

孙晓霞 著

主语关系从句优先加工的非普遍性
The Non-Universality of Subject Preference

语言新专题研究系列

倪传斌 主编

世界
用



司

语言新专题研究系列 倪传斌主编

主语关系从句优先加工的非普遍性

The Non-Universality of Subject Preference

孙晓霞 著



世界图书出版公司

北京·广州·上海·西安

图书在版编目 (CIP) 数据

主语关系从句优先加工的非普遍性/孙晓霞著. —北京:世界图书出版公司北京公司, 2014. 7

ISBN 978-7-5100-7775-3

I. ①主… II. ①孙… III. ① 汉语 - 句法 - 研究 - 英文
IV. ①H146.3

中国版本图书馆 CIP 数据核字 (2014) 第 057951 号

主语关系从句优先加工的非普遍性

著 者: 孙晓霞

责任编辑: 武传霞

出 版: 世界图书出版公司北京公司

出 版 人: 张跃明

发 行: 世界图书出版公司北京公司

(地址: 北京朝内大街 137 号 邮编: 100010 电话: 64077922)

销 售: 各地新华书店和外文书店

印 刷: 三河市国英印务有限公司

开 本: 711mm × 1245mm 1/24

印 张: 9

字 数: 224 千

版 次: 2014 年 7 月第 1 版 2014 年 7 月第 1 次印刷

ISBN 978-7-5100-7775-3

定 价: 28.00 元

版权所有 翻印必究

(如发现印装质量问题, 请与本公司联系调换)

《语言科技文库》总序

李葆嘉

当代语言学已经进入了一个科学与技术的互补时代，信息处理水平成为衡量国家现代化水平的重要标志之一。知识世界的载体是语符系统，信息处理的根本对象是语言信息处理。与计算机的出现使得语言符号有可能成为数据处理对象相似，神经科学实验仪器设备的应用，使得在大脑神经层面探讨语言机制成为可能。这些无疑都引导语言研究走向科技化，“语言科技新思维”（李葆嘉 2001）应运而生。所谓“语言科学”包括理论语言学、描写语言学、历史语言学、应用语言学等分支学科，所谓“语言技术”指语言研究的现代技术手段，包括语言信息处理、语音实验分析，以及语言的神经、心理和行为实验分析的技术手段等。就语言信息处理而言，又可以分为语料库研制技术、知识库研制技术、知识挖掘和抽取技术、句法信息处理技术、词汇信息处理技术、语音信息处理技术、语义信息处理技术、语用信息处理技术等。

2001 年 5 月，南京师范大学文学院创办了史无前例的“语言科学及技术系”，率先迈出了从传统文科教育范型向现代科技教育范型转变的步伐。“十五”期间，南京师大“211 工程”重点学科建设项目“语言信息处理与分领域语言研究的现代化”（陈小荷教授主持），以基础平台建设、资源建设和理论探索等为主，迈出了语言科技研究的一大步。

“十一五”期间，南京师大文学院、外国语学院和国际文化教育学院联袂申报“211 工程”三期重点学科建设项目。该项目以“语言科技”为引导，以“多学科交叉、跨院系整合、开放型营运”为理念，建设具有前瞻性、原创性、成长性的语言科技高级工作平台。以典型课题的工作原理为核心，进行资源开发和系统研制，拓展语

音科技、二语习得的神经机制研究、言语能力受损儿童的语言能力研究等新方向。同时造就新一代学术领军人物和培养一批高层次复合型人才，以期形成一支高水平的交叉学科团队。该项目设计，体现了工作平台建设、理论创新、应用研究、人才培养、团队建设的学科发展一体化思路。其旨趣在于，加速语言研究从传统文科范型向现代科技范型的转变，以引领 21 世纪语言科技的新潮流。

作为新兴交叉学科项目，通过教育部组织的专家匿名评审，“语言科技创新及工作平台建设”（2008～2011）获批，总投入 1 000 万元。总体而言，这一“语言科技创新”团队，分支学科齐全，专业知识互补。涵盖了理论语言学、计算语言学、语义科技、语音科技、实验方言学、历史语言学、神经语言学、二语习得研究、话语行为语言学等领域。这一期间，项目组成员获批的国家级基金项目达 20 多项。该项目理念之前瞻、实力之雄厚、工程之浩大、经费之保障，为学界瞩目。

