

# 语料库翻译学的缘起

## 1.1 理念来源

任何科学研究都不是无源之水，无本之木，语料库翻译学亦不能例外。通常认为，语料库翻译学于1993年被作为一种研究方法正式提出，其诞生标志是曼彻斯特大学学者Mona Baker在论文集《文本与技术：致敬约翰·辛克莱》(*Text and Technology: In Honour of John Sinclair*)第二部分“语料库研究：理论与实践”(Corpus Studies: Theory and Practice)的一篇论文，文章标题为《语料库语言学与翻译研究——启示与应用》(“Corpus linguistics and translation studies - Implications and applications”)。在这篇文章中，Baker初步提出了用语料库来考察翻译语言(文本)的设想及方法，从中我们也可发现其研究理念的一些源头。该文摘要如下：

对任何以语言为主要研究对象的学科来说，语料库语言学的兴起都能为其带来一些重要启示。本文探究利用语料库开展翻译研究可能会对作为经验现象的翻译产生怎样的影响。作者认为，翻译学作为一门新兴学科，语料库语言学领域的技术与方法论发展势必对其——尤其是理论和描写分支——产生直接影响。作者比较详细地讨论了

这一影响的性质，同时也简要介绍了如何将语料库技术运用于应用型翻译研究。

(Baker 1993: 233, 笔者译)

上述摘要提及了语料库语言学的兴起、翻译是一种经验现象、翻译学是一门新兴学科以及翻译学的描写分支等内容。Baker在正文中对这些方面分别作了进一步阐释，重点论述了用语料库方法研究翻译文本的可行性，认为这一可行性主要建立在继承与发扬前人相关研究成果的基础之上。

就语料库语言学的发展历程而言，黄昌宁、李涓子(2002: 3-5)认为，早期的语料库语言学是指20世纪50年代中前期，以一定规模的语言材料为基础的语言研究，涵盖语言习得研究、方言研究、语言教学研究、句法和语义研究及音系研究等方面。这一时期，研究者通常会采集自然生成的一手书面语或口语材料，通过观察、分析和归纳得出关于特定语言现象的一般规律，研究的一大特征是注重真实语料。20世纪60年代以后，一些大规模的语料库开始建立并用于语言研究。20世纪80年代后，第二代语料库兴起，语料库语言学迅速发展。该学科注重语料的真实性、代表性和概率性，遵循“让语料说话”的描写原则，即对语料的呈现方式进行充分描写，然后在描写基础上从不同视角解释语言现象。以上都为后来将语料库应用于翻译研究提供了理论依据。

就翻译学科的发展历程而言，直到20世纪七八十年代，翻译学才逐步发展成为一门独立学科。1972年，美籍荷兰学者James S. Holmes在于丹麦哥本哈根召开的第三届国际应用语言学大会上发表了《翻译研究的名与实》(“The name and nature of translation studies”)一文。以此文为起点，Translation Studies<sup>1</sup>(翻译学，英文简称TS)这一名称逐步得到广泛认可。Holmes(1972/1987: 15)将翻译学界定为一门实证学科(an

1 又译作翻译研究，以下论及学科时统一使用“翻译学”这一名称，论及研究分支时统一使用“翻译研究”这一译法。

empirical science), 并认为其主要目标有两个:

- 1) 描写我们经验世界中有关翻译过程及翻译产品本身的各种现象;
- 2) 建立一般原则, 以解释和预测这些现象。

Holmes将翻译学划分为“纯”翻译学(“Pure” TS)和“应用翻译学”(Applied TS)两大类(参见图1.1)。

“纯”翻译学即我们通常所说的基础型研究, 又可进一步划分为描写翻译学(Descriptive Translation Studies, 简称DTS)和理论翻译学(Theoretical Translation Studies, 简称TTS), 或分别称为翻译描写(Translation Description, 简称TD)和翻译理论(Translation Theory, 简称TTh)。Holmes(1972/1987)强调, 描写翻译学尤其靠近实证研究。

之后, 以色列学者Gideon Toury进一步发展了描写翻译学。Toury(1995: 24)指出, “翻译即目标语文化中的事实”, 这一观点扩大了翻译学研究对象的范围, 也为语料库翻译研究对象的确立奠定了基础。Toury认为, 目标文本可以被作为独立研究对象, 这种脱离源文本单独讨论目标文本的做法在之前传统的翻译研究中比较少见。Baker(1993: 238-239)也认为, 基于Toury的这一观点, “源文本(在翻译研究中)所承担的就是一个刺激源或信息源的角色, 而非分析的起点”。换言之, 翻译研究中, 源文本至高无上的地位可以被打破, 研究者可以将关注重点放在目标文本上。值得一提的是, 兴起于20世纪70年代的翻译功能学派也持类似观点。Vermeer(1983)认为, 翻译文本的功能由(目标语)文本接受者的利益和期待决定, 而非由源文本的功能决定。Baker(1993: 237-238)也指出, Itamar Even-Zohar等人提出的多元系统理论(polysystem theory)将翻译文学视为文学系统中具有独立地位的一个子系统; 据此, 翻译研究的对象就可以不再局限于单个目标文本, 目标语言(文本)整体也可以作为独立于源语言(文本)的研究对象。

