

“V_双 + 时间”搭配难度等级分类与日本汉语教学

李 佳

大阪大学人文学研究科, 日本 大阪

收稿日期: 2022年8月9日; 录用日期: 2022年9月14日; 发布日期: 2022年9月21日

摘 要

为了考察日本汉语学习者的基本词语搭配范畴, 本论文以“V_双 + 时间”为例, 进行了尝试性的实例分析。本论文通过三个方面对“V_双 + 时间”搭配进行了分析。第一, 依据语料库进行统计指标分析; 第二依照日本中国语检定考试的词汇等级分类对“V_双”进行归类; 第三, 考察了“V_双 + 时间”的猜词难度。将主观标准和客观标准相结合, 本论文结论将“V_双 + 时间”尝试性地划分为四个等级。

关键词

语料库, 词语搭配, V_双 + 时间, 互信息, 语义透明度

Classification of the Difficulty of “Disyllabic Verb + Time” and Teaching Chinese in Japan

Jia Li

Graduate School of Humanities, Osaka University, Osaka Japan

Received: Aug. 9th, 2022; accepted: Sep. 14th, 2022; published: Sep. 21st, 2022

Abstract

In order to examine the scope of basic words collocation of CSL in Japan, tentative analyses on the case study with the example of “disyllabic verb + time” were carried out in this paper. In this paper, disyllabic verbs collocating with time were analyzed in three aspects. Firstly, the analysis is conducted into the statistic indices on the basis of corpus; secondly, classification is made for

“disyllabic verb” in accord to the glossary grade classification in Japan’s the Chinese language proficiency test; thirdly, the semantic transparency of “disyllabic verb” is examined. Combining the subjective criteria with the objective ones, this thesis’ conclusion tentatively divides “disyllabic verb + time” into four grades.

Keywords

Corpus, Collocation, Disyllabic Verb +Time, MI Score, Semantic Transparency

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

“搭配是一种意义方式，由词之结伴可知其词” (Firth, 1957) [1]，词语的结伴关系或共现关系是词语搭配的重要表现形式。Halliday 和 Sinclair 从词汇角度研究搭配，进一步发展 Firth 的搭配观念：搭配“体现词项在某种显著的临近范围内组合关系的线性共现” (Halliday, 1976) [2]。该定义把搭配界定为组合关系，将“线性共现”视为搭配最重要或者唯一的界定标准。Cruse 从语义角度理解搭配，将搭配界定为“习惯性共现的词语序列” (Crak & Lockhart, 1972) [3]。

人们在使用语言时，更多的是利用一些预制好的语言片段，这些预制片段作为一个整体被理解、记忆、存储和提取。语言的记忆和存储、产出和使用并不是以单个的词为单位的，那些固定和半固定模式化的板块结构才是人类语言交际的最小单位(Becker, 1975) [4]。这些重复出现的组合或预置的、半预置的词组被称为搭配。Nation 认为“搭配知识是词汇知识的重要组成部分，搭配在语言交际中无处不在，是文本构成的基本要素，是意义表达的基本实体单位” (Nation, 1990) [5]。随着二语习得研究的深入，词语搭配在语言学习中的作用越来越受人们的重视。“词语搭配知识被认为是学习者整体语言能力的重要组成部分” (Stubbs, 1996) [6]。然而 Marton (1977) [7]的研究表明，词语搭配错误在第二语言学习者言语失误中占很高比例。Nesselhauf (2003) [8]指出，二语学习者不能够使用一些习惯性的搭配，语言学习中发生的错误最好从搭配的角度加以解释，即可以通过介绍词项及其习惯性语境学习词语搭配。可见在二语习得中，词语搭配的习得至关重要。

