

英汉互译中的过渡处理

张艳红

中国科学院声学研究所 100080

yanhong_z@hotmail.com

摘要: 根据 HNC 理论的机器翻译观, 语言之间的翻译是一个映射和反映射的过程, 而反映射的过程中包含有过渡处理。本论文对过渡处理的六个环节做了简单的介绍, 并在对英汉对照语料进行 HNC 理论分析的基础上, 对各项过渡处理所占的比例做了必要的统计, 最后, 较为详细的说明英汉句类转换的一些基本规律。

关键词: 概念空间 语言空间 过渡处理 句类基元 句类转换

the transition of MT between English and Chinese

yanhong zhang

institute of Acoustics of CAS 100080

yanhong_z@hotmail.com

ABSTRACT: Translation from the Source Language to the Target Language is a process of mapping and inverse mapping according to the Machine Translation view of the HNC theory, and a transition which includes six items simply, and do necessary statistic of each item's rate on certain corpus after analyzing certain English-Chinese language corpus with HNC theory. Finally we give some rules of Sentence-Category Transformation.

Keywords: the concept space the language space transition Sentence-Category Sentence-Category transformation

1. 关于过渡处理

HNC 理论立足于语义块和句类等基本概念而提出了概念空间和语言空间的说法, 提出了处理自然语言的模型。认为计算机要理解自然语言, 需要建立起语言空间向概念空间的映射关系,

从而达到理解的目的。就机器翻译而言，HNC 理论认为源语言向目标语言的翻译需要映射和反映射的两个过程，即先对源语言进行句类分析，分析的结果是把它映射为概念空间的表示，然后把这种概念空间的表示生成目标语言的形式，这是一个反映射的过程。因为就人的思维而言，理解是基于概念的，概念空间可以说是唯一的、独立的。而语言空间则有形形色色不同语种的语言。HNC 理论的机器翻译的原则就是，先理解，后翻译。

语义块和句类是 HNC 理论分析语言的基本单位。HNC 理论认为，就源语言和目标语言而言，语义块的内涵是不变的，但句类的确定、语句的格式、语义块的构成及位置等由于语种的不同而显示出不同的个性。由此，HNC 理论提出，在机器翻译的模型中，需要一个过渡处理，主要是完成从概念空间向目标语言的反映射。HNC 理论把它总结为“两转换、两变换和两调整”等六个环节。即句类转换、语句格式转换；语义块构成变换、语义块主辅变换；语义块的位置调整和句类排序调整。

2. 六个环节简介

句类是句子的语义类别，HNC 理论在 7 个句类基元的基础上，提出了 57 个基本句类及其一级子类，并有混合句类和复合句类的说法。而 57 个基本句类及其子类和混合句类能完成对句子的类别的判定。但是语言的表达是灵活多样的，可以用这种句类表达，也可以用别的句类来表达，尤其是不同语种的语言，语种个性甚至决定了它不存在某种句类。但是，句类转换又有它的规律，抓住这些规律，就可以实现计算机对句类转换操作。例如，英语经常使用是否判断句 jDJ 而汉语一般采用 S04J 来表达相同的意义。

His methods ||were ||direct and strong. jDJ
他的手段||直截了当，顽强有力。S04J
But they|| were|| still very vague. jDJ
但是这些符号||仍然很不明确。S04J

所以，英语中由属性 u 概念充当表语的这种系表结构，句类分析为是否判断句，翻译到汉语中一般都要进行转换，转换为简明状态句。

HNC 理论给出了基本句类的表示式，是由主语义块串接而成的。但是这些主语义块在具体语言中的位置是可以改动的。而且在不同语言中，这种变化更加明显。所以语句的格式也是需要转换的。HNC 理论根据语义块标记符的有无、语义块的位置等把语句格式分为基本表示式、规范格式、违例格式和省略格式。比如：就英汉都有的被动句而言，语义块 JK1 的位置就不同。

[By September 21] the city ||was ringed ||by Wehrmacht forces;

JK1

[到了九月二十一日], 华沙||被德国军队||包围。

JK1

语义块可以由一个词或短语充当,也可以由一个小句充当。HNC 理论提出了块扩和句蜕的概念,提出了语义块 JK 包含有要素和说明的不同部分,语义块 EK 有复合构成,提出了语义块有分离现象,提出了语义块有良性非良性之别等说法。不同语种的语义块构成是不同的,就变换而言,需要从两个角度来考虑,一个是语义块的分离现象,比如作为 EK 复合构成一部分的 QE 或 HE,在汉语中一般紧紧附着在 EK 前后,而英语中的位置很灵活,可以位于句首、句中和句尾等。如英语的 already、first 以及表示程度等的副词等 u 类概念。另一个方面是就语义块的内部结构而言的,拿 JK 来说,汉语通常是一种偏正结构,先对象后内容,是 BC 的形式,而英语经常是正偏结构,先内容后对象,通常采用不定式、分词和关系代词引导的从句来进行修饰。比如:

.... contempt from English and France.....

