

汉英机器翻译中以谓语为中心的句型分析和转换

陈 圣 信

清华大学外语系

摘 要

由于汉语句型与英语句型存在很多重大的区别,给汉英机器翻译带来许多困难。解决好汉语句型的分析及其到英语句型的转换,对提高汉英翻译系统的准确率影响很大,是汉英机器翻译中的一个关键问题。谓语是句子的中心,对全句起支配作用。本文以谓语为中心,重点进行汉语的句型分析及探索其到英语句型的转换规则。

关键词: 汉英机器翻译; 句型分析; 以谓语为中心; 转换规则

ABSTRACT

The significant differences between the Chinese sentence pattern and English sentence pattern cause a lot of trouble in Chinese-English machine translation. To solve the problems in Chinese sentence pattern analysis and the transformation to English sentence pattern would make great influence in increasing the accuracy of the machine translation system. The predicate is the center of a sentence and it dominates the whole sentence. Therefore this paper takes the predicate as the center and stresses on the analysis of the Chinese sentence pattern and studies the transformation rules from Chinese sentence pattern to English sentence pattern.

Key words: Chinese-English machine translation; sentence pattern analysis; predicate as center; transformation rules

一、引言

在汉英机器翻译中有许多棘手的问题,如词的切分,定词性及解决兼类词,短语分析,虚词分析,句型分析及转换等等。本文重点讨论了以谓语为中心的主谓结构的句型分析,分析谓语的组成,性质及其转换为目标语英语的规律性问题。

汉语与英语的一个重大区别在于汉语是一种语义性语言,而英语是一种语法性语言。例如,在英语的句型结构 NP1-VP-NP2 中,一般 NP1 (主语) 是 VP 的施动者,而 NP2 (宾语) 是 VP 的受动者,由语法结构来决定。而翻成汉语时,其词序基本不变。但在汉语句型结

构 NP1-VP-NP2 中, 却错综复杂, 往往语义起很重要的作用。而翻成英语时, 词序及结构常常改变。请看下面例句:

- ①这本书给他。
- ②天空挂着明月。
- ③这辆车能坐 10 个人。
- ④他正在弹钢琴。

以上的例句都是 NP1-VP-NP2 句型结构, 但翻成英语时, 除④词序结构不变外, 其余三句均发生变化。

①中的“这本书”在汉语句型结构中是名词性主语, 而转换为英语后成为直接宾语。“他”在汉语句型结构中是宾语, 而转换为英语后却成为间接宾语。

NP1-VP-NP2—>VP-NP2-NP1

②中的“天空”在汉语是地点性名词作主语, 而转换成英语后则成为状语“in the sky”而主语却是原来汉语句型中的宾语“明月”。

NP1-VP-NP2—>NP2-VP-(NP1—>DP)

③中的宾语“10 个人”是动词的施动者, 而“这辆车”却是动词的受动者。

NP1-VP-NP2—>NP2-VP-NP1

以上的句型的变化和语义有密切的关系。如①中的“这本书”的语义信息决定它不能作为“给”的施动者, 而能作为“给”的直接宾语。如②中的“天空”是地点性名词不能作为“挂着”的施动者。在③中的“这辆车”是交通工具性名词, 不能作为“坐”的施动者, 而“10 个人”的语义信息决定它能作为“坐”的施动者。由此可见, 在汉语到英语的句型转换中, 不仅要看语法信息, 还要考虑语义信息, 才有可能解决问题。

二、句型的成分和以谓语为中心的句型成分分析

汉语句型成分分析应以谓语为中心。

一个原因是: 一个汉语句子, 一般来说, 主语和状语都在谓语的前面, 而宾语与补语在谓语的后面。从句型的成分次序来看, 谓语处在句子的中心。

另一个原因是: 一般的句子, 主语表示谓语陈述的对象, 状语是修饰、限制谓语的成分, 宾语是动词谓语涉及的对象, 宾语是补充说明谓语的成分。从句型成分之间的关系看, 谓语处于支配的地位。

