

句法与信息结构界面的 优选论分析 (英文版)

An OT Approach to the Interface Between Syntax and
Information Structure

郑继正 著



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



郑继正

上海海关学院外语系副教授，2014年获上海外国语大学英语语言文学博士学位，2017年赴美国佐治亚州奥古斯塔大学访学，现为上海市外文学会会员、上海对外经贸大学英语学院校外导师。

郑继正副教授长期从事语言学、应用语言学、英汉翻译等各领域理论和实践研究，迄今为止，在国内各类期刊上发表论文9篇，其中两篇涉及优选论的论文发表于CSSCI期刊，主编或参编教材共4部，承担市级和校级重点课程建设和教改项目多项，主审或主译国际海关法律法规或文献共10余部，总字数超百万，并承担近百场的海关国际交流和培训现场翻译。目前郑继正副教授正主持教育部社科《英汉关务词典编撰与研究》项目，并负责海关英语英汉平行语料库和单语语料库建设工作。

句法与信息结构界面的 优选论分析 (英文版)

An OT Approach to the Interface Between Syntax and
Information Structure

郑继正 著



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

内容提要

本书主要基于优选论句法(OT-Syntax)的理论框架,通过建立句法和信息结构的约束条件,论述了英汉句法和信息结构的互动界面,从优选论的角度来探索英汉句法生成的规律。本书主要内容涵盖英汉句子信息结构的研究综述及其约束条件的建立、优选论句法研究综述及其约束条件的建立、优选论句法分析等,涉及汉英主语条件、名词移位(NP-movement)、话题化、悬垂话题(dangling topic)、宾语移位(object shift)和疑问词移位(wh-movement)等各种重要句式。本书创新地运用优选论句法的理论,从全新的角度对英汉不同句式的生成进行了理论探索,以剖析英汉语言的生成规律。本书适合语言研究学者、英汉语言专业硕士和博士研究生、计算机语言研究人员和广大英汉语言爱好者。

图书在版编目(CIP)数据

句法与信息结构界面的优选论分析:英文/郑继正著. —上海:上海交通大学出版社,2019

ISBN 978 - 7 - 313 - 22688 - 4

I. ①句… II. ①郑… III. ①句法—研究—英文 IV. ①H043

中国版本图书馆 CIP 数据核字(2019)第 281455 号

句法与信息结构界面的优选论分析(英文版)

JUFA YU XINXI JIEGOU JIEMIAN DE YOUXUANLUN FENXI(YINGWENBAN)

著 者: 郑继正

出版发行: 上海交通大学出版社

邮政编码: 200030

印 制: 江苏凤凰数码印务有限公司

开 本: 710mm×1000mm 1/16

字 数: 328 千字

版 次: 2019 年 12 月第 1 版

书 号: ISBN 978 - 7 - 313 - 22688 - 4

定 价: 90.00 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: 19.75

印 次: 2019 年 12 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: 025 - 83657309

前言

人类进入 21 世纪以来,科学技术日新月异,新技术层出不穷,尤其在计算机和互联网的领域,人类的技术创新高歌猛进,取得了令人眼花缭乱的成果,包括人工智能、机器翻译、大数据、云计算等等。尽管如此,在许多领域,人类的研究仍未能取得实质性的突破。例如,在人工智能领域,如何让机器人能说“人”话,或者说机器的自然语言处理仍无法完美地实现,也许“语言”是人工智能科学家要攻克的最后堡垒。

语言对人类来说,既熟悉又陌生。我们每天都在使用语言,而语言对我们来说许多方面却始终还是个谜。从古至今,人类对语言的探究一直孜孜不倦,众多哲学家、语言学家、科学家等投身其中,想揭开语言神秘的面纱。他们提出了各种不同的理论和观点来描述语言,解释语言生成的规律。其中最具代表性的便是美国学者乔姆斯基在 20 世纪六七十年代提出的转换生成语法理论。该理论将语言视为一个自足的系统,在约束条件下,通过句法的转换规则生成句子。虽然在 20 世纪末期,转换生成语法理论发展为最简方案理论,但其核心思想并未改变,短语结构的合法移动和限制条件仍然是句子结构的生成基础。

