

言语产生中双词素词的语音编码*

周晓林 庄捷于森

(北京大学脑科学与认知科学中心、北京大学心理学系, 北京 100871)

摘要 采用同音判断和音节监控方法,考察言语产生中双词素词语音激活的特点。选用以偏正结构的双词素词为名称的图片和与双词素词中首尾两个词素同音的两组探测字作为实验材料,探测字在图片呈现后 50 毫秒(实验一),或图片呈现前 1300 毫秒(实验二)出现,要求被试判断出现的字是否与图片名称中任何一个词素同音。实验结果与绝大部分言语产生理论的预期相反,对图片名称第一词素的反应慢于第二词素。考虑到首尾两个词素对整词意义贡献的差别,作者把实验结果解释为词素意义对语音激活的作用:第二词素的语义重要性决定了对应音节激活的快速性,双词素词中词素音位激活的速度和时间性受词义和词素意义激活程度的影响,而不完全取决于词素发音的序列性。实验不支持音位编码从左到右、序列进行的观点。

关键词 言语产生,音位编码,语音编码,序列性加工,音节监控,同音判断。

分类号 B842

1 引言

言语产生是组织交流意图、在心理词典激活表达意图的语词的表征、按照句法规则安排词序、进而控制发声器官发出声音的过程。词汇信息的激活是言语产生的关键之一。言语产生中的词汇加工主要包含词义的激活和抽象语音形式(即音位表征)的编码。语义的激活会通过表达词的语法特性的节点,传输到对应的音位表征上,从而使得进一步的具体语音编码和发声成为可能^[1~3]。音位编码严格说来应该被称为词法——音位编码。对多词素词来说,音位编码必须考虑到不同词素之间音位激活的时间和序列关系。虽然我们的具体发音是序列性的,总是先发出一个音节,再发出第二个音节,但抽象的音位编码却不一定有类似的序列性。多词素词不同词素的音位表征也可能是后面的词素先激活,或多个词素同时激活。那么,多词素词的音位表征到底是同时被激活,后面的先激活,还是存在一个“从左到右”、序列激活的时间关系?多词素词的词汇结构、词素之间的语义和语法关系是否会对音位编码的序列性产生影响?这些问题是任何一个完整

的言语产生理论都必须回答的问题,也是汉语言语产生研究所规避不掉的问题。汉语中超过 70% 的词是由两个词素、两个音节组成的合成词^[4],这些词一般说来都具有语义不透明性,整词的意义一般与作为组成成分的词素的意义具有一定的关系,但不简单地是由词素意义组合而成,整词意义有自己的语义特征,是一个独立的语义单位^[5,6]。要想了解汉语言语产生过程,我们就必须探讨汉语双词素词语音编码的属性。

西方学者在对英语和荷兰语的言语产生研究基础上,提出了多词素词音位编码以词素为单位,“从左到右”,依次进行的观点^[7~10]。与此相应,他们也认为,在词素内部,音节之间或音节之内的音位之内也存在“从左到右”的序列编码过程^[11]。支持这种观点的实验证据主要来自内隐启动(implicit priming)法的使用,即先呈现一组启动词—目标词词对(如“书本—图画,箱子—屠夫,酒杯—途径”),要求被试记住每个词对中前后两个词的联结关系。在正式实验时只呈现启动词(如“书本”),要求被试尽可能快地说出它所对应的目标词(“图画”)。当目标词的首音节之间具有同音规律(“图,屠,徒”)时,与无

收稿日期:2001-05-07。

*本研究得到国家攀登计划(批准号:95-专-09)、国家自然科学基金(30070260)、教育部博士点基金(99000127)、科学技术重点项目基金(01002)和高等学校骨干教师基金的资助。