2008 年秋，本项目以南京师范大学语言科技研究所为实施单位正式启动。主要有三大任务：建设一个领先性的语言信息科技实验室、建立一个独创性的语言科技工作平台、撰著一套有特色的语言科技文库。

从实验室方案设计到设备招标采购，再到实验室用房改造，经过 8 个月的努力，2009 年 12 月，语言信息科技实验室建成，为语言研究从传统范型向科技范型的转变提供了基本保障。该实验室划分为实验工作区、科研工作区和管理服务区。实验工作区建有语音实验与计算室、神经认知实验与计算室、课堂话语实录室三个专门实验室。科研工作区建有语义科技工作室、语音科技工作室、方言实验工作室、知识工程工作室Ⅰ（先秦词汇）、知识工程工作室Ⅱ（中古词汇）、知识工程工作室Ⅲ（敦煌俗语言文字）、语言习得神经机制工作室、语言习得中介机制工作室，以及参研工作室。管理区服务包括办公室、管理室、编辑室和交流室。出席“语言科技高层论坛暨语言信息科技实验室落成仪式”（2009 年 12 月 14 日）的专家认为，该实验室体现了语言学跨学科研究的当代性和先进性，具有整体性、科技型、开放型三个特点，处于全国领先地位，是“语言科技新思维”的又一体现。同时认为，该实验室的科研工作涵

盖了四个二级学科、四个博士学位点，有稳定明确的研究方向，有合理的设计规划和很好的科研基础；整体设计合理，功能齐备。以教育部重点实验室建设标准衡量，很多方面超过了指标。

语言科技工作平台是基于工作原理（课题定位—理论方法—技术路线—关键技术—评估方式）而建设的高级平台。一方面，从语言信息、语言知识和语言机制三个层面，围绕典型课题进行设备配置、资源建设和软件开发；一方面，将典型课题研究与工作平台建设融为一体，依据典型课题建设的子平台应具有解决同类课题的功能。

建设语言科技工作平台的目标是要实现语言研究手段的技术化和模型化，总体设计包括三个二级平台和八个子系统。

一、语言信息工作平台 1. 语义科技工作系统（李葆嘉教授主持）：基于词汇语义—句法语义的一体化研究思路，开发“人—机交互语义标注工具”，研制“深度语义标注信息库”；研制“幼儿（2~6）日常话语跟踪语料库”，完成幼儿语义系统和话语行为分析研究。2. 语音科技工作系统（顾文涛教授主持）：研制“多语言、多语境、多语用的语音语料库”，基于声学信号分析、感知实验和数学建模，完善语音韵律理论与相关技术应用。3. 方言实验工作系统（刘俐李教授主持）：完成“网络版汉语方言有声语料库”，拟定系统的可操作性语音、词汇、语法实验模型和研究方法，进一步完善新兴交叉学科“实验方言学”。

二、语言知识工作平台 1. 先秦词汇统计与知识检索系统（陈小荷教授主持）：研制“先秦文献语料库”、“专名知识库”、“汉语词汇档案库”等，开发先秦文献自动分词算法、古籍版本异文自动发现算法、同指专名检索软件工具等，完成“先秦汉语词汇统计与知识检索”。2. 中古词汇统计与知识检索系统（董志翘教授主持）：研制“中古文献语料库”、“专名知识库”、“中古汉语词汇档案库”等，开发中古文献自动分词和标注工具等，完成“中古汉语词汇统计与知识检索”。3. 敦煌俗语言文字统计与检索系统（黄征教授主持）：研制“敦煌文献资料库”、“敦煌文献俗语语档案库”，开发相应工具，完成“敦煌文献资料与知识检索”。

三、语言机制工作平台 1. 二语习得的神经机制研究系统（倪

传斌教授主持):研制“英语受蚀词汇库”等,基于行为学、脑成像和脑电三维度模型,进行中国人英语习得与磨蚀的神经机制研究,完成“基于神经机制的英语个性化学习分析系统”。2. 二语习得的中介机制研究系统(肖奚强教授主持):研制“留学生汉语口语中介语语料库”,基于中介语理论、对比分析理论、偏误分析理论以及二语习得影响因素等,完成“留学生汉语习得的中介机制研究”。