..... 来自英国和法国的轻蔑.....

HNC 理论把语义块分为主辅两种语义块,主语义块与句类密切相联,辅语义块相对独立,有自己的分类和特点。辅语义块一般是有标记符的。辅语义块有方式、工具、途径、参照、条件、因、果、因果不分等八类。句类对主语义块的类型、内涵等是有约束的,但是,辅语义块往往可以成为主语义块的修饰部分,从而改变它辅语义块的地位。在具体语言中,主语义块也可以加上标记符或通过某种形态上的变化而成为辅语义块。如:

The Poles ||was ||very inept at their strategic dispositions

jDJ=DB+jD+DC

波兰人民||[在战略部署上]||十分低能。

SJ=SB+Re+S

波兰人民的战略部署||十分低能。

S01J=SBC+S

语义块的排序调整主要指辅语义块而言的。就英汉两大语言而言,英语中辅语义块可以位于句首,也可以位于句尾,位于句中时,一般在核心 E 语义块之前。而汉语的辅语义块一般都不能位于句尾,而是句首或句中(JK 和 EK 之间)。如:

I || will go|| there|| [by bike].

我||[乘车]去||那儿。

不仅是语义块需要排序调整，句类也是需要调整的。主要体现在两个方面，一个是由句间逻辑说明符引导的语句上，比如，英语中 if 引导的条件句，可以位于句尾，而汉语的“如果”部分一般位于句首；一个是，由于英语有诸多关系词，从而有英语多用长句，汉语多用短句的说法，这样从英语到汉语的翻译中，经常会有一变多句类的情况，这多个句类的顺序排列，就是句类排序的调整问题。

3. 过渡处理的研究方法及结果

对于过渡处理的研究，我们是基于英汉对照的语料的。对两种不同语言的语料进行 HNC 理论的分析，主要是句类语义块的分析。HNC 理论适合于不同语言，对汉语进行过大规模的分析，对英语的分析同样得到了证明。

语料从文体上可以分为表述文、叙述文、评述文和论述文四类。我们主要就叙述文和论述文语料，涉及到小说、新闻、说明和记叙等运用领域。总计 800 句英汉对照语料，四个领域各占 200 句，且以篇章为单位。语料分别选自《战争风云》、《简爱》、《英语世界》中的说明文和《英语沙龙》中的新闻。

在标注分析语料的基础上，对六个过渡处理进行了统计。在总的 800 句双语对照语料中，不需过渡处理的总计有 175 个，需要过渡处理的句子则是 625 个，统计情况如下：

过渡处理	个数	总百分比	次百分比
句类转换	236	30%	38%
格式转换	162	20%	26%
构成变换	349	44%	56%
主辅变换	58	7%	9%
块调整	204	26%	33%
句调整	20	3%	3%

我们可以发现，在六大过渡处理的项目中，语义块的构成变换是最复杂而频繁的，它占的比例最大，然后依次是句类转换、语义块位置调整、格式转换和主辅变换，最后是句类的排序调整。由此也能看出英汉两大语言的 HNC 理论的明显区别特征。并且，大多数需要进行过渡处理的句子，并不是单项需要的，而是好几项都需要，也就是说，它的过渡处理方面的难点是多重嵌套的，有时，一个句子中，光语义块的构成变换就需要进行多次。

4. 句类转换

就英汉不同语种而言，我们分析到一些句类转换的规律。列举如下：

第一，名动意义上的转换。由于英语是形态语言，表示动态概念的词在英语中有名词的形式。往往会采用与汉语不同的句类表达。如：

His fury ||was to the highest. (jDJ)
他||愤怒到了极点。(X20J)

第二，系表结构引起的转换。英语的系表结构中，如果是由 u 类概念充当时，我们把英语分析为是否判断句 jDJ，而汉语一般地采用不含 E 语义块的简明状态句 S04J。

But they ||were ||still very vague. jDJ
但是这些符号||仍然很不明确。

此外，英语中的介词或介词短语，副词也可以有动态概念的信息，而汉语中的介词不象英语丰富，大多数介词也来自动词的演变。所以，表语由这些介词短语或副词充当时，也会有句类转换的现象。英语中是否判断句句蜕转换为汉语何种句类，则需要具体判断。如：

Class || is ||over. jDJ
课||结束了。PJ

第三，被动和非被动引起的句类转换。汉语中，“被”字是从“遭受”的意义演变而来的，逐渐蜕变为一个功能词，HNC 理论把它分析为主语义块的标记符。这样，汉语的被动句同样也可用“挨、受到、遭受、得到”等来表达。而这些词作为句中的特征语义块时，所代表的是不同的句类。所以英语和汉语的翻译中，就出现了基本作用句向被动承受句、反应句及其混合句向一般承受句的转换。如：