关控制组(无规律组)相比会出现明显的内隐启动效应^[10,11],即有规律性的音节发音时间短于无规律性的音节;如果目标词的首音节之间没有任何关系时,则不存在启动效应;当目标词的尾音节之间具有同音关系(如“书本—机器,箱子—尾气,酒杯—放弃”中的“器,气,弃”)时,也不存在启动效应。如果语音编码过程是序列性的,那么被试在发音时只能先准备首音节,再准备尾音节,在内隐启动实验中出现首音节的促进效应也就不足为奇了;不出现尾音节的促进效应是因为发首音节时,尾音节还没有准备好,不会影响命名潜伏期。上述实验结果因而被解释为支持语音序列性编码的观点。Meyer^[11]研究音节内的语音编码的时间进程时发现,词素内语音编码也存在“从左到右”的序列性特点。但内隐启动法存在一个致命的缺陷,即难以保证纯粹的言语产生过程,因为从一个词到另一个词的语音过程,可以像正常言语产生那样经过语义通达语音,也可能不经过语义的提取,而通过在两个词之间建立的机械的联想关系来实现。

有关词长效应的研究也有助于探索语音编码的序列性问题。一般认为,人在发声之前必须激活一定量的语音信息,有人甚至认为只有在全词音位编码完成后才能开始发声^[12]。如果语音编码是按“从左到右”的顺序依次进行,那么多音节词的语音编码过程所需时间就要比单音节词多,开始发声前的潜伏期就较长,即存在词长效应。有人^[13,14]选用两组图片材料,一组图片具有单音节名称,另一组具有双音节或多音节名称,而且双(多)音节名称的起首音节是单音节图片名称,在图片命名任务中发现双(多)音节组反应时明显长于单音节组;而 Bachoud-Levi 等人^[15]使用同样的方法却没有发现类似的词长效应。在汉语研究中,庄捷、周晓林^[16]选取能同时用单音节(如“耳”)和双音节(“耳朵”)命名的图片作刺激材料,发现双音节命名明显慢于单音节命名。作者把这个效应解释为支持语音序列性编码的观点。有关词长的研究似乎还没有定论,更重要的是,图片命名中的词长效应是否能反映语音编码的序列性这一问题值得探讨。图片命名法与内隐启动法一样,要求被试必须发出声音,因而不可避免地涉及具体的、序列性的发声阶段。现在不清楚内隐启动效应和图片命名中的词长效应在多大程度上反映了抽象的音位编码的序列性,在多大程度上反映了具体发音准备(音节化)和发声过程的序列性。

为了克服上述研究方法上的不足,本研究采用

同音判断和音节监控方法,选用与图片双字词名称(如“飞机”)中首词素和尾词素分别同音的两组探测字(“非”和“击”),要求被试判断图片名称中是否包含探测字的语音,考察首词素组和尾词素组两种条件下反应时是否存在差异。如果双词素词的音位编码按“从左到右”的顺序进行,则首音节/fei 1/应比尾音节/ji 1/激活得早,被试判断图片名称中是否有/fei 1/这个音的速度应快于判断是否有/ji 1/这个音。这种方法由于不涉及具体发声阶段,与内隐启动法和图片命名法相比,能更明确、清晰地回答汉语词汇产生中抽象音位编码的序列性问题。

2 方 法

2.1 实验材料和实验设计

本实验共使用 132 幅名称为双字词的图片,除个别图片外,多数来自经舒华等人修订的 Snodgrass 和 Vanderwart 的标准图^[17],所选用的图片具有非常高的图片命名一致性、熟悉性、表象一致性和视觉复杂性。实验材料分为关键材料和填充材料两大类。关键材料包含 48 幅图片,其名称均为偏正结构的双字词(如“鳄鱼”),即第一个词素修饰第二个词素;选取与图片名称首词素、尾词素分别同音的两组汉字作为探测材料,同时匹配这两组探测字的字频和笔画数。探测字与图片名称没有字形、字义等其他关系。这两组探测字中均不存在多音字,避免了发音的不准确性。考虑到汉语一个音节对应多个词素,音节的频率和同音词素的多少会影响词汇加工的特点^[16,18],我们检查了图片名称中词素和探测字的音节频率和同音词素数量。音节频率是指某一音节所对应所有汉字(词素)的频率之和。考虑到实验材料频率的算术平均数易受极端值的影响,我们同时计算了各种频率的中数。本研究所采用的所有频率资料均来自《现代汉语频率词典》,以一百八十分之一为单位^[4]。这些频率和数值并没有完美匹配(见表 1),总的说来偏向第一词素,使得第一词素的反应时更有可能快于第二词素。填充材料分为两类,第一类包含 18 幅图片,名称结构有连绵式(“玫瑰”)、偏义在前(“月亮”)和并列式(“开关”)三种;选取与图片名称首词素、尾词素分别同音(但没有字形、字义等其他关系)的两组汉字作为探测材料。我们很希望把偏义在前的图片作为关键实验材料,但受材料可能性和实验设计要求所限,我们找不到足够数量的此类材料。第二类包含 66 幅图片以及与其图片名称中任何一个词素都不同音(同时没有字

