

基于内容分析法的在线英语学习词典用户评价模型构建

陈贤德 杨玉玲

(北京语言大学 教师教育学院, 北京 100083)

摘 要:词典评价研究对词典编纂具有重要意义,但以往研究鲜少从用户视角来建立词典评价模型。以 App Store 上用户对五部在线英语学习词典的评价为样本,运用 ROST Content Mining 6 软件从在线词典用户积极评价与消极评价文本中提炼高频词并进行社会语义网络分析,挖掘出用户的潜在需求和使用偏好。结合高频词表和社会语义网络图选取 100 个有效分析单元,对分析单元进行编码并进行分类检验后,初步建立起用户评价模型。结合词典用户需求研究,完善评价模型。最后,以有道为调查对象,将模型运用于评价实践中。验证结果显示,模型具有较好的信效度。

关键词:内容分析法;在线英语学习词典;使用偏好;用户评价;评价模型

中图分类号:H06 **文献标志码:**A **文章编号:**1674-6414(2024)01-0156-12

0 引言

词典是语言学习的重要资源,在众多词典类型中,在线双语词典,已成为中国英语学习者的主要选择(Dilin et al., 2021)。但高使用率并不意味着高质量,黄锦鸿(2021)便指出五大在线英语学习词典存在收词差异性较大、系统性不足、及时性不够等缺陷。为指引用户购买和使用词典,引导编者提升词典编纂质量,词典评价研究就显得尤为重要。无论是词典交际观还是认知观,用户需要都应居于重要的地位,表现出学习者倾向(胡文飞, 2015),词典评价中更应如此。但目前词典评价大都局限于介绍和描述,评价性特征略显不足,基于用户视角的评论也相对缺乏(王娇等, 2016)。由于网络空间具有虚拟性、宽松性、平等性等特点,用户能够比较真实地将其对产品的满意之处及失望之处表达出来,可成为研究用户偏好的重要资料。基于此,本研究拟通过 App store 搜集五部在线英语学习词典

收稿日期:2023-10-12

基金项目:国家社会科学基金项目“外向型汉语学习融媒词典的研发与创新研究”(22BYY159)、北京语言大学研究生创新基金项目“基于眼动实验的汉语二语学习者词典图文释义阅读策略研究”(23YCX185)的研究成果

作者简介:陈贤德,男,北京语言大学语言学及应用语言学博士研究生,主要从事词典学研究。

杨玉玲,女,北京语言大学教师教育学院教授,博士生导师,主要从事汉语语法、词汇教学、外向型汉语学习词典研究。

引用格式:陈贤德,杨玉玲. 基于内容分析法的在线英语学习词典用户评价模型构建[J]. 外国语文, 2024(1): 156-167.

用户评论,借助内容分析法,挖掘出用户的潜在需求和使用偏好,并在此基础上构建出用户视角下的在线英语学习词典评价模型。

1 相关研究

词典评价体系包含词典评价理论与实践两个主要组成部分。目前,词典评价理论研究多集中于对评价标准和原则的讨论(罗思明等,2006)。在总体评价标准上,陈楚祥(1994)较早地提出了“收词是否全面”“立目是否合理”“义项是否分明”等十条标准。邹艷(2004)则从编纂角度、使用角度、总体角度和类别角度探讨了辞书的评价标准。有关双语词典的评价标准,姜治文等(1998)提出了三方面构想:从词典编纂过程的角度对词典进行描写性评论;从编纂质量、校对质量、设计质量等角度对双语词典进行具体评论;从历时分析、共时比较角度对词典进行定性评论。章宜华等(2000)则基于双语词典评奖委员们的评奖经验,从“出版质量”“编纂质量”两方面对双语词典的质量进行考察。魏向清(2001)从人类评价实践的“合目的性”与“合规律性”的两大哲学基本原则出发,对双语词典批评的复合理据进行重点分析,提出了“编纂目的明确,有自主性”“编纂理据充分,有针对性”“编纂质量优良,有学术性”等用于评价双语词典内容质量的六大标准。

