

基于语料库的词语搭配研究 ——以“素直”“正直”“率直”为例

赵圣花¹, 刘莉², 邹善军¹

(1. 大连理工大学软件学院, 辽宁 大连 116023; 2. 沈阳建筑大学外国语学院, 辽宁 沈阳 110168)

摘要:基于日本 NLB 语料库考察近义词“素直”“正直”“率直”的词语搭配异同。主要根据搭配词的共现频数、MI 值、LD 系数的数据, 分析了“素直”“正直”“率直”与后续动词和名词的典型搭配特点。考察结果发现, 其后续相同的显著搭配动词较多, 而后续相同的显著搭配名词较少。

关键词:近义词; 词语搭配; 素直; 正直; 率直

中图分类号: H36

文献标志码: A

日语学习者在词汇学习的过程中, 如果只记词典中标注的释义和几个简单的例句, 而不了解该词的典型搭配与特征, 就可能在对日交流的过程中无法进行准确、地道的表达, 容易引起误会。

近年来, 随着大型语料库的开发与建设, 兴起了基于语料库的语言研究。通过检索软件提取语料库本族语者准确、地道的词汇用法, 利用数据驱动学习, 提高学习者的词汇搭配能力。研究词语的搭配特点有助于了解语言的本质, 总结词汇的搭配规律, 有利于输出准确、地道的外语。孙海燕^[1]认为, 如果学生没有掌握足够的词语搭配的方法和技巧, 他们产出的语言就会缺乏流利性和地道性, 影响正常的语言交际。外语学习者对词汇的理解往往来自词典里的中文解释, 并在此基础上进行遣词造句, 因此语言表达不够自然、地道。外语词汇的学习, 需要学习者掌握词汇的搭配特点, 了解目标语语言表达的特征与规律。

一、词语搭配的相关研究

目前, 国内外学者基于语料库的词语搭配研究已有相关数例。

赵恩英^[2]基于《每日新闻》语料库, 研究了“ついに”“とうとう”的不同搭配。该研究基于报刊的书面语料库, 规模小且没有采用数据驱动的方法, 分析的结果难免出现偏颇。

中溝朋子等^[3]分析了意义相近的名词“進歩”与“向上”的搭配特点。王华伟等^[4]介绍了基于语料库的意义相近的一组日语动词搭配的研究方法。他们都采用了数据驱动的词语搭配研究方法, 比较科学, 但需要掌握一定的统计学知识。

李倩^[5]利用语料库对“行為”与“行動”进行了相关搭配研究, 对搭配词按照语义韵进行了归类。

基于日本 NLB 语料库^[6]的近义词研究有郜枫^[7]对于“願う”“祈る”的搭配研究, 赵

圣花等^[8]对于“いきなり”“突然”“急に”的搭配研究。利用 NLB 语料库呈现的搭配信息,研究关键词的搭配规律与特点,弥补了先行研究的不足。

日本 NLB 语料库,解决了统计学的难题,为日语词语搭配研究提供了便利。日语 N2 真题中与词汇搭配相关的题约有 12 道左右。学生对这类题型感到困惑,因此,有必要将高频词搭配的研究成果运用到词汇教学中,以期提高课堂教学效果,促进学生的词汇习得。

“素直”“正直”“率直”这 3 个词都有“坦率”的意思,都属于生活中常用的词汇。但由于语义重叠导致学习者很难区别使用,造成词语搭配不当。

本研究基于日本 NLB 语料库呈现的词语搭配的共现频数、MI 值、LD 系数,辨析“素直”“正直”“率直”的异同。

二、NLB 语料库的介绍及研究方法

NINJAL - LWP for BCCWJ (简称 NLB),是日本国立国语研究所和 Lago 语言研究所为检索《现代日语书面语均衡语料库》共同开发的在线检索系统。该语料库共收录约 1 亿词。

