

·综述·

## 言语产生的成分及时间进程的电生理研究\*

萧演清<sup>1</sup> 宋为群<sup>1,2</sup> 汪洁<sup>1</sup>

语言是人类特有的认知功能,而言语产生是人类语言能力的重要组成部分。从19世纪下半叶,尤其是近三十年认知心理学和认知神经科学兴起之后,随着神经影像学、神经电生理学的发展,为研究语言的产生过程和机制提供了重要手段,获得了重大的成果。目前认为言语产生(说话)是从组织交流意图、激活概念、提取词义、句法和语音信息,到控制发音器官发出声音的过程。词汇产生被划分为5个阶段,在每个阶段的具体细节和不同阶段的实时加工关系上,众多学者由于研究目的、研究方法不同,以致结果不同,甚至矛盾。本文重点对神经电生理在言语产生成分及其时间进程的研究现状进行综述,并分析探讨词汇提取过程的两种模型。

### 1 言语产生过程

目前,关于言语产生的时间进程及其成分较为广泛认可的是Levelt, Roelofs和Meyer(简称LRM)等<sup>[1]</sup>提出的阶段加工学说,具体的各个阶段及其时间进程如下:

#### 1.1 概念准备(conceptual preparation)

要说出一个有意义的词,先要激活相应的词汇概念并进行选择。例如,当进行图命名时必须先认识图所表示的事物,并且选择正确的概念。传统的语误和图命名研究表明,视觉输入后激活词汇概念是一个复杂的过程<sup>[2]</sup>。一幅“绵羊”的图画,激活的不仅是“绵羊”这个概念,可能还会激活诸如“动物”和“山羊”之类的概念,说者应选择哪一个概念来表达,这取决于谈话的情景以及实验的任务。例如在分类任务中,被试会选择层次高的概念如“动物”、“植物”这类属性,而在常规的命名任务中,被试会更多的选择低层次的概念如“羊”、“牛”等。这一选择策略称为“观点采择”(perspective taking)<sup>[3]</sup>。

有的研究者采用反应-不反应(go/no-go)任务范式进行事件相关电位(event-related potential, ERP)研究<sup>[4-5]</sup>。任务分别是要求被试根据图画是动物还是非动物,进行分类和词图匹配,得到属性概念和词概念准备的潜伏期均为150ms。另有研究者报道的潜伏期略为长一些<sup>[6]</sup>。他们用类

似Thorpe等<sup>[4]</sup>的实验方法,根据图画是否具有“动物”的属性作按键或不按键反应,但被试需同时根据词的语义信息和语音信息做决定。其中单侧准备电位(lateralized readiness potential,LRP)与反应准备有关,是用来测量按键这一结果的指标;N200与反应抑制有关,是用来测量不按键这一结果的指标。N200出现的平均时间大概在图画出现后206ms。换句话说,此时,被试已经收集了充足的信息否定了图画具有动物这一属性并抑制了按键的反应。由于抑制的过程本身需要时间,因此206ms这一时间段反映了提取动物这一概念的潜伏期的上限。总结以上实验结果,从视觉输入到词概念的提取,这一时间段约在图画呈现后150—200ms,估计的中位数是175ms<sup>[2]</sup>。

#### 1.2 词条选择(lemma retrieval or lexical selection)

概念表征激活后的下一阶段是词汇表征的通达,词汇表征通常也称作词条(lemma),主要指词汇的句法表征。例如,“goat”的词条内容包括它是一个可数名词、有单复数形式,另外在德语和法语中名词是有性别之分的,都属于句法表征。概念表征的激活传递到心理词典中相应的词汇表征上,即词条被概念输入激活,概念表征激活越多,相应的词条激活也就越多,目标词条选择的潜伏期就越长;反之越短。最后,目标词条因激活最高而被选择。

