，还是以听觉性质为基础，还是还是各种都兼顾？国际音标传统上是根据发音来定义特征的。

Jakobson 等人 (1952) 的特征既用声学的定义，又用发音的定义。Chomsky and Halle (1968) 基本上用的是发音的定义。Ladefoged (1992) 和 (2006) 认为元音特征应该用听觉性质来定义，而辅音特征主要用发音方法来定义。

无论那种分析，恐怕都不能完全不用发音的定义。这是因为语音毕竟是靠发音来产生的。而且，如果靠发音的定义能够解决所有有关的区别，那么就没有必要增加其它的定义方法。下面我们只讨论用发音的定义。

国际音标把辅音分析成发音部位 (places of articulation)、发音方法 (manners of articulation)，和清浊 (voicing)。这可以看成是一种区别特征系统。而且，发音部位、发音方法和清浊都是发音概念。可是严格讲，发音是一种动作。国际音标的特征，是不是都和动作有关呢？下面我们考虑几个具体情况。

国际音标有个发音方法叫“鼻音”，可是这不是一个动作。鼻音实际上是一个效果，即气流通过鼻腔。真正的动作，应该是软腭下降，发音器官是软腭。

国际音标的“清音”或“浊音”，也不是动作，而是效果，即声带无振动，或声带有振动。真正的动作，按照 Halle and Stevens (1971) 的理论，是声带收紧(难以振动)，或声带放松(便于振动)。所以，这里的发音器官是声带，动作是收紧与放松。

国际音标有个发音部位叫“齿” (dental)。这是个什么动作呢？牙齿本身是不会动的。“齿”实际上指的是将舌尖移到齿后(或齿间)，真正产生动作的是舌尖。同样，国际音标有个发音部位叫“齿龈” (alveolar)。齿龈本身也是不会动的，“齿龈”实际上指的是将舌尖移到齿龈，所以真正产生动作的也是舌尖。以上的“齿”和“齿龈”一般叫“被动发音器官” (passive articulator)，而“舌尖”叫“主动发音器官” (active articulator)。

朱晓农 (2006, 12-13) 认为，“发音器官”这个概念不大精确，应该重新命名为“调音器官”。其原因是，在一般情况下，声音的来源 (source) 是声带，而声带以上的器官只是通过改变口腔形状，利用共振原理，来调节声源里的不同频率成分。朱晓农的说法有一定道理，不过不够全面。元音的声源，的确全是来自声带，其它的器官的确只起调节作用。可是辅音的情况不完全一样。有的辅音，比如近音 (approximants) 和鼻音，情况和元音相似。而其它辅音有噪音成分，其声源多数不来自声带，而是来自有关的发音器官。比如 [f] 的声源是噪音，是由气流通过唇齿的开口时产生的，所以唇齿应该叫做发音器官，不应该叫做调音器官。[v] 这个音既有来自唇齿的噪音成分，又有来自声带的周期性成分 (periodic sound)，所以唇齿有双重作用，又是发音器官(发噪音)，又是调音器官(调节声带来的周期性成分)。为了简便，本文仍然使用发音器官，不用调音器官。

国际音标的发音部位基本上都是由一个主动发音器官和一个被动发音器官组成的，只不过主动发音器官往往被省略了，这点麦耘 (2005) 已经仔细谈过。那么，被省略的发音器官，是不是总是很清楚，根本就没有必要明说呢？看来情况不是这样。比如“齿龈后”这个发音部位，是用什么主动发音器官去靠近？是“舌尖”？还是“舌叶”？还是“舌面”？还是“舌尖”加“舌面”？而且，舌尖，舌叶，舌面的区别到底在哪儿？每个器官是不是还可以分“前后”呢 (如“舌尖前”、“舌尖后”、“舌面前”、“舌面后”等)？这些问题，实际上都不清楚，都应该有个明确的阐述。

所以，国际音标作为一个特征系统，虽然是以发音为基础的，可是对发音“动作”的表达，往往只是间接的，有时甚至是含糊的，这点 Ladefoged and Halle (1988) 已经作过批评。我们所需要的，是一个明确以主动发音器官和发音动作为基础的特征系统。这个新系统的两个核心观点阐述于 (18)。

- (18) 以发音动作为基础的特征系统:
- a. 每个特征就是一个发音动作。
 - b. 每个发音动作都应该有一个主动发音器官。

要采用这个系统，我们首先需要确定发音器官的数量。其次，我们需要确定每个发音器官能够作哪些动作。下面我们具体讨论这个系统。

6. 以发音动作为基础的特征系统

我们首先讨论主动发音器官(简称发音器官)。根据 Sagey (1986)、Halle and Ladefoged (1988)、Halle (2003) 等人的研究，六个发音器官就足够了，见 (19)。

- (19) 六个主动发音器官
- | | |
|------|-------------|
| (下)唇 | Labial |
| 舌尖 | Coronal |
| 舌体 | Dorsal |
| 软腭 | Soft-palate |
| 舌根 | Tongue-root |
| 声带 | Vocal-cords |

本文把唇音的主动发音器官列为“下唇”，是因为在发唇齿音时(如发 [f] 时)，主动发音器官是下唇在移动(向后移动)。发双唇音时(如发 [p] 时)，我们可以认为上下唇同时闭合，两者都是主动发音器官。我们也可以认为是下唇往上与上唇闭合，这样的话主动发音器官还是下唇。所以，下唇可以把双唇音和唇齿音的主动发音器官统一起来。