虽然国内关于词典标准的讨论较多,但真正将标准用于评价实践的研究却鲜少。国外出现了部分有借鉴意义的评价实践研究。伊妮德·皮尔森(Enid Pearsons)和温德琳·尼科尔斯(Wendalyn Nichols)(2013)提出包含“界面友好度”“内容质量”“信息来源广度”等八个评价维度的单语英语词典评价框架,山田茂(Shigeru Yamada)(2013)则使用该框架评价了 Merriam-Webster、Macmillan 和 Oxford 等在线英语词典。罗伯特·卢(Robert Lew)和阿格涅什卡·沙罗夫斯卡(Agnieszka Szarowska)(2017)在前人研究的基础上,制定了包含“词汇覆盖面”“词条处理”“词汇信息访问”“词汇信息的呈现”等评价维度的在线双语词典评价框架,并以该框架评价了 MEGAslownik、Diki、bab. la 等六部英语-波兰语双语词典。刘迪麟(Dilin Liu)等(2021)则借用该评价框架评价了 Youdao、Ciba (Power Word)、Bing 等六部汉英在线双语词典并证明该框架具有较强的可操作性。

虽然已有部分研究对词典评价体系的建立进行过探讨,但这些研究更多的是词典学家、语言学家和出版商之间的对话,极大地忽略了用户在评价过程中的作用。因此,纳入用户的评价标准对完善以往的词典评价体系或框架具有重要意义。为弥补前期研究之不足,本文拟基于内容分析法构建起用户视角下的在线英语学习词典的评价框架。

2 研究设计

2.1 研究问题

本研究拟解决以下三个问题:(1)如何对在线英语学习词典用户评论数据进行挖掘,

从而发现用户对词典的使用偏好？(2) 如何根据用户使用偏好构建在线英语学习词典评价模型？(3) 所构建的评价模型是否具有良好的信效度？

2.2 研究过程

内容分析法是一种基于信息特征研究的系统、客观的定量研究方法,其实质是对传播内容中所包含的信息先进行定量的编码分析,然后根据编码结果得出更深层次的定性分析 (Neuendorf, 2016)。本文的研究过程为:收集五部在线英语学习词典用户的积极评论和消极评论,利用 ROST CM6 软件分别对积极评价文本和消极评价文本进行句子拆解和词频计算,再利用社会语义网络分析发现高频词之间的语义连接。结合高频词表和社会语义网络分析图提取出分析单元。对分析单元进行编码设计,划分出一级编码和二级编码。编码完成后,将分析单元进行归类,利用 Holsti 信度值对归类结果进行信度计算,以保证其有效性。结合词典用户需求研究,构建起用户评价模型,最后对模型的可信度进行验证,如图 1 所示。

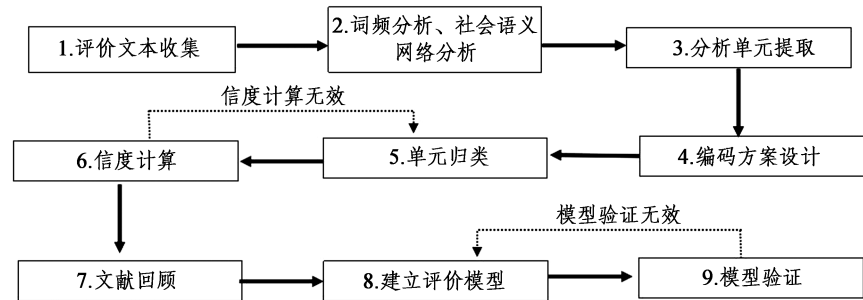


图 1 研究过程与思路

2.3 用户评价文本收集

在 APP Store 中搜索“词典”,选择五部近一年下载量较大、评价数量较多的在线英语词典作为研究对象,分别为:有道词典、金山词霸、欧路英语词典、牛津高阶英汉双解词典和朗文当代高级词典,词典信息如表 1 所示。根据 APP Store 评价类型中“最高评价”和“最低评价”筛选出积极评价和消极评价,在筛选过程剔除“恶意评判”“凑字数”“无效重复”等无效评价,五部词典分别筛选出积极评价和消极评价各 100 条,积极评价总计 500 条,消极评价 500 条,样本共计 1 000 条。

