

基于句本位标注的单一教材语法点分析

朱 雪

北京师范大学 信息科学与技术学院 100875

zx_514@163.com

摘 要: 对外汉语教材是对外汉语教学过程中的重要依据, 而对外汉语教材对语法结构历来重视, 因此对对外汉语教材语法方面的编写情况进行分析研究具有极其重要的意义。黎锦熙先生提出的句本位语法体系具有格局观, 对于分析对外汉语教材语法具有独特的优势。本文将对外汉语教材《长城汉语·生存交际》(以下简称《长城汉语》) 中的语法项目用句本位语法体系形式加以呈现, 进而用计算机技术进行语法点统计分析。

关键词: 对外汉语教材 句本位标注 XPath 检索 语法分析

The Analysis of the Grammar Points of Single Material Based on “Sentence-based” Annotation

Zhu Xue

School of Information Science and Technology, Beijing Normal University, 100875

zx_514@163.com

Abstract: Foreign language teaching materials are an important basis in the teaching process. It is very important and meaningful to analyze preparation of the grammars, for foreign language teaching has attached great importance to the grammatical structure analysis. Sentence-based grammatical system created by Jinxi Li has pattern view, and has unique advantage to analyze grammars. In this paper, the grammars of Great Wall Chinese Essentials in Communication will be presented with the form of sentence-based grammatical system.

Key words: Foreign language teaching material, “Sentence-based” Grammatical annotation, XPath, Retrieving, Grammar points analysis

0 引言

随着国家改革开放日新月异, 汉语逐步成为一种国际化的语言, 对外汉语教学作为一门学科也蓬勃发展起来。伴随着对外汉语教学事业的发展, 各种各样的对外汉语教材在市场上大量涌现, 虽然大多数已经基本满足教学需要的体系, 但却参差不齐, 因此需要有一个对教材各方面进行评价的体系, 使得学习者根据自身需要选择适合自己的教材, 从而进行有效的汉语学习。

目前句法分析中采用的语法体系主要分成以成分为基础的短语结构语法和以词汇为基础的依存语法两大阵营^[19]。两大语法体系用于句型或句式的结构统计时仍有不足之处, 具体体现在两方面: 一是短语结构语法在句法方面太过强调短语结构层次, 缺乏对句子格局的考虑; 二是依存语法在词法方面依据词语间的依存关系工作, 仅强调词与词之间的相互

关系而缺乏词性信息和短语类等句法语义范畴。但是语法教学中语法点的设置很多都和句型或者句式有关,所以需要继续寻找一种适合对外汉语教材语法点分析的语法体系,而句本位语法体系最具特色的地方就是其具有格局观,所以本文试图引入句本位语法体系,使对外汉语语法分析达到更好的效果。

1 语法点分析研究现状

语法点的分析包括两种情况:显性语法点分析和隐性语法点分析。显性语法点是指在课后明确呈现给学习者的语法点,隐性语法点指那些在课后没有明确说明,而存在于正文句子中的需要通过其他方法以将其识别出的语法点。不管是哪种语法点,目前大部分学者采用的方式都为人工分析。两者之中显性语法点的分析比较简单,只需要将课后呈现出来的语法点进行统计分析即可。相比之下隐性语法点由于其出现位置和方式都不明显,所以对这部分语法点分析会比较麻烦。

通过人工方式进行语法点分析,弊端之一是因为规范标准不同所以术语不完全统一,导致语法点统计难度增加。如我们平时所说的兼语句,《王力文集》中将其称为递系式。弊端之二是隐性语法点的统计需要人工逐句统计,需要耗费大量的人力和物力,而且最终分析的结果也并非完全准确。在语料规模非常大的情况下,人工进行语法分析的方式并不现实。

近几十年来,对外汉语教学界的跨学科研究意识越来越强烈,其中与信息科学和现代教育技术的结合受到重视。许多语言学家在对外汉语教学语法点研究方面都做过大量研究工作,他们的理论成果都非常值得借鉴,但是大多只停留在理论研究层面。所以本文尝试运用信息技术手段进行语法分析,旨在减少人工工作量的基础上尽可能保证分析结果的正确性,期望能在之前研究的基础上结合信息技术手段对语法项目进行研究,最终通过量化显示的结果总结适合对外汉语教学的语法项目。

2 教材语法树库构建

2.1 对外汉语教学语料库的选材

对外汉语教材建设已经成为对外汉语教学学科建设和发展的重要标志。近几年我国对外汉语教学事业空前繁荣,伴随发生的是教材的数量以异乎寻常的速度增长,教材的种类不断丰富。据中山大学国际汉语教材基地信息收集情况,截止到2011年3月,共收集含45种媒介语的汉语及文化教材信息15000余册,这些资源的类型包括课堂教材、读物和自学教学、工具书等,并且部分教材标注了学习者的使用水平(零起点和初级、中级、高级等)。虽然教材的数量和种类都较之前有很大发展,但质量却参差不齐,普遍受到国内外师生欢迎教材种类的不多。其原因是现有的教材评价标准模棱两可,缺少一个科学评价教材优劣的大众标准。如果有一个能解决此问题的知识库,将评价标准完善化,进而指导各类教材的编写,那么现在的对外汉语教学发展便会更进一步,教材编写也会突破之前固定的模式,呈现百花齐放的局面。

本文构建对外汉语教材语法树库所采用的生语料来自一部比较具有代表性的现代对外汉语教材:《长城汉语》。《长城汉语》是国家汉办/孔子学院总部规划、组织、研发、运营的重点项目,定位为基于网络多媒体技术的新型对外汉语教学模式。长城汉语总体规划中,将汉语学习分为18个等级,其中1—6级为生存交际级别(初级),7—12级为拓展交际级别(中级),13—18级为自由交际级别(高级)。已经完成的《长城汉语》(生存交际)包括纸板教材和多媒体两个部分,纸板教材由马箭飞主编,包括6册课本和6册练习册,于2005年由北京语言大学出版社出版,该套丛书以培养学习者的汉语交际能力为主要目标,

适用于初级水平学习者。6 册课本中, 每册课本包括 10 个单元, 每个单元包括 3 篇课文, 合计 180 课内容, 所有课文均以对话体形式呈现, 累计有 4382 个句子。由于这部教材在课后专门进行了语法的讲解, 而且每一册的语法讲解都达到了一定数量, 加之发行量大, 所以本文首先以这部教材作为语法分析的代表。

2.3 语料标注

2.3.1 语料预处理

本实验室之前有过一套语料预处理的工具, 可以在这套工具的基础上将对外汉语教材语料进行处理, 包括相同步骤的语料的录入、切句和 XML 格式初始化等步骤。语料的录入由 OCR (Optical Character Recognition, 光学字符识别) 软件将扫描好的 PDF 文档快速转变为可编辑的电子文档形式; 切句是按文本中句末的标点符号 (“。”、“!”、“?”、“;”、“:”) 由 perl 语言编写的切句工具来实现; XML 格式初始化是将切句之后的文本初始化存储为 XML 格式的文本, 便于之后计算机进行处理, 也由 perl 语言编写工具实现。

2.3.2 语料标注规范说明

语料标注过程是后期语法分析阶段实施的基础过程, 语法分析结果的正确性依赖于标注的规范性。如果语料标注这一步出现大量错误, 那么以后语法分析量化显示的结果一定会和实际情况有巨大偏差, 所以标注阶段非常重要。

标注过程仍然加入了人工审核的步骤, 标注者和审核者可以相互转换角色, 进一步保证标注的准确性。遇到意见不一致的情况会从理论和实际方面加以考虑, 尽量选择利于语法分析的一项, 以期分析结果更加科学和完善。

3 检索工具设计

3.1 检索工具设计

3.1.1 工具查询语言——XPath

根据上一章分析, 由于所有语料的存储形式为 XML 格式, 所以需要寻找一种在 XML 文档中检索符合特定结构的查询语言。经过比较分析, 本文决定选取 XPath 作为检索 XML 路径的语言工具。

XPath 为 XML 路径语言 (XML Path Language), 可用于在 XML 文档中查找某一部分信息。XPath 作为一种小型查询语言, 可以使用路径表达式选取 XML 结构中节点 (包括元素和属性) 或节点集。结合本文中所使用 XPath 路径表达式, 现举例如下:

- /ju/dj/zy/m

在 XML 结构树中, 从根节点开始依次向下查找子元素 ju 之下的子元素 dj 之下的子元素 zy 之下的子元素 m。此表达式用于查找单句结构的句子中充当主语的名词元素, 而不是代名词等其他元素。

- //wy[zh/f='也'] 或者 //wy/zh[f='也']

