

句子加工和语义记忆结构的实验研究¹⁾

陈永明 雷晓军 宋均

中国科学院心理研究所 100012

摘 要

本文报告的实验,以句子验证范式为基础,采用常规试验和信号试验相混合的程序,对汉语句子的加工的有关问题作了考察,实验结果表明,句子理解的时间,依赖于人对句中所含概念之间语义关系的认知度;人对句子的理解是一个部分信息逐渐积累的过程。根据实验材料,作者对语义记忆的结构问题作了讨论。

关键词: 语义网络、记忆结构、句子加工、记忆模型

一、前 言

语义记忆的问题,涉及到人类记忆中存储的语言知识和世界知识的结构及其利用过程。它们在知觉、再认、语言理解及人类其它的复杂认知活动中起着重要作用。因此,语义记忆的结构和过程的问题,不但吸引着心理学家的关注,而且也引起了其他领域科学家的兴趣。所以,它们在认知科学研究中占据着十分重要的地位。

目前,研究人类语义记忆的主要方法,是“句子验证”范式。^[1]这种范式的程序是:给被试呈现一个句子,要求他对句子陈述的内容的真伪,尽可能正确和尽可能快地作出判断;同时,对被试反应的正确率和速度进行测量。研究者据此对语义记忆的结构和过程作出推论,提出了不同的解释语义记忆的模型。^[2]

本实验参照 D. E. Meyer 等人^[3]提出的“速度精确性分解”技术的基本思想,对语义记忆的问题作一些探讨。这种技术在一定程度上扩展了句子验证范式,可为语义记忆的结构和过程问题提供一些有用的信息。

二、实验设计和步骤

1. 刺激

实验中呈现给被试的刺激均是一些句子,属于带有全称量词的范畴命题,其形式是:“一切××都是××”。这些作为刺激的句子分三类不同的主谓范畴关系。一类是“子集句”,其主语范畴是谓语范畴的一个子集(如“一切蚂蚁都是昆虫”),它们均属于真句。第二类称为“超集句”。在超集句中,谓语范畴是主语范畴的一个子集(如“一切野兽都是狮子”),所以,它们均属于假句。第三类是“无关联句”,其主语范畴和谓语范畴之间无直接的类属联系(如“一切桔子都是扁豆”)。显然,无关联句也均属假句。

1) 本文于1991年5月6日收到。

此外,为了使句子中每个主语范畴与其谓语范畴的关系均有“真”或“假”两种可能性,以防止被试在对句子作出判断时出现不恰当的偏向,所以,在实验中还插入相应数量的“平衡句”。平衡句均属真句。

实验用的句子共 512 个(子集句、超集句、无关联句和平衡句各占 128 个)。其中有 256 个真句和 256 个假句。此外,还有一组练习用的句子,它们的格式与实验句相同,但陈述的内容不同。

2. 仪器装置、刺激呈现及反应的方式

实验是在 IBM-PC/AT 机上进行的。刺激呈现和反应的记录,均由计算机加以控制和实现。每个句子呈现前,先在屏幕中央呈现一个“·”号,预示被试这是句子将出现的位置。星号持续时间为 1 秒,接着,以它为中心点,呈现一个句子的格式,即“一切____都是____”。1 秒钟后,在句子格式的横线上,分别填入二个词,构成一个完整的句子。

实验时,被试右手的食指和中指分别轻放在键盘的插入(“对”)键和删除(“错”)键上;视距约在 30—40cm 之间,各人可以自由调整。被试若判断刺激句为“真”,就迅即用右手食指按“对”键;若判断刺激句为“假”,则用右手中指按“错”键。

3. 实验安排

对句子的判断分为常规试验和信号试验两种形式。按前述方式呈现一个句子,要求被试对该句子的真伪尽可能正确和迅速地作出适当的反应,这属于常规试验。信号试验是指在刺激句呈现后的一定时间内,给被试一个提示反应的听觉信号;被试在接收到这个信号时,尽管对所呈现的句子还没有进行完全的加工,他必须根据当时从不完全的加工过程中获得的有关信息,立即作出相应的反应。当然,这种反应就会带有“猜测”的性质。在实验中,对提示反应信号与刺激句呈现之间的时间间隔,作了三种长短不同的安排,构成早、中、晚三种不同的信号试验。早信号与刺激句呈现之间的时间间隔,是由另一程序来加以确定的,它相当于被试扫描一个句子所需的时间再加上 80 毫秒。中信号和晚信号与刺激句之间的时间间隔,分别比早信号延长 100 毫秒和 200 毫秒。

整个实验分为四个试验期。每个句子在一个试验期中只出现一次,但是,在整个实验中,每个句子将分别以常规试验(无信号)和早、中、晚信号试验的形式出现 4 次。因此,在一个试验期中,被试需对 512 个刺激句作出“真”或“假”的判断。对一个被试的实验将延续数天时间。

