

日语口语语料库中倒装表现的形态特征

郭 洁

(天津财经大学珠江学院, 天津 301811)

摘要: 本文分析了日语口语语料库中核心模拟讲座的倒装表现, 从形态论信息方面阐明了倒装表现的形态论特征。主要分析了倒装要素末尾词, 本体末尾词和本体次末词的特征, 分析结果主要为(1)本体末尾词中终助词的比例较高;(2)本体次末词+本体末尾词最常见的模式是“助动词+终助词”; (3)本体末尾词+倒装要素末尾词最常见的模式是“终助词——格助词”。(4)最常见的倒装要素末尾词是格助词。

关键词: 倒装; 日语口语语料库; 形态论

一、引言

一般来说, 自然语言的表达都有固定的语序, 日语中的语序为主语在前、谓语在后, 修饰语在前、被修饰语在后。然而, 在实际语言使用中, 尤其是在口语中, 并非所有句子都一定遵循这种语序, 而是由于各种原因导致语序发生变化, 倒装就是其中之一。

日语中的语序在类型学分类中被认定为SOV, 但在实际语言中却表现出多样性, 包括宾语在主语之前和句子要素在谓语之后等。其中, 句子要素在谓语之后的倒装现象是口语的特点。倒装的句子要素主要分为主语、修饰语等, 但没有基于详细形态论信息的分析。因此, 本研究将重点放在倒装表达式的形态特征上, 并在分析口语语料库的基础上研究它们对倒装表达式使用的影响。

二、方法

(一) 分析对象

分析对象是“日语口语语料库”(Corpus of Spontaneous Japanese)。日语口语语料库包含约660个小时的自发语音(约700万字)。语料库的一部分(约45小时(50万字))被称为语料库的“核心”(core), 除转录文本外, 还提供(人工校正的)形态信息、分句单位信息、分段和语调标签。本文的分析包括107篇核心模拟讲座。

(二) 倒装句识别

在日语口语语料库中, “核心”中约500,000个单词的数据被分配为“分句单元”, 与书面日语中的句子相对应。倒装句被视为单个分句单元, 并且被标注出来, 模拟语音数据中共有106个倒装句。

这些倒装句的本体和倒装要素定义如下。

句子“回去了呀, 去老家”是一个倒装句, 分为两部分: 本体“回去了呀”和倒装要素“去老家”。本文重点考察了以下三个词的词形信息。“desu”: 本体的次末词(倒数第二个词); “ne”: 本体的末尾词; “ni”: 倒装要素的末尾词。

同样, 我们确定了9577个非倒装句中的次末词和末尾词, 并将它们与倒装句中“本体的次末词”、“本体的末尾词”和“倒装要素的末尾词”的形态论信息进行了比较。

三、结果

(一) 形态论特征

1. 倒装要素末尾词的品词分布

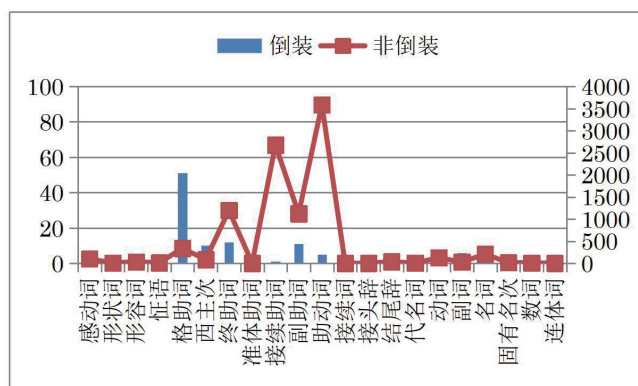


图1 倒装要素末尾词的品词

倒装句中倒装要素的末尾词和非倒装句中末尾词的品词频率如图1所示。条形图代表倒装句, 折线图代表非倒装句, 纵轴(左)代表倒装句中的频率, 纵轴(右)代表非倒装句中的频率。在倒装要素的末尾, 有51个格助词, 如“ni”、“de”和“kara”, 占有倒装句的48.1%。这与非倒装句中末尾词的分布(主要是助动词和连接词)大相径庭。

