

日本学生汉语书面语中的口语化情况研究

汲传波

(北京大学 对外汉语教育学院 北京 100871)

摘要: 本研究通过比较 37 个口语格式在日本学生汉语中介语语料和汉语母语者语料中使用频率的差异,发现日本学生汉语书面语中有部分口语化的情况,这种口语化并没有随着其汉语水平的提高而减弱。另外,本研究通过考察 37 个口语格式在现代汉语词典、HSK 大纲中的收录情况及其在汉语对话语料中的使用情况,对词典、HSK 大纲收录口语格式提出了建议。

关键词: 日本学生; 书面语; 口语化; 口语格式

中图分类号: H195 **文献标识码:** A **文章编号:** 1000 - 579(2017)02 - 0080 - 06

Features of Oral Style in Chinese Compositions of Japanese Students

Ji Chuanbo

(School of Chinese as a Second Language, Peking University, Beijing 100871, China)

Abstract: Features of oral style in Chinese compositions of Japanese students are found by comparing differences in frequency of 37 oral formats in Japanese students inter-language corpus and Chinese native speakers corpus. And also the oral style features of Japanese students are not reduced with the improvement of their Chinese level. On base of examining 37 formats in *Modern Chinese Dictionary*, HSK Syllabus and Chinese discourse corpus, some suggestions about oral formats included in *Modern Chinese Dictionary*, HSK Syllabus are given.

Key words: Japanese students; writing Chinese; oral style features; oral formats

第二语言习得中的语体问题,尤其是书面语中的口语化倾向问题,在国内英语教学界一直受到关注,比如文秋芳等、文秋芳、潘璠利用中介语语料库与英语本族语者语料库进行对比研究,发现了中国英语学习者作文中存在明显的口语化倾向。^{[1][2][3]}近些年,汉语作为第二语言教学界也开始关注留学生语体习得问题。比如,有的利用中介语语料库与汉语母语者语料的对比,研究留学生书面语中的口语化倾向;^[4]有的运用语体量化分析法对留学生教材的语体分布情况进行考察;^[5]有的运用测试和调查了解留学生的语体语法能力;^{[6][7][8]}有的研究高水平留学生学术论文语体使用情况。^[9]这些研究视角各异,研究材料不同,但是结论几乎是一致的,即留学生存在书面语体能力不足的问题。

虽然目前的研究已经发现了汉语第二语言学习者存在语体能力不足的问题,但是大都在研究时不分国别。针对单一母语背景学习者语体习得情况的研究尚比较缺乏,仅见针对泰国学生语体习得情况的研究。^[10]一般认为日语母语的汉语学习者比其他母语背景的学习者汉语书面语能力强,口语表达能

收稿日期:2016 - 09 - 07

基金项目:教育部规划基金项目“面向第二语言教学的汉语语体语法研究”(编号:16YJAZH021);获 2015 年度北京大学桐山教育基金研究资助

作者简介:汲传波(1977 -),男,山东莒南人,语言学博士,北京大学对外汉语教育学院副教授,硕士生导师。研究方向为汉语作为第二语言语法研究,教师信念研究。

力弱。那么是否可以推论:日语为母语的学习者在书面语中不存在口语化情况或者口语化不明显。本研究将在原有研究^[4]基础上重点探讨两个问题:(1)日本学生汉语书面语中是否存在口语化情况;(2)如果存在口语化,这种倾向是否会随着其汉语水平的提高而减弱。

我们将通过口语格式的使用情况来判定日本学生书面语中是否存在口语化情况。首先通过对比日本学生书面语语料(HSK 动态作文语料)和汉语母语者语料(国家语委现代汉语平衡语料)中 37 个口语格式使用的标准化频率均值(每百万字),判断日本学生是否在书面语中多用口语格式,多用哪些口语格式。然后通过考察作文分数、HSK 考试成绩与口语格式使用数量的关系,判断日本学生书面语中的口语化是否会随着其汉语水平的提高而减弱。需要说明的是,本文与原有研究^[4]稍有不同,对中介语语料不再细分为记叙文体和议论文体。主要原因是考虑到现代汉语平衡语料库中记叙文、议论文均有,不再细分语体也可以进行对比,研究结论不会受到影响。

