

言语产生研究的理论框架*

周晓林

庄捷舒华

(北京大学心理学系,100871) (北京师范大学心理学系,100875)

摘要 本文回顾了言语产生的研究历史,总结了言语产生的理论框架和研究方法,对言语产生研究中的一些关键问题进行了系统的分析,这些问题包括:(1)语义如何存储在心理词典中;(2)词条选择和音位编码的独立性;(3)语音和音位编码方式;(4)音位信息与音节框架之间的关系。

关键词:言语产生 交互作用 语音编码 音位编码

0 引言

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。它是人类语言能力的重要组成部分,是人类意识和意志表达的重要途径。生物的进化促进了人类大脑体积和复杂性的增加,使得人类在三十万年前具有说话和符号操作的能力。然而,直到十九世纪下半叶,人们对个体言语产生的过程才有所了解,而真正的科学研究则是在近几十年认知心理学和认知神经科学兴起之后。时至今日,由于研究方法上的困难,我们对言语产生过程及其神经基础还处于探索阶段,我们心理学界和神经科学界对这一重要问题也无系统的研究。本文的目的在于介绍言语产生研究的历史来源和方法、理论模型和框架,分析探讨目前存在的问题,为今后的研究提供支持和帮助。

1 研究的历史来源和方法

对言语产生的研究有两种历史来源,一是对语误的分析,二是以图形命名为主的实验室研究。1895年,Meringer和Mayer发表的关于德语语误的语料库及其理论分析,成为语误研究的开端。弗洛伊德(1901)的《日常生活中的精神病理学》是对语误的心理学研究的经典著作。70年代在洛杉矶加州大学和麻省理工学院出现了两个规模较大的语误语料库。1966年,Brown和McNeill找到诱导被试产生TOT(the tip of tongue,舌尖现象)状态的方法,随后,Baars等人提出在实验控制条件下诱导、产生语误的方法。70年代以后出现了很多英语语料库,随

后荷兰语、德语、西班牙语及法语等都出现了各自的关于语误的语料库。语误分析法是言语产生研究中的基本方法之一,通过搜集人们实际语言交流中出现的口误材料,从不同的角度进行分类,可以揭示语言在大脑中的表征方式和语言产生的具体心理过程。已有的口误研究^{[1][2]}发现,句子的产生首先要确定句子的基本结构,包括句子的主次重音;词项的填入是实词在先,虚词在后。语误分析对于探讨具有汉语特色的一些理论问题也具有重要意义,如沈家煊^[2]通过多方面的口误材料,证明声调在音节中的独立性。

1885年,Cattell发现命名图形所用的时间是命名图形所对应字词的两倍,由此开始了对言语产生的实验室研究。1967年,Fraisse发现,当命名一个小圆为“圆圈”时的反应时是619毫秒,当命名为“oh”时反应时仅为453毫秒。实验在排除图形编码的影响后,说明字词的命名可以从字形直接通达语音,而图形的命名要先激活物体对应的概念,再通达其语音。Rosinski等人将Stroop任务转化成图——词干扰范式,包括两种任务:一是排除图形中字词的干扰作用,命名图形的名称;二是排除图形自身名称的干扰,命名图形中的字词。研究表明,语义相关词对图形命名的干扰作用远远大于图形名称对图形中字词命名的作用。Lupker^[3]发现了词汇产生中的语音促进效应,即干扰词与图形名称有相同韵律关系(包括头韵关系)时,命名速度变快。从八十年代末至今,对言语产生过程及其脑机制的探讨逐渐成为语言心理学和认知神经科学研究的重点和热点之一。目前,对言语产生的研究重点主要集中在词汇产生方面,但言语产生中句子与句法加工亦

* 本研究得到国家自然科学基金(批准号30070260)、教育部博士点基金(批准号:99000127)和高等学校骨干教师基金的资助。

在逐渐为人重视。

就研究的方法论而言,图片命名法是言语产生实验研究中最常用的方法,因为它能够准确地反映言语产生从获得概念,激活相应的语义、语音信息,到控制发声器官发出声音的完整过程。具体做法是,在呈现图片时加入与图片名称具有一定关系的干扰字词,要求被试以最快的速度读出图片名称,通过和控制条件相比,来考察言语产生的具体过程。干扰字词最初以视觉方式呈现,后来出现听觉呈现方式。干扰字命名法在图片命名法的基础上发展而来,要求被试命名图片中呈现的干扰字词,而不是图片名称。