这一工作平台,既是科技研究平台,也是人才培养平台,即一个现代化的科学研究和人才培养工作体系。

作为本项目的文本成果,《语言科技文库》包括计算语言学研究、语义语法学研究、汉语方言学研究、古代汉语学研究、语言教学与研究、语言新专题研究六个系列。其总体特征为:领域的开拓性、理论的原创性、选题的新颖性、方法的交叉性、考据的精确性、成果的应用性。在研究过程中,除了数据采集分析、资源建设和软件开发,更重要的还是要有新思路、新理论和新材料。陈小荷提出的先秦文献信息处理新方法,从先秦典籍注疏文献中挖掘出用于自动分词和词义消歧的知识,再注入已开发的古汉语分词和词性标注工具中去,所取得的前秦古籍版本异文自动发现、先秦词汇知识自动挖掘等成果均具开拓性。李葆嘉提出的语义语法学理论和话语行为理论,基于研制专用语料库或语义信息库和技术手段,开拓了语义网络建构、深度语义分析和话语行为研究等新的领域。刘俐李建构的实验方言学理论和方法,为方言学向现代科技方法的转型研究提供了新路,并取得了一系列新成果。黄征多年来从事敦煌文献及其俗词语文字研究,古代汉语学研究系列中的敦煌文献校录整理,以及敦煌写本字词考释、以古佚和疑伪经为中心的敦煌佛典词语和俗字研究、两汉声母系统研究等新见迭出。肖奚强基于汉语中介语语料库的二语习得研究,在对外汉语教学研究界已经产生了影响。钱玉莲的汉语介词与相应英语形式比较研究等专著各有亮色。倪传斌依据语言测试和认知实验等数据,从行为学、生理学和语言学三个层面分析影响中国英语学习者外语磨蚀的相关因素。刘宇红基于隐喻的理论探讨,对各类隐喻形式的结构、特性和解读规律进行了多视角的深入探讨。

《语言科技文库》所收论著,由作者在2008年12月申报选题,

2011 年始逐步完稿。系列主编审读了书稿，主要就其学术价值、章节安排、内容关联、行文表述、图表绘制等方面，提出审阅意见。此后，作者们对书稿又进行了修改和润色。《语言科技文库》的作者，大多数是具有博士学位的年轻教师。对于我们这些 20 世纪 80 年代走进语言学研究领域的而言，出版论著可能已不足为道。然而，对于年轻学者而言，其论著的出版既是几年来研究的结晶，也是对其继续探索的促进。换言之，“211 工程”重点学科建设的目的之一，就是为年轻教师搭建一个可持续发展的科研和教学平台。学科带头人的主要任务之一就是提携后进。

尽管从根本上来说，科学或学术研究是一种个人的探索行为，然而复杂问题的研究，无疑需要群体协作。“学科建设”或团队合作模式，是 20 世纪 90 年代后期出现的一个新概念。这种模式涉及总体规划、多方协调，是需要付出精力和心血的。2008 年，通过投票方式推举我担任该项目总负责时，就意识到自己成了一个“劳动班委”。2009 年，前往安徽大学拜访黄德宽教授时，曾谈到“学科负责人的任务就是规划设计，争取项目经费和提供科研设备设施”，得到黄教授的赞许。2010 年，申报江苏省高校哲学社会科学重点研究基地时，评审专家柳士镇教授提问的“作为一个交叉学科项目，各学科之间的协调是怎么考虑的，有什么做法”，可谓一语中的。作为后学，深知交叉研究之艰、学科整合之难。相关学科之间的整合协调需要借助行政机制，但凭借行政方式并非就能完成。当时的回答是，目前做到的是建成了一个可以合作研究的场所，至于学科之间的进一步沟通合作应有较长过程。有一点很明确，只有通过交叉项目，相应学科才能渗透，合作者才能逐步磨合。我们只是在一步步探索。

十一五期间的“211 工程”建设项目即将完成，但是学科建设的任务并没有结束。2010 年，“语言信息科技研究中心”被评审为江苏省高等学校哲学社会科学重点研究基地，为“语言科技”这一交叉领域注入了新的建设活力。重点研究基地建设，除了“跨院系整合、多学科交叉、开放型运行”理念，需要凸显“合作性攻关”。围绕交叉性项目，实施计算语言学、语音科技、神经语言学、语义科技等力量的联合攻关计划。只有通过全面开放以及和与国内外同