词语搭配在使用上习惯性很强，规律性相对较弱，用规律难以概括，其中动宾搭配更以其使用的高频性、复杂性、灵活性受到学者的广泛关注。在日本，由于各个大学使用的汉语教材种类各不相同、教学方法存在差异，使得在实际教学中教授的词语搭配的种类、数量、频度也不尽相同。那么，对日本学习者来说，哪些搭配是典型搭配呢？本文以“V_双 + 时间”为例，在前人的研究基础上，进行了尝试性地分析和分类。

2. 现有研究介绍

随着语料库语言学的兴起，强大的语料库证据及其科学有效的研究方法和手段使词语使用的典型语境和形式都呈现在研究者面前，研究者可以依据大量的真实的搭配实例，对各个词语及其搭配的典型意义进行分析、概括和归纳，基于语料库的搭配研究在词汇教学中的重要性日趋明显。借助语料库，出现了一些基于统计模型的词语搭配自动获取方法，分析汉语搭配的统计指标和新标准被学者们所关注。

2.1. 统计指标

孙茂松等(1997) [9]在《汉语搭配定量分析初探》中,借鉴了国外在语言学和语料库语言学两个方面关于词语搭配的研究成果,提出了包括强度、离散度及尖峰值三项统计指标在内的搭配定量评估体系,构造了相应的搭配判断算法。卫乃兴(2002) [10]重点介绍了词语搭配研究中常用的三种统计方法,特别是对词语搭配研究中常见的 MI 值和 T 值的计算方法做了详细的介绍,并对比分析了每一种方法的优点和缺点。卫乃兴(2003) [11]、石川(2008) [12]详细介绍了西方词语搭配的不同研究体系以及各研究体系的特点及其差异,并概括了搭配概念演变的脉络与研究方法的历史。全昌勤(2005) [13]对搭配的定义及搭配在自然语言处理中的应用做了介绍,着重分析了目前广泛使用的其余统计模型的词语搭配自动获取方法,并对衡量搭配强度的六种统计指标进行了详细的分析和比较。

常用的判断词语搭配的统计指标,除了搭配频次以外,还有 T 值(T score)、MI 值(mutual information score)、对数似然函数值(Log-likelihood ratio)、Z 值(Z score)、骰子概率系数、卡方检验等等,其中, T 值和 MI 值最为常用。杨百翰大学 BYU 的 Mark Davies 制作的网上数据库就是依据以上两个统计指标完成的。MI 值是信息论中的一个概念,用来衡量两个事物的相关程度,在语言学中通常用来衡量两个单词之间的可预见度(strength of association),通过 MI 值能找出频度低但黏着度高、特殊性大的搭配。关于这些判断标准,更多的学者认为,不应单独使用,应综合使用或根据研究内容和目的有选择地使用统计指标。

我们依据以上学者的研究成果,将词语搭配的分析标准做以归类,见表 1。

Table 1. Two main criteria for analyzing word collocation

表 1. 分析词语搭配的两大大标准

客观标准	词与词的搭配频次 词与词的搭配强度、搭配可信度 词与词的搭配的连续性和距离
主观标准	构成成分意义的明晰性 构成成分的意义与搭配意义的对应性

2.2. 语言提取原则

以往的语言研究认为语言是人们按照语法规则将词汇组合在一起而形成的。但是研究发现,语言中存在很多预制结构,在提取时不需要由语法规则生成,而是直接提取。基于此, Sinclair (1991) [14]提出了两种模型:惯用原则(idiom principle)和开放选择原则(open-choice principle)。前者认为,意义的实现主要依赖于一次选择各种半预制词组;后者则认为,结构空位的要求使得选择基于单词进行。Sinclair 认为语言学习者记住并使用的大多是可单独使用的半固定搭配,其搭配上还存在一定的限制,在实际的场景中,惯用原则处于优势地位。Hoey (2005) [15]经过研究发现高频搭配是构成语言自然度的关键因素,也就是说在语言学习中,如果掌握了典型搭配即高频搭配,学习者产出语言的流利度和准确度就会增加,因此典型搭配在词语搭配习得中具有重要的作用。

