

基于语料库的中日同形词类联接和语义韵对比研究——以“获得”一词为例

郑依迪

上海海事大学外国语学院, 上海

收稿日期: 2022年9月2日; 录用日期: 2022年10月19日; 发布日期: 2022年10月26日

摘要

中日文的“获得”一词为同形词,但在翻译的时候,有时不能或不宜互译为“获得”。如果我们不仔细区分,很容易在翻译中产生一些偏差。因此本文将从类联接和语义韵两个方面对中日同形词“获得”在中文和日文语料库中的使用实例进行分析,探索使用语料库区分汉语和日语同形词的方法,并探讨其语料库中形成的语义韵的异同。

关键词

语料库, 类联接, 中日文对比, 语义韵

A Corpus-Based Comparison of Colligation and Semantic Prosody between Chinese and Japanese Homographs—A Case Study of the Word “获得”

Yidi Zheng

College of Foreign Languages, Shanghai Maritime University, Shanghai

Received: Sep. 2nd, 2022; accepted: Oct. 19th, 2022; published: Oct. 26th, 2022

Abstract

The word “获得” in Chinese and Japanese is a homonym, but sometimes it cannot or should not be directly translated into “获得” when the two languages are translated into each other. If we do not

distinguish carefully, it is easy to produce some deviations in translation. Therefore, this paper will analyze the examples of the use of Chinese and Japanese homographs “acquired” in Chinese and Japanese corpora from the two aspects of generic links and semantic prosody, explore the methods of using corpora to distinguish Chinese and Japanese homographs, and explore the similarities and differences of semantic prosody formed in their corpora.

Keywords

Corpus, Colligation, Comparison between Chinese and Japanese, Semantic Prosody

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

词汇中存在非常丰富的日汉同形词,大河内康宪(1997)对日汉同形词下的定义是“いずれがいずれを借用したかを問わず、双方同じ漢字で表記されるものを同形語と呼ぶ(不管是汉语借自日语还是日语借自汉语,双方用相同的汉字标记的都称为同形词)。”[1]由于词义的相似性或重合性,中国日语学习者在学习这些词的过程中容易产生理解或使用上的混淆。主要原因是对词汇的关键要素的理解不够。因此,本文以一组词“获得”为例进行调查,将该组词放入特定语料库中,基于数据示例来研究词的特征。

2. 文献综述

Firth (1957: p. 95)定义类联接(colligation)为“语法范畴间在句法结构上的相互关系”[2]。濮建忠(2003)指出类联接用来描述一个词的语法形制(grammatical patterns)。搭配(collocation)指的是在类联接下词与词之间的共现关系,即重复出现两个或多个词,尤其是名词、动词、形容词和副词在某一语法系统下反复出现。[3]。

词汇项的搭配行为表现出一定的语义趋势:某些词汇项会习惯性地吸引具有相同语义特征的某类词汇项形成搭配。因为这些具有相同语义特征的词和关键词在文本中频繁出现,后者会被相关语义特征“感染”,一定的语义氛围会渗透整个语境,这就是语义韵(semantic prosody)。语义韵可以分成积极(positive)、中性(neutral)和消极(negative)三类。根据卫乃兴(2002)的研究,动词“cause”表现出强烈的消极语义韵。在“cause”的V+N类联接下,在N位置上总出现的搭配词包括 accident, trouble, damage, failure, problem, 等。这些词均指在日常生活中人们的消极情绪或令人不快的事。它们在节点词周围形成强烈的消极语义氛围[4]。由此可见,搭配与语义韵也密切相关。

卫乃兴(2002)提出了三种方法语义韵的研究方法。主要包括:1)使用基于数据的方法建立类连接;2)用数据驱动方法计算节点词的搭配;3)使用基于数据和数据驱动的方法[5]。三种研究方法分别揭示了研究思路和具体步骤,分析了每种方法的特点和实用性,为语义韵研究创造了价值。

对语义韵的研究虽然大多集中在英语方面,近年来,也有少数学者在日语方面对语义韵展开了研究。李倩(2018)基于JpTenTen11日语语料库,分析了两个日语近义词「行為」及「行動」的搭配型式,揭示出两者的语义韵差异[6]。王华伟、曹亚辉(2012)选取一组近义词“かえりみる”“ふりかえる”“ふりむく”,采用统计学的方法,以名词搭配为重点,分析语料库中每个动词的搭配特点、搭配词的搭配能力等,以探讨词汇搭配研究在日语教学中的意义[7]。童富智,修刚(2020)以扩展意义单位模型为框架,