在每一步上设置断言对选择的节点集进行筛选, 断言的内容是要查找的限定的具体内容, 这个表达式用于查找任意位置的 wy 元素下状语为副词“也”字的句子。断言内部可以嵌套使用, 用 and、or 逻辑运算符。如不限定元素的具体取值, 可以写为 //wy[zh/f] 或者 //wy/zh[f]。

- //by[@bs='bzjs']

@bs='bzjs' 指任意位置处带属性为 bs, 并且属性值为“bzjs”, 此表达式用于选择变式宾语形式为“宾至句首”的句子。

另外，XPath 表达式还可以使用函数表达，常用的函数有 count() 返回节点数，position() 返回节点在父节点中的索引号，name() 返回当前节点的名称。

3.1.2 XML 文档

与句本位语法体系相应的是黎锦熙先生创立的教学工具“图解法”，“因为汉语是各词孤立的‘分析语’，主要是依靠词的位次来表达意思，这语序一经图解，就把组织规律明白清楚地摆在眼前。”^[8]运用句本位语法体系理论图解析句，图解之后的句子以 XML 格式进行存储，依赖实验室之前开发的可视化标注工具，在此工具中存储的 XML 文档与图解后的图形完全映射。图表 1 是一个 XML 文档的基本框架：

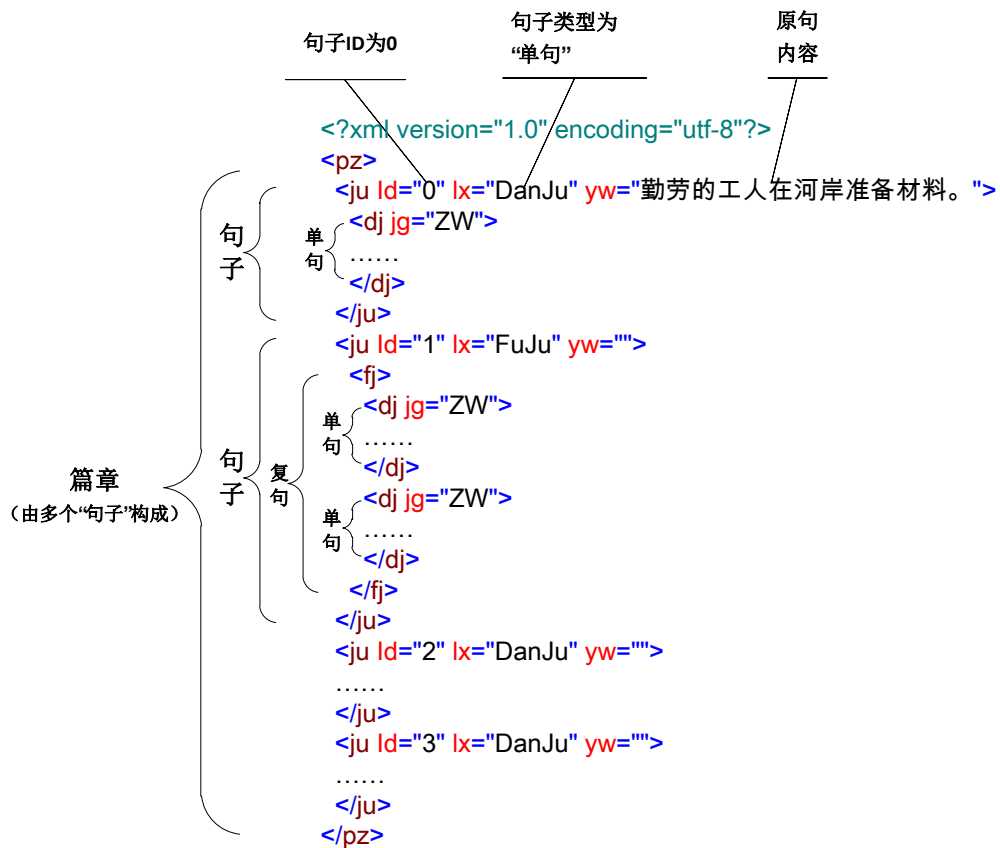


图 1 XML 文档基本框架

在 XML 文档中起始标签（如<pz>）必对应一个结束标签（</pz>），若标签需要嵌套，则必须按照合适的顺序进行，结束标签也需要按照镜像顺序来与起始标签匹配。本文使用的语料所存储的一个 XML 文件，是对篇章进行切句之后的多个句子（ju）构成，句子类型（lx）为单句（dj）或者复句（fj），每一个复句则至少由两个单句构成。

3.1.3 检索工具基本图解功能

本检索工具是在之前实验室开发的可视化标注工具的基础上进一步开发而成。此工具设计了四种检索接口，分别为基于原文的检索、基于词类的检索、基于短语的检索和基于句式的检索。此检索工具所做的具体工作是，对字符串进行直接匹配或者用正则表达式进行匹配。后面三种检索需要借助 XPath 路径查询功能，需要在查询界面输入路径表达式，或者借助可视化图解控件自动生成表达式，本论文运用的主要方式为基于句式的检索。

基于句式的检索由基准路径、包含条件和排除条件三组路径表达式来完成，但句式检索较其他检索的路径表达式的生成规则更为复杂，需要判断句子类型是否为复句，或者句子存在语序变式等方面的信息。

句式检索的图形界面默认为一个单句控件，基准 XPath 设为 “(/ju|/ju/fj)/dj/wy”，因为单句 (dj) 结构中可能存在多个谓语 (wy)，在句本位语法体系中定义为复谓语，所以如果用 dj 作基准 XPath，有可能会出误检。因为 wy 节点是句式的核心，所以本系统将基准路径定在谓语一级，以方便大多数句式结构的检索。例如检索句型“主语+是+名词”：

<input type="checkbox"/> R	基准路径	/ju /ju/fj)/dj/wy
<input type="checkbox"/> V	包含条件	preceding-sibling::zy self::*[@jg='WeiBz' and d='是'] bz[m]
<input type="checkbox"/> X	排除条件	

图2 “主语+是+名词”检索

但是整个基准路径也不能保证所有句式的检索，因为有的句子虽然句型是单句，但是里面包含两个谓语，比如单句“我认为这是对的”。如果检索语法句式：“‘是’字句”，仅仅将基准定为 “(/ju|/ju/fj)/dj/wy” 是不够的，还需要考虑多加限定条件。如图表 3：

<input type="checkbox"/> R	基准路径	/ju /ju/fj)/dj/wy (//mz //xz //fz)/wy
<input type="checkbox"/> V	包含条件	preceding-sibling::zy self::*[d='是']
<input type="checkbox"/> X	排除条件	

图3 基准路径限定条件

这种情况只有将基准路径定为 “(/ju|/ju/fj)/dj/wy|(//mz|//xz|//fz)/wy” 时，才能保证检索的完全性。

本系统在设计时也考虑了复杂的联合检索情况，如增添一项功能，可以对上一次的检索结果进行二次检索。见图 4 从结果中检索”复选框。

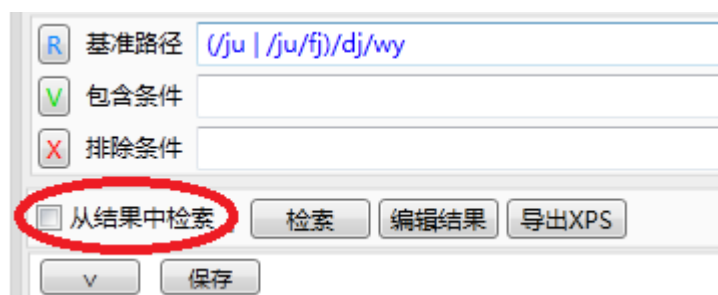


图4 “从结果中检索”位置示意图

检索中分别设置了一个基准路径和两组路径条件（“包含条件”和“排除条件”）。基准路径负责选出一组节点集，然后根据需要要用这两组路径条件对其进行过滤工作。进行检索的具体过程中，“包含条件”中所有项必须完全匹配，并且“排除条件”中的任一项必须为空，这样才能正确匹配，找出符合要求的句子。

设置这两个条件是为了更好地满足用户需要，并且保证检索结果的正确性。例如：在以谓语为中心的短语中，用户可以在检索界面（述说位控件）上添加一个副附，然后点击“包含条件”前面的“√”按钮，则系统自动生成所需的条件表达式，如图表2所示。但是用户还欲限制此中心词不能再有副附修饰，这时候可在控件上再添加一个副附，然后点击“排除条件”前面的“×”按钮，则系统自动生成所需排除条件表达式。因为此检索系统目前还未完善开发，仍然存在自动生成条件表达式的层级限制问题，如二级副附不会自动生成，需要人工手动输入，所以遇到此类问题还需要手工编辑。

