

# 面向汉语国际教育的介词知识库系统设计

郭 静

北京语言大学首都国际文化研究基地 100083

dreamerguo@126.com

**摘 要:** 语法教学是汉语国际教育的难点,在此方面充分借助计算机及网络是一种有效的选择。论文以辅助汉语语法教学为目标,选取虚词语法点中的介词为研究对象,在教材和传统词典数据库基础上研究介词知识库的功能结构和元数据体系,并进一步探索教学语法资源建设的方式方法。

**关键词:** 介词知识库 汉语国际教育 计算机辅助教学

## Design of Prepositional Knowledge Database System for International Chinese Education

Guo Jing

Capital International Culture Research Base, Beijing Language and Culture University, 100083

dreamerguo@126.com

**Abstract:** As grammar teaching is one of the difficulties in international Chinese education, the full use of computer and network will be an effective choice. Aiming to assist Chinese grammar teaching, this study selects prepositions from function words as representative samples, explores the functional structure and the metadata system of the prepositional knowledge database system based on teaching materials and traditional dictionaries, and further explores the resource construction of Chinese teaching grammar.

**Key words:** prepositional knowledge database, international Chinese education, computer aided education

汉语国际教育中存在许多教学上的难点,语法教学是其中比较突出的领域。对此已有不少学者进行了深入研究,虽然已经解决了不少问题,但仍然存在一些难以攻克的困难。利用计算机辅助汉语教学可成为处理这些难题的一个有效选择。本文选择语法中虚词部分的介词为研究对象,研究介词知识库的系统设计,以探索利用计算机辅助语法教学的一种方式方法。

### 1 介词知识库系统设计研究的必要性

#### 1.1 汉语教学对介词知识库的需求

丰富独特的虚词词类是现代汉语的一大特点。与印欧语系不同,汉语没有严格意义上的形态标志和屈折变化,而是主要借助于语序和虚词作为语法手段。虚词的意义大多比较“虚”,且功能丰富、用法复杂,一直以来都是外国学生学习汉语的重点和难点。陆俭明(2000)

特别强调了虚词的重要性：“在任何语言里虚词比实词少得多，但它在语言里起着‘经络’的作用，其重要性大大超过实词。在汉语里，虚词则占有更重要的位置。据不完全统计，外国学生所出现的语法错误，跟虚词相关的要超过 60%。”介词作为虚词的一类，对外国学生学习和掌握的难点可概括为：

(1) 从整体上看，介词虽然是一个封闭词类，但其作用并不容易一语概括，介词整体的分类问题也较为复杂。

(2) 从单个介词来看，各介词语法功能有不少差异，有的介词可以加“着”，有的不可以；有的既可以出现在主语前也可以出现在主语后的位置上，有的则只能出现在主语后；每个介词还可以构成各自的介词框架或固定搭配，用法各异。

(3) 从对比的角度来看，由于介词大多从古汉语动词虚化而来，基本没有实在意义，两个介词表达相似功能时，如“朝”和“冲”、“对”和“对于”等，想解释清楚它们之间的区别就比较困难。还有一些介词没有彻底虚化，如“在”、“给”主要用作介词，但仍保留动词的用法。从语法层面来看，一些易混淆词在句法分布上有时也会有差异，比如“关于……”做状语时只能用在主语前，“对于……”做状语用在主语前后均可；“关于……”还可以做定语，“对于”则不能。还要注重与之共现成分语法特征，如“按”和“按照”的选择与后面名词的音节多寡有关(苏英霞，2010)。

由此可见，对于外国学生来讲准确掌握并且正确使用介词还是非常困难的。目前，有关介词的用法信息、辨析对比等知识散见于各种研究文献和语法、词汇工具书中，且各有侧重，教师和学生需要翻阅大量资料以获得相关知识，无形中也加大了介词教与学的难度。

以上分析显示，介词整体分类的复杂性、单个介词语法功能的多样性、近义介词间的易混淆性等，都是可能成为介词难讲、难学以及偏误率高的原因。所以本文将利用现有的介词本体研究、介词教学研究成果并借助计算机技术，设计开发一个面向汉语国际教育的介词知识库，对介词知识以及相关语法知识进行形式化的归纳与整合，以期为介词的教学提供更好的资源。从应用价值上讲，本研究有助于介词教学质量的提高，有利于介词乃至虚词释义模式的探索，能够为教材、教案中语法点的编写提供便利的素材和资源，更重要的是为探讨计算机辅助语法教学的方式以及教学语法资源建设提供一种思路。