这六个发音器官够不够，要看每个发音器官有多少个发音动作。关于发音动作，有两种观点。一种认为，每个发音器官只有固定的几个发音动作，比如，元音的高度最多只有高中低三个舌体动作 (Sagey 1986; Halle 2003 等)。另一种观点认为，每个发音器官没有固定的发音动作，比如，元音的高度没有固定的数量和位置，各个语言可以自定 (Ladefoged 1972; 1992)。这个争论我们下面还会讨论。现在我们先假设前一种观点，即每个发音器官只有固定的几个发音动作，因为这个理论更加简单。

有的动作不限于某个发音器官，Halle (2003) 将其称为 *articulator-free*。有的动作是局限于具体发音器官的，Halle (2003) 将其称为 *articulator-bound*。局限于具体发音器官的动作见 (20)。按照习惯，我们把特征(即发音动作)放在方括号里，而发音器官不加括号。对于每个特征，我们按传统用加减号表示两个不同的动作。比如，软腭的[+向上]表示升，[-向上]表示降。

(20) 局限于具体发音器官的动作

发音器官	发音动作	传统术语	举例
(下)唇	[+向前]	双唇	[w]
	[-向前]	唇齿	[v]
	[+圆唇]	圆唇	[p ^w]
	[-圆唇]	非圆唇	[p]
舌尖	[+向前]	齿间	[θ]
	[-向前, -向后]	齿(齧)	[s]
	[+向后]	卷舌	[ʃ]
	[+边音]	边音	[l]
	[-边音]	非边音	[t]
舌体	[+向前]	前(元音)	[i]
	[-向前]	后(元音)	[u]
	[+向上]	高(元音)	[i]
	[-向上, -向下]	中(元音)	[e]
	[+向下]	底(元音)	[æ]
软顎	[+向上]	非鼻音	[b]
	[-向上]	鼻音	[m]
舌根	[+向前]	紧(元音)	[e]
	[-向前]	松(元音)	[ɛ]
声带	[+收紧]	清音	[s]
	[-收紧]	浊音	[z]
	[+分开]	送气	[p ^h]
	[-分开, -闭合]	不送气	[p]
	[+闭合]	喉塞音	[ʔ]

特征值应不应该用加减号来区别, 实际上没有统一看法。比如, Ladefoged (2007) 就不用加减号, 而直接用不同的术语表达加减号的意思。比如, 他不用 [+voice] 和 [-voice], 直接用 [voice] 和 [voiceless]。舌尖的动作他不用 [+向前], [-向前, -向后], [+向后], 而直接用 [齿], [齧], [卷舌]。我们用加减号有两个原因。第一, 如果是谈动作, 实际上就是指肌肉的运动, 应该是要么有, 要么无。第二, 加减号可以表示, 哪些动作是相反的一对, 只能选择两者其一。比如, 舌体的 [向上], 只能要么 [+向上] (高元音), 要么 [-向上] (中元音或低元音)。如果直接用 [高] 和 [中], 那么人们也许会问, 有没有 [高中] 元音或 [中高] 元音呢?

特征术语的选择也有一定的任意性。比如, 前元音可以用 [+向前], 也可以用 [-向后], 鼻音可以用软顎 [-向上], 也可以用 [+向下]。因为对肌肉的运动还不是十分清楚,

本文对这类术语随便做了一些选择。还有，上面的表格并不全面，还缺一些特征。比如，内爆音 (implosives) 可能还应该增加声带的 [向下]。而且，声带还应该负责声调特征。不过我们暂时不讨论这些特殊情况。

上面可以看到，有的动作好像反复出现在不同的发音器官里。比如，[向前]在唇，舌尖，舌体，舌根里都有。这是不是说，这些动作有什么共性呢？比如，唇的[向前]，舌尖的[向前]，舌体的[向前]，和舌根的[向前]，都是[向前]，那么 [w、θ、i、ε] 是不是一组“自然音类” (natural class) 呢？可是好像没有这样的证据。实际上，以前的特征系统为了避免这种无意的“共性”，往往采用不同的术语来表达“向前”这个意思。比如，Halle (2003) 用 [+anterior] 指舌尖的[向前]，用 [-back] 指舌体的[向前]，用 [+advanced] 指舌根的[向前]。本文为了方便，没有专门采用不同的术语来表达基本相同的意思(而且中文里好像也不容易找出那么多的不同术语)。但愿只要作出了说明，以上的发音动作都是局限于各个不同发音器官的，互相之间没有关系，就不会引起误解。

现在我们讨论不局限于具体发音器官的动作。这些动作有时也叫做发音方式特征 (manner features)，主要分类特征 (major class features)，或闭合度特征 (stricture features)。前人提出过好几个这样的特征，比如 Halle (2003) 用了五个：辅音([consonantal])，响音([sonorant])，延续音([continuant])，强噪音([strident])，边音([lateral])。本文认为，只要两个就够了，见 (21)。

(21) 不局限于具体发音器官的动作

[塞] (stop)

[擦] (fricative)

这两个特征可以排列出四种情况，也就是四种传统的分类，见(22)。

(22)	传统分类:	塞音	擦音	塞擦音	近音
		(stop)	(fricative)	(affricate)	(approximant)
	特征表达:	[+塞、-擦]	[-塞、+擦]	[+塞、+擦]	[-塞、-擦]

其它的一些传统分类也可以表达，不必再加其它特征，见(23)。

(23)	传统分类	本文分析
	鼻音 (nasal)	软顎是近音(鼻腔的近音)
	响音 (sonorant)	口腔或鼻腔的近音
	近音 (approximant)	口腔近音
	流音 (liquid)	舌尖是近音
	元音 (vowel)	舌体是近音
	介音 (glide)	不在韵腹的高元音