NLB 语料库最大功能之一是根据研究者的需要在系统中输入节点词,就能自动呈现节点词与其他关联词汇之间的共现搭配信息,以及共现频数、MI 值、LD 系数。

本研究将 MI 值 ≥ 3 、LD 系数 ≥ 5 的词判断为显著搭配词。有关 MI 值、LD 系数的具体概念及公式可参考赵圣花等^[9]的相关研究。

三、“素直”“正直”“率直”在词典中的解释

参考日本《大辞泉》^[10]的解释,这 3 个词的常见用法和解释分别如下:

1. 素直

【名・形容動詞】ありのままで、飾り気のないさま。素朴。性質、態度などが穏やかでひねくれていないさま。従順。

2. 正直

【名・形容動詞・副詞】うそやいつわりのないさま。見せかけやごまかしではないさま。率直なさま。本当のところ。

3. 率直

【名・形容動詞】ありのままで、隠すところがないこと。またそのさま。

由《大辞泉》的解释可知,“素直”“率直”都以“ありのままで”来解释,而“正直”以“率直なさま”来解释,由于这 3 个词存在语义重叠,导致其很难区别,因此需要研究这 3 个词汇的搭配特点。

四、语料库中“素直”“正直”“率直”与后续动词的搭配信息

利用 NLB 语料库检索的结果如表 1 ~ 表 3 所示。分别抽出后续动词的频数排在前 25 位的动词,通过 MI 值和 LD 系数探讨词语搭配特点。

1. 后续相同的搭配动词

由表 1 ~ 表 3 可知,“素直”“正直”“率直”之后都与“言う”“認める”“する”“伝える”“答える”“書く”“話す”共现。其中,与“率直”搭配的动词里,“答える”与“話す”的 LD 系数分别为 4.87,4.91,这两个数值可以四舍五入算作 5。因此,按照 MI 值 ≥ 3 、LD 系数 ≥ 5 的标准判断显著性,与“素直”“正直”“率直”搭配显著性高的后续动词依次是“する”“伝える”“答える”“話す”(见表 4)。

2. 后续不同的搭配动词

从表 1“素直に + 動詞”的搭配信息 MI 值和 LD 系数可知,“素直”与“なれる”“認める”“受け入れる”“喜ぶ”“うなずく”“聞く”“する”“従う”“喜べる”“伝える”“答える”“言える”“受け止める”“受け取る”“表現する”“謝る”“信じる”“応じる”“話す”“傾ける”搭配强度较高。

例(1):真子は素直にうなずいている。

汉译:真子老实地点头。

“素直”在这里表示没有反驳的意思、

表1 “素直に+動詞”(425种)共现频数排在前25位的搭配词的搭配信息

序号	动词	频数	MI	LD	序号	动词	频数	MI	LD
1	なる	67	2.95	2.17	14	答える	29	7.13	6.13
2	言う	67	3.71	2.93	15	言える	28	6.05	5.17
3	なれる	62	9.09	7.94	16	受け止める	24	9.20	7.54
4	認める	61	7.52	6.60	17	受け取る	21	8.03	6.74
5	受け入れる	59	9.59	8.28	18	表現する	20	8.09	6.76
6	喜ぶ	57	9.06	7.88	19	謝る	20	9.86	7.70
7	うなずく	56	9.16	7.95	20	信じる	17	6.65	5.61
8	聞く	47	5.57	4.76	21	受ける	15	4.49	3.65
9	する	42	21.16	9.46	22	応じる	15	7.84	6.45
10	従う	42	8.64	7.46	23	書く	15	4.50	3.66
11	できる	38	3.45	2.67	24	話す	15	6.22	5.22
12	喜べる	33	14.13	9.04	25	傾ける	13	8.80	6.89
13	伝える	31	7.36	6.34					