表 1 词典信息及评价文本数量

词典	近一年下载量	评分数量	评分	消极评价量	积极评价量
有道	约 610 万	46 万条	4.7 分	100 条	100 条
金山词霸	约 100 万	47 万条	4.8 分	100 条	100 条
欧路	约 100 万	7.4 万条	4.8 分	100 条	100 条
牛津高阶	约 30 万	7 270 条	4.4 分	100 条	100 条
朗文当代	约 1 万	5 756 条	4.5 分	100 条	100 条

2.4 词频分析

将 500 条积极评价和 500 条消极评价储存为 txt 文本格式,然后将文本导入 ROST CM6 软件中进行分析。先利用 ROST CM6 对评价文本进行分词,分词完成后,再对其进行词频统计。为凸显用户需求,在词频表中剔除了“词典”“有道”“牛津”“欧路”等高频词。表 2 为消极评价和积极评价中排名前 30 的高频词。

表 2 积极评价文本和消极评价文本高频词表

积极评价文本高频词						消极评价文本高频词					
序号	特征	序号	特征	序号	特征	序号	特征	序号	特征	序号	特征
1	单词	11	解释	21	辨析	1	单词	11	视频	21	查询
2	功能	12	全面	22	英英	2	翻译	12	发音	22	收费
3	英语	13	广告	23	视频	3	没有	13	模式	23	使用
4	翻译	14	没有	24	帮助	4	功能	14	软件	24	错误
5	例句	15	权威	25	轻松	5	广告	15	打开	25	页面
6	释义	16	准确	26	文章	6	不好	16	语音	26	自动
7	发音	17	界面	27	详细	7	词语	17	拍照	27	每天
8	学习	18	简洁	28	词汇	8	闪退	18	安排	28	笔记
9	方便	19	语法	29	强大	9	离线	19	横屏	29	复习
10	考试	20	图示	30	适合	10	消失	20	界面	30	不准

据表 2 可初步推知,用户的积极评价更为关注词典的微观结构信息如“翻译”“例句”“释义”“发音”等,而消极评价除关注微观结构信息外,还较关注系统的稳定性如“闪退”“消失”,功能的有无如“没有”“离线”。

2.5 社会语义网络分析

词频分析能通过提取的词典评价文本中的高频词反映研究对象的主要特征,但无法反映词组在特定意义上的联系以及文本深层次的结构关系,而社会语义网络分析则能通过构建概念和语义关系的网络图来直观展现要素之间的关系,有助于了解用户的具体使用偏好,如图 2 中“英英”指向“释义”,表明“英英释义”作为共现词,在总文本中出现频率较高。社会语义网络分析的过程为,首先分别对总的积极评价文本和消极评价文本进行分词处理,提取高频词并过滤掉无意义词汇后,对文本进行特征分析,并形成 VNA 文件,启动 NETDRAW,绘制完成社会语义网络图。积极评价社会语义网络如图 2,消极评价的语义网络如图 3。

积极评价的社会语义网络图中“英语”“例句”“单词”“功能”的节点连接量最大,成为主节点。相较之下,消极评价社会语义网络图更多地展现出用户的情绪,如有大量的“不

用偏好。根据分析单元、词典学理论并结合相关编码研究(韦艳丽等,2022;程昱等,2011),将用户需求划分为七个部分:内容需求、界面需求、资源需求、系统需求、经济需求、服务需求和功能需求。

