

产品语义提取方法及流程研究

李然, 支锦亦, 肖江浩, 秦玉京

(西南交通大学, 成都 610031)

摘要: **目的** 对产品语义的提取方法及流程进行了研究与实践。**方法** 借鉴设计学、语言学中关于语义的提炼方法, 提出了由用户语义采集、文本整理、语义词汇提炼 3 个步骤构成的提取方法及流程。**结果** 按照语义提取方法及流程, 研究小组从调查问卷中获取得到电动汽车的核心语义词汇, 为项目的顺利推进提供了较好的支持。**结论** 基于研究, 构建了一套操作简单, 易于掌握, 适用范围广, 花费成本低, 适用于工业设计师在实践项目中进行产品语义提取的方法与流程。

关键词: 产品语义; 语义采集; 文本整理; 词汇提炼; 语义提取

中图分类号: TB472 **文献标识码:** A **文章编号:** 1001-3563(2018)22-0132-06

DOI: 10.19554/j.cnki.1001-3563.2018.22.022

Product Semantic Extraction Method and Procedure

LI Ran, ZHI Jin-yi, XIAO Jiang-hao, QIN Yu-jing

(Southwest Jiaotong University, Chengdu 610031, China)

ABSTRACT: The work aims to study and practice the extraction method and flow of product semantics. Borrowing the related semantic extraction method from the design science and linguistics, the extraction method and flow composed of such three steps as the users' semantic gathering, the text reorganization and the semantic vocabulary extraction were proposed. According to the semantic extraction method and flow, the research group got the core semantic vocabulary of electric car from the questionnaires, to provide better support for the smooth progressing of the project. Based on the research, a set of method and flow featured by simple operation, easiness to grasp, wide application scope and low cost and suitable for the industry designer to extract product semantics is established.

KEY WORDS: product semantics; semantic gathering; text reorganization; vocabulary extraction; semantic extraction

产品语义^[1]作为产品外观各要素组成的符号结构, 以特定的语言向消费者传达着各种信息^[2], 在设计及其相关研究被广泛应用, 扮演着非常重要的角色。

在设计领域中, 产品语义研究涉及产品造型中情感因素与语义关系研究^[3], 认知过程中的产品语义运行机制研究^[4], 产品语义信息参数化研究^[5], 不同领域知识背景用户的语义差异性研究^[6], 产品语义范畴及分级研究^[7], 计算机辅助设计应用研究^[8], 产品方案评价模型构建^[9]等。总体而言, 产品语义目前的研究重点是, 关注语义在设计过程及其结果中的作用和效果, 通过分析和挖掘产品语义, 获取人的认知感受

与造型特征间的关联性。

语义在产品设计和研究中都占有非常重要的地位, 但是在如何具体提取语义方面, 却没有得到相应的重视。虽然现有设计或研究中所常用的语义提取方法及流程, 架构较为完整, 具有一定的可用性, 但是由于缺乏严谨且科学的理论与方法支持, 获取结果的准确性与可重复性得不到足够的保证, 因此, 本文试图在原有架构基础上, 通过进一步规范和完善, 提出一套操作简单、规范性强, 适于工业设计师掌握的语义提取方法及流程, 以保障语义在产品设计和研究中的可信性和可靠度。

收稿日期: 2018-07-04

基金项目: 理工类科技创新项目 (A0920502051619-129); 四川省哲学社会科学重点研究基地 2018-2019 年度课题 (一般项目) (MD18E004); 西南交通大学 2018 年本科教育教学研究与改革项目 (1804073)