表2 “正直に+動詞”(177种)共现频数排在前25位的搭配词的搭配信息

序号	动词	频数	MI	LD	序号	动词	频数	MI	LD
1	言う	306	6.75	5.13	14	告げる	11	8.35	6.33
2	話す	80	9.48	7.73	15	認める	9	5.60	3.91
3	答える	74	9.32	7.58	16	思う	8	2.01	0.38
4	書く	28	6.25	4.59	17	なさる	7	8.47	6.22
5	生きる	28	7.33	5.63	18	白状する	7	11.61	7.45
6	なる	23	2.25	0.63	19	言える	7	4.90	3.22
7	伝える	22	7.71	5.95	20	述べる	7	5.89	4.15
8	告白する	20	10.88	8.25	21	お願いする	5	5.37	3.64
9	申し上げる	18	7.53	5.77	22	しゃべる	5	7.00	5.03
10	する	17	20.70	9.00	23	なれる	5	6.31	4.47
11	打ち明ける	15	10.87	8.02	24	反映する	4	6.91	4.89
12	語る	14	7.26	5.48	25	教える	4	3.69	2.03
13	申す	12	7.80	5.92					

表3 “率直に+動詞”(158种)共现频数排在前25位的搭配词的搭配信息

序号	动词	频数	MI	LD	序号	动词	频数	MI	LD
1	言う	189	6.63	4.44	14	表現する	6	7.78	5.35
2	申し上げる	48	9.53	7.23	15	話し合う	6	8.91	6.24
3	認める	28	7.82	5.57	16	受け止める	5	8.36	5.77
4	思う	23	4.11	1.91	17	打ち明ける	5	9.86	6.74
5	する	22	21.65	9.95	18	お伝えする	4	10.86	6.96
6	述べる	22	8.12	5.84	19	尋ねる	4	7.34	4.88
7	聞く	14	5.25	3.04	20	感じる	4	4.51	2.28
8	伝える	12	7.41	5.12	21	指摘する	4	6.95	4.56
9	答える	11	7.15	4.87	22	書く	4	4.02	1.80
10	話す	11	7.20	4.91	23	おっしゃる	3	6.33	3.96
11	申す	10	8.12	5.74	24	交換する	3	7.83	5.18
12	語る	10	7.35	5.05	25	吐露する	3	12.06	6.88
13	致す	7	6.62	4.33					

表4 “素直”“正直”“率直”后续相同的显著搭配动词

后续相同的搭配动词	素直に+动词	正直に+动词	率直に+动词
する	42	17	22
伝える	31	22	12
答える	29	74	11
話す	15	80	11

注:数字为频数。

服从。

例(2):本当は大好きなのに、嫌いだよとか言ってしまう。どうしたら素直になれますか。

汉译:明明很喜欢,但是我说了讨厌。怎样才能变得坦率呢?

“素直”在这里表示坦率、真实、实事求是。

例(3):素直に自分のミスを認める。

汉译:我老实承认自己的错误。

“素直”在这里表示坦率、老实。

例(4):先生の勧めを素直に受け入れてたばこを止めた。

汉译:我老实听从大夫的劝告,戒掉烟了。

“素直”在这里表示听从。

例(5):彼は私の質問に素直に答えてくれた。

汉译:他老实地回答了我的质问。

“素直”在这里表示老实。

例(6):親の言うことは素直に従う。

汉译:父母说的话我老实服从。

“素直”在这里表示老实、顺从。

例(7):素直に謝った。

汉译:我坦率地道歉。

“素直”在这里表示温顺、老实。

例(8):幼い時、この小説を素直に信じてしまって失敗したことが何回もある。

汉译:幼小的时候,我单纯地相信这个小说,因此失败过好几次。

“素直”在这里表示单纯、天真。

由例(1)~例(8)可以看出,“素直に”主要表达坦率交代、老实服从,或表达天真、单纯的行为。

由表2“正直に+動詞”的搭配信息 MI 值和 LD 系数可知,“正直”与“言う”“話す”“答える”“書く”“生きる”“伝える”“告白する”“申し上げる”“する”“打ち明ける”“語る”“申す”“告げる”“なさる”“白状する”“しゃべる”“なれる”“反映する”搭配强度较高。

例(9):秘密を正直に話しますか?