此外,研究者还提出一些新的方法,如符号命名法,内隐启动(implicit priming)法、翻译(translation)法等。符号命名法要求被试在实验前通过学习建立特定符号(如“++”)与一些字词(如“土地”)的联结关系,在正式实验中只呈现特定符号,被试尽可能快地命名它所对应的字词。上述方法突破了图形命名方法中图形实验材料的限制,扩展了研究的领域,但它们难以保证纯粹的言语产生过程,即从特定符号到字词语音的过程,可能经过语义通达语音,也可能不经过语义的提取,在特定符号和字词语音之间直接建立联结关系。

2 言语产生的理论框架和模型

言语产生过程大致可分为三个层次^[4]。最高的层次是概念化(conceptualization),此过程确立说话的意图和想表达的概念;中间的过程是言语组织(formulation),即把要表达的概念转换为语言形式;最后一个过程是发声(articulation),涉及到更为具体的语音和发音的计划,即把语音编码转换成发声的肌肉运动程序,并执行这一程序。言语组织包括词汇生成和语法编码两部分。语法编码指句子的选词和排序,即根据词汇的意义和语法性质选择恰当的词汇,并产生一句话的句法框架。而词汇产生或词形式的编码是言语产生研究的中心问题之一。这一过程可细分为五个组成部分^[5],即概念准备(conceptual preparation)、词条选择(lexical selection)、音位编码(phonological encoding)、语音编码(phonetic encoding)和发声(articulation),虽然研究者经常把最后三个部分合为一体。

概念准备是指大脑把思想、观点等转化成概念的过程,这也是言语产生的最初环节。人在准备表达某个思想时,需要从相关的大量信息中挑选最恰

当的词汇概念(lexical concept),这个概念必须对应心理词典中的词或词素。比如在表达“有四条腿、嗅觉灵敏、会汪汪叫的动物”时,最简洁的是“狗”。词汇化(lexicalization)和词条选择在概念准备完成后发生。准备好的概念将会激活心理词典中对应的语义表征(如表达“狗”的语义特征),词汇化把激活的语义表征转换成词的音位表征。一般认为,在词义与音位表征之间有一层抽象的表征,称为特定词汇表征(lemma,或称词条),这一表征涵盖了该词的语法特征。词汇化包含两个阶段(不同的观点请见Caramazza^[6]),第一阶段为语义表征的激活传输到中间层的特定词汇表征上,即多个词条被概念输入激活,经过一段时间后,激活最高的词条被选择,词条的激活将词汇的语法特性提供给句子的句法编码,供其使用;第二阶段是音系形式(phonological form)通达,即中间层的激活进一步传输到特定词汇的语音表征上,使得说话者能够提取词汇的语音。受到激活或提取的音位表征将进一步进行词法—音位编码(morpho-phonological encoding)。在这一阶段,单词的词汇结构、韵律特征和音段组成被展开,词素单位内的音位信息被组合成符合实际发音情况的音节。一般认为,这个音节化(syllabification)是一个“从左到右”,按时间先后顺序逐步加工的过程,而词的音段信息和韵律结构(如CV结构、轻重音)是分别提取的。音位片段将以“从左到右”的方法插入抽象的韵律结构中,结合成词的音节。在语音编码和发声之间,可能还有音节发声程序的提取。一种语言经常使用的音节的抽象发声程序被存储在一个“仓库”中,这些音节程序发声时将“从左到右”依次被提取。激活这些程序也归因于语音编码的激活传播。音节发声程序一旦被提取,一个词的发声就立即开始。发声系统计算出执行抽象音节发声程序的最佳方案,指挥呼吸系统和咽喉系统发出声音。

目前,关于言语产生的理论模型都是符号化的、具体的,而不是分布式的、抽象的。具有重要影响的两个主要模型是:Dell^[4]的两阶段交互激活模型和Levelt^[5]、Roelofs^[7]等的独立两阶段(或序列加工)模型。

两阶段交互激活模型把从语义到语音的提取过程主要分为三个层次(Dell等人^[8]),每一层次上都有大量的结点,最上层为语义层,其结点代表语义特征,中间是lemma层,最下层为音素层,由起首音素(onset phoneme)、核心部分(nucleus phoneme)、结尾音素(coda phoneme)等构成。该模型主要贡献是在