在每个试验期中,属于常规试验和早、中、晚信号试验的刺激句各有 128 个,它们是混合在一起随机地呈现的。实验时,被试不知道将呈现的句子是属于常规试验还是信号试验;所以,这给被试提出了很困难的要求。实验前,每个被试需进行大量的练习,直至被试对常规试验的反应正确率平均达到 90%,对信号试验达到 90% 次在信号出现以后的 350 毫秒内作出反应然后才可进入正式试验。若被试对常规试验的刺激句作出错误的反应,或对信号试验的刺激句在反应信号出现 350 毫秒以后才作出反应,则立即给予一个“反应错误”或“反应太慢”的反馈信号。

4. 被试

这种实验对被试来说是一种极其困难的作业。有的被试经过数天的练习也未能达到要求,从而放弃参加实验。国外用这种方法来进行词汇决策或句子验证的实验时,被试为 3—6 人。参加本实验的被试为 5 人,大专以上文化程度,年龄在 20—28 岁之间。

三、结 果

表1列出了常规试验条件下被试对三种句型的平均反应时和反应正确率。虽然变异数分析表明,句型对平均反应时的影响未达到统计学上的显著水平($F(2,8) = 3, P > 0.05$),但是,从表1列出的数据中还是可以看出一定的趋势:被试对无关联句的反应较快而且正确率较高,对子集句的反应较慢,对超集句的反应更慢一些。这一趋势与J. Kounios的实验所获得的结果基本上是一致的。^[4]

表1 常规试验下三种句型的平均反应时和反应正确率

句 型	平均反应时(ms)	正确率(%)
子集句	843	91
超集句	876	91
无关联句	829	97

表2 信号试验下三种句型的平均反应时和正确率

句 型	早 信 号		中 信 号		晚 信 号	
	反应时 (ms)	正确率 (%)	反应时 (ms)	正确率 (%)	反应时 (ms)	正确率 (%)
子集句	574	58	612	66	670	73
超集句	563	67	610	76	674	79
无关联句	560	76	617	81	678	85

表2给出了信号试验时三种句型的平均反应时和反应正确率。可以看出,信号试验条件下,各种句型的平均反应时和正确率,远低于常规试验条件下的平均反应时和正确率;并且,三种句型在信号试验中的平均反应时和正确率,随着信号时间间隔的延长而相应地增加。对子集句在三种不同信号时间间隔条件下的反应时进行变异数分析表明,它们存在显著差异($F(2,8) = 25, p < 0.01$)。与此相应,由于信号时间间隔的延长,即容许被试有较多的时间对句子加工的情况下,被试对子集句的反应正确率就随之而提高。对常规试验和三种信号试验的子集句的反应正确率作变异数分析表明,反应正确率的差异达到非常显著的水平($F(3,12) = 14.89, p < 0.01$)。

此外,从表2中可以看到,在信号试验条件下,由于对句子的加工是不完全的,被试偏向于对刺激句作出“假”的反应,从而使子集句的反应正确率相对于其他句型来说明显偏低。这种现象在J. Kounios的实验结果中尤为突出。^[4]当然,在他的实验中,三分之二的刺激句属于假句,因此,在句子加工不充分的情况下,更诱使被试对刺激句作“假”的反应。但是,这一事实也表明,在缺乏充分的关于刺激句的信息的情况下,被试易偏向于作出“假”的反应。

在实验中,我们根据句子的主项与谓项范畴之间的距离,把三种类型的句子各分成二种亚型:子集句1(如“一切狮子都是野兽”)和子集句2(如“一切狮子都是动物”);超集句1(如“一切昆虫都是蝗虫”)和超集句2(如“一切动物都是蝗虫”);无关联句1(如“一

切梅花都是水稻”)和无关联句 2 (如“一切梅花都是鸽子”)。每类句子的每种亚型均为 64 个句子。对这些不同亚型句子的实验结果见表 3。

表 3 常规试验中各亚型句子的平均反应时和正确率

句 型		反应时 (ms)	正确率 (%)	平均反应时 t 考验
I	子集句 1	904	88	t=4.315, p<0.05
	子集句 2	782	94	
II	超集句 1	916	88	t=1.60, p>0.05
	超集句 2	837	93	
III	无关联句 1	873	95	t=2.47, 0.1>p>0.05
	无关联句 2	784	98	

从表 3 第 I 项中可以看到,子集句 1 的平均反应时大于子集句 2 的平均反应时,而正确率却低于子集句 2。对这两种子集句的平均反应时的 t 考验表明,两者差异是显著的。表 3 第 II、III 项说明,超集句 1 的平均反应时大于超集句 2,而正确率则低于超集句 2;无关联句 1 的平均反应时大于无关联句 2,而其正确率则略低于无关联句 2。虽然 t 考验的结果未表明它们之间平均反应时的差异在统计学上是显著的,但其趋势是明显存在的。