行的合作交流，才有望建成具有影响的语言科技研究、人才培养和学术交流基地。

十年前，我（2001）曾写道：“语言科技”的内涵是以理论研究为指导，以描写研究为基础，以应用研究为枢纽，促使语言研究向计算机应用、认知科学和现代教育技术领域等延伸，沟通文理工相关学科以实现语言研究过程及其成果的技术化。“语言科技”的外延为语言工程科技、语言教育科技和语言研究科技。其中，“语言研究科技”是将语言研究活动与资源建设、软件开发相结合，其目标是实现语言学自身的科技化。还应包含语言实验、数据处理这些实验语音学、神经语言学研究的科技手段。

虽然语言学家不可能也不必要都转向语言计算或实验研究，尽管描写、考据和内省始终是最基本的方法，但是具有一定的语言科技意识却非常必要。语言学家只有了解有哪些可供利用的资源、软件或仪器，才能提高其研究深度、精度和效率。语言学家也只有了解到信息处理的语言研究需求，才有可能为之提供可资应用或参考的基础成果。“语言科技”是21世纪语言学研究的潮流。

此为出版缘起。是为总序。

2011年8月谨识于南都

成 序

语言加工是一种心理现象。当代语言学研究，无论在哪一个层面，无论是语言形式、语言意义，或语言使用，都离不开对人的认知机制和心理活动的关注，都必须基于人的心智做出解释。语言学研究已经和认知、心理、心智研究紧密结合起来。

20 世纪 60 年代，当米勒（G. Miller）试图为他的朋友乔姆斯基（N. Chomsky）提供支持，以论证语言的心理特质时（用米勒的话说，乔氏理论揭示了语言的心理现实），他使用的是心理学研究方法。在实验中，他设计了从 SAAD 核心句转换到其他结构的句型，让受试进行理解加工。按照他的假设，转换生成中涉及的步骤越多，理解加工的时间就越长。尽管这项实验及其结论引起了许多争议，但是米勒的确在证实语言是一种心理活动，只不过在那个年代，他的论证只能使用间接的方法。

人的心智是最难理解的现象之一。在这个“黑匣子”里，蕴藏着许多至今还不了解的东西。我们能感知到它的能量，但是却无法直接观察（认知）到它的活动，我们只能通过间接证据进行推理。造成这种尴尬和无奈的因素很多：或许因为我们是用自己的大脑来认识自己的大脑，主体和客体纠缠不清；或许因为大脑是生物神经的，而心智则是心理的，心理的研究根本无法采用直接观察法。比克顿（D. Bickerton）曾不无讽刺地说，大脑里哪有什么思想、语言和意象，里面充其量不过是一堆神经元。大脑的活动不过是神经元的放电现象，于是就有了大脑的物理论、功能论等。大脑的本质和工作机制有那么重要吗？既然我们对此无能为力，研究它的功能不就是研究它本身吗？比如把人脑移植到一个容器里（“瓮中之脑”假说），然后将其外接到显示屏上，或移植到计算机里，然后对它们发出一加一等于几的指令，得到的结果不是一样的吗？难道我们不是把计算机也叫“电脑”吗？我们不是按照这个原理模拟人工智

能吗？

也有人不同意这种心智一元论的观点。不言而喻，大脑是物质的，物质是具有展延性的实体。但是心智则是非展延性的实体，不占有任何空间。我们通过感官获取了经验感受，这些主体经验构成了我们的心理世界。笛卡尔（R. Descartes）早在 17 世纪就提出了二元的“身心交感论”。当代哲学家查尔姆斯（D. Chalmers）秉承笛卡尔衣钵，提出了区分心智（意识）的容易问题和困难问题。容易问题可以通过物理主义的还原论方法，如认知科学、神经科学等来描述解决，但是困难问题则不然。回答心智的困难问题，我们只能回到哲学上来，通过心智哲学，建立起新的心理—物理原则，从而把物理过程的属性与主体经验的属性联系起来。