语言产生中提出交互激活的思想。激活的传输方式是瀑布式(cascades)的,即激活可以在网络中不同层次间的任何可能的结点之间传递。“交互”是指所有结点之间的联结是双向的,它们之间的联结只有促进作用,没有抑制作用;而且激活的传递也是双向的,可以从高层次向下传递,也可以从低层次向上传递。受到最大激活的词或 lemma 结点总是得到选择,这种选择是由外部形成的语法结构决定的。被选择的 lemma 结点将把激活传递到对应的语音编码上。这种模型的提出受到解释语误中混合错误的启发和影响。混合型错误是指实际产生的词和目标词之间既存在语音的关系,又存在语义的关系,如误把 cat 说成 rat。按照交互激活模型,cat 和 lemma 结点得到激活后向下传递到 lexeme 层,激活/k/、/æ/、/t/结点,另一方面,在 lemma 层与 cat 语义相关的动物(包括 rat)的结点也被激活,rat 被从 lexeme 层反馈回来的/æ/、/t/再次激活,因此,rat 被提取的机率大大增加,远远超过仅有语义相关的 dog 和仅有语音相关的 mat。在这种模型中,双向联结和反馈使使用频率高的字词更容易得到激活,即它能解释词频效应和音节频率效应。

独立两阶段模型也包括三个部分,最上层的结点表示整个词典概念,其结点表示语义特征。中间是句法(syntactic)层或 lemma 层,表示词的句法及特征,包括语法结构、变化特征、词性等。底层是语音层(phonological 或 lexeme level),包括大量词素结点及音韵、音素结点等。在整个模型中,各结点之间的联结不存在抑制关系。Levelt 等人^[9]假定概念层和 lemma 层为语言产生和语言识别两个系统所共用,但 lexeme 层仅在言语产生系统中存在。概念层和 lemma 层之间的联结是双向的,激活的信息可以相互传递,但 lemma 层和 lexeme 层是独立的,两个从 lemma 层向 lexeme 层的传递只是单向的,不存在反馈,所以被称为独立的两阶段模型。1997年,Roelofs 等人将其总结为 WEAVER (Word-form Encoding by Activation and Verification)模型。

3 存在争论的问题

不同的言语产生理论基本上都同意词汇产生中五个阶段的划分,但在每个阶段的具体细节和不同阶段的实时加工关系上存在着严重分歧。在概念选择和词义激活阶段,一个争论的问题是,词义是如何存储在心理词典中的。词义到底是以局部的词义结点的形式,还是以分解的词义特征的形式存储在心

理词典中。以 Levelt 为首的荷兰学派认为整体概念存储具有优势,能够轻而易举地避开所谓的“上下位概念”问题^[7],他们认为,如果概念以分解的特征方式存储,人在激活“狗”的概念时,“动物”这一概念也应该被激活,因为“狗”的特征是“动物”的一个子集,因此,为了排除“动物”这一概念,大脑中必须存在另外一个抑制机制,而这不符合经济原则。其它研究者(如 Dell^[4])则认为,以语义特征为单位的语义表征方式与有关语言理解中许多词汇加工理论一致,也符合许多来自脑损伤病人的证据。另一个争论的问题是语义在大脑中的表征是否具有范畴性,使得同一范畴(如“动物”)的概念都存在于一个脑区。言语产生和语言理解是否具有共同的语义表征系统,也是争论的问题之一。

在词条选择和音位编码这两个阶段的研究中,存在的争论是词条选择和音位编码是两个完全独立的阶段,还是存在着交互作用。这一争论包括两个方面:一、词条选择和音位编码这两个阶段是否完全独立,即音位编码是否必须在词条选择完全结束后才开始。独立两阶段模型认为,这两个阶段在时间上有明显的分界,语义特征的激活会传输到多个词条上,这些表征之间产生竞争,直至剩下一个最符合语义激活的词条为止,此时,音位编码才能开始。两阶段交互作用理论则认为音位编码在词条选择完成之前就已经开始。检验这两种观点的经验问题就是是否存在同时激活的两个或两个以上的语音表征。二、音位信息是否将激活反馈到词条和词义层次上。交互作用模型认为这种反馈能够很好地解释语误中的混合错误,序列加工模型则否认这种机制的存在。

