

DOI: 10.3969/j.issn.1672-8874.2011.03.030

# 分级教学环境下大学英语分级词表的构建研究

李明富, 燕丽勤, 刘晶

(国防科学技术大学 人文与社会科学学院, 湖南 长沙 410073)

**[摘要]** 教学词汇表是外语教学大纲的重要组成部分,也是教材编写、课程教学和测试的参考依据。但现今使用的词汇表分级不明确,收词也不全面,制约着词汇教学和测试。因此,构建一个分级词汇表意义重大。本文围绕如何构建大学英语分级词表这一研究主题,探讨了运用 RANGE 构建分级词表的依据,制表过程及其对英语教学和研究的重要意义。

**[关键词]** RANGE; 分级词表; 构建

**[中图分类号]** G642.0 **[文献标识码]** A **[文章编号]** 1672-8874(2011)03-0090-04

## The Compilation of Level - based College English Word List for Level - based Teaching

LI Ming - fu, YAN Li - qin, LIU jing

(College of Humanities and Social Sciences, National University of Defense Technology, Changsha 410073, China)

**Abstract:** A word list is the crucial part of English syllabus, and the basis for compiling textbook, teaching and evaluations. Now the word list we use is not appropriate in terms of word gradation and selection, which hinders vocabulary teaching and testing. Therefore, it is necessary to make a level - based college English word list. The article expounds how to make a level - based college English word list for Chinese students. It discusses the principles and procedures of making the word lists, and their significance in English teaching and research.

**Key words:** RANGE; level - based word list; constructing

### 一、研究背景

语言学家 Wilkins 曾说“要是没有语法,很多东西可能无法表达;要是没有词汇,什么东西都无法表达<sup>[1]</sup>。”在英语教学中,词汇在英语听、说、读、写各项技能的培养和训练方面都起到了基础作用。现在,我国英语专业和非专业的学生都有各自的统一词汇表,这些词汇表虽然规定了具体的单词数量,但是却没有“级”的概念,这与学生要面对的四、六级考试以及专业四、八级考试不能衔接起来。如果没有统一的分级词表,那么教材的编撰和选择也就没有统一的原则与方法。

从现有词表制定的过程来看,制表原则和方法缺少量化分析,主观因素较大。大学英语词表工作组成立后,首先选出约1万词的备选词目。然后根据具体词汇量要求,对这1万个词目进行删改,将拟增删的四、六级词目提交给大学外语教学指导委员会英语组扩大会议讨论。根据会议的讨论结果,词表组确定了四、六级词表的词目。《大学英语教学大纲》词汇表共收录单词6500个,其中要求入学前应掌握单词数约为1800个,在1-4级阶段掌握单词数

约为2400个,在5-6级阶段掌握单词数约为1300个,在六级之后的高级英语阶段掌握单词数约为1000个<sup>[2]</sup>。

以这种方法制定的词表在选词原则、词汇分级和词表适应性等方面存在一些问题。国内的专家教授对此做过一些研究,指出了词表存在的主要问题。

#### 1、屈折形式的收录

《大纲》词表在单词的屈折形式收录方面,存在着原则模糊,主观性较强的问题。《大纲》词表同时收录一些单词的屈折形式,比如说,limit/limited; tire/tired; think/thinking; train/training。但是另外一些单词的屈折形式却没有收录,比如,大纲收录了tire; work; teach,却没有收录相应的屈折形式tiring, working, teaching。

#### 2、派生词的收录

在派生词的收录方面,《大纲》词表也存在着收词原则模糊的问题,对前缀和后缀的处理原则也不一致,导致词汇收录不够全面,不够系统。比如,《大纲》词表同时收录了final/finally; certain/uncertain,而对于artificial/artificially; active/inactive等单词,《大纲》词表却只收录了其中一个单词。