按照传统的分析，元音和近音不是一类：近音的开口度比元音略小。可是从对比的角度来看，元音和近音没有对比。常用的有关对比是介音和高元音的区别，如[w]是近音，[u]是高元音，除此以外其它特征完全一样。可是介音和高元音的区别可以在音节的位置

上体现：元音在韵腹(和韵尾)出现，介音在其它位置上出现。所以，没有必要用特征来区别介音和高元音，或近音和元音。

下面我们讨论一下 [擦] 和 [塞] 到底是什么动作。如果 [+擦] 理解为口腔有些开口，而 [+塞] 理解为口腔无开口，那么 [+塞、+擦] 这对动作严格讲是有矛盾的，因为口腔不能同时既有开口，又无开口。能不能把 [+塞、+擦] 理解为有前后时间顺序，即先 [塞] 后 [擦] 呢？可是正如 Duanmu (1994) 指出，这样会引起新的问题，即引出很多不存在的音，使得理论有太多的超量预测 (over-prediction)。比如，假如我们有 10 个特征，每个有加减两个值，那么，如果时间不能分先后，可能组合出来的音是二的十次方，即 1024 个音，这个数基本上符合世界上已知语言的元音辅音的总数量 (Maddieson, 1984, Ladefoged 2001, Mielke 2008)。可是如果时间可以分一先一后，那么可能组合出来的音是 1024 的二次方，即 1048576 个音。显然，后者远远超过自然语言中实际使用的语音的数量。

那么如何理解塞擦音呢？Duanmu (2009) 认为，关键在于重新定义 [擦] 这个概念。[擦] 的传统定义是，口腔的开口很小，以至使气流产生湍流和噪音(林焘、王理嘉 1992, Ladefoged 2006)；这个定义强调的是口腔有开口。Duanmu (2009) 从发音动作来定义，认为 [擦] 的发音动作是 [边阻碍] ([edge closure])，即发音器官的左右两边上下合拢。比如发 [s] 时，舌尖(叶)的左右两边向上关闭，形成边高中低的槽型。如果只发擦音，那么发音器官的中间没有关闭，有个小开口。如果发塞擦音，那么同时有 [塞] 和 [擦] 两个动作。[塞] 的动作将口腔的中央关闭，而 [擦] 的动作增加了左右两边的关闭。开口(除阻)时，发音器官仍然暂时是边高中低的形状，所以产生出短暂的擦音效果。换句话说，[擦] 强调的不是中间有开口，而是[边阻碍]，中间是否开口与其无关，所以 [+塞、+擦] 并无矛盾。

7. 一个音有多少特征需要赋值？

有的特征或特征搭配，在有的语言里是不用的。比如，缅甸语的鼻音要分清浊，而普通话的不分。我们给缅甸语的鼻音标特征时，应该包括清浊，这是无疑的。普通话的鼻音一般都是浊的，那么我们应不应该包括“浊”这个特征呢？有了好像更准确，可是又显得多余。还有，普通话的 [p、p^h]，可以用“送气”来区别，这个特征应该包括。可是，如果两者都是清音，要不要再给它们加“清”这个特征呢？加了，又显得多余。而且，[p] 经常还可以以浊音的形式出现，把它说成“清”也不太合适。

研究特征的赋值程度的理论，叫“非充分赋值理论”(underspecification theory)。一般都认为，“本语有对立”的特征要赋值，“本语无对立”的特征不赋值 (Steriade, 1987, Archangeli 1988)。而且，Keating (1988b) 进一步指出，“本语无对立”的特征，在表层结构(即具体发音时)也往往是空白的，完全不用填入。当然，某个语言里哪个特征有对立性，哪个没有，有时也不容易确定。比如，英语的 [p、b]，区别在于“清浊”还是在于“送气”？有关争论可见 Iverson and Salmons (1995)。限于篇幅，本文不讨论这些细节。本文只是遵照“非充分赋值理论”的大原则，尽量省略“本语无对立”的特征。

如果我们省略“本语无对立”的特征，那么一个音的特征表示是很简略的。下面我们举几个普通话的例子，见(24)。每个音我们列出有关发音器官，以及每个器官的动作。无关的器官和动作一概不列。

(24) 普通话辅音元音举例

- [n] 舌尖— [+塞]
软顎— [-向上] (鼻音)
- [k] 舌体— [+塞]
声带— [-分开] (不送气)
- [k^h] 舌体— [+塞]
声带— [+分开] (送气)
- [a] 舌体— [+向下] (低元音)

鼻音在普通话里无清浊对立，所以 [n] 不用标声带特征。[k、k^h] 用“送气”区分就够了，不用加“清浊”，发音器官也只用舌体和声带，不用列别的。最简单的是普通话的 [a]。因为普通话只有一个低元音，所以不用标“前后”或“圆唇”，标了舌体 [+向下] 就够了，不会和任何其它音混淆。不给 [a] 标“前后”，实际上也给 [a] 打开了一扇大门，让它可以根据语音环境有前后自由移动的能力。

8. 复合音

复合音是相对于简单音而言。比如 [d] 是简单音，而 [d^w] 可以认为是复合音，因为多了个 [w] 的成分。[a] 是简单音，而 [ã] 可以认为是复合音，因为多了个鼻音的成分。

不过，复合音和简单音的区别很难严格定义。比如，汉语的 [y] 是不是个复合音？如果认为 [y] 是 [i] 加 [u] 的组合，那么 [y] 是个复合音。可是，把 [y] 看成 [i] 加 [u] 有没有足够的根据呢？还有，[o] 是不是复合音，能不能看成是 [ə] 加圆唇？[m] 是不是复合音，能不能看成是 [b] 加鼻音？[o] 和 [m] 很多语言都有，发起来好像也没什么麻烦，而且国际音标也是一个符号，没有附加符号，分析成复合音，是不是合理呢？