## 1.2 现有相关知识库应用中适用性的不足及本文的借鉴作用

目前已有一些对汉语语言方面知识库的研究，与本研究直接相关的是北京大学计算语言学研究所牵头研制的综合型语言知识库建设的研究，其中包括与郑州大学合作研制的面向机器的现代汉语虚词知识库。该库是一个包括了现代汉语虚词用法词典、虚词用法规则库以及虚词语料库的三位一体的语言知识库(俞士汶等，2011)。其中虚词用法词典利用关系数据库的形式详细描述了每个虚词用法的具体属性特征。该知识库主要是面向语言处理，并不考虑汉语教与学的需要，因此，本研究在该知识库成果的基础上，从汉语教与学的角度，提出了自己的数据库结构和数据表中的元数据元素，增加了相关的应用功能，如例句检索、辨析检索等。

## 2 介词知识库系统设计的理论基础及介词知识库系统总体结构

目前已有许多学者研究了汉语国际教育中的虚词教学原则、方法和策略,如认为应当注重虚词语义、句法和语用的结合,注重语境,注重虚词分阶段教学等。本文将这些研究成果进行了归纳总结,并据此提出介词知识库的系统总体结构。

### 2.1 介词知识库系统设计的理论基础

本研究拟从词汇—语法教学模式、多维释义与释义优选原则两个方面,探讨建立介词知识库的理论基础。

提出词汇—语法教学模式的代表人物是北京大学的李晓琪,主要观点包括三方面内容:

(1) 以虚词为核心,向外辐射到重要语法点,如介词可辐射到很多重要句式、常用词及由这些常用词组成的固定格式(语言板块),像“依我看”、“按照我的看法”等;(2) 虚词有明显的板块结构(词和短语、搭配词、惯用话语、句子框架和引语);(3) 强化虚词意识,建立有效的虚词学习途径(李晓琪,2004)。

多维释义与释义优选原则认为语义的解释很复杂,一方面要注意从多角度、多层面揭示语词的意义,并综合运用认知概念结构、语义结构、语义分解理论、原型理论和配价理论来进行多维释义的方法和原则(陈伟,2006);一方面针对对外汉语学习要注意词典释义的优选原则(李智初,2012)。

### 2.2 介词知识库系统总体结构

介词知识库系统应能告知每个介词是什么、在汉语语言中的结构是什么以及如何对其进行应用。根据这一目标,本研究设计的介词知识库系统主要包括三个层次:资源层、应用层和展示层,如图1所示。

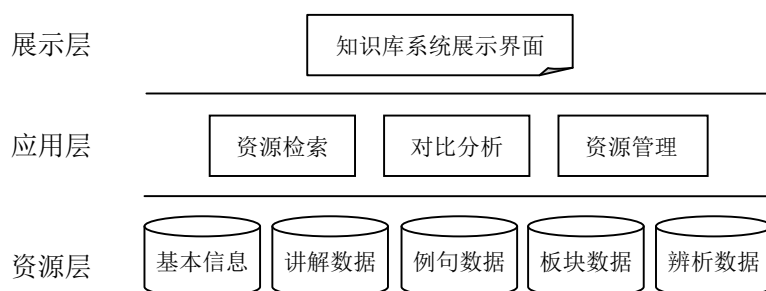


图1 介词知识库系统功能结构

基于关系型数据库的资源层是整个介词知识库系统的基石,负责对底层数据进行日常管理,包括资源的入库、删改和属性设置等。应用层是介于资源层与展示层之间的一个动态网页生成层,通过Web服务器对应用模块进行统一管理,是知识库系统的后台管理层。展示层是面向终端用户的可视化显示,用户通过展示层的Web浏览器与服务器相连,根据实际的教学需要进行相关的介词资源查询与下载等操作。

