

广东省普通高校人文社会科学重点研究基地暨南大学汉语方言研究中心

南方语言学·2015

主编 甘于恩

广东省普通高校人文社会科学重点研究基地暨南大学汉语方言研究中心

南方语言学·2015

主编 甘于恩



暨南大学出版社
JINAN UNIVERSITY PRESS

中国·广州

图书在版编目 (CIP) 数据

南方语言学. 2015/甘于恩主编. —广州: 暨南大学出版社, 2015. 8
ISBN 978 - 7 - 5668 - 1570 - 5

I. ①南… II. ①甘… III. ①汉语方言—方言研究—丛刊 IV. ①H17 - 55

中国版本图书馆 CIP 数据核字(2015)第 177042 号

出版发行: 暨南大学出版社

地 址: 中国广州暨南大学

电 话: 总编室 (8620) 85221601

营销部 (8620) 85225284 85228291 85228292 (邮购)

传 真: (8620) 85221583 (办公室) 85223774 (营销部)

邮 编: 510630

网 址: <http://www.jnupress.com> <http://press.jnu.edu.cn>

排 版: 广州市天河星辰文化发展部照排中心

印 刷: 佛山市浩文彩色印刷有限公司

开 本: 850mm × 1168mm 1/16

印 张: 13

字 数: 421 千

版 次: 2015 年 8 月第 1 版

印 次: 2015 年 8 月第 1 次

印 数: 1—1000 册

定 价: 42.00 元

(暨大版图书如有印装质量问题, 请与出版社总编室联系调换)

汉语方言研究中心

学术委员会名单

学术顾问

詹伯慧 李如龙 张振兴 许宝华 邢福义 邵敬敏 潘悟云 张双庆 鲍厚星
张洪年 单周尧 刘村汉 林立芳 钱曾怡 温端政 平山久雄 陆镜光

主任委员

麦耘

委员

(按姓氏音序排列)

曹志耘 甘于恩 李蓝 林伦伦 麦耘 邵慧君 万波
汪国胜 伍巍 张敏 张屏生 庄初升

《南方语言学》编辑委员会

名誉主编

詹伯慧

主 编

甘于恩

编 委

(按姓氏音序排列)

陈晓锦 范俊军 甘于恩 高然 刘新中 彭小川 邵宜 伍巍

编 辑

刘新中 侯兴泉 曾建生 赵越

目 录

特 稿

“中国音系数据库”的构建及用途	端木三 (1)
-----------------------	---------

南方汉语方言研究

广东中山三乡白石客家话词汇特点述略	高 然 丘金芳 甘于恩 (16)
粤语声调演变机制初探——香港话与广州话声调混同的比较	梁 源 (34)
粤方言字的繁简使用问题	侯兴泉 欧嘉文 (43)
从两岸比较看台湾客家熟语	黄映琼 温昌衍 (49)

地理语言学

乡镇级以下的地理语言调查——两种台湾彰化县地理言论著的比较	董忠司 (58)
广东化州境内粤方言历史地理分布研究	陈李茂 甘于恩 (80)
广佛粤语人称代词“我”的读音地理差异与语音变化	陈卫强 (89)

海外汉语方言研究

海外汉语方言研究刻不容缓	陈晓锦 张敏怡 (96)
--------------------	--------------

岭南语言资源与语言技术

语料转标注工具 ELAN 和 Partitur - Editor 评测	马 松 (99)
---	----------

语音学研究

基于口语语料的粤语疑问句末音高浅析	王茂林	刘新中 (104)
梅县客话和五华客话古浊上字调类分化比较研究	张健雅	邵慧君 (112)

语法学与方言语法

阳江话“敢”字句及其惊奇语气标记的来源	黄高飞 (118)
“吧”字祈使句在口语中的功能分类及其韵律特征	董芳良 郭楚江 (125)

词汇学与词典学

试论度量形容词的不对称性	吴南开 (142)
编写闽南话词典的用字问题——方言词典编写札记之一	周长楫 (149)
浙江九姓渔民方言词汇(三)	刘倩 (154)
“评头论足”类语词的感情色彩——兼论中性、贬义共存词的形成因素	谢晓晖 (185)

评论

赣方言研究的又一力作——评《永新方言研究》	蔡国妹 (192)
-----------------------------	-----------

博士论文撮要

广东新会司前话音韵研究	汤翠兰 (195)
山西北区方言语音研究	崔淑慧 (195)
湘西乡话音韵研究	杨蔚 (196)
海南闽语的语音研究	刘新中 (196)
湖南赣语语音研究	李冬香 (197)
汉语方言动词重叠比较研究	王红梅 (197)
惠安方言动词谓语句研究	陈曼君 (198)
汉语“儿化”研究	王媛媛 (198)