词条选择竞争模型得到了许多实验的支持<sup>[7]</sup>。根据这一模式,Glaser等<sup>[8]</sup>利用不同的SOA的语义干扰实验得出词条选择的时间窗约100—150ms。Schmitt等<sup>[9]</sup>在一项ERP的研究中,要求被试做语义和句法双重选择任务。两种不按键的情况下,N200出现的峰值分别在477ms和550ms。LRM的理论认为N200出现的两峰值的73ms的差别反映了词条提取时间窗的上限,由于没有一个适宜的下限,通过相减得到词条选择所需时间约为0—150ms。Level等<sup>[2]</sup>的估计中位数是75ms。结合之前关于词汇概念提取的估计时间窗150—200ms,词条选择应在图像呈现后150—200ms开始,在150—350ms之间的某个时间点结束。

#### 1.3 词形编码(form encoding)

DOI:10.3969/j.issn.1001-1242.2012.04.023

\*基金项目:国家自然科学基金(30540058, 30770714);北京市自然科学基金(7052030);北京市委组织部优秀人才基金;北京市科技计划项目(Z0005187040191-1);首都医学发展科研基金(2007-2068)

1 首都医科大学宣武医院康复科,100053; 2 通讯作者  
作者简介:萧演清,女,硕士研究生; 收稿日期:2011-08-09

目标词条被选择后将进一步激活词形编码系统。词形编码是一个从提取音位代码(phonological code)到出现词汇发音的过程。目标词条的激活进一步传递到靶词的语音表征上并进行词法-音位提取(morpho-phonological code retrieval),即以词素为单位提取音位代码。词的音段信息和韵律结构分别提取。音位片段按“从左到右”的方法插入抽象的韵律结构中,结合成词的音节。一般认为一些经常使用的音节发音程序储存在称为“心理字音表(mental syllabary)”的“仓库”中。在音节化的过程中,“仓库”里对应的音节发音程序一旦被激活,就马上转化为动作指令。

关于词形编码的整个时间段可以通过相减的方法得到。LRM理论认为,音位代码的提取在词条选择之后开始。根据上述的估计值,词条选择在图画呈现后约150—350ms内结束。用图命名的整个时间段(约600ms)减去这一时间窗(150—350ms),得到首音编码时间段的上限。有研究报道,无干扰图命名的潜伏期分别为680ms、591ms<sup>[10]</sup>、567ms<sup>[11-12]</sup>,平均估计值是600ms,通过减去词条选择时间窗150—350ms,得到词形编码的时间窗为217—530ms<sup>[2]</sup>。这一范围很大,与在相减的过程中,累积了很多不确定的因素有关。

然而,词形编码本身是一个阶段性过程(staged process),其过程可细分为3个部分,并有其相应的独立表现。具体如下:

第一步,语形学编码(morphological encoding),即词法-音位提取(morpho-phonological code retrieval)。说者需提取靶词所有词素(morphemes)的音位代码。例如单词“goats”是一个复数形式,有两个词素的音位代码需进行提取,分别是词干/goat/和词尾变化/s/,对于不规则词“sheep”只需提取一个词素的音位代码。

利用ERP测得音位代码提取的潜伏期相当短<sup>[13]</sup>,测量到单侧化准备电位(LRP)表明了单词的第一个音段(phonological segment)出现在句法内容的性别信息被提取后40ms。如前述,词条选择在性别信息提取最多73ms之前,因此第一个音段出现的时间窗应在词条选择后约40—113ms,估计中位值是80ms<sup>[2]</sup>。

第二步,音节化和音律编码(syllabification and metrical encoding)。音节化是一个按“从左到右”先后顺序逐步加工的过程。例如,要把动词“consist”分成各个音节,首先把头两个音段/k/和/ən/组合在一起,产生了第一个音节/kən/,然后把后面的四个音段组合在一起构成了第二个音节/sist/,最后得到了/kən-sist/这一个由两个音节组成的词。通过LRP实验和音素自我反馈实验得到,音节的提取速度大约为每25ms一个音节。按此估计,一个有五个音节的单词其音节化的时间平均约125ms<sup>[2]</sup>。