3. 分析过程

3.1. 目的

名词“时间”在各语料库中均为高频名词,属于中日同形同义词(日语为「時間」),因此本文将“V_双 + 时间”作为分析对象。从客观标准、主观标准和日本汉语学习者的教育背景三方面出发,以“V_双 + 时间”这一动宾搭配为例,进行分析和归类。

3.2. 分析方法

方法一：高频搭配的收集。本次考察中，首先利用了现代汉语语料库 TORCH2009¹，进行搭配指标的统计。为了找出搭配可预见度高、搭配频次概率高的搭配，我们统计出了共现频次(Co-occurrence frequency)和互信息值(Mutual information score)各居前 15 位的“V_双 + 时间”。其次，我们统计出了北京语言大学现代汉语语料库 BCC (多领域)中，共现频次居前 10 位的“V_双 + 时间”搭配，并计算出该 10 个“V_双 + 时间”与“V_双”在 BCC 中出现频次的百分比。最后，将两个语料库中的高频次、高强度“V_双 + 时间”进行统合。

方法二：日本汉语学习者词汇学习、教学的实际情况。首先，将依据日本中国语检定考试的等级高频“V_双 + 时间”中的“V_双”分类。日本中国语检定考试是立足于日本汉语学习者的学习环境、以日本国内为主实施的一项考试，其词汇等级分类特征在一定程度上反映了日本汉语教育的视点。

方法三：“V_双 + 时间”的猜词难易度分析。

4. 分析结果

4.1. 统计指标

计量统计可以为理论研究提供一定的依据，也可以更加明晰地彰显规律。由于考虑到把字句、被字句、宾语前置等情况，本文设定词间跨度为“时间”的左、右各 4 个词(Span: +/-4)；本文研究的对象是典型搭配，因此我们把搭配频次(Freq.)规定为 ≥ 3 ；Hunston (2002)把搭配强度(MI score) ≥ 3 的词作为显著搭配词，本文亦遵照该理论。基于以上理由，我们以两项统计指标为依据，首先分别统计出了 TORCH2009 语料库中搭配频次(Freq. ≥ 3)的居前 15 位的“V_双 + 时间”、以及搭配强度(MI score ≥ 3)居前 15 位的“V_双 + 时间”。在统计时，将非动宾形式的搭配排除在外。例如“持续时间”、“准备时间”等，这些搭配形式上虽为“V_双 + 时间”，但搭配关系实际为状中结构，此后排列依次提前。如表 2 所示。

Table 2. “V_双 + 时间” in TORCH 2009 with the top 15 frequency
表 2. TORCH2009 中搭配频次居前 15 位的“V_双 + 时间”

排名	搭配频次(Freq ≥ 3)		互信息值(MI score ≥ 3)	
	V _双	Score	V _双	Score
1	需要	28	挤出	5.79
2	没有	28	耗费	5.74
3	抓紧	7	花费	5.66
4	利用	6	缩短	5.47
5	花费	6	浪费	5.47
6	浪费	5	推迟	5.47
7	计算	4	抽出	5.38
8	失去	4	打发	5.15
9	约定	3	压缩	4.97

¹TORCH2009 语料库是由中国 64 所以上高校的 115 位老师和硕士生及博士生参与语料收集和校对，共同创建的现代汉语语料库。该语料库大小为 1,087,619 词, 1,703,635 字。该语料库中所收文本绝大部分为 2009 年出版。语料库的名称 TORCH 为 Texts Of Recent Chinese 的缩略词。<http://111.200.194.212/cqp>

Continued

10	争取	3	抓紧	4.85
11	延长	3	约定	4.53
12	缩短	3	延长	4.36
13	压缩	3	度过	3.85
14	耗费	3	分配	3.53
15	²	-	争取	3.15

根据表 2, TORCH2009 中, 两个统计指标共现于前 15 位的有 9 个, 分别是“抓紧时间”、“花费时间”、“浪费时间”、“约定时间”、“争取时间”、“延长时间”、“缩短时间”、“压缩时间”、“耗费时间”。我们认为这些“V_双 + 时间”在 TORCH2009 中属于高频、高搭配强度。