3.2 图解与 XML 编解码关系

WPF 界面显示应该和 XML 结构设计建立相互对应的逻辑映射关系，程序能够将 WPF 界面上的图形根据这种映射关系转化为 XML 数据储存到数据库中，同时可以读取 XML 数据然后正确地将其转化为 WPF 界面图形。可以理解为 WPF 图形和 XML 数据是一种编解码的关系，语料标注人员可以通过读取 XML 数据的信息在图形界面上进行图解修改，并将其再次保存为 XML 格式的数据。

4 《长城汉语》语法点分析

4.1 语法分析展示平台

4.1.1 设计思路

信息化时代运用计算机技术进行对外汉语语法分析将是一种趋势，这不仅能减少人工处理的工作量，还能降低人工分析结果的错误率。语法分析展示平台用于直观显示检索工具进行检索的结果，这将为以后语法分析方法打开一种新思路。在分析过程中，分析者只需在句本位语法体系的标注规范下，找到对应的语法句式，根据语法点附加的其他属性信息（如语法点序号、章节号等），便能在课文中进行检索。检索工具作用示意图如下：

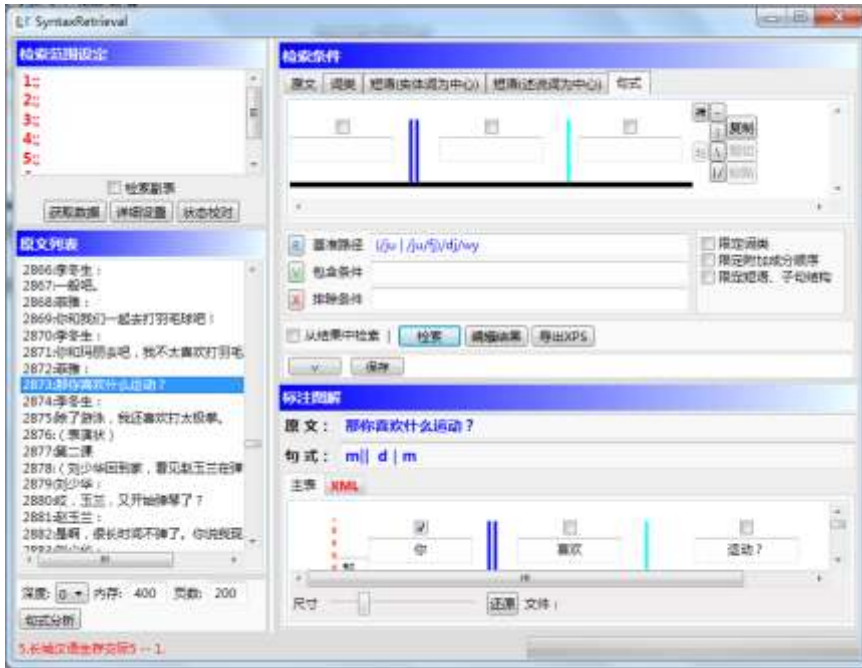


图5 检索工具界面

4.1.2 平台主界面

语法分析展示平台图形界面如下图:



图6 平台分析主界面

图形界面的左侧为所有长城汉语课后的语法点列表,按照讲解顺序从大到小依次排序,右侧为各语法点附加的各种属性,包括语法点 XPath、语法点对应例句、语法点出现章节信息和最终检索句子情况。

为方便用户使用，左侧语法点列表进行了颜色设置，以区分语法点是否正常检索。下表给出了各种颜色表示的相关属性（“√”代表具有属性，“×”代表不具有属性）：

表 1 各种颜色相关属性表

属性 颜色	出现语法点的单元可检索出匹配句子	其他单元课检索出匹配句子	语法点存在 XPath
蓝色	√	√	√
黄色	×	√	√
红色	×	×	√
灰色	×	×	×

颜色的显示只是对检索结果的一种视觉比较，对于语法点不存在的属性并不说明语法点在检索时一定出现了错误。红色显示的语法点在有 XPath 条件下未能检索出句子，是因为课后语法点讲解时虽然是以一个语法点的形式出现，但是提到了两个意思，而课文中出现的句子只符合其中的一种形式，对另一种形式的句子未呈现，比如，《长城汉语》第 3 册第 8 单元语法点：(是+) 时间+动词短语+的，在句本位语法体系中，句尾的“的”，可作为两种情况出现，作语气词和不作语气词，所以必须分情况讨论。灰色显示的语法点不能写出 XPath，这是与句本位语法体系不能对应的情况，目前还无法实现这些语法点的 XPath 检索。

考虑到用户需求，检索出的每一个句子的信息会在最下方的句子列表显示，并显示句子所在单元的所有课文信息。如下图：

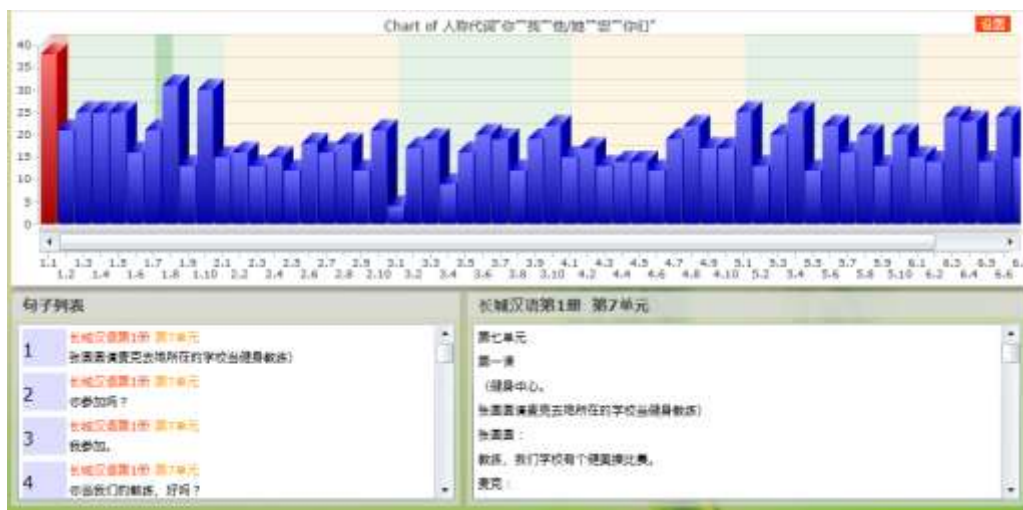


图 7 句子信息显示界面

4.2 语法点分析

本文在分析《长城汉语》的语法点前，首先对其中一些语法点进行了调整。如第一册

第二单元中的语法点“‘什么’的特殊问题”，结合讲解时给出的例句，可以将此语法点分成两种情况：(1)“主语+动词+什么+名词”（“什么”为疑问形容词），(2)“主语+动词+什么”（“什么”为疑问代名词）。这样调整的目的，一是与句本位语法体系图解法相对应，二是部分语法点具体化利于之后语法点的分析。《长城汉语》中未经调整的语法点有 161 个，调整之后为 189 个，本文采用的数据均为调整之后的语法点。

4.2.1 语法点 XPath 形成

单一教材语法点分析，首先将语法点按照在课文出现顺序，逐一存储到 excel 文档中。Excel 表格中设置分析过程中需要的一些属性：语法点序号(id)、语法点名称(gram_name)、语法点对应 XPath(xpath)、语法点所在书号(kehoushu)、语法点所在章号(kehouzhang)、例句(example)等。