在从音位激活向语音编码的转换过程中,一个重要然而很少实验研究的问题是,多词素词是如何进行语音编码的。音位信息应按词素为单位激活,每个词素再进行音节化,把音位组合成音节。然而不清楚的是,同一个词的不同词素是同时激活,还是按“从左到右”的顺序依次激活。另一问题是词素内的音位信息激活是否同时进行。为了阻止同时激活的不同音节间的音位相互串位,产生语误,Dell^{[4][10]}提出了所谓的“时间约定”机制(binding by time),认为音位以词素或音节为单位,依次组合。然而实验研究证明,当说话者同时具有多个词的语音形式时,他们并不把这些语音信息融合,从而产生错误^[11]。

在音位信息向语音编码转换的过程中,另一个重要问题是,如何处理一个词的韵律结构,如轻重音、元音和辅音组合方式和规则等。Schattuck -

Hufnagel^[12]提出了一个“框架与插槽”观点,认为一个词的音位信息和它的音节框架是分别提取的。框架表示了一个词的音节数目和重音位置,框架内有每个音节的抽象结构(如 CV 或 CVC),标志着可接受的语音的性质。提取的音位信息就被放在这些结构的插槽中。这种把语音与结构分开在言语产生过程中临时计算音节的方法,能够解释口语中许多语误现象,在实验室研究中也有一定的支持^[13]。但也有证据说明,有些常用的音节并不是临时计算出来,而是从“仓库”中直接提取的。据统计,母语为英语或荷兰语的人,所说的话语中 80% 仅仅由 500 个音节组成,而英语和荷兰语却拥有一万多个音节,这说明人们经常使用的音节数量非常少,这些有限的音节可能存储在一个“仓库”中,音节化过程中激活的信息将传递到“仓库”中对应的具体音节,使其做好发音准备^[11]。但是否存在这个“仓库”是个有争议的问题^[5]。

在语音编码和发音这两个阶段的研究中也存在争论,即发声是否必须在词的所有音节的语音编码完成之后才能开始。根据语音编码“从左到右”、序列进行的观点,如果发声在词的所有音节的语音编码完成之后开始,则多音节词的发音潜伏期要比单音节词长,出现词的长度效应。反之,则词长效应很难表现出来。这方面的实验研究结果相互矛盾,未能给出一致的答案。Eriksen 等人和 Klapp 等人的早期研究表明,在图形命名任务中,单音节词的产生时间快于双音节词。但 Bachoud - Levi 等人^[14]在严格匹配单音节和双音节词的频率,图形物体熟悉性等因素后,用图形命名和符号命名两种方法,发现单音节词与双音节词的图形命名反应时没有差别,即没有发现长度效应。如前面所述,符号命名法虽然能够排除项目间无关因素的影响,扩大挑选刺激材料的范围,但具有无法保证纯粹的言语产生过程的缺陷,Bachoud - Levi 等的实验二、四、五均采用符号命名方法,结果的可信性值得怀疑。庄捷和周晓林^[15]最新的研究比较了汉语使用者对同一幅图片单音节命名(如“龟”)和双音节命名(如“乌龟”)的命名潜伏期,发现了明显的词长效应。

4 参考文献

1 Fromkin V A. The nonanomalous nature of anomalous ut-

- 2 沈家焯. 口误类型. 中国语文,1992;4:306—316
- 3 Lupker S J. The role of phonetic and orthographic similarity in picture-word interference. *Canada Journal of Psychology*, 1982;36:349—367
- 4 Dell G S. A spreading activation theory of retrieval in language production. *Psychological Review*,1986;93:226—234
- 5 Levelt W J M. Models of word production *Trends in Cognitive Sciences*,1999;3:223—232
- 6 Caramazza A. How many levels of processing are there in lexical access? *Cognitive Neuropsychology*,1997;14:177—208
- 7 Roelofs A. The WEAVER model of word-form encoding in speech production. *Cognition*,1997;64:249—284
- 8 Dell G S, Schwartz M F, Martin N, Saffran E M, Gagnon D A. Lexical access in aphasic and non-aphasic speech. *Psychological Review*,1997;104:801—837
- 9 Levelt W J M, Roelofs A, Meyer A S. A theory of lexical access in speech production. *Behavior and Brain Science*,1999;22:1—38
- 10 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
- 11 Meyer A S, Schriefers H. Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*,1991;17:1146—1160
- 12 Shattuck - Hufnagel S. Speech errors as evidence for a serial-ordering mechanism in sentence production. In: *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, ed. W. E. Cooper & E. C. T. Walker. Lawrence Erlbaum,1979
- 13 Sevald C A, Dell G S, Cole J S. Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory and Language*,1995;34:807—820
- 14 Bachoud - Levi A - C, Dupoux E, Chen L, Mehler J. Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*,1998;39:331—346
- 15 庄捷,周晓林. 言语产生中的词长效应. *心理学报*,2001;33(2):28—32

