

● 语言学

语篇分析与机器翻译

夏家驷, 时 汶

(华南农业大学 人文学院, 广东 广州 510642)

[作者简介] 夏家驷(1956-), 男, 湖北红安人, 华南农业大学人文学院副教授, 主要从事英语
教学与翻译研究; 时 汶(1957-), 女, 河北赞皇人, 华南农业大学人文学院
副教授, 主要从事英语教学与翻译研究。

[摘 要] 机译系统应该建立在具有语篇分析的大型语料库基础之上, 语篇语料库内的原
语和目的语文本不是单向、而应是双向或多向流动的网络系统。文本的译文也应该有历时和
共时的区别, 因为互文性已是翻译的某种本质特征。

[关键词] 语篇分析; 机器翻译; 描写翻译学; 语篇特征

[中图分类号] H059 [文献标识码] A [文章编号] 1671-881X(2003)04-0488-04

2002 年 10 月在美国加州 Tibron 召开的美洲机器翻译协会 (AMTA) 会议提出的第一个中心议题就是: 为什么现行的商用机译系统都不是以数字驱动为主? 这个问题揭示了当前全文自动机译系统的窘况, 预示机器翻译工作者正从交互式翻译向全数据驱动翻译方向努力。以语料库为基础研究机译系统是当前的一个趋势, 因为随着科学技术的发展, 建立大型语料库已经不再是很困难的事, 接下来的是如何建立各种有地域、语体、乃至语篇特色的大型语料库以及如何对现有的语料库进行深层次的开发与研究。对话篇语料库的研究, 是摆在我们面前的一个突出任务。

一、机译语篇的互文性和关联性

典型的语篇应具有 7 个语篇特征: 意图性, 可接受性, 语境性, 信息性, 连贯性, 衔接性, 互文性。

互文性是符号系统的互换, 它可以是一个或几个符号系统与另外的符号系统之间的互换。它体现出文本与文本之间的一种错综复杂的关联, 任何文本都是互文本。原文文本与译文文本可以看做是两维文本, 它们虽然有着同样的作者和同样的内容, 却过着“半独立的生活”。无论对正在翻译的译员, 还是机译系统来说, 原语与译语文本并没有分离, 相反, 两者同时紧密相连, 组成了“bi-text”。“bi-text”不是两份文本, 而是具有两维性质的独立语篇, 每一维度是一种语言。更精确地说, 它是由两种或更多相关联的文本组成的思维结构。这就是机译系统的本质特征。

互文理论始提者克利斯蒂娃就强调互文性既包括文本之间空间上的组合关系, 又包括此时的文本与彼时的文本在时间上的聚合关系, 它体现空间与时间、历时与共时的统一。这个观点对机译的最大启示就是要将所有的相关文本联系起来, 无论它们是现代的还是古代的, 是风格迥异的还是类似的, 是同一种语言符号还是两种语言符号。现代的计算机具有巨大的储存能力, 只要我们给语篇合理地加上附码, 它们之间就可互相联系起来。如果附码的理论与技术得当, 机译系统就能确认各种语篇的特征, 辨认各种风格不同的译文, 回译时当然也会回到原语文本。

确认语篇特征的一个重要意义在于可以建立一部以语篇分析为基础动态网络的机译词典, 词典里面每一个词句及其译文都在特定的语篇里占据特定的位置, 都与相关语篇有着或强或弱的联系, 都有其特定的语境、功能和意义。

关联论认为任何话语都是有关联的,话语的理解过程就是寻找关联的过程,在较小的语境中找不到关联,就得借助大一些的语境,直至找出话语的关联,使话语在这个语境中具有语境效果。语篇分析方法把机译看作是一个自上而下的过程,对原语的理解要从原语所处的文化、情景语境、功能特征、意义到表层语言结构着手,分析译文在目的语文化中的属性或类型及其交际功能,判断译语怎样通过一个个的语言结构来完成目的语语篇。

