

外向型汉语学习词典APP的词语义项设置研究

刘 辰, 徐艳华

鲁东大学人文学院, 山东 烟台

收稿日期: 2026年4月17日; 录用日期: 2026年5月29日; 发布日期: 2026年6月9日

摘 要

义项是词典编纂的关键, 其设置科学与否是衡量一部词典质量优劣的一个重要指标。本文以语料库为研究基础, 分析现有外向型汉语学习词典APP在词语义项设置方面存在的问题。研究表明汉语学习词典APP在义项设置方面存在常用义项缺失、冗余义项添加、语义范围不对等、释义过度限制、义项拆分随意、排序不当等问题, 需借助语料库提升义项设置的客观性、时代性与完备性, 更好地为汉语二语学习者发挥工具性作用。

关键词

外向型汉语词典APP, 词语, 义项设置, 客观性

Research on Lexical Sense Setting of Entries in Chinese Learner-Oriented Dictionary Mobile Applications

Chen Liu, Yanhua Xu

School of Humanities, Ludong University, Yantai Shandong

Received: April 17, 2026; accepted: May 29, 2026; published: June 9, 2026

Abstract

Sense enumeration is a core part of dictionary compilation, and its rationality serves as a crucial indicator for evaluating dictionary quality. Based on corpus data, this paper analyzes the problems in lexical sense design of existing Chinese as a foreign language (CFL) learning dictionary apps. The study shows that these apps suffer from several defects in sense setting, including the absence of frequently used senses, redundant low-frequency senses, mismatched semantic scope, overly restrictive definitions, arbitrary sense division, and inappropriate ordering. It is necessary to employ corpora to improve the objectivity, timeliness and completeness of lexical sense design, so as to better serve as an effective learning tool for second language learners of Chinese.

Keywords

Outward-Oriented Chinese Dictionary APP, Words, Sense Demarcation, Objectivity

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在全球化浪潮推动下,中国与世界各国的交流日益紧密,汉语作为承载中国文化与思想的重要载体,受到了全球范围内学习者的广泛青睐,“汉语热”持续升温。教育部统计数据显示,全球长期学习汉语的人数已突破 16 亿¹,在此背景下,汉语学习工具的需求也随之激增,而 Lew & Szarowska (2017)明确指出,“随着词典编纂数字化革命的持续深化,语言学习者已从传统纸质词典主导的学习模式,大规模转向兼具便捷性与功能性的数字词典”[1]。如今,在线融媒体词典已成为辞书学研究与应用的主流,尤其是适应移动互联网时代学习场景的词典类应用程序 APP,凭借其便携性、交互性和资源整合性等优势,逐渐取代传统纸质词典,成为汉语学习者日常学习中不可或缺的核心工具。

词典的主要作用是描写词义和提供词语的用法,因此建立义项对于词语的解释非常重要,一个词语只有在义项建立之后人们才能真正全面掌握其意义[2]。比如,我们不能说“快”的反义词只有“慢”,虽然“快”在指“速度快;走路、做事等费的时间短”时,跟“慢”相对,如“快步-慢步”,“快车-慢车”。但在“菜刀不快了”等例句中,“快”的反义词是“钝”。义项设置不仅能揭示词义的内涵、反映词义源流经纬的演进面貌[3],更能有效解决跨文化交际中的语义障碍。例如“龙”在汉语中代表着“权威”和“神圣”,衍生出“望子成龙”“龙飞凤舞”等积极表达;而在西方神话传说中,“dragon”则指喷火的邪恶蜥蜴,“my boss is a dragon”不是夸老板,而是说“老板是个凶猛的人”。汉语“龙”和英语“dragon”在所指和文化内涵方面存在很大的差异,如果事先不知道其义项的不同,说汉语的人与说英语的人交流时会产生误解。由此可见,义项设置科学与否直接关系到词典 APP 的质量高低,也关系到学习者实际的理解和使用。