汉译:会如实说出秘密吗?

例(10):正直に言うと怖いです。

汉译:(不骗你)如实说,我真害怕。

例(11):お前はこれから訊くことに正直に答えるんだ。

汉译:现在开始你要如实回答问题。

例(12):正直に悩みを打ち明ける。

汉译:我如实说出烦恼。

例(9)~例(12)中,“正直”都表示不撒谎、如实说出。

例(13):正直に告白します。

汉译:我会如实表白的。

“正直”在这里表示诚实、坦率。

例(14):正直に生きてきた。

汉译:我诚实地活着。

“正直”在这里表示诚实。

例(15):正直に言わせていただきたいんですが……。

汉译:请允许我坦率地说……。

“正直”在这里表示不骗人、实事求是。

由例(9)~例(15)可以看出,“正直に”主要表达诚实、不撒谎、实事求是、坦率。

由表3“率直に+動詞”的搭配信息 MI 值和 LD 系数可知,“率直”与“申し上げる”“認める”“する”“述べる”“伝える”“答える”“話す”“申す”“語る”“表現する”“話し合う”“受け止める”“打ち明ける”“お伝えする”“尋ねる”“指摘する”“交換する”“吐露する”搭配强度较高。

例(16):率直に申し上げてちょっと、お答えいたしかねます。

汉译:坦率地说,有点难以回答。

例(17):どこまで率直に語ってよいのだろう。

汉译:直接说到什么程度才好呢。

例(18):依頼者は弁護士と率直に話し合うよう促されるのである。

汉译:要求当事人和律师不隐瞒事实、如实地交谈。

例(19):率直に打ち明けて差しつかえない。

汉译:有什么话不妨直说。
例(16)~(19)中,“率直”都表示不隐瞒事实、直言不讳。

例(20):彼は私の質問に率直に答えてくれた。
汉译:他爽快地回答了我提出的问题。

例(21):敗北を率直に認めた。
汉译:我坦率地承认了失败。
例(22):批判を率直に受け止める。
汉译:我坦诚地接受批判。

例(20)~例(22)中,“率直”都表示态度

上非常坦率、爽快。
由例(16)~例(22)可以看出,“率直”主要表达态度上非常坦率、爽快、不隐瞒事实、直言不讳。

五、语料库中“素直”“正直”“率直”与后续名词的搭配信息

利用 NLB 语料库检索的结果如表 5~表 7 所示。分别抽出后续名词的频数排前 10 位的名词,通过 MI 值和 LD 系数探讨词语搭配特点。

表 5 “素直な + 名詞”(136 种)共现频数排在前 10 位的搭配词的搭配信息									
序号	名词	频数	MI	LD	序号	名词	频数	MI	LD
1	気持ち	59	9.87	6.51	6	子	9	6.60	3.26
2	心	21	7.46	4.11	7	感情	7	8.52	5.11
3	人	16	4.10	0.76	8	自分	7	4.44	1.11
4	感想	13	10.79	7.27	9	人間	5	5.37	2.03
5	性格	12	9.14	5.74	10	言葉	5	5.46	2.11

表 6 “正直な + 名詞”(90 种)共现频数排在前 10 位的搭配词的搭配信息									
序号	名词	频数	MI	LD	序号	名词	频数	MI	LD
1	ところ	262	9.27	6.84	6	人間	11	6.60	3.26
2	気持ち	67	9.14	6.68	7	男	8	5.17	2.73
3	話	50	7.64	5.20	8	意見	7	6.50	4.01
4	感想	34	11.26	8.51	9	心	6	4.74	2.30
5	人	21	3.58	1.16	10	自分	4	2.72	0.30