二、机译语篇的语域与语境描写方式

用语料库确定和分析语篇特点比确定和分析词汇特征难得多,但对描写语言学和应用语言学十分重要,因为许多词汇和语法特征只有通过它们在语篇中的功能才能得到完满的解释。根据 Biber 等的分析,语篇分析过去没有采取语料库的方法主要有下面两种原因:第一,许多语篇特点不能被自动辨认,分析这些特点需要大量的人力物力,还必须详细考虑较大语境下的语言特点。第二,已有的商业语料库分析工具在研究语篇特点时不能发挥很大的作用。解决这些问题的办法就是使用互动式计算机程序分析语篇特点。用自动分析器也可检索整个文本中某些语法特征。

按照当前语言理论的方向制定语篇分析框架,给各种语篇特点加上附码,再用互动式计算机程序分析语篇特征,作为机译语料库的重要参数,应该是一条可行的道路。

语域分析可以确定语篇类型。情景语境三要素——语场、语旨和语式语就是语域变体,也就是说语场、语旨和语式的变化可产生不同领域、不同正式程度、不同渠道的语域。语场指的是正在发生的事情及其性质、特点。语旨指的是交际者及其特点、地位、角色、关系等。语式指的是语言在交际中所起的作用,包括交际渠道和修饰方式。

Bell 用很简洁的方式提供了语域的分析方式,让我们来看看他的例子:

The dog bit the man.

Domain: +referential

Tenor: +accessible

Mode: -participation

+channel limitation

(written to be read)

+public

对语场的描述, Halliday 列举的是母亲与孩子在洗澡时的对话:家庭洗手间,母亲帮助孩子洗澡;孩子同时在探索 1) 容器原理(把东西放进东西里面);财物的拥有权(得到属于别人的东西)。其中,话语主题、场合、目的角色等都是语场的要素。目的角色可划分为技术性和非技术性。非技术性角色,例如,建立个人关系就有一系列可能相联系的语场,诸如气候、健康、新闻等。而技术性角色与专业语场相关,例如语言学、心理学等。语场的目的性角色具有意图性。

语旨指的是交际过程中相关的参与者之间的角色关系,包括个人语旨和功能语旨。个人语旨涉及到参与者社会角色、地位和个人特点,因此,与语言交际的正式、熟悉、技术性程度有关。例如, Martin Joos 提到的五种使用变体:

Frozen: Miss Jones must keep silent.

琼斯小姐,请务必保持安静。

Formal: Kindly keep quiet, Miss Jones.

琼斯小姐,请您保持安静。

Consultative: Miss Jones, would you mind not talking please.

琼斯小姐,请不要讲话。

Casual: Better not talk now, Mary.

玛丽,最好不要讲了。

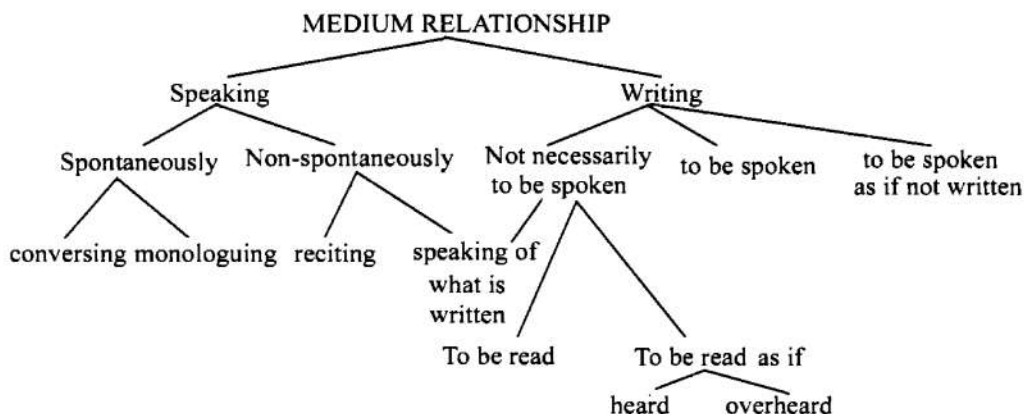
Intimate: Ssh, darling.

嘘,亲爱的!