## 3 介词知识块系统资源层元数据体系设计

### 3.1 资源选择的来源与范围

#### (1) 讲解、固定格式与例句选取来源

本知识库系统主要是针对汉语教师备课、制作课件以及为新手教师补充介词知识、完善介词知识体系而用。讲解与例句对于教师备课以及帮助学生理解语言点非常关键。尤其对于介词这种意义不好把握的虚词来说,在注释给出的释义基础上,例句可以提供简单的语境和实际用法。选取的讲解与例句要通俗易懂,主要来源为现有教材。同时,介词的固定格式(介词框架与固定搭配等)是介词本身的一大特点,在教学中存在着遗漏、混用与杂糅等偏误问题,在知识库中应强化其作为一个固定结构的整体性。本研究借助于国际汉语教材智能标注系统,选取了近九年(2005—2013)、出版地为中国、适用课型为综合、阅读、写作、商务、语法的教材,从中筛选出课后具有明显词汇讲解、语法讲解模块的教材231本,经统计共有200本教材涉及介词知识的讲解。因此合计选取200本教材用于统计介词的讲解、固定格式与例句。

#### (2) 介词收词范围与基本信息选取来源

收词范围可依据汉办颁布的汉语水平考试词汇大纲,该大纲是我国汉语教学总体设计、教材编写、课堂教学以及成绩测试的重要依据,是规范性的大纲。大纲共收入介词56个,固定格式5个。通过考查《现代汉语词典(第五版)》、《现代汉语虚词手册》等工具书,可将这5个固定格式“除了……以外”、“连……都/也”、“到……为止”、“拿……来说”、“连……带”分别归入“除了”、“连”、“拿”、“到”4个介词中(关于“到”是否可视作为有介词用法,有待进一步讨论)。另外还收入汉语水平等级标准与语法等级大纲的丁级介词“论”,因此从两大纲中共收词62个。同时,通过对200本教材的讲解进行归纳统计,除了覆盖上述62个介词外,还存在11个大纲外介词“临”、“管”、“基于”、“依”、“依据”、“因”、“因为”、“出于”、“至于”、“挨”、“应”,故本系统共计收词73个。介词的基本信息主要参考《现代汉语语法信息词典》中的介词库,从中选取可应用于教学的字段。

#### (3) 辨析选取来源

介词与介词间的偏误主要发生在同类或近义类的介词中,加强介词的对比辨析越来越受到重视。有的工具书没有给出易混淆词或近义词内容,有的只是简单列出近义词而没有进行具体辨析,有的将辨析内容涵盖在了释义里;此外,工具书多因纸质限制,一组辨析内容只列在某一个易混淆词下,需要使用者多次翻阅。因此,构建介词辨析数据库,需要考虑对辨析词组的覆盖面、辨析分组的方式与标准、词典本身的权威性等问题。本研究具体考察了10部工具书的辨析内容,从中归纳出50组辨析词组,涉及56个介词。由于各位专家学者分组的方式和标准不尽相同,同时没有一部工具书涵盖了所有需要辨析的词,包含介词辨析词组最多的工具书包括的词组也仅有27个,多数包括的介词辨析词组都在20个以下,平均为15个词组。因此选择任何一部工具书作为辨析词组内容的来源,数据量都较少。同时,这些工具书也有大量内容交叉重复,从工作量和建库必要性考虑,也不适合选择全部工具书的相应内容。为此,本研究权衡了出版时间、辨析内容的风格与长度等因素,最终确定了可包含上述50组辨析内容的5部工具书用于建库。

### 3.2 资源层数据表元数据元素的设计

在介词知识库系统资源层包括的 5 个数据表中，介词基本信息数据表包含着该系统的主要信息，涉及介词的基本含义和用法，后四个数据表是对每个介词的含义和使用进行深度剖析和内容分解。元数据是描述数据的数据，在这里具体是指描述每个介词以及介词之间关系的数据。

(1) 介词基本信息数据表元数据元素的设计

根据介词本身的特性，介词基本内容数据表的元数据元素可划分为三大类，如下：

① 揭示介词是什么的元素，包括介词标识符、词形、拼音、级别、表义类型。其中介词标识符是每个介词的唯一标识，主要用来和其他数据表中的元素建立关联；级别元素的设置是考虑到汉语国际教育教师教学时的顺序与侧重。表义类型元素是展示介词的语义功能，可分为表时间、空间，关涉对象，原因、目的，凭借、依据，施事、受事，排除等六类表义类型。

② 揭示介词的一些规律性使用方法的元素，包括多用于口语还是书面语，可否加“着”，否定词在介词前还是介词后，与主语的位置关系，可否带时间词宾语、处所词宾语、方位词宾语等。介词在使用上有很多的共性，但也一些差异，差异往往是偏误集中、不易讲清的部分。如图 2 所示：