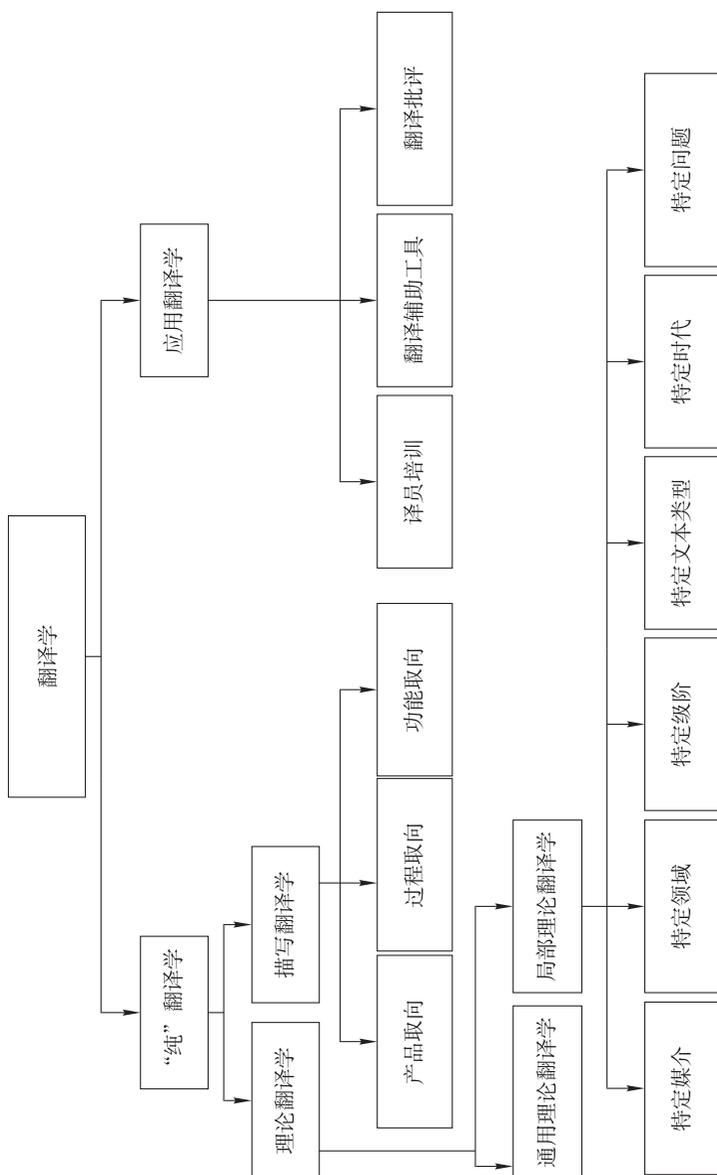


图 1.1 翻译学研究分支概览 (Holmes 1972/1987: 22)

以上这些观点扩大了翻译学研究对象的范围，提升了目标文本的地位，将目标语言(文本)整体同样作为记录各类真实交际事件的手段，这些都为语料库翻译学奠定了理念基础。相关理念主要包括：描写主义(descriptivism)、使用中的语言(language in use)、规律性(regularities or patterns)和概率(probability)，以下作一简要说明。

### 1.1.1 描写主义

语言(langue)与言语(parole)的划分奠定了现代语言学的基础。在很长一段时期内，研究者将静态的(static)语言视为主要研究对象，一定程度上忽略了动态的(dynamic)言语。理论语言学致力于建构统一的模式或理论，来解释个别语言或一般人类语言的性质，相关研究通常会提出一些公理或步骤来检验相关假说或理论，研究方法以演绎法为主；描写语言学是传统语法的继承者，致力于描写语言本身，其所依据的是实际产生的语言数据(既包括书面语也包括口语)，相关研究主要采用分析归纳法(Jackson 1982: ix)。描写的方法在20世纪20年代随着美国结构主义语言学派(也称为美国描写主义语言学派)的诞生逐步兴起，其核心理念是客观描写一手语料，避免主观臆断。20世纪50年代以后，描写语言学的理念已经被广泛接受(参见Joos 1958: v)。

描写研究不是对某一现象或其所带来的影响进行介入式研究，而是对现象本身的情况进行介绍，主要采用描述、比较、对照、分类、分析和解释等方式。其关注对象主要包括：(1)客观存在的条件或关系；(2)主流研究做法；(3)研究者所持看法、观点或态度；(4)过程进展；(5)特定研究效果；(6)发展趋势；等等。除此之外，描写研究还会关注现有情况与先前某些事件之间存在何种关系(Tavakoli 2012: 160)。

