

当代语言学 第 26 卷 2024 年第 6 期

目 录

语言、社会及意识形态研究

论语言与文明 王春辉(791)

数智新时代的语言认同与文化认同 张璟玮(811)

当代语言学前沿

语言开启思维：语言奇点与思维初始基因 赵汀阳(825)

研究性论文

工作空间定义下合并的最新发展 潘俊楠 杜雨桥(842)

指量结构和关系小句位置关系及信息功能差异 寇 鑫(860)

汉语口语叙述语篇中的链式话题结构 许立群(880)

基于语料库统计模型的汉语与格交替实例构式原型性分析 张 懂 许家金(897)

语义的寄生、多重寄生和去寄生——以(非)现实语义的寄生表达为例
..... 周晨磊 夏俐萍(911)

藏语的存在动词与示证范畴 邵明园(927)

其他

致谢 (910)

“递系”的术语翻译问题 (926)

《当代语言学》2024 年总目录 (946)

讣告 (948)

(本期责任编辑：王 伟)

Contemporary Linguistics

Volume 26 Number 6

December 2024

Contents

Studies of Language, Society, and Ideology

On language and civilization

WANG Chunhui (791)

Language identity and cultural identity in the New Era of Digital Intelligence

ZHANG Jingwei (811)

Frontiers in Contemporary Linguistics

Language initiates thinking: Language singularity and thinking's initial gene

ZHAO Tingyang (825)

Research Articles

The latest development of MERGE under the notion of Workspace

PAN Victor Junnan and DU Yuqiao (842)

Demonstrative-classifier positioning in Chinese relative clauses:

A discourse-pragmatic study

KOU Xin (860)

Chain topic structures in Chinese oral narrative discourse

XU Liqun (880)

A corpus-based statistical modeling of instance prototypicality of Chinese dative alternation

ZHANG Dong and XU Jiajin (897)

Parasituality, multi-parasitization, and de-parasitization:

A case study on parasitical expressions of (ir)realis meaning

ZHOu Chenlei and XIA Liping (911)

Interrelationship of existential verbs and evidentiality in modern Tibetan

SHAO Mingyuan (927)

基于语料库统计模型的 汉语与格交替实例构式原型性分析^{*}

张懂 许家金

提要 以往对构式原型性的分析以内省和思辨为主，本文则为分析构式原型性引入一种基于语料库数据的统计建模和可视化方法。本文基于大规模语料库标注数据，使用线性判别分析法，综合考察句法、语义和语用等诸多变量对汉语与格交替的共同影响，分析了与格交替实例构式的原型性。研究发现：1) 最高线性判别值对应的实例为双宾构式的原型实例，最低线性判别值对应的实例为介词与格构式的原型实例；2) 以原型实例为参照，通过比较其他实例与原型实例线性判别值之间的差异，即可确定其他实例的原型性。

关键词 与格交替 线性判别分析法 原型实例 构式原型性 语料库

A Corpus-based Statistical Modeling of Instance Prototypicality of Chinese Dative Alternation

ZHANG Dong and XU Jiajin

Abstract Construction alternation, which refers to the interchangeable use of near-synonymous constructions, has gained considerable attention in the field of Construction Grammar over the past few decades. Dative alternation is a type of construction alternation frequently observed in both English and Chinese, which involves the alternating use of a double object construction and a prepositional dative construction (Bresnan and Ford 2010; Röthlisberger et al. 2017; Zhang and Xu 2019). For example:

- (1) a. Who gave that wonderful watch to you?
b. Who gave you that wonderful watch? (Bresnan and Ford 2010: 169)
- (2) a. Zhang San song Li Si yiben shu.
'Zhang San gave Li Si a book.'
b. Zhang San song yiben shu gei Li Si.
'Zhang San gave a book to Li Si.' (He 2009: 18)

Example (1a) illustrates the prepositional dative construction in English, and (1b) represents the double object construction in English, which together represent the English dative alternation. (2a) depicts the double object construction in Chinese, and (2b) represents the prepositional dative construction in Chinese, where *gei* serves as the

^{*} 本文系国家社科基金青年项目“构式变体模型视角下英汉构式变异与演变的交叉验证研究”(22CYY002)的阶段性成果。《当代语言学》编辑部和匿名审稿专家对本文提出了诸多宝贵意见，谨致谢忱。

recipient marker, introducing the recipient Li Si. These two examples (2a) and (2b) represent the Chinese dative alternation.

While previous studies on the prototypicality of constructions have primarily used introspective methods, this study aims to provide a corpus-based statistical approach to explore the prototypicality of constructions. This study employs linear discriminant analysis to investigate the prototypicality of instances of Chinese dative alternation based on large-scale richly annotated corpus data. The results indicate that instances with the highest linear discriminant scores are prototypical instances of Chinese double object constructions, whereas those with the lowest linear discriminant scores correspond to prototypical instances of Chinese prepositional dative constructions. The prototypicality of the remaining instances can be assessed by the differences between their linear discriminant scores and those of the prototypical instances.

