

不同母语背景学习者能愿动词「会」的偏误类型—以语料库为本的研究*

张莉萍

台湾师范大学国语教学中心

提要

本研究应用 HSK 动态作文语料库与 TOCFL 学习者语料库分析学习者使用能愿动词「会」的情况，发现「会」的缺漏情形很显着，不同于一般文献中所关心的混用或误用。为探讨缺漏是否为普遍现象，分别观察日、英、韩、越、印度尼西亚、泰不同母语学习者的使用情况，并归纳缺漏情形，建议利用语料库辅助教学与学习，以提高学习者使用「会」的正确率。

关键词：能愿动词、偏误、缺漏、学习者语料库

The error types of the modal verb “会”- A corpus-based study

Abstract

The modal verbs are difficult for L2 learners. Most studies of learners' usage focus on the confusion among different modal verbs. This study analyzes the data in both HSK and TOCFL learner corpora and finds the omission of 会 is the most obvious. Furthermore, different language background learners' corpora including Japanese, English, Korean, Vietnamese, Indonesia and Thai are observed to identify whether the omission is universal. The analysis of omission phenomenon is provided and suggests using the corpus tool to assist teaching and learning.

Key words: modal verb, error, omission, learner corpus, Mandarin

一、前言

能愿动词这一类无论在语言学或华语教学中，都有相当多的讨论，在华语学习上，也是一个公认的难点（邓守信 2009: 113）。一般教师或研究者多从能愿动词混用的角度来分析原因（赖鹏，2006），叶信鸿（2009）则从 HSK 动态作文语料库中，实际统计「会、能、可以、想、要」等五个主要能愿动词的偏误类型发现，缺漏(omission)与误加(addition)这两类远较错用(mis-selection)与错序(mis-ordering)等类型比例来得高，一般文献中所关心的混用或误用

*本研究感谢教育部「迈向顶尖大学计划」与科技部「跨国顶尖研究中心计划」(NSC 103-2911-I-003-301)，以及台湾师范大学「华语文与科技研究中心」支持。

情况反而不是偏误较高的类型。在缺漏类的偏误中，又以「会」的偏误率最高。「会」的偏误笔数或频率也是五个能愿动词中最高的（叶信鸿，2009：105）。

在这个研究里，我们把讨论重心放在能愿动词「会」上面，探讨几个问题：

1. 两岸学习者语料库是否呈现同样的趋势，即「会」的缺漏偏误比例最高？
2. 不同母语背景学习者是否呈现同样趋势？
3. 针对学习者「会」的偏误情况提出因应的教学建议。

二、能愿动词「会」的偏误表现

本研究所使用的学习者语料库有两个，一是大陆的 HSK 动态作文语料库（张宝林等，2004），一是台湾的 TOCFL 学习者语料库（张莉萍，2013b）。两者性质接近，都是来自母语非汉语的能力考试作文语料。前者搜集了 1992-2006 年的作文，共 11569 篇，424 万字，皆为中高级水平学习者语料；后者搜集 2006 年至今的作文，语料仍在持续收集中，截至目前为止约 4 千多篇，114 万词（174 万字左右），学习者语言能力涵盖初级到中高级，来自世界 39 种不同母语背景。HSK 语料库已经为语料做了偏误标记，TOCFL 语料库仅为百分之六十语料做了初步的偏误标记，不过，由于标记相当简陋，例如，针对缺漏，只标记[M]，无法得知缺漏成分为何，因此个别研究需要经由人工一笔一笔检视分析。

以往无论是根据教师的经验或语料库，多仅能呈现混用的统计数据，因为标记大部分是针对学习者已经产出的语料，而少有汉语学习者语料库将缺失的成分也标记出来，耗时费力是主要原因。HSK 语料库已做出了重要的贡献，据此，叶信鸿（2009）统计出「会」的偏误以缺漏类最多，引起了研究者的注意。根据他的统计，缺漏、错用、误加的比例分别是 44%、21%、19%，缺漏类很突出，错用和误加两类则没有显著差异。

由于缺漏类的突出与一般教学者的经验不一致，究竟语料库所呈现的数据不可信，不同区域学习者语料库所呈现出来的趋势是否相同是本研究的重点。为了统计的可比性，本研究针对 TOCFL 中高级学习者（CEFR 的 B2 等级，相当于 HSK 语料库中的学习者程度）的语料进行分析。为了进一步分析不同母语背景学习者在「会」的偏误类型上，是否表现一样，在这个研究中也作了国别化的分析观察，分别观察了 TOCFL 前六高母语语料，即日语、英语、韩语、泰语、印度尼西亚语、越南语。表一是统计的结果。