句子的构成成分可以有主语、谓语、宾语、补语、定语和状语, 但并不是这六种成分可以作为句型成分, 我认为, 主语、谓语、宾语和补语可以作为句型成分, 定语不作为句型成分, 而状语是否作为句型成分要进行具体分析。如副词作状语, 并与其修饰的动词、形容词或另一副词紧密联系在一起, 则应与其所修饰成分抱团, 而不能成为句型成分。但如果是修饰全句的状语, 如时间状语、地点状语等, 翻成英语时作为一个成分, 将有词序变化, 这种状语应算句型成分。

三、汉语句型中的谓语性质分类及其转换为英语的规则

在主谓结构的汉语句型中，谓语的性质大致可以分为以下四类：

1、由名词或名词短语（包括代词、数词、量词）充当谓语的句型其表现形式为：NP1-NP2

这类的句型是汉语中所特有的，翻成英语时，其结构不仅决定于 NP2 的语义属性，更重要的是决定于 NP2 的语义属性对应于 NP2 的语义属性的关系，分述如下：

① NP2 与 NP1 的语义属性相当时，如 NP1 与 NP2 同时具有“人”、“时间”等的语义属性时，可以把原语句看成是 NP1-“是”-NP2，其转换为目标语的规则为：

NP1-NP2-->NP1-“be”-NP2

如：他日本人。——>He is a Japanese.

她大学生。——>She is a college student.

今天国庆。——>Today is National Day.

昨天星期五一——>Yesterday was Friday.

② NP1 与 NP2 的语义属性关系是整体与局部的关系，NP1 与 NP2 可以是同一语义属性，如“一年 365 天”，“年”和“天”都具有同一的时间性语义属性，“年”和“天”又是整体与局部的关系。NP1 与 NP2 也可以是非同一语义属性，如“兔子长耳朵”，“耳朵”是“兔子”的器官，仍然是局部与整体的关系。这种句型的原语句可以看成是：

NP1-“有”-NP2，其转换为目标语的规则为：

NP1-NP2-->NP1-“have”-NP2

③ NP2 是带数量词的体词而与 NP1 的语义属性不同时，如：

爸爸一米八。桥长五十米。小王 25 岁。等等。

这种句型的原语句可以看成是：NP1-“是”-NP2，其转换为目标语的规则为：

NP1-NP2-->NP1-“be”-NP2

④ NP1 与 NP2 都是数量词的体语，但 NP1 与 NP2 的语义不相当，如：

一斤两元。两元一斤。

如果 NP1 和 NP2 当中之一是“钱”的语义属性，另一个是另外的度量词，其规则为：

NP \$(NP \$)-NP \$(NP \$!) ——> “It costs”-NP \$-“for”-NP \$! 或

“The price for”-NP \$!-“be”-NP

(NP \$ 为带“钱”语义属性的数量词，NP \$! 为不带“钱”语义属性的数量词。

其他的规则就不一一历数了。但有的情况非规则所能解决问题的。例如一些成语、谚语等的 NP1-NP2 句型，如，一分价钱一分货；一本万利；中国人民万岁等等，应采取直进直出的办法，完全不必进行句型分析及转换。

2、由形容词或形容词短语充当谓语的句型，又分以下几种情况：

① 由独个形容词或带状语的形容词充当谓语。如，墙壁乳白色的；她很美丽；穷人生活真困难，等等。这类句型的转换规则为：

NP-(DP)-AP-->NP-“be”-(DP)-AP (DP 为状语，AP 为形容词或形容词短语)

形容词所带的状语有：很；十分；非常；特别；真；多么；无比；无限；分外；格外；太；越来越；等等。

② 由带补语的形容词充当谓语的句型。如，我高兴极了。其句型转换规则为：

NP-AP-CP-->NP-“be”-CP-AP (CP 为补语)

③ 形容词所带补语为带“得”的形式。如，

这朵花艳丽得很；这纸白得象雪；水清得看见水底的石头；等等。其转换规则视“得”后的成分而不同。

NP-AP-“得”-DP—>NP-“be”-DP-AP

NP1-AP-“得象”-NP2—>NP1-“be as”-AP-“as”-NP2

NP-AP-“得”-VP—>NP-“be so”-AP-“that”-VP 目标语的 VP 应是帮助动词的动词短语。

3. 由动词或动词短语作谓语的句型。

汉语句子中的动词可以是谓语，也可能非谓语，这要看动词或动词短语的性质及其在句子中的位置来决定。例如助动词通常不单独作为谓语，而是与后面另一谓语动词结合起来才能构成谓语。又如补语动词，也不能单独作为谓语，必须与前面另一被修饰的动词结合起来才能作为谓语。试看下面的句子：