Keywords Chinese dative alternation, linear discriminant analysis, prototypical instance, prototypicality of construction, corpus

1. 引言

近年来,与格交替(dative alternation)已成句法研究热点(Bresnan 和 Ford 2010; Röthlisberger 等 2017; 张懂 2019, 2020, 2021; 张懂、许家金 2019)。与英语类似,汉语与格交替也包含双宾构式和介词与格构式两种变体。例如:

- (1) a. Who gave that wonderful watch to you?
b. Who gave you that wonderful watch? (Bresnan 和 Ford 2010:169)
- (2) a. 张三送李四一本书。
b. 张三送一本书给李四。(何晓炜 2009:18)

(1a)和(1b)分别为英语介词与格构式和双宾构式,二者构成了英语与格交替。(2a)为汉语双宾构式,论元实现方式为与事(即“李四”)在受事(即“一本书”)之前,句法结构为 $NP_1 V NP_2 NP_3$; (2b)为汉语介词与格构式,句法结构为 $NP_1 V NP_3$ 给 NP_2 ,“给”充当与事标记成分,位于受事之后,用来介引与事论元(顾阳 1999; 何晓炜 2008, 2009),二者构成了汉语与格交替。

以往对构式原型性的分析以内省和思辨为主,主观性较强。本研究拟基于大规模真实语料,对汉语与格交替现象的语境制约因素进行综合量化分析,通过线性判别分析法(linear discriminant analysis),全面考察句法、语义、语用等诸多因素对与格交替的影响,进而借助可视化方法直观勾勒出与格交替实例构式的原型性分布模式。

2. 文献综述

2.1 原型范畴理论及其新进展

范畴化是认知语言学研究的重要议题。经典范畴理论使用一组充分必要条件定义范畴,认为:1)范畴成员地位平等;2)范畴间边界清晰(Taylor 1989: 79)。原型范畴理论则认为:1)范畴成员地位不一,原型是最典型的成员,其他成员因与原型相似得以进入该范畴(Rosch 1975; Rosch 和 Mervis 1975; Lan-

gacker 1987; Ellis 2020), 与原型相似度越高, 则原型性越强, 越居于中心 (Lakoff 1987:60; Taylor 1995:60; Croft 和 Cruse 2004:81); 2) 范畴间边界模糊, 边缘成员的范畴地位不易确定 (Labov 1973:353; Blank 和 Koch 1999:92)。

认知语言学内部与范畴化相关的研究进展是“样例理论”(Exemplar Theory) (Medin 和 Schaffer 1978; Nosofsky 1988; Nosofsky 等 1992)。该理论认为: 1) 范畴表征为独立存储的样例集; 2) 相似性是范畴化的基础, 范畴化倚赖新样例与已存储样例之间的相似性, 且相似性并非恒久不变, 而是因语境而变化的 (Nosofsky 1986:39)。样例理论和原型理论的主要差异在于对概念表征方式有不同看法: 前者认为概念通过独立存储的样例进行表征, 后者认为概念通过对所有范畴成员进行抽象和概括来表征 (Murphy 2002:4)。然而, 越来越多的认知语言学研究倾向于将两种理论结合起来 (Langacker 2010:132-138; Divjak 和 Arppe 2013:226; 张伟伟、王芳 2017:23), 认为具体的样例知识 (item-specific exemplar knowledge) 和抽象的图式知识 (schematic knowledge) 在语言学习过程中共同起作用。说话者既存储关于构式用法的具体知识, 也存储对构式进行抽象和概括形成的图式知识 (Taylor 1989:67; Goldberg 2006:55, 2019:64; Bybee 2010:18)。

目前, 学界对“原型”的界定并不统一, 概言之, 主要有两种解释: 1) 指范畴中具体的典型代表和中心成员 (Rosch 1975; Langacker 1987:371); 2) 指范畴核心特征的抽象图式表征 (Taylor 1995:59)。本研究采用 Langacker (1987:371) 对“原型”的界定, 认为原型是范畴的典型实例, 其他成员依据与原型的不同相似性而表现出不同程度的原型性。本研究将原型范畴理论与样例理论加以结合, 探究与格交替实例构式的原型性。

2.2 已有汉语与格交替研究

此前汉语与格交替研究主要从结构主义、转换生成语法和构式语法等视角展开。结构主义视角下的研究聚焦动词语义特征及其分布。朱德熙 (1979) 讨论了汉语双宾构式及其相关构式之间的句法语义关系, 分析了上述构式间的细微差异。马庆株 (1992) 将双宾构式分为给予类、取得类、准予取类、原因类、结果类等, 探讨了双宾构式的构成及其组成成分的语义特征。转换生成语法视角下的研究分析了双宾构式和介词与格构式之间的句法转换关系 (顾阳 1999; 邓思颖 2003; 满在江 2003)。构式语法视角下的研究聚焦于双宾构式和介词与格构式的构式义。张伯江 (1999:176) 分析了双宾构式的引申机制, 认为其构式义为“有意的给予性转移”。沈家煊 (1999:98) 区分了与“给”相关的构式, 认为介词与格构式表达的构式义为“惠予事物转移并达到某终点, 转移和达到是两个分离过程”。熊学亮 (2007) 分析了双宾构式的原型性, 认为给予类动词构