In the past decade, two million children||have been seriously injured. XJ
在过去 10 年中，有 200 万儿童||遭受||重伤。 X12J

另外，英语中绝大多数及物动词的短语都有被动式，被动式借助形态变化来表达。而汉

语的被动式的运用就有很多限制。很多情况下，由于英语的被动式由系表结构演变而来，既可以表示动态，也可以表示静态。汉语的被动句基本上用来表示动态的，HNC 理论的分析，汉语中，广义作用句才有可能含有被动句，而表示静态的句子主要属于效应句。所以英汉的翻译中，就存在广义作用句向效应句的转换现象。如：

Glass || is broken. XJ
玻璃杯||破了。 YJ

第四，物称与人称所引起的句类转换。英语较常用物称表达法，即不用人称来叙述，而让事物以客观的口气呈现出来。汉语则常用人称表达。这一特点主要体现在如何使用主语和动词这两方面。这样就导致了句类转换。如：

An idea || suddenly struck || me. XJ
我||突然想到了||一个主意。 D01J

第五，使动引起的句类转换（自动和他动引起的句类转换）。及物动词（又称自动词），是可以带宾语的，而不及物动词（又称他动词）是不可以带宾语的。英语中动词的及物、不及物与汉语是不相对应的，主要体现在英语中的不及物动词加介词后可以有及物动词的功能，翻译为汉语同样有及物的功能，而英语中有一部分及物动词在汉语中是不及物动词，所以这类动词在英语中一般是作用句，汉语中一般用效应句来表达，翻译时，尤其是从英语的这种表达到汉语的语句时，一般用使动句来表达，成为作用效应句的句类。如：

They || would paralyze || the German nation forever. XJ
他们||会永远使|| 德国民族||瘫痪。 XYJ

第六，一变多句类引起的句类转换。英语多用长句，汉语好用短句。原因很多，英语语法形态丰富，词语的意义、地位等从形态上一目了然。而且英语有关系词，借助这些关系词可以把句子按照主从、同位等不同性质拈连起来，仍然是一个整句。汉语没有关系词，缺乏把单句拈连起来的手段。另一方面，动词对任何语言来说，都是很重要的，在汉语中更为突出。其中一个原因是，与英语比较而言，汉语的介词很贫乏，而且汉语没有分词。在英语中使用介词和分词的地方，汉语经常都直用动词，且成为独立的动词结构，成为一个简短的句子。根据 HNC 理论，语义块或语义块的一部分变为句子。所以，在英汉的互译中，会有句类个数的问题。从英语到汉语，往往是一个主句类变为多个主句类，而汉语到英语，则是多个主句类糅合到一个主句类中。其中，突出表现在，复杂因果句或果因句向复句的转换、基本状态句和基本判断句向无特征语义块句类的转换。如：

The Poles || elected to stand || all their indefensible border, thus inviting || quick fragmenting, encirclement, and reduction. P21J

波兰人||把兵力||分散在全部无法防守的边境线上, 结果很快就被分割、包围、消灭。XT0J+XJ

另外, 上面主要就英汉不同语种的角度分析句类转换的现象, 单纯就句类而言, 也有一些句类转换的规律。比如, 任何句类都可以转换为是否判断句, 任何句类都可以转换为存在判断句等。

最后, 就句类转换的意义而言, 主要是为利用起两类关联知识, 特征语义块与广义对象语义块之间的关联和广义对象语义块之间的关联。每一个句类都含有的特定的约束知识, 而且这些知识也主要就这两类关联知识出发的。比如, 就任何句类向是否判断句的转换而言, 是否判断句的广义对象语义块之间的关联性很弱, 需要还原到原来的句类, 才能引出语义块之间的关联知识, 实现句类分析的目的。就英汉不同语言引起的句类转换而言, 语言不同, 对句类的选择和使用也是不同的, 通过这种句类转换的知识, 才能达到翻译语言的目的。

参考文献

- [1] 黄曾阳 《HNC理论与自然语言语句的理解》 《中国基础科学》, 1999 2-4
国家科技部高技术研究发展中心 中国科学院文献情报中心
- [2] 黄曾阳 《句类分析的20项难点》 内部资料 2000
- [3] 汪康懋 肖研 《英语语序的比较研究》 《英汉对比研究论文集》 上海外语教育出版社 1997 3
- [4] COLLINS COBUILD 《英语语法大全》 商务印书馆 1999
- [5] 刘宓庆 《文体与翻译》增订本 汇总国对外翻译出版公司
- [6] 连淑能 《英汉对比研究》 高等教育出版社 1993 3