表 7 “率直な + 名詞”(97 种)共现频数排在前 10 位的搭配词的搭配信息									
序号	名词	频数	MI	LD	序号	名词	频数	MI	LD
1	意見	45	10.36	6.73	6	感じ	8	7.33	3.71
2	感想	37	12.56	8.81	7	発言	6	9.35	5.65
3	ところ	21	6.80	3.20	8	願い	6	10.35	6.55
4	気持ち	12	7.84	4.22	9	疑問	5	8.67	5.00
5	態度	9	9.16	5.50	10	見解	5	9.41	5.68

1. 后续相同的搭配名词

由表 5~表 7 可以看出,“素直”“正直”“率直”都与名词“感想”“気持ち”共现。其中,“感想”是这 3 个词的显著搭配名词。

2. 后续不同的搭配名词

由表 5“素直な + 名詞”的搭配信息 MI 值和 LD 系数可知,“素直”与“気持ち”“感想”“性格”“感情”搭配强度较高。
例(23):今日あなたに素直な気持ちを告白します。
汉译:今天我向你坦白真实感情。
“素直”在这里表达真实的情感。

例(24):息子の素直な感想がおかしいやら可愛いやらでたまらない。
汉译:儿子坦率的感想非常好笑可爱。
“素直”在这里表达坦率、单纯。
例(25):その素直な性格は、私の想像をはるかに超えていた。彼女には、無心に遊ぶ子供の一面があった。
汉译:她那天真的性格远远超出了我的想象。她有天真玩耍的孩子的一面。
“素直”在这里表达天真无邪。
例(26):子供らしく素直な感情を出して甘えることができできなかった。

汉译:我没能像孩子一样表现出天真的感情撒娇。

“素直”在这里表达孩子般的天真、纯朴的感情。

由例(23)~例(26),可以看出“素直な”主要表达真实的、天真的、纯朴的、单纯的。

从表6“正直な+名詞”的搭配信息 MI 值和 LD 系数可知,“正直”与“ところ”“気持ち”“話”“感想”搭配强度较高。

例(27):正直なところ、参っています。

汉译:老实说,我很为难。

“正直”在这里表示不撒谎、实话实说。

例(28):これが私の正直な気持ちだ。

汉译:这是我真实的想法。

例(29):正直な話、なにを言っているんだか、さっぱりわからない。

汉译:说实话,我完全不知道你在说什么。

由例(27)~例(29),可以看出“正直な”主要表达诚实、真实、实话实说,实事求是。

从表7“率直な+名詞”的搭配信息 MI 值和 LD 系数可知,“率直”与“意見”“感想”“態度”“発言”“願い”“疑問”“見解”搭配强度较高。

例(30):率直な意見で結構です。よろしく願います。

汉译:坦率的意见就可以。拜托您如实告诉我。

例(31):彼は少し内気ですが、時々、びっくりするほど率直な発言をすることもあります。

汉译:他虽然有点腼腆,但有时也会直接地说出令人吃惊的话。

例(30)和例(31)中,“率直”都表示不客气、坦率、直接。

例(32):文部大臣に率直な感想をお伺いしたいと思います。

汉译:我想请教一下教育部大臣您的真实感想。

“率直”在这里表示不隐瞒事实、真实。

由例(30)~例(32),可以看出“率直な”主要表达坦率、不隐瞒事实、真实、直言不讳。

六、研究结果

通过考察“素直”“正直”“率直”的搭配信息,得出了以下结论。

1. 相同搭配特点

(1)后续相同的搭配动词

“素直”“正直”“率直”的显著搭配动词有4个,分别是“する”“伝える”“答える”“話す”。

(2)后续相同的搭配名词

“素直”“正直”“率直”之后都与名词“感想”“気持ち”共现。根据搭配信息可知,“感想”是后续显著搭配名词。

2. 不同搭配特点

(1)“素直”后续显著搭配动词与名词

动词有:“なれる”“認める”“受け入れる”“喜ぶ”“うなずく”“聞く”“する”“従う”“喜べる”“伝える”“答える”“言える”“受け止める”“受け取る”“表現する”“謝る”“信じる”“応じる”“話す”“傾ける”。