id	gram_name	xpath	kehoushu	kehouzhang	example	comment
1	人称代词“你”“我”“他” “她”“您”“你们”	//*[m or l] self::[-'你' or i-'我' or i-'他' or i-'她' or i-'您' or i-'你们']	1	1	你好! 我是菲雅。 他/她是老师。 您好! 你们好!	
2	人称代词+是+谁	/j/j1/j2/j3/j4/j5/j6/j7/j8/j9/j10/j11/j12/j13/j14/j15/j16/j17/j18/j19/j20/j21/j22/j23/j24/j25/j26/j27/j28/j29/j30/j31/j32/j33/j34/j35/j36/j37/j38/j39/j40/j41/j42/j43/j44/j45/j46/j47/j48/j49/j50/j51/j52/j53/j54/j55/j56/j57/j58/j59/j60/j61/j62/j63/j64/j65/j66/j67/j68/j69/j70/j71/j72/j73/j74/j75/j76/j77/j78/j79/j80/j81/j82/j83/j84/j85/j86/j87/j88/j89/j90/j91/j92/j93/j94/j95/j96/j97/j98/j99/j100/j101/j102/j103/j104/j105/j106/j107/j108/j109/j110/j111/j112/j113/j114/j115/j116/j117/j118/j119/j120/j121/j122/j123/j124/j125/j126/j127/j128/j129/j130/j131/j132/j133/j134/j135/j136/j137/j138/j139/j140/j141/j142/j143/j144/j145/j146/j147/j148/j149/j150/j151/j152/j153/j154/j155/j156/j157/j158/j159/j160/j161/j162/j163/j164/j165/j166/j167/j168/j169/j170/j171/j172/j173/j174/j175/j176/j177/j178/j179/j180/j181/j182/j183/j184/j185/j186/j187/j188/j189/j190/j191/j192/j193/j194/j195/j196/j197/j198/j199/j200/j201/j202/j203/j204/j205/j206/j207/j208/j209/j210/j211/j212/j213/j214/j215/j216/j217/j218/j219/j220/j221/j222/j223/j224/j225/j226/j227/j228/j229/j230/j231/j232/j233/j234/j235/j236/j237/j238/j239/j240/j241/j242/j243/j244/j245/j246/j247/j248/j249/j250/j251/j252/j253/j254/j255/j256/j257/j258/j259/j260/j261/j262/j263/j264/j265/j266/j267/j268/j269/j270/j271/j272/j273/j274/j275/j276/j277/j278/j279/j280/j281/j282/j283/j284/j285/j286/j287/j288/j289/j290/j291/j292/j293/j294/j295/j296/j297/j298/j299/j300/j301/j302/j303/j304/j305/j306/j307/j308/j309/j310/j311/j312/j313/j314/j315/j316/j317/j318/j319/j320/j321/j322/j323/j324/j325/j326/j327/j328/j329/j330/j331/j332/j333/j334/j335/j336/j337/j338/j339/j340/j341/j342/j343/j344/j345/j346/j347/j348/j349/j350/j351/j352/j353/j354/j355/j356/j357/j358/j359/j360/j361/j362/j363/j364/j365/j366/j367/j368/j369/j370/j371/j372/j373/j374/j375/j376/j377/j378/j379/j380/j381/j382/j383/j384/j385/j386/j387/j388/j389/j390/j391/j392/j393/j394/j395/j396/j397/j398/j399/j400/j401/j402/j403/j404/j405/j406/j407/j408/j409/j410/j411/j412/j413/j414/j415/j416/j417/j418/j419/j420/j421/j422/j423/j424/j425/j426/j427/j428/j429/j430/j431/j432/j433/j434/j435/j436/j437/j438/j439/j440/j441/j442/j443/j444/j445/j446/j447/j448/j449/j450/j451/j452/j453/j454/j455/j456/j457/j458/j459/j460/j461/j462/j463/j464/j465/j466/j467/j468/j469/j470/j471/j472/j473/j474/j475/j476/j477/j478/j479/j480/j481/j482/j483/j484/j485/j486/j487/j488/j489/j490/j491/j492/j493/j494/j495/j496/j497/j498/j499/j500/j501/j502/j503/j504/j505/j506/j507/j508/j509/j510/j511/j512/j513/j514/j515/j516/j517/j518/j519/j520/j521/j522/j523/j524/j525/j526/j527/j528/j529/j530/j531/j532/j533/j534/j535/j536/j537/j538/j539/j540/j541/j542/j543/j544/j545/j546/j547/j548/j549/j550/j551/j552/j553/j554/j555/j556/j557/j558/j559/j560/j561/j562/j563/j564/j565/j566/j567/j568/j569/j570/j571/j572/j573/j574/j575/j576/j577/j578/j579/j580/j581/j582/j583/j584/j585/j586/j587/j588/j589/j590/j591/j592/j593/j594/j595/j596/j597/j598/j599/j600/j601/j602/j603/j604/j605/j606/j607/j608/j609/j610/j611/j612/j613/j614/j615/j616/j617/j618/j619/j620/j621/j622/j623/j624/j625/j626/j627/j628/j629/j630/j631/j632/j633/j634/j635/j636/j637/j638/j639/j640/j641/j642/j643/j644/j645/j646/j647/j648/j649/j650/j651/j652/j653/j654/j655/j656/j657/j658/j659/j660/j661/j662/j663/j664/j665/j666/j667/j668/j669/j670/j671/j672/j673/j674/j675/j676/j677/j678/j679/j680/j681/j682/j683/j684/j685/j686/j687/j688/j689/j690/j691/j692/j693/j694/j695/j696/j697/j698/j699/j700/j701/j702/j703/j704/j705/j706/j707/j708/j709/j710/j711/j712/j713/j714/j715/j716/j717/j718/j719/j720/j721/j722/j723/j724/j725/j726/j727/j728/j729/j730/j731/j732/j733/j734/j735/j736/j737/j738/j739/j740/j741/j742/j743/j744/j745/j746/j747/j748/j749/j750/j751/j752/j753/j754/j755/j756/j757/j758/j759/j760/j761/j762/j763/j764/j765/j766/j767/j768/j769/j770/j771/j772/j773/j774/j775/j776/j777/j778/j779/j780/j781/j782/j783/j784/j785/j786/j787/j788/j789/j790/j791/j792/j793/j794/j795/j796/j797/j798/j799/j800/j801/j802/j803/j804/j805/j806/j807/j808/j809/j810/j811/j812/j813/j814/j815/j816/j817/j818/j819/j820/j821/j822/j823/j824/j825/j826/j827/j828/j829/j830/j831/j832/j833/j834/j835/j836/j837/j838/j839/j840/j841/j842/j843/j844/j845/j846/j847/j848/j849/j850/j851/j852/j853/j854/j855/j856/j857/j858/j859/j860/j861/j862/j863/j864/j865/j866/j867/j868/j869/j870/j871/j872/j873/j874/j875/j876/j877/j878/j879/j880/j881/j882/j883/j884/j885/j886/j887/j888/j889/j890/j891/j892/j893/j894/j895/j896/j897/j898/j899/j900/j901/j902/j903/j904/j905/j906/j907/j908/j909/j910/j911/j912/j913/j914/j915/j916/j917/j918/j919/j920/j921/j922/j923/j924/j925/j926/j927/j928/j929/j930/j931/j932/j933/j934/j935/j936/j937/j938/j939/j940/j941/j942/j943/j944/j945/j946/j947/j948/j949/j950/j951/j952/j953/j954/j955/j956/j957/j958/j959/j960/j961/j962/j963/j964/j965/j966/j967/j968/j969/j970/j971/j972/j973/j974/j975/j976/j977/j978/j979/j980/j981/j982/j983/j984/j985/j986/j987/j988/j989/j990/j991/j992/j993/j994/j995/j996/j997/j998/j999/j1000/j1001/j1002/j1003/j1004/j1005/j1006/j1007/j1008/j1009/j1010/j1011/j1012/j1013/j1014/j1015/j1016/j1017/j1018/j1019/j1020/j1021/j1022/j1023/j1024/j1025/j1026/j1027/j1028/j1029/j1030/j1031/j1032/j1033/j1034/j1035/j1036/j1037/j1038/j1039/j1040/j1041/j1042/j1043/j1044/j1045/j1046/j1047/j1048/j1049/j1050/j1051/j1052/j1053/j1054/j1055/j1056/j1057/j1058/j1059/j1060/j1061/j1062/j1063/j1064/j1065/j1066/j1067/j1068/j1069/j1070/j1071/j1072/j1073/j1074/j1075/j1076/j1077/j1078/j1079/j1080/j1081/j1082/j1083/j1084/j1085/j1086/j1087/j1088/j1089/j1090/j1091/j1092/j1093/j1094/j1095/j1096/j1097/j1098/j1099/j1100/j1101/j1102/j1103/j1104/j1105/j1106/j1107/j1108/j1109/j1110/j1111/j1112/j1113/j1114/j1115/j1116/j1117/j1118/j1119/j1120/j1121/j1122/j1123/j1124/j1125/j1126/j1127/j1128/j1129/j1130/j1131/j1132/j1133/j1134/j1135/j1136/j1137/j1138/j1139/j1140/j1141/j1142/j1143/j1144/j1145/j1146/j1147/j1148/j1149/j1150/j1151/j1152/j1153/j1154/j1155/j1156/j1157/j1158/j1159/j1160/j1161/j1162/j1163/j1164/j1165/j1166/j1167/j1168/j1169/j1170/j1171/j1172/j1173/j1174/j1175/j1176/j1177/j1178/j1179/j1180/j1181/j1182/j1183/j1184/j1185/j1186/j1187/j1188/j1189/j1190/j1191/j1192/j1193/j1194/j1195/j1196/j1197/j1198/j1199/j1200/j1201/j1202/j1203/j1204/j1205/j1206/j1207/j1208/j1209/j1210/j1211/j1212/j1213/j1214/j1215/j1216/j1217/j1218/j1219/j1220/j1221/j1222/j1223/j1224/j1225/j1226/j1227/j1228/j1229/j1230/j1231/j1232/j1233/j1234/j1235/j1236/j1237/j1238/j1239/j1240/j1241/j1242/j1243/j1244/j1245/j1246/j1247/j1248/j1249/j1250/j1251/j1252/j1253/j1254/j1255/j1256/j1257/j1258/j1259/j1260/j1261/j1262/j1263/j1264/j1265/j1266/j1267/j1268/j1269/j1270/j1271/j1272/j1273/j1274/j1275/j1276/j1277/j1278/j1279/j1280/j1281/j1282/j1283/j1284/j1285/j1286/j1287/j1288/j1289/j1290/j1291/j1292/j1293/j1294/j1295/j1296/j1297/j1298/j1299/j1300/j1301/j1302/j1303/j1304/j1305/j1306/j1307/j1308/j1309/j1310/j1311/j1312/j1313/j1314/j1315/j1316/j1317/j1318/j1319/j1320/j1321/j1322/j1323/j1324/j1325/j1326/j1327/j1328/j1329/j1330/j1331/j1332/j1333/j1334/j1335/j1336/j1337/j1338/j1339/j1340/j1341/j1342/j1343/j1344/j1345/j1346/j1347/j1348/j1349/j1350/j1351/j1352/j1353/j1354/j1355/j1356/j1357/j1358/j1359/j1360/j1361/j1362/j1363/j1364/j1365/j1366/j1367/j1368/j1369/j1370/j1371/j1372/j1373/j1374/j1375/j1376/j1377/j1378/j1379/j1380/j1381/j1382/j1383/j1384/j1385/j1386/j1387/j1388/j1389/j1390/j1391/j1392/j1393/j1394/j1395/j1396/j1397/j1398/j1399/j1400/j1401/j1402/j1403/j1404/j1405/j1406/j1407/j1408/j1409/j1410/j1411/j1412/j1413/j1414/j1415/j1416/j1417/j1418/j1419/j1420/j1421/j1422/j1423/j1424/j1425/j1426/j1427/j1428/j1429/j1430/j1431/j1432/j1433/j1434/j1435/j1436/j1437/j1438/j1439/j1440/j1441/j1442/j1443/j1444/j1445/j1446/j1447/j1448/j1449/j1450/j1451/j1452/j1453/j1454/j1455/j1456/j1457/j1458/j1459/j1460/j1461/j1462/j1463/j1464/j1465/j1466/j1467/j1468/j1469/j1470/j1471/j1472/j1473/j1474/j1475/j1476/j1477/j1478/j1479/j1480/j1481/j1482/j1483/j1484/j1485/j1486/j1487/j1488/j1489/j1490/j1491/j1492/j1493/j1494/j1495/j1496/j1497/j1498/j1499/j1500/j1501/j1502/j1503/j1504/j1505/j1506/j1507/j1508/j1509/j1510/j1511/j1512/j1513/j1514/j1515/j1516/j1517/j1518/j1519/j1520/j1521/j1522/j1523/j1524/j1525/j1526/j1527/j1528/j1529/j1530/j1531/j1532/j1533/j1534/j1535/j1536/j1537/j1538/j1539/j1540/j1541/j1542/j1543/j1544/j1545/j1546/j1547/j1548/j1549/j1550/j1551/j1552/j1553/j1554/j1555/j1556/j1557/j1558/j1559/j1560/j1561/j1562/j1563/j1564/j1565/j1566/j1567/j1568/j1569/j1570/j1571/j1572/j1573/j1574/j1575/j1576/j1577/j1578/j1579/j1580/j1581/j1582/j1583/j1584/j1585/j1586/j1587/j1588/j1589/j1590/j1591/j1592/j1593/j1594/j1595/j1596/j1597/j1598/j1599/j1600/j1601/j1602/j1603/j1604/j1605/j1606/j1607/j1608/j1609/j1610/j1611/j1612/j1613/j1614/j1615/j1616/j1617/j1618/j1619/j1620/j1621/j1622/j1623/j1624/j1625/j1626/j1627/j1628/j1629/j1630/j1631/j1632/j1633/j1634/j1635/j1636/j1637/j1638/j1639/j1640/j1641/j1642/j1643/j1644/j1645/j1646/j1647/j1648/j1649/j1650/j1651/j1652/j1653/j1654/j1655/j1656/j1657/j1658/j1659/j1660/j1661/j1662/j1663/j1664/j1665/j1666/j1667/j1668/j1669/j1670/j1671/j1672/j1673/j1674/j1675/j1676/j1677/j1678/j1679/j1680/j1681/j1682/j1683/j1684/j1685/j1686/j1687/j1688/j1689/j1690/j1691/j1692/j1693/j1694/j1695/j1696/j1697/j1698/j1699/j1700/j1701/j1702/j1703/j1704/j1705/j1706/j1707/j1708/j1709/j1710/j1711/j1712/j1713/j1714/j1715/j1716/j1717/j1718/j1719/j1720/j1721/j1722/j1723/j1724/j1725/j1726/j1727/j1728/j1729/j1730/j1731/j1732/j1733/j1734/j1735/j1736/j1737/j1738/j1739/j1740/j1741/j1742/j1743/j1744/j1745/j1746/j1747/j1748/j1749/j1750/j1751/j1752/j1753/j1754/j1755/j1756/j1757/j1758/j1759/j1760/j1761/j1762/j1763/j1764/j1765/j1766/j1767/j1768/j1769/j1770/j1771/j1772/j1773/j1774/j1775/j1776/j1777/j1778/j1779/j1780/j1781/j1782/j1783/j1784/j1785/j1786/j1787/j1788/j1789/j1790/j1791/j1792/j1793/j1794/j1795/j1796/j1797/j1798/j1799/j1800/j1801/j1802/j1803/j1804/j1805/j1806/j1807/j1808/j1809/j1810/j1811/j1812/j1813/j1814/j1815/j1816/j1817/j1818/j1819/j1820/j1821/j1822/j1823/j1824/j1825/j1826/j1827/j1828/j1829/j1830/j1831/j1832/j1833/j1834/j1835/j1836/j1837/j1838/j1839/j1840/j1841/j1842/j1843/j1844/j1845/j1846/j1847/j1848/j1849/j1850/j1851/j1852/j1853/j1854/j1855/j1856/j1857/j1858/j1859/j1860/j1861/j1862/j1863/j1864/j1865/j1866/j1867/j1868/j1869/j1870/j1871/j1872/j1873/j1874/j1875/j1876/j1877/j1878/j1879/j1880/j1881/j1882/j1883/j1884/j1885/j1886/j1887/j1888/j1889/j1890/j1891/j1892/j1893/j1894/j1895/j1896/j1897/j1898/j1899/j1900/j1901/j1902/j1903/j1904/j1905/j1906/j1907/j1908/j1909/j1910/j1911/j1912/j1913/j1914/j1915/j1916/j1917/j1918/j1919/j1920/j1921/j1922/j1923/j1924/j1925/j1926/j1927/j1928/j1929/j1930/j1931/j1932/j1933/j1934/j1935/j1936/j1937/j1938/j1939/j1940/j1941/j1942/j1943/j1944/j1945/j1946/j1947/j1948/j1949/j1950/j1951/j1952/j1953/j1954/j1955/j1956/j1957/j1958/j1959/j1960/j1961/j1962/j1963/j1964/j1965/j1966/j1967/j1968/j1969/j1970/j1971/j1972/j1973/j1974/j1975/j1976/j1977/j1978/j1979/j1980/j1981/j1982/j1983/j1984/j1985/j1986/j1987/j1988/j1989/j1990/j1991/j1992/j1993/j1994/j1995/j1996/j1997/j1998/j1999/j2000/j2001/j2002/j2003/j2004/j2005/j2006/j2007/j2008/j2009/j2010/j2011/j2012/j2013/j2014/j2015/j2016/j2017/j2018/j2019/j2020/j2021/j2022/j2023/j2024/j2025/j2026/j2027/j2028/j2029/j2030/j2031/j2032/j2033/j2034/j2035/j2036/j2037/j2038/j2039/j2040/j2041/j2042/j2043/j2044/j2045/j2046/j2047/j2048/j2049/j2050/j2051/j2052/j2053/j2054/j2055/j2056/j2057/j2058/j2059/j2060/j2061/j2062/j2063/j2064/j2065/j2066/j2067/j2068/j2069/j2070/j2071/j2072/j2073/j2074/j2075/j2076/j2077/j2078/j2079/j2080/j2081/j2082/j2083/j2084/j2085/j2086/j2087/j2088/j2089/j2090/j2091/j2092/j2093/j2094/j2095/j2096/j2097/j2098/j2099/j2100/j2101/j2102/j2103/j2104/j2105/j2106/j2107/j2108/j2109/j2110/j2111/j2112/j2113/j2114/j2115/j2116/j2117/j2118/j2119/j2120/j2121/j2122/j2123/j2124/j2125/j2126/j2127/j2128/j2129/j2130/j2131/j2132/j2133/j2134/j2135/j2136/j2137/j2138/j2139/j2140/j2141/j2142/j2143/j2144/j2145/j2146/j2147/j2148/j2149/j2150/j2151/j2152/j2153/j2154/j2155/j2156/j2157/j2158/j2159/j2160/j2161/j2162/j2163/j2164/j2165/j2166/j2167/j2168/j2169/j2170/j2171/j2172/j2173/j2174/j2175/j2176/j2177/j2178/j2179/j2180/j2181/j2182/j2183/j2184/j2185/j2186/j2187/j2188/j2189/j2190/j2191/j2192/j2193/j2194/j2195/j2196/j2197/j2198/j2199/j2200/j2201/j2202/j2203/j2204/j2205/j2206/j2207/j2208/j2209/j2210/j2211/j2212/j2213/j2214/j2215/j2216/j2217/j2218/j2219/j2220/j2221/j2222/j2223/j2224/j2225/j2226/j2227/j2228/j2229/j2230/j2231/j2232/j2233/j2234/j2235/j2236/j2237/j2238/j2239/j2240/j2241/j2242/j2243/j2244/j2245/j2246/j2247/j2248/j2249/j2250/j2251/j2252/j2253/j2254/j2255/j2256/j2257/j2258/j2259/j2260/j2261/j2262/j2263/j2264/j2265/j2266/j2267/j2268/j2269/j2270/j2271/j2272/j2273/j2274/j2275/j2276/j2277/j2278/j2279/j2280/j2281/j2282/j2283/j2284/j2285/j2286/j2287/j2288/j2289/j2290/j2291/j2292/j2293/j2294/j2295/j2296/j2297/j2298/j2299/j2300/j2301/j2302/j2303/j2304/j2305/j2306/j2307/j2308/j2309/j2310/j2311/j2312/j2313/j2314/j2315/j2316/j2317/j2318/j2319/j2320/j2321/j2322/j2323/j2324/j2325/j2326/j2327/j2328/j2329/j2330/j2331/j2332/j2333/j2334/j2335/j2336/j2337/j2338/j2339/j2340/j2341/j2342/j2343/j2344/j2345/j2346/j2347/j2348/j2349/j2350/j2351/j2352/j2353/j2354/j2355/j2356/j2357/j2358/j2359/j2360/j2361/j2362/j2363/j2364/j2365/j2366/j2367/j2368/j2369/j2370/j2371/j2372/j2373/j2374/j2375/j2376/j2377/j2378/j2379/j2380/j2381/j2382/j2383/j2384/j2385/j2386/j2387/j2388/j2389/j2390/j2391/j2392/j2393/j2394/j2395/j2396/j2397/j2398/j2399/j2400/j2401/j2402/j2403/j2404/j2405/j2406/j2407/j2408/j2409/j2410/j2411/j2412/j2413/j2414/j2415/j2416/j2417/j2418/j2419/j2420/j2421/j2422/j2423/j2424/j2425/j2426/j2427/j2428/j2429/j2430/j2431/j2432/j2433/j2434/j2435/j2436/j2437/j2438/j2439/j2440/j2441/j2442/j2443/j2444/j2445/j2446/j2447/j2448/j2449/j2450/j2451/j2452/j2453/j2454/j2455/j2456/j2457/j2458/j2459/j2460/j2461/j2462/j2463/j2464/j2465/j2466/j2467/j2468/j2469/j2470/j2471/j2472/j2473/j2474/j2475/j2476/j2477/j2478/j2479/j2480/j2481/j2482/j2483/j2484/j2485/j2486/j2487/j2488/j2489/j2490/j2491/j2492/j2493/j2494/j2495/j2496/j2497/j2498/j2499/j2500/j2501/j2502/j2503/j2504/j2505/j2506/j2507/j2508/j2509/j2510/j2511/j2512/j2513/j2514/j2515/j2516/j2517/j2518/j2519/j2520/j2521/j2522/j2523/j2524/j2525/j2526/j2527/j2528/j2529/j2530/j2531/j2532/j2533/j2534/j2535/j2536/j2537/j2538/j2539/j2540/j2541/j2542/j2543/j2544/j2545/j2546/j2547/j2548/j2549/j2550/j2551/j2552/j2553/j2554/j2555/j2556				