成的双宾构式原型性较高。

综上所述，已有研究主要采取内省法分析构式原型性，鲜有研究基于大规模真实语料对与格交替构式原型性进行综合量化考察。鉴于此，本研究拟基于大规模汉语语料库，使用线性判别分析法，综合考察句法、语义和语用等诸多变量对汉语与格交替的共同影响，量化分析与格交替构式的原型性。

3. 研究语料与研究方法

3.1 研究语料

基于已有研究(朱德熙 1979；马庆株 1992；张伯江 1999；Liu 2006 等)，本文总结出 43 个与格动词(包括“拨、补、补偿、传、传授、赐、赐予、递、发、分、分配、付、告诉、给、归还、过继、还、汇、寄、奖励、交、交还、教、借、捐赠、卖、赔、赔偿、扔、赏、捎、送、提供、退、退还、献、移交、赠、赠送、转交、转卖、转送、租”)。在 BCC 语料库(综合语料)中逐个检索上述与格动词，每次检索随机抽取包含每个动词的 1000 条索引行，经过人工筛选，得到 1092 句与格交替实例，其中包括 861 句双宾构式和 231 句介词与格构式实例。汉语与格交替构式实例的分布见附录。

3.2 变量选取与标注

基于已有研究(何晓炜 2008；张懂、许家金 2019)，本文从句法、语义和语用等不同层面，对上述 1092 句与格交替实例的影响因素进行定性分析和特征标注。标注变量见表 1。

| 因素类型 | 因素名称 | 因素水平 | 标注名称 |
|------|---------|--------------|------------|
| 自变量 | 与事有生性 | 有生；无生 | ReciAnim |
| | 受事有生性 | 有生；无生 | ThemeAnim |
| | 与事可及性 | 已知；未知 | ReciAcc |
| | 受事可及性 | 已知；未知 | ThemeAcc |
| | 与事代词性 | 代词；非代词 | ReciPron |
| | 受事代词性 | 代词；非代词 | ThemePron |
| | 与事有定性 | 有定；无定 | ReciDef |
| | 受事有定性 | 有定；无定 | ThemeDef |
| | 受事具体性 | 具体；抽象 | ThemeCon |
| | 与事句法复杂性 | 简单；复杂 | ReciCom |
| | 受事句法复杂性 | 简单；复杂 | ThemeCom |
| | 长度差异 | 受事长于/短于/等于与事 | LengthDiff |
| 因变量 | 与格构式变体 | 双宾构式；介词与格构式 | Dative |

表 1 标注变量

本文将有生性标为“有生”或“无生”两类。其中,有生名词短语包括人和动

物,其他名词短语标为无生。可及性标注方案为:当名词短语为第一或第二人称代词,或其所指在其所在句子的前十句已有提及时标为已知,其他名词短语标为未知(Bresnan 和 Ford 2010:174)。将人称代词、不定代词、指示代词和反身代词等标为代词,其他成分标为非代词。当听话者能将名词短语的所指对象与该语境中可能存在的其他实体区别开来时,将其标为有定,否则标为无定(陈平 1987:82)。具体性标注方案为:能被五官感知的原型具体物质标为具体,其余标为抽象(Bresnan 和 Ford 2010:175)。句法复杂性标注方案为:对于有限定成分且限定成分具有句子功能的名词短语,标为复杂成分;对于无限定成分或其限定成分不具有句子功能的名词短语,标为简单成分(Wasow 和 Arnold 2003)。长度差异标注方案为:受事包含的词数多于与事包含的词数、受事包含的词数等于与事包含的词数或受事包含的词数少于与事包含的词数(张懂、许家金 2019)。需要说明的是,在标注长度差异之前,本文使用了中科院计算所汉语词法分析系统 ICTCLAS 3.0 对语料进行分词,并对分词的结果进行人工核查。因变量为句中实际使用的构式变体——双宾构式或介词与格构式。

3.3 研究方法

已有原型性研究的主要不足在于,仅依据成员具有的原型特征数量判断原型性,未量化考察不同特征对原型性的影响权重,造成对原型性的分析主观性较强(Taylor 1995:61;彭睿 2020:341)。本研究采用线性判别分析法,其原理是基于汉语与格交替实例的多项语境特征构建线性判别函数,对汉语与格交替使用进行多变量统计建模,统计出汉语与格交替实例的构式特征权重及其线性判别值,通过线性判别值区分实例的原型性。具体而言,具有最高线性判别值的实例体现了构式最重要的特征,可视为原型实例(Gries 2003b:10)。以原型实例为参照,通过比较其他实例与原型实例线性判别值的差异,即可判定其原型性。与原型实例线性判别值越接近,原型性越强。例如 Gries (2003a, 2003b)使用线性判别分析法,对英语动词-小品词构式交替(如 Fred picked up the book 与 Fred picked the book up)和英语与格交替的原型性进行了考察,证明了线性判别分析法可用于分析构式的原型性。

总之,线性判别分析法可同时考察影响构式使用的多个语境特征,统计实例的线性判别值,量化分析实例的原型性,对构式原型性的分析更加全面综合。而且该方法的分类原理与原型范畴理论和样例理论较为一致,一定程度上能够反映人类范畴化的结果(Gries 2003a:137)。

4. 结果与讨论

4.1 汉语与格交替原型实例分析

构式实例往往呈现出原型效应(prototype effect)(Taylor 1995, 1998;

