

可比语境下关系代词取舍的多因素分析*

北京邮电大学 房印杰 北京航空航天大学 梁茂成

提要:本文以中国英语学习者和英语本族语者的关系代词取舍为研究对象,通过双回归模型分析,发现:1)关系代词取舍并非随机选择,而是受到与认知加工负载相关因素的制约;2)与英语本族语者相比,中国英语学习者过度使用显性关系代词,但对隐性关系代词学习较好;3)中国英语学习者与英语本族语者的组间差异体现为关系代词取舍制约语境的典型性,在典型语境下,二者差异较小;在非典型语境下,二者差异增大。研究表明,基于“可比语境”的双回归分析能够准确剥离关系代词取舍中的显著制约因素,提取组间差异。

关键词:可比语境、关系代词取舍、过度概括、概率性预测

[中图分类号] H319.3 [文献标识码] A [文章编号] 1000-0429 (2019) 03-0435-12

1. 引言

前人研究显示,关系代词取舍(如例(1))受到多种因素的共同制约,体现出多因素特征,该特征适用于英语母语者(Diessel & Tomasello 2005; Wasow *et al.* 2011)和二语学习者(房印杰 2017)。鉴于语言使用的多因素特征,Gries & Deshors (2014)对中介语对比研究中的语境进行了更加精确的界定,提出了“可比语境”(comparable context)的概念,将其量化为对二语学习者与英语本族语者均具有显著制约作用的一系列语言学因素。他们提出,中介语对比研究应当在充分论证二语学习者与英语本族语者间存在一系列共同制约因素的前提下开展。基于该论断,本研究尝试将语言使用的多因素特征与“可比语境”相融合,探究中国英语学习者与英语本族语者在关系代词取舍中的组间差异,并尝试从认知固化和过度概括两个方面阐释该差异性的内在动因。

* 本研究为教育部人文社会科学研究青年基金项目“基于使用的中国学生英语关系从句认知统计模型研究”(17YJC740018)的阶段性成果。衷心感谢《外语教学与研究》编辑部及匿名审稿专家提出的宝贵意见。

- (1) a. The radiation that mobile phones give out can do harm to our health (WECCL 1.0).
b. Freedom is the first thing I need. (WECCL 2.0)

2. 研究背景

2.1 可比语境

Péry-Woodley (1990: 143-144) 提出:中介语研究所开展的二语学习者与英语本族语者的对比分析应当在“可比语境”下进行。传统研究往往默认学习者语料与英语本族语者语料存在可比性,而未对该可比性加以详尽阐述。其基于频次的对比分析精细度不足(Granger 1996: 45),无法对语言使用的制约机制进行有效解释。语言现象的多因素特征要求使用相对复杂的统计方法开展相关研究(Gries 2013)。在该前提下, Gries & Deshors (2014) 基于复杂统计模型对“可比语境”进行了实证性阐述,将其具化为可比的语言学特征,换言之,在一系列语言学特征的共同制约下,探讨二语学习者与英语本族语者的语言使用异同(同上: 113)。该“可比语境”1)摆脱了传统中介语对比基于原始/正态频次的范式,将语言使用的多因素特征纳入研究范畴;2)采用复杂统计模型对语言学因素进行准确剥离,并确定二语学习者和英语本族语者的共同制约因素集合;3)具有语言学理论支撑(Péry-Woodley 1990)。

2.2 关系代词取舍研究

关系代词取舍研究作为学界研究的一个重点问题,经历了三个维度的转变。首先,在研究路径上,由内省式思辨转向基于真实语言产出。在早期研究中,由于缺乏大规模语料和科学量具, Jespersen (1933: 359-361) 基于个人内省提出关系从句紧凑性假设。在基于语言产出的研究路径下,学界先后提出了从句扩展假设(Diessel 2004)、单句性倾向假设(Fox & Thompson 2007)和关系从句可预测性假设(Wasow *et al.* 2011)。其次,在研究焦点上,由单因素分析向多因素研究过渡。越来越多的研究开始同步考察前人研究所涉及的多种因素,如依存成分的长度(Gibson 1998)、不同句法成分在认知层面的可及性(Ariel 1990)、语义模糊(Temperley 2003)、先行名词和关系从句主语的有灵/无灵(Mak *et al.* 2002)等。再次,在研究对象上,从聚焦英语本族语者逐步转向探究二语学习者的关系代词取舍使用特征,对比两个群体在关系代词取舍使用中的异同(如房印杰 2017)。

前人研究探讨了关系代词取舍的制约机制,但仍存在四点不足:1)单因素分析占据相当比重,这与语言使用的多因素特征相悖(Gries 2013);2)前人提出的理论假设往往基于关系代词取舍的某几种影响因素,涵盖范围有限;3)现有研究及假设均以英语本族语者为研究对象,对二语学习者关注不足;4)前

人的多因素研究 (Fox & Thompson 2007; Wasow *et al.* 2011) 在研究方法上仍有待深入。鉴于此,本研究以中国英语学习者为研究对象,基于复杂统计模型,探究“可比语境”在中介语研究中的适用性,提取中国英语学习者与英语本族语者在“可比语境”下的组间差异。研究问题如下:在可比语言特征下,中国英语学习者与英语本族语者的关系代词取舍有何差异? 该差异呈现何种规律?