对词义义项的设置,从现有的研究方法看大致有这样几种:1)从历时的角度来分析词义结构,主要研究词义的历史演变,如义素产生的先后,以及相互之间的派生关系,根据词的历史性即词源,词义出现的先后顺序分析词义。例如《辞海》的义项就是根据历史的原则划分安排的。但是由于有些词的意义在历史的长河中失去,多义词的意义并不总是按时间的先后分解,我们很难再按时间顺序编排义项。2)从认知语言学的角度研究义项,认知语言学家关注的是词义之间的联系,认为词义的变化固然是因为历史、社会的变化,但这仅是外部因素,其内因是语言使用者的认知思维发生变化,“词义的转移、扩大,缩小是建立在人的隐喻或转喻的认知基础上的”[4]。认知语言学家认为,最基本的、最先产生的是中心义项,其它义项由于与中心义项具有某种关联(隐喻或转喻关系)而成为同一词的义项,这是根据词义的逻辑顺序设置义项。认知语言学家对义项设置的原则是:考虑如何触发学习者的认知图式,给出简单而有效的释义,使学习者付出最小的努力,获得最佳的效果。认知词典学的主要目标不是追求释义的“圆满”,而是释义的最佳触发效能,优点是对词语的主要义项和次要义项的划分有较强的解释力,其缺陷是这种义项设置方法有时会受编纂者主观因素的影响。《柯林斯 OBL 英语学习词典》(上海外语教育出版社)就

¹此次数据来自中国教育部 2024 年 9 月发布的数据,目前全球已有 85 个国家将中文纳入国民教育体系,国际中文学习者和使用者累计超过 2 亿人。

是按这种原则编排的。鉴于上述情况, 我们建议以语料库为基础来设置词语的义项。

在语料库与义项研究领域, 国外研究聚焦词义消歧与义项标注语料库建设, 如 Agirre & Edmonds (2007) 梳理了自然语言处理领域词义消歧的算法研究[5], Yarowsky (2000)通过机器学习算法提升多义词义项识别准确率, 同时构建了 SemCor、Senseval-2 等义项标注语料库[6], 但此类研究的义项划分多面向机器, 难以适配人类学习者的需求。国内研究则侧重语料库在汉语词典义项设置中的应用, 《现代汉语大词典》编纂团队依托大型语料库与文献资料, 实现了义项设立的全面性与准确性, 兼顾规范性与描写性; 葛晓帅、王梦(2025)提出“基于词典的义项表”概念[7], 验证了 ChatGPT 等大语言模型辅助生成义项表的可行性, 为义项设置提供了新的技术路径, 但目前尚未形成成熟的语料库与义项设置相结合的标准化方法。

2. 现有外向型汉语词典在义项设置方面存在的问题

目前面向汉语学习者的词典 APP 数量众多, 但质量参差不齐, 用户体验与学习效果差异显著。根据开发地区, 本文选取了 Pleco (海外)、trainchinese (中国香港)和 JUZI 汉语(中国大陆)三款代表性软件进行考察, 发现目前外向型汉语学习词典 APP 的义项设置主要存在以下几个问题:

2.1. 多义词的常用义项缺失

例如, JUZI 汉语将“翻盖”解释为: (动)重建; 翻新(房屋); trainchinese 将其解释为: 实质性翻修(建筑), 重建(拆除现场旧建筑后的建筑)。通过检索语料库, 我们检索到这类例句:

- (1) 它是一款内置数码相机功能的翻盖 CDMA 手机。
- (2) 功能最先进的一款手机: 独创 270 度翻盖立体设计。

“翻盖”的本义是“翻修房屋”, 但随着手机的发明, 人们开始用“翻盖”指手机的一种设计款式, 来区分“直板”手机。

再如, trainchinese 将“方便”解释为: (形)方便; 适当的。通过检索语料库, 我们得到以下几个例句:

- (3) 这给数据库管理和程序设计人员带来很大的方便。
- (4) 大家入席以后, 李赤起身去方便。
- (5) 如果手头方便, 就给我四分钱; 手头不方便, 就算我老汉请客。

“方便”除了“方便”“适当的”这两个义项外, 还有(名)某种便利条件; (动)婉辞, 指排泄大小便; (形)婉辞, 跟“手头儿”搭配使用, 表示有富余的钱等义项。