①他去上海是肯定的。

②这是他写的报告。

③你自己去拜访他更好。

④我们飞抵北京的时候，他已经走了。

①中有两个动词，“是”是作谓语的动词，而“去”只是主语“他去上海”的一个成分，不能作全句的谓语。

②中的“写”是宾语的一个成分，不能作全句的谓语，而“是”才是谓语动词。

全句中只有一个动词，这个动词也不一定是谓语。如③中的“拜访”是主语的一个成分，谓语是形容词。

④中的“飞抵”是状语结构的一个成分，“走了”才是真正的谓语。

动词作谓语的句型，重点讨论转换后的英语句型与原有的汉语句型不同的情况。

①一般情况下，不带宾语或以名词或代词作宾语的动词谓语句型转换后的句型结构和词序不变，但有些情况值得注意。例如：

①代词作宾语时，要用宾格，如 me; us; him; her; them,等等。

②名词作宾语时，如果通过语义分析发现宾语是谓语动词的施动者，如“这辆车能坐十个人。”这时或者改变词序，或者谓语动词要改变，以使主语为施动者，宾语为受动者，如把原句看成是“十个人能坐这辆车。”，或者“这辆车能载十个人。”。其转换规则为：

NP1-VP-NP2—>NP2-VP-NP1 或

NP1-VP-NP2—>NP1-VP'-NP2

②形容词作宾语的动词谓语句型，如“这座建筑显得更加雄伟。”

“显得”译成英语时为 look; seem; appear 等等。这几个词都是多义词，意思不同时，词性也异。但作为与“显得”意思最接近的“看起来”的词性看，根据 Longman 的《当代英汉双解词典》，其词性均为 L7，即为后接形容词的系动词。这类动词谓语句型，原语与目标语的动词分类不同，但其词序不变。

③动词作宾语的动词谓语句型，如“爷爷喜欢闭目养神。”

“喜欢”是谓语动词，另一动词短语“闭目养神”作宾语。“喜欢”的英语译义也是多义的，其中有一词性为 T3，即后接带“to”的不定式动词或动词短语作为其宾语。这种句型的转换规则为：

NP-Vt-VP—>NP-Vt-Vb (Vt 为及物动词, Vb 为带“to”的不定式动词或动词短语

有些象是动词作宾语, 其实是同形名词或名词短语, 如“我们班参加了观摩演出。”; “老师提出了批评。”等等。这种句型应将名词作为宾语处理。

④数量词作宾语、补语的动词谓语句型, 如“我走了 10 分钟。”; “他跑了 10 公里。”等等。这种句型转换为英语时, 一般谓语动词用完成时态, 数量词前加“for”, 如“I have walked for ten minutes.”; “He has run for ten kilometers.”。这种句型的谓语动词一般为不及物动词。

如果谓语动词为及物动词, 则谓语动词不用变为完成时态, 而数量词前也不用加介词。如“奶奶邮寄了 100 元给我。”英语译为“Grandma mailed 100 yuan to me.”

谓语动词为“有”、“是”、“在”, 也有其转换规则, 这里就不一一论述了。

⑥宾语表示工具的动词谓语句型, 如“我用钢笔写。”; “他用铁锹挖。”等。转换句型词序不变, 只是在宾语前加介词“with”。

⑦带双宾语的动词谓语句型, 如果原语句型的顺序为: 主语-谓语动词-间接宾语-直接宾语, 如“他给我一本书。”; 小李问小王一个问题。”等, 英语句型的词序不变。如果第二个宾语是动宾的形式, 谓语动词是“给”、“让”等, 则要作具体分析。请看下面句子:

①老人给小孩讲故事。

②老人让小孩讲故事。

③黄山给游人留下了深刻的印象。

①与②的差别只在谓语动词“给”和“让”, 但实际意思差别很大。①中“讲”的施动者是“老人”, ②中“讲”的施动者却是“小孩”。“小孩”是一个兼语, 它是“让”的宾语, 又是“讲”的主语。①的英语译义是“The old man tells the children stories”, 而②的英语译义是“The old man lets the children tell stories”。①与③的主要差别在主语“老人”和“黄山”的语义属性不同。“老人”可作“讲”的施动者, 而“黄山”却不能作“留下了”的施动者, 只有“游人”可作“留下了”的施动者

从这一点看②与③实际更接近。③中的“给”用“让”来代替, 其意思不变。“让”的施动者的语义属性可以是“人”, 也可以是“具体物”, 甚至可以是“抽象物”。但“讲”的施动者是“人”。这种以“给”或“让”等作谓语动词, 动宾结构作第二宾语的句型结构中, 如果谓语动词是“给”, 而且不能用“让”替代, 全句的主语又能作动宾结构的施动者, 其转换规则是:

NP1-“给”-NP2-Vt-NP3—>NP1-Vt-NP2-NP3

这种句型是带双宾语的动词谓语句型。

如果谓语动词是“让”, 或者是“给”, 但可以用“让”替代。全句的主语又不能作动宾结构的施动者, 其转换规则是:

NP1-“让”-NP2-Vt-NP3—>NP1-“let”-NP2-Vt-NP3

根据 Longman 的《当代英汉双解词典》, “let”的词性为 V2, 即后接宾语加不带“to”的不定式动词短语结构(动宾结构)。这种句型结构和带双宾语的动词谓语句型在英语句型中是不同的。

如果谓语动词是“把”, 而动宾结构的动词是“给”, 则其转换规则为:

NP1-“把”-NP2-“给”-NP3—>NP1-“give”-NP3-NP2

这种句型结构也是带双宾语的动词谓语句型。

4. 由主谓结构作谓语的句型

①主谓结构的谓语是形容词或形容词短语, 如“小朋友思想很单纯。”, 其意思和下面两个

句子相当。

①小朋友有单纯的思想。

②小朋友的思想很单纯。

这种句型的转换规则是：

NP1-NP2-AP → NP1-"have"-AP-NP2 或
→ NP2-"of"-NP1-"be"-AP

②主语是主谓结构的宾语

从语义分析得出，主语是主谓结构的受动者。如“花生米人人爱吃。”；“你的意思我领会。”。其句型转换规则为：

NP1-NP2-Vt → NP2-Vt-NP1

③全句的主语实际是状语，如“炎热的夏天，食品很容易变质。”；“山村空气新鲜。”。这时全句的主语往往是时间性名词或地点性名词作状语，其转换规则为：

NP1,NP2-VP → NP2-VP-DP

或 NP1-NP2-AP → NP2-"be"-AP-DP

(DP 是 NP1 前加或不介词形成的状语)

四、结束语

本文仅就汉英机器翻译中以谓语为中心的句型分析和转换进行了一些初步的探讨，就这个问题而言，也还有不少问题没有进行分析和讨论。所讨论的问题肯定还有不全面和值得斟酌的地方。由于汉语句型结构的多样性和灵活性，有些规则在某些场合可以行得通，在另一些情况下就可能行不通。希望得到同行的批评指正，我们将在翻译系统的调试和运用中，对这些规则不断进行修改和完善，逐步做出准确率较高，实用性较强的汉英机器翻译系统。

参考文献

1. 朱德熙 《现代汉语语法研究》，商务印刷馆，1985
2. 王力 《中国现代语法》，商务印刷馆，1986
3. 马真 《简明实用汉语语法》，北京大学出版社，1988.6
4. 赵淑华 《谈 80 年代与 90 年代的句型研究》，北京语言学院出版社，1992
5. 黄昌宁 《计算语言学讲义》，清华大学计算机系，1991.2
6. 马庆株 《汉语动词和动词性结构》，北京语言学院出版社，1992.10
7. 田高库 《现代汉语基础知识释疑》，北京语言学院出版社，1992.8
8. 张开勤 《小学生句典》，国际文化出版公司，1992.6
9. 郑碧霞 《汉语句型自动分析和分布统计模型的研究与实现》，
清华大学计算机系硕士论文，1993.6