另外,我们将通过考察 37 个口语格式在现代汉语词典、HSK 大纲中的收录情况分析日本学生书面语中的口语化是否与词典、大纲相关;利用中国传媒大学有声媒体语料库,考察 37 个口语格式在汉语对话语料中的使用情况,分析词典、大纲在口语格式收录方面的得失。

一、日本学生汉语书面语中口语格式使用情况统计分析

(一) 日本学生 37 个口语格式总体使用情况

频率均值的计算方法与汲传波、刘芳芳^[4]相同,首先统计出 37 个口语格式在所有日本学生中介语语料中出现次数(309),然后推算出日本学生的作文总字数。HSK 动态作文语料库总计有 11569 篇文章,总字数为 424 万字,其中日本学生共有 3211 篇文章,其总字数推算方法为: $3211 \div 11569 \times 424$ 。最后用次数除以字数,计算出总频率均值。结果见表 1。

表 1 37 个格式在不同语料库中的总频率均值比较

	日本学生	各国留学生	汉语母语者
37 个口语格式出现次数	309	1094	2571
语料总字数(百万字)	1.18	4.24	20
频率均值(每百万字)	262.57	258.02	128.55

与汲传波、刘芳芳^[4]统计结果进行比较,发现日本学生使用 37 个口语格式的频率均值稍高于各国留学生总体频率均值,但是远远高于汉语母语者的频率均值($262.57 > 128.55$)。这一结果说明我们在引言中所做出的预设不成立。虽然日本学生有汉字的优势,但是其书面语中仍然存在口语化情况,并且还略高于各国留学生的平均值。日本留学生尚未意识到 37 个汉语口语格式常用于口语语体中,因此在书面语中过多地使用。虽然我们只是从单一维度考察口语化问题,但这一结果可以说明日本学生的汉语书面语体能力尚存在不足。

(二) 每个口语格式的具体使用情况

下面将根据每个口语格式在日本学生中介语语料中的分布情况进行排序,并且将研究结果与汲传波、刘芳芳^[4]中各国留学生的分布情况进行比较。

表 2 每个口语格式在中介语中的频率均值

口语格式	日本学生	各国学生	均值差	口语格式	日本学生	各国学生	均值差
不好意思	47.59	50.94	-3.35	没事儿	0.85	1.65	-0.8
无所谓	41.64	20.99	20.65	巴不得	0.85	1.42	-0.57
没问题	27.19	20.75	6.44	(就)得了	0.85	1.18	-0.33
不得了	25.49	19.10	6.39	你说呢	0.85	1.18	-0.33
就这样	24.64	34.91	-10.27	说真的	0.85	1.18	-0.33
不用说	22.94	19.10	3.84	没什么	0.85	0.71	0.14
不怎么样	16.99	15.57	1.42	有的是	0.85	0.71	0.14
对了	11.05	14.39	-3.34	说了算	0.85	0.47	0.38
你看	6.80	8.02	-1.22	说白了	0.85	0.24	0.61

口语格式	日本学生	各国学生	均值差	口语格式	日本学生	各国学生	均值差
话说回来	5.95	6.37	-0.42	怪不得	0.00	1.42	-1.42
罢了	4.25	6.37	-2.12	怎么搞的	0.00	0.71	-0.71
算了	3.40	4.25	-0.85	话是这么说	0.00	0.47	-0.47
好了	2.55	8.49	-5.94	说不好	0.00	0.47	-0.47
(不)要说	2.55	4.95	-2.4	要我说	0.00	0.47	-0.47
就是了	2.55	3.77	-1.22	一句话	0.00	0.47	-0.47
不像话	2.55	2.12	0.43	本来嘛	0.00	0.24	-0.24
不要紧	2.55	1.89	0.66	别看	0.00	0.24	-0.24
没劲	2.55	0.94	1.61	够朋友	0.00	0.24	-0.24
再说吧	1.70	1.65	0.05				

