

# 词汇语义表征编码中的抽象性和像似性关系研究

林立红, 卢 姝, 江婷婷

(宁波大学 外国语学院, 浙江 宁波 315211)

**摘要:** 词汇语义编码中存在抽象性和像似性两种制约关系。主要讨论了与词汇形义关系相关的三个问题: 第一, 抽象词义的来源问题, 认为语言的内在属性和词汇分布规律是抽象词义知识的主要来源; 第二, 像似性和抽象性在词汇形义关系中的功能差异问题, 像似性因其形义理据而有助于词汇学习或交际, 而抽象性则有助于新范畴或形义关系的构建; 第三, 词汇形义关系中一直存在像似性和抽象性两种关系的竞争。任意性特征能赋予语言表述更多的抽象意义, 抽象性也因此词汇形义关系发展中占据上风。

**关键词:** 像似性; 抽象性; 语义表征; 形义关系

**中图分类号:** H03

**文献标识码:** A

**文章编号:** 1001-5124(2019)02-0068-05

## 一、引言

词汇形式和词义之间的关系有两种: 一种是任意性关系, 另一种是像似性关系。任意性关系指词汇形式和词义之间缺乏显性理据; 比如, 汉语中被称为“猫”的动物在英语中则被称为“cat”, 这反映了语言符号与其所指之间的任意性关系。像似性关系指词形(音)与其意义之间具有相似或相同的信息; 比如, bang、click 等拟声词的语音理据。有学者提出,<sup>[1]</sup>形式和语义之间的像似性关系在促进语言学习和加工的同时, 也是抽象词义建构的一个障碍, 即: 词与所指之间的像似性关系若过于紧密, 则不利于抽象概念的形成。形式和语义之间的高像似度不仅使人们难以表述抽象意义, 而且也让读者或听者难以理解抽象意义。语言的抽象性和像似性关系涉及语言的基本设计原则。

## 二、抽象性及形式—语义表征关系的两种观点

词汇的语义编码类型可分为具体词汇和抽象词汇。抽象词汇的语义内容难以直接感知, 其抽象性特征显著; 而具体词汇的语义内容可直接为我们身体的感官系统所感知。

Brysbaert 等人把具体词汇定义为“现实中的事物或行为, 或可以通过身体通道系统直接体验的词”,<sup>[2]</sup>如 Lemons、dog 等为具体名词; fly、cut 等为具体动词。抽象词汇则指那些无法直接为身体感官系统所感知、但其词义能被理解的词。抽象词汇的指称没有直接可见的外在特征, 如 freedom、democracy。对抽象词汇的解释也只有通过使用其他词汇来进行。词汇中的抽象性与具体性具有相对性特征, 词汇在“抽象—具体”连续体内部有不同的位置。

人们对词义表征的构成问题也有两种不同的观点。一类是具身认知(embodied cognition)观照下的具身语义观, 认为感知运动、情感体验和各种通道信息均属于语义的重要内容;<sup>[3]</sup>另一类是非通道语义观, 强调符号和非知觉表征的作用, 不承认知觉、运动和情感知识是语义知识的组成部分。<sup>[4]</sup>具身语义观认为, 人们阅读或听到“lemons”“dog”时, 会利用柠檬的视觉表征和猫叫的听觉表征。<sup>[5]</sup>这类知觉表征构成了词义和概念知识本身。人们对“柠檬”和“猫”的知识主要来自现实世界与“柠檬”“猫”有关的经历; 文字只是触发了我们身体的感知系统对其所指对象进行感知运动模拟。

收稿日期: 2018-10-08

基金项目: 国家社科基金项目“二语违实结构加工中的层级意动研究”(17BYY088)

第一作者简介: 林立红(1968-), 女, 浙江宁波人, 教授, 主要研究方向: 语言学。E-mail: linlihong@nbu.edu.cn

非通道语义观则认为，词义表征内容是通过文字提取的，单词映射的是概念的基本状态；而概念状态具有非通道性质，所以单词的语义表征也是非通道的。显然，非通道观淡化了现实经验对语义表征内容形成过程中的作用。比如，人们读到 the eagle in the sky 和 the eagle in the nest 时，<sup>[5]</sup> 大脑中对 eagle 会有两种不同的具体感知表征。这说明知觉模拟（perceptual simulations）在发挥作用；人们理解语言也是通过感知模拟来实现的。研究显示，<sup>[5]</sup> 包含通道信息的单词可以快速激活与单词所述语义表征相关的通道区域。知觉、运动和情感系统与语言理解有因果关系。非通道观反驳具身语义观的一个证据是：不同感知体验的人为何能够具有相同的概念内容？天生失明者的语义概念结构为何与视力正常者相似，甚至连颜色等视觉概念也非常相似。<sup>[6]</sup> 非通道语义观对具身语义观提出的另一个挑战是：如何充分解释抽象概念的具身语义结构？