下面我们出于方便，给复合音一个简单的定义，见(25)。

(25) 复合音的定义

- a. 使用两个(或更多)的发音器官(不算软顎和声带)
- b. 比一个简单音多出一个(或更多)特征

不算软顎，是为了可以把 [m、n] 等算作简单音。不算声带，是因为大多数音都有声带的动作，即清浊特征，如 [p、b、t、d] 等，我们不想把它们都叫做复合音。(25b) 是为了把 [ã] 这样的音算作复合音。因为 [ã] 只比 [a] 多了一个软顎特征 (鼻音)，如果软顎不算，[ã] 就会被分析成简单音。

复合音的定义不管怎么下，其关键是，一个音可以由两个(或更多)发音器官来产生。这个概念非常重要。试考虑 (26) 的例子。为了强调它们是复合音，我们把第二个音符使用上标。

(26) 复合音举例

[k^w] 舌体— [+塞]
唇— [+圆唇]
声带— [+收紧] (清)

[k^l] 舌体— [+塞]
舌尖— [+边音]
声带— [+收紧] (清)

[k^s] 舌体— [+塞]
舌尖— [+擦]
声带— [+收紧] (清)

[k^w] 是个常见例子，除声带以外有唇和舌体两个发音器官，有时叫做“双重发音” (double articulation)。需要指出的是，因为声带只能取一个清浊特征，[k^w] 中的 [w] 应该是个清音，不是浊音。英语的 [kw] 可以看成是 [k^w]，其中的 [w] 的确是清音 (Ladefoged 2006)。

[k^l] 这样的结构，往往看成是两个音。可是因为 [k] 和 [l] 用的是不同的发音器官，两组发音动作可以同时进行，形成一个复合音，这和 [k^w] 的情况是类似的。英语的 [kl] 也可以看成是 [k^l]，其中的 [l] 也是清音 (Ladefoged 2006)。

[k^s] 是个塞擦音。一般的塞擦音是 [+塞, +擦] 在同一个发音器官，如 [ts、dz] (或 [t^s、d^z])，而 [k^s] 的 [+塞] 和 [+擦] 在不同的发音器官。有的特征理论，如 (Sagey 1986)，排除了这种塞擦音的可能。有的语言学家却认为应该承认这样的塞擦音 (如 Prinz and Wiese 1991, Wiese 1996)。从发音动作的角度来看，[k^s] 这种复合音没什么奇怪。同样的例子还有 [p^s]，[t^x]，等。

上面的讨论里，我们还没有提到过颞音 (palatals)。现在我们讨论颞音的分析。有了复合音的概念，我们可以把颞音分析成舌尖和舌体的结合，不必再增加 [颞音] ([palatal]) 或 [延伸闭合] ([distributed]) 这样的特征。这种想法 Keating (1988a) 曾经提出过。为了简化讨论，我们这里只考虑非塞音。可能的情况见(27)，声带特征省略不写。

(27) 舌尖、舌体、颞部的非塞音

舌尖	+擦	+擦	+擦	-擦	-擦
舌体		+擦	-擦	+擦	+擦
国际音标	[s]	[x]	[s ^j]	[ç]	[x ^j]/[ç]

[s] 只用舌尖，[x] 只用舌体。如果舌尖舌体都用，就产生四种颞音或颞化音，其中的 [x^j] 也就是 [ç]。如果我们再区别舌尖是否[向后]，又会多出四个音，见(28)。

(28) 颚部的非塞音

舌尖	+擦	+擦	-擦	-擦	+擦	+擦	-擦	-擦
	-向后	-向后	-向后	-向后	+向后	+向后	+向后	+向后
舌体	-擦	+擦	+擦	-擦	-擦	+擦	+擦	-擦
国际音标	[sʲ]	[ç]	[xʲ]	[j]	[ʃʲ]	[ʃ]	[xʲ]	[rʲ]

如果舌尖、舌体的下面分别再考虑[+塞]，包括舌尖舌体取不同的[塞]值，那么情况就更多了。比如，如果我们考虑鼻音，把[擦]换成[塞]，那么[sʲ]就对应于[nʲ]，[ç]对应于[n̥]，[xʲ]对应于[n̥ʲ] (或 [n̥])。限于篇幅，这里我们不能一一举例。

麦耘(2005)说，苏州话的[s]在一般情况下是舌尖音，可是在介音[i]前，如在[sian]“箱”里，是“舌叶”音。Duanmu(2009)认为，在[sian]里，[si]实际上是个复合音，即上面的[sʲ]，其中的[s]已经不是单独的音了。而所谓的舌叶音，也就是本文颚音的一种。在谈到发音部位的术语时，麦耘还提出，使用被动发音器官比使用主动发音器官更好，更清楚。麦耘的看法，好像和本文的建议有矛盾，因为本文恰恰是要强调主动发音器官。不过，麦耘的说法实际上没错：他指的是，如果只用一套发音器官来表示发音部位的话，那么被动器官总的来说比主动器官清楚。本文的建议，不光是要强调主动发音器官，而且还要考虑每个发音器官的动作，这也就涵盖了不同的被动发音部位。

Recasens(1990)曾经批评Keating(1988a)，说Keating的分析所包括的颚音太少。Recasens提出，颚音不应该分析成由舌尖和舌体组成的复合音，而应该独立分类。而且还应该把颚音细分成四个小类，见(29)。