[收稿日期] 2010-11-08

[作者简介] 李明富(1988-),男,湖南浏阳人,国防科学技术大学人文与社会科学学院硕士研究生,研究方向:军事外语教育。

### 3、词汇分级

对于确定收录的词汇,《大纲》词表没有合理地进行词汇分级,因此,出现同一词目的不同词形出现在不同的级别,这既不利于学习,也限制了词表词汇量的增加。比如,destroy; destruction; destructive 这三个单词同属一个词目,却被分成三个级别的词汇,分别对应中学词汇、四级词汇和六级词汇<sup>[3]</sup>。

我国现有的各种英语教学大纲词表,从基础教育到研究生教育,从大学英语专业到公共英语等级考试,纵横交错,不下七八种。由于这些词表都是由不同的词表工作组制定,缺乏协调统筹,导致词表有的相互重叠,有的上下脱节<sup>[4]</sup>。为解决这一问题,本研究就尝试在 RANGE 框架下建立专业与非专业的统一分级词表。

## 二、理论基础

### 1、词表构建的原则和方法

目前,国内外比较认同的词表构建原则与方法主要有以下几点:

#### (1) 代表性

代表性指的是词汇的语料来源要具有代表性。现在,许多词表的制定都是依托于语料库的支持。在制定词表时,为保证词汇选择的全面性,最好选择大型的语料库作为参照,比如, British National Corpus (BNC), English Bank, 这些语料库包含的材料内容广泛,词汇容量都达到1亿以上。

#### (2) 词频

词频是词表制定的重要依据,《大纲》词表在选择1万个备选词时,就是选取了词频靠前的词汇作为备选词。依托大型语料库进行的词频统计结果具有较强的可靠性、权威性以及实用性。词频不仅指的是某个词出现的频率,也包括这个词的变化形式在语料中出现的频率。词频作为一种选词原则,越来越受到词表和词典编纂者的重视。

#### (3) 词族

词族包括词头及其所有屈折变化形式。现在,我们有不同的词表词汇量定量标准。定量标准可以是以基础词为基本收词和统计单位,也可以是以词族或词目为基本单位。比如说,able/ability/unable,如果按照词形统计方式,这是三个词。但是以词族统计,则只能看成一个单词,他们同属于able这个基础词,或者说词头,而ability和unable只是词头下的屈折变化形式,不能看成不同的单词。基于RANGE制作的词表都是以词族作为统计分类单位,这样以词族或词目为收录统计单位的词表词汇定义能力虽然小于以基础词为单位的词表,但前者实际词汇量远远大于后者,前者收词比较系统化,也便于学习和研究。

#### (4) 其他原则

在制定词表中,除了以上三条原则外,我们还要考虑其他一些原则,比如说,词汇表的适应性。要删掉一些“不适用”的词汇,同时增加一些“鲜活的”词汇,比如说,card phone (磁卡电话),ATM (自动取款机),LD (激光光盘)<sup>[5]</sup>。另外,美式英语和英式英语的一致性,词汇意

义的收录都是需考虑的问题。

### 2、Nation 的研究

Nation 认为词汇的重要性是不一样的,我们应该先学习频率高的词汇,然后再学习频率低的词汇。而判断一个词的重要性最典型的方法就是参照词频和出现范围,也就是说,语言学习者了解和掌握高频词汇比低频词汇更有用。因此,运用计算机技术来确定这些低频词汇意义重大,这同样可以帮助确定文章的难易程度,确定教学中需要提前学习的词汇<sup>[6]</sup>。

Nation 还发现,从第一个1000词到第三个1000词,每掌握1000词,所认识的文本单词覆盖率会提高5% - 6%,但是,词汇量达到4000 - 8000时,即便词汇量增加100%,文本单词覆盖率也只会提高3.1%。由此可知,英语学习者应该按照单词词频的顺序学习和掌握词汇<sup>[7]</sup>。为此,他设计制作了RANGE词频统计软件以及三个基础词表,如表1所示<sup>[8]</sup>。

表1 RANGE 的三个基础词表

词频等级文件	词族	词形	描述
Basewrd1	998	4119	常用词
Basewrd2	988	3708	次常用词
Basewrd3	570	3107	非常用词
Off the list			生僻词