形、字义等其他关系)的一组探测字。

表 1 关键材料和探测字字频、音节频率和笔画数的平均数(括号内为中数)

材料性质	图片名称		探测字	
	首词素	尾词素	首词素	尾词素
材料举例	飞	机	非	击
词素频率	1294 (582)	441 (239)	323 (223)	353 (191)
音节频率	3822 (2199)	3591 (1570)	3822 (2199)	3591 (1570)
同音词素数量	7.4 (7.5)	9.7 (8)	7.4 (7.5)	9.7 (8)
笔画数			8.3 (8)	8.5 (8)

以拉丁方设计把 48 项关键材料和 18 项第一类填充材料按探测字是探测首词素(音节)还是尾词素(音节)分别交叉分成两个测验组。每个测验组包含所有图片,一半是探测首词素,另一半是探测尾词素,每个被试只接受一个测验组和所有的实验条件(首词素和尾词素组探测条件)。被试对以上 66 项材料应做“ Yes ”判断。66 幅第二类填充图片及相应的单字探测材料均相同地出现在两个测验组中,被试对其应做“ No ”判断。

2.2 实验程序和步骤

本研究由两个实验组成,实验步骤根据汉字是在图片前出现还是在图片后出现而有所不同。根据已有研究结果^[19],被试在看到图片 50 毫秒后,图片名称的语音信息就已经开始得到激活,因此,实验一选取 SOA 为 50 毫秒。实验材料在计算机屏幕正中间顺序呈现。首先呈现“ + ”300 毫秒,接着空屏 300 毫秒,然后呈现图片,图片呈现 50 毫秒时,在图片中间加入探测字(图片保持不变),图片及探测字在被试作出反应的同时在屏幕上消失。被试对探测词素是否与图片名称中任何一个词素同音作出判断,即,如果探测词素和图片名称中的首词素或尾词素同音,则按下反应盒的“ Yes ”键,如果探测词素和图片名称中的两个字都不同音,则按“ No ”键。被试必须在 2 秒之内做出反应,否则算错,每两个项目之间的时间间隔为 4 秒。实验实施使用 DMDX 系统,该系统呈现与计时精度均为 1 毫秒。

实验二采用音节监控方法,在正式实验中先呈现探测字 1000 毫秒,汉字消失后空屏 300 毫秒,然后呈现图片,要求被试在看到图片后立即判断探测字与图片名称中任何一个汉字是否同音。被试反应及计时方式同实验一。

两个实验在具体操作时均分为两个阶段,第一为训练阶段,要求被试熟悉实验中所有图片并记住其相应的双词素名称,被试有充足的时间学习图片名称。第二阶段为正式实验,要求被试在看到图片中间出现的探测字(实验一)或图片(实验二)时立即作出判断。计算机记录下被试的反应时和错误率,主试对被试的反应情况进行详细的记录。