表 1 TOCFL 语料库「会」使用偏误类型统计

	篇数	使用次数	偏误小计	缺漏次数	%	误加次数	%	错用次数	%
日	260	396 (0.53)	305	204	66.89	81	26.56	20	6.56
英	122	418 (1.07)	127	54	42.52	37	29.13	36	28.35
韩	130	268 (0.68)	42	16	38.10	9	21.43	17	40.48
泰	41	96 (0.89)	28	10	35.71	4	14.29	14	50.00
印度尼西亚	112	335 (1.05)	38	2	5.26	3	7.89	33	86.84
越	188	401 (0.73)	28	5	15.15	12	36.36	16	48.48
平均					33.94		22.61		43.45

从平均偏误率来看，TOCFL 显示的是错用类型最高（43%），其次是缺漏类型（34%），与 HSK 的统计结果缺漏类型最高不同。张莉萍（2013a: 79）利用语料库工具中的主题关键性方法（keyword-keyness）分析日、韩、英、印度尼西亚、泰、越南六个不同母语背景语料，发现英语、印度尼西亚语母语者使用「会」的高关键性（即使用频率较其他四种母语者高，见表 1 使用次数中的圆括号数字），推测因为英语助动词和汉语能愿动词有相近的语义和语法范畴，可能英语者缺漏类偏误较少，因此相对使用频率高。然而从表 1 的统计数据可以看出这个推测并不成功，虽然英语者缺漏频率较日语者低（43%: 67%），但还是高于其他语种。表 2 的 HSK 统计数据，甚至显示英语者缺漏较其他语种高，不过，由于 HSK 语料库并没有纪录考生的母语，而仅记录国籍。因此本研究仅能将英、美、加拿大国籍语料归为英语母语，未必反映英语母语者的真实情况，此数据仅能做为参考。

表 2 HSK 语料库「会」使用偏误类型统计

	偏误小计	缺漏次数	%	误加次数	%	错用次数	%
日	158	81	51.27	32	20.25	45	28.48
英	13	10	76.92	0	0.00	3	23.08
韩	237	150	63.29	52	21.94	35	14.77
泰	18	13	72.22	1	5.56	4	22.22
印度尼西亚	23	6	26.09	10	43.48	7	30.43
越	6	4	66.67	0	0.00	2	33.33
平均			59.41		15.20		25.39

从上面两个表，可以看出这两个学习者语料库所呈现的偏误类型统计数据虽然有些差异，但也有一致性，一致的是日语、英语学习者皆以缺漏的偏误类型最显著。虽然 TOCFL 语料库呈现缺漏并非最高频率的偏误类型，不过，从两岸语料库的数据，还是足以让我们意识到学习者对「会」缺漏的严重性，以往研究者多偏重在能愿动词之间的混用辨析，以语料库为本的研究结果则让我们看到学习者更全面性的使用情况。

至于不同母语者的偏误表现，HSK 语料库显示的是除了印度尼西亚以外，其他五种母语者以缺漏类最高；TOCFL 语料库则是除了日、英外，其他四种母语者以错用为多。由于泰、印度尼西亚、越南偏误笔数不多，统计结果不一定可信。不过，还是可以看到在这些不同母语背景者间，似乎印度尼西亚学习者偏误表现特别不一样，不论在 HSK 或 TOCFL 语料库中缺漏类比例都是三类中最低的。原因为何，未来可以进一步探讨。

由于讨论能愿动词混用文献较多，本文限于篇幅，以下仅归纳缺漏发生的情况并举学习者用例以做为未来教学建议的基础。（以下仅标示「会」缺漏处，不更正其他偏误）

1. 出现在假设句或条件句的后一分句，这种用例在语料库中最多，我们仅举 3 例：
 - (1) 如果你工厂可以进步声音和臭味的问题，我们都{会}很快乐
 - (2) 在现代生活上，由只有先生的薪水来生活{会}很辛苦，
 - (3) 儿童、青少年的想法是未成熟的，需要许多经验来让他们思考，最后才{会}有正确的想法

2. 出现在保证言语行为的，通常以「会...的」形式出现：
 - (4) 我和家人觉得非常惊讶，因我通常{会}注意饮食安全的。
 - (5) 希望李先生能够了解一个小小的消费者的情境，并相信政府当局{会}妥善处理。
3. 推测未来可能发生的事：
 - (6) 我认为我们难免{会}遇到这样的误会。
 - (7) 这样做说不定公司{会}听你们的要求。
 - (8) 以后一定可以从事理想的工作，薪水也{会}高多了吧！
4. 表达结果发生的可能性：
 - (9) 我觉得最好妳跟李大年多加讨论为什么这个误会{会}发生了。
 - (10) 噪音与臭味除了{会}伤害附近居民之外，也{会}伤害环境。