作者简介: 李然 (1983—), 男, 贵州人, 博士, 西南交通大学讲师, 主要研究方向为交通工具设计创新与辅助设计方法。

1 产品语义获取方法与流程研究

按照时序,产品语义提取一般包括语义素材收集、语义处理、语义提炼3个步奏。第1步,收集产品相关的描述、说明或评价信息,是整个语义提取的前提和准备;第2步,对收集信息做适当删简和概括整理,将零散、繁杂的语句转换为便于理解和统计分析的初级语义信息;第3步,根据实际情况及其研究需要,对初级语义信息再次提炼,以获得设计或设计研究所需的语义信息。针对这3个步骤,本文将从用户语义采集、文本整理、语义词汇提炼,提出如下具体操作。

1.1 用户语义采集

口语报告是产品语义收集的重要手段之一,可分为同时性口语报告和追述性口语报告两种^[10]。两种方式各有其特点^[11],可适用于不同目的的设计与研究。

1)同时性口语报告。因为报告其思维活动或描述其思考过程是同时进行的,没有受到长时间记忆的太多影响,能够较为准确地反映出当前信息的加工和储存情况,直接和有效地呈现出当时活动思维的内容,因此,同时性口语报告侧重于过程性分析,适用行为研究,可用性评价,概念草图分析等被访者不必深入思考,不需要早期经验辅助的研究类型。

2)追述性口语报告。给予被访者回顾时间较为充裕,完整性更好,有利于对短时记忆中记忆痕迹或长时记忆中有关信息的提取,可适用于品牌印象、风格趋势等需要结合原有经验对思维内容推论和补充的研究类型,但此种方式受遗忘、干扰等因素影响,会造成信息不准确、细节模糊,易失真的情况。

综上所述,两种报告各有擅长点,并无优劣之分。在实际语义采集时,应当采用哪一种口语报告,将取决于报告方式是否适合于实际设计研究任务的目的与需要。

1.2 文本整理

口语报告内容多属于自然句形式,其中还包括语义不完整的词句,或词句具有多义性、上下文相关性、模糊性等特点,因此,需要将口语报告中的自然语言信息进行文本规范化处理,将短语与句子分割,并转换为适合研究分析的词汇形式。

1)文本规范化。原始报告中包括名词、动词、形容词等实词与虚词部分。很多高频虚词对语义提炼并无意义,需要对其进行过滤操作:去除文本格式标记,去除包括介词、冠词等语义内容很少或者是文本集的每个文本中都可能出现的,且对区分文本内容价值不大的停用词,并通过词干提取或词形还原,将一个词的不同形式统一为一种具有代表性的标准形式(词干或原形)^[12]。值得注意的是,词与词之间没有

明显的切分标志,例如“我期待能够有一辆mini路虎式样的电动车”,应该划分为“mini 路虎”还是“mini”和“路虎”,则需要同时根据上下文、分词目的进行权衡。此外,产品语义词汇不仅包括形容词或形容词性的短语,还应该包括其他可用于描述产品造型的非形容词性词汇。例如,“肌肉”、“老龄”词汇,反映了用户对产品在功能、用途等方面所产生的印象。为使得研究易于操作,便于理解,可将部分非形容词性词汇做形容词化处理,例如将“肌肉”、“老龄”转换为“肌肉感”、“老龄化”,将“不要太大差异”转换为“低差异化”。

2)词汇聚类。文本规则化后得到的语义词汇,不仅数量繁多,彼此间还存在明显的语义关联、语义冗余和语义重叠等现象,不利于进行研究操作,也难以清晰、准确、直观地表达出设计的语义特征。词汇的聚类处理,则可继续缩减词汇数量并突出语义表达重点。需要强调的是,聚类应与用户考虑他们需求的方式一致,而不是与设计师考虑其产品的方式一致^[13],因此,聚类形成的语义词汇类是与用户的观点相对应。

词汇间关联性包括相对、相离、相接、相交、重叠、包含,6种基本关系^[14]见图1。按照6种关联度,可将彼此语义相同、相近或对立的词汇聚类到同一个语义词汇类中。例如,“商务的”与“正式的”、“流线型的”与“流畅的”或“可爱的”、“萌萌的”与“可亲的”,由于彼此间存在包含或相离关系,可聚类到一个语义词汇类里。虽然聚类将大量词汇缩减到少量词汇类,但是在语义表达方面仍显笼统和模糊。