表3 分析单元

积极评价			消极评价		
1 功能全面	21 训练口语	41 例句有用	56 没有离线	76 流量消耗	96 呈现空白
2 没有广告	22 医学学习	42 发音方便	57 经常闪退	77 词义太难	97 客服效率
3 翻译准确	23 例句全面	43 知识很多	58 不能查询	78 文件消失	98 发音图标
4 图示英语	24 单词词缀	44 插图有趣	59 单词没有	79 翻译速度	99 客服差劲
5 例句发音	25 单词拼写	45 发音好听	60 开屏广告	80 语音翻译	100 词义关联
6 功能方便	26 拼写正确	46 释义全面	61 翻译不好	81 单词错误	
7 释义权威	27 训练听力	47 单词词根	62 完善发音	82 夜间模式	
8 英语文章	28 托福考试	48 医学翻译	63 广告太多	83 发音差劲	
9 翻译全面	29 查询速度	49 四级考试	64 翻译不出	84 界面杂乱	
10 英英释义	30 单词打卡	50 四级复习	65 不能翻译	85 界面花哨	
11 解释详细	31 单词复习	51 查询方便	66 释义错误	86 复习安排	
12 语法解释	32 图示美观	52 轻松学习	67 单词颜色	87 学习视频	
13 界面简洁	33 字体美观	53 帮助发音	68 音频延迟	88 电话骚扰	
14 单词考试	34 单词很多	54 拍照翻译	69 功能收费	89 经常卡顿	
15 单词齐全	35 考研学习	55 专业单词	70 词语收费	90 经常黑屏	
16 英语阅读	36 例句收藏		71 发音卡顿	91 语音识别	
17 高中学习	37 发音标准		72 没有横屏	92 词义不好	
18 视频有趣	38 近义词辨析		73 视频杂乱	93 翻译出错	
19 例句翻译	39 单词辨析		74 翻译质量	94 没有翻译	
20 语法详细	40 例句权威		75 单词复习	95 消耗电量	

内容需求指用户期望词典能提供准确、详尽、权威的结构信息。界面需求指用户期望在进行查阅时词典界面能带来良好的审美体验。资源需求指词典内置的学习资源能够满足用户不同阶段英语学习的需求或满足用户提升不同言语技能的需求。系统需求指用户对系统稳定性的需求。经济需求则指用户期望在使用词典时能够花费较少的流量、电量,同时期望词典没有额外的付费功能。服务需求指用户期望客服具有友好、及时的服务态度。功能需求指用户期望词典能够具有多样的功能,这些功能是指在词典结构信息之外的额外功能,如是否具有横屏模式、夜间模式、离线模式等。使用需求下可划分出不同的使用偏好,如表4所示。

表 4 单元归类

使用需求	使用偏好	编号	占比	总计
内容需求	释义	7、10、66、77、92	5%	58%
	义项	11、46	2%	
	词形	25、26、81、	3%	
	例证	23、40、41	3%	
	发音	5、37、42、45、62、68、71、83	8%	
	翻译	3、9、19、48、54、61、64、65、74、79、80、93、94	13%	
	辨析	39、38	2%	
	语法	12、20、24、47、100	5%	
	多模态呈现	4、18、32、33、44、67、73、87、98	9%	
	检索	29、51、58、91	4%	
	收词	15、34、55、59	4%	
界面需求	界面美观	13、84、85、2、60、63	6%	6%
资源需求	学习资源多样	8、14、16、17、21、22、27、28、 0、31、35、43、49、50、52、53、75、86	18%	18%
系统需求	系统稳定	57、78、89、90、96	5%	5%
经济需求	消耗较小	76、95	2%	4%
	功能免费	69、70	2%	
服务需求	客服友好	88、99、97	3%	3%
功能需求	功能多样	1、6、36、56、72、82	6%	6%

2.8 单元归类与分析

编码划分完成后,需将 100 个分析单元归类。归类必须完全、彻底,能适用于所有分析材料,使所有分析单位都可归入相应的类别,不能出现无处可归的现象,且每一分析单位只能归入一个类别。由于部分分析单元表意较为模糊,还需结合原始评价文本了解具体含义。如“11 解释详细”“46 释义全面”,其典型评价为“解释很详细,很多意思都有”“释义比较全面,可以知道单词的很多意思”,可知用户所要表达的应是义项丰富,具体分类如表 4 所示。分类统计结果显示,在线词典用户关注的使用需求顺序依次为内容需求(58%)、资源需求(18%)、功能需求(6%)、界面需求(6%)、系统需求(5%)、经济需求(4%)、服务需求(3%)。