20 世纪 90 年代,美国学者 Grimshaw 首先基于转换生成语法的理论,利用音系学中优选论(Optimality Theory, OT)的分析框架,对句子的生成进行了优选论分析,取得了较好的解释效果,从而吸引了众多学者纷纷加入该研究领域,构建了优选论句法理论基础。

本书基于优选论句法的理论框架,将语言视为一个多层构造的复杂系统,认为语言研究的主要目的之一在于揭开各个界面的互动规律。在优选论的框架下,本书通过建立句法和信息结构的约束条件来探讨句法和信息结构的互动界面。本书的分析主要基于句法成分的排序,不涉及音系、形态和词汇的语义实现方式。本书采用标准 OT 句法的假设,认为语言输出的表层结构受制于普遍语

法(UG)的窄式 X' 理论,所以,不涉及 CP 和 IP 范畴以外的功能词组范畴。与标准 OT 句法理论不同的是,本书在输入系统的赋值上,采用词汇函项语法(LFG)的框架,在输入系统中纳入语法功能和信息结构的赋值。

本书认为信息结构属语言系统的一个独立层面,与句法、语义和语法功能等结构一起共同参与竞争,生成最优的表层形式。根据 Choi(1996)提出的信息结构系统,本书认为依据不同的句法表现,信息结构应包含话题(topic)、尾部(tail)、对比焦点(contrastive focus)和信息焦点(presentational focus)。话题以认知的方式定义为知识库中的链接和参照点,在句法结构中通常嫁接于最高投射的左侧或右侧。焦点被认为是表达形式中最具信息值且带有重音的句法成分,可位于动词词组 VP 的左右侧,不同的语言对此有不同的参数选择。同时,本书认为对比焦点常比信息焦点更易于移位。根据对信息结构成分的排序分析,本书建立了一个单独层面的信息配置结构(IS-configuration),并认为其是句法结构所需匹配的语言系统层面之一。

本书主要从汉英的主语条件、名词移位(NP-movement)、话题化、悬垂话题(dangling topic)、宾语前置和疑问词移位(wh-movement)等方面,以 OT 的形式分析信息结构和句法的界面。英语中,主语不可或缺的条件表明 OB-SPEC 的约束条件支配 IDENT-IO 的约束条件。而汉语中该两个约束条件的排序刚好相反,从而禁止其他词项在句法生成过程中的插入,以填补空缺的主语句法成分。本书假定在汉语中 OB-SPEC 的约束条件位于层级体系的底部,从而使汉语出现主语脱落的现象。通过对汉英存现句和倒装句的分析,本书认为在汉英中 ALIGN-PRES-FOCUS 的约束条件支配 SPEC-LEFT, ALIGN-SUBJECT 和 OB-SPEC 的约束条件,从而使句子中的信息焦点成分向右侧移位,包括潜在的主语成分。汉英的话题结构表明 ALIGN-TOPIC 约束条件在体系中高于 ALIGN-PRES-FOCUS,从而使焦点成分也可话题化,出现在最高投射的左侧。但在英语中,ALIGN-TOPIC 约束条件受限于 SUBJACENCY 和 PROJ-PRIN 的约束条件,从而使英语中的话题化必须符合邻接条件。相比之下,汉语的话题化无须满足 SUBJACENCY 和 PROJ-PRIN 的约束条件。本书认为该两个句法约束条件在汉语中同样位于体系结构的底部,从而使汉语中出现违背孤岛条件的话题以及话题出现在多层包孕的分句中。但汉语的话题化必须符合 IDA 的语义约束条件和 LIR 的语法约束条件,以确保句法结构能在逻辑形式(LF)中被成功释义。

本书认为 UG 的格核查理论(case-checking theory)未能充分解释名词移

位现象,尤其当带有主格形式的名词无法出现在主语位置时,格核查理论未能给出满意答案。因此本书提出结合 **ALIGN-PRES-FOCUS** 的信息约束条件来考察名词移位现象。由于该约束条件在体系中支配 **CASE** 的句法约束条件,从而阻止主格形式的名词移至主语位置,空缺主语位置只能由无语义的填充词 (**expletive**) 来弥补。同时本书认为在汉语中 **CASE** 的句法约束条件也位于体系的末端,在表层形式汉语无须进行格核查。在英语中疑问词的移位主要受 **OP-SPEC** 句法约束条件的限制,本书认为该约束条件在体系中高于 **ALIGN-PRES-FOCUS**,与 **ALIGN-TOPIC** 并列。汉语中,**OP-SPEC** 句法约束条件也位于体系的下方,但 **OB-HEAD** 和 **IDENT-IO** 的句法约束条件共同阻止了汉语的疑问词出现在 **SpecCP** 的位置,从而使汉语出现疑问词不移位的现象。