通常而言，描写研究具有如下特征：

- 1) 讨论各个变量间的关系，检验假设，提出具有普遍效度的一

般结论、原则或理论。

2) 通过明晰的描写来证明特定现象的存在, 测算该现象的频次。描写研究以问题或假设为出发点, 但对所描述的现象不作任何人为诱导或操纵。

3) 同时兼具定性研究 (qualitative research) 与定量研究 (quantitative research) 的特征。其与定性研究的相同之处在于, 二者讨论的均是自然发生的现象, 均使用一手数据或既有数据, 且不介入实验或对实验进行其他任何人为处理; 而其与定性研究的不同之处在于, 描写研究往往从事先假设入手, 考察范围较为有限, 通常采用归纳或演绎的逻辑推理法得出一般结论。从这一角度来讲, 描写研究具有定量研究的一些特点, 但其与定量研究的主要区别在于, 描写研究考察语言的途径既可是综合的 (synthetic) 也可是分析的 (analytic), 而定量研究考察语言的途径则必须是分析的。两种方法都属于以理论或研究问题为出发点的假设驱动型研究, 但描写研究中不存在对自然数据的人为操控, 而定量研究中的变量控制则是保证研究内部效度和外部效度的重要手段。

(Tavakoli: 160-161, 笔者译)

由此可以看出, 描写研究最突出的特点在于: 借助各类参数客观描述一个或多个现象的呈现方式及特征, 由此得出关于该现象的一般性结论或规律, 且避免作价值判断。也就是说, 描写研究一般不会对各个现象作出孰优孰劣的区分, 而会在客观、充分地描述现象的基础上, 分析和解释现象背后的动因。此外, 描写研究主要关注形式, 注重形式的规律性, 因为形式是客观的、可复制的、可验证的, 也可借助计算机技术自动识别并加以统计分析。

### 1.1.2 使用中的语言

像一般言语交流一样, 翻译活动同样涉及语言的使用, 且使用的通常

是两种或多种语言。通常认为,只有使用中的语言才有意义。而翻译研究讨论的对象——即使用中的语言——包括源文本所使用的语言和目标文本所使用的语言。这里涉及一个意义观问题,即对意义及意义转换的认识。纵观翻译理论的发展,不难看出,翻译研究中对意义的认识有一个发展变化的过程。Baker(1993: 236)曾指出:翻译研究话语在很长一段时间受这样一种思想主导,即源文本中的意义或信息应当完完全全地转移到目标文本中,就如同将酒从一个杯子倒入另一个杯子;直译与意译的二分法即是此种意义观的产物。

从历时角度看,翻译研究中的意义观经历了一个从静态到动态的变化过程。在以《圣经》或佛经等宗教文本为主要翻译对象的时期,源文本至高无上,译者的自由度非常有限,译者不可以随意改动源文本的意义,甚至不能改变源文本的形式。这一时期,源文本的意义被视为一种静态存在,包含于文本中;不同语言就如同一个个语言存储器,译者的任务就是将意义从一种语言存储器中取出,放到另一种语言存储器当中,意义本身不发生任何变化。因此,字对字(word-for-word)的直译成为这一时期的主流策略,意思对意思(sense-for-sense)的意译则被认为会破坏对源文本意义的传达。然而,直译的后果就是,目标文本所使用的语言不一定符合目标语读者的语言使用习惯,或者说不符合目标语读者对此类文本的阅读期待。因此,只传达静态的意义并不一定是一种成功的交际方式。

英国诗人、文学评论家、翻译家兼剧作家John Dryden提出了直译(metaphrase)、意译(paraphrase)和仿译(imitation)的三分法,其中仿译的提法为译者在翻译活动中享有一定的自由度提供了依据。按照Dryden的观点,译者不仅可以改变源文本的语言和意义,而且可以根据场合需要同时放弃二者,仅从原作中获取一些提示信息,按照自己的意愿进行翻译,从而将译作与原作从根本上区别开来(参见Robinson 2002: 172)。Dryden的三分法相对于直译与意译的二分法有一定进步,因为在实际的翻译实践中,译者所采用的策略或方法并非泾渭分明;但需要指出

的是, Dryden所讨论的主要是诗歌翻译。按照这一思路, 诗歌翻译过程中, 意义并非恒定不变, 而是会发生一定变化; 正如把酒从一个特定形状的容器倒入另一个不同形状的容器时, 酒会有一定的挥发量一样; 译者在将一种语言转换成另一种语言时, 也会产生意义的“挥发”, 这就需要译者将这部分“挥发”的意义在译文中进行创造性的弥补。这种动态的意义观认为, 译者在译作中可以赋予原作新的意义, 文本意义在从源文本到目标文本的转换过程中可以发生变化, 但这种变化必须有理有据。