### 三、词汇中的抽象性等级与抽象语义表征的来源

Brysbaert 等人曾对英语中的 4 万个词条进行五级量表的分类研究，把那些单词分成最抽象（1=Most abstract）和最具体（5=Most concrete）五个等级。<sup>[2]</sup> 例如，freedom、democracy 和 justice 的具体性赋值分别为 2.34，1.78 和 1.45。名词在所有词类中的具体性等级赋值最高（M=3.53）。不过，也有很多具体性等级比 freedom（2.34）赋值更低的名词，如 idea（1.61）、chance（1.64）等。动词具体性等级赋值比较高的，如 skate（4.6）；具体性等级较低的动词，如 imagine（1.53）。副词的具体性赋值大多位于抽象度较高一端，如 especially、maybe、already、never（<1.60）等。形容词中有赋值等级较高的，如 wooden（4.61）、bald（4.69）；也有较低的，如 normal（1.40）、irrelevant（1.50）。Juhász 和 Yap 也收集了英语本族语者对 5 千个词汇所涉及感知体验程度的相关数据。他们采用七级量表判断；其中量表中的 1 为无感知体验，即抽象词汇；7 为最大感知体验，即涉及身体感知运动和通道信息最多的具体词汇。<sup>[7]</sup> 在其七级量表等级分类中，garlic（6.56）、walnut

（6.50）等为高感官体验单词；choice（1.0）、mere（1.08）等则为低感觉体验词汇。该感知体验量表与 Brysbaert 等人的五级具体性量表之间有适度关联关系（ $r=0.4$ ）。<sup>[2]</sup>

从现有文献我们不难发现，无论是通道语义观或无通道语义观，他们所举例子大多位于“具体-抽象”连续体内部的两个极端。比如，通道语义观所举的例词或句子，其通道特征非常明显，如 kick、pick。非通道观则喜欢列举 freedom 这样的例词，但对 fun、idea 等这种有较多语境关联的词语往往关注不够。如果词义知识不是源自我们对世界的直接体验，那么它从何而来？除了无处不在的高抽象实词之外，语言中还充满了高度抽象的功能词，如 the、it、to 等。对高度抽象词汇来说，其抽象词义恰恰源自语言的内部属性，即抽象概念知识源自语言本身。首先，语言是各种命题的来源，大量语义知识来自语言使用中的各种命题。其次，语言是范畴线索（categorical cue），它为我们提供了具有分类叠加关系世界的一种语言标签。这种语言标签有助于婴儿的早期学习，直至成年期各种不同范畴的形成。语言具有使表征更具范畴化并与范畴中具体例子切分的功能。<sup>[8]</sup> 再次，语言统计本身就是语义知识。词汇分布或搭配结构为我们提供了丰富的知识来源。比如，blue 与 sky 共现的频率高于其他颜色词。具有相似意义的单词产生相似的向量，而且这些向量最终代表了更抽象的关系，如事件之间的时间关系。同一个单词在不同语境中的使用有助于人们创建新概念或新范畴。要建立一个此前根本不存在的新范畴，就需要那些与其所指没有相似关系的单词来填充。也就是说，词形和词义之间具有任意性关系，而任意性特征又为词的抽象性特征提供了空间；这同时意味着，任意性关系是抑制词汇形义像似性的主要力量。

### 四、语言结构设计中的像似性和抽象性

像似性使词形和词义之间的关系更自然、也更符合逻辑，具有像似性的词汇形式本身就暗示了词义。人们对像似性关系明显词汇的理解也更容易。声音和意义之间的像似性还可体现在语音或音位当中。比如，许多语言中表示