第三步,语音编码(phonetic encoding)。如前述,音节化

是一个逐步提取的过程,音节发音程序一旦被提取,则马上转化为动作指令。一般认为,在发音之前必须激活一定量的语音信息。对于多音节词,发音动作可以在所有语音编码完成之前执行。Bachoud-Le'vi等<sup>[14]</sup>通过实验有力的论证了这一观点。至少第一个音节的语音编码要在发音启动前完成。Meyer等<sup>[15]</sup>认为发音启动前语音编码的数量由说者决定,不是一个固定的常数。由于没有一种独立测量语音编码的精密计时的方法,只能通过相减得到估计的时间窗,因此必然会有一些不定因素的累积效应。表1列出了命名任务的几个阶段的时间窗的估计值,整个的平均时间是600ms。

表1 言语产生的连续的各加工阶段的估计时间窗

加工阶段	持续时间(ms)
概念准备(从图片的呈现到目标概念的选择)	175
词条选择	75
<b>词形编码:</b>	
音位代码提取	80
音节化	125
语音编码(至发音的启动)	145
总计	600

## 2 言语产生的词汇通达理论

目前,关于言语产生的词汇通达理论有两种,一种是Level提出的独立两阶段模型,另一种是Dell提出的交互激活模型。两种模型的分歧主要在于:独立两阶段模型认为词汇层和语音层是分离的、独立的,信息的激活在这两层之间的传递是单向的,只能从词汇层传递到语音层,即不存在从语音层到词汇层的反馈。因此,独立两阶段模型预测在言语产生的早期只有语义的激活,在晚期阶段只有字形语音的激活;而交互激活模型认为激活在各个表征水平之间的扩散是双向的,语义层可以将激活同时传递到词条节点和音位节点,而音位信息又可以将激活反馈到语义层次上,从时间的角度来看,就意味着不同的编码阶段不是各自分离的时间进程,或者说不同的阶段之间存在很大的重叠。因此,交互激活模型预测在言语产生的早期阶段只有语义的激活,在晚期阶段同时存在语义和语音的激活。也就是音位编码与词条选择的产生时间存在重叠,前者可以在后者完成之前进行。

### 2.1 正常人的词汇通达研究

对舌尖现象(tip of the tongue, TOT)现象(舌尖现象或唇边现象,是指人们知道想要说什么词,但在说话时却一时无法正确提取该词的状态)的研究<sup>[16]</sup>,以及以拼音文字为对象的言语产生ERP研究和对汉语词汇的语义、语音信息提取的时间进程的ERP研究<sup>[17-20]</sup>均大多支持独立两阶段模型。然而有人认为<sup>[21-22]</sup>,对TOT现象,以及神经心理学研究观察到的语法和语音信息分离的现象的原因的猜测是没有时效性的,这些现象只是表明了词汇句法与音位信息有各自的表

征,并不能表明词汇提取过程的时间结构。因此这些观察结果并不能确定这些编码过程是相继还是同时进行。

有人认为利用 go/no-go 任务范式来研究探讨言语产生的时间进程时采用 LRP 和 N200 作为反应指标并不能直接得到这些阶段的时间窗<sup>[23-24]</sup>。原因有以下几点:首先,LRP 只是反映了语言信息传送到初级运动皮质的瞬间,N200 也只是表明具体信息处理的终点时刻,两个指标都不是信息处理的过程。其次,单侧化准备电位涉及相减计算,目的是得到语义编码,句法编码,或音位编码的相对时限。Camen C 等<sup>[23]</sup>通过计算发现被试作出反应的时刻与句法或音位信息出现的时刻(根据单侧化准备电位测量得到)并不相同。这一计算结果再次表明了句法或音位信息出现的时间并不是这些信息处理编码的时间。因此这些研究结果只是反映了信息处理的终点而不是信息处理的时限。他们通过改变实验的方法研究信息的处理过程,行为学结果表明性别监测任务(属句法内容)与首音监测任务的反应时间相同,而监测第二个音节的反应时间更长。时段分析结果表明性别监测与首音监测都在相同的时间窗,约在图画出现后 270—290ms,而监测第二个音节的的时间窗在更后,约为 480ms 处。通过直接比较性别监测任务与首音监测任务发现,两者只在约 200ms 处有 10ms 的区别。所有这些结果均提示词条的提取过程与音位的编码过程是同时或者是有重叠的。

