

# On the Construction and Application of Personality Vocabulary Corpus

Ying Chang\*, Chun Yong#

School of Psychology, Inner Mongolia Normal University Hohhot, Hohhot Inner Mongolia  
Email: changyingpsy@163.com, #yongchun0580@sina.com

Received: Jul. 24<sup>th</sup>, 2019; accepted: Aug. 8<sup>th</sup>, 2019; published: Aug. 15<sup>th</sup>, 2019

---

## Abstract

Healthy and sound personality education is an inevitable requirement to comprehensively promote psychological health education and realize students' physical and mental development. This paper adopts document analysis, lexicology, semantics and database description to collect and encode personality words. The steps of building the corpus of personality vocabulary are as follows: corpus collection, corpus standardization, corpus labeling, updating and perfecting, and forming the corpus of personality vocabulary. Vocabulary is the basic material for the study of personality, including the language information base of personality description vocabulary, personality evaluation vocabulary, personality trait vocabulary and personality state vocabulary, which aims to promote the development of students' positive personality and cultivate the personality characteristics needed by the development of modern society in the mental health education.

## Keywords

Psychological Health Education, Personality Vocabulary, Corpus, Big Data, Information Processing

---

# 论人格词汇语料库的创建及应用

常颖\*, 永春#

内蒙古师范大学心理学院, 内蒙古 呼和浩特  
Email: changyingpsy@163.com, #yongchun0580@sina.com

收稿日期: 2019年7月24日; 录用日期: 2019年8月8日; 发布日期: 2019年8月15日

---

## 摘要

健康健全的人格教育是全面推进心理健康教育, 实现学生身心发展的必然要求。文章采用文献分析法、

---

\*第一作者。

#通讯作者。

词汇学研究法、语义学研究法和数据库描述法, 对人格词汇进行采集与编码。人格词汇语料库的建库步骤为: 语料采集, 语料规范, 语料标注、更新与完善, 形成人格词汇语料库。词汇是用于人格研究的基本材料, 包括人格描述词汇和人格评价词汇、人格特质词汇和人格状态词汇的语言信息库, 其目的是在学生心理健康教育中促进其积极人格的发展和培养现代社会发展所需要的人格特征。

## 关键词

心理健康教育, 人格词汇, 语料库, 大数据, 信息处理

Copyright © 2019 by author(s) and Hans Publishers Inc.

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

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



Open Access

## 1. 引言

语料库是指“有目的、有计划、收集具有代表性的、一定结构的、可以被计算机所识别的具有一定规模的语言的集合”。从 20 世纪 90 年代起, 语料库和语料语言迅速发展, 多种形式的语料库已经在诸多方面发挥了作用(王利娜, 2012)。通过对语言中与人格特征相关的形容词的分析, 可以得到人格结构的具体维度, 并反映出文化内涵(王登峰&崔红, 2000)。目前我国人格词汇语料库的建设情况尚不完善, 文章就学生心理健康教育的人格词汇语料库的建设问题, 对人格词汇语料库的建设进行探讨, 并对语料库的应用提供建议。

## 2. 人格词汇语料库的创建意义

英国学者高尔顿在 1884 年提出了人格词汇学假说。高尔顿首先把词典作为一种工具来估计所有词汇中描述人格的词汇数量, 并确定这些词汇能够在何种程度上传达这种特征的含义。“人格心理学之父”G.W.奥尔波特用传记文献的词汇进行分析, 较早地对个人特质进行了深入的研究, 为人格特质的研究做出了贡献, 这也使“词汇分析法”成为人格研究的主要方法之一(谭磊, 2017)。“词汇学假设”认为, 对自然语言中习得的人格词汇进行分析, 会得到一组数目有限的特质, 这些特质代表了这种语言环境中人最重要的特征, 人格词汇学的研究很难独立于其存在的语境。由此可见, 建设人格词汇语料库能够弥补这一不足。20 世纪 90 年代初, 王登峰、杨国枢等中国学者对中国人格的词汇学进行了探索, 促进了词汇学研究的本土化进程。词汇分析法假设是人格心理学研究的常用方法, 其含义为在人们生活中, 与人格有关的一切活动及个体差异最终都会被编码到他们的语言中去, 其活动和差异越重要, 就越有可能由一个单独的词来描述(赵伶俐, 2010)。

人类的社会日益认识到了国民心理素质是一个国家长治久安、繁荣昌盛的根本性基础因素, 心理健康教育是人类社会真正幸福的必要内涵。由此可见, 心理健康教育是学生在人格发展过程中必不可少的重要因素。将人格词汇语料库的创建辅助心理健康教育是一项全新的探索与尝试, 尚由于此方面没有过多的研究, 在研究中理论支撑较少、研究较少等问题比较常见, 但这也是研究的创新性与独特性。

