

語言學論叢

第三十八輯

北京大学汉语语言学研究中心《语言学论丛》编委会编



商務印書館

語言學論叢

(第三十八輯)

北京大学汉语语言学研究中心

《语言学论丛》编委会编

商務印書館

2008 年 · 北京

图书在版编目(CIP)数据

语言学论丛. 第38辑/北京大学汉语语言学研究《语言学论丛》编委会编. —北京:商务印书馆, 2008
ISBN 978-7-100-05959-6

I. 语… II. 北… III. 语言学—丛刊 IV. H0-55

中国版本图书馆 CIP 数据核字(2008)第 131954 号

所有权利保留。

未经许可,不得以任何方式使用。

YŪYÁNXUÉ LÙNCÓNG

语 言 学 论 丛

(第三十八辑)

北京大学汉语语言学研究《语言学论丛》编委会编

《语言学论丛》编委会编

商 务 印 书 馆 出 版
(北京王府井大街36号 邮政编码100710)
商 务 印 书 馆 发 行
北京市白帆印务有限公司印刷
ISBN 978-7-100-05959-6

2008年12月第1版 开本787×960 1/16
2008年12月北京第1次印刷 印张23¼

定价:39.00元

《语言学论丛》编委会

主 编：陆俭明

编 委（按姓氏音序排列）：

贝 罗 贝	丁 邦 新	郭 锡 良	何 九 盈	何 莫 邪
江 蓝 生	蒋 绍 愚	鲁 国 尧	梅 祖 麟	平 山 久 雄
裘 锡 圭	唐 作 藩	王 福 堂	王 洪 君	王 士 元
余 霭 芹	郑 锦 全	朱 庆 之	邹 嘉 彦	

编辑部成员（按姓氏音序排列）：

陈保亚	董秀芳	耿振生	郭 锐	李小凡	宋绍年
王洪君（主任）	詹卫东	朱庆之（副主任）			

本辑执行编辑：董秀芳

执行编辑助理：曹牧春

目 录

语言比较择词的思辨与实证方法·····	黄 行 (1)
关于上古汉语核心词表的确定的几个问题 ——评《论核心语素表的确定——以上古汉语为例》(陈、 汪 2006)·····	汪维辉 (11)
基本层次范畴与核心词集构建·····	江 荻 (30)
Grammaticalization as an Areal Phenomenon	
— The Case of East and Mainland Southeast Asian Lan- guages and Its Consequences for Concepts of Com- plexity and Maturation ·····	Walter Bisang (64)
从隐含到呈现(下)——词汇变化影响语法变化·····	胡敕瑞 (99)
言说和求索对象标记的来源·····	李宗江 (128)
谓词重叠疑问句的语言共性及其解释·····	刘丹青 (144)
“有+VP”的功能特性 ·····	周国正 (165)
《经典释文》“或音”“一音”研究·····	张 冰 (182)
赵少咸《经典释文法氏校语录》批校之音读考辨·····	万献初 (225)
王质《诗总闻》叶音性质与宋代语音史问题·····	张民权 (246)
清代来华西方人所描述的官话音及基础方言·····	黄灵燕 (261)
从《书学慎余》看康熙年间的河南柘城话·····	周赛华 (281)
两汉鱼侯分部·····	孙 顺 (290)
“美”字的归部问题·····	郑张尚芳 (312)

闽蜀同风与方言 刘晓南 (318)

《直说通略》和它的语言特色 祖生利 (335)

ABSTRACTS(提要) (353)

CONTENTS

Lexical Selection for Language Comparison:

Reasoning and Empirical Approaches Huang Xing (1)

Issues on the Identification of Basic Words

in Ancient Chinese — On the Definition

of Kernel Morphemes—The Case of Old

Chinese by Chen Baoya and Wang Feng

..... Wang Weihui (11)

Reconstructing Swadesh Basic-word List with

the Basic-Level Category Jiang Di (30)

一种区域性的语法化——东亚和东南亚大陆

语言个案以及它对“复杂性”和“成熟性”概

念的影响 Walter Bisang (64)