介词标识符	词形	拼音	级别	表义类型	口语/书面语	加“着”	否定词位置	与主语位置关系	能否做补语	名词宾语	时间词宾语	处所词宾语	方位词宾语	介词短语单说	易混淆词
an	按	àn	乙级	引出凭借、依据	口语/书面语	能	前	兼	否	能	否	否	否	能	照
anzhao	按照	ànzhào	乙级	引出凭借、依据	口语/书面语	否	前	兼	否	能	否	否	否	能	按、依照、根据
ba	把	bǎ	甲级	引出对象	口语/书面语	否	前	后	否	能	能	能	能	否	被、将、拿
bei	被	bèi	甲级	引出对象	书面语	否	前	后	否	能	否	否	否	否	把、由、叫、让
benzhe	本着	bènzhe	丁级	引出凭借、依据	书面语	否	兼	兼	否	否	否	否	否	否	
bi	比	bǐ	甲级	引出对象	口语/书面语	否	兼	后	否	能	能	能	能	否	
chao	朝	cháo	甲级	引出方向/引出对象	口语/书面语	能	前	后	否	能	否	能	能	能	往、向、对、顺、沿
chen	趁	chèn	乙级	引出凭借、依据	口语	能	前	兼	否	能	能	否	否	否	乘
chong	冲	chòng	丙级	引出对象	口语	能	前	兼	否	能	否	能	能	否	向
chu	除	chú	丙级	引出对象	书面语	否	后	兼	否	能	能	能	能	否	除了

图 2 基本信息数据表

例如：

介词一般都可以在口语和书面语中使用，但有些语体色彩很明显的介词需要在口语/书面语元素中标识出来，如表示被动意义的“叫”、表示“从”的北方方言色彩介词“打”等。

同是表示引出方向，介词“向”、“朝”、“沿”可以有“向着”、“朝着”、“沿着”的用法，而有的则不能，如“\*往着”。

“从”和“自从”都可表示时间的起点，但“自从”的宾语则只能是时间词，“从”的宾语还可以是处所词和方位词。

有的介词结构可以单说，如“她昨天是跟谁出去玩儿？”“跟她舍友。”中的“跟她舍友”可以单独作为答句。有的则不能单说，如“昨天大爷把谁骂了一顿？”“\*把他。”

与主语的位置关系和能否做补语反映了介词与其他成分在语言中的位置结构关系。同为引出对象的“关于”和“对于”，“关于”做状语时必须放在主语前，“对于”则在主语前后都可。几乎所有介词结构都可作状语，但只有个别介词结构能作补语（通过介词与动词的位置关系判断），如“于”、“自”，这种位置的分布已经定型。

③ 揭示介词之间关系以深度解析介词含义和使用方法的元素，主要包括易混淆词。很

多介词在使用中容易与其他介词相混，如“按、按照”、“从、自、自从”。本研究以现有的词典、工具书为基础，结合各种研究文献，提取出容易导致汉语学习者困惑的易混淆词组。易混淆词字段主要用来与辨析数据表中的辨析词组建立关联。

## (2) 其他四个数据表元数据元素的设计

讲解数据表中的元数据元素主要包括讲解 ID、介词标识符、词形、书 ID、课 ID 和讲解内容。讲解内容收录自 200 本教材中涉及到介词知识的部分（教材课文后的词语例解、语法讲解、注释等），每本教材中的每一个介词讲解作为一条唯一记录，共计 1420 条。如图 3 所示：

讲解ID	介词标识符集	词形	书ID	课ID	讲解内容
83	ti,xiang_k	替、向	5478	2742	替我向他/她问好
84	weile	为了	5424	2745	为了 介词，后接名词或动词短语，表示动作或行为的目的，“为”后也可不用“了”。例如： （1）为了看病，他跑了很多次医院。 （2）为了解决这个问题，我查了大量资料。
85	anzhao	按照	5424	2753	按照（介词） 根据，依照。与后边的名词或名词性词组构成介宾结构，用在动词前或主语前作状语，引出后边的陈述或动作行为的依据。例如： （1）我们应该按照实际情况考虑应该用什么办法。 （2）按照规定，他应该向老师请假。 （3）按照李教授的看法，中国目前的经济发展不是太慢了，而是太快了。
86	bi	比	5451	2776	数量补语 ①在用“比”表示比较的形容词谓语句中，如果要表示两个事物的具体差别，就在谓语后边加上数量词作补语。例如： （1）他比我大两岁。（2）他家比我家多两口人。 ②要表示大略的差别程度，可以用“一点儿”、“一些”、“多了”或用“得”加程度补语“多”。例如： （1）他比我大一点儿（一些）。（2）那儿比这儿冷多了。