2.9 信度检验

内容分析中的信度可以定义为不同编码员对内容归类的一致性(Daniel et al. ,2010: 124)。为检验分类的信度,选择了 A、B、C 三位硕士研究生作为编码人员,三位编码人员皆有词典编纂和研究经历。对编码人员进行编码培训后,三位编码员独立完成归类任务。归类完成后,采用 Holsti 信度值计算内容分析中编码员归类的一致性,其中:

$$\text{相互同意度} = \frac{2M}{N1+N2}, \text{holsti's 信度} = \frac{n \times \text{相互同意度均值}}{1 + [(n-1) \times \text{相互同意度均值}]}$$

M 为两位编码人员编码一致数量,N1 和 N2 分别为编码单位数,n 为编码人员数量。据表 5 数据显示,Holsti's 信度值达到 0.983%,满足理论要求(夏长杰 等,2017:28)。

表 5 编码信度检验

统计	编码一致数量 (N = 100)	相互同意度	相互同意度均值	Holsti's 信度值
编码员 A 与编码员 B 一致数	95	0.95		
编码员 A 与编码员 C 一致数	94	0.94	0.953	0.983
编码员 C 与编码员 B 一致数	97	0.97		

3 评价模型建立

评价维度确立之后,开始纳入具体的评价项。虽有 100 个分析单元作为评价项目确立的依据,但仍遗漏部分项目。遗失的项目多为词典用户不常使用或不关注的微观结构信息,如语用、标记、文化等。雍和明(2003)的调查便指出,有些词典用户甚至没意识到英汉词典中包括有词源信息,52%的被访者从未有过或很少查阅过有关“美国英语”或“正式用语”之类的语言或文体变体方面的信息。

因此,为使评价框架更为完整,还需要纳入“专家引导下”的用户需求调查的研究成果。这些调查多以发放问卷、访谈的方式来进行,而问卷和访谈的题目多是专家依照自身使用或编纂词典经验而确立。一定程度的专家介入使得整个评价框架具有了完整性,最后结合雍和明(2003)、陈玉珍(2007)、胡文飞(2013)等人的研究制定评价模型,如表 6 所示。

表 6 用户视角下在线英语学习词典评价模型

一级编码	二级编码	评价项
	释义	①释义权威、准确。②释义易理解。③有多语释义。④释义对等词语义准确、形式实用。⑤释义对等词排序符合需求。
	义项	⑥多义词义项划分详细。⑦多义词义项排序符合需求。
	词形	⑧单词拼写形式正确。⑨有易拼写错单词的正误对照。
	例证	⑩例证数量丰富。⑪例证权威、准确。⑫例证实用。⑬例证配有准确、地道的翻译。⑭例证典型。⑮例证形式多样。
	发音	⑯提供了多种发音变体。⑰发音标准。⑱发音优美、流畅。
	翻译	⑲内置翻译功能翻译准确、地道。⑳翻译速度快。㉑为待翻译文本提供了多种语篇识别功能。㉒语篇识别功能能准确识别信息。
	辨析	㉓单词辨析全面、准确。
内容需求	语法	㉔单词语法信息全面、准确。㉕语法解释易理解。㉖语法标注符号易理解。
	多模态呈现	㉗多模态信息丰富、准确。㉘多模态信息生动有趣。㉙多模态信息美观。㉚多模态信息未干扰查词过程。
	语用	㉛单词语用信息全面、准确。
	标记	㉜统一、规范地标记了单词的词类、专业、词源等信息。
	文化	㉝提供了准确、丰富的词语文化信息。
	超链接	㉞有多样的超链接参见。
	检索	㉟具备多样的检索功能。㊱检索功能能准确识别所录入的信息。㊲单词检索速度快。
	收词	㊳收词量大。㊴收词全面,收录了全面的专业词汇、新词等。

一级编码	二级编码	评价项
界面需求	界面美观	④⑩界面美观,排版合理。
资源需求	学习资源多样、有效	④⑪有多样的学习资源。④⑫学习资源有助于我通过各类型语言考试。④⑬学习资源有助于我提高言语技能。
系统需求	系统稳定	④⑭系统稳定,鲜少出现卡顿、黑屏、闪退等问题。
经济需求	消耗较小 功能免费	④⑮运行时消耗的流量、电量少。 ④⑯无额外付费解锁功能要求。
服务需求	客服友好	④⑰客服服务态度好。④⑱客服无推销行为。
社交需求	分享与监督	④⑲能满足我多样的社交需求。
功能需求	功能多样	④⑳功能多样,如有夜间模式、收藏词语、离线查词等功能。