通过文本规范化与词汇聚类,将口语报告转换为设计或研究可用的词汇形式。在这一阶段内,大量初级信息经缩减得到数量更少、价值更高的次级信息,为后一阶段的语义提炼做好准备。

1.3 语义提炼

1)词汇类概括。研究表明,产品语义由核心语义层和解释语义层两个层次组成^[15],见图2。核心语义层是词汇聚类后每类语义范围中心的词(简称“核心语义词汇”),是最为肯定、最简练的描述,具备共识性和完整性的语义,能够从认知的角度形成产品开发设计准则,评判设计提案的标准;解释语义层中的词汇体现了每类语义的范围,并从语义的空间范围上反映了语义的跨度。由此可见,通过对词汇类内语义词汇的概括和提炼,才能获得最具操作性、概括性、一般性,且有良好语义表达能力的核心语义词汇。核心语义词汇优先选取语义词汇类中已有的,且能够概括整个类的语义词汇,如词汇类中有“男人”、“男性化”、“猛男”、“汉子”,优先选取“男性化”一词;如无合适现成词汇,就需基于类中词汇做语义概括,例如,

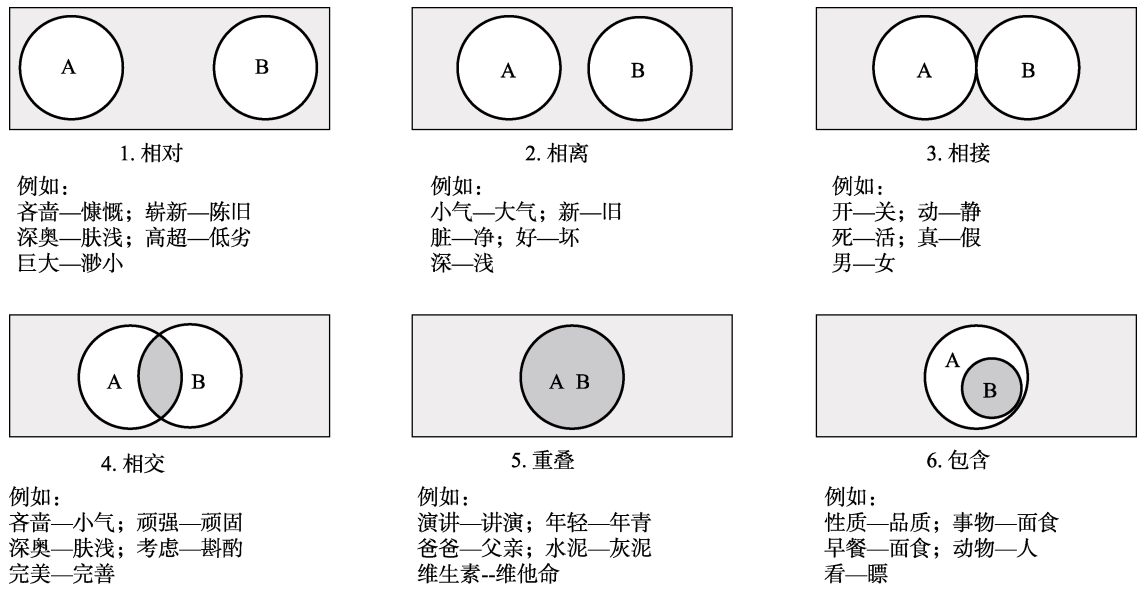


图 1 词汇间的 6 种关联度
Fig.1 Six types of correlations between vocabularies

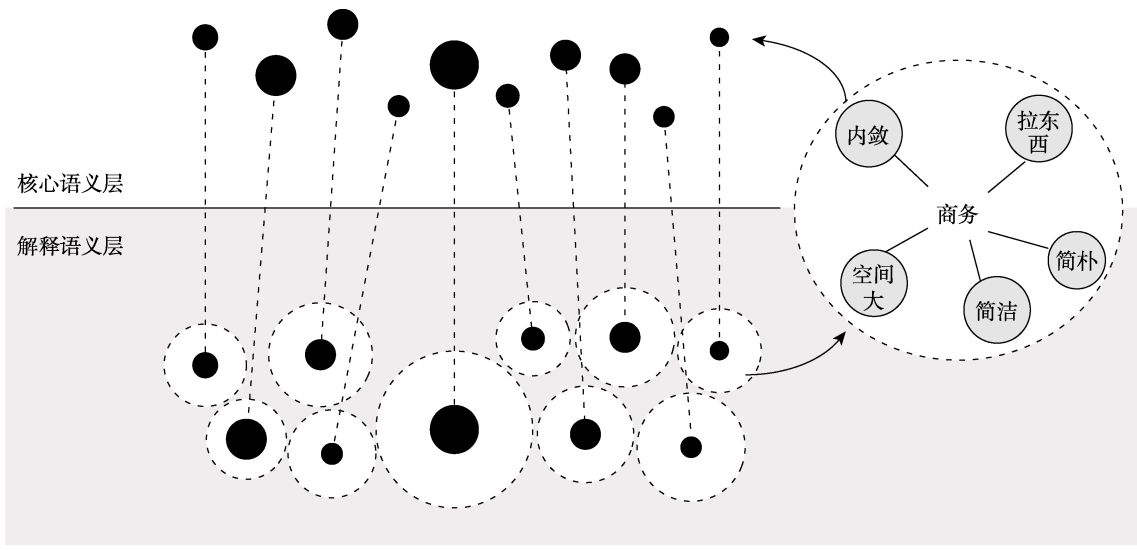


图 2 产品语义词库的层次结构