分析比较了动词“坚持”及其日语对应词“堅持(する)”构成的多词序列语义韵的异同,发现这两个词在每个语料库中的搭配形成了不同的共同选择模式,呈现不同的语义韵[8]。

3. 研究方法

本研究是一项基于语料库的跨语言语义韵对比研究,旨在探索和讨论中日同形词。研究以汉语和日语通用语料库为基础,以汉语动词“获得”和日语对应词“獲得(する)”(以下简称为“獲得”)为节点词,观察其在两个语料库中各自形成的类链接与搭配词,总结各自的语义韵。

3.1. 语料库

本文运用的汉语语料库为中央研究院汉语平衡语料库,收录了哲学、文学、生活、社会、科学艺术等多方面的语料实例;日语语料库为日本国立国语研究所开发的现代日本語書き言葉均衡コーパス(BCCWJ)语料库,包含书籍、杂志、报纸等十余种类型的文本,库容近一亿五千万日文词。

3.2. 研究步骤

本研究采用的是建立并参照节点词的类链接,并基于数据来概括和描述节点词语搭配和语义韵的方法。首先,在汉语平衡语料库和 BCCWJ 日语语料库中分别以两个动词作为节点词检索其语言实例。之后对节点词在索引行中出现的语法形制进行分析和总结,建立关于“获得”和“獲得”的类链接。由于两个节点词均是两语言中的常用词,在其语料库中检索到的实例过多,不利于对节点词的搭配和语义韵进行详细的分析,因此,笔者采取了随机抽样的方法对中日语料库分别抽出 200 条索引句。数据显示,“获得”在汉语平衡语料库中共出现 2488 次,于是随机抽取编号为 1~100, 501~600 的 200 条索引句,再排除其非动词的实例(例如:“获得的民间捐款”)共得到 189 条索引句,全部用于数据分析;“獲得”在 BCCWJ 日语语料库中共出现 3631 次,采取与汉语相同的方法随机抽取编号为 1~200 的 200 条索引句,排除其非动词的实例(例如:“地方参政権の獲得”地方参政权的获得),共得到 152 条索引句,全部用于数据分析。接着再对节点词的搭配词进行语义分类,并归纳它们的语义韵特点。通过对语料库中实例的样本处理,两组文本形成对应的定量关系,可以减少数据处理量,使研究更具可操作性。

4. 研究结果

4.1. “获得”的类链接及语义韵

由于篇幅限制,笔者从在汉语平衡语料库中随机抽取的“获得”的 189 条索引句中,列举 10 例如下。

1) 的女人婚前对其丈夫的称呼一样。后来,在他在世界各地获得辉煌胜利后,他们就开始称他为卡斯楚,那是出于严肃的

2) 学教堂是与华裔建筑师贝聿铭的共同创作,双曲面的设计获得全球建筑界肯定,近年他参与多项公共工程规划,但他觉

3) 的文建会专门委员杨式昭说明,艺术村构想方案已于日前获得行政院的核准,因此她日前邀约国内知名建筑设计家黄永

4) 一道中文翻译的手续,两岸学者皆希望文字歧异性能早日获得解决。前海基会文化处处长周逸衡表示,他曾于七月底赴

5) 今年的林本源中华文化教育奖,由师大美术系教授陈银辉获得,其他获奖人有林星华、郑管麟、简锦清(国画类),萧

6) 岸艺术中心展览部主任吉尔表示,朱铭的作品,自展出后获得热烈反应,因此他在获得朱铭与赞

助单位港英友谊协会的

7) 作品, 挑出一百件送往日本大坂国际交流中心参加展览, 获得金牌奖的是台北县永福国小的江暖暖。

8) 品参加日本廿一世纪协会主办的亚洲儿童画家绘画比赛, 获得一面金牌、十六面银牌。吴隆荣指出, 此次该会向全国国

9) 英国皇家南岸艺术中心举行雕塑展, 自展出后, 作品不断获得当地媒体好评, 他目前人已前往法国布置展出的作品。

10) 位同学, 四年一班陈柏仁小朋友, 代表全校参加台南市, 获得全市第三名, 现在请校长上台颁奖。哇! 太好了, 我又得奖了

索引证据是研究的基础。索引提取后, 可以观察关键词的搭配。上述 10 个索引显示, “获得”在汉语平衡语料库中主可用于 N + Adv + V + N, V + N, N + V + N, N + V 四种形式的类链接。第一种形式的类链接是多数, 有索引 1, 3, 4, 6, 9 这 5 例, 占全部索引的 50%。后三类搭配合计占 50%。其中, 可以观察到其名词搭配词有金牌, 好评, 胜利等, 这些名词大部分都是一些具有积极语义特点的词汇。由于上述四种类链接中动词 V 左边的名词 N 位置上的搭配词皆为表示人物的名词, 在句子中处于主语的位置。因此下面观察以更细致的视角分析在上述四种类链接下动词 V 右边 N 位置上的处于宾语位置上的名词搭配词, 它们的语义特征及出现次数等特征见表 1:

Table 1. Noun collocation of the word “获得”

表 1. “获得”名词的搭配情况

搭配词的语义分类	搭配词举例	搭配词出现次数	共现附加语举例	语义韵分类
表示知识, 技能, 金钱等意义	知识, 奖励, 利息, 支助, 经费, 补助, 稿酬, 报酬, 利润,	35		积极语义韵
表示学历, 名次, 地位, 成功等意义	金牌, 胜利, 免试, 晋级, 冠军, 成功, 表扬, 专利, 前三名	51	终于, 不断, 已经, 早日, 今天,	积极语义韵
表示认可, 关注, 满意, 支持等意义	认同, 配合, 重视, 鼓舞, 肯定, 信赖, 喜爱, 尊重, 重用, 好评	57		积极语义韵
其他意义或意义模糊	回应, 报道, 印证, 答案, 变更, 观念	46		中性语义韵
合计		189		

从表 1 中可以看出, “获得”左侧常见“一段时间”或“新的时间点”的共现附加语, 这类词语突出“获得”行为的时间特征是“长久的”或“新的, 第一次的”, 这类时间搭配词与“获得”一起共现, 说明“获得之物”是经过长时间努力达到的或具有开创意义的事。并且“获得”在四种类链接下, 右侧搭配名词在前三种语义范围内大多表示人们希望获得的金钱, 地位, 认可和支持, 共 143 个搭配词, 约占全部 189 条索引句的 76%, 具有明显的积极语义特征(例句 11.12.13)。其次, 从回应, 报道, 印证等词汇来看, 这些搭配词与“获得”共现, 使得句子弥漫着中性语义韵, 约占全部索引句的 24% (例句 14.)。

11) 在一九九二年与 Dr. Edmond Fischer 共同获得诺贝尔奖, 更肯定了他在这方面的贡献。这次他与本所研

12) 奖助学金颁发要点修正案, 凡获得奥运金牌的选手, 将可获得新台币一千万元的奖金。行政院会昨天通过教育部所提中

13) 岸艺术中心展览部主任吉尔表示, 朱铭的作品, 自展出后获得热烈反应, 因此他在获得朱铭与赞助单位港英友谊协会的

14) 动, 读者可在网路上用电子信件针对个别新闻提出问题并获得答案。网路互动式大学报的刊登和管理方式尚在研拟中

4.2. “獲得”的类链接及语义韵

与汉语的研究方法相同, 利用 BCCWJ 日语语料库将之前随机抽取到的 152 条索引句列举 10 例如下 (由于语料库中索引句前后并不完整, 因此笔者仅对索引句中关键词“獲得”所在句进行中文释义。):

15) ろいろな症候群であるといえる。したがって、高山病に上手にかかり、高度順応を確実に獲得することが、高山病の予防ならびに治療の最良のほうほうとなる。(因此, 要善于治疗高山病, 确保高度适应是预防和治疗高山病的最佳方法)

16) ヨナルスクールに通うすべての子どもたちの実態ではありません。日本語と英語の両方を獲得し、自由に使いこなしている人も大勢います。ただ、外国語の環境に置いても複数の言語(也有很多人获得了日语和英语两种语言, 并且运用自如。)

17) に細胞にあらかじめ温熱処理をすると、細胞は次の温熱処理に対して一時的な温熱耐性を獲得する。温熱耐性を獲得した細胞は、熱ショックや種々のストレスによる細胞障害を防御し(细胞对下一次加热处理获得暂时的热能耐性。获得热能的细胞可以防止因热休克和各种压力引起的细胞障碍)

18) るにつれて、住民の信頼が遞減する傾向にある。三者のなかで住民のもっとも高い信頼を獲得したのは、香港政庁であった。かくて、一九八四年の英中共同声明以後、香港住民のな(在三者中, 获得居民最高信赖的是香港政府机关。)