3. 研究设计

3.1 语料抽取与标注¹

本研究选取 WECCL 1.0 (文秋芳等 2005) 和 WECCL 2.0 (文秋芳等 2008) 语料库中的议论文部分作为中国英语学习者语料。选取 LOCNESS²、NESSIE³ 中的议论文部分和 ICNALE⁴ 中的英语本族语者部分作为英语本族语者语料。获得上述语料后,首先以 Stanford parser 句法剖析器对其进行句法剖析,而后基于 Tregex 树库检索软件,提取非主语型限定性关系从句,之后进行人工判别,最终获得 1726 个合格语句 (见表 1)。

表 1. 隐性 / 显性关系代词从句分布

语料类型	语料库	显性关系代词	隐性关系代词	合计
NS	LOCNESS	133	253	386
	NESSIE	86	123	209
	ICNALE.ns	36	41	77
NNS	WECCL 1.0	184	277	461
	WECCL 2.0	179	414	593
	合计	618	1108	1726

在抽取出 1726 个非主语型限定性关系从句语料后,将全部语料按照表 2 中设定的语言学特征集进行人工标注⁵。该特征集在参照前人研究发现的基础上归纳而成。

表 2. 语言学特征标注集

变量类型	变量
先行词	生命性、具体性、语义丰富度、确定性、修饰语数量、先行词在主句中成分、先行词在从句中成分
从句主语	生命性、具体性、语义丰富度、确定性、词性
主句	时态、语态、谓语、否定、句子长度、谓语动词长度
从句	时态、语态、谓语、否定、句子长度、谓语动词长度

3.2 统计分析

3.2.1 双回归模型

为了客观提取关系代词取舍的组间差异,本研究基于 R (3.5.1 版本) 统计编程语言,使用双回归分析 (Gries & Deshors 2014; Wulff & Gries 2015), 分别对英语本族语者和中国英语学习者的语料进行回归分析,这样做的理论动因在于尽可能保证两个语言群体的对比在可比语境 / 语言特征下展开。当该回归方程在两个语言群体的语料中均拟合较好时,则以二者的预测差异值为因变量开展第二种回归分析。与二项逻辑斯蒂回归不同,双回归分析聚焦于中国英语学习者与英语本族语者在关系代词使用中的差异,特别是关系代词的组间差异程度。换言之,双回归分析可以较为客观地挖掘出隐性 / 显性关系代词在中国英语学习者和英语本族语者两个群体中的组间差异,进而判定两种关系代词构式的组间差异是否具有统计学意义。

3.2.2 统计分析流程

第一轮回归: 1) 首先对英语本族语者开展二项逻辑斯蒂回归分析,并检验该模型对语料的拟合程度 (预测准确率, R^2); 2) 如果回归方程对英语本族语者语料拟合良好,则将该回归方程用于对中国英语学习者语料进行回归分析,并检验模型拟合程度 (预测准确率, R^2); 3) 当同一回归方程对英语本族语者语料和中国英语学习者语料拟合程度均较好时,提取新因变量: 同一回归方程对中国英语学习者与英语本族语者关系代词取舍预测值的差异值 (deviation), 将其标记为 Dev。

第二轮回归: 1) 以 Dev 为因变量,以第一轮回归中自变量和原因变量 (RC_type) 为变量,进行线性回归分析; 2) 对模型进行优化,提取显著主效应和因素交互效应。

4. 研究发现

4.1 模型检验

第一轮回归: 经对英语本族语者语料进行混合效应模型 (mixed-effects model) 分析,发现该回归方程对数据拟合良好: 理论预测值 (C) 达到 0.833, 实际预测值达到 0.768, 模型边际 $R^2 = 0.329$, 条件 $R^2 = 0.491$ 。这说明回归方程中的变量对英语本族语者的关系代词取舍有较强的预测力。回归方程对英语本族语者语料拟合良好,这说明可以尝试将其对中国英语学习者的语料进行混合效应模型拟合。混合效应模型分析表明: 该回归方程对中国英语学习者语料拟合同样良好,理论预测值 (C) 为 0.874, 实际预测值为 0.802, 模型边际 $R^2 = 0.504$, 条件 $R^2 = 0.536$ 。回归方程对英语本族语者和中国英语学习者的语料拟合度均