2.3 被试

被试为北京大学 62 名本科生(实验一、二各 31 人),北方人,普通话标准,裸视或矫正视力正常,以前没有参加过类似的实验,实验结束后获得少量报酬。

2.4 结果分析

实验一所有被试在关键材料上的错误率及所有关键项目的错误率没有超过 25%,在对实验数据进行分析时没有删除任何被试或项目。在实验二中,对图片“袋鼠”和探测字“属”作判断时,错误率达 100%,故删除该项目。所有被试在关键材料上的错误率均低于 25%,进行数据分析时保留所有被试的数据。被试的反应时和错误率见表 2。

表 2 被试在首词素组和尾词素组两种条件下的平均反应时(ms)和错误率(%)

实验	指标	首词素组	尾词素组	效应量
实验一	反应时	967	904	63
	错误率	13.7	6.9	6.8
实验二	反应时	743	703	40
	错误率	9.2	4.1	5.1

对实验一反应时进行被试内(t_1)和项目间(t_2)的 t 检验,发现首词素组和尾词素组条件之间的反应时存在显著差异, $t_1(30) = 3.93$, $p < 0.01$, $t_2(47) = 2.87$, $p < 0.01$,被试对尾词素组条件的反应要明显地快于首词素组。在错误率分析中表现出相同的模式,即首词素组和尾词素组条件之间的错误率存在显著差异, $t_1(30) = 4.22$, $p < 0.01$, $t_2(47) = 3.32$, $p < 0.01$,被试在尾词素组条件下的错误率明显少于首词素组。

在实验二中,对反应时进行 t 检验,结果表明首词素组和尾词素组条件下的反应时之间存在显著的差异, $t_1(30) = 3.82$, $p < 0.01$, $t_2(46) = 2.49$, $p < 0.05$,被试对尾词素组条件材料的反应明显快于首词素组。错误率分析也表现出相同模式,首词素组和尾词素组条件下的错误率也存在显著差异, $t_1(30) = 4.45$, $p < 0.01$, $t_2(46) = 2.78$, $p <$

0.01,被试在尾词素组条件下的错误率明显低于首词素组。

3 讨 论

本研究在即时的同音判断(实验一)和音节监控(实验二)两种任务下均发现,对尾词素组材料的反应时明显快于首词素组。这一结果与双词素音位信息序列激活的观点不符。在讨论实验结果的理论含义前,我们首先排除下列可能性,即实验结果是由图片名称中首词素与尾词素或探测字中首字与尾字性质匹配不当造成的。从表1中可以看到,图片名称首词素的词素频率略高于尾词素,说明首词素的语义表征和音位信息的激活速度应快于尾词素。而首词素的音节频率高于尾词素,同音词素数量低于尾词素,表明首词素的音位信息激活也应快于尾词素,而且音位激活少受同音词素之间相互竞争的影响^[16,18](音节频率与同音词素数量应对探测字音位的加工影响较小)。如果这些特性的不完美匹配会影响不同词素的反应时间,那只会使得首词素快于尾词素,实验结果显然与此相反,说明实验结果不是这些因素造成的。

如何解释现在的实验结果?如前文所述,双词素词的词素音位信息激活有三种可能性:从左到右序列激活,同时平行激活,或尾词素先激活,首词素随后激活。本研究结果不支持从左到右序列激活的观点,似乎支持尾词素先激活的看法,但这一看法违背了大多数人的直觉预期。可以看出,我们的实验结果既不符合西方流行的序列性编码的观点,也不完全符合平行加工观点的预期^[20]。

我们认为,音位信息的激活不单是受语音本身特性的影响,也受其他因素的制约,本研究结果实际上是反映了语义激活对音位激活的作用。本实验所用关键材料都是偏正结构(如“鳄鱼”),前一个词素修饰后一个词素,第二词素即是语言学上所谓的“词头”,它的语义和语法特性在很大程度上决定了整个词的语义和语法特性。换句话说,第二词素和整词共用了许多语义特性,而第一词素与整词的语义重叠相对较少^[5]。考虑到双词素词语义表征与语音表征之间的关系,词素的音位表征不但接受整词的语义激活传输,也接受词素的语义激活的传输,词素的语义激活强度越大、速度越快,所对应的音位表征的激活强度就越大,速度就越快。因此,偏正结构双词素词第二个词素的音位表征的激活就可能强于、快于第一个词素的音位。