表2结果显示,日本学生使用最多的十个口语格式依次为:不好意思、无所谓、没问题、不得了、就这样、不用说、不怎么(样)、对了、你看、话说回来。37个口语格式中有9个格式在日本学生语料中一次也未出现:怪不得、怎么搞的、话是这么说、说不好、要我说、一句话、本来嘛、别看、够朋友。

通过与各国留学生使用情况的比较(均值差)发现,日本学生多用的格式(均值差为正值)有13个,如“无所谓、没问题、不得了、不用说、没劲”等。少用的格式(均值差为负值)有24个,如“就这样、好了、不好意思、对了、(不)要说”等。这一发现比较有意思,说明虽然总体上留学生的书面语中存在多用汉语口语格式的情况,但是不同母语背景的留学生在具体格式的使用上存在差异。这一问题值得今后深入研究。

(三) 口语格式在中介语与平衡语料库中的频率均值比较

为进一步比较日本学生在哪些具体口语格式的使用上与汉语母语者有差异,我们需要计算每个口语格式在中介语语料库和现代汉语平衡语料库中的频率均值差。结果见表3。

表3 口语格式在日本学生中介语与平衡语料库中的频率均值比较

口语格式	日本学生	汉语母语者	均值差	口语格式	日本学生	汉语母语者	均值差
不好意思	47.59	9.4	38.19	本来嘛	0	0.15	-0.15
无所谓	41.64	5.25	36.39	说不好	0	0.2	-0.2
没问题	27.19	2.25	24.94	够朋友	0	0.4	-0.4
不得了	25.49	3.55	21.94	你说呢	0.85	1.45	-0.6
不用说	22.94	4.9	18.04	巴不得	0.85	1.5	-0.65
不怎么(样)	16.99	2.2	14.79	说了算	0.85	1.7	-0.85
对了	11.05	0.6	10.45	说真的	0.85	1.75	-0.9
话说回来	5.95	1.65	4.3	没什么	0.85	2.15	-1.3
不像话	2.55	0.45	2.1	就是了	2.55	4	-1.45
就这样	24.64	22.55	2.09	怎么搞的	0	2.65	-2.65
再说吧	1.7	0	1.7	一句话	0	2.95	-2.95
没劲	2.55	0.85	1.7	算了	3.4	6.6	-3.2
没事儿	0.85	0.1	0.75	(不)要说	2.55	6.35	-3.8
说白了	0.85	0.35	0.5	怪不得	0	5.25	-5.25
好了	2.55	2.35	0.2	别看	0	6.05	-6.05
(就)得了	0.85	0.85	0	不要紧	2.55	9.1	-6.55
有的是	0.85	0.85	0	你看	6.8	17.2	-10.4
话是这么说	0	0.1	-0.1	罢了	4.25	17.95	-13.7
要我说	0	0.1	-0.1				

表3呈现了37个口语格式在日本学生中介语与平衡语料库中的频率均值差,并且按照均值差从大到小的顺序进行排列。结果显示,书面语中日本学生比汉语母语者多用的口语格式有15个:不好意思、无所谓、没问题、不得了、不用说、不怎么(样)、对了、话说回来、不像话、就这样、再说吧、没劲、没事儿、说白了、好了。另外表3也显示,书面语中日本学生比汉语母语者少用的口语格式有20个:罢了、你看、不要紧、别看、怪不得、(不)要说、算了、一句话、怎么搞的、就是了、没什么、说真的、说了算、巴不得、你说呢、够朋友、说不好、本来嘛、要我说、话是这么说。