较好,这表明:该回归方程的变量可以代表英语本族语者和中国英语学习者在使用关系代词中的“可比语境”。

第二轮回归:首先设置新的因变量:中国英语学习者与英语本族语者在“可比语境”下对关系代词的预测差异值。该因变量表示为 Dev,当在中国英语学习者的语料中,英语本族语者在“可比语境”下与中国英语学习者的关系代词实际选择相同时, $Dev = 0$,表示二者无差异;实际选择不同时, $Dev = p - 0.5$ (Gries & Deshors 2014: 128),其中 p 代表英语本族语者对关系代词的预测概率。新因变量 Dev 的中数 (median) 为 0,平均数 (mean) 也接近 0 (-0.01236)。

线性回归模型显示其对因变量 Dev 拟合良好,调整 R^2 为 0.644, $p < 0.0001$,模型不存在较强的共线性。模型中显著自变量包括:从句主语确定性 (RC_subj_def)、从句主语词性 (RC_subj_POS)、先行词生命性 (Head_ani)、从句主语具体性 (RC_subj_conc);与关系代词取舍产生显著交互效应的因素包括:先行词语义丰富度 (Head_cont)、先行词具体性 (Head_conc)、关系从句长度 (RC_length)、主句谓语动词类型 (MC_predicate)、先行词修饰语数量 (Head_modi_number) 和先行词在主句中充当的句法成分 (Head_ext)。

表 3. 线性回归模型显著变量

变量	F 值 (自由度)	变量	F 值 (自由度)
RC_subj_def	10.05 (1) **	Head_modi_number:RC_type	5.13 (1) *
RC_subj_POS	149.16 (1) ***	Head_ext:RC_type	2.76 (4) *
Head_ani	5.19 (1) *	MC_predicate:RC_type	3.31 (5) **
RC_subj_conc	4.50 (1) *	RC_negation:RC_type	3.75 (1) M.S.
RC_embed	3.64 (1) M.S.	RC_length:RC_type	141.89 (1) ***
Head_cont: RC_type	11.29 (1) ***	Head_conc:RC_type	25.77 (1) ***

注: * 表示 $p < 0.05$; ** 表示 $p < 0.01$; *** 表示 $p < 0.001$; M.S. 表示 p 值边缘显著。

4.2 关系代词取舍的组间差异

关系代词取舍的组间差异围绕两个维度展开:1)隐性/显性关系代词在中国英语学习者和英语本族语者使用中的差异程度;2)隐性/显性关系代词各自在交互因素不同变量水平上的差异。本研究基于 R 语言,对七种语言学因素与关系代词取舍的交互效应 (见表 3) 逐一进行了可视化分析。分析结果显示,七种交互效应体现出相似的分布特征:显性关系代词在中国英语学习者与英语本族语者间的使用差异均大于隐性关系代词。限于篇幅,本节主要汇报与关系代词取舍具有高度交互效应 ($p < 0.01$) 的四种因素:先行词语义丰富度、从句长度、先行词具体性和主句谓语动词类型。

4.2.1 先行词语义丰富度

在线性回归模型(见表 3)中,先行词语义丰富度与关系代词取舍(RC_type)的交互效应非常显著($F = 11.29, p < 0.001$)。在对该交互效应进行可视化分析时可以发现:中国英语学习者与英语本族语者在隐性/显性关系代词使用上存在不同的差异度。在图 1 中,横轴表示中国英语学习者与英语本族语者在使用关系代词上的差异程度,该差异值在 $-0.5 \sim 0.5$ 之间摆动。观察图 1 可以发现:在先行词语义丰富度的两个变量水平(空乏、具体)上,显性关系代词的预测区间分别为 $-0.20 \sim -0.18$ 和 $-0.16 \sim -0.14$,隐性关系代词的预测区间分别为 $0.03 \sim 0.04$ 和 $0.04 \sim 0.06$ 。这说明中国英语学习者在显性关系代词的使用上与英语本族语者的差异程度更大。当先行词语义丰富度较低时,显性关系代词的使用差异程度大于先行词语义丰富度高时。Fox & Thompson (2007)提出:当先行词语义较为空乏时,关系代词省略的概率随之显著上升。这说明:当先行词语义丰富度较低时,隐性关系代词的使用倾向更大,而中国英语学习者在这种情况下仍然会更多地使用显性关系代词,出现了显性关系代词的过度使用。