Nation 制作的RANGE词频统计软件以及三个基础词表功能齐全,操作简便,在测试词汇深度,编撰教材方面作用突出。RANGE 可让编撰者很容易发现各个不同程度学习者的词汇掌握情况,通过RANGE测试文章词汇覆盖率可以有效地避免教材编撰者通过经验感知和粗略估计挑选和标注生词。

## 三、研究方法

### 1、研究问题

在本研究中,围绕大学英语分级词表的构建,重点回答了以下几个问题:(1)专业词表与非专业词表单词重合率是多少?(2)如何构建大学英语分级词表(LUWL)?(3)在RANGE框架下,如何使用LUWL评估文本词汇覆盖率?

### 2、使用工具

本研究基于RANGE词频统计软件以及RNAGE BNC基础词表。

### 3、RANGE 使用说明

第一步,双击RANGE图标,打开运行界面。

第二步,在运行界面中点击File选项,点击Open项。

第三步,选择需要运行的文件。这些运行文件必须是TXT格式,如果是Word等其他格式的文件,则需要先转化成TXT格式再调入RANGE运行。

第四步,选定文件后,点击Open,再次点击File,选择

Save 选项, 并对运行结果命名。

第五步, 选择 RANGE 窗口中的需要运行的选项, 点击 Process Files 按钮, 开始运行。

#### 4、研究设计

第一阶段, 测试专业与非专业词表单词重合率。研究语料来自《大学英语课程要求(试行)》和《英语专业四、八级词汇表》。具体步骤如下:

(1) 将全部语料转换成 TXT 格式文档, 并保存。

(2) 将保存好的 TXT 文档调入 RANGE BNC 运行, 保存运行结果。

(3) 查看运行结果, 即可得到专业与非专业单词重合率。

第二阶段, 构建大学英语分级词汇表(LUWL)。建表过程可以分为以下几个步骤:

第一步, 词汇输入工作。将《大学英语课程要求(试行)》和《英语专业四、八级词汇表》的词汇全部输入到 TXT 文档中。

第二步, 将制作好的 TXT 文档调入 RANGE BNC 中运行, 剔除《大学英语课程要求(试行)》和《英语专业四、八级词汇表》两个词汇表中重复的单词。

第三步, 将没有被剔除的单词替换 RANGE BNC 中 16 个基础词表的单词, 并保留这 16 个词表的命名, 即 Basewrd1, Basewrd2...

第四步, 将替换好的词表与 RANGE 软件放在同一个文件夹中, 将替换后的大学英语分级词表命名为 Level-based University Word List (LUWL), 词表即制作完成。

第五步, 词表测试与修改。

第三阶段, 用 RANGE LUWL 评估文本词汇覆盖率。此研究中, 选取 2009 年 12 月份的全国大学英语四级考试(CET-4) Reading Four 作为测试对象。具体步骤如下:

(1) 将 Reading 转换成 TXT 格式文档, 并保存。

(2) 将保存好的 TXT 文档调入 RANGE LUWL 运行, 保存运行结果。

(3) 将运行结果制作成图表, 并做出相应的分析。

## 四、结果讨论

### 1、专业与非专业词汇重复率

经过 RANGE BNC 测试, 专业与非专业词表词汇重复分布情况如表 2 所示。

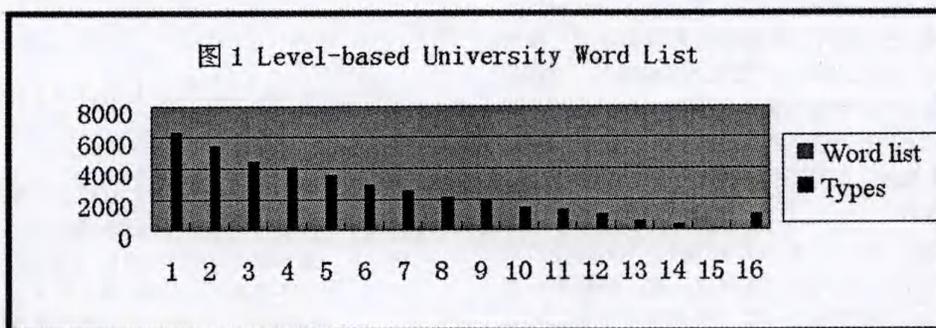
表 2 专业与非专业词表词汇重复分布情况

Word List	Number of Words	Word List	Number of Words	Word List	Number of Words	Word List	Number of Words
Basewrd1	1422	Basewrd5	629	Basewrd9	135	Basewrd13	15
Basewrd2	1118	Basewrd6	420	Basewrd10	86	Basewrd14	6
Basewrd3	922	Basewrd7	309	Basewrd11	58	Basewrd15	0
Basewrd4	790	Basewrd8	185	Basewrd12	22	Basewrd16	0
Not in any list				22			
Total number of words				6139 (6579)			
Total Percentage				93.31%			

从表 2 可以清楚地看出专业与非专业词表词汇重复在 RANGE BNC 中 16 个基础词表的分布情况。其中, 单词重复数总共为 6139 个, 而非专业词表总共单词数为 6579 个(不包括同形不同意的单词), 因此, 重合率为 93.31%。

专业与非专业词表词汇的高重合率也为两者结合为统一的分级词表提供了有利条件。

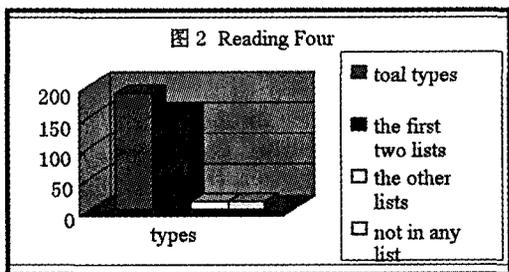
### 2、LUWL 词汇分布情况



基于 RANGE 构建的大学英语分级词表 (LUWL) 词汇基本分布情况与数据反映在图 2 中。我们知道, 专业与非专业的词汇数总共为 12906 个, 而他们的词形数为 39749 个, 词族数为 10136 个。如果以词族为统计单位, LUWL 的词汇容量则为 10136 个, 而如果以词形为统计单位, 那么 LUWL 的词汇容量将达到 39749 个, 远远大于 Nation 制作的三个基础词表的词族数 2556 个和词形数 10934 个。

### 3、词汇覆盖率测试

经过 RANGE LUWL 测试, 2009 年 12 月份 CET - 4 Reading Four 的词汇覆盖率情况如图 2 所示。



从 CET - 4 Reading Four 的运行结果来看, 在这篇阅读理解中有 14 个单词超出了 Basewrd1 和 Basewrd2 的范围, 大致相当于大学英语四级的词汇要求。其中有两个单词 *poise* 和 *alcoholic* 在文中分别给出了相应的中文意思, 这两个单词分别出现在 LUWL 中的 Basewrd10 和 Basewrd3 之中。剩下的 12 个超纲词汇也就是答题者在 Reading Four 中遇到的大纲词表之外的生词, 所占比率为 3.75%。这个生词比例完全符合出题要求, 因此, 从词汇覆盖率上来看, 这篇阅读理解难易度比较合理。由此也可说明 LUWL 的可靠性和适用性比较好。

### 4、小结

在量化分析的基础上, 本研究提出的三个研究问题都得到解决。首先, 通过 RANGE BNC 测出了专业与非专业词表的词汇重合率。然后, 在 RANGE 框架下, 完成了大学英语分级词表的构建。最后, 通过个案分析, 检验了 LUWL 的可靠性和适用性, 并熟悉了其在测试词汇广度方面的基本操作程序与步骤。

## 五、研究启示

本研究在 RANGE 框架下, 构建了大学英语分级词表 (LUWL)。该表的构建给英语词汇相关研究提供了新的思路和方法, 也为同学们的英语词汇学习提供了良好的教材。就 LUWL 的作用而言, 它有以下几个方面的启示, 这几个方面启示也是今后研究的重点方面。