id	gram_name	xpath	kehoushu	kehouzhang	example
1	人称代词 "你..."	//*[m or i]self...	1	1	你好！我是非...
2	人称代词+是+谁	(/ju /ju/fj)/dj/...	1	1	他是谁？你是...
3	人称代词+是+...	(/ju /ju/fj)/dj/...	1	1	我是山口和子...
4	主语+动词+什...	(/ju /ju/fj)/dj/...	1	2	你叫什么名字？
5	主语+动词+什...	(/ju /ju/fj)/dj/...	1	2	他姓什么？
6	主语+副词+动...	(/ju /ju/fj)/dj/...	1	2	他是老师吗？...
7	主语+也+动词...	(/ju /ju/fj)/dj/...	1	2	他也是学生。...
8	主语+也+副词...	(/ju /ju/fj)/dj/...	1	2	我也很高兴。
9	主语+不+动词...	(/ju /ju/fj)/dj/...	1	3	李冬生不是学...
10	主语+动词+哪...	(/ju /ju/fj)/dj/...	1	3	玛丽是哪儿/什...
11	主语+动词+哪...	(/ju /ju/fj)/dj/...	1	3	NULL
12	主语+介词+哪...	(/ju /ju/fj)/dj/...	1	3	你从哪儿/什么...
13	主语+动词+名...	(/ju /ju/fj)/dj/...	1	3	他是你的老师...
14	主语+时间副词...	NULL	1	4	金太成每天来...
15	名词/代名词+...	//*[m or i]di[...	1	4	罗森的护照公...