日本学生比汉语母语者多用部分口语格式,可能表明日本学生对这些口语格式比较熟悉,但是并不了解其口语语体特征,因此在书面语中使用较多。日本学生比汉语母语者少用部分口语格式,可能表明日本学生对这些口语格式不是非常熟悉,而不是因为他们知道这些格式的口语语体特征而少用。这是基于日本学生对37个口语格式使用频率远远高于汉语母语者所做出的推测,因为总体而言日本学生的汉语书面语体能力比较弱。

(四) 日本学生汉语水平与口语格式使用的关系

我们预设随着汉语水平的提高,日本学生书面语中的口语格式使用频率应该越低,但从表4来看,日本学生口语格式的使用并没有随着其作文成绩、证书级别的提高而减少。比如作文分数在70~80之间的口语格式使用频率反而比80分以上的低,获得证书B的口语格式使用频率反而比A的低。该结果说明日本学生汉语水平与其书面语中口语格式的使用没有直接的关系,也说明语体能力的习得难度非常大。另外有趣的现象是作文最高分组与证书级别最高组的频率均值都是本组中最高的,分别为330.24、430.82。这说明汉语水平高的日本学习者掌握的口语格式越多,越愿意在写作中使用。但是日本学生并不了解这些口语格式的语体属性。“语言习得大致有个时限,可以规定出个习得标准;而语体意识习得则不然,既无时限,也很难找出个共同的标准。”^[11]这说明语体习得与语言具体要素的习得有重要差异,语体习得更为复杂。

表4 汉语水平与口语格式使用的关系

作文分数	频率均值	证书级别	频率均值
<60	209.89	无	244.09
≥60, <70	273.93	C	282.19
≥70, <80	245.41	B	262.36
≥80	330.24	A	430.82

本研究结果与文秋芳不同,她选择14个典型语体特征分析英语专业学生一年级与四年级英语口语语体特征的变化,同时考察其变化与本族语者口笔语语体差异的方向是否一致。^[2]其研究结果表明经过4年学习,差异程度朝着本族语者口笔语差异的方向变化,但与本族语者仍旧有很大区别。本研究与文秋芳结果不同的原因可能在于,由于目前汉语学界对于区分口语、书面语语体的核心语体特征研究成果非常少,我们仅能从口语格式使用这一维度考察,因此结果并不能清晰地显示汉语水平与口语格式使用之间的相关性。

二、词典、大纲及对话语料中收录、使用口语格式的情况分析

上文研究发现,日本学生的汉语书面语中存在口语化。下面,我们将通过考察37个口语格式在现代汉语词典、HSK大纲中的收录情况,分析日本学生书面语中的口语化是否与词典、大纲相关;另外通过考察37个口语格式在汉语对话语料中的使用情况,分析词典、大纲在口语格式收录方面的得失。

(一) 词典、大纲收录情况分析

在现代汉语词典(第6版)中检索37个口语格式的收录情况,收录记为Y,没有收录则记为N。在HSK新旧词汇大纲中检索口语格式的级别情况,不在大纲中记为N,收入大纲则直接记为大纲中标示的级别,结果见表5。

表5 37个口语格式在词典、大纲中收录情况

口语格式	词典	旧大纲	新大纲	口语格式	词典	旧大纲	新大纲
不好意思	Y	N	5	没事儿	Y	乙	N
无所谓	Y	N	N	巴不得	Y(口)	N	6
没问题	N	N	N	(就)得了	Y	丙	N
不得了	Y	N	5	你说呢	N	N	N
就这样	N	N	N	说真的	N	N	N
不用说	N	N	N	没什么	Y	乙	N
不怎么(样)	Y	N	N	有的是	Y	乙	N