这种分类只是一个基础,我们还可以加上 informal, familiar, impersonal, colloquial, technical 等。功能语旨与社会功能和言语角色相关,能辨别语言因何目的而使用。

语式指的是语言交际的渠道或媒介,可以按下图划分(Bell, 2001: 191):

语域的情形有时会很复杂, Hatium 和 Mason 指出某些文本中会出现语域转移(shifts of register)的现象,在他们分析的一篇一百多字的短文里,出现了至少四种语域。虽然这种情况给语域分析带来很大困难,但是他们还是详细地分析



了其中语场、语旨和语式的变化。

郭著章教授认为,译者要正确地理解原文的意思和风格就要认真分析原文的语域。他把语域标志分为四种:音层、词语层、句层和段层语域标志,而汉语的比较明显的语域标志在音、词和句三个层次中,其中以词层标志最为明显。这些现象也充分体现了语篇的语境性。

三、机译语篇的功能与语义描写方式

语篇功能指在语义层中把语言成分组织成语篇,它包括主位结构,信息结构,衔接方式等。语篇功能至少可以回答语篇特征的两个问题,即信息性和衔接性。

语义与功能是语言不可分割的两个方面,互为对方而存在。无论何种功能都是通过语言意义的实现来完成,意义的实现过程即是功能,语篇是功能和意义的载体,分析功能离不开意义,分析意义离不开功能。语义—符号学模式比较适宜于微观分析,如词汇、词组或语篇的某个截面上意义的对等关系。此模式正好可作为功能模式与语言形式的衔接桥梁。

根据美国哲学家和逻辑学家查尔斯·皮尔斯(Charles Peirce)的观点,符号被理解为一种三角关系,从它的各个组成部分的同时性出发,符号是同时作为媒介、被表征对象和释素的。皮尔斯是这样说的:“A sign, or representamen, is something which stands to somebody for something in some respect or capacity.” (1931—58) (“符号或表现体是某种对某人来说在某一方面或以某种能力代表某一事物的东西。”)用美国哲学家和行为语义学家查尔斯·莫里斯(Charles Morris)的话来说,任何符号都是由三个部分组成的实体:(1) a sign vehicle (符号载体)(2) a represent of the sign (符号的所指)(3) interpretant (释素)。据此,任何符号都不可能单独地与其它符号不相关联地出现,每个符号若要能被解释,它需要至少通过另一个符号来说明,即每一个符号都是可以被替代的,符号的可替代性或置换性是翻译的前提。莫里斯发展了皮尔斯的符号学理论,将符号学划分为三个分支:(1)句法学,研究符号与符号之间的关系;(2)语义学,研究符号与指涉对象之间的关系;(3)语用学,研究符号与使用者之间的关系。由此,语言意义也分为三种(1)言内意义,即符号相互之间所体现的意义;(2)指称意义,即符号与所指对象之间的关系所体现的意义;(3)语用意义,即符号与使用者之间的关系所体现的意义。

指称意义的翻译可从原文与译文指称的四种关系——相符,包孕,交叉和空缺关系着手,对两种语言词层之间的关系特征附码。另外,词的多义、广义、狭义、变义、反义等都应是考虑的因素。

语用意义包括联想意义,情感意义(例如词语的褒贬等),以及言外意义(intended meaning)等。何自然教授指出:“Pragmatic force, or illocutionary force in speech act theory of pragmatics, is the intended meaning for a given message.”当前语用学的许多成果可以在这里得到体现,例如明示推理交际(Ostensive - Inferential Communication),说话人是否或怎样用“显映的”(manifest)方式进行编码、表达意图的?