名词有:“気持ち”“感想”“性格”“感情”。

“素直”主要表达坦率、老实、服从,或天真、纯朴、简单不复杂、单纯。

(2)“正直”后续显著搭配动词与名词

动词有:“言う”“話す”“答える”“書く”“生きる”“伝える”“告白する”“申し上げる”“する”“打ち明ける”“語る”“申す”“告げる”“なさる”“白状する”“しゃべる”“なれる”“反映する”。

名词有:“ところ”“気持ち”“話”“感想”。

“正直”主要表达诚实、正直、不撒谎、坦率、真实、实事求是。

(3)“率直”后续显著搭配动词与名词

动词有:“申し上げる”“認める”“する”“述べる”“伝える”“答える”“話す”“申す”“語る”“表現する”“話し合う”“受け止める”“打ち明ける”“お伝えする”“尋ねる”“指摘する”“交換する”“吐露する”。

名词有:“意見”“感想”“態度”“発言”

“願い”“疑問”“見解”。

“率直”主要表达坦率、不隐瞒事实、真实、直言不讳。

七、结 语

通过对近义词“素直”“正直”“率直”的搭配研究,介绍了关于近义词辨析的一种便捷可行的方法。以期更多的日语学习者,通过利用现代语料库检索工具掌握词汇的搭配规律和特点,使自己的外语输出更准确、地道,更接近本族语者的水平。

参考文献:

[1] 孙海燕. 语境论视角下学习者英语搭配的认知特点[J]. 外国语,2013(6):143-146.
[2] 趙恩英. 毎日新聞データからみた複文における「ついに」「とうとう」の差について:共起表現を中心に[J]. 日本語研究,2010(30):13-29.
[3] 中溝朋子,坂井美恵子. 漢語名詞「進歩」と「向上」のコロケーションの異同について[J]. 大

学教育,2011(8):88-95.
[4] 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.
[5] 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.
[6] 日本国立国語研究所, Lago 言語研究所. NINJAL-LWP for BCCWJ[EB/OL]. [2021-04-01]. <http://nlb.ninjal.ac.jp/>.
[7] 郜枫. 词汇搭配视阈下「願う」和「祈る」的语义特征研究[J]. 日语学习与研究,2016(2):23-30.
[8] 赵圣花,刘艳伟. 基于语料库的词语搭配研究:以副词「いきなり」「突然」「急に」为例[J]. 外语教育研究,2020(1):22-30.
[9] 赵圣花,邹善军,李光赫. 基于语料库研究表示“盖”动作的日语词语搭配[J]. 语文学刊,2014(1):3-5.
[10] 松村明. デジタル大辞泉[EB/OL]. [2021-04-01]. <http://www.daijisen.jp/>.

Corpus – Based Collocation Study on Adverbs:
Taking “Sunao”, “Syoujiki” and “Sottyoku” as an Example

ZHAO Shenghua¹, LIU Li², ZOU Shanjun¹

(1. School of Software, Dalian University of Technology, Dalian 116023, China; 2. School of Foreign Languages, Shenyang Jianzhu University, Shenyang 110168, China)

Abstract: Based on the Japanese corpus of NLB, this paper examines the similarities and differences of collocation of three Japanese near-synonyms: “sunao”, “syoudjiki” and “sottyoku”. According to the data from three aspects: co-occurrence frequency, MI value and LD coefficient, the typical collocation characteristics of collocation words with their follow-up verbs and nouns are analyzed. The results indicate that there are more similarly significant collocation verbs, but fewer significant collocation nouns.

Key words: near-synonym; collocation; “sunao”; “syoudjiki”; “sottyoku”

(责任编辑:高旭 英文审校:林昊)