口语格式	词典	旧大纲	新大纲	口语格式	词典	旧大纲	新大纲
对了	Y	N	N	说了算	N	N	N
你看	N	N	N	说白了	N	N	N
话说回来	N	N	N	怪不得	Y	丙	5
罢了	Y	N	N	怎么搞的	N	N	N
算了	N	N	6	话是这么说	N	N	N
好了	Y	N	N	说不好	N	N	N
(不)要说	N	N	N	要我说	N	N	N
就是了	N	N	N	一句话	N	N	N
不像话	Y	N	6	本来嘛	N	N	N
不要紧	Y	N	5	别看	Y	N	N
没劲	Y	N	N	够朋友	Y	N	N
再说吧	N	N	N	收录总计	18	5	7

表5结果显示,现代汉语词典中收入了18个口语格式,占所有考察格式的近一半。日本学生在书面语中使用的口语格式前十位的有5个收入现代汉语词典(不好意思、无所谓、不得了、不怎么、对了),有5个未收入(没问题、就这样、不用说、你看、话说回来)。由此看来,日本学生是否多用口语格式与词典是否收入关系不大。

另外,现代汉语词典中收录的18个口语格式中,仅有一个词“巴不得”标注了口语,其他都未标注。这一现象说明“口语格式”的认定在现代汉语词典中更为严格。虽然如此,现代汉语词典作为最权威的汉语辞书,在收入哪些口语格式,哪些口语格式应该标注口语方面,应该更为精确、细致,标准应该更清晰、统一。这样对汉语母语者或者二语学习者会更有价值。

在旧HSK大纲中,37个口语格式中仅有5个收入:没事儿、没什么、有的是、(就)得了、怪不得。新HSK大纲中,仅有7个收入:不好意思、不得了、巴不得、算了、不像话、不要紧、怪不得。两个大纲中仅有一个词是相同的(怪不得)。这两个HSK大纲是目前对外汉语教材编写是最为重要的参考,可是对于口语格式的收录差异非常大,并且收入比例也很低。我们同意关于构建对外汉语口语教学语法大纲的观点,“在考虑使用频率、参与到句法结构中所表现出的难易程度以及通用度等三个因素基础上”,^[12]把常用度高的口语格式收入HSK词汇大纲。

(二) 37个口语格式在对话语料中实际使用情况考察

我们在中国传媒大学有声媒体文本语料库中搜索中央电视台的对话语料,共计1584个文本,5958906词次。对37个口语格式在这些文本中出现次数进行统计并排序,结果见表6。

从表6可知,在近600万词次的对话文本中,有18个口语格式使用次数为0。其中“没问题、就这样、对了”三个格式在日本学生的书面语中的频率均值大于10,但是在母语者的对话语料中使用次数为0。虽然语料规模不大,但是也说明本文所研究的37个口语格式常用度不同。词典、大纲在收录口语格式时如果能够以语料库统计为基础,将会更科学。另外,在口语教学中,如果过于强调一些常用度不高的口语格式,可能未必对学生有利。因为留学生学习中文的时间有限,要提高学习效率,在教学安排时就要有所取舍。我们建议今后建设大规模的现代汉语对话平衡语料库(字数最好能够达到1亿以上),对目前口语教材中涉及的口语格式进行统计、排序,据此筛选出口语教学中的重点格式,使留学生既能理解,也能表达。而对于常用度不高的口语格式,可以只要求理解,不要求表达。

表6 37个口语格式在对话语料中的使用情况

格式	次数	格式	次数	格式	次数	格式	次数
不得了	94	就是了	13	没问题	0	你说呢	0
不好意思	65	说真的	10	就这样	0	没什么	0
不用说	64	没事儿	9	对了	0	怎么搞的	0
说了算	47	算了	8	你看	0	话是这么说	0
无所谓	39	不像话	5	话说回来	0	说不好	0
不怎么(样)	29	(就)得了	4	好了	0	要我说	0

格式	次数	格式	次数	格式	次数	格式	次数
怪不得	26	够朋友	4	(不)要说	0	一句话	0
不要紧	23	巴不得	3	没劲	0	本来嘛	0
说白了	21	有的是	2	再说吧	0	别看	0
罢了	16						