翻译研究的语言学转向进一步改变了研究者对翻译中意义的认识。Jakobson (1959: 232) 指出, “无论是语言学家还是普通语言使用者, 对我们而言, 任何语言符号的意义都通过将其翻译成其他替代符号来表达”。也就是说, 要理解或认识一种符号的意义, 就必须用另一种符号来“翻译”它, 意义必须经过翻译才能被认识。据此, Jakobson (1959: 233) 提出了三种翻译类型: (1) 语内翻译 (intralingual translation), 即在同一种语言内, 将一种语言符号用另一种语言符号来阐释; (2) 语际翻译 (interlingual translation), 指的是将一种语言符号用另外一种语言符号来进行跨语言阐释; (3) 符际翻译 (intersemiotic translation), 即将语言符号用非语言符号来阐释。在应用语言学理论开展的翻译研究中, 意义被认为具有一定层级性, 包含词汇层面、短语(搭配)层面、语法层面、语篇层面、语用层面等, 而翻译就是在各个层面找寻源语与目标语之间的对应项。然而, 由于源语与目标语之间并非总是存在一一对应关系, 因而必须由译者在各个层面建立起两种语言间的对应关系。这样一来, 个体意义的简单相加并不一定等于整体意义, 而整体意义往往会大于个体意义之和。对等 (equivalence) 成了这一时期讨论翻译的核心概念工具, 而翻译就是建立两种语言在各个层面的对等关系。不过, 由于大多数情况下两种语言都不对等, 因此译者的任务就是“异”中求“同”。这一时期源文本的地位依然举足轻重, 但翻译中的意义建构转换成了格式塔 (Gestalt) 式的建构, 即整体建构。这一思想被称为整体意义观。

翻译研究的文化转向带来了更为多样的研究视角，翻译中的意义观也随之变得更为多样化；意义不再是恒定的，而是会随着文化阐释、读者解读、时代背景、目标语语言观与文化观等因素的变化而发生一定变化。换言之，对于同一个文本来说，不同文化对它的解读不同，同一文化中不同时期的读者或译者对它的解读也不同，同一文化中同一时期的不同读者群体对它的认识也不尽相同。

语料库视角下翻译研究的关注焦点是译者(包括口译员和笔译员)所使用的翻译语言(文本)，此类研究更侧重于译者本身无意识或下意识的语言行为，或者说单凭研究者肉眼无法察觉的语言行为特征及其规律性。研究者会在描写语言行为规律性的基础上，来阐释这些规律性的语言行为背后所蕴含的意义或动机。这一意义观已经超越了语言或文本本身，它是将影响翻译行为的社会、文化、历史、意识形态等各种限制性因素纳入进来，在目标语社会文化内部进行解读的语言行为意义观。

### 1.1.3 规律性

Toury (1995: 259) 指出，科学的一个本质特征就是对规律或法则的求索，即将某一领域内相关变量之间的关系尽可能地理论化。尽管这些理论的性质、地位、规范性、可行性及可接受性都会随时间推移而变化，但科学研究的首要目标始终是找寻现象背后的规律性，语言研究与翻译研究也不例外。

实际上，翻译研究在很长一段时间内仅限于归纳一般性规则，并不热衷于探究规律。将具有科学性质的语言学理论应用于翻译研究，在一定程度上提高了翻译研究的科学性，传统翻译研究中的忠实观被对等观所取代。忠实这一概念比较主观、模糊，不同的人很难对其形成统一认识；对等这一概念则主要聚焦不同语言在词汇、语法、语篇及语用等层面的对应关系，这些对应关系的梳理均建立在语言共性与个性的相互关系之上。描写翻译研究进一步推进了翻译研究对规律的探求，其目标在于描写翻译现

象，在描写基础上建立一般原则，并借助这些原则来解释和预测翻译现象。由此，借助翻译语言(文本)，找寻翻译行为的规律性(包括共性与个性)，便也成为翻译学的目标之一。

将语料库方法应用于翻译研究，从方法论方面进一步促进了翻译研究对规律性的探寻。扩大的语料规模与计算机技术相结合，可以帮助研究者探究特定类型翻译语言(文本)的某一(些)特征、特定译者对某类语言形式的个性化或规律性使用、特定时期译者群体的翻译语言整体特征、性别差异与译者翻译策略的关系、翻译文本对源文本或源语言所涉及文化形象的构建等，这些都是对翻译行为背后规律性的探究。

#### 1.1.4 概率

语料库翻译学从相关学科借鉴的另一个概念就是概率。概率是一个数学概念，通常被用来测算或度量一系列可能结果中随机事件发生的可能性。与概率相关的一个典型例子就是掷骰子游戏：骰子的六个面各有一个数字，每次掷骰子时每个数字出现的概率是 $1/6$ ，或者说概率值约为 $0.1667$ 。与此相类似的是，我们使用语言时，通常针对同一意义有多种表达方式，因而对每一种表达方式的选择就是一种概率行为。我们的选择会受各种因素的限制，不同因素的搭配与组合会促使我们作出不同的选择。翻译活动同样如此，译者在目标语中选择用某种方式来呈现源文本，也是在各种社会、文化、历史等限制性因素作用下的一种概率性选择行为。Toury(1995: 265)指出，每条关系法则通常都可以用这样一种基本形式呈现：