动态

中国民族语言学会第11届全国研讨会在暨南大学举行	彭婧 (200)
--------------------------------	----------

Contents

Feature

- The Construction and Application of “Chinese Phonetic Data Base” Duanmu San (1)

Chinese Dialects in South China

- A Brief Research on Vocabulary Features of Hakka Dialects in Sanxiang, Zhongshan, Guangdong
..... Gao Ran, Qiu Jinfang & Gan Yuen (16)
- Mechanism of Tone Merger in Cantonese: Comparison of Tone Merging in Hong Kong and Guangzhou
..... Liang Yuan (34)
- On the Use of Traditional and Simplified Characters of Cantonese Hou Xingquan & Ou Jiawen (43)
- A Comparative Study on Idioms of Hakka Dialects between Mainland and Taiwan
..... Huang Yingqiong & Wen Changyan (49)

Geolinguistics

- Geolinguistics Investigation below Township—A Comparative Study between Two Geolinguistics
Books of Zhanghua, Taiwan Dong Zhongsi (58)
- A Study of Historical Geographical Distribution of Yue Dialect in Huazhou, Guangdong Province
..... Chen Limao & Gan Yuen (80)
- Geographical Differences and Phonetic Changes of Personal Pronoun “Me” in Guangzhou
and Foshan Yue Dialects Chen Weiqiang (89)

Oversea Chinese Dialects

- Seize Overseas Chinese Dialect Research’s Arena Chen Xiaojin & Zhang Minyi (96)

Ling Nan Language Resources and Technique

- Corpus Annotation Tool ELAN and Partitur – Editor Evaluation Ma Song (99)

Phonetics

- A Brief Analysis in Pitch at the End of Interrogative Sentence of Yue Dialect Based on Spoken Corpus
..... Wang Maolin & Liu Xinzhong (104)
- Comparative Study on the Differentiations of the Tone Sorts of Ancient Voiced Shangsheng (上声) in
Hakka Dialects between Meixian and Wuhua Zhang Jianya & Shao Huijun (112)

Syntactics and Dialect Grammar

- The Sentence of “Gan” (敢) in Yangjiang Dialect and Its Origin of the Surprised Modality Mark
..... Huang Gaofei (118)
- The Functional Classification of Imperative Sentences Ending with the Modal Particle “Ba” (吧)
and Their Prosodic Features in Oral Chinese Dong Fangliang & Guo Chujiang (125)

Lexicology and Lexicography

- A Study on the Asymmetry of the Adjectives for Measurement Wu Nankai (142)
- The Problems of Diction in Compiling Min Dialect Dictionary—One Note of Compiling Dialectal
Dictionaries Zhou Changji (149)
- A Lexicon of the JiuXing Yumin (Nine Surnames Fishermen) Dialects, Zhejiang Province (III)
..... Liu Qian (154)
- The Emotive Meaning of the Kind Words of “PingTouLunZu” (“评头论足”) —And the Forming
Factors of “the Same Word with both Derogatory and Neutral Meanings” Xie Xiaohui (185)

Comments

- Comments on *A study of Yongxin Dialect* Cai Guomei (192)

Ph. D. Dissertation Abstracts

- Research on Phonology of Siqian, Xinhui, Guangdong Tang Cuilan (195)
- A Study of Phonology in the North Part of Shanxi Dialects Cui Shuhui (195)
- A Phonological Study of the Xianghua Dialect in West Hunan Yang Wei (196)
- Phonetics and Phonology of the Min Dialect in Hainan Liu Xinzhong (196)
- A Phonological Study of Gan Dialect in Hunan Province Li Dongxiang (197)
- A Study of Verb Overlapping in Chinese Dialect Wang Hongmei (197)
- A Study of the Verb Predicate Sentence of Huian Dialect Chen Manjun (198)
- A Study on “Erhua” (儿化) in Chinese Wang Yuanyuan (198)

News

- The 11th National Seminar was Held by Zhongguo Minzu Yuyan Xuehui (Linguistic Association
of Chinese Ethnic) in Jinan University Peng Jing (200)

“中国音系数据库”的构建及用途*

端木三

(密歇根大学 美国密歇根州)