(29) Recasens(1990)的四类颚音

类名	举例
龈后(post-alveolar)	[ʃ]
龈颚(alveolopalatals)	[ɲ]
前颚(front palatal)	[j]
后颚(back palatal)	[kʲ]

Recasens的[ɲ]，可能是指[nʲ]，相当于我们的[sʲ]例。Recasens的[kʲ]，可能相当于我们的[xʲ]例。我们的例子说明，复合音能代表的颚音是很多的，区别是很细微的，远远超过Keating和Recasens的分析。

9. 发音动作是天生的还是约定俗成的？

发音动作(或特征)是天生的还是约定俗成的，这是语言学界经常争论的一个问题。天生的动作，一般的含义是指生来(或长到一定程度)就能作这个动作，不管有没有人教，就像鸭子生来就会游泳，人长大后都会走，会跑。约定俗成的动作，一般的含义是指生来不会作这个动作，或者不主动作，要跟别人学了才做，比如人生来不会做饭，写字，或倒退着走路，长大了也不自动会(或主动作)，要学了才会。不过，上述两种观点有时候很难区别。Mielke(2008)对有关争论作过一些详细的讨论。下面我们只是简单谈谈地这个问题。

首先，恐怕所有的人都同意，有的动作是天生的，比如哭，笑，走，跑等。而且，所有的人都认为，有的动作不是天生的，是学的，比如写字，僵着腿跨步行走，抬腿 90 度地行走，举臂 90 度地行走，等等。游泳，对大多数人来说，可能也不是天生的。

所有的人恐怕也都同意，有的动作，既不完全是天生的，也不完全是学来的。比如跳芭蕾舞，人天生有能力站立，抬腿，举手，保持平衡等。需要学习的是，如何把这些能力进行搭配，或者提高。又比如游泳，人天生有能力控制呼吸，控制手脚的运动，理解水的浮力和对手脚运动的反作用力等。需要学习的是，如何协调这些动作，如何利用水的浮力和反作用力。

再回到行走这个例子上来。大家都同意，很多行走方式都不是天生的，如倒退着走，抬腿 90 度地走，举臂 90 度地走，双臂不动地走，体育上的竞走，等等。这些行走方式，可能是从人的基本走路方式上改变出来的，也可能是从人的总的平衡能力和手脚控制能力上发展出来的。那么，既然承认人有总的平衡能力和手脚控制能力，能不能说，就连一般的走路和跑步，也不是天生的，也是从基本的平衡能力和手脚控制能力上发展出来的，和其它的行走方式没什么不同呢？这种看法，忽略了自然的走路和跑步与其它移动方式之间的一个根本区别，即前者是现成的，不用学，而其它移动方式不是现成的，一般需要学习才会。

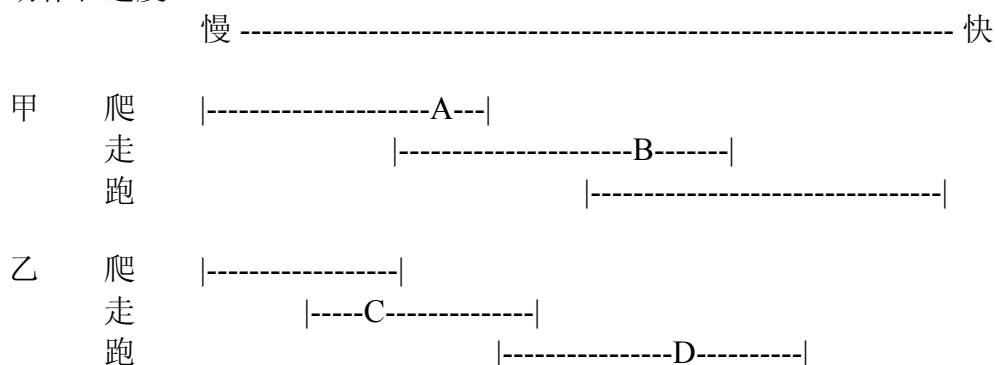
回到发音动作上，我们先考虑舌位高低。一种观点是，舌体有两个天生的基本动作，即或向上，或向下。向上是高元音，向下是低元音，不上不下是中元音。这当然不是说舌体不能产生其它的高低，而是说三种高低是最自然的，最简单的动作，不用学，而其它高低不是最自然的动作，要学，要下功夫去控制。第二种观点是，舌体的上下没有固定动作，可以自由调节，产生无限的高低变化，或延续的变化。国际音标的元音表，一般认为舌位的高度是延续的，起码是不止三等。比如，传统的“基本元音位” (cardinal vowels) 就把舌位高度分成四等。而一般的国际音标表，经常不把某个语言的元音放在某个固定的高度，而是放在任何看起来合适的高度。我们把两种观点概括于 (30) 和 (31)。

- (30) 特征是自然的(天生的)发音动作
- a. 每个发音器官的自然动作是有限的。
 - b. 语音特征一般只使用自然动作。
 - c. 不同语言之间的特征有本质上的对应关系。
- (31) 特征是学来的(非天生)发音动作
- a. 每个发音器官的动作是无限的。
 - b. 语音特征不是(或不必要是)自然动作。
 - c. 不同语言之间的特征没有本质上的对应关系。

我们已经讨论了 (a) 和 (b) 这两点。下面我们举个例子解释一下 (c)。如果我们考虑人的三种自然运动方式跑、走、爬，不同人之间的速度很不一样。有的人走得比别的人跑得还快。同一个人的三种运动的速度范围也有交叉，有时跑得比走得还慢。理论上，也不能排除有的人可能爬得比有的人跑得还快。可是跑就是跑，走就是走，爬就是爬，不管速度怎样，我们还是能够确定它们之间的本质上的对应关系。如果我们不从动作着眼，而从