创建专用语料库不但有利于翻译效率的提高, 而且有利于推动语言服务的发展。人格词汇语料库的创建属于专用语料库, 建设人格词汇语料库能规范人格词汇的参差不齐的现状。人格词汇的信息化处理是建设人格词汇语料库的重要组成部分, 为语料库的跨学科做出一定的贡献, 人格词汇的建设是当代信息社会发展的必要举措。当遇到有关人格词汇不明确时, 能够便捷、自由的检索语料, 并对语料进行了

充分的理解, 这突出了语料库的便捷性与实用性。人格词汇具有一定代表性、一定规模和结构的语料集合, 人格词汇可以成为人格发展研究的基础, 基于现有的语料库建设大部分为语言学领域研究, 缺少了跨学科的理论指导, 人格词汇语料库的选材不仅限于某一领域, 语料库建设大部分成为选材面向所有教材中有关人格的词汇学研究, 尤其是为语文教材中的人格词汇语料库对人格教育的具体实施提供便利。我国人格词汇语料库的发展并不完善, 有关人格词汇双语资料较少, 因此创建人格词汇语料库有利于推动语言服务的发展, 创建人格词汇语料库还有利于解决教育教学中人才培养的问题。

截至目前, 国内还没有专门服务于人格系统的语料库。只有在一些语料库的子库中, 能够零星的发展有关于人格的英汉/汉英双语语料。人格词汇语料库的空白不利于有关人格课题研究以及心理学、社会学服务人才的培养。本研究为语料数据库提供跨学科的理论指导。由此可见, 创建人格词汇语料库非常必要。

### 3. 人格词汇语料库的建设

#### 3.1. 建设步骤

人格词汇语料库的系统是以语料库为核心, 包括计算机硬件、软件、语料库用户、语料采集和加工规则、语料库管理和应用程序的一个完整系统(何婷婷, 2003)。其建库过程可以表示为图 1:

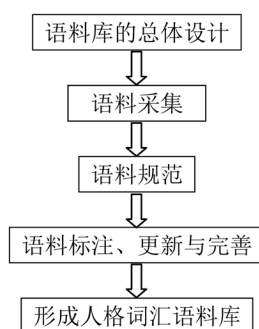


Figure 1. Research flow chart

图 1. 研究流程图

##### 3.1.1. 语料采集与标注

通过文献分析法、词汇学研究法、语义学研究法和数据库描述法对人格词汇进行采集, 最终形成人格词汇文本。文本人格语料库的构建, 为本研究提供更全面的语料库, 取材要保证多样化, 从《现代汉语词典》、中小学语文教材、名著、小说集和诗歌集等书籍采集人格有关词汇, 精心加工, 设计词汇属性和标记, 为撰写文本打好基础。确定人格词汇标注是一个非常庞大又基础性的工作, 根据人格词汇特点对语料不断的进行标注与更新, 并在具体的调查分析调查中不断地完善和扩充标注集。通过调查分析可以探索最核心的人格特质, 调查其人格词汇的喜好度、利用度、熟悉度等, 通过频次统计并确定核心人格词汇, 建设一个覆盖面较广的人格语料数据库, 建设多维度、多变量、多属性的人格语料数据库。

##### 3.1.2. 语料规范

人格词汇总数据库包括一级分库、二级分库和三级分库。一级分库人格理智特征词汇数据库, 包括二级分库感知觉词汇分库、记忆词汇分库、想象词汇分库、思维词汇分库、语言词汇分库和行为词汇分库; 一级分库人格态度特征词汇数据库, 包括二级分库他人态度词汇分库、劳动态度词汇分库、自我态

度词汇分库、集体态度词汇分库、社会态度词汇分库和家庭态度词汇分库；一级分库人格情绪特征词汇数据库，包括二级分库愉快和激活水平高词汇分库、愉快和激活水平低词汇分库、不愉快和激活水平高词汇分库，不愉快和激活水平低词汇分库；一级分库人格意志特征词汇数据库，包括二级分库自觉有关词汇分库、坚韧性有关词汇分库、自制力有关词汇分库、果断性有关词汇分库。将以上的词汇分库按性质属性、强度属性、动力性属性、利用度属性、性别属性、年龄属性、稳定性属性、词频和喜好度对人格词汇进行规范，将规范后的词汇纳入信息库。