Fig.2 The hierarchical structure of semantic lexicon of the product

“美女”、“娘娘腔”、“娇气”、“甜蜜”都不适于直接用作类的核心语义词汇，可从中概括出“女性化”一词，以作核心语义词汇。按照词汇关联关系，类中各语义词汇存在相对或相离关系的，核心语义词汇可采用词对形式；而只存在同义词或上下位词关系的，核心语义词汇采用单个语义词汇形式则更为合适。而在实际操作中，不少设计研究在遇到后一种情况时，往往会将单个语义词汇补全为语义词汇对，使得提炼结果在形式上更加齐整。例如，类中有“越野的”“运动化”“动感”，提炼出核心语义词汇“运动感”，可补齐与之对应的语义词汇“休闲感”(也可以是“商务感”)，然而，由于补齐受补齐者主观影响，所以这一做法仍值得商榷。

2) 语义甄选。一般情况下，提炼获得的核心语

义词汇数量还较多，但其中只有少数词汇在产品语义表达中起到关键作用，因此，为满足研究或设计项目的实际需要，核心语义词汇还应再做精简。甄选的意义是找出少数关键核心语义词汇，以更为经济的形式开展研究或设计项目。甄选操作包括：通过统计解释语义层词汇数量，保留高关注度的核心语义词汇；剔除有明显态度倾向性的核心语义词汇，如“美的、丑的”。通常情况下，关键核心语义词汇应保持在 10 个/对以下。此外，根据实际研究情况，核心语义词汇也可做分级处理，其目的是通过判断词汇的语义级别，明确各语义在产品中的重要性 and 地位。