提 要 本文先介绍现有的音系数据库 UPSID 和 P-base 的收录方法和内容,再指出它们的一些缺点,然后介绍新建的“中国音系数据库”,包括收录方法、内容及应用。本文还通过一些例子,比较详细地讨论音系数据库如何用于理论研究,特别是对特征理论、音节理论和音位分析理论的研究。

关键词 中国音系数据库 语音特征 音位 音节

一、现有的音系数据库

音系数据库包含多种语言的音系信息,是探讨语言规律的主要工具。常见的音系数据库有 3 个,基本信息见 (1)~(2)。

(1) 常见的 3 个音系数据库的名称。

中文	英文名	简称
斯坦福音系档案	Stanford Phonology Archive	SPA
加州洛杉矶大学音系数据库	UCLA Phonology Segment Inventory Database	UPSID
音系数据库	P - base	P - base

(2) 常见的 3 个音系数据库的细节。

简称	创建时间	音系数量	文献
SPA	70 年代	200	Crothers et al, 1979
UPSID	80 年代	451	Maddieson, 1984; Maddieson & Precoda, 1990、2011
P - base	2004—2007	628	Mielke, 2004 - 2007

* 笔者作为富布赖特研究学者,2014 年夏天在北京大学现代汉语研究中心进行了 3 个月的访问,本项目的部分工作是在访问期间完成的。笔者对富布赖特基金会、北京大学现代汉语研究中心表示感谢。中国农业大学的衣莉作为中国留学基金会访问学者,在 2013 年至 2014 年访问密歇根大学中国研究中心期间,为“中国音系数据库”的建立做了很多工作。笔者对衣莉、中国留学基金会、密歇根大学中国研究中心表示感谢。本项目还得到了不少其他学者提供的宝贵意见和帮助,特别是中国社会科学院民族研究所的孙宏开、江荻、燕海雄、王海波,成都理工大学的何家勇,暨南大学汉语方言研究中心的刘新中、侯兴泉,南开大学的曾晓渝,以及密歇根大学的苏婷、张瀛月,笔者也在此表示感谢。

SPA 创建于 20 世纪 70 年代, 包括 200 个音系 (其中一两个点的资料不全)。当时的音系研究很少大量引用资料或量化数据, 所以 SPA 的建立具有开创性意义。SPA 的内容可见 (3)~(4)。

(3) SPA 的内容举例: 音位特征 (Crothers et al., 1979: 5)。

PAGE 005 STANFORD PHONOLOGY ARCHIVE			Hausa		01
VOLUME 1 -- SEGMENT INVENTORIES, GENERAL COMMENTS, FOOTNOTES					
	010 Hausa		010 Hausa		010 Hausa
010	01 b (b-labialized) ⁶⁰ (free)		14 k-ejective-palatalized ^{02 65} (restricted, free) */k-ejective/		
010	02 b-implosive (b-implosive-labialized) ⁶⁰ (free)		15 t/s-hacek		50 iota
010	03 t ⁰¹		16 d/z-hacek		51 i-long
010	04 d		17 phi ⁰³ (pl) ³⁰ (free)		52 e-mid ³³
010	05 d-implosive		18 phi-palatalized ⁰⁰		53 e-mid-long
010	06 k ⁰¹ *(k-labialized) *(k-palatalized)		19 s ⁰⁴		54 schwa ¹¹ (epsilon) ⁶³ (allo, free)
010	07 k-labialized ^{02 60} (restricted, free) */k/		20 z ⁰⁴		55 a-long
010	08 k-palatalized ⁰² (restricted, free) */k/ 1651		21 s-glottalized ⁰⁵ lt/s-ejective ³¹ (free)		56 upsilon
			22 s-hacek ⁰⁴		57 u-long
			23 m		58 o-mid ³³
			24 n ⁰⁶		59 o-mid-long
					60 yod