小物体和小动物的单词倾向于使用前高元音 /i/,<sup>[9]</sup> 表示 nose 的单词使用鼻音 /m/ 或 /n/。此外, 像似性还有感知生动性 (sensory vividness) 的功能。比如, 用 squealing 这样的像似性单词来描述声音, 会比其他任意性关系词语更生动。这是因为, 像似性与通道感官意象 (sensory imagery) 相互联系。与任意性相比, 有像似性关系的词语会引起大脑感知运动区域的更大激活。像似性因为具有词义线索功能, 所以, 儿童和成人对具有像似性关系的词汇学习会更容易。有研究显示, 像似性可以预测词汇习得年龄: 像似性越多, 儿童对词汇的习得越早。像似性的这种优势在语言进化过程中起着重要作用。<sup>[10]</sup> 而任意性也曾被认为是一种有待纠正的缺陷。结构主义视角下, 任意性是语言理论化进程中的默认状态。任意性和像似性是语言结构设计中的两个原则, 二者相互补充、相互制约。抽象性具有创建新的思维空间的功能。<sup>[11]</sup>

#### (一) 像似性关系在词汇中受抑制的动因

如上文所述, 具有形义像似性关系的词汇在语言学习中能更快、更好地习得和使用。既然如此, 语言为何没有变得更具像似性特征呢? 像似性除了在手语和口语中扮演重要角色之外, 语言似乎越来越远离像似性而向抽象性方向发展。其原因有二: 第一, 单词和意义之间的像似性范围没有声音模仿、手势形状模仿那么广泛。第二, 如果形式与词义相似, 那么相似的词也会有相似的意义, 从而出现形义之间的混乱。换句话说, 如果像似性的交叉式对应关系 (如 /i/ 和 “小” 对应) 被语言使用者有效地使用, 那么像似性就会引起系统扩展。<sup>[1]</sup> 随着词汇量的增长, 这种系统会导致形义之间的混乱。可见, 只有在稀疏的语义社区 (semantic neighborhoods), 词汇形义之间的像似性才不会导致混乱。<sup>[12]</sup> 而任意性所具有的加工优势表现在, 它允许语义相似的语素在语音上保留差异, 如前缀 dis- 和 mis-; 这样就避免了形义之间因为系统的增加而导致的干扰和混淆。然而, 像似性并不是一个 “有” 或 “无” 的问题, 而是存在一个程度问题。例如, 人们将元音与物体大小联系起来, 这种关联能够使

人拥有一个新的系统。在这个系统中, 大型动物或物体的名称包含后元音, 如 max-; 而小型动物或物体的名称则包含前元音。在这样一个系统中, 单词可以包含额外的音位片段 (phonological segments), 该音位片段可以在不提及大小的情况下区分特定的动物。像似性和任意性在同一个词形上的结合, 可大大降低混淆的可能性。

#### (二) 语义表征中像似性与抽象性之间的制约关系

如前文所言, 像似性与抽象性相互抑制。像似性具有选择性, 仅在词义的某个方面通过像似性关系表达。我们对 green 的理解是通过感知模拟某些典型的绿色物体或动物而完成的; 该词也只能帮助我们唤起一个具体 “绿色” 的例子, 或所指对象中与颜色无关的其他特征。具有任意性形义关系的单词 green 通过与一系列的绿色物质相关联, 并从特定的绿色中抽象出其词义表征。虽然像似性的形义关系更容易学习, 但在某种程度上它们近似于特定的原型或局限在很狭窄的语境范围。所以, 最初难以形成较为抽象的词义表征。英语中, 我们可以将动词 eat 的词义扩展至 The acid ate the iron key 中 ate 的抽象词义。ate 的这种词义只有在学习者摆脱 “吃” 是人类用嘴完成饮食事件中的像似性关系之后才会出现。

表声音词汇在英语和其他语言中都是像似性最高的单词类型,<sup>[1]</sup> 这种表声音词汇因为其形义之间的像似性关系而无法实现抽象化。正如 Classen 所言,<sup>[13]55</sup> 若单词的拟声性太强或对其语音的暗示性太强, 它们便不能用于描述其他感官信息。像似性等级较高的形容词也不太可能与描述其他感觉域 (sensory domains) 的名词共现。由此可见, 语言没有朝像似性方向发展的原因有两个: 其一, 像似性关系在方便词汇学习的同时, 也把词与特定的具体词义固化下来, 这就杜绝了该词进一步抽象和泛化的可能性; 其二, 因为成年人谈论的内容大多是抽象的, 所以, 语义域中的像似性也受到了任意性关系的抑制。