第一, 教材编撰。现在有些教材的编写往往急功近利, 对其重视不够<sup>[9]</sup>。就教材文本的词汇覆盖率来说, 缺少量化分析, 凭主观估计选定文本及生词。而 LUWL 正好解决了这一问题, 在 RANGE 运行环境下, LUWL 可以帮助我们清楚地了解测试文本的单词覆盖率情况。正如刘润清所说“如何确定词频等级、选择词频标准, 以有利于教材开发, 是一项有意义的课题”<sup>[10]</sup>。

第二, 词汇教学。Nation 在词汇教学方面有深入系统的研究。他认为, 英语词汇可以分为两大类, 一类就是 2000 - 3000 个高频词汇, 另一类就是几千个低频词汇。这两类词汇属性上差异很大, 也就意味着需要采取不同的教学策略与方法<sup>[11]</sup>。我们需要细心地去学习和掌握高频词汇, 而对于低频词汇, 最好的教学方法就是在上下文中进行学习, 而不是单纯地教词汇<sup>[12]</sup>。LUWL 的构建可以快速高效地确定文本中的低频和高频词, 这也就提高了教学效率。

第三, 词汇测试。关于词汇量的测定有多种测定标准, 因为词形和词族等单位都可以作为测定标准。而 RANGE LUWL 可以一次性地计算出文本中的词形和词族数, 能比较方便快速地评估文本的词汇分布情况。我们也可以利用 LUWL 测定阅读某一文本需要掌握的单词量, 或者用 RANGE LUWL 对学生的词汇量进行调查与分析。

第四, 试卷评估。无论是学校课程学习考试, 还是英语专业和非专业过级考试, 或是其他各种公认的英语过级考试, 出题者都要考虑所选文本的词汇覆盖率或者说试卷难易程度。在探讨试卷文本词汇覆盖率时, 往往缺乏快捷、方便、高效、并且贴近我国英语教学实际的测试工具, 而本研究所制定的大学英语教学词表 (LUWL) 正好弥补了这一不足, 为英语教学和研究提供了方便高效的研究工具。

## [参考文献]

- [1] 钟志英等. 测试英语词汇的广度、深度和强度[J]. 华南农业大学学报, 2005(4): 131 - 136.
- [2] 黄建滨. 关于《大学英语教学大纲(修订本)》词汇表的说明[J]. 外语界, 1999(4): 27 - 29.
- [3] 邓昭春. 《大纲》词表局限性与分类词族的教学意义[J]. 中国大学教育, 2004(9): 45 - 48.
- [4] 孙文抗. 试论基础英语教学词表的编制原则和方法[J]. 课程·教材·教法, 2005(3): 61 - 65.
- [5] 韩冰, 冯奇. 论《大学英语教学大纲》词汇收录问题和原则[J]. 上海大学学报, 2004(5): 96 - 101.
- [6] Nation. Using small Corpora to Investigate Learner Needs [M]. Small Corpus Studies and ELT Theory and Practice. John Benjamins Publishing Company Philadelphia 2001.
- [7] Nation. Vocabulary Size, Text Coverage and Word Lists [M]. New York: Cambridge University Press. 1997.
- [8] 郑志恒. 报刊英语语料库概论 [M]. 南京: 南京大学出版社, 2009.
- [9] 冯辉, 张雪梅. 英语专业教材建设的回顾与分析[J]. 外语界, 2009(6): 63 - 69.
- [10] 刘润清. 关于英语教学大纲改革[J]. 外语教学与研究, 2002(6): 403 - 404.
- [11] Nation, I. S. P. 1983. Teaching and Learning Vocabulary [M]. Wellington: English Language Institute.
- [12] Xue Guo - yi, Nation, I. S. P. 1984. A University Word List [J]. Language Learning and Communication, 2003(2): 93 - 242.

(责任编辑: 范玉芳)