2.2. 增加了不常用的义项

例如 Pleco 和 JUZI 汉语在解释“化”时, 出现“僧侣乞讨(化缘)”这一义项, 通过检索 BCC 语料库, 30,359 个含“化”的结果中仅 10 条句子使用这一义项, 添加不常用的义项反而给学生增加了负担。

2.3. 语义范围不对等

语义范围不对等, 尤其体现在中英文互译上。例如《现代汉语词典》(第 7 版)将“否定”解释为: (动)否认事物的存在或事物的真实性(跟肯定相对); 属性词, 表示否认的、反面的, 而 Pleco 中用“negative” (坏的、有害的; 消极的、负面的; 否定的)对译, 显然不够准确;

2.4. 过度限制

由于对词义缺乏足够的分析, 导致描述词的区别成分时概念过窄, 有些义项仅仅反映了词典编纂者

的语言观, 因而不具有代表性。例如, Pleco 将“提问”限制为“老师向某人提问”, 释义过窄。实际语料表明, “提问”可广泛用于网民、专家、普通学习者等多种主体:

(6) 报告团成员回答网民提问 30 多个。

(7) 评审会分为部门陈述申请预算理由、专家现场提问、财政部门作出情况说明、评审小组闭门评议、合议等阶段。

(8) 只要有需求, 官兵随时可以通过网络向指定教授提问、咨询。

2.5. 同一义项随意拆分为几个义项

同一个词在不同的词典中的义项的确立、分合、排列有差异, 比如 Pleco 中“省”的义项为:

- ① (动)节约, 保存(跟“费”相反)。~劲, ~着点儿用
- ② (动)省略。~一道工序, 这两个字不能~
- ③ (名)缩写。“佛”是“佛陀”之~
- ④ (名)省份。~报, ~长
- ⑤ 省会。~抵省

trainchinese 中“省”的义项为:

- ① (动)节约。为了~钱, 他每天乘巴士上班
- ② (动)省略。这样子可以~去很多麻烦
- ③ (动)缩写
- ④ (名)省份。~长, 全~
- ⑤ (名)省会。

JUZI 汉语中“省”的义项为:

- ① (名)行政区域划分的一个单位, 直接由中央政府管理。我在中国的广东~
- ② (动)<非正式>能少用就少用; 节约。
我们坐公共汽车去那儿吧, 可以~钱
- ③ (动)不要或者减少。
我还是想住到离公司近的地方, 这样可以~去很多麻烦

同一个词“省”, 在三部词典里所设的义项却有所不同, Pleco 和 trainchinese 中“省”有 5 个义项, 都将名词“省”拆分为“省份”和“省会”, 而 JUZI 汉语中只有 3 个义项。像这种问题, 倘若使用语料库检索, 我们可以更精确地设置义项, 从而减少分歧。

2.6. 义项顺序排列不当

对国内 8 所高校国际教育学院的 295 名留学生开展的调查显示: 在词典义项呈现方式的选择上, 118 位受访者选择“按频率”呈现义项, 占比 40.00%; 101 位受访者选择“按难度”呈现义项, 占比 34.24%; 52 位受访者选择“按时间”呈现义项, 占比 17.63%; 24 位受访者表示“不清楚”, 占比 8.13%。但词典在释义时, 经常将不常用的义项最先展示, 例如“扭”有五个义项, 按使用顺序排列为“1) (头、身体)向另外一个方向转; 2) 因为意外或者不正确的动作, 伤了腰、脚和腿连接的地方、手和胳膊连接的地方等; 3) 跳舞、运动等活动身体的时候, 身体左右转动; 4) 《非正式》向一个方向用力拉; 5) 《非正式》用两只手抓住一个人, 往一个方向按住, 让他不能动”

Pleco 中“扭”的义项排列顺序为:

- ① (动)转身, 扭转。他~过身子看了一下
- ② (动)扭曲。把树枝子~断
- ③ (动)扭伤。~了筋; ~了腰
- ④ (动)摆动。走路一~一~
- ⑤ (动)抓住。两人~在一起
- ⑥ (形)不直。七~八歪

trainchinese 中“扭”的义项排列顺序为:

- ① (动)扭伤。我的胳膊好像~了
- ② (动)扭曲。~曲
- ③ (动)转身。~身退出
- ④ (动)摆动。~秧歌
- ⑤ (动)扭打。互相~打
- ⑥ (形)弯曲的, 不直的。歪歪~~

可以看出, Pleco 和 trainchinese 均不是按使用频率进行义项排序, 将核心高频义项后置或穿插在低频义项之间, 既不符合用户查询逻辑, 也不利于学习者优先掌握常见词义, 直接影响词典的使用效率。

3. 外向型汉语学习词典 APP 义项设置优化建议

3.1. 发挥语料库在词典义项设置中的作用

语料库在词典编纂中发挥着重要的作用, 比如说有助于新词目的及时发现; 使真实例句的提供成为可能; 凸显词汇的重要语义和语法信息, 为典型例证的选取和提供创造了条件; 丰富现有词条信息等等。其中, 在词典的义项设置方面发挥的作用尤为突出。具体说来, 大致包括这样几个方面:

1. 语料库增加了义项设置的客观性

由于语料库可以检索到词语所出现的所有文本, 并且词语所出现的语言材料是真实可靠的, 义项的处理不再依靠词典编纂者对词语使用的直觉, 或不真实的例句, 这增加了义项设置的客观性。

例如, Pleco 中光棍的释义为: (名)地痞, 流氓。

(9) 草菇场场长王庆盛逗他: “小张, 你一天 300 多元进账, 今年不仅要娶老婆, 还要盖小洋楼了” “还不多亏你们承诺包销草菇, 不然我可要打一辈子光棍”。

(10) 刚摘掉国家级贫困县帽子的重庆市忠县, 曾有个远近闻名的“光棍村”。

(11) 家家户户盖新房, 老光棍个个娶娇娘。

(12) 600 人的小村, 25 岁到 45 岁的老光棍就有 108 个。

从检索的语料看, “光棍”不只“地痞, 流氓”这一个义项, 还有“未结婚的单身汉”的含义, Pleco 中应该补充这个义项。

再如, Pleco 中裂纹的释义为: (名) ① 裂纹, 分割标记; ② 裂纹(陶器、瓷器等)

(13) 满是老茧、满是裂纹的双手。

(14) 松树皮十分坚硬, 而且有粗阔的裂纹。

从检索的语料看, 词典上所列出的两个义项, 并不是“裂纹”最常用的含义, 我们还可以从“裂纹”出现的语料中概括另一个常见义项——“裂缝”。

再如, Pleco 中打包的释义为: (动) ① 捆包, 打包; ② 打开, 从(箱、包等中)取出: ~检查; ③ 在餐馆吃饭后把剩菜装进袋子里: 这个菜请~

(15) 常用中文设计字体打包下载。

(16) 经典 mp3 打包下载。

(17) 将应用程序和所有数据全部打包到一个 exe 软件大小。

从检索的语料看, “打包”在客观现实中还有一个词典未概括的义项, 即用来指“把电脑中的文件所占的空间压缩, 以便于传输”。

2. 语料库增加了义项设置的时代性

随着社会的进步和发展, 有些词语又出现了新的义项, 这些义项在词典编纂中都应该有所体现。例如:

包装的释义为: ① (动)包裹商品; ② 包装商品用的纸、盒子等。

(18) 包装明星策划书。

(19) 如果你是一名歌手, 因身边琐碎的事情而忙, 没时间包装自己。

“包装”本义是“包裹商品”, 通过语料检索发现, “包装”的语义外延由商品扩展到明星、艺术等方面, 引申出另一种含义, 即指“对演艺界明星的培养, 使其快速成名”, 词典应及时概括出这个义项。