图9 语法点索引库

4.2.3 数据库设计

在数据库中新建语法分析库 (GramStat)，并将存储为 excel 形式的语法点信息读取为其中一表 (grampoint)。

id	gram_name	xpath	kehoushu	kehouzhang	example
1	人称代词 "你" "我" "他"	//*[m or i]self+*[is' or is'	1	1	你好！我是非... 他/她是老师。您...
2	人称代词+是+谁	(/ju /ju/fj)/dj/...	1	1	他是谁？你是谁？
3	人称代词+是+名词	(/ju /ju/fj)/dj/...	1	1	我是山口和子。他是老师。
4	主语+动词+什么+名词	(/ju /ju/fj)/dj/...	1	2	你叫什么名字？
5	主语+动词+什么 (什么作宾语)	(/ju /ju/fj)/dj/...	1	2	他姓什么？
6	主语+副词+动词+名词/代名词	(/ju /ju/fj)/dj/...	1	2	他是老师吗？你也认识克/姆吗？
7	主语+也+动词+名词	(/ju /ju/fj)/dj/...	1	2	他也是学生。我也认识克/姆。

图10 数据库中 grampoint 表示例

之前标注好的句子信息存储为语料库 (YuLiaoKu)，读取为另一表 (Juzi_gwc)，此表内容设置各列属性：句子内容 (yuanwen)、句子序号 (juhao)、句子所在书号 (shuhao)、句子所在章号 (zhanghao) 等。

id	juhao	yuanwen	biaoshu	shuhao	zhanghao	yehao	juhao	beizhu	biaoshuhao	shenbushijian
1	1	第一单元	<ju_id="0">_je="PeiJu">_ye="第...	1	1	0	0		NULL	2011-12-25 20:36:11.96
2	2	第一课	<ju_id="1">_je="PeiJu">_ye="第...	1	1	0	0		NULL	2011-12-25 20:36:11.96
3	3	刘少华:	<ju_id="2">_je="PeiJu">_ye="第...	1	1	0	0		NULL	2011-12-25 20:36:11.96
4	4	我是刘少华。	<ju_id="0">_je="PeiJu">_ye="第...	1	1	0	0		hijing	2011-12-30 22:14:22.05
5	5	赵玉兰:	<ju_id="4">_je="PeiJu">_ye="第...	1	1	0	0		NULL	2011-12-25 20:36:11.96
6	6	我是赵玉兰。	<ju_id="0">_je="PeiJu">_ye="第...	1	1	0	0		NULL	2012-08-03 13:56:00.75
7	7	黄森:	<ju_id="6">_je="PeiJu">_ye="第...	1	1	0	0		NULL	2011-12-25 20:36:11.96