另外对于 Van Turenhout 等<sup>[13]</sup>实验结果:句法信息比语音信息早出现 40ms。Abdel-Rahman 等<sup>[24]</sup>认为另一种解释也支持这一结果,即如果语义信息的提取与音位信息的提取是同时进行的,但是前者完成的时间比后者要早,因此也能得到两者的单侧化准备电位相差 40ms 的结果。所以这一结果同时符合独立两阶段模型与交互激活模型。Abdel-Rahman 等<sup>[24]</sup>通过提高分类任务的难度来控制语义的编码过程,通过对单侧化准备电位的间隔的分析作为对语义编码和对音位编码的持续时间的估计,结果却支持交互激活模型,因为随着分类任务难度的提高,语义任务与音位任务的间隔不是逐渐延长而是缩短。

## 2.2 失语症的词汇通达研究

无论是独立两阶段模型还是交互激活模型,实验的结果都是来自对健康人的研究,然而,利用失语患者对言语的产生过程进行研究的实验却很少,结果也很不一致。Dobel 等<sup>[27]</sup>利用句法和语义分类的静音图命名任务对 18 例失语症患者以及 23 例健康人进行大脑半球在语言功能康复中的作用的研究时发现,这两组被试在句法分类任务和语义分类任务中,ERPs 波形的区别均出现在图画呈现后 300—600ms,这一时间窗与 Levelt 等对音位信息编码的时间窗估计值相一致但却不是句法或语义的时间窗。然而,Cornelissen 等<sup>[28]</sup>通过对 3 例失语症患者言语治疗前后研究发现,患者治疗前后的变化

不仅表现在行为学上的改善,并且出现大脑皮质激活区域的改变,集中表现在图画呈现后 300—600ms 的时间窗,部位在左侧顶叶下。Cornelissen 等<sup>[28]</sup>认为,大量的局部大脑损伤研究以及功能影像学研究表明左侧顶叶下部是言语的工作记忆即音位信息的提取和编码的重要部位,从而推断出这 300—600ms 的时间段是音位信息提取编码的时间窗。这一结果与 Levelt 对音位信息编码的时间窗估计值是一致的。Laganaro 等<sup>[27]</sup>认为不同类型的失语症患者可以归因于言语产生过程不同阶段水平的损伤,不同类型的失语症患者在不同的时间窗应该有相应异常的电生理表现。因此可通过对失语症患者的研究来探索言语产生的过程。Laganaro 等<sup>[28]</sup>对两名音位性错语和两名语义损伤失语症患者进行电生理研究,波形分析及地形图分析结果发现:与正常人比较,两名音位性错语的患者异常 ERPs 波幅出现在 300—450ms 之间,两名语义损伤的患者异常的 ERPs 波幅约出现在 110—310ms 之间。因此语义损伤表现为早期出现 ERP 异常而音位损伤为晚期的 ERP 异常,这一结果提示了语义编码在前,音位编码在后。这一研究结果支持独立两阶段模型。既然言语生成过程的不通环节有其相应的时间窗及其对应的电生理特征,通过言语康复治疗促使受损环节功能恢复,在其受损环节的时间窗就应该有相应的电生理改变。Laganaro 等<sup>[28]</sup>利用 ERPs 对两名音位性错语和两名语义损伤的患者进行治疗前后电生理研究,结果表明:两名音位性错语的患者治疗前后电生理变化表现为 300ms 之前是异常波幅的正常化以及 300ms 以后异常波幅的增大。而两名语义损伤的患者治疗后电生理的变化出现在另一个时间窗,表现为 300ms 以前出现异常波幅的增大。这一结果与上述结果一致,即语义编码的启动在前而音位编码的启动在后。一名被试曾作为健康对照组而作过图命名的 ERPs 研究,一年后因左半球的脑卒中而表现为找词困难和音位损伤性失语,Laganaro 等<sup>[29]</sup>对这一患者患病前后的 ERPs 进行对比,结果发现 ERPs 电生理变化出现在 250—400ms 之间。这一结果表明了词汇通达(也就是从语义到语音的激活)在 300ms 前后,因此这些结果似乎都更支持独立两阶段模型。