冷场: (名)戏剧、曲艺等演员演出时因演员迟到或忘记台词造成的局面。出现~。

(20) 从这个角度讲, 画摊儿的冷场, 又令人欣慰。

(21) 动漫玩具热卖, 人才招聘冷场。

(22) 和女朋友出去, 总不知道聊些什么好, 老是冷场, 谁可以帮帮我。

(23) 在和客户电话沟通时, 怎么才能很好地和客户交流? 不会出现冷场。

“冷场”本义来源于演艺界, 但现在又出现新的义项。在例句(20)(21)中有“冷清”的意思, 在例句(22)(23)中有“在聊天中没有话题, 双方都沉默不语的局面”, 词典应及时概括出这个义项。

这一类的词语很多, 比如“龙头”要增加一个“某一方面、某一行业的重点”的义项; “程序”则应增加“计算机专业术语, 是用计算机语言编写的命令”这个义项; “土豪”一词, 本有义项为“旧时农村中有钱有势的地主或恶霸”, 新义项有两个: “泛指有钱人”与“让人羡慕嫉妒恨的有钱阶级”。通过语料检索, 我们发现, 由于文化的交流, “便当”成为日语“弁当”(音: ben tou)的译音词, 即“盒装餐食”的意思, 这是新出现的词义, 应该在词典中体现; “前台”除了现有词典所列的三个义项之外, 还应该增加“酒店等服务行业或公司的接待处”这个义项; “模特”不再专指艺术创作的原型或对象, 它的使用范围扩大了。比如展示自身美的, 如“走秀模特”“耳朵模特”“发型模特”和“手部模特”等; 展示消费品的, 如“时装模特”“汽车模特”“手机模特”等。“模特”的语义越来越丰富, 字典应及时更新陈旧的义项, 添加新出现的、当代人们广泛使用的词语含义。

3. 语料库增加了义项设置的完备性

通过语料库的检索大量真实的文本, 我们可以知道一个词的所有含义, 及其用法, 进而概括出所有的义项, 而不会有遗漏。

局限: (动)限制在狭小的范围里。

(24) 阻碍足球训练水平提高的关键自然是教练员水平的局限。

(25) 这套经济模式日益显出它的局限。

(26) 认识自己的局限, 无论何人都有局限性。

通过检索语料库, 我们不难看出“局限”还可以作名词, 表示“缺陷, 不足”。

家居: (动)没有就业, 在家里闲逛。

(27) 业内人士认为, 家居风格是主人修养、品位、职业、习惯的综合表现。

(28) 北京大钟寺家居广场。

从这两个例句可以看出, “家居”还可以指“居住条件和设施”。

非礼: (形)不合礼节; 不礼貌。

(29) 我虽有一片痴心, 却不敢妄行非礼。

(30) 换句话说, 人们可能以礼的名义实施非礼, 有可能以义的名义实施非义。

从这些例句中我们不难发现, “非礼”也可用作动词, 表示“对某人进行冒犯”。

小班: (名)幼儿园里由三到五周岁儿童编成的班级。

(31) 学校现有骨干教师 50 名, 学校实行小班授课, 课堂教学效果好。

从这个句子中可以看出, “小班”还有另一种含义, 即指“一种教学组织形式, 班级的人数在 15~20 人之间”。

红杏出墙: 形容春意盎然, 春光明媚。

(32) 如果单从这个笑容来看, 修平绝不会怀疑妻子红杏出墙。

(33) 他对夫人红杏出墙采取了特别冷静(应该说是非常温和)的态度。

“红杏出墙”出自诗《游园不值》(叶绍翁, 南宋诗人)里面的诗句“春色满园关不住, 一枝红杏出墙来”, 本义是形容“春意盎然, 春光明媚”。但现在越来越多地被用来比喻“有男朋友或丈夫的女性与第三者另有恋情”, 因此应将此义项编入词典。