#### (三) 抽象性和像似性相互抑制的语言基础

像似性和抽象性之间的反向关系表明, 抽

象词汇更倾向于任意性。而像似性因为在形式和具体/生动意义之间的一个强关联而被限定在特定的语境之中，这同时也限制了该类词汇抽象性关系的扩展空间。抽象性和像似性相互制约的语言基础主要表现在以下三个方面。

首先，高像似性单词的抽象性特征相对较弱。像似性和抽象性之间的反比关系从对具体性、感知生动性 (sensory vividness) 和像似性的测量中体现出来。根据受试对词汇及其词义之间关联程度的判断，像似性评级高的词，其词义表征也更具体、感知生动性更高。比如，“形声词” (ideophones) 是一种专门描述感官意象 (sensory imagery) 的词汇。<sup>[1]</sup> 还有一些词具有感官意义，如 glimmer、glitter、glitz、glisten 都与特定光反射相关。这些单词中的像似性关系表述的信息大多都非常具体。<sup>[14]</sup> 比如，与 /o/、/u/ 等后元音相比，前元音 /i/ 与概念的精确性有关。即使如此，语义表征也与感知特征有关。这种效应似乎源于语音线索所传达的“大小”差异。因此，各种形式的像似性本质上都带有感官性质。说明像似性不仅与知觉表征相关，而且与其他语义表征也密切相关；这限制了词汇的抽象性空间，与高抽象性单词正好相反。

其次，具有像似性关系的抽象单词其词义会更具体。像似性和具体性之间的反向关系表明，具体词语更具有像似性。为了检验具体词语是否更具有像似性，Lupyan & Winter 对 Pexman 等人研究中的数据进行了再次分析。<sup>[10]</sup> 在 Pexman 等人的研究中，<sup>[15]</sup> 受试需要尽快判断所给目标词汇是否表述具体概念、是具体词还是抽象词。如果有像似性关系的抽象词语激活了更具体的语义表征，那么受试在将像似性抽象单词视为具体抽象单词时的错误比率会高于把像似性抽象单词视为任意性抽象词语的比率。回归分析显示，具体性和像似性的交互作用明显。具体词的准确率未受到像似性的影响。抽象词语的准确性与像似性有显著的负相关。像似性值低于 25% 的抽象词被归类为“抽象”，准确率达到 86%。像似性值超过 75% 且在离线评级时被认为抽象单词的判断准确率为 79%。这个结果证明，像似性程度高的词语比像似性程度低的词语更能激发出具体

意义。

最后，像似性词汇与具体语境更关联。如果像似性词语抑制其抽象性，那么它们应该仅出现在相对有限的语境范围之内。根据 Lupyan 和 Winter 的语料库研究发现，<sup>[10]</sup> 语境多样性与像似性有显著关联。像似性高的词语出现在较少的语境中。在控制语料库词频情况下，这种负相关性仍然显著。若同时控制具体性、词频和词类时，也会出现类似的显著负相关关系。当排除封闭类词汇时 (closed-class word)，该相关关系会变得更加显著。另一种对语境多样性的测量不是简单地计算语境，而是使用分布统计来评估这些语境的异构性 (heterogeneity)。在控制词频和词类的情况下，语境多样性与像似性也有负相关关系。但在进一步控制具体性时这种负相关关系消失。可见，像似性限制了单词再使用 (reusability) 的活性，将它们与更少的语境联系在一起。

## 五、结语

本文讨论了英语词汇中的像似性和抽象性关系问题，并从语言基本属性关系讨论了抽象词汇语义知识的来源问题。认为语言的像似性和抽象性特征是两个相互抑制的特征：像似性强化了词汇形式和词义之间的联系，同时也排除了抽象词义派生的可能；而抽象性特征则有利于建构新的范畴或系统。现有研究中，具身观由于过多关注具体概念而忽略了语言中意义的抽象性。抽象词汇在语言使用中出现的频率和范围都很高，日常语言中不仅经常出现像 democracy、freedom 这种表抽象概念的词汇，还有更多的诸如 fun、enjoy 等可具象性较高的抽象词汇。每五个字当中，我们有 95% 的几率遇到像 freedom 这样的抽象词汇。<sup>[10]</sup> 词汇形式—词义的任意性有助于抽象概念和新范畴的形成。而词汇形式和意义之间的像似性关系则妨碍了抽象性的拓展空间。