同时, 我们统计出 BCC 语料库(多版本) [16] 中居前 10 位的“V_双 + 时间”搭配, 并按照与“V_双”的出现频次在“V_双 + 时间”出现频次中所占的百分比, 从高到低进行排列, 如表 3 所示。

Table 3. Percentage of “V-double + time” and “V_双 + 时间” frequencies of the top 10 pairings in BCC
表 3. BCC 中搭配频次居前 10 位的“V_双 + 时间”与“V_双”频次的百分比

	V _双	与“时间”的搭配频次 ²			“V _双 ”的出现频次	%
		Span = 0	Span > 0	合计(Span ≥ 0)		
1	耽误	535	1413	1948	2097	92.9
2	浪费	5439	3668	9107	10761	84.6
3	消磨	940	239	1179	1578	74.7
4	拖延	898	477	1375	2269	60.6
5	打发	1726	391	2117	3925	53.9
6	约定	364	828	1192	2489	47.9
7	抓紧	4402	490	4892	16381	29.9
8	节省	846	1414	2260	8338	27.1
9	赢得	296	860	1156	8765	13.2
10	安排	702	1935	2637	21282	12.4

我们将 TORCH2009 和 BCC (多领域) 中高频、高搭配强度的“V_双 + 时间”进行统合, 去掉重复后, 共计 16 个。分别是: 抓紧时间、花费时间、争取时间、节省时间、拖延时间、约定时间、延长时间、缩短时间、压缩时间、压缩时间、耗费时间、赢得时间、浪费时间、打发时间、消磨时间、耽误时间、安排时间。我们将这些“V_双 + 时间”作为接下来考察的对象。

4.2. “V_双”在中检词汇等级大纲中的分布

目前, 日本中国语检定(以下简称“中检”)的等级共有 6 级, 从低到高分别是: 准 4 级、4 级、3 级、2 级、准 1 级和 1 级。

²在 TORCH2009 中, Freq ≥ 3 的“V_双 + 时间”仅统计出 14 个。

“V_双”的分布考察依据了《キクタン》³汉语系列丛书。该书为中检词汇备考书籍，在对以往考试内容进行分析的基础上，对照中检的级别，编制了《入门篇——中检准4级水平》、《初级篇——中检4级水平》、《初中级篇——中检3级水平》、《中级篇——中检2级水平》、《上级篇——中检准1级水平》等一系列丛书，是目前中检备考书籍中较权威的一种。为了更明确的说明词汇等级，我们将本次考察对象的16个动词在该书中的分类与中检考试等级大纲相结合，如下表4：

Table 4. The syllabus and competency descriptions for each level of the Chinese test and the corresponding level of “V-Double”

表4. 中检各等级大纲和能力描述表及“V_双”的对应等级

类别	等级	考试内容	V _双
入门篇 610 词	准4级	1. 基础单词约 500 个 2. 拼音(表音字母)的读法及拼写方法 3. 短句的基本语法 4. 简单的日常会话约 50~80 句	
初级篇 560 词	4级	1. 单词释义、汉字注拼音(表音字母)及拼写改汉字 2. 在常用的 500~1000 个单词范围内的汉语单句日译及日语汉译	安排
初中级篇 1059 词	3级	1. 单词释义、汉字注拼音(表音字母)及拼写改汉字 2. 在常用的 1000~2000 个单词范围内的汉语单句日译及日语汉译	抓紧、延长
中级篇 896 词	2级	1. 单词、复合词及惯用的日汉互译 2. 多音字、轻声问题及语句用法改错 3. 100~300 字左右文章的日汉互译	花费、节省、缩短、赢得、压缩
上级篇 1008 词	准1级	报纸、杂志、文学作品及应用文等难度较高的汉日互译	拖延、耽误、耗费
无对应书籍	1级	包括时事用语在内的高难度汉日互译：包括复合词、惯用语等在内的综合试题。	
书中未现词汇		浪费、争取、打发、消磨、约定	