English Abstracts

A STUDY OF 6- to 12-YEAR-OLD CHILDREN'S BRAIN POWER DISTRIBUTION OF ULTRA-SLOW FLUCTUATING GRAPH WITH ITS DEVELOPING TRAITS

Lin Chongde, Wo Jianzhong

(Institute of Developmental Psychology, Beijing Normal University)

In order to explore children's brain power distribution of ultra-slow fluctuating-graph and its developing traits, 228 subjects aged from 6 to 12 were analyzed through an ET-encephalofluctuograph Technology System. The following were some of our results: 1) The superiority of staff power was increased with the age increasing while the most significant decline appeared in the brain domain of F3, C3, P3 O1 F7, T5, which showed a slow trend after 8 years; 2) There were changes with an up-and-down wave (at no regular interval) in that of F4, C4, P4, O2, F8 and T6 in the right brain; 3) The distributions of Pmax—Pmin in each age group were in regular order. Pmax had an absolute superiority in the occipital lobe, especially in O2. Pmin distributes in all the temporal lobe, especially in the fore-temporal lobe. But Pmin seldom appeared in the domain of T6, and the gender difference was found significant ($P < .01$). For example, Pmin of the boys was mainly in the left temporal lobe while that of the girl mainly was in the right one; 4) The ratio of the right power compared with that of the left showed the male children generally had the superiority in the right brain and the female children in the left one.

Key words: brain power, ultraslow fluctuating-graph, child, age.

PROGRESS IN THE RESEARCH OF SPEECH PRODUCTION

Zhou Xiaolin

(Department of Psychology, Beijing University)

Zhuang Jie, Shu Hua

(Department of Psychology, Beijing Normal University)

This paper evaluated theories and experimental methods in the research of speech production. It was pointed out that any general theories of speech production had to consider how semantic information was represented in the lexicon, the interactivity between semantic and phonological activation, and

the way phonetic and phonological information was encoded. The influence of the characteristics of a specific language on speech production was briefly discussed.

Key Words: speech production, interactivity, phonological encoding, phonetic encoding.

TACTILE SPATIAL ACUITY OF CHILDREN AND ADULTS

Qian Xiuying, Xu Xiaomei, Hu Lingyan, Xu Min

(Department of Psychology, Zhejiang University)

The relationship between the tactile sensitivity and the density of afferent fiber was focused by measuring the tactile spatial acuity of 6—7 years old children and college students in using a grating orientation task. The result showed that adults performed better than children with wider groove and children did better than adults with narrower groove.

Key Words: tactile sensitivity, density of the afferent fiber units, grating orientation task.

THE CAPACITY OF IOR IN SUCCESSIVE AND SIMULTANEOUS CUEING PROCESSES

Zhou Jianzhong, Wang Su

(Department of Psychology, Beijing University)

The capacity of inhibition of return (IOR) was studied in two experiments with college students as the subjects. The cue-target paradigm and the detecting task were used. Experiment 1 was conducted by successively cueing of 1-5 peripheral locations. The results showed that the capacity of IOR was 4 when the successively cueing locations were adjacent and the capacity of IOR was 1 when the successively cueing locations were spaced. Experiment 2 was conducted by simultaneously cueing of 1-5 peripheral locations. The results showed that the capacity of IOR was 3 when the simultaneously cueing locations were adjacent and the capacity of IOR was 1 when the simultaneously cueing locations were spaced. Our results supported the idea that there were two kinds of IOR, one being dispersed with large capacity, and the other being concentrated with very small capacity.

Key Words: selective attention, successive