(4) SPA 的内容举例: 音位细节 (Crothers et al., 1979: 6)。

PAGE 006 STANFORD PHONOLOGY ARCHIVE			Hausa		010
VOLUME 1 -- SEGMENT INVENTORIES, GENERAL COMMENTS, FOOTNOTES					
	Foreign Service Institute \$q second author/informants \$r unknown (first author)				
010	\$a Branner, Siegmund, and Michael Ashiwaju \$b 1966 \$c Lehrbuch der Hausa - Sprache \$g Muenchen: Max Hueber Verlag \$q second author				
010	\$a Kraft, Charles H. and A.H.M. Kirk-Greene \$b 1973 \$c Hausa \$f (Teach Yourself Books) \$g London: Collins \$q informant \$r unknown				
010	\$a INTONATION \$A With regard to interrogative intonation, "typically...the last high tone in an utterance will jump to a pitch level of the previous high tone syllable. A following low tone will drop only one step rather than all the way down to pitch level 1 as in the narrative contour." Utterances which contain interrogatives have the basic pattern of the narrative contour but are distinguished from both narrative and interrogative intonation in that "(1) the whole utterance is typically on a slightly higher pitch, and (2) the final tone of the utterance, if a high tone, becomes a falling tone." This intonation is also used for vocative utterances. See K&K, p.30-36 for examples.				
010	\$a INTONATION \$A "The total pitch component (tone plus intonation) of a Hausa utterance can be described...in terms of five pitch levels." K & K give three basic intonational types: 1) narrative intonation 2) interrogative intonation 3) vocative intonation. Within the posited five pitch levels, "the pattern, therefore, of narrative intonation-JL is to drop 2 steps from a high tone to a following low tone and to jump up one step from a low tone to a following high tone with two qualifications: (1) an initial low tone will actualize on pitch level 3 but the following high tone...will still jump up to pitch level 5; and (2) a final low tone will, no matter what the length of the utterance, always fall on pitch level 1." They add, however, that long utterances are "frequently broken into more than one intonational contour, especially in slow, more deliberate speech." p.30, 31)				

例 (3) 所示的是 SPA 中的第一种语言——Hausa 的音位及其特征。例 (4) 所示的是 Hausa 的音位细节说明, 包括文献来源。可以看出, SPA 细节繁多, 使用不便, 而且没有公开的电子版本, 因此现在已经基本上没人使用 SPA 了。

UPSID 的创建紧随 SPA 之后，它几乎全盘吸收了 SPA 所收录的音系，而且还增加了许多音系。UPSID 的初版有 318 个音系（Maddieson, 1984），后来又扩展到 451 个音系（Maddieson & Precoda, 1990、2011）。UPSID 的内容见（5）~（7）。

（5）UPSID 的音位表（共 451 种语言、13 966 个音位）。

语言代码	音位
2000	p
2000	b
2000	k
2000	g
2000	ts
2000	dz
2000	f
2000	v
2000	0D
2000	6D
...	...

（6）UPSID 的音标表（共 921 项）。

序号	音标	次	特征
1	kpW	1	labialized voiceless labio-velar plosive
2	gbW	1	labialized voiced labio-velar plosive
3	Nmgb	6	prenasalized voiced labial-velar plosive
4	kph	1	voiceless aspirated labial-velar plosive
5	kp	35	voiceless labial-velar plosive
6	gb	39	voiced labial-velar plosive
7	pW-h	1	labialized velarized voiceless aspirated bilabial plosive
8	bW-	1	labialized velarized voiced bilabial plosive
...	...	...	...

（7）UPSID 的特征表（共 64 项）。

序号	特征
1	plosive
2	implosive
3	ejective stop
4	click
5	fricative
6	ejective fricative
...	...

(8) UPSID 的元音、辅音表。

语言代码	语言名称	音位总数	元音组	元音数	辅音组	辅音数
2000	Greek	26	i E u O a_	5	p b k g t s d z f v 0D 6D s z x g F m j t Dd Dn Dl Dr	21
2001	Irish	69	I" ea U" oI ~ " e ~ a ~ U ~ " o ~ i; e: u: o: a_: i ~ : e ~ : u ~ : o ~ : a_ ~ : @ i @ ui @ u @	24	p W - h p J h b J t Dh d Dk J h g J k h g t S d Z P W - P J B J S C x g F m W - m J n D N J Nl - hl - lJh l Jh r [- r [Jhr [J h j w b W - lD - t h d s n Jn - h n J h n - n_ r [- l_	45
2002	Breton	45	i y u a_ E" e" o/o" o" e ~ E ~ a_ ~ O ~ a E a u a o Ei Eo a_ O ~ EO ~	20	b g f v S Z x m n j l j w p t D k c d D d j g F s z n D N h l j	25
2004	German	41	i: y: e: o/: E: a: u: o: IYE" @ 4UOE) ai Oi au	19	p b k g p f f v s z S Z x m N l R h j d t s n t	22
...	...	...	...	...	...	...