四、讨 论

A. M. Collins 和 M. R. Quillian 提出的记忆模型,^[6]是探讨人类记忆结构中迈出的重要一步。根据该模型的基本思想,人类记忆中的知识存储方式可以由一种网状的图结构来表示,其基本的结构成分是一集节点和一集把这些节点联结起来的链。节点代表概念,链则表示概念之间的关系。节点和链相互交织,构成一个错综复杂的网络。该模型的要点之一是采用了层次结构的形式,概念在网络中是按类别和外延的大小一层一层地排列的,所以,它通常被称为“层次网络模型”。该模型的另一要点是:人对句子的理解过程,就是对网络的搜索过程,而这种搜索过程被认为带有“全或无”的性质,^[3,4]即或者搜索到需要的信息,或者未搜索到需要的信息。

根据层次网络模型的假设,验证一个句子所需的时间的长短,取决于搜索过程所花的时间;而搜索的时间则取决于句子主项范畴和谓项范畴在网络中相隔的链的数量。对这一假设,学者们已有一些异议。^[6,7]本实验也为评价该模型提供了一些依据。本实验结果表明,被试对子集句 1 (如:“一切蜜蜂都是昆虫”,在网络中仅相隔一条链)的反应时,大于对子集句 2 (如:“一切蜜蜂都是动物”,在网络中相隔二条链)的反应时。这说明,验证句子的时间,实质上并不是取决于句子包含的两个范畴在层次网络中相隔的链的数量。事实上,被试对子集 1 所包含的两个范畴(如:“大雁”与“飞禽”、“野兔”与“野兽”)之间的语义关系的认知度(即对两个概念范畴之间语义关系的熟悉程度和联结强度;这种熟悉程度和联结强度在客观上则依赖于概念范畴的典型性和同时出现的频率)较低,而对子集句 2 所包含的两个范畴(如:“大雁”与“动物”、“野兔”与“动物”)之间语义关系的认知度相对来说较高,这才可能是被试对子集句 1 反应慢而对子集句 2 反应较快的真正原因。

以往的某些实验表明,被试对假句的验证时间大于对真句的验证时间。根据层次网

络模型的预测也是如此,因为假句所包含的两个概念范畴在网络中相隔的链的数量较多。事实上,这种预测并不具有普遍意义。从表1中可以看到,被试对无关联句的反应时并不大于对子集句的反应时。

此外,从表3中还可看到,同是无关联句,但由于句中两个概念范畴间的语义关系不同,被试的反应时也不同。被试对无关联句1(如“一切荷花都是小麦”)的反应时大于对无关联句2(如“一切荷花都是老虎”)的反应时。对超集句的两种亚型句子的反应时也有类似情况。被试对超集句1(如“一切昆虫都是蜜蜂”)的反应时大于对超集句2(如“一切动物都是蜜蜂”)的反应时。这种结果与我们的预期是一致的。虽然它们在统计学上尚未达到显著水平,但这种趋势是明显的。这些现象,显然是由被试对句中两个概念范畴之间的语义差别的认知度所决定的,是不能用层次网络模型的预测来加以解释的。

根据层次网络模型的假设,人对句子的理解是对网络的搜索过程;这种搜索过程的特点,带有“全或无”的性质。但Kounios的实验首次表明,人对句子的理解是一个连续的、渐进的过程,是一个部分有效信息逐渐积累的过程。^[4]本实验的结果表明,在提示性反应信号出现较早的情况下,被试用以加工的时间较短,从而信息的积累较少,所以,反应的正确率就较低;随着提示性反应信号的延迟,或者在无信号出现的情况下,被试用于加工句子的时间较充分,信息的积累也随之较多,从而表现出被试的判断正确率相应地提高。可见,人对句子的理解,不是一个“全或无”的过程。

在E. E. Smith等人提出的“特征比较模型”中,^[2]概念是用一集松散的语义特征来表示的;概念范畴之间的关系不是预存的,而是通过特征比较计算出来的。因此,从特征比较模型来看,人对句子的加工可以是一个渐进的、信息逐步积累的过程。但是,该模型的基本假设与我们对句子加工的直觉是很难一致起来的。在通常情况下,人在验证一个句子时,决不是通过对两个概念的特征集进行计算以后才作出判断的。只是在某种特殊的、缺乏有关知识的情况下,这种把两个概念的特征集进行比较或计算的过程才或许会发生。此外,根据该模型的预期是:人验证“麻雀是动物”这类句子所需的时间,必然会少于验证“麻雀是鸟”这类句子所需的时间,因为验证前一句所要比较的特征较少,而验证后一类句子所要进行比较的特征较多。然而,生活经验和某些实验材料均说明,实际情况并不是这样的。^[8]