From Covert to Overt: Influence of Lexical

Change on Syntactic Change Hu Chirui (99)

Origin of the Speaking and Requesting Objects

Marker Li Zongjiang (128)

Language Universals of Verb-reduplicating

Questions and Their Explanation Liu Danqing (144)

The Functions and Characteristics of “you

(有) + VP” Chow Kwok Ching (165)

A Study on the “*huoyin*(或音)” and “*yiyin*(一

音)” of *Jing Dian Shi Wen*(《经典释文》) Zhang Bing (182)

A Study on Zhao Shaoxian's Collation of Fa's Collation of <i>Jing Dian Shi Wen</i> (《经典释文》)	Wan Xianchu (225)
On Wang Zhi's (王质) <i>Shi Zong Wen</i> (《诗总 闻》) and the Phonetic History of the Song Dynasty	Zhang Minquan (246)
Standard Mandarin and Basic Dialects as Described by Westerners to China in the Qing Dynasty	Wong Ling Yann (261)
The Zhecheng Dialect of Henan Province during the Kangxi Period in <i>Shu Xue Shen</i> <i>Yu</i> (《书学慎余》)	Zhou Saihua (281)
The Distinction between <i>yu</i> (鱼) and <i>hou</i> (侯) Rhyme Groups in the Han Dynasty	Sun Shun (290)
Notes on Which Rhyme Group the Character <i>mei</i> (美) belongs to in Old Chinese	Zheng-Zhang Shangfang (312)
The Saying "Min and Shu Have the Same Customs" and the Dialects	Liu Xiaonan (318)
On <i>Zhi Shuo Tong Lüe</i> and Its Linguistic Characteristic	Zu Shengli (335)
ABSTRACTS	(353)

语言比较择词的思辨与实证方法

黄 行

提要 选择语言比较用词以说明语言的同源关系是重要的问题。本文提出语音对应、普适性和标记性是核心同源词应具备的三个方面的品质,并用理据的思辨和语料的实证方法论证了语言比较择词的合理性。

关键词 语言比较 核心词 思辨方法与实证方法

语言中的词汇单位通常比形态单位更加古老,因此如果是亲属语言,语言之间一定会存在一些有语音和意义联系的同源词。另一方面,语言在长期的历史演变过程中,词汇系统会发生明显的音义变化或词语替换,语言分化时间越久,亲缘关系越疏远,保留的同源词越少。亲属语言之间同源词的去留应该是有规律的,比如产生年代较早,与人的基本社会活动、认知活动关系密切的核心词是同源词的可能性,要大于那些产生年代较晚,不具有人们交际认知普遍性的一般词汇。在进行语言历史比较时,通常会采用 Swadesh 为测定语言分化年代而制订的核心词表。该词表的确定遵循了两个原则,一是选取稳定性、普适性强的词汇,尽量避免选择限于特定民族和特定时期形成的特殊词汇;二是被选的核心词需要根据已确定亲属语言的同源词加以验证。前一个原则是思辨的方法,后一个原则是实证的方法,二者相互补充,缺一不可。

本文也拟从思辨的角度阐释语言比较研究用核心词应具备语音对应、普适性和标记性三个方面的品质,并通过实证的方法用已知的语言事实(主要是广义的汉藏语系语言)验证择词的合理性。

1 语音对应规律

同源词是具有语音和意义联系的词,只有语音对应的词才可能是同源词,因此语音对应规律是确定语言比较用核心词应该坚持的一个原则。

例如“眼睛”是人体的重要器官,因此该词是公认的非常基本的核心词,它在已经确定亲属关系的语言里通常是同源词。以下列举“眼睛”在侗台语族和苗瑶语族诸语中的语音形式与对应关系。

侗台语:

壮语	傣语	侗语	水语	黎语	构拟(梁敏、张均如)
ɣa ¹	ta ¹	ta ¹	nda ¹	tsha ¹	*mp ^{ra}

苗瑶语:

苗语	布努语	炯奈语	勉瑶语	标敏瑶语	构拟(王辅世、毛宗武)
mɛ ⁶	mɔŋ ⁶	mɔ ⁶	mwei ⁶	mi ⁶	*mwæi ⁶

因为是同源词,所以可以得到以下语音对应规律的支持。

侗台语声母对应:壮语 ɣ: 傣语 t: 侗语 t: 水语 nd: (黎语 tsh)

词汇	壮语	傣语	侗语	水语	黎语
眼睛	ɣa ¹	ta ¹	ta ¹	nda ¹	tsha ¹
竹篾	ɣuk ⁷	tɔk ⁹	—	ndjuk ⁷	—
香	ɣa:ŋ ¹	—	ta:ŋ ¹	nda:ŋ ¹	—

侗台语韵母对应:壮语 a: 傣语 a: 侗语 a: 水语 a: 黎语 a

词汇	壮语	傣语	侗语	水语	黎语
眼睛	ɣa ¹	ta ¹	ta ¹	nda ¹	tsha ¹
腿	ha ¹	xa ¹	pa ¹	pa ¹	ha ¹
狗	ma ¹	ma ¹	ŋwa ^{1'}	ma ¹	pa ¹
厚	na ¹	na ¹	na ¹	?na ¹	na ¹

苗瑶语声母对应:苗语 m: 布努语 m: 炯奈语 m: 勉瑶语 mw:
标敏瑶语 m

词汇	苗语	布努语	炯奈语	勉瑶语	标敏瑶语
眼睛	mɛ ⁶	mɔŋ ⁶	mɔ ⁶	mwei ⁶	mi ⁶
细	moŋ ⁴	mɔŋ ⁴	—	mwan ⁶	mun ⁶
你	moŋ ²	—	mɛŋ ²	mwei ²	mɛi ²
蜜蜂	—	—	—	mwei ⁴	mi ⁴

苗瑶语韵母对应: 苗语 ɛ : 布努语 ɔŋ : 炯奈语 ɔ : 勉瑶语 ei : 标敏瑶语 i

词汇	苗语	布努语	炯奈语	勉瑶语	标敏瑶语
眼睛	mɛ ⁶	mɔŋ ⁶	mɔ ⁶	mwei ⁶	mi ⁶
软	mɛ ⁶	—	mɔ ⁶	mwei ⁶	mi ⁶
拿	mɛ ¹	—	—	—	—

再来看“眼睛”在语源关系不确定的语言之间的分布。许多人认为侗台语的“眼睛”与汉语“眈”^{*b. ra:} 来母鱼部音近, 侗台语该词的韵母^{*a} 的确可以和汉语鱼部^{*a} 建立在诸多词语中重复出现的语音对应关系。但是, 侗台语的“眼睛”也和南岛语的“眼睛”(如阿美语 mata、排湾语 matsa 等) 严格对应, 因此就“眼睛”一词而言, 侗台语和南岛语关系更近。因为汉语的“眈”是由“黑”的本义引申为“瞳人”义, 根据词义简单匹配的原则, 汉语如果与它有关系, 也是比较间接的。侗台语的“瞳人”一词显然也与南岛语 mata“眼睛”有关。

壮语	临高语	毛南语	标语	黎语	构拟
pat ⁸	fət ⁸	pat ⁸	pat ⁸	but ⁸	*bət

苗瑶语的“眼睛”和汉语的“目”、“眉”音近, 如果用语音对应验证, 它和觉部的“目”大概只是貌似, 而与脂支部的“眉”等有可能同源。

词汇	苗语	布努语	炯奈语	勉瑶语	标敏瑶语	汉语声韵
眼睛	mɛ ⁶	mɔŋ ⁶	mɔ ⁶	mwei ⁶	mi ⁶	眉明母脂部
来	ta ²	tɔ ²	mɔ ⁶	tai ²	ta ²	抵端母脂部
屎	qa ³	kɔ ³	ka ³	gai ³	kai ³	屎晓母脂部
剖	pha ⁵	phɔ ³	pha ⁵	phai ⁵	pha ⁵	裨帮母支部