优选论句法的理论研究目前在国外吸引了众多的学者,也取得了较多的成果。由于该理论较为复杂,涉及语言学多种理论,从而在一定程度上影响了该理论的普及度。相比国外的语言学界,国内语言学界对优选论句法的研究更是不足。本书是在作者博士论文的基础上修订完成的,希望本书的出版能在一定程度上推动国内学术界对优选论句法的研究,也能吸引更多的学者加入该领域的研究队伍。

感谢上海交通大学出版社同意出版本专著,也感谢臧燕阳老师为本书的出版所付出的艰辛努力。感谢上海外国语大学李基安教授,对我在就读博士期间所给予的细心指导和照顾。感谢上海外国语大学 **Jeroen van de Weijer** 教授对我在优选论领域研究的启发、指导和无私的帮助。感谢上海海关学院对我的学术研究所给予的大力支持。最后,感谢我的家人,他们的理解、包容、鼓励和陪伴给了我随时能停靠的港湾和不断前行的动力。

由于我学术水平有限,书中的观点难免有些偏颇或错误之处,请各位同行专家和广大读者批评指正。

Abbreviations

ADJ	Adjective
ADV	Adverbial
AgrP	Agreement Phrase
AP	Adjective Phrase
AVMs	Attribute Value Matrices
CD	Communicative Dynamism
COMPL	Complement
CP	Complementizer Phrase
DP	Determiner Phrase
Eval	Evaluator
FEM	Feminine
FinP	Finiteness Phrase
FocP	Focus Phrase
ForceP	Force Phrase
GEND	Gender
Gen	Generator
GPSG	Generalized Phrase Structure Grammar
HPSG	Head Driven Phrase Structure Grammar
IP	Infinitive Phrase
IS	Information Structure
LFG	Lexical Function Grammar
LF	Logical Form
LSC	Layered Structure of the Clause

MASC	Masculine
MP	Minimalist Program
NP	Noun Phrase
NUM	Number
OBJ	Object
OT	Optimality Theory
PERS	Person
PF	Phonetic Form
PL	Plural
PP	Preposition Phrase
PRED	Predicate
PROM	Prominence
RRG	Role and Reference Grammar
SD	Structural Description
SG	Singular
SPEC	Specifier
SUBJ	Subject
TNS	Tense
TopP	Topic Phrase
TP	Tense Phrase
UG	Universal Grammar
VP	Verb Phrase

List of Constraints

* NC̣: No nasal-voiceless obstruent sequences.

* VOICED-CODA: Obstruents must not be voiced in coda position.

ALIGNFOCUS (XP, Left, VP, Right): Align the left edge of the focused constituent XP with the right edge of VP. Failed by non-aligned focus constituents.

ALIGN-FOCUS: Align focused constituents with the right edge of CP.

ALIGNTOPIC (ALGNTOP): A topic is the right or left adjoined initial or final constituent of a [+verbal] Extended Projection, except when it serves as a subject and occupies the SpecIP position.

CASE-MARKING (CASE): DPs must be marked.

COMPLEMENTLEFT (COMPLFT): Every complement is at the left edge of an X-max.

ECONOMY OF MOVEMENT (* trace/STAY): Traces are not allowed.

FULL-INTERPRETATION (FULL-INT): Lexical conceptual structure is parsed.

HEAD LEFT(HD-LFT): The head is leftmost in its projections.

HEAD RIGHT (HD-RT): The head is rightmost in its projections.

HEADLEFT (HDLFT): Every X-zero is at the left edge of an X-max.

IDENT-IO(ObsVce): An input and output obstruent should have the same value for the feature [voice].

IDENT-IO(voice): The specification for the feature [voice] of an input segment must be preserved in its output correspondent.