速度着眼，那么不同动作之间的对应关系就很难说了。试考虑 (32) 中的情况。其中甲乙是两个人，点线表示速度域，ABCD 是四个速度点。

(32) 动作和速度



如果我们看动作，那么甲的 B 应该对应于乙的 C，两者都是“走”，不管走得快不快。A 和 C 虽然速度相近，但是动作上没有对应关系。同样，B 和 D 虽然速度相近，动作上也没有对应关系。如果我们只看速度，那么甲的 A 基本对应于乙的 C，甲的 B 基本对应于乙的 D。可是 A 和 C 之间，除了速度相近以外，没有本质上的对应关系。同样，B 和 D 之间，除了速度相近以外，也没有本质上的对应关系。

如果我们考虑舌位的高中低三个动作，那么情况也很相似。每个舌位动作有一定的上下范围。三个上下范围之间有一定的交叉。每个人对三个范围的划分不会完全一样。如果我们从发音动作着眼，那么不管人与人之间有什么不同，舌体向上就是向上，向下就是向下，不动就是不动，本质上的对应关系应该是有的，而且可能十分清楚。可是如果我们不从动作着眼，而只从舌位的高度着眼，那么不同动作时间的对应关系就很难说了。两个动作之间，除了高度相不相近以外，很难说有什么本质的对应关系。

以上的两种理论，有没有什么办法用事实来进行验证和选择呢？发音动作是否是自然动作恐怕不容易直接验证。不同语言之间的音有没有什么本质上的对应关系，恐怕也不容易验证。比较容易验证的是每个器官的有对立的发音动作的数量。如果发音动作的数量确实十分有限，远远少于每个器官可以作出的动作，那么发音动作就可能是天生的，自然的。比如，人人可能都会唱歌(或能学会唱歌)，而且一般起码可以掌握十个音阶。可是在自然语言里，声调的高低远远少于十个。同样，舌位的高低可以自由上下，可是自然语言的元音舌位，如果排除“松紧”(tense)的影响，最多分三个高低。从这些事实来看，发音动作有可能的确是天生的自然动作。

下面我们再讨论一个理论，即 Stevens (1972) 的“突变理论”(quantal theory)。这个理论的基本观点是，人的发音部位上有一些“突变点”(quantal positions)，当主动发音器官越过这些突变点时，口腔形状的变化会引起语音的突然变化。当主动发音器官不越过突变点时，口腔形状的变化不会引起很大的语音变化。因此，突变点为语音特征建立了一些自然分界线。根据这个理论，语音特征在很大程度上是天生的，特征的数量往往是有限的，而且不同语言之间的特征是有对应关系的。不过，天生的成分主要不在于天生的动作，而在于天生的口腔形状，特别是一些固定结构的形状，以及它们产生的声学效果。当然，从理论上讲，不能排除既有天生的固定结构，也有天生的发音动作，可是突变理论很少讨论后者。

突变理论的一个关键假设是语音的变化有大小。而它的一个弱点是没有提出严格的标准来定义语音变化的大小。比如，音调高低的变化多少才算大？如果很多人都能唱歌，都能起码区别十几个音阶，说明音调的频率只要变一点就够大了。可是为什么自然语言的声调高低远远少于十几个等级呢？

总之，发音动作是天生的还是约定俗成的这个问题，题目太大，暂时不能作出结论。不过，它不影响我们对特征的讨论。而且，一旦特征理论清楚了，可能还有助于解决这个争议。

10. 对立特征为什么可以有不同的体现？

一个对立特征的物质体现经常会有很多变化，这是一个常见的事实。我们具体讨论三种不同的变化，见(33)。每一种变化都可以举出很多例子。在(34)里我们各举一个。

(33) 对立特征的各种变化

- a. 同一个特征在音与音之间的变化。
- b. 同一个特征在人与人之间的变化。
- c. 同一个特征在人群与人群之间(语言与语言之间)的变化。

(34) 对立特征的变化举例

- a. “浊”这个特征体现为声带振动，可是在塞音里振动很少，在擦音里振动稍多，在元音里振动最多。
- b. “声调”这个特征体现为声带振动的频率，可是同一个声调不同的人发出来的频率各不一样。
- c. 美国加利福尼亚的人发 [θ] 这个音时，舌尖一般是伸出来的，而英国南部的人发 [θ] 的时候，舌尖一般不伸出来 (例子来自 Ladefoged 1992: 173-174)。

同一个特征，如果有不同的变化，即不同的语音表达形式，是不是说特征理论有缺陷呢？Ladefoged (1972; 1992) 认为正是这样。他指出，两分法的，或离散的 (discrete) 特征值，不足以区别一个特征的多种表达形式，特别是语言之间的细微差别。他提出的解决办法是抛弃离散的特征值，取而代之以程度性的或延续的特征值。比如，“清浊”不用加减来两分，而是具体说明声带振动的程度。[θ] 的发音部位，不能仅仅说是齿，还应该具体说明舌尖到底在什么准确的位置。

Ladefoged 的建议可以有两种理解。一种是完全不要特征，取而代之的是一些物理“尺度” (scales)。比如，不要划分发音部位，而是把口腔从前到后看成一个延续的区域，一个辅音可以处于任何一个位置，这样可以区分出无限多的辅音位置。同样，舌位的高低不是分成有限的三级或四级，而是看成一个延续的区域，一个元音可以处于任何一个位置，这样可以区别出无数的元音高度。这种分析类似于在研究人的移动方式时，不用爬走跑来分类，而只用速度来分类。其主要缺点是“超量预测” (over-prediction)。比如，如果元音的高度是延续的，那么有的语言也许会用四个对立性的高度，或五个对立性的高