图 3 教材讲解数据表

例句数据表中的元数据元素主要包括例句 ID、讲解 ID、介词标识符、词形和例句。例句元素是从讲解数据表中的讲解内容中摘取出来的，每条例句作为一个记录。如图 4 所示：

例句ID	讲解ID	介词标识符	词形	讲解内容
1948	452	rang	让	孩子让人打了。
1949	883	rang	让	我的车让汤姆开走了。
1950	53	rang	顺	小明住在我家附近，我回家时顺路去看小明。
1951	53	rang	顺	我顺着河边走。
1952	53	rang	顺	眼泪顺着脸颊流下来。
1953	638	shun	顺	我和丈夫顺着他手指的方向看过去……
1954	638	shun	顺	昨天晚上，我们俩顺着河边一边走一边聊，谈得特别开心。
1955	638	shun	顺	你顺着这条路一直往东，过两个红绿灯就到了。
1956	638	shun	顺	外边的雨很大，雨水顺着窗户往下流。

图 4 例句数据表

介词板块结构数据表中的元数据元素主要包括结构 ID、板块结构、介词标识符、讲解 ID 和例句 ID。介词可辐射到的语言点非常丰富，如介词框架“从……到……”、“除了……以外”等，又如与动词形成的固定搭配“来自”、“不亚于”等。根据词汇一语法教学模式，本研究将介词的框架结构、介词与搭配词形成的固定格式等统称为板块结构。板块结构的内容主要来源于对讲解内容的统计分析。通过介词标识符字段的关联，可建立基本信息数据表和板块结构数据表的 1 : n 的关系。如图 5 所示：

板块ID	板块结构	介词标识符	讲解ID集	例句ID集
1	从……到……	cong	132, 135, 179, 188, 206, 220, 234, 260, 267, 623, 729, 818, 872, 888, 963, 996, 999, 1072, 1169, 1181, 1205, 1207, 1246, 1287, 1442, 1471, 1557	1143, 1144, 1145, 1146, 1147, 1148, 1149, 1150, 1151, 1155, 1156, 1157, 1158, 1159, 1160, 1161, 1162, 1163, 1164, 1165, 1166, 1167, 1168, 1169, 1182, 1185, 1199, 1200, 1202, 1203, 1204, 1205
2	从……起	cong	107, 217, 326, 399, 793, 889, 1112, 1208, 1233, 1261, 1414	1131, 1133, 1152, 1153, 1154, 1181, 1184
3	从……做起	cong	4, 1186	1134
4	从……来	cong	306	1174, 1175, 1176

图5 板块结构数据表

辨析数据表中设置的元数据元素主要包括辨析 ID、辨析词组、词典 ID、辨析内容和关联辨析词组。根据之前所确定的 50 组辨析词组在 5 部工具书中进行统计，同时利用《现代汉语词典》和教材作为对工具书涵盖内容不全之处的补充，共计有 82 条辨析记录。由于每组辨析至多有 5 条记录而已，本系统不再对辨析内容进一步整合，显示以及检索结果会默认按照辨析 ID 顺序呈现出来。另外，由于出自五部不同的词典，在辨析词组构建方面有内容的上交叉，因此需要建立介词辨析词组之间的关联，即关联辨析词组元素，增强“启发”学习和“联想”学习的作用。