4 模型信效度检验

4.1 测试对象及方法

以五部词典中下载量最大的有道作为考察对象。为防止部分用户因之前从未关注过部分项目而导致信度降低,提前告知问卷测量维度,用户完成对未关注项目的尝试后,对 50 个项目进行评分(五度量表,“1”表示很不符合,“5”表示“很符合”),问卷通过问卷星系统发送至用户。共收回 210 份有效问卷,用户年龄为 16-27 岁,词典平均使用时长为 12.5 个月,平均使用频率为每星期 3 次。

4.2 信效度检验

本研究采用 SPSS 26.0 对量表的信效度检验。50 个测试题目,内部一致性信度 Cronban’sα 系数为 0.765>大于参考值 0.6,信度较高。进行探索性因素分析,样本适当性和 Bartlett 球形检验的结果显示,KMO 值为 0.761,近似卡方值为 9 000.90, $P<0.001$,表明数据适合进行探索性因素分析。

4.3 结果分析

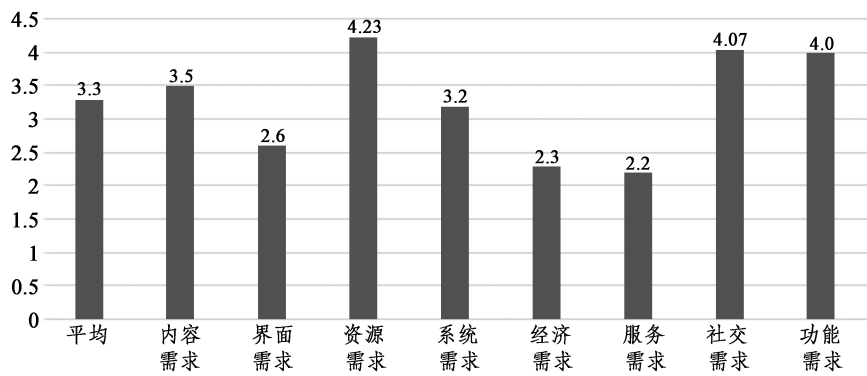


图 4 有道各使用需求得分情况/分

据图 4 可知,有道最终得分为 3.3 分。在所有需求维度中,“资源需求”均值得分最高,为 4.23 分,表明有道向用户提供了较为丰富的学习资源,这些学习资源有助于用户提高言

语技能并通过各类英语考试。“功能需求”“社交需求”评分均值得分靠前分别为 4.0 分和 4.07 分,表明用户认为有道的功能较为丰富,自己的社交需求也能得到一定程度的满足。“经济需求”和“服务需求”均值得分较低分别为 2.3 分和 2.2 分,表明有道在使用时消耗了较多的流量、电量等,同时部分用户反映有道客服友好性欠佳,存在回复不及时、回复内容无效及强制推销产品等问题。

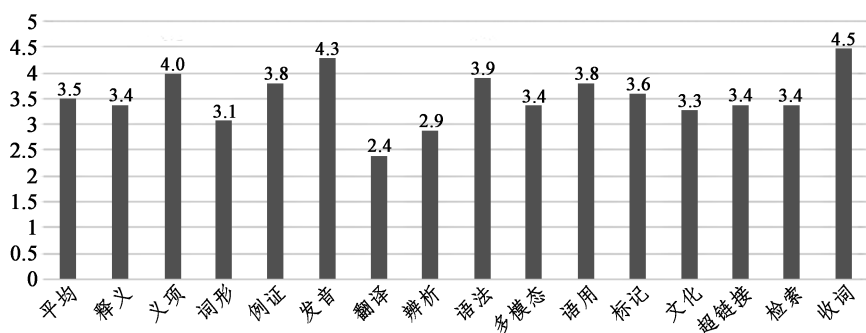


图5 有道内容需求各评价项得分情况/分

“内容需求”是词典评价的重心。据图 5 可知,有道“内容需求”中“收词”“发音”“义项”评分均值得分较高,这表明,有道收词量较大且全面,词语、例证等的发音较为标准、流畅、优美,多义词义项较为丰富且排序较符合用户需求。“翻译”“辨析”“词形”评分较低,这表明,有道内置翻译功能所译文本质量有待提高,拍照翻译、语音翻译在识别文本时存在一定问题。同时,比较缺乏对词语全面、准确的辨析。词形上,因未提供易拼写错单词的正误对照导致得分较低。