Bybee 2010), 换言之, 与格交替实例体现出的双宾构式或介词与格构式的原型性存在差异。已有研究多从理论层面对构式实例的原型性进行定性分析 (Taylor 1998; 张伯江 1999; 彭睿 2020), 而本文基于大规模语料, 使用线性判别分析法, 对与格交替实例构式的原型性进行定量研究和可视化分析。

具体分析步骤如下: 使用线性判别分析法, 统计出与格交替的原型实例。本研究使用 R 语言软件^①(R Core Team 2017), 安装 {MASS} 程序包 (Venables 和 Ripley 2002), 调用 lda 函数, 采用线性判别分析法对与格交替使用特征数据进行统计分析^②。统计结果显示, 本文构建的线性判别分析模型的分类准确率约为 83%, 即在同时考虑所有标注特征的情况下, 该模型能准确预测 83% 的实例属于双宾构式还是介词与格构式。可见该模型对与格交替实例具有较好分类能力。使用线性判别分析法, 对不同特征的影响权重进行统计, 结果见表 2。

| 特征 | 水平 | 线性判别系数 |
|---------|--------|--------|
| 与事可及性 | 可及 | 1.367 |
| 受事可及性 | 可及 | -1.072 |
| 与事代词性 | 代词 | 0.436 |
| 长度差异 | 受事长于与事 | 2.624 |
| 长度差异 | 受事等于与事 | 1.987 |
| 与事有定性 | 有定 | -1.548 |
| 受事有定性 | 有定 | 0.234 |
| 受事具体性 | 具体 | -1.090 |
| 受事句法复杂性 | 复杂 | 0.237 |

表 2 特征影响权重

通过与英语与格交替的影响因素对比发现, 上述特征对汉语与格交替也具有显著影响, 但其影响权重以及相对重要性与英汉与格交替有明显差异。具体而言, 对汉语与格交替具有最大影响权重的特征是长度差异, 而对英语与格交替具有最大影响权重的特征是与事有生性 (Bresnan 等 2007:20)。这反映了英汉与格交替最主要的跨语言差异, 其成因有待于进一步探讨。

使用线性判别分析法, 对所有与格交替实例的线性判别值进行统计, 结果见图 1。图 1 中每个黑点代表每个实例的线性判别值。根据线性判别分析法的统计原理可知: 具有最高和最低线性判别值的实例分别对应双宾构式和介词与

① R 语言软件是目前国际学界主流的开源统计分析软件, 在 R 中可以安装程序包实现各种统计分析。本研究使用的线性判别分析法需借助 {MASS} 程序包中的 lda 函数来实现。

② 由于本研究数据集中, 与事有生性、受事有生性、受事代词性和与事句法复杂性的分布极不平衡, 因此在进行线性判别分析之前删除了上述四个变量。

格构式的原型实例，线性判别值约为 0 的实例对应边缘实例。由图 1 可知，最高线性判别值为 2.40，对应的实例如(3)；最低线性判别值为-5.30，对应的实例如(4)。

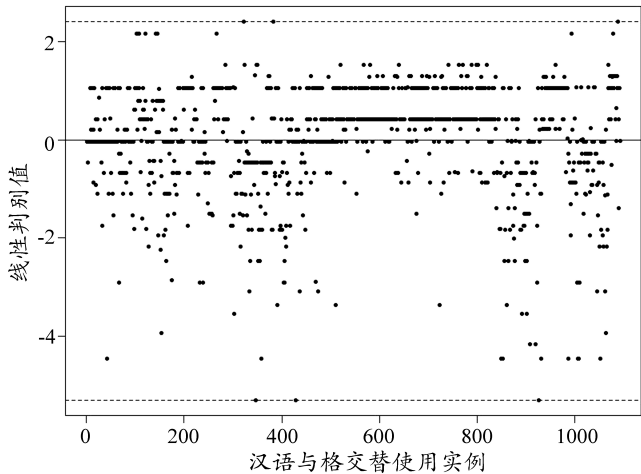


图 1 与格交替实例的线性判别值分布

- (3) 提高山民的科技水平，增强就业能力；请来专家、教授指导和传授山民各种种植和养殖技术，帮助山民开发、创业。
- (4) 我认为从法律上讲，我们虽然不需赔偿那些钱给我们受损失的客户，但是我敢肯定过不了多久他们就不再是我们的客户了。

因此，(3)为双宾构式的原型实例，其使用特征包括：与事表征已知信息，长度短于受事；受事指称抽象事物，表征未知信息 and 无定事物，长度长于与事。(4)为介词与格构式的原型实例，特征包括：受事指称具体事物，表征已知信息和定指事物，长度短于与事；与事为有生事物，表征未知信息，长度长于受事。在本文搜集的所有与格交替实例中，双宾构式原型实例有 3 个，占双宾构式实例的 0.4%；介词与格构式原型实例也有 3 个，占介词与格构式实例的 1.3%。

由此可见，构式范畴表征既包含抽象的构式使用特征，又包含具体的构式实例。换言之，范畴既可以通过抽象图式表征，又可以通过具体样例表征。此外，范畴化需参照其他成员特征与原型的相似性，不同特征对范畴的贡献权重不同，因此判断原型性时，不仅要考虑范畴成员具有原型特征的多寡，还要考虑特征的权重。

