

LANGUAGE TEACHING | 语言教学研究文库 |

张 乐◎著

**基于语料库的
学术文本篇章性句干研究：
型式、意义及跨语言等值
(英文版)**



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

LANGUAGE TEACHING | 语言教学研究文库 |

基于语料库的学术文本 篇章性句干研究：型式、 意义及跨语言等值

(英文版)

张 乐 著



上海交通大学出版社

内容提要

篇章性句干指在学术英语中反复出现、实现评价和组织功能的半固定句级序列,是短语学领域里较为新颖的概念和语言实体。全书紧密围绕两大话题。其一,从形式、意义、功能三个方面提出了篇章性句干的界定标准。其二,借助可比语料库和平行语料库,确立了一部分篇章性句干的汉语等值语,是前一个话题在对比短语学视角下的拓展和延伸。本书有助于专业人员进一步认识半固定句级评价语在学术文本中的价值和英汉学术文本在表达态度意义方面的异同。研究成果对语言教学研究、对比语言学研究、翻译研究、词库建设等有一定的启示意义。

图书在版编目(CIP)数据

基于语料库的学术文本篇章性句干研究:型式、意义及跨

语言等值/张乐著. —上海:上海交通大学出版社,2015

ISBN 978-7-313-13924-5

I. ①基… II. ①张… III. ①英语—短语—研究

IV. ①H314.3

中国版本图书馆 CIP 数据核字(2015)第 240468 号

基于语料库的学术文本篇章性句干研究: 型式、意义及跨语言等值

著 者:张 乐

出版发行:上海交通大学出版社

邮政编码:200030

出 版 人:韩建民

印 制:虎彩印艺股份有限公司

开 本:787mm×960mm 1/16

字 数:246 千字

版 次:2015 年 10 月第 1 版

书 号:ISBN 978-7-313-13924-5/H

定 价:49.00 元

地 址:上海市番禺路 951 号

电 话:021-64071208

经 销:全国新华书店

印 张:18

印 次:2015 年 10 月第 1 次印刷

版权所有 侵权必究

告读者:如发现本书有印装质量问题请与印刷厂质量科联系

联系电话:0769-85252189

序

张乐博士所著《基于语料库的学术文本篇章性句干研究：型式、意义及跨语言等值》即将付梓，我感到由衷的喜悦！数年的寒窗之苦和漫漫的治学之路终于结出了硕果，实在可喜可贺！

张乐这部书所讨论的篇章性句干(textual sentence stem, TSS)指的是英语学术文本中实施态度表达和篇章组织的高频半固定句级序列。这个概念最初源自 Pawley & Syder(1983)提出的“词汇化句干(lexicalized sentence stem, LSS)”。根据 Pawley & Syder,人类语言交际所使用的相当一部分小句是作为整体存储于大脑之中的。其中有些小句已成为代表社会惯例、表达程式化意义的标准形式和最小单位。这些小句便是 LSS。需要注意的是, LSS 有别于一般所说的成语或习语。大多数 LSS 都具备一定程度的标准化、惯用语化及任意性,是一种半固定短语。

LSS 无疑是短语学领域的重要概念和实体之一。但与此同时, LSS 仍有一系列悬而未决的问题亟待研究,例如,如何判定它们的惯例化程度?如何确定它们的变异范围及其规律?它们所实现的交际功能适用于哪些语境?在其他语言中是否存在类似的词语手段?等等。Pawley & Syder 并未详述这些问题。而眼前这部著作是张乐运用语料库语言学的新视角和新方法,对 TSS——一类特殊的 LSS——所作的有益探讨。

在我看来,这部著作有如下几个鲜明特点。

一、本书依循 John Sinclairian corpus linguistics 的框架。具体来说,是 Sinclair 提出的“成语原则”(1991)指导着本书的探索。该原则认为,语言使用者有相当数量的半构筑词组可供使用。这些短语构成了多词序列的意义单位,在语言组织中占主导地位。相应地,在跨语言视角下,张乐采用的是 Sinclair 倡导的语料库驱动的翻译