5 总结与讨论

通过对在线英语学习词典用户评论的内容分析,我们了解到用户的使用偏好,并结合相关文献研究构建出用户视角下的在线词典评价模型。最后,以有道为调查对象,将模型运用于评价实践中,结果显示,模型具有较好的信效度。同时,评价模型的构建过程也可为在线学习词典评价研究提供启示。

评价从本质上来说是客体与主体之间价值关系的体现,价值关系是以主体尺度为依据的关系内容,客体满足主体自身的需求是形成这种价值关系的前提和基础(李守福, 2005)。对于在线学习词典而言,其评价主体的构成、需求,主客体之间的关系以及客体的内容构成等都是复杂、多元的,故对其评价应改变以往单一、片面的一元评价模式而应形成系统、全面的多元评价模式,这主要体现在:

评价主体的多元化。以往的词典评价多是以词典学界的专家学者们为评价主体的单一主体评价。接受美学理论认为“在作者、作品与读者的三角关系中,读者绝不仅仅是被动的部分或者仅仅作出一种反应,相反,他自身就是历史的一个能动的构成”(金元浦,

2002)。同样,词典编纂是一个传播和分享知识与信息的交际过程,而用户在这一过程中发挥着重要的能动作用。同时,学习词典的“学习性”是为了使学习者获得好的学习效果并有助于教师提高语言教学效果。因此在线学习词典评价的主体不应只局限于专家学者,用户、教师等也应参与到词典评价过程中才能保证评价结果的权威性、客观性和有效性。具体而言,各评价主体可依据自身评估优势有侧重地对词典进行评价。专家学者可借助专业知识评判词典文本编纂的科学性与规范性,如验证是否存在循环释义等问题。教师则基于语言教学经验评判词典的学习性,而用户则基于使用感受可评价词典客服的友好性、界面的美观性等。

评价内容的多元化。传统的词典评价研究主要是“以文本为中心”,但随着信息技术的发展和词典用户需求的扩大,现如今在线词典所能囊括的功能、内容等也更为多样,故其评价内容应不只局限于文本。

评价阶段的多元化。在线词典的编纂和运行是分阶段进行的,大致可分为编写阶段、试用前阶段、小范围运行阶段、运行阶段等,不同阶段的评价主体、评价目的、评价方式都应有所不同。

评价方式的多元化。总体而言,词典评价应坚持定性和定量结合的方法。从评价阶段上看,编写阶段以词典编者团体为评价主体,主要以同伴互评的方式监控词条编写质量。在试用前阶段以非编者的专家学者为评价主体,侧重对词典内容的评价,可采用定量分析法、对比法、综合分析法等。在运行阶段则以用户、教师等为评价主体,侧重对词典使用效果的追踪评价,可采用问卷调查法、访谈法、内容分析法等。

最后,我们须知词典评价的目的不是为了给词典贴标签、排名,而是为了促进词典理论与实践的进步,引导用户选择适合、优质的词典,指导编写团队对词典的功能、内容等进行优化升级。

参考文献:

- Dilin, L, D. Yaochen & Y. Shiyan 2021. Evaluating Popular Online English-Chinese E-dictionaries in China by Applying Lew and Szarowska's (2017) Evaluation Framework[J]. *International Journal of Lexicography*(34):157-182.
- Howard, Gardner. 1983. *Frames of Mind: The Theory of Multiple Intelligences*[M]. New York: Basic Books.
- Lew, R. & A. Szarowska. 2017. Evaluating Online Bilingual Dictionaries: the Case of Popular Free English-polish Dictionaries[J]. *Recall*(PT.2): 138-159.
- Nerendorf, K. A. 2016. *The Content Analysis Guidebook*[M]. Los Angeles: Sage.
- Pearsons, E. and W. Nichols. 2013. Toward a Framework for Reviewing Online English Dictionaries[J]. *Dictionaries Journal of the Dictionary Society of North America*(34): 201-210.
- Riffe, D. S. Lacy & F. G. Fico. 2010. *Analyzing Media Messages: Using Quantitative Content Analysis in Research*[M]. trans, by Ji Meiyun, Beijing: Tsinghua University Press.
- Yamada, S. 2013. A Test of the Proposed Framework for Reviewing Online Dictionaries: M-W. com, Dictionary. com, Macmillan