### 3.2. 建设过程中出现的问题

#### 3.2.1. 理论研究匮乏

在中国，研究人格词汇语料库的专业人员很少，由于缺乏相应的理论依据，将会导致语料库相关建设中存在诸多问题。对知网中有关人格词汇语料库的文献检索中，未发现有关人格语料库的相关研究，尽管语料库的相关研究较多，但人格词汇语料库的研究较少。从数据上看，国内相关研究者针对人格词汇的研究十分匮乏，这会对人格语料数据库的创建将产生不良影响。

#### 3.2.2. 研究方法内容局限

在上世纪 60 年代时语料库才开始发展，最开始人们只用语料库做简单的分析，例如词频统计等，之后才增加了词的语法属性的标注。但时至今日，语料标注没有实质性突破。语料库研究的目的是深入研究语言的结构和进化规律。语料库中标注的不成熟和不完善会使相关研究停滞不前，语义研究不够深入。语料库与其他学科的交叉研究尚不成规模，语料库在心理学、社会学、文学和翻译学等方面的研究数量少、规模小、影响力较弱(石志亮, 2017)。跨专业语料库的创建越来越成为信息、科技发展的必然选择。

## 4. 人格词汇语料库在心理健康教育中的应用

在数字化时代背景下，心理学健康教育更加需要适应时代变化，将人格词汇进行信息化处理是人格词汇语料库建设的重要组成部分，为心理健康教育中的人格测评、人格结构的确定和人格特点的分析提供大数据分析技术。人格词汇的信息化处理是建设人格词汇语料库的重要组成部分，为语料库的跨学科做出一定的贡献，人格词汇语料库的建设是当代信息社会发展的必要举措。建设人格词汇语料库，能够改善在心理健康教育中人格词汇的参差不齐的现状。当遇到有关人格词汇不明确时，能够便捷、自由地检索语料，并对语料进行充分的理解，突出了语料库的便捷性和实用性。

积极心理学作为当代西方心理学三大研究取向之一，对国内的心理学发展起到了高度的借鉴作用。积极的价值观取向特征主张心理学应把自己的工作重心放在研究和培养人固有的积极潜力上，通过培养或发掘人的这种积极力量而使之成为真正健康并生活幸福的人(况志华&叶浩生, 2005)。积极价值观取向的特征完全与心理健康教育相辅相成，而人格词汇语料库的创建能够非常便捷的在心理健康教育中呈现积极心理学的使用，在语料库中检索积极的人格特质词汇就会出现很多相关语料，从而方便教师的教育教学更加有效的运用积极心理学，使积极心理学的渗透到教育的各个方面和各个环节中，并且能够为学习者研究提供量化研究的基础。

## 5. 结论

随着社会的经济快速发展，学生的心理健康教育应与时代相结合，为培养学生适应现代社会发展的健全人格，建设完备的人格词汇语料库势在必行。在语料库和语料库语言学迅速发展的大环境下，人格词汇语料库的建设会更加顺利，这为语料库的跨学科和跨领域做出重要贡献。

## 基金项目

内蒙古自治区高等学校科学研究重点项目 NJSZ19020, “面向信息处理的蒙古语人格词汇库框架设计”阶段性研究成果。

## 参考文献

- 何婷婷(2003). 语料库研究. 博士学位论文, 武汉: 华中师范大学.
- 况志华, 叶浩生(2005). 当代西方心理学的三种新取向及其比较. *心理学报*, 37(5), 702-709.
- 石志亮(2017). 语料库语言学研究现状分析. *吉林省教育学院学报*, 33(1), 98-100.
- 谭磊(2017). 颜真卿人格特质相关研究. 硕士学位论文, 重庆: 西南大学.
- 王登峰, 崔红(2000). 文化、语言、人格结构. *北京大学学报(哲学社会科学版)*, (4), 38-45.
- 王利娜(2012). 专门用途英语语料库建设的现状及对策. *海外英语*, (5), 121-122.
- 赵伶俐(2010). 人格特质词汇学的美学审视. *西南大学学报(社会科学版)*, 36(1), 16-22.

### 知网检索的两种方式:

1. 打开知网首页: <http://cnki.net/>, 点击页面中“外文资源总库 CNKI SCHOLAR”, 跳转至: <http://scholar.cnki.net/new>, 搜索框内直接输入文章标题, 即可查询;  
或点击“高级检索”, 下拉列表框选择: [ISSN], 输入期刊 ISSN: 2160-7273, 即可查询。
2. 通过知网首页 <http://cnki.net/>顶部“旧版入口”进入知网旧版: <http://www.cnki.net/old/>, 左侧选择“国际文献总库”进入, 搜索框直接输入文章标题, 即可查询。

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [ap@hanspub.org](mailto:ap@hanspub.org)