通过上表4可知，语料库中与时间搭配的高频动词，约半数为中检中的初中级动词，4级、3级、2级动词共8个，占考察对象的50%，其中以2级居多，共5个。

4.3. 词语搭配猜词策略和结果

猜词策略，是指学习者在阅读过程中遇到生词时所采用的猜测词义的手段，比如运用文字学、词汇学知识、语境线索、背景知识等推测生词的意思。一般认为，阅读过程中的策略使用是一种问题解决的过程，作为阅读策略的猜词自然也是阅读过程中的问题解决，即是如何利用已有的知识来解决生词问题。猜词是阅读中跳跃生词障碍的方法之一，猜词所能用到的策略反映了学习者所掌握的语言知识与非语言知识水平(李慧, 2012) [17]。阅读活动涉及学习者所掌握的有目的语的文字、语言、文化知识水平以及认知能力与方式。作为成年的第二语言学习者，他们至少已有母语阅读的经验，已有的语言文字知识及其认知方式都会影响他们的第二语言阅读(马文颖, 2003) [18]。因此，不同母语背景的学习者在汉语阅读学习中既存在普遍性又有特殊性，尤其是汉字背景的有无，势必影响学习者的汉语阅读能力的发展。而学习者的汉语水平将影响他们汉语的汉字与汉语知识结构，在猜词能力上也会有所区别。

³日本关西大学出版社出版、关西大学中国语教材研究会编制。《キクタン入门篇——中检准4级水平》(2008)、《キクタン初级篇——中检4级水平》(2008)、《キクタン初中级篇——中检3级水平》(2009)、《キクタン中级篇——中检2级水平》(2010)、《キクタン上级篇——中检准1级水平》(2013)。

吴门吉(2008) [19]对欧美汉日学生的阅读猜词策略进行了问卷调查研究,对8项猜词策略进行了讨论,并对比分析了初级和中级日本学习者的猜词策略。下图1引用于吴(2008)一文:

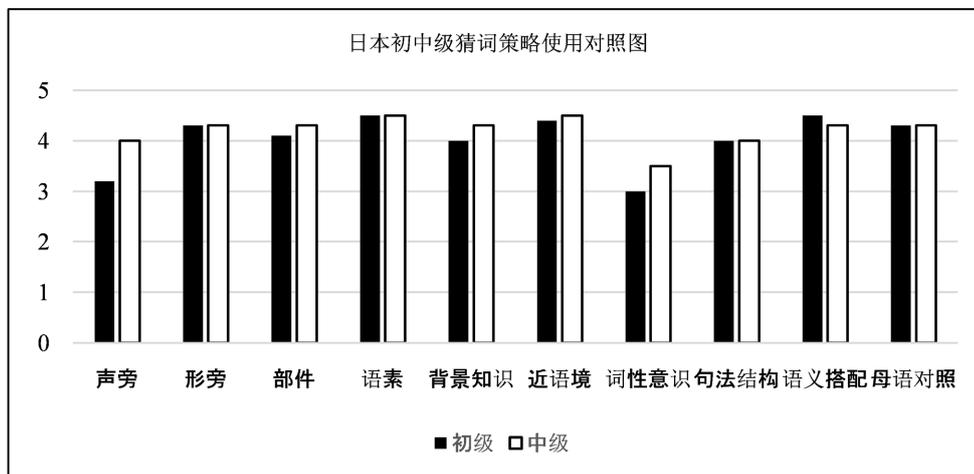


Figure 1. Comparison chart of word guessing strategies used by Japanese beginners and intermediate learners