对等研究方法。其核心主张为，观察对比词项的文本环境是确立语言间等值关系的主要途径。总的来说，理论框架明确，有理有据。

二、方法上，张乐兼用平行语料库和英、汉可比语料库。平行语料库汇聚了大量翻译实例，可比语料库则集合了英汉语最自然、地道的语言使用特征。二者各有千秋，相得益彰。以这三种语料库为基础，张乐共检查了 19 个型式、45 个篇章性句干、408 个词汇序列（类符数）和 10,821 个具体实例（形符数），并为 24 个篇章性句干确定了汉语等值语。工作量巨大，数据翔实，具有极高的研究效度。

三、虽然篇章性句干这个术语并非张乐首创^①，但他在前人基础上，对其形式、意义、功能等特征作了比较全面而又深入的阐述。可以说，这是第一次对 TSS 作出的相对完整的工作定义和特征探索。其中，对 TSS 的功能分类和解释是全书重要而富有特色的贡献。

四、从以往文献来看，有关 TSS 的英汉短语对比研究尚无明确报道。从这个意义上说，张乐利用单语、双语语料库来探索 TSS 的汉语等值语，确属一种创新之举。此外，张乐描述了篇章性句干及其汉语等值语在词汇、语法、意义、功能等层面上的共选特征，也具有重要的理论与实践意义。

五、讨论学术论文语言特征时，我们经常提及“客观”、“公正”、“中立”，却不愿多谈“态度”、“情感”、“个性”这样的词语。而这部著作却无疑佐证了另一趋势。在撰写学术论文时，专业人员往往有多重情感诉求：凸显作者权威，重视读者意见，避免同行苛责，指引读者认同作者观点，邀请读者参与研究过程，等等。张乐的研究表明，篇章性句干在实现这一系列人际功能方面起到了至关重要的作用。

综上所述，本书针对篇章性句干这一新颖概念进行了积极探索，得出了一系列有价值的结论，对学术英语教学、词库建设、机器翻译等领域均有一定的启示意义。

张乐是我指导的一名博士研究生。2008 年，由我主持的国家社科

① 根据我们掌握的文献，Granger & Paquot(2008)最早使用了 textual sentence stem 这一术语，但仅仅只作了简要定义，没有深入讨论。

基金项目“基于平行语料库的英汉对应意义单位研究”(07BYY004)正处于数据收集阶段。张乐一入校即投身于“政治学子库”和“科技子库”的建库工作,为之做出了宝贵的贡献,同时他也开始关注英汉学术语篇的元话语手段,并具有浓厚的兴趣。可以说,本书是张乐近年来学术积累和沉淀的集中体现。另一方面,语料库驱动的短语学以及句干研究也在持续发展,新的方法和思想不断涌现。这部著作的出版既展示了张乐一个阶段内的研究成果,也揭示了研究征途上的又一个起点。希望张乐以此为契机,坚持语料库驱动的短语学研究路线,问津新的问题,开拓新的路径,尝试新的方法,向新的目标迈进,取得新的成果。作为青年学者,张乐所述难免疏漏和不当之处,诚望学界同仁不吝指正,扶助后学。

是为序。

卫乃兴

2015年7月于北京

前 言

本书是我在博士论文的基础上修订完成的,主要反映我就读于上海交通大学期间的研究成果。

篇章性句干作为本书所使用的专门术语,直接受到 Pawley & Syder 于 1983 年提出的词汇化句干的启发。所谓词汇化句干,指的是存储于记忆中、以整体形式使用于日常交际的半固定小句。词汇化句干在提出之初是就一般交际文本的意义实现单位而言的,对于书面语中的小句特征考察是否同样具有理论和实践意义呢?答案是显而易见的。在 Pawley & Syder 看来,词汇化句干有着鲜明的话语社团属性:受到社会和文化制约,特定的话语社团趋于使用特定范围内的词汇化句干。同时,根据 Swales(1990:25-26)的观点,隶属于一定话语社团的专业人员有一套具有高度因循性的交流方式,包括反复使用的词语组合。因此,专业人员频繁使用的词汇化句干无疑是值得探索的短语学实体。