4.2 汉语与格交替实例构式的原型性分析

原型实例是最具典型性和代表性的构式成员，其他实例与原型实例的相似性越高，原型性越强。换言之，其他实例与原型实例的距离可作为衡量其原型

性的手段(彭睿 2020:343)。因此,使用线性判别分析法,判断实例构式原型性的思路是:首先,根据实例的线性判别值,找出原型实例;其次,以原型实例为参照,通过比较其他实例与原型实例线性判别值的差异,确定其原型性,线性判别值越接近原型实例的,原型性越强。具体而言,判断双宾构式实例具有的原型性,可将双宾构式的原型实例作为参照点,比较其他实例与该原型实例线性判别值的差异。与原型实例线性判别值越接近,则与原型实例的相似性越高,具有的双宾构式原型性越强。例如:

- (5) 她说就算她心里有的人给不了她未来,她也不相信能给她未来的他,因为他们很久没联系。
- (6) 那位漂亮动人的小姐急忙走出来,亲自搬了一张椅子,请白朗宁坐下,说:“白朗宁先生,久仰您的大名了。”白朗宁只好送了她一个笑脸。

(5)的线性判别值为 0.42,其使用特征包括:与事为代词,表征有生事物和已知信息,指称有定事物;受事指称抽象事物,表征未知信息 and 无定事物。其与双宾构式原型实例相似的特征包括:与事为已知信息;受事指称抽象事物,表征未知信息 and 无定事物。(6)的线性判别值为 1.06,其使用特征包括:与事为代词,表征有生事物和已知信息,指称有定事物;受事指称抽象事物,表征未知信息 and 无定事物,长度长于与事。其与双宾构式原型实例相似的特征包括:与事表征已知信息;受事指称抽象事物,表征未知信息 and 无定事物,长度长于与事。将上述双宾构式实例的线性判别值可视化,可得图 2。

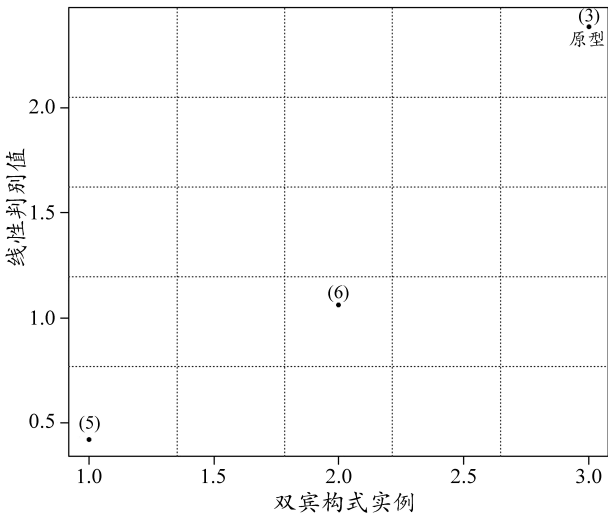


图 2 双宾构式实例的原型性可视化图

由此可知,相较于(5),(6)的线性判别值与双宾构式原型实例判别值更接近,与双宾构式原型实例相似性更高。因此,(6)的原型性比(5)更强。同理,判定介词与格构式实例具有的原型性,则以介词与格构式的原型实例为参

照，比较其他实例与该原型实例线性判别值的差异，与该原型实例的线性判别值越接近，则与原型实例的相似性越高，其原型性越强。例如：

- (7) 他的身边有人递了一杯酒给他，他接过来……
- (8) 她躺在床上看书时，脑海里居然浮现了这个名叫大力的男人递名片给她丈夫的情形……

(7)的线性判别值为-1.11，其使用特征包括：受事指称具体事物，表征已知信息和无定事物，长度长于与事；与事为代词，表征有生事物和已知信息，指称有定事物。其与介词与格构式原型实例相似的特征包括：受事指称具体事物，表征已知信息；与事为有生事物。(8)的线性判别值为-2.86，其使用特征包括：受事指称具体事物，表征未知信息和有定事物；与事表征有生事物和已知信息，指称有定事物，长度长于受事。其与介词与格格式原型实例相似的特征包括：受事指称具体事物，表征有定事物；与事为有生事物，长度长于受事。将上述介词与格构式实例的线性判别值可视化可得图 3。