如果把表5和表6结合起来观察,会发现比较有意思的结论。对话语料中出现的格式倾向于在词典中收录(比例为19:13),而在对话语料中没有出现的格式则倾向于在词典中不收录(比例为18:13)。这说明现代汉语词典在收录这些口语格式时可能部分考虑到了常用度问题。当然,我们希望作为权威的汉语词典,在收入口语格式时如果能够有大规模口语语料的统计结果的支持会更科学。

三、结论及余论

本文通过对比日本学生和汉语母语者在书面语中使用37个口语格式的差异,发现日本学生汉语书面语中存在口语化情况。另外,还发现随着日本学生汉语水平的提升,其口语化倾向水平并未随之减弱。通过考察37个口语格式在现代汉语词典、HSK大纲中的收录情况,发现日本学生书面语中的口语化与词典、大纲是否收录关系不大。但根据37个口语格式在汉语对话语料中的使用情况,我们建议词典、大纲在收录口语格式时标准应该更为清晰、统一,标注应该更为精确、细致。

本研究结论与英语作为第二语言教学界的发现(如文秋芳等^[1]、潘璠^[3])一致,说明二语学习者书面语中都会存在一定的口语化倾向。既然这种口语化倾向是客观存在的,那么如何看待这种口语化倾向,在汉语教学大纲设计、教材编写、课堂教学实践中如何操作,这些问题更为重要。

书面语中的口语化倾向如同中介语的偏误一样,有一定的规律性,也是二语学习者必然要经历的过程。在二语教学中,应该正确看待这种现象,不能将之看作洪水猛兽、因噎废食,由此取消口语语法或口语词汇的教学。相反,应该加强口语语法、词汇的用法教学,尤其是语体的教学。同时,我们也应该重视和加强书面语的教学,加强对留学生区分不同语体能力的培养。

在具体课程设置上如何保证留学生语体能力的培养能够落到实处?我们赞同废止“综合+小四门”,推广“听说+读写”课堂教学模式。^[13]因为读写课可以侧重书面语体的教学,听说课可以侧重口语语体的教学,因此“听说+读写”模式对于留学生二语语体能力的培养是有利的。

参考文献:

- [1]文秋芳,丁言仁,王文字.中国大学生英语书面语中的口语化倾向[J].外语教学与研究,2003,(4).
- [2]文秋芳.学习者英语语体特征变化的研究[J].外国语,2009,(4).
- [3]潘璠.中国非英语专业本科生和研究生书面语体的多特征多维度调查[J].外语教学与研究,2012,(2).
- [4]汲传波,刘芳芳.留学生汉语书面语中的口语化倾向研究[J].语言教学与研究,2015,(1).
- [5]黄婧.《博雅汉语》语体分布情况的考察分析[D].广西民族大学,2014.
- [6]周芸,张永芹,张婧.论泰国学生汉语语体能力培养的基本原则[J].云南师范大学学报(对外汉语教学版),2011,(2).
- [7]王楨.留学生汉语语体转换能力考察与分析[D].复旦大学,2012.
- [8]刘婕.中级阶段汉语留学生的语体意识及教学对策[D].复旦大学,2012.
- [9]主贵芝.留学生本科毕业论文中的语体特征和语体意识研究——以复旦大学留学生毕业论文为例[D].复旦大学,2014.
- [10]周芸,张婧.泰国学生汉语谈话语体能力习得调查[J].云南师范大学学报(对外汉语教学版),2010,(3).
- [11]丁金国.对外汉语教学中的语体意识[J].烟台大学学报,1997,(1).
- [12]徐晶凝.对外汉语口语教学语法大纲的构建[J].语言教学与研究,2016,(4).
- [13]宗世海.我国汉语教学模式的历史、现状和改革方向[J].华文教学与研究,2016,(1).

(责任编辑:刘伏玲)