2. 本体末尾词的品词分布

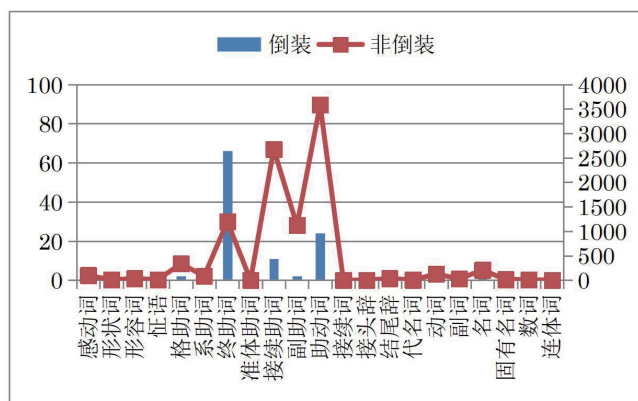


图2 本体末尾词的品词

倒装句和非倒装句本体末尾词的品词频率见图2。在本体末尾词中，“yo”和“ne”等终助词词尾共62个，占倒装句总数的58.5%。此外，助动词占22%，接续助词占13%。而在非倒装句中，只有12.5%的句子的最后一个词是终助词，倒装句中终助词的比率是非倒装句的4.68倍。

3. 本体中次末词的品词分布

倒装句本体次末词和非倒装句本体次末词的品词频率如图3所示。从图中可以看出，“desu”、“masu”、“da”等66个助动词占本体次末词的绝大部分（62.3%）。然而，在非倒装句中，助动词的比例也是46.2%，因此不能说助动词在倒装句中特别常见。与此相反，倒装句中出现频率第二高的终助词（12个）占有所有倒装句的11.3%，远高于非倒装句中终助词占比的2.3%。因此可以说，与非倒装句相比，倒装句本体次末词具有终助词比较多的特点。

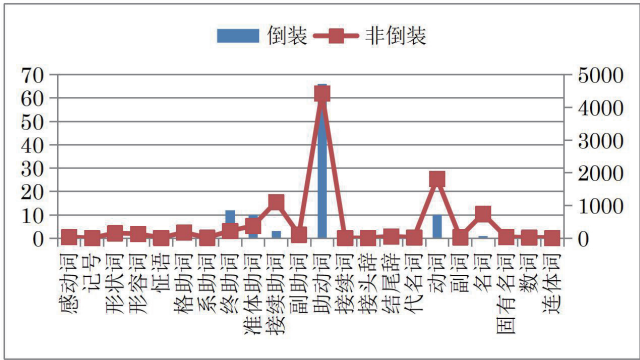


图3：本体中次末词的品词

4. 依存关系

到目前为止基本分析了倒装句的形态特征，但仅对每个位置进行独立研究是不够的。通过厘清这三个部分之间的依存关系，可以更好地阐明倒装句的形态特征。

首先，我们考察了“本体次末词”和“本体末尾词”之间的关系（表1）。最常见的模式是“助动词+终助词”。

表1 本体次末词与末尾词之间的关系

本体次末词	本体末尾词							總計
	格助词	係助词	終助词	助動词	接統助词	副助词	名词	
格助词		1	1				1	3
形状词				1				1
形容词				2				2
終助词			8					8
準体助词	1		1	1				3
助动词			51	11	11	1		74
接統助词			1	1		1		3
代名词				1				1
動词				4	2			6

副助词				1				1
名词	4							4
總計	5	1	62	22	13	2	1	106

其次是“本体末位词”和“倒装要素末尾词”之间的关系（表2）。最常见的模式是“终助词--格助词”，其次是“终助词--终助词”，第三多的是“接续助词--格助词”。

表2 本体末尾词和倒装要素末尾词的关系

倒置要素的末尾词	本体末尾词							统计
	格助词	系助词	终助词	助动词	接续助词	副助词	名词	
格助词	4		28	8	9	1	1	51
系助词			4	3	2			9
终助词			10	1		1		12
助动词			4	1				5
接续词		1		1				2
接续助词				1				1
接尾辞			1					1
代名词			1					1
副词	1		5					6
副助词			7	3	1			11
名词			2	4	1			7
统计	5	1	62	22	13	2	1	106

四、讨论

在日语口语语料库中，倒装表现的形态论特征包括：（1）本体末尾词中终助词的比例较高；（2）本体次末词+本体末尾词最常见的模式是“助动词+终助词”；（3）本体末尾词+倒装要素末尾词最常见的模式是“终助词--格助词”。（4）最常见的倒装要素末尾词是格助词。本体以终助词结尾的句子，特别是“masune”和“desune”结尾的句子，说话者是关注听话者的反应的，如果此时听话者没有点头之类的反应，说话者就会追加倒装要素用来补充内容。因此，在以“终助词”结尾的句子中，倒装要素很可能经常跟在本体之后。今后，我们希望利用会话数据来分析倒装表达与使用者的关系。

参考文献：

[1] 毛文伟. 日本的语料库文体学研究：进展、问题及展望 [J], 外国语, 2024, 44 (3) : 82-90.

[2] 腾玥. 中日对译语料库中终助词的性别差异中日对比研究 [D], 大连外国语学院, 2023.

[3] 高霞. 基于中西学者学术论文可比语料库的情态动词研究 [J]. 解放军外国语学院学报, 2020, 43 (5) : 1-9, 127, 159.