

多词素词的通达表征: 分解还是整体^{*}

王春茂

彭聃龄

(中国科学院心理研究所, 100101) (北京师范大学心理学系, 100875)

摘要 有关通达表征的结构主要有三种观点: 词素、整词和混合的观点。分解存储的通达表征认为, 在通达表征层次中只有不可再分的词素, 而没有彼此独立的词条。也就是说, 词语是以词素分解形式存储在通达表征中的。整词存储的通达表征认为, 在通达表征中存储的都是整词, 每一个词都有其独立的词条。混合的通达表征认为词素和整词都有可能是通达表征中的单元。它具有更大的灵活性。综合考虑已有的文献和进一步的理解思考, 作者提出影响词语通达表征的两个因素: 语义透明度和词频。

关键词: 多词素词, 通达表征, 心理词典

词素通常被定义为形态和意义的最小语言单位。自从 Taft 和 Forster^[1] 将词素这一语言学概念引入心理语言学的研究领域以来, 它日益受到了研究者的注意, 并得到了极其热烈的讨论。现在, 几乎所有的心理语言学家, 无论他持的是整词的观点还是词素分解的观点, 都承认词素或词素结构在人的词汇表征和/ 或加工过程中起着一定的作用。词素的效应是独立于形、音、义的一种比较稳定的现象^[2]。有关词素的争论主要集中于, 词素信息的加工是在哪一个层次、哪一个阶段起作用, 词素(或词素结构) 是否存在于心理词典的表征之中, 或是存在于表征的哪一个层次上。

这些争论基本上都同意这样一个假设, 即, 词汇的加工机制存在着这样两个层次^[3]: 一是通达表征层次(the level of access representation), 它是通达达到词条所必经的通道和目的, 又可分为视觉的和听觉的, 一般认为词汇判断作业发生在这个层次; 一是语言学表征层次(the level of linguistic representation), 其中包含了有关的语义、句法等信息。

有关通达表征层次的结构主要有三种观点: 词素、整词和混合的观点。下面将分别进行讨论。

1 分解存储的通达表征

最早的分解存储模型是 Taft 和 Forster^[1] 提出的词缀剥离的模型。该模型认为, 单词的识别有两个层次: 通达文件(access file) 和主文件(master file), 也就相当于通达表征和语言学表征。主文件中存储了与单词的语言学相关的一些信息; 通达文件中存储了通达一个单词所必须的信息。对于多词素词, 通达文件中存储的是其词干。通达文件的编码是依据形的相似程度和频率的高低, 加工过程是序列的搜索。Taft 和 Forster 认为, 前缀词在识别过程中要经过一个词缀剥离的过程, 单词的识别是经过在通达表征中的词干而实现的。

表征经济性是激励人们支持这种理论的主要动机之一, 尽管它并不是 Taft 和 Forster^[1, 4] 提出词缀剥离模型的出发点。对于语言中极其繁冗的词汇在头脑中的表征, 直觉上应该有一个经济的形式。如果多词素词是以词素分解的形式存储的, 那么一个较小的由词素组成的心理词典就可表征大量的多词素词。但是, 经济性原则并不一定是人脑组织的特点, 它只是人们在科学研究中的一种思维方法, 思考问题时的一个模式: 似乎经济的就是好的。这种思维模式可能严重地使研究偏离

^{*} 本研究得到国家自然科学基金会 39470250, 397711120158 项目和教育部九五规划项目的资助。

正确的方向。人脑的容量极其巨大,冗余的表征也许根本就不是一个问题。

真正激励 Taft 和 Forster 提出词素分解模型的是加工的有效性。

利用词素的通达也许要比利用整词的通达更快。但是,为了实现这种通达加工,必须假设在此之前的词素分解过程以及在此之后的整合过程(将通达表征中的信息整合、映射到语言学表征中)。这两个过程又会影响词汇加工的效率。Bergman, Hudson, 和 Eling^[5]发现前缀词和单词素词的词汇判断时间并无差异。可见,即使存在词素分离的通达表征,也对词汇识别的速度无影响。并且由于假设缀词的存在,对它们的错误的词素分解将会延长加工时间。Schreuder 和 Baayen^[6]通过对真前缀词和假前缀词的统计发现,基于前缀剥离的搜索步骤要比相同前缀词存在一起的搜索步骤多 8 倍。