借词也是有语音对应的,甚至会对应得更整齐,但是同源词和借词的语音对应规律有所不同。李方桂先生(1976)曾指出壮语龙州话和武鸣话中有两类汉语借词:一类在不同的台语中有整齐的对应,而与汉语读音表现出音韵上的分歧;另一类词在诸台语中不一致,而与汉语读音表现出音韵上的一致性。李先生倾向如果有两种分歧的音韵规则,与台语一致与汉语不一致的词为母语词,与汉语一致与台语不一致的词为汉语借词。因此关系词的语音特征比较符合母语还是汉语的音韵规律可以作为区分同源词和借词的标准。

比如瑶语 fai^5 “小”、 $fi:\eta^3$ “醒”、 fje^3 “写”、 $fu\eta^5$ “送”等词基本不见于苗语,因此无法建立苗瑶语之间的语音对应,但是瑶语诸词的声母 f 显然与汉语的心母整齐对应,诸词的韵母 ai 、 $i:\eta$ 、 e 、 $u\eta$ 也显然与汉语的蟹(fai^5 “小”借自汉语的“细”)、梗、麻、通诸韵对应,因此可以断定它们为较晚的汉语借词。

2 普适性

具有同源地位的核心词应该是语言分布广泛、历史存留久远的词汇,同时从认知上应该是代表比较朴素的概念,较少文化引申意义,语言之间词义可以简单匹配的词汇。能够满足这些条件的词不是很多,所以核心词在数量上应该是比较有限的。

用已知同源语言词汇去检验上述核心词的普适性,会发现思辨的假设有时会与实证的事实之间并不一致。

2.1 基本核心词的多源性

如果一个语言群体高度同源,其基本的核心词应该高度同一,即核心词在语群内部应该具有唯一性。但是事实上并不是这样简单。

词语“人”无疑是一个基本核心词,但是其在诸多亲属语言中却呈多源和多样分布。

藏缅语中至少包括三个不同来源的“人”。

藏缅语“人”1,多见于藏、羌语支:

藏语	门巴语	普米语	嘉戎语	达让语
mi	mi	mi ⁵⁵	rmi	me ⁵⁵

藏缅语“人”2,多见于彝语支:

毕苏语	阿依语	格曼语	傈僳语	怒苏语
tshaŋ ⁵⁵	tshaŋ ³¹	tson ³⁵	tsho ³³	tshō ³³

藏缅语“人”3,多见于缅语支。在下面缅彝语支语言中“人”3、“人”2 并存:

载瓦语	浪速语	波拉语	勒期语	怒苏语
pju ⁵¹	pju ³¹	pju ⁵⁵	pju ³¹	su ³³
tso ²¹	tso ³⁵	ta ³¹	tso ³³	tshō ³³

苗瑶语的“人”至少有两个词:

苗瑶语“人”1:

苗语	布努语	畲语	勉瑶语	标敏瑶语
ne ²	no ²	ne ²	mjen ²	min ²

苗瑶语“人”2:

养蒿苗语	吉卫苗语	先进苗语	石门苗语
ṁhu ¹	ɕoŋ ¹	ṁoŋ ¹	ṁau ¹

苗瑶语“人”1 可以与汉语的“人”或“民”对应,所以有可能是汉语借词。“人”2 是苗族的自称,其本义应为苗语的“人”。

侗台语表示“人”的词:

壮语	傣语	侗语	水语
vun ²	kun ²	kən ²	zən ¹

此外,侗台语还有两个指人的量词,其本义当为“人”的自称。

其一主要分布于壮傣语支:

武鸣壮语	龙州壮语	版纳傣语	德宏傣语	布依语
pau ⁴	phu ³	phu ³	phu ³	pu ⁴

其二主要分布于侗水、仡央、黎语支:

水语	毛南语	佯僂语	布央语	黎语
ʔai ³	ʔai ¹	ʔai ³	ʔja:ŋ ⁴²	ɬai ¹

以上诸汉藏语系语族基本核心词“人”的多样性分布,说明这些语言群体的基本词汇可能有不同的来源,或者曾经发生过基本词汇的替换。

2.2 基本核心词词义非简单匹配

核心词是认知上比较朴素的概念,较少转义,语言之间应该可以简单匹配。但是实际上即使是亲属语言,也会比较普遍地发生词语的转义和替换。

例如核心词“手”在以下语言群体之间的简单匹配是看不出联系的。

汉语上古	藏缅藏语	苗瑶苗语	侗台壮语	南岛排湾语
s-nuʔ	lag	pi ⁴	fau ²	lima

但是这些语言在与“手”有关的一些词语中却可能存在着语音关系。譬如:

汉语“手”*s-nuʔ: 藏缅语“手指”*s-nuʔ。词例:

汉语上古	藏缅载瓦语	藏缅彝语	藏缅羌语
s-nuʔ	-ŋju ²¹	-nɿ ³³	-su

汉语“把”*p-laʔ: 藏缅语“手”*lak。词例:

汉语上古	藏缅藏语	藏缅门巴语	藏缅缅语
p-laʔ	lag	laʔ	lak

汉语“扶”*ʔ-pa: 苗瑶语“手”*q-bu: 侗台语“手掌”*k-paʔ: 南岛语“手掌”*dapa。词例:

汉语上古	苗瑶苗语	侗台壮语	南岛泰雅语
ʔ-pa	pi ⁴	pa ³	qəbaʔ

侗台语“手”*muula: 南岛语“手”*(qa)lima。词例:

侗台武鸣壮语	侗台龙州壮语	南岛排湾语	南岛邵语
fau ²	mu ²	lima	rima

南岛语“手”、“五”*(qa)lima : 苗瑶语“五”*pra。词例:

南岛排湾语 南岛邵语 苗瑶高坡苗语 苗瑶标敏瑶语

lima 五 rima 五 pla¹ 五 pla¹ 五

lima 手 rima 手 — —

南岛语“手臂”: 汉语“五”*ŋaʔ; 藏缅语“五”*p-ŋaʔ。词例:

南岛排湾语 南岛邵语 汉语上古 藏缅藏语 藏缅景颇语

lima 手臂 rima 手臂 ŋaʔ 五 lŋa 五 ma³¹ ŋa³³ 五

2.3 基本核心词的语言分布

亲属语言的核心同源词因为有共同的起源,因此应该在亲属语言中有广泛的分布;而后起的一般词汇会因语言而异地零散分布。所以做语言历史比较,仅用较少的核心词汇进行比较即可。

“舌头”一词可以支持这个假设,该词在以下亲属语言中基本是唯一的同源词。

汉语“舌头”*s-det : 藏缅语“舌头”*s-blet:

汉语上古 藏语 羌语 景颇语 缅语 彝语

s-det ltøe zəχ -let³¹ hlja:² lo⁵⁵

苗瑶语的“舌头”*m-blet:

苗语 布努语 炯奈语 畲语 勉瑶语 标敏瑶语

ŋi⁸ ntɬa⁸ mpli⁸ pi⁸ bjet⁸ blin⁸

侗台语的“舌头”*linʔ / *C-ma:

壮语 傣语 临高语 黎语 侗语 水语 仡佬语

lin⁴ lin⁴ lin⁴ ɬin³ ma² ma² mau³¹

南岛语的“舌头”*dilaq / *Sema:

Mansaka 语 Imorod 语 排湾语 邵语

dilaq lila səma ðama

根据“舌头”一词的语源可将上述语言分为两个语群,汉语、藏缅语、苗瑶语的“舌头”有同源关系,侗台语、南岛语的“舌头”有平行的同源关系。

然而,这一基于思辨的假设并不能得到普遍的推广,有些即使是非常稳定的核心词,在亲属关系密切的语言之间也通常会有缺项。核心词在已经确定有亲缘关系语言中的分布缺项,以及核心词多种来源和无法简单匹配等现象给我们的启示是,所谓同源语言或亲属语言实际上都是包含了异源成分的,纯粹而理想的原始共同语事实上是不存在的。