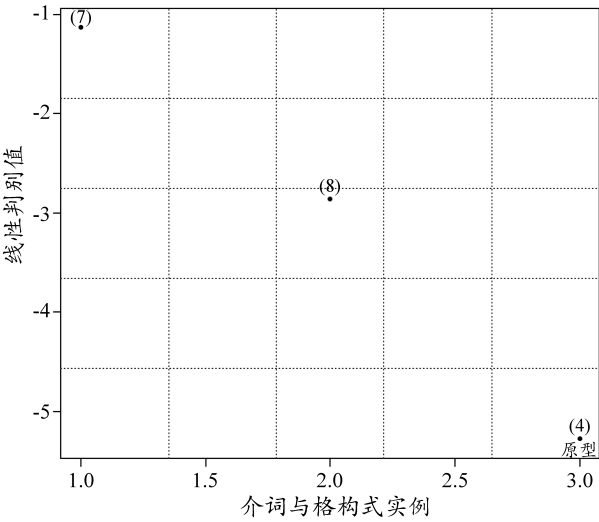


图 3 介词与格格式实例的原型性可视化图

由此可知，相较于(7)，(8)的线性判别值与介词与格构式原型实例判别值更接近，与介词与格构式原型实例的相似性更高。因此，(8)具有的介词与格构式原型性更强。

可见，相似性是范畴化的基础，范畴化依赖于新样例与原型样例之间的相似性，且相似性因语境而变化(Nosofsky 1986:39)。通过实例的线性判别值，可以量化比较不同实例间的相似性程度，从而确定其原型性。换言之，其他实例与原型实例线性判别值的差异是判断原型性的依据，与原型实例线性判别值差距越小，其原型性越强。具有不同程度原型性的实例构成连续统。

如前所述,本文构建的线性判别分析模型的分类准确率约为83%,换言之,该模型未能准确预测17%的与格交替实例属于双宾构式还是介词与格构式。观察模型分类错误的句子可以发现,其线性判别值均接近分类阈值(即0)。例如:

(9)“从心,我送一件礼物给你。”燕阳自枕头下取出一本深色小册子……

(10)他轻轻地说:“小榛,我想再送你一件衣服。”

(9)的线性判别值为-0.03,说明线性判别分析模型预测其为介词与格构式,该句中实际使用的也是介词与格构式。然而,(10)的线性判别值也为-0.03,模型预测其为介词与格构式,该句中实际使用的却是双宾构式。观察发现,上述两个实例既具有双宾构式原型实例的特征,如与事为已知信息,长度短于受事,受事表征未知信息和无定事物,又具有介词与格构式原型实例的特征,如受事指称具体事物,与事为有生事物。相应地,上述两例的线性判别值均接近分类阈值,说明在上述语境下使用介词与格构式和双宾构式均有很大概率,此时模型较难对其进行准确预测。譬如,若不考虑语用功能上可能存在的细微差异,(10)中说话者也可以使用介词与格构式“送一件衣服给你”。由此可见,有些实例兼具双宾构式和介词与格构式的部分特征,造成两个构式范畴之间边界模糊、互有交叉,此类边缘实例的范畴类别较难确定。换言之,双宾构式和介词与格构式的语义在某些情况下相近,无法截然分开。

综上所述,使用线性判别分析法,借助句法、语义和语用等诸多特征分析与格交替的影响权重和线性判别值,可清晰揭示与格交替范畴的结构特点,量化呈现构式原型性,为判断实例构式的原型性提供一种客观、有效的统计建模和可视化方法。

5. 结语

本研究基于大规模汉语与格交替实例,使用线性判别分析模型,考察了诸多句法、语义、语用特征对汉语与格交替构式的影响权重,统计了与格交替实例的线性判别值,分析了汉语与格交替构式的原型性。研究发现,使用线性判别分析法,有助于量化分析构式原型性,揭示构式范畴内部结构特征以及不同构式变体之间的关系。

本研究验证了线性判别分析法可有效分析汉语构式的原型性。然而,本文基于语料库数据,通过线性判别分析法对构式原型性的分析是否具有心理现实性,尚需结合心理实验进行交叉验证。采用行为实验或脑电实验等汇流证据对语料库统计模型进行补充验证是未来构式研究的重要拓展方向。此外,未来还可使用线性判别分析法对汉语其他构式以及跨语言构式对比展开研究。

附录：汉语与格交替构式实例分布

| 与格动词 | 双宾构式 | 介词与格构式 | 总计 |
|------|------|--------|-----|
| 拨 | 0 | 11 | 11 |
| 补 | 2 | 1 | 3 |
| 补偿 | 14 | 0 | 14 |
| 传 | 9 | 3 | 12 |
| 传授 | 22 | 2 | 24 |
| 赐 | 57 | 9 | 66 |
| 赐予 | 315 | 1 | 316 |
| 递 | 0 | 18 | 18 |
| 发 | 0 | 16 | 16 |
| 分 | 2 | 7 | 9 |
| 分配 | 2 | 1 | 3 |
| 付 | 10 | 1 | 11 |
| 告诉 | 52 | 1 | 53 |
| 给 | 27 | 3 | 30 |
| 归还 | 15 | 0 | 15 |
| 过继 | 0 | 5 | 5 |
| 还 | 16 | 6 | 22 |
| 汇 | 0 | 9 | 9 |
| 寄 | 0 | 5 | 5 |
| 奖励 | 21 | 2 | 23 |
| 交 | 3 | 7 | 10 |
| 交还 | 4 | 1 | 5 |
| 教 | 36 | 0 | 36 |
| 借 | 9 | 4 | 13 |
| 捐赠 | 1 | 4 | 5 |
| 卖 | 4 | 5 | 9 |
| 赔 | 14 | 1 | 15 |
| 赔偿 | 40 | 2 | 42 |
| 扔 | 1 | 4 | 5 |
| 赏 | 38 | 2 | 40 |
| 捎 | 0 | 39 | 39 |
| 送 | 69 | 27 | 96 |
| 提供 | 0 | 1 | 1 |
| 退 | 6 | 0 | 6 |
| 退还 | 34 | 0 | 34 |
| 献 | 0 | 5 | 5 |
| 移交 | 2 | 1 | 3 |
| 赠 | 14 | 2 | 16 |
| 赠送 | 17 | 6 | 23 |
| 转交 | 2 | 12 | 14 |
| 转卖 | 0 | 1 | 1 |
| 转送 | 2 | 4 | 6 |
| 租 | 1 | 2 | 3 |