3) 语义润色。语义提取的最终目的是为设计及其研究所用，这意味着概括出的语义词汇不仅应当是准确的，还应当是易懂的、有启迪性的、具有积极设

计价值的。带有贬义，生僻或容易发生歧义的语义词汇，不便于实际中的运用。语义润色的目的是在不改变词汇语义的前提下，对文字做如下调整：第一，将生僻词汇替换为人们较为熟悉的常用词汇，如将“峥嵘”替换为“高大”；第二，将贬义或不雅词汇转换成中性词汇，如用“厚重”替代“笨重”；第三，调整部分词汇的语体色彩，使词汇既通俗易懂、平易朴素，又不失严谨缜密性，如用“曲”代替“弯”。在上一阶段文本整理的基础上，再进行词汇类概括，语义甄选与语义润色处理，能够进一步保证关键核心语义词汇易于理解且具有良好的设计与研究价值，使之可以成为一种重要的设计资源和工具提供给产品设计师或设计研究者，以用于情境构建、方案评价等环节中。

2 实际案例——电动汽车造型的产品语义获取

项目目标是设计一款针对目前中国城市居民市区内出行的电动汽车概念方案。现有电动汽车造型大致可分为两类，一类是基本沿用传统汽车造型，仅在某些细节或局部设计作出调整，例如特斯拉系列，宝马 i 系，众泰知豆等；另一类力图突破传统车型，造型独特新颖，比较典型的有雷诺 Twizy，大众 I.D.，丰田 i-Road。正是由于电动汽车同时具有两类造型的

发展方向，使得人们尚未对电动汽车的造型形成较为一致的认同感，因此，设计师首先需要获取用户对电动车的认知与认同情况，以准确了解用户对电动车的接受度与期望，把握设计方案的大致方向。

2.1 电动车口语报告

根据项目定位，研究小组将调研对象年龄控制在 18~45 岁，具有或潜在具有电动汽车购买能力的城市居民。调研目的是获取电动汽车的语义词汇，为之后的设计情景版制作、方案概念发展和方案评选奠定良好的基础。电动汽车的语义词汇获取，包含着人们对电动汽车的已有印象与未来期待，因此采用追述性口语报告更为适合。鉴于项目经费、时间等因素，项目组短期内收回了 30 份有效问卷，问卷包括对现代电动汽车的评价与看法，购买关注点，以及对未来电动汽车设计期望等内容。

2.2 问卷文本整理

按照文本整理规则，将问卷内容规范化处理，得到 128 个独立词汇。参考词汇间关联性，独立词汇缩减到 22 个词汇类，见图 3。聚类后，与电动车造型设计直接相关的词汇类共计 13 组，间接相关 9 组。间接词汇类涉及蓄电量，充电桩等技术因素，不在本次项目主要关注范围之内，因此将暂时不做余下分析工作。



图 3 语义词汇聚类为词汇类

Fig.3 Semantic vocabularies classified as vocabulary category

2.3 电动车语义提炼

研究小组基于词汇类内解释语义层的词汇，概括提炼出 13 组词汇类的核心语义词汇，见表 1。词汇类 1 包括“小巧的”、“轻巧的”、“玲珑的”、“大气的”和“大方的”等语义词汇。研究小组综合词汇出现频次，表达蕴意，描述对象等因素，从中选取语义词汇对“小巧的—大气的”，作为本词汇类的核心语义词汇。而词汇类 5 则是通过语义概括得到核心语义词汇“共享式—自用的”。在 12 组词汇类的核心语义词

汇中，有 8 类是语义词汇对，4 类为独立语义词汇。

综合考虑本项目的研究目的，包括词汇类解释语义层的词汇数量，可操作性，覆盖范围等因素，研究小组保留 9 类核心语义词汇，去除了“圆润的”，“美观的”与“时尚的—老气的”3 类核心语义词汇。再按照语义润色调整方法，将略有贬义感的“廉价的”转换为“实用的”，“高调的”转换为“张扬的”。最终，获得得到电动汽车的核心语义词汇，见表 2。基于产品语义提取方法及流程获得的核心语义词汇，为项目后期的方案造型设计、提案评判提供了较好的支持。