19) 知らず、乱れた日本語を使う私たちの方が滑稽に感じられることさえあります。言語を獲得し始める時期の子どもには、自我の芽生えによる強い「衝動」や「感情」に伴って発語が(对于刚开始学习语言的孩子来说, 伴随着自我萌生的强烈“冲动”和“感情”而发言)

20) てはすまされません。さらに言えば、高いインセンティブを示して優秀な人材を積極的に獲得していききたいという思いがあります。成果主義を導入していない証券会社は早晚つぶれる(更进一步说, 我想通过展示高额激励来积极地获得优秀人才。)

21) 交通網の整備は、国内市場を深化させてきたが、加えて重商主義的抗争のなかで覇権を獲得したイギリスは、広大な海外市場に恵まれ、これらがマニュファクチャーの限界性の突破(交通网的整備, 使国内市场深化, 再加上在重商主义的抗争中获得了霸权的英国, 拥有广阔的海外市场)

22) のものが突然変異で生じた瞬間に、その神経細胞ネットワークを持った個体は無限概念を獲得したのではないだろうか。そうすると、有限な宇宙が無限に張られた空間のなかの微小な(拥有神经网络网络的个体不是获得了无限概念吗。)

23) 問題が詳細な学習にふさわしいか、生徒がプロセスを通して最適のレベルの知識・技能を獲得するよう、十分にやりがいがあるかどうか注意すべきである。評価は、アイデアをテ(问题是否适合详细学习, 要注意是否有足够的工作价值, 以便学生利用这个过程获得最佳水平的知识和技能)

24) いる点に特徴がある。彼らの多くも「地方」出身の非学歴エリートである。雑誌の読者を獲得するためにとった彼らの戦略が、自分と同じような地方勤労青年を対象に、自身の出自や(为了获得杂志读者而采取的他们的战略, 以和自己一样的地方劳动青年为对象)

上述索引显示,“獲得”可用于 N+ を + V, N+ を + 形容动词 + V 两种形式的类链接。由于日语拥有特殊的语法结构, 宾语一般出现在谓语动词的前面。N+ を + V 形式的类链接占了大多数, 有索引 16.17.18.19.21.22.23.24.共 8 例, 占全部索引的 80%。N+ を + 形容动词 + V 是在动词 V 前面加上修饰性词汇, 形容动词属于形容词的一种, 这类形式的类链接虽然不多, 在 10 条索引中仅出现 2 例, 但其前面的修饰词比如“積極的に(积极地)”, 表达出“主动的”等意思, 与动词 V 搭配呈现出积极地语义氛围。另外, 动词 V 前面的处于宾语位置 N 的搭配词的情况有: 信赖, 霸权, 知识技能, 日语英语, 热能耐性等。可以发现日语跟汉语有所不同, 日语中“獲得”的搭配词十分丰富, 难以仅仅通过 10 个索引句来判断, 因此下面将分析两种类链接下全部索引中“獲得”的名词搭配词的出现情况, 它们的语义及出现次数等特征见表 2。

Table 2. Noun collocation of the word “獲得”

表 2. “獲得” 名词的搭配情况

搭配词的语义分类	搭配词举例	搭配词出现次数	共现附加语举例	语义韵分类
表示知识, 技能, 金钱等意义	知識, 利益, 經濟効果, 賞金, 収益	33	積極的に, 十全に	积极语义韵
表示学历, 名次, 地位, 成功等意义	金メダル(金牌), MVP, 一位, 特 A, 最高賞	30		积极语义韵
表示科技, 生物等领域的专门用语	体細胞, 弛緩状態, 温熱耐性, 耐性, 免疫, 抵抗性	34	確實的に, 後天的に, すぐに, 一度	中性语义韵
表示违法, 犯罪等意义	霸権不法,	9		消极语义韵
表示其他意义或意义模糊	抽象語, 個室, 両足歩行, 土地近代性	46		中性语义韵
合计		152		

从表 2 可以看出“獲得”在两种类链接下, 在表示知识, 技能, 金钱与表示学历, 名次, 地位, 成功语义范围内, 与汉语相同, 其搭配词具有明显的积极语义特点, 大多表示人们希望获得成功, 金钱等, 共 63 个, 占全部索引句的 41% (如例句 25.26)。搭配词在表示科技, 生物等领域的专门用语和其他意义或意义模糊类语义范围中, 80 个搭配词无明显的积极或消极氛围, 占全部索引的 50% 以上, 其中表示科技, 生物等领域的专门用语大多是高科技技术领域内的词汇, 多出现在自然科学类书籍中(如例句 27)。这类词汇占全部索引句的 22%。搭配词在表示违法, 犯罪等意义时则明显表现出消极语义特征(如例句 28), 约占全部索引的 6%。对这几个百分比进行对比可以发现, “獲得”搭配词的积极语义韵占 41%, 中性语义韵占 53%, 消极语义韵占 6%。因此, 可以判断其呈现出中性语义韵。