另外,多词素词,如派生词、合成词,它们的意义不能简单地由其组成的词素的意义推知。汉语中大量的不透明词的存在也表明,要准确地从通达表征映射到相应的语言学表征,对于这些词来说,一定要有独立的通达表征。即使是透明词,其语义也不是两个词素义的简单相加。如“钱包”除了是装钱用的包外,还有其它的一些限制,如,应该比较小巧,很可能是皮革或塑料做的等。合成词中包含了大量的语义功能,如, wineglass 是 glass FOR wine, 马路是马“走的”路。分解的通达表征是如何包容这些信息的,尚无满意的回答。

2 整词存储的通达表征

整词存储的通达表征假设一般认为,在通达表征中存储的都是整词,每一个词都有其独立的词条。单词的识别就是通达到其对应的词条。但这并不等于说,词素在其中不起作用。单词间词素关系可通过某种方式蕴藏在通达表征层次的结构之中。其中一个模型是为塞尔维亚-克罗地亚语的名词建立的卫星模型^[7],该模型中的名词的每一个词尾变化形式都有各自的词条。其中,没有词尾变化的那个单词(另外还有一些由此单词通过词尾变化派生的)具有优势,作为一个核心,和家族其它成员——卫星——相连。

有的模型假设在词条上标记有词素信息,这些词素信息可以通过某种形式组织起来,形成一定的结构。Cole, Beauvillain, 和 Segui^[8]发现后缀词的词汇判断时间受词根的累积频率影响,但前缀词无此效应。同一词素家族(词根相同)的派生后缀词的词汇判断时间受整词的频率影响。他们认为,前、后缀词在心理词典中以整词的形式存储,然而其词素结构受其家族组织的影响。家族成员通过编码词缀词的词素、句法、语音等特质,使其相互联系。词根组成该词素家族的“头”(head)^[9]。词汇是经过“头”通达的,家族中的成员按频率进行搜索。前、后缀词在加工上有所差异:在后缀词的加工过程中,由于自左而右的加工,词根先被通达,从而词根的频率对其识别产生影响;前缀词的词根不先于整词加工,其词根的频率也就没有影响。

整词存储的通达表征的主要问题是,词汇表征和词素连接的约束极少。所有的单词都列在通达表征中。尤为可疑的是,如果假设所有的屈折变化词都列在通达表征中,这将导致认为一种有大量屈折变化的语言(如土耳其语)的说话者有一个巨大的通达表征。Hankamer^[10]的工作更进一步表明,对于许多语言来说,这种假设从计算角度是不可信的。

3 混合的通达表征

考虑到语言的多样性,以及每种语言中词的多样性,用同一种通达表征为刻画所有单词的表征是不太现实的。不同种类的多词素词,如屈折变化词、派生词和合成词,以及派生词中的前缀词和后缀词,它们在通达表征中的形式可能是不同的。有人认为,规则的屈折变化词的通达表征是以词

素为基础的,而其它的词的通达是以整词的形式进行的。Niemi, Laine, 和 Tuominen^[11]的 SAID (Stem Allomorph/ Inflectional Decomposition) 更明确地指出,屈折变化词在输入层和中心词典中的存储方式是分解的,而派生词不是。同样,前缀词和后缀词也可能有不同的表征形式。

语义透明度也有可能影响单词的表征。如果假设在通达表征中全为不可再分的词素,那么不透明的词是如何准确地映射到语义学表征中就是一个难以解决的问题。因此,不透明词的通达表征应是整词形式的。Sandra^[12]发现,透明合成词的两个词素对该合成词的词汇判断时间都有促进作用,而不透明词和假合成词无此现象。听觉加工中也有类似的现象。Marsler Wilson 等^[13]用跨通道的重复启动发现,语义不透明词之间无启动现象,而透明词之间有。

词频也有可能影响单词的表征。高频的多词素词,由于常在一起呈现,很有可能在通达表征层次上建立起自己独立的表征;而对于低频的多词素词,特别是对于没见过的多词素词,它们的通达可能只是词素。一个多词素词要重复多少次才能形成稳定的整词通达表征呢?在此之前也许要先通达其词干和词缀的表征。土耳其语中有成百万的屈折变化词,有些词在个体的一生中也难得碰到几次。对这些词,个体不可能建立稳定的整词通达表征,只能依赖其分解的途径。