高压: (形)残酷迫害, 极度压制。

(34) 露天高压管道, 架设高压电线。

(35) 有一种洁碧洗牙器形成每分钟 1200 次的脉冲高压水柱。

(36) 固态氢、高温超导体和高压下合成新物质。

(37) “肝硬化门静脉高压症”为世界肝病三大难题之一。

从语料库中检索的例句来看, “高压”还可用来指“高的气压”或者“高的血压”的简称。

3.2. 从用户视角出发编纂义项

外向型汉语学习词典 APP 的核心用户是汉语学习者, 其义项设置的核心目标是帮助学习者快速理解词义、准确运用词语, 因此必须摒弃“一刀切”的通用模式, 贯彻“以用户为中心”的编纂理念, 提升用户友好度。

1. 义项设置需贴合用户汉语水平

汉语学习者存在明显的水平差异, 初等学习者侧重掌握基础常用义项, 中等学习者需要补充延伸义

项和搭配用法, 高等学习者则需要了解特殊义项、语用差异及文化内涵。当前多数汉语词典 APP 未对学习水平进行区分, 将所有义项混排呈现, 导致初等学习者被复杂义项困扰, 高等学习者又无法获取足够的深度信息。因此, 义项编纂应采用“分层呈现”模式, 先明确标注义项的难度等级(如初等、中等、高等), 优先展示对应水平的核心义项, 再通过超链接呈现延伸义项。例如, 对于“省”的义项, 初等学习者仅需掌握“省份”“节约”两个核心义项, 中级学习者补充“省略”“缩写”, 高级学习者再了解“省会”的特定用法, 避免不同水平用户的学习负担失衡。同时, 针对初等学习者, 义项释义需简洁通俗, 避免使用复杂的汉语术语, 必要时搭配简单的英文释义辅助理解, 而高等学习者则可补充义项的语用色彩、语法功能等拓展内容。

2. 义项设置需契合用户实际使用场景

通过对 295 名留学生的调查可知, “写作业时”“阅读时”和“上课时”是受访者使用汉语学习词典 APP 的高频场景, 分别被选择了 183 次、138 次、118 次; 其次是“日常交流”“影视娱乐”等生活场景, 分别被选择了 94 次、77 次。因此义项编纂需紧密结合这些场景, 优先呈现场景化义项, 补充高频搭配和实用例句。前文提到的“翻盖”义项缺失问题, 本质上是词典编纂者未考虑到学习者的实际使用需求——初等学习者会在购物和日常交流中接触到“翻盖手机”, 却无法从词典中找到对应义项, 影响交际效果。因此, 在义项设置时, 需通过语料库筛选出不同场景下的高频义项, 优先纳入词典。例如, “方便”的义项中, “婉辞, 指排泄大小便”“与‘手头儿’搭配, 表示有富余的钱”均是日常口语中的高频用法, 应优先呈现, 并搭配贴合生活场景的例句, 如“我去方便一下, 马上回来”“他最近手头不方便, 暂时没法借钱”, 让学习者不仅能理解词义, 还能直接运用到实际交流中。此外, 针对不同母语背景的学习者, 还可结合其文化差异补充义项说明, 例如“红杏出墙”的比喻义, 对于不了解中国传统文化的学习者, 需简要说明其文化渊源, 避免误解词义。

3. 义项排列需遵循用户查询习惯

结合前文对 295 名留学生的调查, 超过 70% 的学习者希望义项按“使用频率”或“难度”排列, 这与学习者的认知规律高度契合——学习者在查询词语时, 首要需求是掌握最常用、最基础的词义, 再逐步了解延伸义项。当前多数词典 APP 的义项排列缺乏逻辑性, 将低频义项前置、高频义项后置, 违背了用户的查询习惯, 降低了词典的使用效率。因此, 义项排列应确立“高频优先、由易到难”的原则, 通过语料库统计词语各义项的使用频率, 将高频核心义项放在最前面, 再按难度梯度排列延伸义项、特殊义项。例如, “扭”的义项应优先排列“(头、身体)向另外一个方向转”“扭伤”等高频义项, 再排列“扭曲”“扭打”等低频义项, 同时标注各义项的使用场景(如口语、书面语), 帮助学习者快速定位所需义项。此外, 还可结合用户的查询数据, 动态调整义项排列顺序, 例如某一义项被查询的频率较高, 可适当前置, 让义项排列更贴合用户的实际使用需求。