## 参考文献

- [1] DINGEMANSE M, BLASI D E, LUPYAN G, et al. Arbitrariness, iconicity, and systematicity in language[J]. Trends in Cognitive Sciences, 2015, 19 (10): 603-615.
- [2] BRYSAERT M, WARRINER A B, KUPERMAN V. Concreteness ratings for 40 thousand generally known English word lemmas[J]. Behavior Research Methods, 2014, 46 (3):

- 904-911.
- [ 3 ] PULVERMULLER F. How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics[J]. *Trends in Cognitive Sciences*, 2013, 17 ( 9 ) : 458-470.
- [ 4 ] METEYARD L, CUADRADO S R, BAHRAMI B, et al. Coming of age: a review of embodiment and the neuroscience of semantics[J]. *Cortex*, 2012, 48 ( 7 ) : 788-804.
- [ 5 ] ZWAAN R A, STANFIELD R A, YAXLEY R H. Language comprehenders mentally represent the shapes of objects[J]. *Psychological Science*, 2002, 13 ( 2 ) : 168-171.
- [ 6 ] SHEPARD R N, COOPER L A. Representation of colors in the blind, color-blind, and normally sighted[J]. *Psychological Science*, 2010, 3 ( 2 ) : 97-104.
- [ 7 ] JUHASZ B J, YAP M J. Sensory experience ratings for over 5000 mono- and disyllabic words[J]. *Behavior Research Methods*, 2013, 45 ( 1 ) : 160-168.
- [ 8 ] EDMISTON P, LUPYAN G. What makes words special? Words as unmotivated cues[J]. *Cognition*, 2015, 143: 93-100.
- [ 9 ] ULTAN R. Size-sound symbolism[C]//GREENBERG J H. *Universals of human language*. Stanford: Stanford University Press, 1978: 525-568.
- [ 10 ] LUPYAN G, WINTER B. Language is more abstract than you think, or, why aren't languages more iconic?[J] *Philosophical Transactions of The Royal Society B*, 2018, 373( 1752 ): 1-11.
- [ 11 ] BARTOLOTTI J, BRADLEY K, HERNANDEZ A E, et al. Neural signatures of second language learning and control[J]. *Neuropsychologia*, 2017, 98: 130-138.
- [ 12 ] SIDHU D M, PEXMAN P M. Lonely sensational icons: semantic neighborhood density, sensory experience and iconicity[J]. *Language Cognition & Neuroscience*, 2017, 33: 25-31.
- [ 13 ] CLASSEN C. *Worlds of sense: exploring the senses in history and across cultures*[M]. London: Routledge, 1993.
- [ 14 ] TZENG C Y, NYGAARD L C, NAMY L L. The specificity of sound symbolic correspondences in spoken language[J]. *Cognitive Science*, 2016, 41 ( 2 ) : 113-121.
- [ 15 ] PEXMAN P M, HEARD A, LLOYD E, et al. The calgary semantic decision project: concrete/abstract decision data for 10,000 English words[J]. *Behavior Research Methods*, 2017, 49 ( 2 ) : 1-11.

## The Abstraction and Iconicity in the Semantic Representations of Words

LIN Li-hong, LU Shan, JIANG Ting-ting

( Faculty of Foreign Languages, Ningbo University, Ningbo 315211, China )

**Abstract:** This paper probes into the mutual constrains of abstraction and iconicity in the semantic encoding of words. Three points are discussed. Firstly, the ubiquity abstract meanings with abstract words evolve from the property of language, and prove the best source of abstract knowledge. Secondly, the roles and functions of iconicity and abstraction are discussed: iconicity facilitates language learning and communication owing to the form-meaning resemblance, and the abstraction or arbitrariness of human languages are useful in the establishment of new categories and new form-meaning mappings. Thirdly, Form-meaning arbitrariness allow language to better convey abstract meanings, and iconicity is inimical to abstraction in that iconic forms are so connected to specific contexts and sensory depictions that generalization is hard to derive.

**Keywords:** iconicity; abstraction; semantic representations; form-meaning mapping

( 责任编辑 周 芬 )