无论是一元的还是二元的，心智研究发展到今天，认知科学家和心智哲学家已经达成一个共识，那就是——研究心智必须研究语言。“语言是心智的一扇窗口。”（Language is the window to the mind. 注意 window 前面是定冠词。）人的主体经验意识就是语言活动。除此之外，难道还有别的途径来研究心智吗？

随着语言研究的深入和科学技术的发展，神经语言学、神经影像学、电生理等研究方法被引入语言学研究领域，特别像 PET（正电子发射断层扫描术）、fMRI（功能性核磁共振技术）、ERP（脑事件相关电位技术）等，已被越来越广泛地应用于语言加工的研究中。尽管这些新技术研究的是大脑的神经活动，尽管它们提供的仍然是心智或语言运作的间接证据，但从米勒到单纯心理研究发展到探究认知的大脑神经基础，这就给我们对心智的认识提供了新的、非语言的佐证。目前国内外许多学者都在致力于此类研究，并已取得丰硕成果。

孙晓霞博士就是运用 ERP 技术来探讨汉语是否遵循主语关系从句加工的优先性，论证学界具有争议的主语关系从句优先加工的共性是否对类型学各异的语言都具有普遍性的问题。这无疑是语言与心智研究的有益尝试。在我们还无法解开人类心智的神秘面纱，无法直面其机制与本质的今天，这种研究具有重要价值。

关系从句是心理语言学中重要的研究领域。它在句法上具有复杂性，在类型学上具有多样性。选择汉语作为研究对象，是因其在

类型学上独一无二的特征。汉语是迄今所知的唯一结合了主谓宾语序和中心词后置双重特征的语言，为检验、修正或补充句子加工理论提供了丰厚的条件。已经具有实证研究数据的语言，如中心词居前的英语、德语、法语、荷兰语，中心词居后的日语、韩语等表明主语关系从句的加工具有普遍优势。如真像孙晓霞博士的研究所论证的那样，被试在对汉语不同类型的关系从句进行加工时，与其他的语言不同，其特点是宾语关系从句具有加工优势，主语关系从句优先加工并不具有普遍性，那么在某种意义上来说，我们也许就真的通过脑电波接近了语言加工甚至是心智活动的本质。

2008年至2011年，孙晓霞在东北师范大学英语语言文学专业攻读博士学位。2009年至2010年，在美国亚利桑那大学毕弗（T. G. Bever）教授主持的心理语言学实验室从事研究，她的博士论文也是在亚利桑那大学的一项研究成果。毕业后，到南京师范大学外国语言文学博士后流动站从事博士后研究，同时在德国科隆大学系统音乐学专业攻读第二个博士学位。近年来，在核心期刊上发表了十余篇论文。我对青年学者这种孜孜不倦的追求和旺盛精力钦佩不已。

本书在博士论文基础上修改而成，即将付梓，专此祝贺。

是为序。

成晓光

2013年11月26日

List of Abbreviations

ABS	absolute
ANOVA	analysis of variance
ASP	aspect marker
BOLD	blood-oxygen-level dependent
CL	classifier
CREL	correlative
CWO	canonical word order
DLT	dependency locality theory
DO	direct object
e	empty categories
EEG	electroencephalogram
ELAN	early left anterior negativity
ERG	ergative
ERP	event-related potential
FGD	filler-gap domain
fMRI	functional Magnetic Resonance Imaging
GEN	genitive marker
GES	Geodesic EEG System
GSN	Geodesic Sensor Net
NPAH	noun phrase accessibility hierarchy
OO	object-modifying, object-extracted
ORC	object-extracted relative clause
OS	object-modifying, subject-extracted
PART	participle
PET	positron emission tomography
RC	relative clause
REL	relativizer
ROI	regions of interest
RT	response/reading time
SO	subject-modifying, object-extracted
SPS	syntactic positive shift
SRC	subject-extracted relative clause
SS	subject-modifying, subject-extracted
SU	subject
t	trace
TO	topicalized, object-extracted
TS	topicalized, subject-extracted