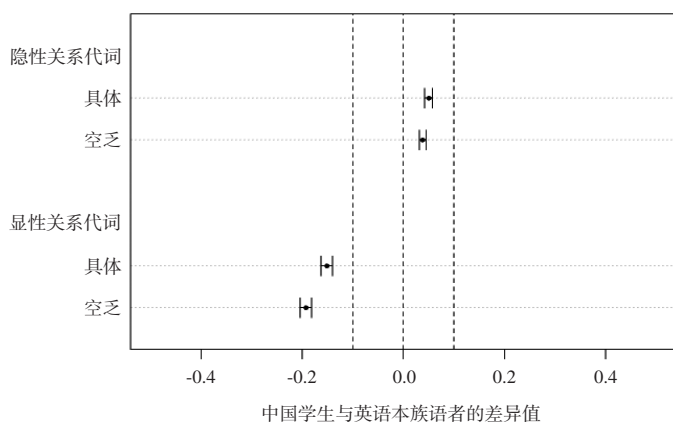


图 1. 先行词语义丰富度与关系代词交互效应

4.2.2 从句长度

双回归分析表明,从句长度与关系代词存在交互效应($F = 141.89, p < 0.001$),即显性/隐性关系代词与从句长度的复合权重存在差异⁶。图 2 显示:两个群体在隐性关系代词的使用差异区间为 $0 \sim 0.14$;其在显性关系代词的使用差异区间则为 $-0.34 \sim 0.13$ 。这说明从句长度对隐性关系代词的影响权重显著低于对显性关系代词的影响。当从句长度增大时,中国英语学习者与英语本族语者在显性关系代词使用上的差异程度迅速缩小。这一发现符合 Hawkins (1994) 和 Rohdenburg (1996) 的研究发现:句子长度与认知加工负载呈现正相关。按

照该论断,当从句长度增大,认知加工负载随之增大,两个群体对显性关系代词的使用差异程度缩小。这似乎表明,在从句长度较短时,中国英语学习者仍然较多地使用显性关系代词。换言之,中国英语学习者对显性关系代词的使用较为均衡,没有因从句长度缩短而导致显性关系代词使用量急剧下降。

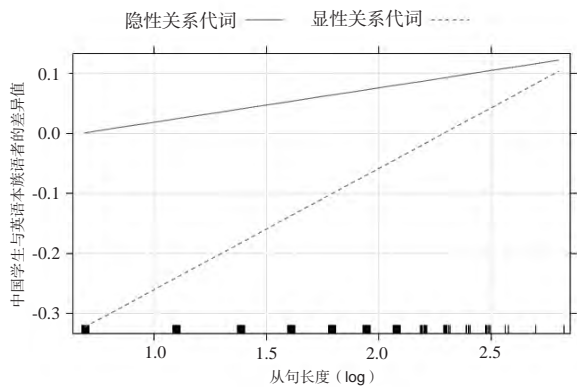


图 2. 关系从句长度与关系代词交互效应

4.2.3 先行词具体性

先行词具体性与关系代词的交互效应 ($F = 25.77, p < 0.001$) (见图 3)同先行词语义丰富度与关系代词交互效应 (见图 1) 的表现较为相似。在先行词具体性的两个变量水平 (抽象、具体) 上,显性关系代词的差异区间为 $-0.2 \sim -0.18$ 和 $-0.14 \sim -0.11$;隐性关系代词的差异区间为 $0.03 \sim 0.05$ 和 $0.04 \sim 0.07$ 。显性关系代词在两个语言群体中的使用差异显著大于隐性关系代词。如果把英语本族语者对关系代词的使用作为参照,那么中国英语学习者对显性关系代词的掌握程度显然较差。在先行词具体性的两个维度上,当先行词较为抽象时,中国英语学习者对显性关系代词的使用差异程度更大。这再次说明:在该变量水平上,中国英语学习者较之英语本族语者过度使用了显性关系代词。