如果X存在，那么Y出现的可能性就较大/小。

上述法则所采用的就是一种概率化表现形式。也就是说，在一定范围内，如果就某一现象对特定变量加以控制，那么结果就会呈现出某种趋

势或倾向。基于这样一种理念, Toury(1995: 265-266)认为, 我们所需要的翻译研究不仅应当是可观察的(observational), 还应当是实验性的(experimental); 在实验性研究中, 一些变量是可控的, 因此研究过程具有可复制性(replicability)。这样的研究可以帮助研究者不断修正现有理论, 进而深入认识译者的行为以及翻译产品的呈现方式及功能。Toury(1995: 266-267)继而指出, 每一条法则都应当处于多重条件下, 其基本形式可归纳为:

如果 $X_1$ 和 $Z_1$ 存在, 那么 $Y$ 出现的可能性就比在 $X_1$ 和 $Z_2$ 条件下出现的可能性大, 也比在 $X_1$ 和 $Z_3$ 条件下出现的可能性大。

这样一来, 翻译理论最终就会表现为一系列相互关联的假设群; 研究者可以通过验证假设, 实现对翻译现象的全面描写、解释和预测。在此基础上, Toury(1995: 267-279)提出了两条翻译法则:

1) 渐进标准化法则(law of growing standardization)。即在翻译中, 源文本的文本成分(textemes)往往会被转化为目标语(或目标语文化)的语言集成分(repertemes); 与此同时, 从源文本中获得的文本关系往往会(在目标文本中)被改变, 有时甚至会被完全忽略, 文本关系转向目标语言集合提供的(更)惯常选项。

2) 干扰法则(law of interference)。即在翻译中, 与源文本构成方式有关的一些现象, 往往会被转移到目标文本当中; 文本构成方式越是被作为目标文本表述方式的影响因素, 目标文本就越会表现出受源文本干扰的痕迹。

(笔者译)

以这两条法则为出发点, 研究者可针对自己的研究目标对众多变量

进行排列组合，提出一系列假设，然后在一定范围内对假设进行检验，得出结论，再进一步修正原有假设或提出新假设；如此周而复始，一步一步实现对翻译的深入认识。例如，研究者可以提出这样一个假设：相对于同时期的原创英语小说来讲，汉语母语者翻译的中国现当代小说要比英语者翻译的此类小说标准化趋势更显著。为了验证这个假设，研究者可建立由汉语母语者英译中国现当代小说、英语母语者英译中国现当代小说和现当代原创英语小说三类文本组成的类比语料库，以一定形式化参数或具体语言项为切入点，借助语料库工具提取数据并加以统计分析，进而对此三类文本进行类比考察。考察结果可能支持原假设，也可能只是部分支持甚至完全推翻原假设。无论产生何种结果，研究者都要对其进行合理解释，以实现对这一现象的深层认识，并在此基础上修正原假设或提出新假设。

## 1.2 理论基础

作为翻译学与语料库语言学的交叉领域，语料库翻译学的理论基础自然主要来源于两方面：(1) 翻译学理论，更确切地讲是描写翻译学理论；(2) 语料库语言学理论。需要强调的是，语料库翻译学的主体依然是翻译学，语料库只是这一研究分支所采用的一种方法或手段，因此研究的重心依然放在翻译问题上，而不是单纯的量化统计与分析方面。

### 1.2.1 描写翻译学理论

描写翻译学最早由Holmes(1972/1987)提出，借由Even-Zohar、Toury等人的深入阐释，得到了进一步发展。描写翻译研究在研究对象、研究方法和研究目标等方面均与先前的规定性翻译研究有所区别。

描写翻译研究的研究对象是“目标语文化中的事实”(Toury 1995:

24);也就是说,只要目标语文化认为是“翻译”的各种现象,无论是全译,还是节译、改译,甚至是伪译(pseudo-translation)或误译,均是描写翻译研究的研究对象。这里所强调的是真实的、客观存在的翻译产品或翻译过程。规定性翻译研究必定会对现有翻译文本作出孰优孰劣式的价值判断,而描写翻译研究不关注这样的价值判断,只将重点放在对纷繁复杂的翻译现象进行客观描写上,以期在描写基础上对翻译现象作出合理解释。具体而言,描写翻译研究的任务包括:(1)描写一定范围内某类翻译现象以何种形态或方式存在;(2)根据对这些翻译现象的分类、归纳、总结与分析,建构各种变量之间的因果关系模式;(3)基于所归纳的因果关系模式,解释这些翻译现象为什么会以现有状态或方式存在或呈现;(4)在此基础上,实现对同类翻译现象的预测。总之,描写翻译研究的关注对象是特定社会文化中相同或不同时期真实存在的翻译现象。