Pawley & Syder 是在理论驱动的短语学研究范式下提出的词汇化概念。尽管他们注意到词汇化句干的数量惊人,却难以说明人类的日常会话究竟涉及多少个常用句干,也无从确定哪些句干比其他句干“更常用”。本书试图在 Sinclair 的“成语原则”(1991)指导下,采用他所倡导的语料库驱动的短语学研究方法来探索词汇化句干。我们认为,人们在交际中使用大量半固定预制件作为单一选择单位。唯有利用大数量语料库和自下而上的数据描述方法,方可揭示这些预制件的使用特征。这也是这本书的理论框架和技术手段。

我对篇章性句干的兴趣始源于对学术英语语料库的观察。我发现,学术英语中的许多常见句子,表面上看起来结构殊异、意义复杂、功能繁多,但通过检查语料库证据,可概括出若干隐藏于千变万化之中的规律:在学术英语语料库中高频发生,说明反复为专业人员使

用;包含主语、谓语两个基本要素及其他可选元素,从形式上看可称为“句”,却又不是完整的“句”,而是“句干”;句干的每一个组成部分都有一定范围内的具体实现形式,稳定性和多样性并驾齐驱;以各种各样的方式帮助专业人员组织语篇、表达态度。本书将这些半固定小句统称为篇章性句干——实际上是词汇化句干发生于学术语篇中的变体。学术英语中有多少例篇章性句干?它们隶属于哪些结构?其形式和意义存在怎样的关联?它们在科技英语中实施哪些有别于普通英语中的语篇功能?为何篇章性句干是专业人员不可或缺的词语手段?本书将在第四和第五章详细讨论这些问题。

如果说篇章性句干是英语学术论文中广泛使用的惯例语化用语,汉语学术论文中是否存在相对应的半固定表达手段呢?要探索这一问题,我们秉持 Sinclair(1996a, b)关于语言对等的基本观点:双语之间存在对等关系,但并非词语本身的对等,而是词语的共现文本(语境)的对等。在双语语料库还未得到长足发展之前,关于跨语言对等问题的争论难以脱离词典的束缚。然而,双语词典的主要关切是单词和固定短语、习语,难以有效呈现半固定序列的对等形态。国家社会科学基金项目“基于平行语料库的英汉对应意义单位研究”(07BYY004)试图取得一点突破:依照事先制定的标准,收集一定数量的源文本和高质量的翻译文本,通过人工和机辅方式实现短语对齐,最终借助专门研制的检索统计软件开展对比短语学研究。我们试图厘清:汉语学术语篇中有哪些英语篇章性句干的对等语?英汉对等短语从形式和功能上看有哪些异同?形式、意义、功能的共选特征如何体现于英汉对等短语上?对等关系的确立受到哪些语境因素的影响?这些是本书第六章探讨的话题。

在我敲入上述文字时,寒窗五年的点点滴滴如电影画面般掠过眼前。如今小书即将付梓,千头万绪一时不知从何说起。

七年前,我的博士生导师卫乃兴先生将我领入到语料库语言学的研究阵地中。那时,我还是一个对这一新奇事物一无所知的顽童,在黑暗中踽踽独行、胡乱摸索。先生没有嫌弃学生的愚钝,而是悉心指点我做人、做事、做学问。多年来,先生倾注了大量心血教导我沿

着 John Sinclair 等前辈的足迹兢兢业业、埋首工作,让我体会这个学科的宏大和精密。我折服于先生严谨治学的态度和永不懈怠的精神。同时,我逐渐感觉到, Sinclair 语言学总能以一个特殊的视角来观照万千变化的语言世界。从语料库中所提取的原始数据庞大无比,令人眼花缭乱。语料库学者不厌其烦地检查索引证据,统计分析概率信息,进而归纳概括。最终,平时不易察觉的语言现象和规律跃然于纸上——恰似探险:入之愈深,得之愈奇。这样的研究怎能不让人击节赞叹?这本书中的一些观点直接来自先生的指点,也有一些则是在与他的一次次讨论中酝酿而成。全书融入了先生在学术英语领域的学术卓识和智慧,我要向他表达崇高的敬意。