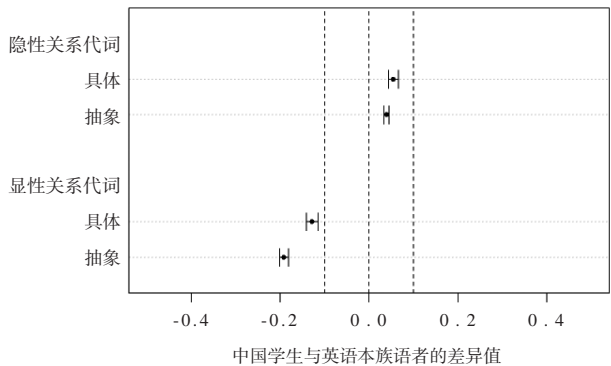


图 3. 先行词具体性与关系代词的交互效应

4.2.4 主句谓语动词类型

在主句谓语动词类型与关系代词的交互效应中(见图 4),隐性关系代词的差异区间主要置于 0~0.1 之间,只有无谓语动词的差异区间大于 0.1 (0.16~0.28);显性关系代词的差异区间则全部小于 -0.1。这表明,显性关系代词的差异程度高于隐性关系代词。在显性关系代词的主句谓语动词的六个变量水平中,认知加工负载较高的双宾语动词和无谓语动词的差异程度却最小,不及物动词的差异程度最大。这反映出中国英语学习者可能存在的一种显性关系代词使用趋势:当主句谓语动词类型的认知加工负载较大时,中国英语学习者和英语本族语者对显性关系代词的使用倾向均会相应增大,二者的使用差异随之缩小;而当主句谓语动词的认知加工负载较小时,英语本族语者使用显性关系代词的倾向会减弱,但中国英语学习者对显性关系代词的使用倾向依然较强,从而造成二者的使用差异程度增大。

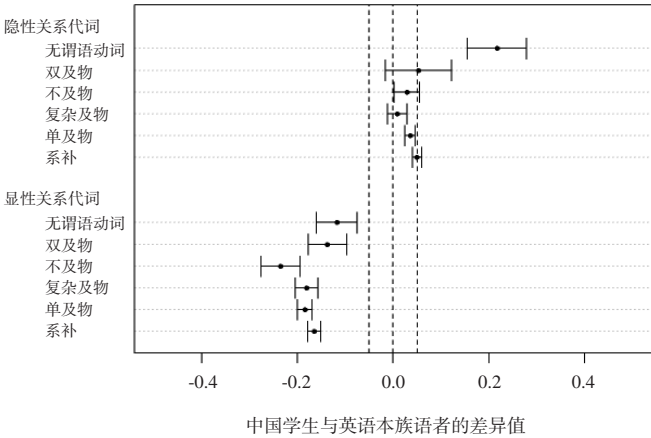


图 4. 主句谓语动词类型与关系代词交互效应

通过上述交互效应分析,本文尝试回答前文提出的两个研究问题:1)显性关系代词与隐性关系代词在中国英语学习者和英语本族语者中存在显著差异,前者的组间差异程度大于后者。这说明,中国英语学习者对隐性关系代词的使用特征掌握较好。2)显性关系代词在两个群体中的组间差异程度与交互因素的变量水平变化存在负相关性。造成这一负相关性的原因在于:在显性关系代词的使用中,当不同因素变量水平的认知加工负载发生变化时,中国英语学习者对其感知程度弱于英语本族语者(Goldberg 2019: 111-115)。换言之,在一些英语本族语者使用隐性关系代词的情况下,中国英语学习者依然使用显性关系代词。

5. 关系代词组间差异的动因分析

5.1 频次与认知固化

针对关系代词取舍双回归模型分析发现,中国英语学习者过度使用了显性关系代词,但对隐性关系代词的使用较好。造成这一使用特征的潜在动因在于隐性/显性关系代词不同的使用频次以及随之导致的认知固化差异。频次效应在语言学习中的作用受到广泛关注(Madlener 2015; Ellis *et al.* 2016),高使用频次可以在学习者头脑中形成较强的认知固化程度(Langacker 2008: 238)。作为一种语言产出数据,语料的使用频次能够表征其认知固化程度(Stefanowitsch & Flach 2017: 104-105)。

在本研究中,隐性/显性关系代词在中国英语学习者与英语本族语者中均呈现相似的分布特征,隐性关系代词的使用频次显著高于显性关系代词(见表1),但是隐性关系代词的使用类型则显著低于显性关系代词。换言之,隐性关系代词的高频次、低能产性使得其在学习者头脑中形成较高度度的认知固化;显性关系代词的低频次、高能产性则形成较低程度的认知固化。针对关系代词取舍的双回归分析(见第4节)进一步支持了上述论断,即显性关系代词的组间差异显著大于隐性关系代词。Bayley (2013: 374)发现,针对低频次、高能产性结构的习得效果要远低于针对高频次、低能产性结构的习得效果。本研究针对关系代词取舍的双回归分析进一步支持了 Bayley 的研究发现。

5.2 概率性预测与过度概括

语言学习的本质在于掌握语言结构的适度能产性(Ambridge 2013: 509),在该过程中,学习者首先要注意到语言结构的概率性和多因素特征,然后才能逐步抽象、概括不同语言结构的适用性,并最终掌握语言的适度能产性(Goldberg 2019: 84)。在这一过程中,基于认知固化的概率性预测发挥着重要作用。当特定语言结构无法在学习者头脑中形成较强的认知固化时,学习者无法准确掌握该结构的语义/句法适用性,因而出现对特定语法规则的过度概括。在本研究中,中国英语学习者对显性关系代词的过度使用便源自其对关系从句语法规则(先行词+关系代词+关系从句)的过度概括。