A. M. Collins和E. F. Loftus提出的扩散激活模型,^[9]与经验材料比较适应。第一,扩散激活模型是根据语义类似性线索组织起来的,某些概念范畴由于它们的共同特性而群聚在一起,从而摒弃了呆板的层次结构。这样,句子验证的时间,就不受层次距离的约束。当然,单纯地认为语义类似性决定网络的组织结构,这也是有局限性的。例如,“蝙蝠”与“兽”的语义类似性显然低于“老虎”与“兽”的语义类似性,所以,一般而言,人验证句子“蝙蝠是兽”的时间较长,而验证“老虎是兽”的时间相对较短。但是,若人们在生活中对“蝙蝠是兽”多次反复认知,那么,它的验证时间将会快于“老虎是兽”的验证时间。这是因为,由于多次的反复认知,它们在网络中的距离就会缩短,连接它们的链的畅达性就会提高。所以,概念范畴之间语义关系的认知度决定着网络结构的细节。由此可以推出,语义网络是一种动态的结构,而决非是一种静态的配置。只有这样来设想网络结构,模型才能适应多样化的经验材料。

第二,扩散激活模型以扩散激活的机制代替了简单的搜索过程。当一个概念被加工

时,相应的节点就进入激活状态。并且,激活沿着由该起始节点伸出的链以不同的强度和速度向外扩散,形成一个激活圈;当与来自另一起始节点的激活圈相交并在相交点上的激活总量达到某个标准时,就获得了作出合理反应的充分根据。这种扩散激活是一个渐进的过程,它与心理学实验中获得的部分信息逐渐积累的现象是相适应的。因此,摆脱了层次结构的网络模型及其扩散激活的加工机制,在说明人的语义记忆结构和过程方面是较有力的。

五、结 论

1. 句子验证的快慢,不取决于概念范畴之间的层次距离,实际上是取决于人对句子所包含的两个概念范畴之间语义关系的认知度。

2. 人对句子的加工,是一种部分信息逐步积累的、动态的过程,而不是一个“全或无”的过程。

3. 与层次网络模型和特征比较模型相比,扩散激活模型对心理学实验中获得的材料具有较好的适应性。

4. 人对概念范畴之间语义关系的认知度,影响着网络结构的细节;所以,人的语义记忆的结构也具有动态的性质,而不是一种静态的配置。

参 考 文 献

- [1] Kintsch, W., Semantic memory, A tutorial. In R. Nickerson (Ed.), Attention and Performance, VIII, p. 595—620, 1980.
- [2] Smith, E. E., Theories of semantic memory. In W. K. Estes (Ed.), Handbook of Learning and Cognitive Process, Vol. 6, p. 1—56, 1978.
- [3] Meyer, D. E. and et al, The Dynamics of Cognition and Action: Mental processes inferred from speed-accuracy decomposition. Psychological Review, 1988, Vol. 95, No. 2, p. 183—237.
- [4] Kounios, J. et al, Structure and process in semantic memory: New evidence based on speed-accuracy decomposition. Journal of Experimental Psychology: General, 1987, Vol. 116, No. 1, p. 3—25.
- [5] Collins, A. M. and Quillian, M. R., Retrieval time from semantic memory, Journal of Verbal Learning and Verbal Behavior, 1969, 8, p. 240—247.
- [6] Conrad, C., Cognitive economy in semantic memory. Journal of Experimental Psychology, 1972, 92, p. 149—154.
- [7] 缪小春, 汉语词汇语义记忆的实验研究.《心理科学通讯》, 1988年, 第6期, 1—5页.
- [8] 陈永明、彭瑞祥, 汉语语义记忆提取的初步研究.《心理学报》, 1985年, 第2期.
- [9] Collins, A. M. and Loftus, E. F., A spreading-activation theory of semantic processing. Psychological Review, 1975, Vol. 82, No. 6, p. 407—428.

AN EXPERIMENT ON SENTENCE PROCESSING AND SEMANTIC MEMORY STRUCTURE

Chen Yongming Lei Xiaojun Song Jun

Institute of Psychology, Academia Sinica

Abstract

This paper described an experiment on Chinese sentence comprehension based on the paradigm of sentence verification, taking the hybrid procedure of normal (no signal) tests and signal tests. The experimental results showed: (1) The time needed for sentence comprehension depends on the subject's cognitive degree to semantic relation between concepts in the sentence; (2) Sentence comprehension is a process of partial information accumulation. From the experimental materials, the authors discussed the structure of human semantic memory and evaluated some models concerned.

Key words: semantic network, memory structure, sentence processing, memory model