描写翻译研究采用目标文本取向的(target-text oriented)研究方法,更关注翻译产品及其功能,主要考察翻译决策过程、翻译规范、第三类语码和翻译共性等(Shuttleworth & Cowie 1997: 39)。按照Toury的观点,目标文本的呈现方式不仅仅取决于源文本或翻译过程,更取决于目标语文化对目标文本生成的作用(Shuttleworth & Cowie 1997: 165)。也就是说,目标文化中的各种限制性因素在很大程度上决定了目标文本的呈现方式,或者说目标文本的呈现方式或其所履行的功能主要是由目标语文化赋予的。因为同一目标语文化在不同时期对翻译的认识不同,所以不同时期的目标文本也会产生差异。按照这一思路,目标文本可以被视为独立于源文本之外的翻译存在,可以被作为独立的研究对象。Baker(1993: 234)也指出,“目标文本记录了真实的交际事件,依此而论,它既不优于也不亚于任何其他语言交际事件”。因此,无论是单个目标文本还是目标文本整体,均可作为独立研究对象。

描写翻译研究的目标不是评判翻译文本的优劣,而是要通过描写找到翻译现象的规律,分析此类规律产生的原因或所蕴含的译者动机,并根据

规律建立一般原则，将其应用于对类似现象的解释与预测当中。

### 1.2.2 语料库语言学理论

语料库语言学兴起于20世纪60年代，是随着现代机读语料库的发展而产生的一个语言学分支。语料库语言学主要借助对一定规模的语料库的检索与统计分析，对特定语言现象进行量化研究，从而找寻语言使用的规律。它强调文本的真实性 (authenticity)，即研究中要以真实使用的或自然生成的书面语或口语文本而不是语言学家杜撰的语言样例为研究对象。

语料库语言学主要采用文本取向的 (text-oriented) 实证研究方法，包括语料库驱动的 (corpus-driven) 研究途径、基于语料库的 (corpus-based) 研究途径和语料库辅助的 (corpus-assisted) 研究途径。其中语料库驱动的研究途径通常采用自下而上的研究方法，从对话料的统计分析出发，归纳总结出关于语言的一般性结论；而基于语料库的研究途径通常以假设为出发点，借助特定语料库对假设进行验证；语料库辅助的研究途径则更为灵活，既让语料库直接发挥其优势，但又不完全依赖于语料库，会同时关注语料库方法无法直接观察的语言现象。语料库语言学既注重定量统计，也强调定量统计与定性分析相结合，其研究目标在于借助大规模真实语料发掘使用中的语言所表现出的规律性特征。

### 1.2.3 描写翻译学与语料库语言学理论的糅合

正是由于描写翻译学与语料库语言学在研究对象、研究方法、研究目标与研究过程等方面存在共性特征或共享界面 (见表1.1)，二者才能够交叉融合，催生出语料库翻译学这一分支学科。

表 1.1 描写翻译学与语料库语言学的特征

	描写翻译学	语料库语言学
研究对象	目标语文化中客观存在的翻译(类)语言(文本);将翻译语言或文本作为独立的研究对象	特定范围内真实使用的或自然生成的语言或文本
研究方法	目标文本取向;定量统计与定性分析相结合	文本取向;定量统计与定性分析相结合
研究目标	通过描写找到翻译现象的规律	发掘使用中的语言所表现出的规律性特征
研究过程	可观察、可复制	可观察、可复制

根据表 1.1, 在研究对象方面, 二者均关注真实文本, 将真实文本作为研究基础; 在研究方法上, 二者都采用文本取向的研究途径, 定量统计与定性分析相结合; 就研究目标而言, 二者均以挖掘所关注的文本的规律性特征为目标; 此外, 二者的研究过程均是可观察、可复制的。

语料库翻译学具有三个特点: (1) 对译文呈现方式的规模化、可视化描写。大规模可机读电子化文本从量上改变了传统翻译研究的语料数据来源, 计算机技术及各类语料库软件可以帮助研究者识别或发现过去单凭肉眼无法观察到的现象或特征。(2) 对翻译语言(文本) 各类规律性特征的定向描写。借助对语料的标注, 研究者可以精确提取词类、搭配、句法、语法甚至语义等方面的语言数据, 并以带语境的检索行(concordance) 形式来呈现检索结果, 从而有针对性地分析研究目标。这样的分析几乎可以穷尽某一语言现象在特定范围内实际使用的数据, 保证研究信度。(3) 基于语料的量化描写。语料库方法注重定量统计, 弥补了之前描写翻译研究在这方面的不足。定量统计不仅让研究结果更具有客观性和科学性, 其与定性分析的结合也是翻译研究的必然趋势。

Williams & Chesterman (2002: 48-57) 指出, 翻译研究中主要有三种基本模式: 比较模式(comparative model)、过程模式(process model)

和因果模式 (causal model)。

作为描写翻译学的延伸, 语料库翻译学仍以比较模式为主, 但语料库语境下的比较形式, 除了传统的单个目标文本与其对应源文本的比较外, 还可包括以下几种: (1) 源语言与目标语言整体的语际对比; (2) 不同目标文本之间的类比; (3) 译文与目标语语言中自然生成的文本的类比。