根据房印杰(2017)总结的隐性关系代词原型性结构:不定代词+隐性关系代词+人称代词,不难发现,在该维度下,认知加工负载较低,易于形成较高度度的认知固化,中国英语学习者与英语本族语者的使用差异较小。该原型性从句结构利于学习者从中抽象、概括其适用性,做出概率性预测。根据 Jaeger (2010)的观点,语言结构内部成分的信息值能够预测该结构中其余部分的使用倾向。具体而言,先行词的信息值能够预测关系从句中关系代词的取舍倾向。

例如,当先行词为不定代词(如 something)时,其自身的语义空乏、信息值较低,这使得学习者倾向预测隐性关系代词;当从句主语为人称代词时,人称代词的信息值同样较低,进一步增强对隐性关系代词的预测。对显性关系代词原型性结构的概括则要困难得多。显性关系代词的低频次、高能产性使得其在学习者头脑中形成的认知固化程度较低。这使得学习者无法有效归纳其使用特征,无法进行概率性预测,从而转向依托抽象的关系从句语法规则。

中国英语学习者与英语本族语者在关系代词取舍使用特征层面的差异本质上是关系代词取舍对其所造成的认知加工负载差异。当隐性/显性关系代词处于理想语境时,两个群体表现出较大的相似性。这一发现支持了 Wulff (2016: 132) 对理想语境的论述,即在对特定语言结构具有典型预测倾向的语境下,二语学习者与英语本族语者的组间差异最小。本文将理想语境进一步解释为认知加工负载的极值性,当认知加工负载对二语学习者和英语本族语者所造成的认知压力均极大或极小的时候,便形成了理想语境。例如,当关系从句长度较大时,其对两个群体均造成了极大的认知加工压力,从而使得二者均倾向使用显性关系代词。当面对非极值性语境时,两个群体表现出较大的差异性。例如,当从句长度较小时,英语本族语者倾向使用隐性关系代词,而中国英语学习者在未见不定代词充当先行词时,仍然倾向选择显性关系代词。

二语学习者的语言使用较为保守(Wulff *et al.* 2018: 118)。这种保守策略在本质上源自其自身对于二语的认知加工能力较弱,无法掌握与本族语者同等的适度能产性(Goldberg 2019)。只有面对原型性语言结构时,他们才体现出较强的概率性预测;在面对非极值性语境、非原型性结构时,他们则倾向借助抽象语法规则。概言之,其在隐性/显性关系代词从句的认知加工过程中,分别体现出基于范例的概率性预测和基于抽象语法规则两种路径。

针对关系代词取舍的双回归分析较为客观地说明:二语学习是一个不断向目标语迈进的渐进过程;二语学习者在语言产出中面临更大的认知加工负载,其与英语本族语者体现出认知加工机制层面整体的相似性以及具体认知加工权重的差异性。

6. 结语

通过基于大规模语料的双回归模型分析,不难发现:1)与英语本族语者相比,中国英语学习者过度使用了显性关系代词,但对隐性关系代词习得较好。显性/隐性关系代词不同的习得效果源自二者迥异的使用频次和能产性;2)中国英语学习者在面对隐性/显性关系代词的原型性结构时,体现出基于范例的概率性推测,在面对非原型性结构时,则呈现对显性关系代词的过度概括。语言使

用自身的多因素特征与中介语对比研究中“可比语境”概念的精细化界定均要求在中介语对比研究中使用复杂统计模型,从而有效剥离各个语言学因素,抽取二语学习者与英语本族语者的组间差异。本研究对中国英语学习者关系代词取舍特征的探讨初步论证了基于复杂统计模型的“可比语境”观在中介语对比研究中的价值。伴随大规模语料库和复杂统计模型的不断涌现,未来的中介语对比研究必将转向基于“可比语境”的研究范式,从而进一步细化统计分析的颗粒度,更加客观地检验前人的相关理论。

注 释

1. 语义类特征标注由两个标注员分别独立完成,而后计算标注一致性,计算结果显示二者一致性良好。
2. 该语料库由比利时鲁汶大学 Sylviane Granger 建立。
3. 该语料库由北京外国语大学许家金教授主持建立。
4. 该语料库由日本神户大学 Dr. Shin Ishikawa 主持建立。
5. 各个因素特征的具体标注过程请见房印杰(2016: 66-86)。
6. 本研究将从句长度全部转化为对数值。在中国本族语者的 1054 句语料中,从句长度大于 2 的语句一共只有 85 句,因此,对该区间的语句的预测可能不够准确。