我的母亲为了帮助我实现学业理想,独自背负家庭重担,倾尽一切。母亲一度深受眼疾折磨,事业遭挫,却从未动摇培养我成才的念想和决心。这份恩情是支撑我在求学路上一步步前行的不竭动力,只愿这本小书能带给她一份宽慰和一丝骄傲。我的妻子杨倩不惧生活艰辛,远离最爱的父母和家乡,坚持陪伴我在上海度过艰难岁月,让我在日复一日的研究工作之余感受到家的温暖。岳父和岳母义无反顾地支持我的每一个决定,给予我们这个小家庭最大限度的宽容和理解。女儿张伶语的出世带给我无穷欢乐,是我的力量之源。

我要感谢参与我论文评阅和答辩的各位专家学者。他们是上海交通大学的胡全生教授和王同顺教授、复旦大学的曲卫国教授、华东师范大学的邹为诚教授、华东理工大学的邵志洪教授。老师们对论文提出了许多宝贵建议,令我受益匪浅,也是对我莫大的鼓舞和鞭策。

上海交通大学的同窗既是谆谆善诱的良师,也是谈心论道的益友。我们一道在东川路 800 号发愤忘食,互相关照,当中的深情厚谊是我一辈子的精神财富。语言文字工程研究所的同门在实验室中夜以继日,苦中作乐。不断迸发出来的集体智慧之光照亮着我追寻真理的小路,令我永远无法忘怀。

本书的出版得到了上海市教委高校青年教师资助计划、上海理工大学博士启动基金和上海理工大学人文社科基金(15HJSK -

YB26)的资助,在此表示诚挚的谢意。特别感谢上海理工大学外语学院的刘芹教授和上海交通大学出版社教材图书出版中心的臧燕阳主任对本书的出版提供的大力帮助。责任编辑逢亚萍老师对本书进行了细致的编辑工作,使本书增色不少,在此一并表示感谢。

学术英语中的常用句式之多,远非这本小书所能涵盖,而英汉对比的有关内容又进一步增加了研究的复杂性和艰巨性。由于水平所限,在全书完成后,深感有许多问题有待进一步研究,书中也难免存在不妥之处。我衷心希望各位读者和专家能够不吝批评,使本书的内容不断完善。

谨以此书献给即将迎来百年校庆的母校——河海大学。诚挚祝愿河海大学外国语学院的未来更加美好灿烂!

张 乐

2015年6月于上海

List of Abbreviations

CARS	Create a Research Space
CEISR	Cause-Effect Intersentential Relation
EFL	English as a Foreign Language
ESP	English for Specific Purposes
F1, 2, ...	the first function, the second function, ...
IMRD	Introduction-Method-Results-Discussion
JDEST	JiaoDa English of Science and Technology
JDPC	JiaoDa Parallel Corpus
KWIC	Key Word in Context
L2	Second Language
LS	Lexical Sequence
LSS	Lexicalized Sentence Stem
MCC	Modern Chinese Corpus
RA	Research Article
T1, 2, ...	the first translation, the second translation, ...
TSS	Textual Sentence Stem

Contents

Chapter 1	Introduction	1
1.1	Research Background	1
1.2	Objectives of this Book	4
1.3	Significance of this Book	5
1.4	Organization	6
Chapter 2	Literature Review	8
2.1	Pawley & Syder's Lexicalized Sentence Stem	8
2.2	Granger & Paquot's Textual Sentence Stem	12
2.3	Textual Sentence Stem and Evaluation	14
2.3.1	The Multifunctionality of TSSs	14
2.3.2	Research on Evaluation in Academic Writing	17
2.3.2.1	Hedging	18
2.3.2.2	Metadiscourse	20
2.3.2.3	Model of Interaction in Academic Discourse	23
2.3.3	Evaluation from a Cross-cultural Perspective	25
2.4	Corpus-driven Studies of Phraseology	28