例（5）显示每种语言的音位音标。每种语言都有个代码，从 2000 开始，不过代码意义与本文无关。音标采用的不是国际音标，而是普通的 ASCII 编码，不同操作系统的电脑都能使用。例（6）是音标表，共 921 项，每项都有该音标的特征注解。例（7）是特征表，共 64 个特征术语。例（8）是例（5）信息的重新组合，包括每种语言的代码、名称、音位总数、元音组、元音数、辅音组和辅音数。

SPA 和 UPSID 的建立都有多人参加，而且耗时多年。P-base 则是由博士研究生 Jeff Mielke 一人所建。除了提供音位以外，P-base 还详细收录了每种语言的音变规则（allophonic rules）和音位分布规率（distribution patterns），P-base 的详细内容见（9）。

(9) P-base 的内容。

序号	语言	内容
...	...	...
8791	147	Language, English
8792	147	Reference, Jensen, John T. <i>English Phonology</i> . Philadelphia: John Benjamins Publishing Company, 1993. McMahon, April. <i>An Introduction to English Phonology</i> . Oxford: Oxford University Press, 2002.
8793	147	Family, INDO-EUROPEAN (Germanic)
8794	147	Location, India, USA, UK, Canada, Australia, South Africa, New Zealand, Ireland, Tanzania, Ghana, Nigeria, Uganda, Zimbabwe, Singapore, etc.
8795	147	Langcode, ENGL
8796	147	Inventory, Core
8797	147	p, t, tʃ, k, b, d, ɔ̯, g, ai, au, f, θ, s, ʃ, h, i:, u:, v, ð, z, ʒ, ɪ, ʊ, m, n, ŋ, eɪ, ə, ɜ:, ɒ, l, ɛ, ʌ, ɔ:, ɪ, æ, aɪ, w, j
8798	147	Inventory, Marginal
8799	147	
8800	147	Target, X→ nasalized / _ nasals
8801	147	Segments, ai, au, i:, u:, ɪ, ʊ, eɪ, ə, ɜ:, ɒ, ɛ, ʌ, ɔ:, æ, aɪ
...	...	...

每种语言的内容包括语言名称、资料来源、语言归类、语言使用地区、语言代码、核心音位、边缘音位以及音变规则和分布规律。例（9）只显示了一条语音规则，即英语的元音在鼻音前会产生鼻化。还可以看出，英语没有边缘音位，而且音位全部采用国际音标标注。

以上介绍了现有音系数据库的内容。现在介绍它们的收录方法。SPA 和 UPSID 采用的都是均衡采样（balanced sampling），即以每种语言分支取一种语言作为代表的方法进行收录。而且，因为 SPA 和 UPSID 的重点是音位，所以很多资料来源都是短文。P-base 为了收集音变规则和分布规律，必须依赖更详细的资料。因此，P-base 的资料来源限于专著和博士论文。不过，P-base 穷尽收集了图书馆所有的英文出版书目，所以也不失为一份客观采样。

二、音系数据库的用途

音系数据库的建立有力地推动了一系列理论研究，特别是对语言规律的探讨。下面我们举一些例子，见（10）。

（10）使用音系数据库的研究举例。

作者	刊物	数据库
Trubetzkoy, 1939	专著	200 种语言
Liljencrants & Lindblom, 1972	<i>Language</i>	SPA 等
Lindblom, 1986	论文集	SPA
Maddieson, 1984	专著	UPSID
Haspelmath et al., 2005	专著	2 559 种语言
Hay & Bauer, 2007	<i>Language</i>	216 种语言
Donohue & Nichols, 2011	<i>Linguistic Typology</i>	UPSID
Atkinson, 2011	<i>Science</i>	504 种语言
Wang et al., 2012	<i>Science</i>	579 种语言
Moran et al., 2012	<i>Language</i>	SPA, UPSID 等

Trubetzkoy（1939）是探讨音系普遍规律的主要创始人。他本人收集了 200 种语言的资料，书中的理论多是以这些语言数据为依据。其他列举的研究多在一流学术刊物上发表，如 *Language*、*Science*。Liljencrants & Lindblom（1972）提出的元音分布理论，影响深远。他们的音系语料来自 Trubetzkoy（1939）以及 SPA。Lindblom（1986）对 Liljencrants & Lindblom（1972）的理论提出了一些修改，其语料数据也是来自 SPA。Haspelmath et al.（2005）更是对世界语言的普遍特征作了一系列的概括，其中有关音系的各种论文共计 19 篇。Hay & Bauer（2007）、Donohue & Nichols（2011）、Moran et al.（2012）等对音位数量跟人口数量是否有相关性这一问题展开了激烈的讨论。他们的论点有的来自 SPA、UPSID 的数据，有的来自自己收集的语言数据。而 Atkinson（2011）、Wang et al.（2012）则通过音位数量的多少来探讨人类语言的起源地，两篇论文都发表于自然科学的一流刊物 *Science*。

与音位系统有关的研究很多。我们将对主要的一些进行总结，见（11）。其中不少课题来自 Haspelmath et al.（2005）的音系部分。

（11）跟音位系统有关的研究课题举例：

- 最常见/罕见的元音/辅音有哪些？
- 最常见/罕见的音位系统有哪些？
- 辅音数量跟元音数量是否有互补关系？

音位数量跟音节长短是否有互补关系?