参考文献

- Ambridge, B. 2013. How do children restrict their linguistic generalizations? An (un-) grammaticality judgement study [J]. *Cognitive Science: A Multidisciplinary Journal* 37: 508-543.
- Ariel, M. 1990. *Accessing Noun-phrase Antecedents* [M]. New York: Routledge.
- Bayley, R. 2013. Data analysis: Quantitative approaches [A]. In M. Salaberry & L. Comajoan (eds.). *Research Design and Methodology in Studies on L2 Tense and Aspect* [C]. Berlin: Mouton de Gruyter. 357-390.
- Diessel, H. 2004. *The Acquisition of Complex Sentences* [M]. Cambridge: CUP.
- Diessel, H. & M. Tomasello. 2005. A new look at the acquisition of relative clauses [J]. *Language* 81: 882-906.
- Ellis, N., U. Römer & M. O'Donnell. 2016. *Usage-based Approach to Language Acquisition and Processing: Cognitive and Corpus Investigations of Construction Grammar* [M]. Ann Arbor: Wiley.
- Fox, B. & S. Thompson. 2007. Relative clauses in English conversation: Relativizers, frequency and the notion of construction [J]. *Studies in Language* 31: 293-326.
- Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies [J]. *Cognition* 68: 1-76.
- Goldberg, A. 2019. *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions* [M]. Princeton: Princeton University Press.
- Granger, S. 1996. From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora [A]. In K. Aijmer, B. Altenberg & M. Johansson (eds.). *Language in Contrast: Text-based Cross-linguistic Studies* [C]. Lund: Lund University Press. 37-51.
- Gries, S. 2013. *Statistics for Linguistics with R*. (2nd edition) [M]. Berlin: Mouton de Gruyter.
- Gries, S. & S. Deshors. 2014. Using regressions to explore deviations between corpus data and a standard/target: Two suggestions [J]. *Corpora* 9: 109-136.

- Hawkins, J. 1994. *A Performance Theory of Order and Constituency* [M]. Cambridge: CUP.
- Jaeger, F. 2010. Redundancy and reduction: Speakers manage syntactic information density [J]. *Cognitive Psychology* 61: 23-62.
- Jespersen, O. 1933. *Essentials of English Grammar* [M]. London: Allen and Unwin.
- Langacker, R. 2008. *Cognitive Grammar: A Basic Introduction* [M]. Oxford: OUP.
- Madlener, K. 2015. *Frequency Effects in Instructed Second Language Acquisition* [M]. Berlin: Mouton de Gruyter.
- Mak, W., W. Vonk & H. Schriefers. 2002. The influence of animacy on relative clause processing [J]. *Journal of Memory and Language* 47: 50-68.
- Péry-Woodley, M. 1990. Contrasting discourse: Contrastive analysis and a discourse approach to writing [J]. *Language Teaching* 23: 143-151.
- Rohdenburg, G. 1996. Cognitive complexity and increased grammatical explicitness in English [J]. *Cognitive Linguistics* 7: 149-182.
- Stefanowitsch, A. & S. Flach. 2017. The corpus-based perspective on entrenchment [A]. In H. Schmid (ed.). *Entrenchment and the Psychology of Language Learning* [C]. Berlin: Mouton de Gruyter. 101-127.
- Temperley, D. 2003. Ambiguity avoidance in English relative clauses [J]. *Language* 79: 464-484.
- Wasow, T., T. Jaeger & D. Orr. 2011. Lexical variation in relativizer frequency [A]. In H. Simon & H. Wiese (eds.). *Expecting the Unexpected: Exceptions in Grammar* [C]. Berlin: Mouton de Gruyter. 175-195.
- Wulff, S. 2016. A friendly conspiracy of input, L1, and processing demands: That-variation in the language German and Spanish learners of English [A]. In L. Ortega, A. Tyler, H. Park & M. Uno (eds.). *The Usage-based Study of Language Learning and Multilingualism* [C]. Washington, D. C.: Georgetown University Press. 115-136.
- Wulff, S. & S. Gries. 2015. Prenominal adjective order preferences in Chinese and German L2 English: A multifactorial corpus study [J]. *Linguistic Approaches to Bilingualism* 5: 120-148.
- Wulff, S., S. Gries & N. Lester. 2018. Optional *that* in complementation by German and Spanish learners [A]. In A. Tyler, L. Huan & H. Jan (eds.). *What is Applied Cognitive Linguistics? Answers from Current SLA Research* [C]. Berlin: Mouton de Gruyter. 99-120.
- 房印杰, 2016, 中国学生英语限定性关系从句中关系代词取舍的多因素分析 [D]. 博士学位论文. 北京外国语大学。
- 房印杰, 2017, 关系代词取舍的共现词分析——中国英语学习者和英语本族语者的对比研究 [J], 《外语与外语教学》(3): 100-108。
- 文秋芳、梁茂成、晏小琴, 2008, 《中国英语学习者英语口语笔语语料库(2.0)》[M]。北京: 外语教学与研究出版社。
- 文秋芳、王立非、梁茂成, 2005, 《中国英语学习者英语口语笔语语料库(1.0 修订版)》[M]。北京: 外语教学与研究出版社。