过程模式的语料库翻译研究集中于口译研究 (Corpus-based Interpreting Studies, 简称 CIS) 领域。研究者从认知语言学和心理语言学视角入手进行语料库口译研究, 可更多地揭示口译的语言心理过程。这一研究领域以实验诱导数据为基础, 多采用过程模式, 即借助有声思维法 (think-aloud protocol, 简称 TAP)、事件相关电位 (event-related potentials, 简称 ERPs) 等实验方式获取关于翻译过程的数据, 研究者据此推理译者头脑中“黑匣子”的情况。

因果模式的翻译研究关注三个层面: (1) 译者的认知 (translator's cognition), 包括译者的知识状态、情感状态、对待翻译任务的态度、自我认识, 甚至译者的个性和生活经历等内在因素; (2) 翻译任务 (translation task) 的外在条件, 如源文本、雇主要求、译者所用的计算机程序及词典、任务期限, 甚至翻译酬劳等因素; (3) 社会文化因素 (socio-cultural factors), 如翻译规范、翻译传统、历史及意识形态、经济目标、所涉及语言的相对社会地位等 (Williams & Chesterman 2002: 54)。所有这三个层面涉及的因素都会对最终翻译产品的生成产生影响。语料库翻译学可以借助两种途径对这些因素进行探索: (1) 借助语料库辅助方法对部分因素, 如译者的自我认识、源文本、翻译规范、意识形态等加以考察。例如, 可以将译者关于翻译的观点 (包括译本中的译者序、前言、后记等副文本, 以及译者发表的其他关于翻译的言论等) 加工成副文本语料库, 作为研究参照; 再如, 可借助特定时期的翻译文本语料库或翻译话语语料库对该时期的翻译规范和意识形态进行考察, 得出一般性结论作为参照。(2) 将这些因素作为解释语料库研究发现的重要视角。如前所述, 翻译研

研究中的语料库方法主要关注真实使用的翻译文本，使用描写的方法对翻译语言特征、翻译共性、译者风格等进行考察；除此之外，还要在描写基础上，对这些规律性现象作出合理解释，这一点尤为重要。除语言或文本因素外，译者个人内在因素、翻译任务的外在条件以及翻译行为所处的社会文化大环境都是解释翻译现象的重要视角。此外，翻译产品的功能也可作为因果模式研究的一个方面，如在考察翻译带给目标语语言的历时变化、带给翻译规范的历时变化以及特定时期的目标语翻译话语时，均可采用语料库方法。

由此可见，翻译研究中的语料库方法可用于对翻译产品、翻译过程和翻译功能作描写和解释，甚至在一定程度上作预测；它能够充分体现翻译学的实证学科性质，有助于完善其方法论体系。

### 1.3 研究性质

通常认为，在翻译学的发展过程中，其研究范式不断演变，从方法论特点上划分，这些研究范式包括语文学范式 (philological paradigm)、语言学范式 (linguistic paradigm)、哲学范式 (philosophical paradigm)、认知范式 (cognitive paradigm) 和文化范式 (cultural paradigm) 等。这些范式很多情况下都是交叉产生、共存发展的。根据不同的科学研究在研究对象、研究目标和研究方法等方面的差异，可以对其性质作不同的划分。那么，语料库翻译学的性质该如何界定呢？

1) 语料库翻译学属于描写性研究。如前文所述，语料库翻译学是描写翻译学与语料库语言学交叉融合的产物，因此与之前的规定性翻译研究相较而言，语料库翻译学属于描写性翻译研究。规定性翻译研究是指带有规约性的研究，其标志性话语包括“正确的翻译”“好的翻译”“忠实的翻译”“对等/等值的翻译”等；规定性翻译研究还注重翻译方法的实用

性，如直译与意译、归化与异化；除此之外，这类研究通常会设立一些翻译标准来衡量译文，只有符合标准的才是好的译文（参见 Shuttleworth & Cowie 1997: 130-131）。应用型翻译研究通常带有规定性研究的色彩，如翻译教学中通常会告诉受训者在翻译中应当遵守什么原则，使用什么技巧，采取什么策略，才能达到理想的翻译效果。而描写性翻译研究是指对翻译现象进行“就事论事”的客观描述，其标志性的话语包括“目标文本的呈现方式”“翻译规范”“规律性”“频次”等。例如，某一源文本在不同时期有不同的译本，研究者若采用描写性分析方法，就会借助一些参数，对每个译本的呈现方式所体现的规律性进行描述，找出不同译本间的差异，然后再根据不同时期的翻译规范对这一历时变化作出解释。由此可见，语料库翻译研究注重文本事实，重在描述现象，不以作出价值判断为目标，因此属于描写性研究。