图1. 日本初、中级学习者猜词策略使用对照图

其研究结果发现:日文中的汉字对日本学习者的汉语阅读猜词有极大帮助。表现在“语义搭配、语素猜词、形旁猜词、与母语对照、近语境”等一系列策略使用上的显著优势,同时研究也发现,从日本学习者的猜词策略从初级到中级并没有太大进步。同为汉字文化圈,日语中的汉字对学习产生母语迁移,日本汉语学习者既有二语学习者的共性,也有其独特性,因此笔者以上文的16个双音节词为对象进行了猜词策略的问卷调查,以便结合日本汉语初学者的母语迁移情况进行分析。

- 1) 调查目标:对“V_双 + 时间”的意思进行猜测。
- 2) 被试:15位日本汉语初级学习者(均为大阪大学中文系一年级在校大学生:21岁13名22岁2名)
- 3) 实验时间:2017年6月
- 4) 实施方法:问卷调查法

5) 实验内容:对本次考察对象的16个“v + 时间”的意思进行推测⁴。实验前,笔者首先简单说明实验内容,然后在实验开始以前,调查被试的汉语学习履历和已知单词,确认16个单词中是否含有被试的已知单词,如含有,即使是一个词,也将此被试从分析对象中排除。实验时,给出简体字和繁体字对照,如“节(節省)”、“延长(長)”等。如果推测结果与搭配意思相近或相符的话,那么该动宾搭配属于较易推测搭配,并说明推测理由;如果不符或根本无法推测的话,该动宾搭配属于较难推测搭配。投票对象为较易推测搭配,以投票数8票以上(含8票)为基准,将16个“V_双 + 时间”分成较易预测和较难预测两种,列表如下。()内为被试所投票数,如表5所示。

Table 5. Results for “v + 时间”

表5. “v + 时间”的猜词结果

较易预测	延长(15)、浪费(15)、压缩(15)、赢得(10)、缩短(15)耗费(15)、节省(13)、消磨(13)、争取(12)、拖延(11)、约定(9)
较难预测	抓紧(0)、花费(3)、耽误(0)、安排(0)、打发(0)

⁴对于中日同形同义词,在判断时向被试提示了简体和繁体两种书写。如:“节约(節約)”、“计算(計算)”。

当我们把学习对象限定为日本学习者时,有两大方面需要调研。第一,日本学习者的母语迁移。本次问卷调查中发现,“浪费”、“延长”、“压缩”、“缩短”、“耗费”因母语正迁移,极易推测,票数为满票 15;而“抓紧”、“打发”、“安排”在『日中・中日辞典』[20]中并无同形词,因此较难推测,票数为 0。第二,则是中日同形近义词。汉语词汇中的“中日同形近义词”被认为是日本汉语学习者表达时误用最多、习得困难的一类词。究其原因,首先,这类词多为多义词,与日语在词义和用法上,既存同又存异;其次,日语与汉语的词义、词性差异、以及汉语本身的难易度差异也都对日本汉语学习者习得“中日同形近义词”产生影响。根据 Laufer (1991)提出的“母语与目标语词义对应关系影响词汇习得难易度”的观点,“花费”中的“花”属于分散类(divergence),即 L2 中的多个词义与 L1 的一个单词的词义相对应,由于学习者要扩大他的母语范围(再扩展),因此这类词会对学习者的产出造成困难。“花”除了有与日语相同的名词“供观赏的植物”的意思,还有日语中没有的动词词性“在支付或支出上用掉钱”(日语为「消費する」「つかう」),因此有 3 位被试在“花费时间”的回答栏中写到“花に費やした時間(用在花上的时间)”;“打发”一词中的“打”和“发(発)”,排除简繁书写的区别,在日语中都有汉字,但是对于“打发”却很难进行联想和推测。再议“约定”,15 位被试中有 9 位正确推测出了“约定时间”的意思是“時間を約束する”,这 9 位在问卷调查的推测理由大体一致,例如其中一例为「約:約束する、定:決める。だから、約束して決める。(“约”是“约定”的意思,“定”时“决定”的意思,所以,约定应该是“约定后做出决定”),但是其余 6 位被试或是空白或是猜想错误(把“约定”翻译为“大体决定”)。因此,我们认为中日同形近义词的搭配与习得难易度并不成正比。