按照我们这种语义观点,双词素词的词素音位激活虽然可能是同时、平行开始,但其最终的表现并不一定有固定的顺序,而是可能取决于语义激活传输的强度和速度。对双词素词来说,无论词头是第一词素(如“耳朵”),还是第二词素(“香烟”),它总是先激活。由此推论,如果我们使用词头在前的双词素词,在同音判断和音节监控实验中,对第一词素的反应应该快于第二词素。遗憾的是,能用图形表达的以第一词素为词头的具体词在汉语中虽然较多,但多数以轻音“子”为第二词素(如“锤子”),不能用在本研究实验设计中。但我们提出的语义观点说明了在今后言语产生研究中区分不同结构词汇的重要性,为新的实验设计和验证提供了理论框架。

从语义观点出发,我们可以进一步推论说,如果图片双词素词名称的词素语音表征都得到充分激活,则在同音判断中,我们就应该得到另一种反应模式,即第一词素反应时间或是与第二词素相同,或是快于第二词素。为此,我们重做了实验一,但把图片呈现与探测字呈现之间的时间间隔(即SOA)加长到350毫秒(根据已有证据,言语产生中,此时语义信息不再激活,而语音信息已经得到充分激活),结果发现,首词素组反应时(759毫秒)略快于尾词素组反应时(778毫秒),但二者之间的差异(19毫秒)没有达到显著性水平($t_1(13) < 1$; $t_2(46) < 1$),与我们的预期相同。但首词素组的错误率(20%)显著高于尾词素组(9%)($t_1(13) = 4.08$, $p < 0.01$; $t_2(46) = 3.24$, $p < 0.01$)。考虑到反应速度与错误率之间的权衡(trade off),我们认为,在这样的时程下,首词素和尾词素二者音位激活之间实际上没有差异,支持我们的语义观点。

在结束本文之前,我们有必要探讨一下我们以前有关言语产生中词长效应的研究与本研究的关系。在词长效应研究中^[16],我们发现图片命名时间随图片名称的长度(以词素数量来衡量)的加长而延长,我们把这种结果解释为支持语音编码的序列性加工观点,这个结论表面上看与本研究的结果和结论相矛盾,但在词长效应研究中,被试必须发出声音,而本研究不需要,因此这两个研究探测的一个是具体语言,一个是抽象的音位表征,所反映的问题并不完全一样。综合这两个研究,我们认为词长效应体现了序列性的、具体语音的准备过程,而本研究则反映了抽象的音位编码阶段的加工特点,但这个看法有待于进一步实验验证。