## 3 小结

结合神经影像学、神经电生理学,利用不同的研究言语产生的实验范式,目前基本上确定了言语产生的五个阶段:概念准备、词条选择、音位编码、语音编码和发声。而神经电生理由于其具有时间分辨率高的优势,在言语产生的研究中,不仅提供了靶词整个准备过程的时间(例如从图片呈现到命名的启动,整个过程的时间大约为 600ms),而且提供了详细的有关中间环节的其他时间信息。但是,由于研究方法上的困难,言语产生各阶段的时间进程以及各阶段之间的时

间关系尚未达成统一认识。除了上述讨论的词条选择和音位编码这两个阶段存在争议以外,关于概念选择和词义激活阶段是自上而下,还是自下而上的关系<sup>[30-33]</sup>,音节是否存在心理词典中,所提取的音位代码是否经过音节化,以及发声是否必须在词的所有音节的语音编码完成之后才开始等<sup>[34]</sup>,仍存在着争议。在今后的研究中,尽可能通过统一的研究范式对言语产生过程作更深入的研究,使得研究结果具有更多的可比性。另外,应用认知神经科学,对失语症患者的言语障碍表现和各种表征与加工之间的分离现象进行分析,结合ERP动态研究脑活动,fMRI进行功能定位以及PET研究功能区神经分子的分布和代谢等,推断言语产生所需要的加工过程,这将是未来研究的一个主要方向。