上述 3 和 4 的情况很接近，差别只在于 3 是对于未来的推测，4 则不一定指未来，例如 (9) 中「为什么会发生误会」，是已经发生的事。从上面的例子也可以看出，关于表示能力的「会」，学生几乎没有偏误用法，有问题的都是表示可能性的「会」。

三、能愿动词「会」的教学建议

叶信鸿调查了海内外四本常用华语教材对于「会」的教学说明，并根据「会」的偏误情况，就教材不足之处，提出了中肯的补充 (p.124)。建议教材在呈现假设句时，需要特别说明「会」使用的情况，例如，「如果...的话」后面的子句多带有「会」，表示事件发生的可能性；另表示「保证类」言语行为的直述句，如果句尾有「的」，也要加上「会」。除此之外，我们建议可以利用语料库来辅助教学与学习。因为虽然教材中均对表示能力的「会」与表示可能性的「会」做了说明，部分教材也对不同能愿动词之间的使用做了辨析，但「会」的缺漏情况是一般学习者的普遍现象，显示仅针对「会」做语义说明或辨析是不够的，如何建立学习者的语感，知道在什么样的语境之下，需要使用「会」，才能有效减少缺漏的情形。

而语料库可以具体应用在教材编辑中，就上一节所归纳的四种缺漏情况，除了第一、二种，有形式上的线索来帮助学习者学习，如「如果」「...的话」「会...的」外，第三、四种情形也可以有有形的形式协助学习者接触各式表示可能性「会」的语境。例如利用语料库工具中词语搭配关系(collocation)以及互见讯息(mutual information, MI value)，可以找出可能性「会」的语境。

首先，我们找出「会」左边 5 个词语的搭配关系，发现互见值高（也就是共同出现关系紧密）的词有「迟早、总有一天、久而久之、偶尔、难免、一定、有时候、多多少少、说不定」等等，也就是这些词汇的后面常常伴随者「会」共同出现，这么一来，在提供教材例句时，就可以选用这些词语造出适当的语境。如果运用数位教学，也可以整理出含有这些关键词语的句子让学习者大量阅读以建立语感。另外，我们试图寻找「会」的右边 5 个词的搭配关系，发现「危害到、拖垮、好起来、变好、影响到、伤害到、流出来、意想不到、卡住、伤到」等词语互见值很高，例如，我们在 1000 万词的语料库中实际搜寻，「好起来」出现 15 次，其中 9 次都和「会」共同出现，而这些高互见值的词有一个共性，即都是动补结构动词，和「会」共同出现，表达的是结果发生的可能性。如(11)-(13)：

(11) 我们当然担心农地上不当的兴建农舍，会危害到社会上的公与义...

(12) 转诊制度若是不能严格执行，不仅会拖垮全民健保的财务...

(13) ...当地的货币、股市都会跟着好起来。

在教学上，多安排动补结构动词与「会」共同出现的阅读材料，能帮助学习者快速有效地掌握「会」的使用语境。

四、结语

透过学习者语料库的观察分析，我们看到了学习者使用目标语的实际情形，不论从 HSK 或 TOCFL 语料库都让我们注意到了「会」的缺漏现象。希望这个发现可以让研究者或教学者稍微转移对能愿动词这个领域关注的方向，未来不只着力在不同能愿动词之间的辨析，也能进一步探讨能愿动词使用的语境，以有效协助学习者在该用的情况下正确使用。

参考文献

- 邓守信（2009）对外汉语教学语法（修订二版）。台北市：文鹤出版社。
- 赖鹏（2006）汉语能愿动词语际迁移偏误生成原因初探。语言教学与研究，2006(5)，67-74。
- 叶信鸿（2009）现代汉语助动词的界定与教学应用。国立台湾师范大学华语语文教学研究所硕士论文（未出版）。
- 张宝林、崔希亮、任杰（2004）关于“HSK 动态作文语料库”的建设构想。2010 年 3 月 15 日，取自 http://china-language.gov.cn/54/2007_6_20/1_54_1356_0_1182321111359.html
- 张莉萍（2013a）欧洲共同架构下的汉语词汇语法特征。国立台湾师范大学华语语文教学系暨研究所博士论文（未出版）。
- 张莉萍（2013b）TOCFL 作文语料库的建置与应用，载于崔希亮、张宝林（主编），第二届汉语中介语语料库建设与应用国际学术讨论会论文集（页 141-152）。北京：北京语言大学出版社。