5. 对“V_双 + 时间”的分类

依据本次考查的结果,综合“V_双 + 时间”的猜词难易度和 V_双在中检中的等级分类,我们尝试性地将考察对象分为四个等级。这四个等级分别为,等级一:“V_双”的语义较易推测,为中日同形同义词。对学习来说,这类“V_双 + 时间”既是易接受,也易产出;等级二:“V_双”的语义较易推测,含有中日同形同义词或同形近义词,这类搭配较易输入,较难产出;等级三:“V_双”为初中级动词,语义较难推测;等级四:“V_双”为中高级动词,语义较难推测,这些动词既难输入,也难产出。具体分类如表 6 所示:

Table 6. Four levels of “V_双 + 时间” basic pairing (from easy to difficult)

表 6. “V_双 + 时间”的基本搭配四个等级(从易到难)

等级	分类标准	V _双
一级搭配	“V _双 ”较易推测,为中日同形同义词	浪费 ⁵ 、消磨、延长、缩短、压缩
二级搭配	“V _双 ”较易推测,含有中日同形同义词或同形近义词	耗费、约定、节省、花费、拖延、赢得、
三级搭配	“V _双 ”较难推测,为初中级词汇	安排、抓紧、
四级搭配	“V _双 ”较难推测,为中高级词汇	打发、耽误、争取、

6. 余论

本文考察了日本汉语学习者在学习“V_双 + 时间”这一搭配时的需要关注的问题点。除了判断“V_双 + 时间”的客观条件和语义透明度,还以日本汉语学习者的教育背景——中检为依据,进行了尝试性的

⁵“浪费”一词在『キクタン』系列丛书中并未出现。依据『日本中国語初級段階学習指導ガイドライン 学習語彙表』,将“浪费”归为初级词汇。

分类。通过上述分析我们认为,对于日本学习者来说,一级搭配只需注意简繁体书写,无需特意习得,如“延长、浪费、缩短(日语为「短縮」)”等。对于含有中日同行同义词或同形近义词的二级搭配,词义透明度较高,在输入时难易度较低,如“约定”中的“约(約)⁶”和“定”、“花费”中的“费(費)”、“节省”中的“节(節)”和“省”、“耗费”中的“耗”,在日语中均有意思相同或相近的汉字对应,同为汉字圈的学习者可通过猜词策略较易推测。但是二级搭配中的动词属于中高级词汇,掌握起来有一定难度,尤其是在产出时。教师应在搭配的产出、运用方面多下功夫,采取循序渐进的方式,在学习者积累了一定汉语词汇之后再作指导,精讲多练,帮助学习者正确产出。对于三级搭配,三级动词的语义透明度不高,但由于动词本身的难易度较低,因此学习者在掌握了动词语义后,可以自行搭配,进行产出。本次考察中,笔者认为四级搭配是教学中的难点。对日本学习者来说,四级搭配中的“V_双”本身属于高级词汇,“V_双+时间”在输入和产出时都存在一定困难。这些动词在初级教材和会话中的出现频次很低,但日常使用的频度却较高,需要教师适当扩充、日本学习者自己主动涉猎。我们虽把“打发、耽误”归入第四级,但通过表3可以得知,“耽误+时间”、“打发+时间”与“V_双”出现频次的百分比均很高,分别为92.9%、53.9%,因此在讲解“时间”一词时,教师可适时适度地对“耽误”、“打发”进行介绍和补充。如果教师能通过明确的指导使学生准确掌握的话,在一定程度上能有效地提高学习者的汉语产出流畅度和准确度,加强他们的交际能力。