4. 义项设置需简化冗余信息

汉语学习者的汉语认知能力有限, 过多的冗余义项、复杂的释义表述, 会增加其学习难度, 甚至引发抵触情绪。前文提到的 Pleco 和 JUZI 汉语在“化”的义项中添加“僧侣乞讨(化缘)”这一低频义项, 就是典型的冗余信息, 不仅对多数学习者无实际帮助, 还会干扰其对核心义项的掌握。因此, 义项编纂需坚持“实用优先、精简冗余”的原则, 通过语料库筛选, 剔除使用频率较低、对学习帮助不大的义项, 聚焦核心义项和实用义项。同时, 释义表述需简洁明了, 避免冗长复杂, 尽量使用短句, 减少专业术语的使用, 对于复杂义项, 可通过拆分、举例等方式辅助理解。例如, “否定”的释义, 可简化为“(动) 否认事物的真实性(与‘肯定’相对): (属性词) 否认的、反面的”, 并搭配简单例句“他否定了这个观点”, 让学习者快速理解词义, 避免被复杂表述困扰。此外, 还可在义项中补充简单的搭配提示(如“提问 + 对象”“节约 + 钱财/时间”), 帮助学习者快速掌握词语的使用方法, 提升学习效率。

4. 结语

基于语料库的研究方法为义项设置注入了新鲜的元素, 语料库语言学更加重视词频, 注重单词的搭配关系和词语的用法, 更加注意语言的变异, 关注词汇在语法中的作用, 更加重视例子的真实性, 语料库辅助义项设置不失为解决当前义项所存在问题的办法。

同时, 外向型汉语学习词典 APP 的核心价值在于服务用户, 词典的使用价值, 主要由它对读者的有用程度, 即适当地满足读者对象的求知欲的程度来决定的[8]。义项设置是否科学合理, 直接决定了词典能否满足汉语学习者的需求与期望, 唯有立足用户视角, 才能达到读者的需求和期望, 进而使外向型汉语学习词典 APP 具有较高的使用价值。

另外, 需要指出的是“语料库语言学仅是一种方法和手段”[9], 语料库即使再大, 也不可能完全反映语言面貌。并且, “语料库只能提供语言事实的例证, 但不能对之进行解释, 不能进行推理, 也不能为文本数据直接提供文化和社会背景方面的信息。它在辅助人们的语言‘直觉’和‘内省’的同时, 也离不开研究者本人的语言‘直觉’和‘内省’”。

基金项目

本文得到 2023 年度国家语委科研项目“融媒视域下汉语学习词典 APP 现状调查及优化研究”(项目编号: ZDI145-91)资助。

参考文献

- [1] Lew, R. and Szarowska, A. (2017) Evaluating Online Bilingual Dictionaries: The Case of Popular Free English-Polish Dictionaries. *ReCALL*, **29**, 138-159. <https://doi.org/10.1017/s0958344016000252>
- [2] 李尔钢. 现代辞典学导论[M]. 北京: 汉语大词典出版社, 2002: 87.
- [3] 邹酆. 辞书学论稿[M]. 武汉: 湖北辞书出版社, 2004: 232.
- [4] 章宜华. 语义学与词典释义[M]. 上海: 上海辞书出版社, 2002: 43.
- [5] Agirre, E. and Edmonds, P. (2007) *Word Sense Disambiguation: Algorithms and Applications*. Springer.
- [6] Yarowsky, D. (2000) Hierarchical Decision Lists for Word Sense Disambiguation. *Computers and the Humanities*, **34**, 179-186. <https://doi.org/10.1023/a:1002674829964>
- [7] 葛晓帅, 王梦. 大语言模型辅助下基于词典的义项表生成及应用研究[J]. *语料库语言学*, 2024, 11(1): 26-37, 147.
- [8] 李开. 现代词典学教程[M]. 南京: 南京大学出版社, 1990: 242.
- [9] Susan Hunston. 应用语言学中的语料库[M]. 上海: 世界图书出版社, 2006: 15.