3 标记性

我们(孟和达来、黄行 1997)曾经提出,核心词与非核心词之间存在标记性差别:核心词的词阶划分从语言的外部关系看,与词的社会化程度有关;从语言结构内部看,与词项的标记性程度有关。词汇的标记性通常表现为,相互对立的词项中分布限制小、出现频率高、形态简单者是无标记的;反之分布限制大、出现频率低、形态复杂者是有标记的。在 Swadesh 的 200 核心词中,可以列举出以下符合上述标准的无标记与有标记对立的词项。(“/”前后分别为无标记词项和有标记词项)

形容词:大/小,长/短,热/冷,好/坏,新/旧,干/湿,远/近,多/少,宽/窄,厚/薄。

名词:人/孩子,男人/父亲,男人/丈夫,女人/母亲,女人/妻子,鸟/翅膀,树/树林,树/树皮。

代词:我/你/他,我/我们,我们/他们,这/这里,那/那里。

数词:一、二/三、四、五。

如果把 Swadesh 的 200 核心词分为高阶词和低阶词,无标记词所处的词阶都高于或至少不低于与之相对立的有标记词,因此推论出反映词汇稳定性的词阶也有标记性差别。

将这样思辨的假设用实际语言的事实去检验,可以证实核心词基本上是符合上述标记性的分布、词频和形态标准的。限于篇幅,以下仅举侗台语、苗瑶语的几个词例说明。

侗台语词例：（－表示无标记词，＋表示有标记词）

标记性	词项	壮语武鸣	侗语南部	临高语	黎语	仡佬语
－	远	kjai ¹	kai ¹	loi ¹	lai ¹	lai ⁴⁴
＋	近	－	－	lo ³	plau ³	lau ¹³
－	鸟	ɣok ⁸	mok ⁸	nok ⁸	taɬ ⁷	ntau ³¹
＋	翅膀	－	－	bik ⁷	phik ⁷	－
－	这	nai ³	nai ⁶	nia ²	nei ²	ni ³¹
＋	这里	i ⁵ nai ³	ai ² nai ⁶	nia ² mɔɯ ⁸	nei ²	nan ⁴⁴ ni ³¹
－	一	deu ¹	dɛu ² 毛南	－	tsheu ³	si ³³
＋	三	sa:m ¹	sa:m ¹ 毛南	－	－	－

苗瑶语词例：

标记性	词项	苗语养蒿	布努语	炯奈语	畲语	瑶语龙胜
－	好	ɣu ⁵	ɣaŋ ⁵	ɣwaŋ ⁵	ɣaŋ ⁵	luŋ ⁵
＋	坏	－	－	－	－	－
－	人	ne ²	nu ²	nai ²	ne ²	mjen ²
＋	孩子	－	－	－	－	tɕwei ³
－	这	noŋ ³	nau ³	ne ³	ne ³	na:i ³
＋	这里	haŋ ³ noŋ ³	khi ³ nau ³	naŋ ¹ ne ³	ne ³ tat ⁸	na:i ³ dau ¹
－	二	o ¹	au ¹	u ¹	u ¹	i ¹
＋	四	lu ¹	tɕa ¹	ple ¹	pi ⁷	pjei ¹

当然对于具体语言，也可能缺少判断词汇标记性的统一标准。

比如，“女人”是人的生物属性，因此可以推论该词应该比作为社会角色的“母亲、妻子”等词语更加无标记。但是在共时的苗瑶语中，“女人”、“母亲”、“妻子”往往形式相同，难以区分哪个是原始的词义，哪个是后来的引申义或替换义。在侗台语中，“女人”、“妻子”在诸亲属语中都是多源分布的，而“母母亲、母牛”却在各种语言中普遍同源，是否其为最原始的无标记词，在共时语言中也无从考证。

再如各语言的数词因为反映的是比较抽象的概念，很可能都是后起的或借用的文化词，是否应该属于可供历史比较的核心词是有