图11 语料库中 Juzi_gwc 表示例

将 grampoint 表与 Juzi_gwc 表在数据库中建立联系，将每一语法点在课文原句中进行

查找，找出与之对应的句子，并读取相互对应的语法点序号（id）与句子序号（jid），生成 Gramstat 中的表 relation。

	id	gid	jid
1	59854	1	1100
2	59855	1	1101
3	59856	1	1103
4	59857	1	1107
5	59858	1	1113
6	59859	1	1115
7	59860	1	1124
8	59861	1	1127

图 12 关联表 relation 示例

4.2.4 语法等级大纲

语法等级大纲出自《汉语水平等级标准与语法等级大纲》一书，本书是国家对外汉语教学领导小组办公室汉语水平考试部制定的对外汉语教学的等级标准和水平大纲，是对外汉语教学总体设计、教材编写、课堂教学和课程测试的主要依据。^[1]

这本书在语法等级大纲部分，根据连续继承原则、纳新发展原则、分级定位原则、循环递进原则、系统协调原则和综合过渡原则等进行多方面多层次的考察、统计、比较和筛选，将主要的语法点分为甲乙丙丁四级。因此将语法等级大纲作为分析单一教材的理论标准可以保证一定的科学性、规范性与实用性。

4.2.5 教材语法分析

将在数据库中生成的表 relation，借助可视化语法分析展示平台进行直观显示，可看到大部分的语法点可以正常检索。以语法等级大纲所分等级和其具体呈现的语法点内容对长城汉语课后语法点进行分析，发现有些语法点符合对外汉语教材语法项目的编写原则，也有语法点在编排上面仍然存在很多问题。本文将据语法等级大纲提出以下几个方面的评价标准，从语法数量、所属等级、语法点对语法大纲各等级覆盖率和语法点的复现等逐一进行举例分析：

1. 语法点数量

《长城汉语》全册书籍共六册，每册十单元，语法总数为 189 个，平均每单元语法点数目为 3 个。有关语法点数量的问题，吕文华曾指出：“每课的语法教学时候适度是影响教学效果直接因素。它取决于教学总量的控制。总量大，分布在备课的教学量就大。同样也取决于每个语法项目的含量，不同语法项目的含量差别悬殊。这既与语法项目本身的简繁有关，也与教材中展示语法项目的内容简繁有关。”^[17]吕文华和杨德峰曾在每课语法含量方面做过相关论述，认为每课的语法数量以 2-3 个为宜，过多反而不利于学习者掌握语法。根据他们的建议来分析，《长城汉语》每课语法数量的安排趋于合理性。

2. 语法点等级

表 2 《长城汉语》语法点等级数量分布

甲级	乙级	丙级	超纲	总和
154	16	10	9	189

其中甲级语法点在书中的覆盖率为 81.48%，超纲语法点一般为具体副词的用法和大纲未涉及的口语格式，如：“…什么的、副词‘一直’”的用法等，这部分覆盖率仅为 4.76%。超纲语法点是作为非关键性语法点扩展讲解，而《长城汉语》一书面向的对象为初级学者，所以即使初级学者对此部分语法点未重点掌握，在系统学习完《长城汉语》全六册之后，已经可以进行简单的日常交流了。

3. 对大纲各等级覆盖率

表 3 《长城汉语》语法点对大纲等级覆盖率

等级项目	《长城汉语》	大纲	覆盖率
甲 级	68	129	52.71%
乙 级	11	123	8.94%
丙 级	10	400	2.5%
丁 级	0	516	0

据从上表可以看到，甲级和乙级语法项目对大纲的覆盖率达到 60% 以上，虽然甲级语法点在书中的覆盖率为 80% 以上，但是存在部分语法点为大纲中同一语法点的不同表达或具体举例表达。国家汉办颁布的文件规定，高等学校外国留学生教学大纲一年级教学语法项目共 102 个，从这个角度来分析，《长城汉语》在语法项目的选择上仍然存在缺陷。

4. 语法点的复现

语法复现是为了巩固学习者对知识的进一步掌握，利于知识的长时记忆。如果一个语法点学过以后在后面课文中很少甚至不再出现，不但不符合学习规律，而且也不利于学生掌握语法。语法点的复现位置包括课文、语法、练习等几个主要部分，课文是构成教材的主要内容，也是学生学习语言依据的主要材料，所以本文主要研究语法点在课文当中的复现。语法点的课文复现除了考虑在之后课文中语法点的复现，还需要考虑语法点在本课课文中的自然复现。

(1) 本课课文复现

从分析结果来看，大部分的语法点在本课中仅出现一次，其次是出现次数 2-4 次的语法点，极少数的语法点能在本课中出现 10 次以上，如语法点“主语+谓语部分(名词)”(《长城汉语》第一册第五单元)出现次数达到 16 次。

(2) 之后课文复现

一些简单的在日常生活中经常能用到的语法点，很容易在课文中反复出现，所以此类语法点应当在课文前面讲解。如《长城汉语》第一册第一单元语法点“人称代词‘你’‘我’‘他/她’‘您’‘你们’”，在课文中出现情况如图 13：

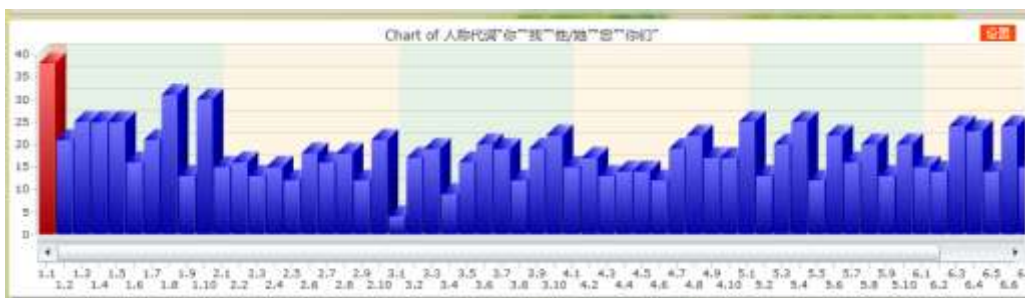


图 13 语法点“人称代词‘你’‘我’‘他/她’‘您’‘你们’”分布图

《长城汉语》全六册的语法点大部分的语法点都能在之后课文进行合理复现，只有极少数的语法点在本课中出现之后的后续课文当中没再出现，如语法点“位置+是+名词短语”（《长城汉语》第二册第二单元），如图表 14：



图 14 语法点“位置+是+名词短语”分布图

5 存在的问题

5.1 平台功能完善

语法分析展示平台在直观展示和量化语法分析方面已经实现了部分功能，但是并不完善。因为要做好语法分析工作并不是用一两个数据展示就能完成的，其自主分析的功能目前还没有完成。在当下阶段笔者想到的可以为此平台添加的功能有以下几点：

1. 给每项语法点附加语法所属等级（甲、乙、丙、丁）属性，自动计算各等级语法点所占有的比例。寻找一个能说明教材语法点等级分布合理的数据，根据得出的数值来分析此种教材是否是一部比较适合学习者的教材。

2. 此平台展示的为每一语法点在课文中出现的位置情况，根据这部分数据，可以统计具体的句子所对应的语法点有哪些，将交集在两个以上语法点的句子进行对比，对其研究分析以找出更好的编排教材的方法。

3. 语法分析研究工作中，语法点的复现是一个很重要的研究工作，所以平台应该根据量化的结果能统计每个语法点在课文中的复现率，并使结果在界面进行直观展现。

5.2 标注工具调整

句本位语法体系应用于对外汉语教材语法点分析有其独特的优势，但是并不能说明这个体系在分析语法时尽善尽美，其中仍然存在一些问题是当下还不能一一解决的。这部分问题大多针对于标注阶段的工作，运用现有的句本位语法体系还不能很好的对应每个语法点的结构。为了使后期结果准确性进一步提高，笔者在自己研究和理解的基础上对标注体系提出以下几点改进的意见：

1. 现有的标注工具对名词、动词、形容词、助动词、语气词等做了区分（或者添加标签进行区分，或者以出现位置进行区分），但是对每种词类中细分的集中词却无法分辨，比如疑问代词与人称代词，地点副词、时间副词和频率副词。像这种涉及到大词类中的具体词类，需要对其进行区分。

2. 目前黎氏检索工具尚无法实现句子语气的检索，像语法点中涉及到夸张语气、感叹句等句子不能很好的进行分辨。陈述句、感叹句、疑问句等几种句型笔者认为可以通过将句末标点单独进行标注来解决，但具体实施需要进行更深入细致的考虑，防止对检索的其他功能产生影响。夸张语气、肯定语气、揣测语气等的检索可通过句末的语气词来判断语气，但是同样的语气词可以表达不同的语气，同一种语气可以选用不同的语气词，所以这给准确检索带来了很大的不便之处。笔者认为可以在句子标注的阶段增加对句子的语气属性进行人工填写，虽然增加了一部分工作量，但是可以保证后期检索的准确性。