辨析 ID <sup>①</sup>	辨析词组 <sup>②</sup>	词典 ID <sup>③</sup>	辨析内容 <sup>④</sup>	关联辨析词组 <sup>⑤</sup>																																				
zi-cong-1 <sup>⑥</sup>	从/自 <sup>⑦</sup>	CDC <sup>⑧</sup>	<p><b>[辨析] 自 从<sup>⑨</sup></b></p> <p>在表示时间或处所的起点这个意思时，这两个词的意思和用法基本相同。不同的是：<sup>⑩</sup></p> <ol style="list-style-type: none"> <li>“自”只用于书面语，“从”书面语口语都常用。<sup>⑪</sup></li> <li>“自……”可以用于动词后，“从”不能这样用。如： 他来自韩国/他从韩国。<sup>⑫</sup></li> <li>“自”没有“从”的其他用法。<sup>⑬</sup></li> </ol> <p><b>[提示]</b>“自”还是副词，是“自然、当然”的意思，用于书面语。如：<sup>⑭</sup></p> <p>老朋友见面，自有许多话要说。<sup>⑮</sup></p> <p>他这样做自有他的道理。<sup>⑯</sup></p>	从/自从；从/自/自从；自/自从 <sup>⑰</sup>																																				
zi-cong-2 <sup>⑱</sup>	从/自 <sup>⑲</sup>	CDD <sup>⑳</sup>	<p><b>自[介] 从[介]</b></p> <p>○<b>词义说明 Definition<sup>㉑</sup></b></p> <p>自[from; since]从；由。<sup>㉒</sup></p> <p>从[from (a time, a place or a point of view)]起于；“从……”表示“拿……做起点”。[via, through, or past (a place)]表示经过，用在表示处所词的前面。<sup>㉓</sup></p> <p>○<b>词语搭配 Collocation<sup>㉔</sup></b></p> <table border="1"> <tr> <td></td> <td>~小<sup>㉕</sup></td> <td>~古<sup>㉖</sup></td> <td>~远而<sup>㉗</sup></td> <td>来~<sup>㉘</sup></td> <td>~小到<sup>㉙</sup></td> <td>~北京<sup>㉚</sup></td> <td>~现在<sup>㉛</sup></td> <td>~桥上<sup>㉜</sup></td> </tr> <tr> <td></td> <td></td> <td></td> <td>进<sup>㉝</sup></td> <td>大<sup>㉞</sup></td> <td>出发<sup>㉟</sup></td> <td>起<sup>㊱</sup></td> <td>过<sup>㊲</sup></td> <td></td> </tr> <tr> <td>自<sup>㊳</sup></td> <td>./<sup>㊴</sup></td> <td>./<sup>㊵</sup></td> <td>./<sup>㊶</sup></td> <td>./<sup>㊷</sup></td> <td>./<sup>㊸</sup></td> <td>./<sup>㊹</sup></td> <td>./<sup>㊺</sup></td> <td>./<sup>㊻</sup></td> </tr> <tr> <td>从<sup>㊼</sup></td> <td>./<sup>㊽</sup></td> <td>./<sup>㊾</sup></td> <td>./<sup>㊿</sup></td> <td>×<sup>㋀</sup></td> <td>./<sup>㋁</sup></td> <td>./<sup>㋂</sup></td> <td>./<sup>㋃</sup></td> <td>./<sup>㋄</sup></td> </tr> </table> <p>○<b>用法对比 Usage<sup>㋅</sup></b></p> <p><b>用法解释 Comparison<sup>㋆</sup></b></p> <p>“自”和“从”是同义词，“自”多用于书面，“从”常用于口语。<sup>㋇</sup></p>		~小 <sup>㉕</sup>	~古 <sup>㉖</sup>	~远而 <sup>㉗</sup>	来~ <sup>㉘</sup>	~小到 <sup>㉙</sup>	~北京 <sup>㉚</sup>	~现在 <sup>㉛</sup>	~桥上 <sup>㉜</sup>				进 <sup>㉝</sup>	大 <sup>㉞</sup>	出发 <sup>㉟</sup>	起 <sup>㊱</sup>	过 <sup>㊲</sup>		自 <sup>㊳</sup>	./ <sup>㊴</sup>	./ <sup>㊵</sup>	./ <sup>㊶</sup>	./ <sup>㊷</sup>	./ <sup>㊸</sup>	./ <sup>㊹</sup>	./ <sup>㊺</sup>	./ <sup>㊻</sup>	从 <sup>㊼</sup>	./ <sup>㊽</sup>	./ <sup>㊾</sup>	./ <sup>㊿</sup>	× <sup>㋀</sup>	./ <sup>㋁</sup>	./ <sup>㋂</sup>	./ <sup>㋃</sup>	./ <sup>㋄</sup>	从/自从；从/自/自从；自/自从 <sup>㋈</sup>
	~小 <sup>㉕</sup>	~古 <sup>㉖</sup>	~远而 <sup>㉗</sup>	来~ <sup>㉘</sup>	~小到 <sup>㉙</sup>	~北京 <sup>㉚</sup>	~现在 <sup>㉛</sup>	~桥上 <sup>㉜</sup>																																
			进 <sup>㉝</sup>	大 <sup>㉞</sup>	出发 <sup>㉟</sup>	起 <sup>㊱</sup>	过 <sup>㊲</sup>																																	
自 <sup>㊳</sup>	./ <sup>㊴</sup>	./ <sup>㊵</sup>	./ <sup>㊶</sup>	./ <sup>㊷</sup>	./ <sup>㊸</sup>	./ <sup>㊹</sup>	./ <sup>㊺</sup>	./ <sup>㊻</sup>																																
从 <sup>㊼</sup>	./ <sup>㊽</sup>	./ <sup>㊾</sup>	./ <sup>㊿</sup>	× <sup>㋀</sup>	./ <sup>㋁</sup>	./ <sup>㋂</sup>	./ <sup>㋃</sup>	./ <sup>㋄</sup>																																