表 1 初级产品语义词汇
Tab.1 The primary product semantic vocabulary

词汇类	数量	词对/独立词汇	核心语义层	解释语义层
			核心语义词汇	词汇（次数）
1	19	词对	小巧的—大气的	小巧的（4），轻巧的（1），小巧型的（1），小的（1），不占地方的（1），太大的（1），娇小的（1），小体积（1），小型化（1），玲珑的（1）；大气的（4），大方的（2）
2	8	词对	可爱的—霸气的	可爱的（1），萌萌的（1），活泼的（1），很萌的（1）；威武的（1），霸气的（3）
3	8	词对	廉价的—高端的	低劣的（1），过于高端的（1），太低端的（1）；高端（3），上档次（1），尊贵（1）
4	9	词对	低调的—高调的	低调的（1），含蓄的（1），简洁的（1）；豪华的（2），浮夸的（1），炫酷的（1），屌炸天的（1），很屌的（1）
5	10	词对	共享式的—自用的	共享的（1），出租车似的（1），校园小白车似的（1）、商用的（1），正式的（1）；家的感觉（1），家庭备用的（1），两座式的（1），一座或两座式的（1），温馨的（1）
6	11	词对	稳重—轻盈的	结实的（1），重的（1），安全的（1），安全感（1），稳定的（1），结构稳固的（1），质量好的（1），浅色感（1），轻便的（1），灵敏的（1），轻质化（1）
7	14	词对	个性化—常规化	不失个性的（1），有突破性的（3），超凡的（1），有所区别的（1），特别的（1）；中规中矩的（1），死板型（1），非另类的（1），低差异化（1），造型瓶颈化（1），大众化（1），千家万户的（1）
8	11	词对	时尚的—老气的	别致的（2），格调化（1），时尚的（3），优雅的（2）；偏老龄化（1），反老年代步车式的（1），老年代步型（1）
9	4	独立词汇	圆润的	圆润的（4）
10	8	独立词汇	美观的	美观的（3），赏心悦目的（1），漂亮的（3），高颜值的（1）
11	6	独立词汇	流线型	流线型（2），速度感的（1），修长的（1），流畅的（1），低阻力型（1），
12	8	独立词汇	运动感	越野型（1），野马感觉的（1），跨越感（1），运动感（1），力量感（1），肌肉感（2），偏 SUV 式的（1）

表 2 核心产品语义词汇
Tab.2 The core product semantic vocabulary

序号	词对/独立词汇	核心语义词汇
1	词对	小巧的—大气的
2	词对	可爱的—霸气的
3	词对	实用的—高端的
4	词对	低调的—张扬的
5	词对	共享式的—自用的
6	词对	稳重—轻盈的
7	词对	个性化—常规化
8	独立词汇	流线型
9	独立词汇	运动感

3 结语

语义提取易受主观、偏见、措辞不当等因素的影响，导致提取结果缺乏足够的解释性、代表性和一致性，因此，本文针对产品语义的提取，从语义形式及

其获取方法的特点和适用范围，提出了一套操作简单，易于掌握，且经济实用的产品语义提取方法与流程。此方法与流程进一步规范了语义提取的具体操作过程，并提供了具有可操作的方法与工具，使设计师或设计研究者能够较为客观、可靠地从用户评价、关注点、产品期望等信息中提炼出所需的适量关键产品语义词汇，在一定程度上保证了设计研发或研究的正确性与准确性。

参考文献：

[1] PETIOT J F, YANNOU B. Measuring Consumer Perceptions for a Better Comprehension, Specification and Assessment of Product Semantics[J]. International Journal of Industrial Ergonomics, 2004, 33(6): 507—525.

[2] 武艳芳. 产品语义及其生成方法探析[J]. 装饰, 2006 (4): 16—17.