参 考 文 献

- 1 Dell G S. A spreading activation theory of retrieval in language production. *Psychological Review*, 1986, 93: 226 ~ 234
- 2 Garrett M F. Processes in language production. In: Newmeyer F J ed. *The Cambridge survey of linguistics. Vol. 3. Language: Psychological and biological aspects.* Cambridge, MA: Harvard University Press, 1988. 69 ~ 96
- 3 Levelt W J M. *Speaking: From intention to articulation.* Cambridge, MA: MIT Press, 1989
- 4 *Modern Chinese Frequency Dictionary (in Chinese).* Beijing: Beijing Language College Publisher, 1986
(现代汉语频率词典. 北京:北京语言学院出版社, 1986)
- 5 Zhou X, Marslen-Wilson W, Taft M, Shu H. Morphology, orthography, and phonology in reading Chinese. *Language and Cognitive Processes*, 1999, 14: 525 ~ 565
- 6 Zhou X, Marslen-Wilson W. Lexical representation of compound words: Cross-linguistic evidence. *Psychologia*, 2000, 43: 47 ~ 66
- 7 Kempen G, Hoenkamp E. An incremental procedural grammar for sentence formulation. *Cognitive Science*, 1987, 11: 201 ~ 258.
- 8 Dell G S. The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, 1988, 27: 124 ~ 142
- 9 Levelt W J M. Accessing words in speech production: Stages, processes and representations. *Cognition*, 1992, 42: 1 ~ 22
- 10 Roelofs A. Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*, 1996, 35: 854 ~ 876
- 11 Meyer A S. The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, 1991, 30: 69 ~ 89
- 12 Levelt W J M. Models of word production. *Trends in Cognitive Sciences*, 1999, 3: 223 ~ 232
- 13 Klapp S T, Anderson W G, Berrian R W. Implicit speech in reading, reconsidered. *Journal of Experimental Psychology*, 1973, 100: 368 ~ 374
- 14 Santiago J, MacKay D J, Palma A, Rho C. Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes*, 2000, 15: 1 ~ 44
- 15 Bachoud-Levi A-C, Dupoux E, Cohen L, Mehler J. Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, 1998, 39: 331 ~ 346
- 16 Zhuang J, Zhou X L. The word length effect in speech production of Chinese (in Chinese). *Acta Psychologica Sinica*, 2001, 33(3): 214 ~ 218
(庄捷, 周晓林. 言语产生中的词长效应. *心理学报*, 2001, 33(3): 214 ~ 218)
- 17 Shu H, Cheng Y S, Zhang H C. The naming consistency, familiarity, representation consistency and visual complexity of 235 pictures (in Chinese). *Acta Psychologica Sinica*, 1989, 21(4): 389 ~ 396
(舒华, 程元善, 张厚粲. 235 个图形的命名一致性、熟悉性、表象一致性和视觉复杂性评定. *心理学报*, 1989, 21(4): 389 ~ 396)
- 18 Zhou X, Marslen-Wilson W. Words, morphemes and syllables in the Chinese mental lexicon. *Language and Cognitive Processes*, 1994, 9: 393 ~ 422
- 19 Zhou X, Zhuang J. Lexical tone in the speech production of Chinese words. In: Yuan B, Huang T, Tang X ed. *Proceedings of the 6th International Congress of Spoken Language Processing. Vol. 2.* Beijing: China Military Friendship Publisher, 2000. 51 ~ 54
- 20 Yu L. Phonological representation and processing in Chinese spoken language production (in Chinese). Dissertation for Doctor's degree, Beijing Normal University, 2000
(余林. 汉语语言产生中的语音表征与加工. 北京师范大学博士论文, 2000)

PHONOLOGICAL ACTIVATION OF DISYLLABIC COMPOUND WORDS IN THE SPEECH PRODUCTION OF CHINESE

Zhou Xiaolin , Zhuang Jie , Yu Miao

(Center for Brain and Cognitive Science , and Department of Psychology , Peking University , Beijing 100871)

Abstract

A homophone judgment task (Experiment 1) and a syllable monitoring task were used to investigate the sequentiality of phonological activation in the speech production of Chinese disyllabic compound words. In these tasks, a picture with a disyllabic compound name was presented, followed (Experiment 1) or preceded (Experiment 2) by a Chinese character homophonic to one of the constituent morphemes. The SOA between presentation of the picture and of the character was 50 ms for Experiment 1 and 1300 ms for Experiment 2. Subjects were asked to judge whether a morpheme in the picture name was homophonic to the character. If phonological activation of compound names is conducted sequentially, from left to right, the reaction time for the first constituent morphemes should be shorter than the time for the second morphemes. In both experiments, however, the reaction time for the second constituents was shorter than for the first constituent, and the response error rate was also lower for the first than for the second, in contradiction to predictions of most existing theories of speech production. These findings were interpreted as reflecting the effect of semantic activation of constituent morphemes on phonological activation. Critical compound words used here had a structure in which the first constituents modify the second constituents, with the second constituents as lexical head and determining the semantic and syntactic properties of the whole words. We propose that phonological activation of constituent morphemes in the speech production of Chinese compound words depends on the efficiency of semantic activation of constituent morphemes, but not on the fixed order between morphemes.

Key words speech production, phonological encoding, sequential processing, syllable monitoring, homophone judgment.