5.3 体系适应性调整

将句本位语法体系引入对外汉语教材语法点分析是一个对其进行检验和改造的过程，其中必然有一定程度的扬弃，进而寻求其在语法分析方面加以运用的突破。黎锦熙先生创立的图解法在格局上存在和教学语法体系不对应的地方，并且很多术语和教学语法也不完全对应。比如语法教学中的双宾语句，如“我送你礼物”，图解法中主干成分只允许一个直接宾语“礼物”的存在，而间接宾语“你”做了谓语动词“送”的补语副位，其图解示例如下图：

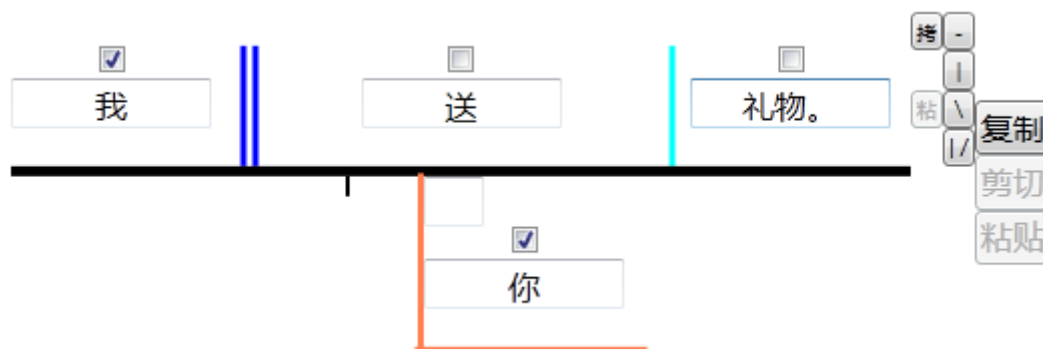


图 15 双宾语示例图

像“我送你礼物”这种涉及双宾语的句子，很容易检索出“斩之于马下”等古文性质一类的句子。而且黎氏语法中定义的补足语成分，在教学语法中已不存在。语法分析中黎氏图解法的弊病已有显现，而一个体系的出现并不是孤立的，而是为了解决实际问题提供一种标准，所以这个体系应该随着其实际运用进一步对格局进行调整。

以上这几点问题，有的是因为笔者自己对句本位语法体系研究不够深入造成的，有的则是因为黎氏图解格局自身存在的缺陷。这些问题都是在检索阶段准确性检测的大量工作

基础上进行总结分析发现的。在本文接下来的工作当中将继续对句式系统进行研究,以期句本位语法体系能更好地适用于对外汉语教材语法点分析。

6 总结与展望

本文将黎锦熙先生提出的句本位语法体系引入对外汉语教材的语法分析研究,设计出与XML标注体系相应的一套图解控件,并通过检索工具对存储的XML格式数据进行检索,实现了《长城汉语》教材语法点的初步量化分析。

语法分析一直是对外汉语教材分析中的一个重要方面,在此之前国内外许多学者一直致力于教材的语法分析工作,但是多数都采用人工统计与选取的方式。本文尝试使用计算机技术进行语法分析,分别设计出可视化检索工具和语法分析展示平台,不仅使分析结果得以量化,还提高了工作效率,为以后语法点的进一步定量分析奠定了发展基础。

将句本位语法体系用于对外汉语教材语法分析虽然目前只针对一种教材进行了应用,语法分析展示平台目前能提供的信息量还不是很完善,但按这种模式进行其他教材的语法分析,甚至扩展为对不同教材语法点的对比分析,都不失为一次很好的尝试。通过对单一教材语法项目的分析,使学习者能根据自身情况选择适合自己学习的教材,也为实际生活中教材语法选择与编排提供了一些较好的方法。

参考文献

- [1]. 段姣平. 初级对外汉语教材《汉语教程》中的补语编排研究[D]. 上海师范大学, 2012.
- [2]. 冯志伟. 自然语言处理的形式模型[M]. 合肥: 中国科学技术大学出版社, 2010.
- [3]. 傅爱平. 语料库研究与应用综述[EB/OL].
<http://ling.cass.cn/yinyong/courses/corpusbase.htm>, /2012-4-2.
- [4]. 耿直. 改革开放以来对外汉语教材编写研究综述[J]. 河南社会科学, 2011, 19(4).
- [5]. 胡明扬. 现代汉语语法的开创性著作——《新著国语文法》的再认识和再评价[J]. 语言科学, 2002, (01).
- [6]. 黄昌宁. 关于大规模真实文本的谈话[J]. 语言文字应用, 1993.
- [7]. 黎锦熙. 汉语语法图解总公式[A]. 黎锦熙选集[C]. 长春: 东北师范大学出版社, 2001: 397-399.
- [8]. 黎锦熙(1955). 新著国语文法[M]. 北京: 商务印书馆.
- [9]. 黎锦熙, 刘世儒. 汉语语法教材[M]. 北京: 商务印书馆, 1957.
- [10]. 黎锦熙, 刘世儒. 汉语语法教材(第一编 基本规律)[M]. 北京: 商务印书馆, 1957.
- [11]. 黎锦熙, 刘世儒. 汉语语法教材(第二编 词类和构词法)[M]. 北京: 商务印书馆, 1959.
- [12]. 李泉. 对外汉语教材研究[M]. 北京: 商务印书馆.
- [13]. 李泉. 近20年对外汉语教材编写和研究的基本情况述评[J]. 语言文字应用, 2002, 3.
- [14]. 李晓琪, 章欣. 新形势下对外汉语语法教学研究[J]. 汉语学习, 2010, 1.
- [15]. 刘若云. 基础汉语教材编写中语法点的复现问题[J]. 中山大学学报论丛, 2004, 24(1).
- [16]. 吕叔湘. 汉语语法分析问题[M]. 北京: 商务印书馆, 1979.
- [17]. 吕文华. 对外汉语教材语法项目排序的原则及策略[J]. 世界汉语教学, 2002(4): 92.
- [18]. 吕文华. 对外汉语教学语法探索[M]. 北京: 北京语言大学出版社, 2008.
- [19]. 宋继华, 杨尔弘, 王强军. 中文信息处理教程[M]. 北京: 高等教育出版社, 2010.
- [20]. 宋柔. 对外汉语教学中的信息资源和信息处理[M]. 北京: 北京大学出版社, 2008.
- [21]. 王兰野. XML数据流上的XPath查询处理研究[D]. 山东大学, 2011.

- [22]. 吴蔚天, 罗建林(1994)《汉语计算语言学—汉语形式语法和形式分析》[M], 北京: 电子工业出版社.
- [23]. 徐子亮, 吴仁甫. 实用对外汉语教学法[M]. 北京: 北京大学出版社, 2005.
- [24]. 杨德峰. 初级汉语教材语法点确定、编排中存在的问题——兼议语法点确定、编排的原则[J]. 世界汉语教学, 2001(2): 84-87.
- [25]. 宗成庆, 曹右琦, 俞士汶. 中文信息处理 60 年[J]. 语言文字应用, 2009, (04).
- [26]. 张宏伟. XML 数据流的 XPath 查询算法研究[D]. 中国石油大学, 2007.
- [27]. 张麟, 郝忠孝. 基于 XML Schema 的 Xpath 查询优化[J]. 电脑知识与技术, 2009, 5(29).
- [28]. 张普. 论汉语信息处理技术与对外汉语教学[J]. 语言教学与研究, 1991.
- [29]. 张普. 关于大规模真是文本语料库的几点理论思考[J]. 语言文字应用, 1999.
- [30]. 张普, 谢天蔚, 徐娟. 数字化对外汉语教学理论与方法研究[M]. 北京: 清华大学出版社, 2004.
- [31]. 张旺熹, 崔永华. 对外汉语教学语法问题研究的基本态势[A]. 世纪之交的中国应用语言学研究——第二届全国语言文字应用学术研讨会论文集[C];1998 年.
- [32]. 张西平, 柳若梅. 世界主要国家语言推广政策概览[M]. 北京: 外语教学与研究出版社, 2008.
- [33]. 张羽. 近十年对外汉语语法教材研究综述[J]. 语文学刊, 2011.
- [34]. 赵金铭. 基于中介语语料库的汉语句法研究[M]. 北京: 北京大学出版社, 2008.
- [35]. 周建儒. Xpath 查询 SQL Server 2000 数据库[J]. 信息技术教学与研究, 2008, 30.