度。可是，如果自然语言里，排除“松紧”(tense)因素以后，对立性的高度最多只有三个，那就需要解释为什么理论和实际的差别那么大。

对 Ladefoged 的第二种理解，是保持现有特征，不过在每一个特征上，进一步指出其物理性质的程度。比如“浊”这个特征要保留，但我们还可以进一步标明浊的物理范围，以及每个音在这个范围里的具体程度。同样，我们可以保留元音的高中低三个分类，但进一步把每类定一个物理范围，再把每一个音定到这个范围里的某一个点。这种分析类似于在研究人的移动方式时，我们既要用爬走跑来分类，还要把每类加一个速度值。这样，对每一个动作，我们都可以决定是爬，是走，还是跑，而且还能说是多快。这种分析实际上承认了动作和动作的物理效果是两个不同的问题，特征和特征的物理体现也是两个不同的问题。特征有对立性，特征的物理程度没有对立性。不过，如果是两个问题，那么我们可以把它们分开处理，不必同时处理。

可是同一个特征，为什么可以有不同的物理效果呢？其实这并不奇怪。我们仔细考虑一下特征变化的三种情况。音与音之间的变化，实际上是不同物理环境造成的变化。声带的振动受两个因素影响，一个是声带的松紧，一个是声带前后的气压差。塞音的气压差最小，擦音的气压差略大，元音的气压差最大。即使声带做的是同一个动作“放松”(即“浊”)，可是因为气压差不同，声带振动的程度也不同。研究特征，强调的应该是动作的一致性，而不是其效果的一致性。同样，走路的速度也会受到环境影响。假如一个人用同样的方式(力度)走路，下坡一般会快些，上坡一般会慢些。如果我们研究人的走路，也应该区别动作和效果。人与人之间的变化，也是不同物理环境造成的变化，也可以同样理解。比如，声带长的人声调偏低，声带短的人声调偏高。腿长的人也许走得较快，腿短的人也许走得较慢。

人群和人群之间的动作区别怎样理解呢？这个问题比较复杂，有很多因素，包括物质因素和社会意识因素，下面我们举一些例子。先考虑走路。假如有人住山上，有人住海边，那么，山上的人走路和海边的人走路有可能有不同，或者姿势不同，或者速度不同。又假如有人住的远，每天走很多路上班，有人住的近，上班不走什么路，那么，前者也可能总的来说比后者走得快。还有，人们有时说四川人个子矮小，东北人个子高大，这可能归因于气候，饮食习惯，生活方式，等等。不管怎样，如果四川人真的矮小，东北人真的高大，那么他们的走路方式和快慢也可能不同。这些例子说明，物质环境可以使同一个动作在一群人与另外一群人之间出现系统的不同。

再考虑另一个情况。假如一个地方的人健康意识强，那么大家可能都尽量多走路，而且走得快。另外一个地方的人，健康意识也强，不过认为打太极拳，做愈加最有效，不用走什么路。那么，前一个地方的人可能总的比后一个地方的人走得快。而且，这两个地方的其它人，不管健康意识怎样，很可能也跟着大家走得快，或跟着走得慢。这个例子说明，社会文化意识也可以使同一个动作在一群人与另外一群人之间出现系统的不同。

回到发音动作上来，人群和人群之间的区别，什么是物质因素的影响，什么是社会文化意识的影响呢？因为有关的研究不多，我们试探性地提一些看法。假如两个语言都有[k]，一个语言里的[k]经常出现在[i]前，另外一个语言的[k]经常出现在[u]和[a]前。那么，前一个语言的[k]很可能就比后一个的稍微靠前一些。这是物质因素的影响。

关于社会意识对语音选择的影响，社会语言学有不少讨论。比如，纽约市有儿化韵和非儿化韵两种方言。Labov (1966) 的研究发现，同一个人根据不同的说话对象，经常改变儿化韵的使用程度。这说明，对语音形式的选择，可以根据社会意识和说话人所求的个

人形象而改变。又比如，一个地方的人，不管什么原因，[θ] 稍微靠前一点，另外一个地方的人，不管什么原因，[θ] 稍微靠后一点。那么，别的人去了，也可能跟着把 [θ] 靠前，或跟着把 [θ] 靠后。这也是社会意识因素的影响。

如果语音的对立性差别远远少于实际上发得出，听得出的差别，那么这可能不但不说明对立特征有问题，反而说明对立特征有道理，即对立性的发音动作，只是所有动作的一部分，很可能是其中最自然的，最基本的动作。

11. 结束语

语音特征是音系学的一个基本概念。本文讨论了常见的特征系统所存在的一些问题，比如国际音标系统存在的问题。这些问题包括分类不明确，术语不同一，概念不清楚，等等。本文认为，语音特征应该以对立为基础，并且以主动发音器官和发音动作来定义。本文还讨论了新的特征系统的主要构架和一些细节，包括六个主动发音器官，每个发音器官的动作，哪些传统特征可以省略，复合音的结构，各种颞音的细微区别(如 [ɕ] 和 [ç] 的区别)，等等。最后，本文对一些理论问题进行了探讨，比如特征是天生的动作还是约定俗成的，同一个特征为什么会有不同的语音体现。

*致谢:

以上有关内容在多年教学中经常和同学们进行讨论，使得一些看法不断改进。后来根据王洪君先生的建议写成本文，张慧丽提过一些修改意见，王洪君和吴君如帮助勘误校正，在此表示感谢。