收稿日期: 2018-09-18; 修改稿: 2019-04-17

通讯地址: 100876 北京市 北京邮电大学人文学院外语系(房)
100083 北京市 北京航空航天大学外国语学院(梁)

in the free recall test than English coding group, and the results were significant in both subjects with low-level and high-level second language proficiency, which is consistent with the prediction of the model. When the subjects with low-level second language proficiency retrieved abstract information across languages, the reaction time was significantly higher than intra-language retrieval, which was consistent with the prediction of the model; there was no significant difference in the reaction time of subjects with high-level second language proficiency, which did not conform to the model's prediction.

Conceptual structures modulate structural priming in L2 complex sentence production (p. 422)

ZHAO Chen (National Key Research Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou 510420, China)

JIANG Shenjun (Faculty of English Language and Culture, Guangdong University of Foreign Studies, Guangzhou 510420, China)

Two experiments are conducted to explore the relationship between conceptual structures and syntactic structures in L2 complex sentence production. Two types of primes which are similar in syntactic structures but different in the number of thematic roles (object raising sentences with infinite complements and object control sentences) were adopted to test whether an additional conceptual role would prime the participants to recall the target sentences (object raising sentences with finite complements) as infinitive complements. The results showed that the high proficiency learners tended to recall the finite complements of object-raising verbs as infinitive complements, the low proficiency learners, however, recall the infinitive complements as finite ones. The differences are discussed in terms of usage-based theories.

A comparable context-based multifactorial analysis of relativizer omission by Chinese EFL learners and native speakers of English (p. 435)

FANG Yinjie (School of Humanities, Beijing University of Posts and Telecommunications, Beijing 100876, China)

LIANG Maocheng (School of Foreign Languages, Beihang University, Beijing 100083, China)

A double regression analysis of relativizer omission by Chinese EFL learners and native speakers of English reveals that a) relativizer omission is not random but constrained by a variety of cognitive processing related factors; b) Chinese EFL learners, compared to native speakers of English, exhibit an overuse of explicit relativizer and a better command of implicit relativizer; c) the inter-group discrepancy by Chinese EFL learners and native speakers of English is closely related to ideal constraining conditions on relativizer omission with a closer similarity under ideal conditions and a greater discrepancy under non-ideal conditions. It is demonstrated that the double regression analysis guided by "comparable context" plays a significant role in teasing apart salient constraining factors and retrieving inter-group discrepancy in relativizer omission.

Effects of different explicit instructions on the learning of subjunctive mood by EFL learners: A depth-of-processing perspective (p. 447)

LI Fei (School of Foreign Languages, China University of Geosciences, Wuhan 430074, China)

SUN Yunmei (School of Foreign Languages, Huazhong University of Science and Technology, Wuhan 430074, China)

Studies revolving around the effectiveness of different explicit instructions (i.e., explicit deduction and explicit induction) have produced inconsistent findings. From the depth-of-processing perspective and employing the think-aloud protocols, this study examined the effects of different explicit instructional approaches on the learning of subjunctive mood by EFL learners. Results showed that different explicit instructions promoted differential levels of the depth of processing and the depth of processing had significant positive correlations with L2 learning outcomes. Given these, different explicit instructions had differential effects on learning outcomes.

An empirical study of the teaching effect of different Japanese grammatical systems in Chinese universities (p. 459)

CAI Yan & LIN Zhang (College of Foreign Languages, Fujian Normal University, Fuzhou 350007, China)

In Chinese universities, Japanese grammar teaching includes two grammatical systems, namely school-based grammar and Japanese teaching grammar. In the sense of pedagogy, the differences of these two systems mainly lie in the ways of thinking, vocabulary learning and verb learning. This paper provides an empirical study on the teaching effect of the two grammatical systems. The results show that there is no difference between these two systems when the teaching content contains a small number of grammatical points, simple grammatical forms and a lower level of difficulty. However, the teaching effect of school-based grammar is better than that of Japanese teaching grammar when the teaching content contains more grammatical points, more complex forms and a higher level of difficulty. These differences are closely related to the logical thinking, knowledge background and attention allocation of second language learners.