- Dictionary. com, Dictionary. Cambridge. org, and Oxford Dictionaries. com[J]. *Dictionaries Journal of the Dictionary Society of North America*(1):211-224.
- 陈楚祥. 1994. 词典评价标准十题[J]. 辞书研究(1):10-21.
- 程罡, 高辉, 余胜泉. 2014. 基于真实用户评论信息构建移动学习资源的评价模型[J]. 现代远程教育(1):43-53.
- 黄锦鸿. 2021. 在线英语学习词典拼合词的收录研究[J]. 辞书研究(6):36-45+130.
- 胡文飞. 2013. 用户需求与汉英学习词典微观结构的构建:基于中国 EFL 学习者的实证研究[J]. 外国语文(2):72-78.
- 胡文飞. 2015. 二语习得的涌现特征与新型汉英学习词典表征模式的构建[J]. 外国语文(2):70-76.
- 姜治文, 文军. 1998. 关于拓展双语词典批评领域的思考[J]. 外国语(上海外国语大学学报)(6):67-71.
- 金元浦. 2002. 接受反应文论[G]. 济南:山东教育出版社.
- 郎建国, 李晋. 2003. 英语学习词典使用情况调查与思考[J]. 北京第二外国语学院学报(6):54-60+63.
- 李守福. 2005. 论大学评价的价值取向[J]. 比较教育研究(12):57-61.
- 罗思明, 曹杰旺. 2006. 词典批评类型与理论构建[J]. 山东外语教学(5):60-63.
- 王娇, 胡文飞. 2016. 汉英词典评论研究:现状分析与结构描述[J]. 辞书研究(6):48-57+94.
- 夏杰长, 刘奕. 2017. 中国服务业发展报告(2016—2017): 迈向服务业强国: 约束条件、时序选择与实现路径[M]. 北京: 经济管理出版社.
- 韦艳丽, 蒋小庆, 李安, 徐曦. 2022. 基于内容分析法的阅读类 APP 用户使用偏好研究[J]. 包装工程(8):157-163+216.
- 魏向清. 2001. 关于构建双语词典批评理论体系的思考[J]. 外语与外语教学(1):59-61.
- 雍和明. 2003. 英汉双语词典与英语单语词典用户语言需求与信息检索的比较[J]. 辞书研究(6):123-133.
- 章宜华, 黄建华. 2000. 关于双语词典评奖的几点思考[J]. 辞书研究(4):65-72.
- 邹酆. 2004. 辞书学丛稿[M]. 武汉:崇文书局.

Construction of Users' Evaluation Model of Online English Learning Dictionary Based on Content Analysis

CHEN Xiande YANG Yuling

Abstract: The evaluation research of dictionary is of great significance to the dictionary's compilation, but the previous studies rarely established evaluation models from the perspective of users. Taking users' evaluations of five online English learning dictionaries on the App Store as samples, using ROST Content Mining 6 to extract high-frequency words and perform social network semantic analysis from online dictionary users' positive and negative evaluation texts, this study has discovered users' potential needs and usage preferences. Combined with high-frequency vocabulary and social network semantic map, 100 effective analysis units are selected, and after the analysis unit is coded and the classification test performed, a user evaluation model is initially established. Combined with the researches of dictionary users' needs, the evaluation model is improved. Finally, taking Youdao Dictionary as the object of investigation, the model is applied to the evaluation practice. The verification results show that the model has good reliability and validity.

Key words: content analysis method; online English learning dictionary; usage preference; users' evaluation; evaluation model

责任编辑:朱晓云