中文参考资料

- 林焘, 王理嘉 1992 《语音学教程》, 北京: 北京大学出版社。
- 麦耘 2005 “对国际音标理解和使用的几个问题”, 《方言》2005年02期, 168-174页。
- 游汝杰, 钱乃荣, 高钰夏 1980 “论普通话的音位系统”, 《中国语文》1980年第5期(总158), 328-334页。
- 朱晓农 2006 《音韵研究》, 北京 商务印书馆。

英文参考资料

- Archangeli, Diana. 1988. Aspects of underspecification theory. *Phonology* 5: 183-207.
- Blevins, Juliette. 2003. The independent nature of phonotactic constraints: an alternative to syllable-based approaches. In *The syllable in Optimality Theory*, ed. Caroline Féry and Ruben van de Vijver, 375-403. Cambridge: Cambridge University Press.
- Cho, Tae-hong, Sun-Ah Jun and Peter Ladefoged. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30: 193-228.
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. N.Y.: Harper and Row.
- Duanmu, San. 1994. Against contour tone units. *Linguistic Inquiry* 25.4: 555-608.
- Duanmu, San. 2009. *Syllable structure: the limits of variation*. Oxford: Oxford University Press.
- Halle, Morris. 2003. Phonological features. In *International encyclopedia of linguistics*, volume 3, ed. William J. Frawley, 314-320. 2nd ed. Oxford: Oxford University Press.
- Halle, Morris. 2005. Palatalization/velar softening: what it is and what it tells us about the nature of language. *Linguistic Inquiry* 36.1: 23-41.
- Iverson, Gregory K. and Joseph C. Salmons. 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12.3: 369-396.
- Jakobson, Roman, Gunnar Fant, and Morris Halle. 1952. *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, Mass.: MIT Press.
- Jakobson, Roman, and Morris Halle. 1964. Tenseness and laxness. In *In Honour of Daniel Jones*, ed. David Abercrombie et al, 96-101. London: Longman, Green and Co.
- Keating, Patricia 1988a. Palatals as complex segments: X-ray evidence. *UCLA Working Papers in Phonetics* 69: 77-91.
- Keating, Patricia. 1988b. Underspecification in phonetics. *Phonology* 5:275-292.
- Kim, Chin-Wu. 1965. On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). *Word* 21: 339-359.
- Kim, Mi-Ryoung, and San Duanmu. 2004. Tense and lax stops in Korean. *Journal of East Asian Linguistics* 13: 59-104.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Ladefoged, Peter. 1972. Phonetic prerequisites for a distinctive feature theory. *Papers in linguistics and phonetics to the memory of Pierre Delattre*, ed. Albert Valdman, 273-286. The Hague: Mouton.
- Ladefoged, Peter. 1992. The many interfaces between phonetics and phonology. In *Phonologica 1988*, ed. Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer, and John R. Rennison, 165-179. Cambridge, UK: Cambridge University Press.
- Ladefoged, Peter. 2001. *Vowels and consonants: an introduction to the sounds of languages*. Malden, MA: Blackwell.
- Ladefoged, Peter. 2006. *A course in phonetics*. 5th edition. Boston, MA: Thomson and Wadsworth.

- Ladefoged, Peter. 2007. Articulatory features for describing lexical distinctions. *Language* 83.1: 161-180.
- Ladefoged, Peter, and Morris Halle. 1988. Some major features of the International Phonetic Alphabet. *Language* 64.3: 577-82.
- Maddieson, Ian. 1984. *Patterns of sounds*. London: Cambridge University Press.
- Mielke, Jeff. 2008. *The Emergence of Distinctive Features*. Oxford University Press.
- Port, Robert F., and Adam P. Leary. 2005. Against Formal Phonology. *Language*, 81.4: 927-964.
- Prinz, Michael and Richard Wiese. 1991. Die Affrikaten des Deutschen und ihre Verschriftung. *Linguistische Berichte* 133: 165-189.
- Recasens, Daniel. 1990. The articulatory characteristics of palatal consonants. *Journal of Phonetics* 18.2: 267-280.
- Sagey, Elizabeth. 1986. *The representation of features and relations in nonlinear phonology*. Doctoral dissertation, MIT, Cambridge, Mass.
- Steriade, Donca. 1987. Redundant values. In *Papers from the 23rd Annual Regional Meeting of the Chicago Linguistic Society, Part 2: Parasession on Autosegmental and Metrical Phonology*, 339-362.
- Steriade, Donca. 1999. Alternatives to syllable-based accounts of consonantal phonotactics. In Osamu Fujimura, Brian D Joseph, and Bohumil Palek, eds., *Proceedings of LP'98: Item Order in Language and Speech* (Columbus, the Ohio State University, September 15-20, 1998), Vol. I, 205-245. Prague: Karolinum Press (Charles University in Prague).
- Stevens, Kenneth. 1972. The quantal nature of speech. In *Human communication: A unified review*, ed. Edward E. David, Jr., and Peter B. Denes, 54-66. New York: McGraw-Hill.
- Wiese, Richard. 1996. *The phonology of German*. Oxford and New York: Oxford University Press.

工作单位:

美国 密西根大学 语言学系

北京大学 汉语语言学研究中心 2008 年度兼职研究员

Contrast, Features, and Articulatory Gestures
San Duanmu

Abstract

This article reviews some problems in current feature theory and then introduces a feature system that is based on active articulators and their gestures. It is proposed that the number of features can be significantly reduced. Theoretical implications are also discussed, such as the structure of complex sounds, the representation of some fine distinctions between sounds (such as between [ç] and [ç̣]), and whether features are innate or learned.

Keywords: contrast, features, active articulators, gestures, complex sounds