言内意义涉及到语言的音系、语法层、词汇层以及句子和话语层,涉及到许多与它们相关的修辞手段,例如头韵、半韵、和音、尾韵、格律、双关、一语双叙、平行结构、反复强调等。词语搭配也是其中的一个重要方面。现代语料库技术在检索语言的符号形式方面有着巨大的优势,以对英语语料库的研究为例(Thomas & Short, 2001: F25),它可以检索出:反映英语词汇在实际使用中的词频表;以英语某种句法、语义或词性为检索单位的分类信息;在英语中反复出现的习语或惯用搭配模式;在英语中反复出现的某些词汇及其特定语法结构;与某种语义场相关的词汇结构;反映某类语篇特征的词汇,语法结构和语篇结构词的使用特征;对附码语料库中两个/多个层次标记之间相互关系(如语音和句法的制约和互补)的信息;双语或多语种语料对译时的匹配和非匹配信息等等。

原语与译语的语篇形式匹配也是很重要的,两种文本都应该在音层、词层、句层、段层到话语层寻找匹配,例如平行语料库。语境分析的重点在语篇层次上,功能分析的重点在句子层次上,意义分析的重点在词汇层次上。语言形式只是这个巨大网络的一部分,情景语境、功能、意义模式是神经中枢,由他们指挥调度,完成自动翻译的伟大战役。

[参 考 文 献]

- [1] 巴尔胡达罗夫. 语言与翻译[A]. 蔡毅,等译. [M]. 北京:中国对外翻译出版公司,1985.
- [2] 陈宏薇. 新实用汉译英教程[M]. 武汉:湖北教育出版社,1996.
- [3] 郭著章. 语域与翻译[A]. 杨自俭,刘学云. 翻译新论[C]. 武汉:湖北教育出版社,1994.
- [4] 何兆熊. 新编语用学概要[M]. 上海:上海外语教育出版社,2000.
- [5] 何自然. Pragmatics and CE/EC Translation[A]. 杨自俭,刘学云. 翻译新论[C]. 武汉:湖北教育出版社,1994.
- [6] 胡壮麟,等. 语言学教程[M]. 北京:北京大学出版社,1988.
- [7] 柯平. 英汉与汉英翻译教程[M]. 北京:北京大学出版社,1993.
- [8] 李努尔. 科技英语的语域特征[A]. 朱永生. 语言·语篇·语境[C]. 北京:清华大学出版社,1993.
- [9] 秦文华. 在翻译文本新墨痕的字里行间——从互文性角度谈翻译[J]. 外国语,2002,(2).
- [10] 夏家驷,时汶. 社会符号学语义—功能翻译模式初探[J]. 华中师范大学学报(哲社版专辑),1996:10-14.
- [11] 张美芳,黄国文. 语篇语言学与翻译研究[J]. 中国翻译,2002,(3).
- [12] Bell, R. T. Translation and Translating: Theory and Practice [M]. Foreign Language Teaching and Research Press & Longman Group UK Limited, 2001.
- [13] Biber, D. et al. Corpus Linguistics[M]. Foreign Language Teaching and Research Press & Cambridge University Press, 2000.
- [14] Halliday, M. A. K. An Introduction to Functional Grammar[M]. Edward Arnold (Publishers) Ltd, 1985.
- [15] Joos, M. The Five Clocks[M]. The Hague, Mouton, 1962.
- [16] Peirce, C. Collected Papers (ed. C. Hartshorne)[M]. Cambridge, Mass.: Harvard University Press, 1931-58.
- [17] Sperber, D. & Wilson D. Relevance: Communication and Cognition[M]. Oxford: Basil Blackwell, 1986.
- [18] Thomas, J. & Short, M. Using Corpora for Language Research[M]. Foreign Language Teaching and Research Press & Longman Group UK Limited, 2001.

(责任编辑 何良昊)

Discourse Analysis & Machine Translation

XIA Jia-si, SHI Wen

(College of Humanities, South China Agricultural University, Guangzhou 510642, Guangdong, China)

Biographies: XIA Jia-si (1956-), male, Associate professor, College of Humanities, South China Agricultural University, majoring in English teaching and translating; SHI Wen (1957-), female, Associate professor, College of Humanities, South China Agricultural University, majoring in English teaching and translating.

Abstract: A data-driven Machine Translation system should be based on discourse analysis in a gigantic computer corpus, intertextualized diachronically and synchronically with all the information available from SL and TL, annotated with marks of register, situation types, function and meaning. The tagged information can be retrieved, further analyzed and added all the time, thus forming a dynamic model.

Key words: discourse analysis; MT; descriptive translation studies; textuality