音位数量跟人口数量是否有相关性?

音位数量跟语言是否古老有无相关性?

世界的语言一共有多少不同的元音、辅音?

下面我们对 (11) 的最后一个问题进行更深一层的讨论。首先, 我们同意 Maddieson (2005) 的说法: 元音的数量可以分成两个部分, 一部分是基本元音 (basic vowels), 另一部分是非基本元音 (non-basic vowels), 见 (12)。

(12) 基本元音与非基本元音。

基本元音: 由舌位高低、前后及圆唇变化构成。

非基本元音: 双元音, 或由基本元音加鼻化、送气、长短构成。

不少人 (如 Ladefoged, 1972; Flemming, 1995) 认为, 很多语音的特征或维度 (phonetic features or dimensions) 都是延续的。比如, 舌位的前后、高低都是延续的。每种语言可以自由决定每个维度分几个对立、每个对立的划分边界在哪里。下面我们用舌位举例。

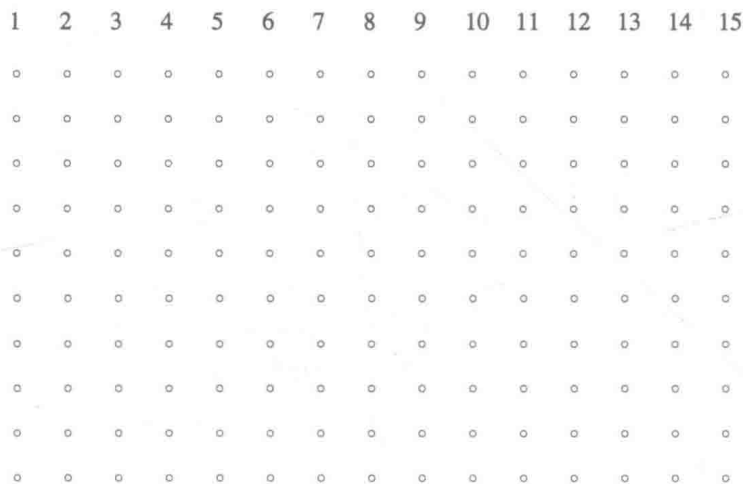
Flemming (1995) 提出, 舌位前后、高低的对立由两个因素决定, 一个是最小可辨区别 (just noticeable difference)、一个是最小听辨距离 (minimal perceptual distance), 见 (13) ~ (16), 其中两个圆点间的距离表示最小可辨区别, 整个前后空间假设分为 15 级可辨区。

(13) 舌位前后、高低对立的两个因素 (Flemming, 1995):

a. 最小可辨区别决定最多对立可能。

b. 最小听辨距离 (即两个对立点之间的最小可辨区别数) 决定实际对立数。

(14) 舌位前后、高低对立的图示: 圆点间距离 = 最小可辨区别。



(15) 舌位前后对立的图示: 最小听辨距离: 前后 = 2。



(16) 舌位前后对立的图示：最小听辨距离：前后 = 4。

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	。	。	。	B	。	。	。	C	。	。	。	D	。	。

根据最小可辨区别，舌位的前后、高低可以分出许多对立等级，两组等级又可以产生许多组合。假如前后对立分 15 级、高低对立分 10 级，一共就有 150 种可能组合，如 (14) 所示。当然，以最小可辨区别作为对立的分级依据未必实际，难免产生错误。所以，为了听觉的方便，实际要求的最小听辨距离往往大于最小可辨区别。如果我们将前后对立的最小距离定为 2（即两个最小可辨区别），那么前后可以有 8 级对立，见 (15) 的 A ~ H。如果我们将前后对立的最小距离定为 4，那么前后可以有 4 级对立，见 (16) 的 A ~ D。