2.4.1	The Firthian Tradition	29
2.4.2	Lexis and Grammar	30
2.4.3	Idiom Principle	31
2.4.4	Extended Unit of Meaning	33
2.4.5	Pattern Grammar; a New Approach to the Lexical Grammar	35
2.4.6	Corpus-driven Study of Cross-linguistic Equivalence	39
2.5	Summary	46
Chapter 3	Methodological Considerations	48
3.1	Research Questions	48
3.2	Corpora Used in this Book	49
3.2.1	The English Comparable Corpus; JDEST	49
3.2.2	The Chinese Comparable Corpus; MCC	51
3.2.3	The Parallel Corpus; JDPC	52
3.3	Analytical Concepts and Procedures	56
3.3.1	Defining TSS, Pattern and Lexical Sequence	56
3.3.2	Searching for TSSs	60
3.3.2.1	Cluster and N-gram	60
3.3.2.2	Phrase Frame	62
3.3.3	Describing Evaluation	65
3.3.4	Looking for Translations	69

3.3.5	Identifying Equivalents	70
3.3.6	Other Terminological Issues	71
3.4	Summary	72
Chapter 4	Overall Frequency Distribution of TSSs	73
4.1	Overall Frequency Distribution of TSSs across Patterns	73
4.2	Content Subject Patterns	75
4.3	Dummy Subject Patterns	79
4.4	Stability and Variability of TSSs	85
4.5	Summary	90
Chapter 5	Functional Characteristics of TSSs	91
5.1	Functional Taxonomy of TSSs	91
5.2	Cognitive Acts	94
5.3	Research Acts	104
5.4	Inference Acts	111
5.5	Attitude Marking	119
5.6	Hedging	126
5.7	Directive Acts	130
5.8	Intertextual Acts	135
5.8.1	Averral and Attribution	135
5.8.2	Categories of Intertextual Acts	136
5.8.3	TSS of Agentative Integral Citation	138
5.8.3.1	Tense of Reporting Verbs and	

	Evaluation	138
5.8.3.2	A Special Case: the Unspecified	
	Author	140
5.8.3.3	Two Layers of Evaluation in	
	Attribution	143
5.8.4	TSSs of Non-agentative Integral Citation	151
5.8.5	TSSs of Identifying Non-integral Citation ...	152
5.8.6	TSSs of Non-identifying Non-integral	
	Citation	156
5.9	Intratextual Acts	160
5.10	Discussion: Strategic Balancing in Academic	
	Writing	162
5.11	Summary	169
Chapter 6	Cross-linguistic Equivalence of TSSs in English	
	and Chinese	170
6.1	Preliminaries	170
6.2	Content Subject TSSs and Chinese Equivalents	173
6.2.1	pro v that	173
6.2.2	pro modal v that	175
6.2.3	n v that	178
6.2.4	n v-link that	181
6.3	Dummy Subject TSSs and Chinese Equivalents	183
6.3.1	Sentence-medial Chinese Equivalents	184

6.3.2	Sentence-initial Chinese Equivalents	189
6.3.2.1	<i>it</i> v-link v-ed that	189
6.3.2.2	<i>it</i> v-link adj that	193
6.3.2.3	<i>it</i> v-link adj to v that and <i>it</i> modal <i>be</i> v-ed that	195
6.3.3	Sentence-final Chinese Equivalents	196
6.4	Translation Units and Equivalence	201
6.4.1	Equivalence of Lexical Sequences	202
6.4.2	Equivalence of Patterns	202
6.4.3	Equivalence of TSSs	205
6.4.4	Equivalence of Phrases around a Node Word	208
6.5	Contextualizing Description and Translation	215
6.5.1	Levels of Context and Co-selection	215
6.5.2	Identifying Functions and Equivalence in Context	218
6.6	Cross-linguistic Equivalence: Possible New Light	231
6.6.1	New Light on ESP Teaching and Learning ...	231
6.6.2	New Light on Contrastive Lexical Studies ...	234
6.6.3	New Light on Lexicon Construction	234
6.7	Summary	236
Chapter 7	Conclusion	238
7.1	Major Findings	238

7.2 Implications 244

7.3 Limitations 247

References 249

Index 261