25) て第一線の選手としてギャラリーからの人気を得つづけ、総額一〇〇万ドル以上の賞金を獲得した。やがて、トーナメントで競技するのがだんだんむずかしくなってきた。身長一七(作为第一线选手, 他一直在画廊里拥有人气, 获得了 100 万美元以上的奖金)

26) ばれた。世界の 3 大映画祭(ほかにカンヌ、ベネチア)でアニメーション映画が最高賞を獲得したのは初めて。ベルリン映画祭での日本映画の最高賞受賞は 1963 年の今井正監督の最高賞受賞は 1963 年の今井正監督の(在世界 3 大电影节上, 动画电影获得最高奖尚属首次)

27) のものが突然変異で生じた瞬間に、その神経細胞ネットワークを持った個体は無限概念を獲得したのではないだろうか。そうすると、有限な宇宙が無限に張られた空間のなかの微小な(拥有神经网络网络的个体不是获得了无限概念吗)

28) 市民の日常生活や経済取引に民事上の権利者や関係者の形で介入、関与し、不法な利益の獲得を図る「民事介入暴力事案」は、年々増加の傾向を示しており、また、政治活動や社会活(以民事权利人和相关人员的形式介入、参与市民的日常生活和经济交易，谋求获得非法利益的“民事介入暴力事件”呈现逐年增加的倾向)

4.3. 讨论

综上所述，从对汉语平衡语料库和 BCCWJ 语料库的分析看，中日同形词“获得”的使用可以从以下两个方面进行辨析：

A 类联接的形式上有所不同。汉语有四种类联接，并且节点动词的左边通常为主语人物事物，右边为各种意义不同的搭配词。日语的类联接种类比较单一，N + を + V 是基本形式，另一种形式是在基本形式的基础上在节点动词左边附加上修饰性的形容词。

B 语义韵方面有区别。汉语在四种类联接下，其名词搭配词大多具有强烈的积极语义特征。这些搭配词使其语境内充满了强烈的积极语义氛围。相比之下，日语则与表示金钱、知识、地位等积极意义的词语搭配远不及汉语那么频繁。此外，日语中还出现了与消极意义的词语搭配。简言之，汉语的“获得”是具有积极语义韵的词，而日语的“獲得”是具有中性语义韵的词。

5. 结语

本文首先基于汉语平衡语料库和 BCCWJ 日语语料库研究了中日同形词“获得”在各自语料库中的使用情况，对二者的类联接语义韵特点进行了描述和分析。虽然这项研究的规模较小，但它表明，尽管一些中日同形词在意义上看起来非常相似，但在词汇搭配和语义韵方面仍存在一些差异。通过语料库和语料库索引工具，我们可以从大量的实际用例中发现和总结词汇的特征，并观察其上下文和搭配的特征。由于篇幅限制，本文仅分析了一对动词。未来的研究可以基于更大的汉语和日语可比语料库，进一步探索其他中日同形词的问题。

参考文献

- [1] 大河内康憲. 日本語と中国語の同形語[C]//日本語と中国語の対照研究論文集. 東京: くろしお出版, 1997.
- [2] Firth, J.R. (1957) *Papers in Linguistics 1934-1951*. Oxford University Press, London.
- [3] 濮建忠. 英语词汇教学中的类联接、搭配及词块[J]. 外语教学与研究, 2003, 35(6): 438-445.
- [4] 卫乃兴. 语义韵研究的一般方法[J]. 外语教学与研究, 2002(4): 300-307.
- [5] 卫乃兴. 语料库数据驱动的专业文本语义韵研究[J]. 现代外语, 2002(2): 165-175.
- [6] 李倩. 基于大规模语料库的日语近义词辨析——以「行為」与「行動」为例[J]. 日语学习与研究, 2018(3): 25-33.
- [7] 王华伟, 曹亚辉. 日语教学中基于语料库的词语搭配研究——以一组动词近义词为例[J]. 解放军外国语学院学报, 2012(2): 71-75+91.
- [8] 童富智, 修刚. 基于语料库的中日对应词语语义韵对比研究——以中央文献日译「堅持(する)」为例[J]. 日语学习与研究, 2020(2): 1-8.