Caramazza, Laudanna, 和 Romani^[14]的 AAM (Augment Addressed Morphology) 模型是一个典型的混合模型。它认为在通达表征中既有整词,又有词素;各个元素之间是交互激活的。比如,呈现 WALKED 可激活整词 WALKED、WALK、TALKED 等,也可激活词素 ED 等。谁先达到阈限谁就获胜。因此,对于熟悉的词语,整词先被激活,然后才是词素;对于新词或假词,在通达表征中没有相应的整词表征,激活的只是相应的词素。这就表现出熟悉词的通达是整词形式的,没有前词汇的分解;新词(或低频词)的通达是词素分解的。

AAM 模型中,词素和整词的表征是处于同一层次。也有人认为,词素和整词虽然都存在于通达表征中,但是它们处于不同的层次。例如, Taft^[15]以及 Zhou 和 Marsler Wilson^[16]等的多层激活模型都假设存在一个整词层,也就是说,他们同样承认在通达表征中包含整词,且是对于每一个词都存在自己的独立的表征。与以往的整词通达理论不同之处在于,他们将通达表征层次化了,在通达到整词之前都要经过词素这样一个层次。同一层次元素以及各个层次之间的元素按照一定的方式连接起来,遵循一定的激活扩散规则。

4 进一步的思索

从以往看似矛盾的研究中我们意识到,单词的整词形式以及它的词素分解形式在通达表征中都是存在的。这也是近年来混合模型兴起的原因。作为一种建立在旧的模型(分解以及整词表征模型)的基础之上的新模型,它具有较强的可塑性,能广泛地解释已有的实验现象。

从以上的分析我们可以看到,影响单词表征的因素至少有两个:语义透明度和词频。语义透明的词的意义可以由其词素的意义建构而来,将其表征为分解的词素比较好理解。但是,语义不透明的词的意义并不是其词素义的简单相加,而是有其特定的新的意义,甚至完全和词素的意义无关。如果仅仅表征为分解的词素,是无法完成向义的正确映射。因此,其表征应是整体的。

词频也是影响多词素词表征的一个因素。我们都认识“海马”、“海豹”、“海狮”这些词语,它们甚至可以在我们脑中激活相应的视觉形象,而这些形象决不是在海中的马、海中的豹、海中的狮子,可见必然存在与之对应的整体单元。但当我们遇到“海鹿”这两个词素,并确知这是一个表示一种动物的词语时,我们会从“海”和“鹿”这两个词素去猜测这个新词的意义。如果它是一个不透明的词语,比如说“海鹿”是指计算机的一种芯片,那么我们会在这两者之间建立某种联系。由于我们从未遇到过“海鹿”这个词语,不可能存在表征它的独立的单元,因此这个“新词”在最初必然是以分解

形式存储的。更一般地说就是,词频特别低的词语是以分解形式表征的。而词频高的词语有可能在脑中建立稳定的表征。

如果我们承认在通达表征中同时具有整词和词素两种形式,那么另一个必须说明的问题就是这两者的关系。有的混合模型,如 AAM 模型,认为所有的表征都在同一个层次。熟悉的词在通达表征中有独立的单元,不熟悉的多词素词是以其分解的词素表征的。有的混合模型认为词素和整词分别在两个不同的层次^[15, 16]。我们认为,对于后一种模型,应该加一个限制,就是,并非所有的多词素词在整词层都有独立的单元。但是,究竟通达表征的结构是怎样的,尚需作更深入的研究。

5 参考文献

- 1 Taft M, Forster KI. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 1975; 14: 638- 647
- 2 Feldman L B, Andejejkovic D. Morphological analysis in word recognition. In: Frost R, Katz L eds. *Orthograph, Phonology, Morphology, and Meaning*. North holland: Elsevier Science Publishers B. V, 1992: 361-376
- 3 Sandra D. The morphology of the mental lexicon: Internal word structure viewed from a psycholinguistics perspective. *Language and Cognitive Processes*, 1994; 9(3): 227- 269
- 4 Taft M, Forster KI. Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, 1976; 15: 607- 620
- 5 Bergman MW, Hudson PTW, Eling PATM. How simple complex words can be: Morphological processing and word representation. *The Quarterly Journal of Experimental Psychology*, 1988; 40: 41- 72
- 6 Schreuder R, Baayen RH. Modelling morphological processing. In press
- 7 Lukatela G, Gligorijevic B, Kostic A, Turvey M. Representation of inflected nouns in the internal lexicon. *Memory & Cognition*, 1980; 8: 415- 423
- 8 Cole P, Beauvillain C, Segui J. On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language*, 1989; 28: 1- 13
- 9 Culter A. Lexical complexity and sentence processing. In: Flores d Arcais GB, Jarvella RJ eds. *The processes of language understanding*. London: John Wiley & Sons, 1983
- 10 Hankamer J. Morphological parsing and the lexicon. In: Marslen Wilson W ed. *Lexical representation and process*. Cambridge, MA: MIT Press, 1989
- 11 Niemi J, Laine M, Tuominen J. Cognitive morphology in Finnish: Foundations of a new model. *Language and Cognitive Processes*, 1994; 9(3): 423- 446
- 12 Sandra D. On the representation and processing of compound words: Automatic access to constituent morpheme does not occur. *The Quarterly Journal of Experimental Psychology*, 1990; 42(A): 529- 567
- 13 Marslen Wilson W, Tyler LK, Waksler R, Older L. Morphology and meaning in the English mental Lexicon. *Psychological Review*, 1994; 101: 3- 33
- 14 Caramazza A, Laudanna A, Romani C. Lexical access and inflectional morphology. *Cognition*, 1988; 28: 297- 332
- 15 Taft M. Interactive activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, 1994; 9(3): 271- 294
- 16 Zhou X, Marslen Wilson W. Morphological structure in the Chinese mental lexicon. *Language and Cognitive Processes*, 1995; 10(6): 545- 600