2) 语料库翻译学属于实证性研究。翻译学领域的实证性研究通常会对一定范围内真实存在的翻译现象或行为样本进行观察和描写，研究对象可以是某个文本，也可以是某类文本整体，还可以是翻译的动态过程。这类研究通常要求研究者尽可能地减少对研究过程的人工干预，采用归纳分析方法得出总体规律（参见姜秋霞、杨平 2005: 23-24），避免对结果作孰优孰劣式的价值判断。此外，实证性翻译研究注重数据采集与分析，研究所依赖的数据从性质上可分为自然数据和诱导数据；自然数据包括语料库提供的语言数据等，诱导数据则指通过实验过程获取的数据。也就是说，研究数据不仅包括翻译产品性质的文本，还包含借助计算机软件或工具采集的音频、视频等电子数据，以及通过实验心理学和神经科学方法采集的实验数据，这些多类型、专业化语料与数据的结合进一步证明了语料库翻译学的实证性质。此外，实证性研究方法具有可复制性，即一项实验研究设计可在类似实验中反复应用，且这些实验的研究结果之间存在一定相关性。在此类研究中，研究者能够清楚地看到提出假设、界定变量、界定受试或研究对象、确定实证研究类型、采集并分析数据、讨论并解释数据等

每一个步骤(参见Neunzig 2011: 35)。这种透明性不仅保证了研究的客观性,还提高了翻译研究的科学性。

3) 语料库翻译学兼具定量与定性的研究性质。语料库翻译研究涉及数据提取与统计,因此通常带有显著的定量研究特征,如平均词长(average word length)、平均句长(mean sentence length)、类符/形符比(type-token ratio,简称TTR)、词汇密度(lexical density)、可读性(readability)、特定语言项的频次等形式参数都可借助语料库软件进行统计。这种定量研究不同于翻译学以往纯粹的定性讨论,它提高了研究的客观性和科学性,在一定程度上使研究结果更具说服力。尽管如此,在强调语料库翻译研究方法客观性的同时,也不能忽略其中的定性分析。Tymoczko(1998: 654)就反对一味强调语料库方法的客观性。她认为,在语料库的构建和任何实验、研究及调查的设计背后是研究者的直觉与判断在起作用;翻译研究中的语料库是人类思想的产物,是真实人类的产物,不可避免地反映了人类的观点、预设和局限性。因此,在设定研究目标、构建与研究目标相匹配的语料库、提出假设、讨论统计数据及解释考察结果等环节中,都应包含定性分析的成分。综上所述,语料库翻译学采取定量统计与定性分析相结合的研究方法,扬长避短,进一步完善了翻译研究方法论。

4) 语料库翻译学既可是基础型研究,也可是应用型研究。基于语料库的翻译研究不仅关注翻译产品,也对翻译过程以及翻译功能进行考察,对这三方面翻译规律性的探究,是这一领域基础型研究的主要内容。语料库翻译教学研究以及基于语料库的翻译教学平台开发与应用则属于应用型研究的范畴。变化的数据和新的对比模式能够帮助研究者重新审视研究对象,而研究过程的可观察性和可复制性以及统计手段的应用则使翻译研究更加直观、科学。此外,近年来基于翻译类语料库的国家形象建构研究还借鉴了语料库的舆情监控功能,运用语料库来讨论实际问题;基于翻译类语料库的中国对外话语体系研究也充分体现了语料库翻译学宏观层面的应

用型研究性质。从根本上说，语料库翻译学是对之前各种翻译研究方法的继承、整合与延续，是翻译研究领域的进步，其一大明显优势就是基础型研究与应用型研究并重。

## 1.4 小结

本章从语料库翻译学的界定入手，指出这一交叉分支学科的两大学术支撑领域：语料库语言学和描写翻译学。前者是这一分支学科的手段，后者则是这一分支学科的主体与目标，即借助语料数据对翻译现象进行描写，从而发现翻译行为的一般规律，并借助总结的规律对翻译现象进行解释与预测。作为一种实证性研究，语料库翻译学建立在如下理念基础上：注重描写理念；注重语言运用，关注使用中的语言，强调文本语料的真实性；借助语料数据表现出的模式化特征，探究翻译行为背后的规律；从假设出发，借助统计分析手段，通过验证假设来突出翻译行为特征的概率性。这些理念均同时体现在语料库语言学与描写翻译学当中，正是上述共性特征促成了两个研究领域的交叉，催生了语料库翻译研究这一分支领域。

从研究性质来看，语料库翻译学属于描写研究，其实证性特征也比较突出，还往往兼具定量研究与定性研究的特征。随着语料库翻译学研究范围的扩大，其关注内容已从翻译产品延伸到了翻译过程和翻译功能，其中既包括基础型研究，也包括基于语料库优势的一些应用型研究。

如今，语料库翻译学所采用的研究方法既包括基于语料库的方法，也包括语料库驱动的方法和语料库辅助的方法。在某些研究中，研究者会综合运用多种方法。在这类研究中，语料库既是研究工具，也是基础方法论之一。

语料库翻译学对翻译产品、过程和功能的全面关注，势必促使其发展成为实证翻译研究甚至实证翻译学的主流。