参考文献 [References]

- 陈平, 1987, 释汉语中名词性成分相关的四组概念 [On four groups of concepts related to nominal components in Chinese]。《中国语文》第 2 期, 81-92 页。
- 邓思颖, 2003, 《汉语方言语法的原则和参数理论》[*A Parametric Theory of Chinese Dialectal Grammar*]。北京: 北京大学出版社。
- 顾阳, 1999, 双宾语结构 [Double object construction]。见徐烈炯编, 《共性与个性——汉语语言学中的争议》。北京: 北京语言文化大学出版社。
- 何晓炜, 2008, 双及物结构句式选择的制约因素研究 [A study on the factors influencing the selection of the ditransitive constructions]。《语言教学与研究》第 3 期, 29-36 页。
- 何晓炜, 2009, 双及物结构的语义表达研究 [A semantic study of the ditransitive construction]。《外语教学与研究》第 1 期, 18-24 页。
- 马庆株, 1992, 《汉语动词和动词性结构》[*The Chinese Verb and Verbal Constructions*]。北京: 北京语言学院出版社。
- 满在江, 2003, 生成语法理论与汉语双宾语结构 [Generative grammar and the double object construction in Chinese]。《现代外语》第 3 期, 233-240 页。
- 彭睿, 2020, 图式性构式的边界: 边缘构例和变异构例 [On the marginal instances and mutating instances of schematic constructions]。《世界汉语教学》第 3 期, 339-353 页。
- 沈家煊, 1999, “在”字句和“给”字句 [“Zai” construction and “gei” construction]。《中国语文》第 2 期, 94-102 页。
- 熊学亮, 2007, 英汉语双宾构式探析 [An analysis of the prototypicality of ditransitive constructions in English and Chinese]。《外语教学与研究》第 4 期, 261-267 页。
- 张伯江, 1999, 现代汉语的双及物结构式 [Ditransitive constructions in Chinese]。《中国语文》第 3 期, 175-184 页。
- 张懂, 2019, 英汉与格交替的多因素对比分析 [A multifactorial contrastive analysis of dative alternation between English and Chinese]。北京外国语大学博士学位论文。
- 张懂, 2020, 英语与格格式变体的多变量分析及其心理现实性研究 [The multivariate analysis of English dative alternation and its psychological reality]。《外语教学与研究》第 1 期, 40-52 页。
- 张懂, 2021, 英语与格交替原型使用特征的多变量统计建模 [A multivariate statistical modeling of the prototypical usage features of English dative alternation]。《现代外语》第 1 期, 49-61 页。
- 张懂、许家金, 2019, 英汉与格交替现象的多因素研究 [A multifactorial study of dative alternation in English and Chinese]。《外国语》第 2 期, 24-33 页。
- 张炜炜、王芳, 2017, 从基于样例的概念空间看构式交替——以“让”和“给”的被动用法为例 [Construction alternation: An exemplar-based conceptual space of Chinese passives with “rang” and “gei”]。《外语与外语教学》第 6 期, 22-33 页。
- 朱德熙, 1979, 与动词“给”相关的句法问题 [Syntactic problems related to the verb “gei”]。《方言》第 2 期, 81-87 页。
- Blank, Andreas and Peter Koch, eds., 1999. *Historical Semantics and Cognition*. Berlin: Mouton de Gruyter.
- Bresnan, Joan et al. 2007. Predicting the dative alternation. In Gerlof Bouma et al., eds., *Cognitive Foundations of Interpretation*. Amsterdam: Royal Netherlands Academy of Arts and Sciences.
- Bresnan, Joan and Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86, 1: 168-213.
- Bybee, Joan. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Croft, William and Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Divjak, Dagmar and Antti Arppe. 2013. Extracting prototypes from exemplars: What can corpus data tell us about concept representation? *Cognitive Linguistics* 24, 2: 221-274.