Lexical Insertion Rule (LIR): Insert item X under terminal node Y, where Y corresponds to the categorical features of X, and YP corresponds to the subcategorisation properties of X.

NO MOVEMENT OF A LEXICAL HEAD (NO-LEX-MVT): A lexical head cannot move.

OBLIGATORY HEAD (OB-HD): A projection has a head.

OBLIGATORY SPECIFIER (OB-SPEC): A projection has a specifier.

OPERATOR IN SPECIFIER (OP-SPEC): Syntactic operators must be in specifier position.

PROJECTION PRINCIPLE (PROJ-PRIN): No adjunction to subordinate clauses; and no movement into the head of a subordinate clause.

SPECIFIERLEFT (SPECLEFT): Every specifier is at the left edge of an X-max.

SUBJACENCY: Movement cannot cross more than one bounding node in a single step, where bounding nodes are IP and DP.

SUBJECT (SUBJ): Clauses have subjects.

TRACE GOVERNMENT (T-GVT): A trace is governed.

TRACE IS LEXICALLY GOVERNED (T-LEX-GOV): A trace is lexically governed.

ALIGN PRESENTATIONAL FOCUS (ALIGN-PRES-FOCUS) in English: A presentational focus is aligned to the right edge of the VP.

ALIGN-CONTRASTIVE-FOCUS (XP, Right Edge, YP, Left Edge): \forall XP, \exists YP such that the right edge of XP and the left edge of YP coincide, where XP is a contrastive focus, and YP is a verbal projection of the extended projection. Violated by non-aligned focus constituents.

ALIGN-SUBJECT: SUBJ should be structurally more prominent than (e. g. “c-command”) non SUBJ functions.

Contents

Chapter 1	Introduction	001
1.1	Theoretical Background	001
1.2	Research Objectives	005
1.3	Methodology and Data	006
1.4	Research Scope	007
1.5	Structure of the Book	008
Chapter 2	Brief Historical Survey of Information Structure	010
2.1	Information Structure as a Distinct Level	010
2.1.1	Bipartite Approach	011
2.1.2	Tripartite Approach	015
2.1.3	Quadripartite Approach	017
2.2	Approaching Is from Phonology and Syntax	019
2.2.1	Deriving Focus from Stress	020
2.2.2	Deriving Information Structure from Syntax	023
2.3	Multi-Dimensional Architectures	027
2.3.1	Generalized Phrase Structure Grammar (GPSG) and Head Driven Phrase Structure Grammar (HPSG)	028
2.3.2	Role and Reference Grammar (RRG)	031
2.3.3	Lexical Function Grammar (LFG)	034
2.4	Topical Research in Chinese	039

2.4.1	Topic as Part of Grammatical Structure	041
2.4.2	Approaching Topic from Functional Grammar	044
2.4.3	Approaching Topic from Syntax	047
2.5	Argument for Multi-Dimensional Construction	048
Chapter 3	Optimality Theory and Syntax	058
3.1	Basic Concepts of OT	059
3.1.1	Language as a System of Conflicting Constraints	060
3.1.2	Constraints; Universality and Violability	062
3.1.3	OT Grammar as an Input-Output Device	064
3.1.4	Architecture of an OT Grammar	065
3.2	OT-Syntax	070
3.2.1	Constraint Inventory of OT-Syntax	071
3.2.2	Lexicon and Input in OT-Syntax	075
3.2.3	Generation in OT-Syntax	083
3.3	Conclusion	094
Chapter 4	Information Structure and Its Constraints	096
4.1	Information Structure as a Separate Representation	096
4.2	Primitives of Information Structure	101
4.2.1	Traditional Binominal Primitives	101
4.2.2	Vallduví's and Choi's Primitives	103
4.2.3	Discussion on Primitives	107
4.3	Topic	111
4.3.1	Definition of Topic	112
4.3.2	Topic and Subject	117
4.3.3	Syntactic Realization of Topic	125
4.3.4	Alignment of Topic	132
4.4	Focus	140
4.4.1	Definition of Focus	140