此外,我们应该注意的是,仅以统计指标为基准归类出的词语搭配用于教育是片面的。在实际教学中,我们依然存在三个问题:Sinclair(1991)(同[14])提倡的习语原则和统计指标之间应该怎样统合;哪些搭配应该是日本汉语学习者需要学习的;这些需要学习的基本词语搭配与统计指标又应该怎样统合、如何在教材中反映,这些都是值得继续考察的问题。

最后,本研究还存在许多不足之处需要改进。首先本文是一次尝试性考察,仅以“V_双+时间”为切入点,研究范围有限;其次,针对日本汉语学习者,我们还有必要从学习者的汉语水平、词语搭配习得规律、产出策略、猜词策略等方面出发,进一步进行讨论。希望这些课题能在今后不断展开,使词语搭配教学更加有的放矢。

参考文献

- [1] Firth, J.R. (1957) Applications of General Linguistics. *Transactions of the Philological Society*, 56, 1-14. <https://doi.org/10.1111/j.1467-968X.1957.tb00568.x>
- [2] Halliday, M.A.K. and Hasan, R. (1976) Cohesion in English. Longman, London.
- [3] Craik, F.I.M. and Lockhart, R. (1972) Levels of Processing: A Framework for Memory. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- [4] Becker, J.D. (1975) The Phrasal Lexicon. *Proceedings of the 1975 Workshop on Theoretical Issues in Natural Language Processing*, Cambridge, 10-13 June 1975, 60-63. <https://doi.org/10.3115/980190.980212>
- [5] Nation, I.S.P. (1990) Teaching and Learning Vocabulary. Newbury House Publishers, New York.
- [6] Stubbs, Michael. (1996) Text and Corpus Analysis. Blackwell, Oxford.
- [7] Marton, W. (1977) Foreign Vocabulary Learning as Problem No. 1 of Language Teaching at the Advanced Level. *Interlanguage Studies Bulletin*, 2, 33-57.
- [8] Nesselhauf, N. (2003). The Use of Collocations by Advanced Learners of English and Some Implications for Teaching. *Applied Linguistics*, 2, 223-242. <https://doi.org/10.1093/applin/24.2.223>
- [9] 孙茂松, 黄昌珉, 方捷. 汉语搭配定量分析初探[J]. 中国语文, 1997(1): 29-38.
- [10] 卫乃兴. 词语搭配的界定与研究体系[M]. 上海: 上海交通大学出版社, 2002.
- [11] 卫乃兴. 其余语料库和语料库数据驱动的词语搭配研究[J]. 当代语言学, 2003, 4(2): 101-114.

⁶)中为日语汉字,后同。

-
- [12] 石川慎一郎. コロケーションの強度をどう測るか—ダイス係数、Tスコア、相互情報量を中心として—[C]// 言語処理学会第14回年次大会チュートリアル資料. 2008: 40-50.
- [13] 全昌勤, 刘辉, 何婷婷. 基于统计模型的词语搭配自动获取方法的分析与比较[J]. 计算机应用, 2005, 22(9): 55-57.
- [14] Sinclair, J. M. (1991) *Corpus, Concordance, Collocation*. Oxford University Press, Oxford.
- [15] Hoey, M. (2005) *Lexical Priming: A New Theory of Words and Language*. Routledge, Taylor and Francis Group, London.
- [16] BCC: 北京语言大学现代汉语语料库[EB/OL]. <http://bcc.blcu.edu.cn>
- [17] 李慧. “V_非+NP”语块的衍生途径及其制约因素[J]. 语言教学与研究, 2012(4): 53-60.
- [18] 马文颖. 词汇重复模式理论与英语阅读中的猜词策略[J]. 外语研究, 2003(6): 62-65.
- [19] 吴门吉. 对欧美韩日学生阅读猜词策略的问卷调查研究[J]. 云南师范大学学报: 对外汉语教学与研究版, 2008, 6(4): 17-23.
- [20] 中日·日中辞典[M]. 第2版. 東京: 日本小学馆出版, 2010.