值得注意的是，确定了一个最小听辨距离后，仍然可能有多种不同结果。比如，如果我们将前后对立的最小距离定为 4，仍然起码有 5 种对立系统，见 (17)。

(17) 舌位前后最小听辨距离等于 4 时，仍然可能有多种对立系统。

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a.	A	。	。	。	B	。	。	。	C	。	。	。	D	。	。
b.	。	A	。	。	。	B	。	。	。	C	。	。	。	D	。
c.	。	。	A	。	。	。	B	。	。	。	C	。	。	。	D
d.	A	。	。	。	。	B	。	。	。	。	C	。	。	。	D
e.	A	。	。	。	B	。	。	。	C	。	。	。	。	。	D

Disner (1983) 研究了各种语言的元音舌位，发现 (17) 所示的这种语言的对立系统间的确存在细微区别。因此，在 (17) 的几个对立系统中，即使每个系统的对立数相同（都是 4 级对立），我们仍然不能将一个系统中的元音等同于另外一个系统中的相对元音。比如，(17a) 的 A 跟 (17b)、(17c) 的 A 位置不一样。(17a) 的 A 跟 (17d) 的 A 位置虽然一样，但是 (17a) 中 A ~ B 的距离跟 (17d) 中 A ~ B 的距离不一样。可是，音位数据库中，遇到 (17a) ~ (17e) 这种情况，人们往往用同样的一组音标来代表 A ~ D。因此，不少学者认为，按照音标的异同来统计音位数据库中的元音数量在方法上是错误的，在实际上基本是不可能的（如 Ladefoged, 1972; Disner, 1983; Flemming, 1995）。

国际音标系统的结论也基本上一样。比如，仅仅围绕 [i] 这个基本符号，国际音标中起码有 9 种不同的元音，见 (18)。

(18) 舌位前后、高低对立的图示：两点间距离 = 最小可辨区别。

i	i	i	。	。	。	。	。	。	。	。	。	。	。	。	。
i	i	i	。	。	。	。	。	。	。	。	。	。	。	。	。
i	i	i	。	。	。	。	。	。	。	。	。	。	。	。	。
。	。	。	。	。	。	。	。	。	。	。	。	。	。	。	。
。	。	。	。	。	。	。	。	。	。	。	。	。	。	。	。
。	。	。	。	。	。	。	。	。	。	。	。	。	。	。	。

对每个元音符号，国际音标可以附加“略前、略后、略高、略低”等符号，因而产生 9 种对立。如

果这些对立不在同一语言里出现,人们往往只选用最简单的符号,即 [i], 无论它是否略前、略后、略高、略低,那么语音的实际值就被忽略了。

不过,以上有关舌位前后最大对立数的讨论基本属于推测。下面我们讨论音系数据库中实际出现的舌位前后对立。首先,我们讨论音系理论最根本的一个原则,即“对立原则”(Principle of Contrast)。我们用(19)定义对立(contrast),用(20)定义对立原则。这两项定义是音系理论的基础,均无争议。

(19) 对立(contrast):

- a. 如果两个音 A、B 在任何一种语言里都能够区别词义, A、B 就有对立。
- b. 如果两个音 A、B 在任何一种语言里都不区别词义, A、B 就无对立。

(20) 对立原则(Principle of Contrast):

- a. 如果两个音 A、B 有对立, 它们必须加以区别(用不同音标表示)。
- b. 如果两个音 A、B 无对立, 它们不须加以区别(可用相同音标表示)。

下面我们讨论如何运用对立原则来分析实际出现的舌位前后对立。我们提出如下步骤, 见(21)。

(21) 分析舌位前后的步骤:

- a. 从音系数据库提取所有有关舌位前后的不同音标。
- b. 将这些音标按舌位前后排序。
- c. 逐一排查每个音标是否跟其左右的音标有对立。
- d. 去除无对立的音标。

下面我们演示(21)的操作。假如我们从某个音系数据库收集到9个前后不同的音标, 如(22)。这些音标只有舌位前后区别, 没有其他区别(即其他特征皆为“高”、“非圆唇”、“紧”)。

(22) 只有舌位前后区别的音标:

[i̥ i̥ i̥ i̥ i̥ i̥ i̥ i̥ i̥ i̥] (9级对立)

首先, 我们排查每种语言的元音系统, 看 [i̥ i̥] 是否有对立。如果没有对立, 我们可以认为 [i̥ i̥] 不必区别, [i̥] 就是 [i̥]。所以, 我们排除 [i̥], 留 [i̥], 见(23)。

(23) 第一步: 检查 [i̥ i̥] 是否有对立。

- a. 排查每种语言的元音系统, 看 [i̥ i̥] 是否有对立。
- b. 如果 [i̥ i̥] 无对立, 则 [i̥ i̥] 不必区别, [i̥] 可以等同于 [i̥]。
- c. 排除 [i̥], 简化最大对立组: [i̥ i̥ i̥ i̥ i̥ i̥ i̥ i̥ i̥] (8级对立)。