4.4.2	Topic and Contrastive Focus	144
4.4.3	Syntactic Realization of Focus	147
4.4.4	Alignment of Focus	155
4.5	Canonical Information Structure	157
4.6	Conclusion	163
Chapter 5	The Subject Condition in English and Chinese	166
5.1	Emergence of Unmarked Construction	166
5.2	Constraint Hierarchy on the Canonical Structure	169
5.3	Subject Condition and Focus Alignment in English	174
5.4	Subject-Drop and Focus Alignment in Chinese	179
Chapter 6	Topicalization in English and Chinese	188
6.1	Topicalization in English and Chinese	188
6.2	NP Movement in English and Chinese	197
6.2.1	UG Analysis of NP Movement	198
6.2.2	An IS Approach to NP Movement under OT	203
6.3	Dangling Topics	208
6.4	Super-Raised Topic Structure	217
6.5	Object Shift	230
6.6	Conclusion	240
Chapter 7	<i>Wh</i>-Movement in English and Chinese	243
7.1	UG and Traditional OT Analysis	243
7.2	Interrogatives and Information Units	250
7.2.1	<i>Wh</i> -Movement and Information Structure in English	251
7.2.2	<i>Wh</i> -Phrase <i>in situ</i>	255

Chapter 8 Overall Conclusion	265
Bibliography	271
Index	295

Chapter 1 Introduction

1.1 Theoretical Background

Language is one of the faculties that distinguish human beings from other creatures. It helps the growth of the human brain and makes people what they are, as intelligent and social creatures. People think in language and communicate in language, either orally or orthographically. Through millions of years of evolution, language based on sound patterns won out over other means as the optimal instrument of communication.

Since ancient times, language has been of great interest to scholars. Broadly speaking, there are two orientations in the linguistic study: formalism and functionalism. Following Aristotelian logic and philosophy and Descartes' rationalism, the former inherits Saussure's advocacy to study language as "langue", or an abstract system of rules on the basis of individuals in an ideal environment isolated from the context. The formalistic approach can be characterized by way of three "autonomies": the autonomy of syntax, the autonomy of knowledge of language with respect to use of language and the autonomy of grammar as a cognitive system (Newmeyer 1998: 18 - 19). This means the linguistic system is considered as a self-contained and distinct module in the cognitive system, which can be captured by the internal investigation into the syntax alone. This orientation was pioneered by Noam Chomsky since the 1950s (Chomsky 1955, 1957) and has exerted overwhelming influence in linguistic study over more than five decades.

In contrast, the functional approach denies the autonomy of the

linguistic system. Functional linguists either perceive the linguistic grammar as dependent on semantics or pragmatics, or link the syntactic structure with the context of use (Foley and Van Valin 1984; Dik 1980; Halliday 1978, 1985, 1994; Halliday & Mathiessen 2004; Hudson 1976, *inter alia*). Though divergent in perspectives, functional grammarians uniformly argue that the linguistic grammar is not an isolated system. In fact, the traditional formalistic studies are not so self-contained as is claimed. The semantic, functional and information components are frequently resorted to in Universal Grammar (UG), like semantic roles, subject, object, topic, focus, etc. However, more often than not, UG resorts to syntactic rules and principles to explicate the syntactic structures affected by the external uses, or simply incorporate them into part of the syntactic phrase structure, such as the breakdown of CP into ForceP, FocusP, TopicP, and FiniteP (Rizzi 1997). As analyzed in Section 2.2.2 below, this extended syntactic structure fails to explain the typological differences of languages and the duplicate roles of a single syntactic constituent.

In UG, the principles and rules are considered inviolable; they are either switched off or on in the parameterization of a language. But this absolute and extreme hold may face some interesting but tricky problems. Let's take a look at the following common structures in English:

- (1) a. [_{IP}It is [_{VP}raining]].
 b. [_{IP}Is it [_{VP}raining]]?
 c. [_{IP}There I [_{VP}is [_{DP}a computer on the desk]]].
 d. [_{IP}John was [_{VP}given [_{DP}a gift [_{PP}by Mary]]]].

The surface form (1a) is a typical example of the English structure with an expletive, which is assumed to fulfill the condition of EPP in the syntax. It is argued in UG that an English clause must have a subject due to the lack of morphological agreement, and in those languages with rich morphological agreement a clause is parameterized with a null subject. Interestingly,