WU Yan-fang. Analysis to Product's Semantic Meaning

- and Its Creating Method[J]. Zhuangshi, 2006(4): 16—17.
- [3] 赵丹华. 产品造型情感类型与情感价值的研究框架[J]. 包装工程, 2016, 37(20): 1—8.
- ZHAO Dan-hua. Research Framework of Emotion Classification and Value of Product Modeling[J]. Packaging Engineering, 2016, 37(20): 1—8.
- [4] 李然, 赵江洪. 基于类型分析的语义驱动汽车造型生成设计[J]. 包装工程, 2015, 36(6): 71—76.
- LI Ran, ZHAO Jiang-hong. Semantic-driven Car Generating Design Based on Type Analysis[J]. Packaging Engineering, 2015, 36(6): 71—76.
- [5] GIANNINI F, MONTANI E, MONTI M, et al. Semantic Evaluation and Deformation of Curves Based on Aesthetic Criteria[J]. Computer-Aided Design and Applications, 2011, 8(3): 449—464.
- [6] ZHAO Dan-hua, LI Ran, ZHAO Jiang-hong. Affective Computing for Product-form Design: an Approach to the Co-evolution of Problem/solution Spaces[C]. The University of Leeds: 6th International Kansei Engineering & Emotion Research Conference(KEER 2016), 2016.
- [7] CHANG H C, LAI H H, CHANG Y M. Expression Modes Used by Consumers in Conveying Desire for Product Form: a Case Study of a Car[J]. International Journal of Industrial Ergonomics, 2006, 36(1): 3—10.
- [8] GIANNINI F, MONTI M, PODEHL G. Styling Properties and Features in Computer Aided Industrial Design[J]. Computer-Aided Design and Applications, 2004, 1(1—4): 321—330.
- [9] 周浩, 焦艺. 基于消费者语义感性信息的新产品方案评价研究[J]. 工业工程与管理, 2014, 19(6): 110—116.
- ZHOU Hao, JIAO Yi. Research on Appraisement New Product Alternatives Based on Semantic Sensory Information[J]. Industrial Engineering and Management, 2014, 19(6): 110—116.
- [10] ERICSSON K A, SIMON H A. Verbal Reports as Data[J]. Psychological Review, 1980, 87(3): 215—251.
- [11] 蹇佳君, 张文鹏. 口语报告法的分类及其在语言认知研究中的应用[J]. 电子科技大学(社科版), 2010(2): 50—53.
- JIAN Jia-jun, ZHANG Wen-peng. Classification of Verbal Report and Its Application to Language Cognition Studies[J]. Journal of University of Electronic Science and Technology of China(Social Sciences Edition), 2010(2): 50—53.
- [12] 吴思竹, 钱庆, 胡铁军, 等. 词形还原方法及实现工具比较分析[J]. 现代图书情报技术, 2012, 28(3): 27—34.
- WU Si-zhu, QIAN Qing, HU Tie-jun, et al. Contrast Analysis of Methods and Tools for Lemmatization[J]. Data Analysis and Knowledge Discovery, 2012, 28(3): 27—34.
- [13] ULRICH K T, EPPINGER S D. 产品设计与开发[M]. 詹涵菁, 译. 北京: 高等教育出版社, 2005.
- ULRICH K T, EPPINGER S D. Product Design and Development[M]. ZHAN Han-jing, Translate. Beijing: Higher Education Press, 2005.
- [14] 石毓智. 同义词和反义词的区别和联系[J]. 汉语学习, 1992 (1): 28—34.
- SHI Yu-zhi. Differences and Links between Synonyms and Antonyms[J]. Chinses Learning, 1992 (1): 28—34.
- [15] 赵丹华, 何人可, 谭浩, 等. 汽车品牌造型风格的语义获取与表达[J]. 包装工程, 2013, 34(10): 27—30.
- ZHAO Dan-hua, HE Ren-ke, TAN Hao, et al. Research on Semantic Acquisition and Expression of Car Brand Styling[J]. Packaging Engineering, 2013, 34(10): 27—30.