Contents

Chapter One: Introduction	1
1.1 Motivation of the study	1
1.1.1 The importance of studying relative clauses	1
1.1.2 Universal processing preference and Current gaps	2
1.2 Aims and research questions of the study	4
1.3 Structure of the book	5
Chapter Two: Typology of relative clauses	7
2.1 Types of relative clauses	8
2.1.1 External-headed RCs vs. Internal-headed RCs vs. Correlative RCs	8
2.1.2 Postnominal/head-initial RCs vs. Prenominal/head-final RCs	9
2.1.3 Subject-extracted RCs vs. Object-extracted RCs	9
2.1.4 Subject-modifying RCs vs. Object-modifying RCs	10
2.2 Word order type and RC type	11
2.3 Typological characteristics of English relative clauses	14
2.3.1 Restrictive vs. non-restrictive clauses in English	14
2.3.2 Derivation of English RCs	15
2.4 Typological characteristics of Chinese relative clauses	16
2.4.1 Typological characteristics of Chinese	16
2.4.2 The Uniqueness of Chinese RCs	20
Chapter Three: Sentence Processing Models	24
3.1 Universal subject preference	24
3.1.1 Noun Phrase Accessibility Hierarchy	24
3.1.2 Experience/Frequency-based accounts	25
3.2 Working memory accounts (Gibson, 1998; 2000)	26
3.2.1 Storage Resources	27
3.2.2 Integration Cost	29

3.2.3	A brief comment: a bridge connecting structural theories and processing theories	31
3.3	Structure-based theories	33
3.3.1	Filler-Gap Domain (FGD) (Hawkins, 1999; 2004)	34
3.3.2	Phrase-Structural Distance Hypothesis (O'Grady, 1997)	38
3.4	Word order theory	40
3.4.1	Canonical word order vs. non-canonical word order	40
3.4.2	Prediction about the English RC processing	40
3.4.3	Prediction about Chinese RC processing	41
3.5	Pragmatic functional perspective shift	41
3.5.1	Subject as a default perspective of readers	42
3.5.2	Prediction about English RC processing	42
3.5.3	Prediction about Chinese RC processing	43
3.6	An evaluation of predictions about English and Chinese RC processing	45
Chapter Four: Previous studies and findings		47
4.1	Processing relative clauses	47
4.1.1	Basic observations	47
4.1.2	Neurophysiological evidence	48
4.1.3	Neuroanatomical evidence	49
4.2	Consistent results of universality from English RC processing	50
4.2.1	General remarks	50
4.2.2	Eye-movement monitor	51
4.2.3	Self-paced reading	52
4.2.4	Event-related potential studies	54
4.2.5	Brain imaging studies using fMRI and PET	56
4.2.6	Lesion studies: Aphasia	59
4.3	Mixed results from Chinese RC processing	60
4.3.1	Self-paced reading	60
4.3.2	A maze task—mixed results	72

4.3.3	Corpora study—supporting universal processing preference	74
4.3.4	Event-related potential studies	78
Chapter Five: An ERP experiment on Chinese RC Processing		80
5.1	ERP methodology	80
5.1.1	An overview of ERP	80
5.1.2	Geodesic EEG System (GES)	81
5.1.3	Software-relating ERP experiments	83
5.1.4	ERP language-related components	84
5.2	The ERP experiment design on Chinese relative clause processing	93
5.2.1	Aims	93
5.2.2	Participants	93
5.2.3	Materials and task	94
5.2.4	Design	96
5.2.5	Procedure	97
5.2.6	Electrophysiological recordings and preprocessing	97
5.2.7	Data analysis	98
5.3	Results	101
5.3.1	Behavioral results	101
5.3.2	Spatial Aspects of the ERP data	101
5.3.3	Statistical results	102
5.4	Discussions	135
5.4.1	Testing prior results of processing asymmetry	136
5.4.2	Involvement of topicalized sentences	139
5.4.3	Minimal pair analysis	142
5.4.4	Evaluation of sentence processing models	145
5.5	Conclusions	149
5.5.1	Non-applicability of universal subject preference to Chinese	149
5.5.2	General phenomenon of object relatives preference in Chinese	149
5.5.3	Non-existence of a universal processing model	150

5.5.4 Other findings contributing to a comprehensive model	150
Chapter Six: Concluding remarks	153
6.1 General findings	153
6.2 Significance of the study	154
6.3 Limitations of the study	155
6.4 Suggestions for future research	156
References	158
Appendix: Experimental Materials	175
后记	183