图6 辨析数据表

## 4 介词知识库系统应用层的设计

应用层主要包括资源检索模块、对比分析模块以及资源管理模块，其中资源管理模块是支持系统管理员对系统的建设和维护，前两个模块是服务于用户。以下主要对这两个模块进行说明。

### 4.1 介词资源查询与检索模块

资源检索模块主要包括介词内容检索、例句检索和辨析检索等。

介词内容检索有两种检索方式：检索词查询和浏览查询（也称导航式）。检索词查询方

式是最基本的查询方式,检索词包括基本信息数据表中的所有介词,检索结果包括该介词的基本信息、注释列表、板块结构、易混淆词和教材例句等。浏览查询方式包括音序导航模式和分类导航模式。音序导航模式提供了拼音查询的途径,分类导航查询模式则是从介词语义类型角度查询。例句检索提供检索词查询方式,检索结果为对应介词在教材中的例句。辨析检索同样含有检索词查询和浏览查询两种方式,浏览查询方式提供了辨析分组导航。借助计算机技术,辨析检索的设计可以较为有效地解决纸质工具书辨析的不便之处,以更为清晰、全面的方式展现易混淆介词间的异同点。

#### 4.2 介词对比分析模块

区别于辨析检索模块,介词对比分析模块提供了一个可供用户自由选择所需对比介词的平台。可以通过检索词方式添加所需对比的多个介词,也可以通过二级联动菜单(一级菜单为介词表义类型)来选择添加,如图7示意图。平台将列出所添加介词的基本信息数据和例句数据,同一元素在横向进行对比,用法的相同点与不同点展示更加直观、一目了然。

介词信息	选择一个介词	选择一个介词	选择一个介词
	选择表义类型	选择表义类型	选择表义类型
	选择介词	选择介词	选择介词
显示设置: <input type="checkbox"/> 显示不同项			
<b>基本信息</b>			
拼音			
级别			
口语/书面语			
能否加“着”			
否定词的位置			

图7 对比分析模块示意图

## 5 结语

上述内容对面向汉语国际教育的介词知识库系统设计进行了探索和研究,并已展开了部分数据的统计分析。在接下来具体构建该系统时,还需要进一步对知识库中各数据表的数据(包括图2—图6所列数据)进行审核校对和深入研究,如拟定筛选原则并对例句库中的例句进行筛选,为讲解数据表中的讲解添加标签以便于检索,确定介词的板块结构和辨析词组等,并逐步完善该知识库系统的各项功能。

## 参考文献

- [1] 陈伟. “多维释义”理论学理探析[J]. 广东外语外贸大学学报, 2006(1):69-73
- [2] 李晓琪. 关于建立词汇—语法教学模式的思考[J]. 语言教学与研究, 2004(1):23-29
- [3] 李智初. 对外汉语学习型词典释义的优化[J]. 辞书研究, 2012(6):37-45
- [4] 陆俭明. “对外汉语教学”中的语法教学[J]. 语言教学与研究, 2000(3):1-8
- [5] 苏英霞. 汉语学习者易混淆虚词的辨析视角[J]. 汉语学习, 2010(2):91-96
- [6] 俞士汶, 穗志方, 朱学锋. 综合型语言知识库及其前景[J]. 中文信息学报, 2011(6):12-20