### 参考文献

- [1] Levelt WJ, Roelofs A, Meyer AS. A theory of lexical access in speech production[J]. *Behav Brain Sci*, 1999, 22(1): 1—38; discussion 38—75.
- [2] Indefrey P, Levelt WJ. The spatial and temporal signatures of word production components[J]. *Cognition*, 2004, 92(1—2): 101—144.
- [3] Clark EV. Conceptual perspective and lexical choice in acquisition[J]. *Cognition*, 1997, 64(1): 1—37.
- [4] Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system[J]. *Nature*, 1996, 381(6582): 520—522.
- [5] Jescheniak JD, Levelt WJM. Word frequency effects in speech production: retrieval of syntactic information and of phonological form[J]. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1994, 20: 824—843.
- [6] Schmitt BM, Munte TF, Kutas M. Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming[J]. *Psychophysiology*, 2000, 37(4): 473—484.
- [7] Roelofs A. A spreading-activation theory of lemma retrieval in speaking[J]. *Cognition*, 1992, 42(1—3): 107—142.
- [8] Glaser WR, Dungenhoff FJ. The time course of picture-word interference[J]. *J Exp Psychol Hum Percept Perform*, 1984, 10(5): 640—654.
- [9] Schmitt BM, Schiltz K, Zaake W, et al. An electrophysiological analysis of the time course of conceptual and syntactic encoding during tacit picture naming[J]. *J Cogn Neurosci*, 2001, 13(4): 510—522.
- [10] 周晓林, 于森. 言语产生中双词素词的语音编码[J]. *心理学报*, 2002, 34(3): 242—247.
- [11] Levelt WJ, Praamstra P, Meyer AS, et al. An MEG study of picture naming[J]. *J Cogn Neurosci*, 1998, 10(5): 553—567.
- [12] Damian MF, Vigliocco G, Levelt WJ. Effects of semantic context in the naming of pictures and words[J]. *Cognition*, 2001, 81(3): B77—86.
- [13] Turenout M, Hagoort P, Brown CM. Brain activity during speaking: from syntax to phonology in 40 milliseconds[J]. *Science*, 1998, 280(5363): 572—574.
- [14] Bachoud-Levi AC, Dupoux E, Cohen L, et al. Where is the length effect? A cross-linguistic study of speech production[J]. *Journal of Memory and Language*, 1998, 39, 331—346.
- [15] Meyer AS, Roelofs A, Levelt WJM. Word length effects in picture naming: The role of a response criterion[J]. *Journal of Memory and Language*, 2003, 48, 131—147.
- [16] Kemmerer D. Auxiliary selection in Italian: a comment on Miozzo and Caramazza's "On knowing the auxiliary of a verb that cannot be named: evidence for the independence of grammatical and phonological aspects of lexical knowledge" [J]. *J Cogn Neurosci*, 1998, 10(3): 421—423.
- [17] 杨闰荣, 韩玉昌, 曹洪霞. 言语产生过程中语义、语音激活的ERP研究[J]. *心理科学*, 2006, (6): 1444—1447.
- [18] 郭桃梅, 卢春明. 汉语词汇产生中的义、音信息提取时间进程的ERP研究[J]. *心理学报*, 2005, (5): 569—574.
- [19] Zhang Q, Damian MF. The time course of semantic and orthographic encoding in Chinese word production: an event-related potential study[J]. *Brain Res*, 2009, 1273: 92—105.
- [20] Zhang Q, Damian MF. The time course of segment and tone encoding in Chinese spoken production: an event-related potential study[J]. *Neuroscience*, 2009, 163(1): 252—265.
- [21] Schriefers H, Jescheniak JD. Representation and processing of grammatical gender in language production: A review[J]. *Journal of Psycholinguistic Research*, 1999, 28(6), 575—600.
- [22] Caramazza A. How many levels of processing are there in lexical access[J]? *Cognitive Neuropsychology*, 1997, 14: 177—208.
- [23] Camen C, Morand S, Laganaro M. Re-evaluating the time course of gender and phonological encoding during silent monitoring tasks estimated by ERP: serial or parallel processing[J]. *J Psycholinguist Res*, 2010, 39(1): 35—49.
- [24] Abdel Rahman R, Sommer W. Does phonological encoding in speech production always follow the retrieval of semantic knowledge? Electrophysiological evidence for parallel processing[J]. *Brain Res Cogn Brain Res*, 2003, 16(3): 372—382.
- [25] Dobel C, Pulvermuller F, Harle M, et al. Syntactic and semantic processing in the healthy and aphasic human brain [J]. *Exp Brain Res*, 2001, 140(1): 77—85.
- [26] Soros P, Cornelissen K, Laine M, et al. Naming actions and objects: cortical dynamics in healthy adults and in an amnic patient with a dissociation in action/object naming[J]. *Neuroimage*, 2003, 19(4): 1787—1801.
- [27] Laganaro M, Morand S, Schwitler V, et al. Electrophysiological correlates of different anomic patterns in comparison with normal word production[J]. *Cortex*, 2009, 45(6): 697—707.
- [28] Laganaro M, Morand S, Schwitler V, et al. Normalisation and increase of abnormal ERP patterns accompany recovery from aphasia in the post-acute stage[J]. *Neuropsychologia*, 2008, 46(8): 2265—2273.
- [29] Laganaro M, Morand S, Michel CM, et al. ERP correlates of word production before and after stroke in an aphasic patient [J]. *J Cogn Neurosci*, 2010, 22(2): 374—381.
- [30] Grill-Spector K, Kanwisher N. Visual recognition: as soon as you know it is there, you know what it is[J]. *Psychol Sci*, 2005, 16(2): 152—160.
- [31] Murphy GL, Brownell HH. Category differentiation in object recognition: typicality constraints on the basic category advantage[J]. *J Exp Psychol Learn Mem Cogn*, 1985, 11(1): 70—84.
- [32] Fei-Fei L, Iyer A, Koch C, et al. What do we perceive in a glance of a real-world scene[J]. *J Vis*, 2007, 7(1): 10.
- [33] Mace MJ, Joubert OR, Nespoulous JL, et al. The time-course of visual categorizations: you spot the animal faster than the bird[J]. *PLoS One*, 2009, 4(6): e5927.
- [34] 周晓林, 于森. 言语产生中双词素词的语音编码[J]. *心理学报*, 2002, (3): 242—247.