第二步, 我们检查 [i̥ i̥], 方法相同。如果 [i̥ i̥] 无对立, 我们可以排除 [i̥], 见(24)。如果 [i̥ i̥] 有对立, 我们进一步检查 [i̥ i̥]。如果 [i̥ i̥] 无对立, 我们还是可以排除 [i̥]。

(24) 第二步: 检查 [i̥ i̥] 是否有对立。

- a. 排查每种语言的元音系统, 看 [i̥ i̥] 是否有对立。
- b. 如果 [i̥ i̥] 无对立, 则 [i̥ i̥] 不必区别, [i̥] 可以等同于 [i̥]。

c. 排除 [i], 简化最大对立组: [i i i i w w w] (7 级对立)。

如果我们用 UPSID、P-base 这两个音系数据库, 重复 (23)、(24) 的步骤, 所得到的结果将是 (25)。

(25) 最大舌位前后对立结果 (“高”、“非圆唇”、“紧” 元音):

[i ɯ] (2 级对立)

我们发现，元音舌位前后的对立最多只有两级，高元音如此，其他元音也如此。还可以证明，如果排除长短、舌根前后（有时也叫“松紧”或“tense”），元音的高低也只分两级。比如，英语的前元音有5个 [i I e ε æ]，其中 [e] 实际上是长元音 [e:] 或双元音 [ei]。如果排除 [e]，英语的前元音只有4个 [i I ε æ]（甚至只有3个 [i I ε]，因为 [æ] 经常是个长元音），它们可以用两极高低加两级舌根前后来区别。同样，(18)所示的跟 [i] 有关的9个元音 [i̥ i̯ ī̯ ī̯̄ ī̯̄̄ ī̯̄̄̄ ī̯̄̄̄̄ ī̯̄̄̄̄̄]，牵涉三级前后区别、三级高低区别，但是可以证明，这些区别都无对立性，因此没有区分的必要。

我们得出的结果有两个理论意义：第一，实际的对立系统比前人想象的要简单得多；第二，如果只有两级对立，不同语言的元音就有了可比性，因为对任何一个特征来说，只有这两级对立才是有音系意义的，其他的细微区别都是无音系意义的。

三、现有数据库的缺点

现有音系数据库有不少缺点。比如, UPSID 的语料来源很多是几页的短文, 缺乏足够的细节。而且, UPSID 基于采样收录, 音系数量有限 (共 451 个音系)。P-base 虽然是穷尽采样, 但是因为限于英语专著, 音系数量也有限 (共 628 个音系)。对于世界上现有的数千种语言来说, 这样的覆盖面并不算大。如果用照片来作比喻, 现有的数据库可以说为世界上的语言提供了一个低清晰度的全景, 但是如果我们想把镜头拉近, 观察某个地区的细节, 就会发现信息既不全面, 又不平衡。比如, P-base 收录的 628 个音系里, 在中国境内使用的只有 10 个, 其中汉语 3 个 (官话、广州话、客家话)、少数民族语 7 个 (藏语、朝鲜语、毕苏语、鄂温克语、柯尔克孜语、图瓦语、达斡尔语)。从汉语来看, 吴语没有收录; 从非汉语来看, 中国有约 130 个少数民族语言 (孙宏开等, 2007), 而 P-base 只收录了 7 个, 标注也有错误。比如, 在 P-base 里, 蒙古语的使用地区是 “蒙古”, 没有包括 “中国” 或 “内蒙古”。

不少非英语国家和地区，如中国、法国、俄国，有很悠久的语言学传统和丰富的语言调查资料，但现有的音系数据库也没有利用这些资源。

四、“中国音系数据库”

为了克服现有音系数据库的缺点，我们计划建立一个“中国音系数据库”（Phonology Inventory Database of China）。我们的收录原则如（26）。

(26) “中国音系数据库”的收录原则:

穷尽收录所有已出版的，所有语言、方言的音系资料。

尽量收录学者们已收集的未出版的语言、方言的音系资料。

可以看出，我们的工作从前人的抽样收录转向了穷尽收录。我们之所以认为这项工作能够成功，是因为中国有很悠久的语言学传统和丰富、全面的语言调查资料，只是这些资料尚未集中起来，成为方便学者们使用的研究工具。从长远来说，我们希望我们的工作只是一个开头，希望这项工作能推动其他地