- Ellis, Nick. 2020. Usage-based theories of construction grammar; Triangulating corpus linguistics and psycholinguistics. In Jesse Egbert and Paul Baker, eds., *Using Corpus Methods to Triangulate Linguistic Analysis*. New York: Routledge. Pp.239–267.
- Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Goldberg, Adele. 2019. *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions*. Princeton: Princeton University Press.
- Gries, Stefan. 2003a. *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement*. New York: Continuum.
- Gries, Stefan. 2003b. Towards a corpus-based identification of prototypical instances of constructions. *Annual Review of Cognitive Linguistics* 1, 1:1–27.
- Labov, William. 1973. The boundaries of words and their meaning. In Charles-James Bailey and Roger Shuy, eds., *New Ways of Analyzing Variation in English*. Washington, DC: Georgetown University Press.
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: The University of Chicago Press.
- Langacker, Ronald. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald. 2010. How not to disagree: The emergence of structure from usage. In Kasper Boye and Elisabeth Engberg-Pedersen, eds., *Language Usage and Language Structure*. Berlin: Mouton de Gruyter.
- Liu, Feng-hsi. 2006. Dative constructions in Chinese. *Language and Linguistics* 7, 4:863–904.
- Medin, Douglas and Marguerite Schaffer. 1978. Context theory of classification learning. *Psychological Review* 85, 3:207–238.
- Murphy, Gregory. 2002. *The Big Book of Concepts*. Cambridge, MA: The MIT Press.
- Nosofsky, Robert. 1986. Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115, 1:39–61.
- Nosofsky, Robert. 1988. Similarity, frequency and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14, 1:54–65.
- Nosofsky, Robert et al. 1992. Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning Memory and Cognition* 18, 2:211–233.
- R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- Rosch, Eleanor. 1975. Cognitive representation of semantic categories. *Journal of Experimental Psychology* 104, 3:192–233.
- Rosch, Eleanor and Carolyn Mervis. 1975. Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology* 7, 4:573–605.
- Röthlisberger, Melanie et al. 2017. Cognitive indigenization effects in the English dative alternation. *Cognitive Linguistics* 28, 4:673–710.
- Taylor, John. 1989. *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Oxford University Press.
- Taylor, John. 1995. *Linguistic Categorization: Prototypes in Linguistic Theory*, second edition. Oxford: Oxford University Press.
- Taylor, John. 1998. Syntactic constructions as prototype categories. In Michael Tomasello, ed., *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*. Hillsdale, NJ: Lawrence Erlbaum. Pp.177–202.
- Venables, William and Brian Ripley. 2002. *Modern Applied Statistics with S*. fourth edition. New York: Springer.
- Wasow, Thomas and Jennifer Arnold. 2003. Post-verbal constituent ordering in English. In Günter Rohdenburg and Britta Mondorf, eds., *Determinants of Grammatical Variation in English*. Berlin: Mouton de Gruyter. Pp.119–154.

第一作者简介

张懂,男,博士,北京航空航天大学外国语学院副教授。主要研究兴趣为语料库语言学和认知语言学。代表作:“英语与格构式变体的多变量分析及其心理现实性研究”。电子邮件: zhang_dong@buaa.edu.cn

ZHANG Dong, male, PhD, is an associate professor at the School of Foreign Languages, Beihang University. His main research interests include corpus linguistics and cognitive linguistics. His major publication is “The multivariate analysis of English dative alternation and its psychological reality”. E-mail: zhang_dong@buaa.edu.cn

其他作者信息

许家金 北京外国语大学中国外语与教育研究中心 电子邮件: xujiajin@bfsu.edu.cn

致谢

2023年11月至2024年10月,以下在各自学术领域中成绩卓著的学者,在百忙之中为本刊匿名审稿,《当代语言学》编辑部同仁在此表示最诚挚的感谢!

巴丹、蔡淑美、曹道根、曹秀玲、陈前瑞、陈榕、陈树雯、陈伟蓉、陈禹、陈哲、陈振宇、程工、储泽祥、崔山佳、崔玉珍、邓盾、邓思颖、丁健、董秀芳、范晓蕾、方迪、高军、顾介鑫、郭继懋、郭龙生、郭锐、郭熙、郝琦、何彦诚、何雨殷、贺川生、洪波、侯兴泉、胡方、胡建华、胡苏红、胡旭辉、黄成龙、黄河、黄洁、黄立鹤、黄文林、黄瓚辉、蒋勇、焦一和、柯航、赖云帆、黎皓、李葆嘉、李兵、李桂兰、李慧敏、李劲荣、李可胜、李倩、李强、李若凡、李湘、李小军、李晓婷、李新良、李旭平、李宇凤、李战子、李章吕、李宗江、栗华益、梁丹丹、廖巧云、林华勇、刘畅、刘明明、刘探宙、刘须宽、龙国贻、卢达威、陆丙甫、罗琼鹏、罗天华、罗颖艺、马文、满海霞、敏春芳、潘俊楠、彭睿、饶宏泉、任荷、任鹰、邵洪亮、沈家煊、沈明、石毓智、史金生、史文磊、史有为、束定芳、宋文辉、宋作艳、孙崇飞、覃凤余、谭滢滢、汪化云、王晨、王春辉、王冬梅、王海波、王丽玲、王璐菲、王敏、王倩倩、王文斌、王鑫、温宝莹、吴炳章、吴福祥、吴早生、伍雅清、夏全胜、项开喜、谢心阳、邢向东、熊文新、熊子瑜、休佳诺伊、徐宇航、许立群、玄玥、薛嗣媛、杨大然、杨坤、杨萌萌、杨牧、杨炎华、姚小平、叶婧婷、应学风、于秀金、余光武、袁梦溪、袁毓林、乐耀、詹卫东、张安琪、张伯江、张聪、张懂、张华、张姜知、张璟玮、张竞婷、张黎、张莉、张庆文、张谊生、张永伟、张云秋、张智义、张子豪、赵彬、赵长才、赵川兵、赵春利、赵绿原、赵明鸣、赵青青、真大成、郑娟曼、钟华、周晨磊、周韧、朱冠明、朱佳蕾、朱晓农、朱祖德、邹玉华、祖生利、祖漪清。