ENGLISH ABSTRACTS

THE INFLUENCE OF MATERIAL SCHEMATA AND COGNITIVE LOADS ON THE LEARNING TRANSFER OF PRIMARY SCHOOL CHILDREN

Mo lei, Zhou Yanchun, Jin Suping

(South China Normal University)

This research probed into the influence of the schemata of learning materials and the cognitive loads in the learning course on the transfer of the learning in the analogical reasoning of pupils. Two experiments were conducted. In Experiment 1, the influence of the different models of learning materials on the transfer of the learning in the analogical reasoning of pupils was studied. Having learned the materials organized in schemata or in nonschemata respectively, the two group of subjects were asked to finish four kinds of items which had varying distance of transfer. The results indicated that the achievement in the condition of schemata significantly excelled that in the condition of nonschemata. In Experiment 2, the influence of the different cognitive loads during the learning on the transfer of the learning in the analogical reasoning of pupils was studied. Having learned the materials in the high load condition or in the low load condition respectively, the two group of subjects were asked to finish the same four kinds of items. The results indicated that the achievement in the high load significantly excelled that in the low load. According to the results of both experiments, it could be concluded that for the learning in the analogical reasoning of pupils, the learning in the conditions of nonschemata and high cognitive loads was advantageous to the far transfer, which supported the active responding theory.

Key words: analogical reasoning, pupils, transfer, schemata, cognitive load.

AN INVESTIGATION ON THE GENERAL CHINESE LANGUAGE COMPREHENSION SKILL

Yang Lixia, Chen Yongming, Cui Yao

(Institute of Psychology, Chinese Academy of Sciences)

Two tests, the Multi-Media Language Comprehension Test (MMLCT) and the Text Reading Comprehension Test (TRCT), were used to study the Chinese

language comprehension skill. MMLCT included two written, two auditory and two picture stories that reflected reading (R), listening (L) and viewing picture story (V) comprehension respectively, and was used to test the general Chinese language comprehension skill. TRCT including four texts of different styles was used to test the Chinese reading comprehension skill. The results showed that: 1) MMLCT was significantly correlated with TRCT. 2) There were two components in the general Chinese language comprehension skill, external and internal language components. 3) Reading and listening were significantly correlated with each other, so the external language comprehension was independent of the information modalities in some degree.

Key words: general language comprehension skill, Multi-Media Language Comprehension Test, external language, internal language.

THE ACCESS REPRESENTATION OF POLYMORPHIC WORDS: DECOMPOSED OR WHOLE?

Wang Chunmao

(Institute of Psychology, Chinese Academy of Sciences, Beijing)

Peng Danling

(Department of Psychology, Beijing Normal University, Beijing)

Generally there were three opinions on the structure of the access representation. The first was decomposed-morphological representation, which assumed that there were only inseparable morpheme units but no isolated word units. The second was whole-word representation, which assumed that the units were only words. The third was a mixed one, which assumed that both morphemic and word units existed in the access representation. It was more flexible. Based on empirical literature and further consideration, the authors stated two factors which might affect the structure of representation: semantic transparency and word frequency.

Key words: polymorphic word, access representation, mental lexicon.

SOME CRITICAL ISSUES IN THE RESEARCH ON THE DEVELOPMENT OF CHILDREN'S EARLY UNDERSTANDING OF THE MIND