

硕士研究生学位论文

**题目：构式视角下的现代汉语四字表达的能产性研究**

**姓　名**：　　　田　骏

**学　号**：　　1301211280

**院　系**：　中国语言文学系

**专　业**：　 汉语言文字学

**研究方向**：　 中文信息处理

**指导教师**：　 詹卫东　教授

二〇一六 年 六 月

**版权声明**

任何收存和保管本论文各种版本的单位和个人，未经本论文作者同意，不得将本论文转借他人，亦不得随意复制、抄录、拍照或以任何方式传播。否则，引起有碍作者著作权之问题，将可能承担法律责任。

# 摘要

A Productive Research on Modern Chinese Four-Word Expression: from a Constructional Perspective

Tian Jun (Chinese Linguistics)

Directed by Professor Zhan Weidong

# ABSTRACT

KEY WORDS:

# 目　录

摘要 I

ABSTRACT II

目　录 III

第一章　引言 1

1.1 问题的提出 1

1.2选题背景及意义 1

1.3文献综述 2

1.3.1 不同的术语体系 2

1.4研究方法 2

1.5论文结构安排 2

参考文献 3

致谢 4

北京大学学位论文原创性声明和使用授权说明 5

# 第一章　引言

## 1.1 问题的提出

四字表达在汉语中占据着重要地位。粗看下，四字表达可分为两种：成语和常规短语。“喜极而泣”是成语，需要收进词典；而“中国制造”用一个偏正式短语规则就能很好地处理。但仔细考察，则会遇到一些困难。例如：“悲极而泣”是成语词，还是自由短语呢？

如果看作自由短语，组成这个短语的成分却并不都是词；我们也会感觉到这些成分的组合并不是完全自由的，成分间结合的紧密程度也不一样。而如果看作词，这样的词“临时性”很强，不是那种会进入我们的“语言素材库”中、被我们频繁运用的词。

类似的表达数量非常之多，因为它们的表达效果非常强，甚至盖过了它们作为信息载体、传递信息的作用。随着个性的解放和互联网的发展，这种表达非常符合“网络虚拟社区”的交际需求：为了避免常规表述的乏味，引起听话者的兴趣，人们乐意创造并使用它们。

《现代汉语词典》，及一些前人的研究，已经开始通过阐释“格式”的概念，如“千…万…”，来规整这些表达，并启发人们创造。但我们认为，这种处理方式只是在帮助人们对已有的表达进行认知，并未触及它们繁荣的本质。“更准确地描述并论证这些四字表达的能产规律，即人们究竟是如何创造出它们的”，是本文最核心的问题。

（我们认为这种表达偏向词的性质，但不应该收入词典中[[1]](#footnote-1)。因为这种表达足够临时、足够低频，而意义又足够清晰，并且足够能产；收进词典，既起不到太多“正视听”的作用，更重要的是“收之不尽”，没有办法启发人们造出更多的表达。

（但我们认为这样的努力作用是很有限的。一方面，这些表达的意义是很直白的，非要提炼出一个格式义，只会把现象复杂化，我们也不相信人们在理解“千言万语”、“千刀万剐”的基础上会提炼出一个“千…万…”的上位概念；更重要的是，在缺乏语感积累的前提下，有多少人能在遇到一个急切的表达需求时，准确地想到并使用这些格式呢？）

## 1.2 选题背景及意义

## 1.3 文献综述

有关四字表达的研究很丰富，光术语体系就分好几套：以（陆志韦，1956）为开端的“四字格”派，以（文炼[[2]](#footnote-2)，1988）为开端的“类固定格式”派，以（张国宪，1993）为开端的“对举格式”派[[3]](#footnote-3)，以（周荐，1999）为开端的“待嵌格式”派等等。此外，还有一些学者的相关术语也为后人借鉴，如（张拱贵，1985）提出的“语汇格式”，（李宇明，1999）提出的“词语模”，（邵敬敏，2008）提出的“框式结构”等。从发表时间、研究角度和研究方法上进行比较，这些研究呈现出从定性到定量、从现象到体系的阶段性特点，显示出学者对问题的认识逐渐深入、具体化。在1.3.1小节中，我们首先按术语体系对前人研究进行综述；在1.3.2小节中，我们再依阶段对前人研究进行总结；在1.3.3小节中，我们最后指出前人研究的不足之处。

1.3.1 不同的术语体系

“四字格”派

（陆志韦，1956）首称所有完整的有意义的四字组合为“四字格”，并把“并立四字格”定义为内部紧凑、形式上为“2+2”式并立的四字格。他对四字格的分类依据首先是形式上是否具有重复成分（即叠字，无论是否相邻），其次再按前后两个成分各自的语法结构分类（如动宾式、主谓式等等），考察了许多北京话的实例。这篇文章的优势有二，一是注意到了并立，也就是后人所谓的“对举”，是相当有能产性的；二是注意到并立四字格的并立特征要远大于其内部结构特征，是一个整体性的视角。值得一提的是，作者注意到了并立四字格可以变换字的顺序（如：开天辟地——开辟天地），但没有深入下去，反而认为后者失去了并立性而排出在讨论范围之外，有些遗憾。

（姚殿芳，潘兆明，1985）进一步分析了某些四字格的产生方式，如“扩展”：短兵（相）接，或“压缩”：（愿）乘（长）风破（万里）浪，利用虚字补足音节：‘不正当的风气’->‘不正之风’等等。作者认为四字格的应用场景广泛，无论是新中国诞生初期还是改革开放以来，都诞生了一大批新的四字格，应当重视。

（马国凡，1987）也提到了某些四字格的形成途径，如词的AXAB式扩展、词组的紧缩与节缩等，其中最具启发意义的是他称之为的成语的衍化格式。他认为成语的衍化有非固定的和框架格式衍化，前者如从“人云亦云”衍化出“‘帮’云亦云”（指帮派、派性）、“人云不云”（不随便附和）、“人倒亦倒”的说法，后者主要通过向一些“A＿B＿”式的框架中填入临时需要的成分构成。

“类固定格式”派

（文炼，1988）指出，判断固定短语的标准，一方面在于其整体义是否大于部分义之和，另一方面在于其是作为语言材料供遣词造句，而非根据交际需要临时创造的。但他随后指出类固定短语既无意义引申，又是临时创造，却在格式和功能上类似固定短语。他认为这样的表达“在在汉语中有稳定、庄重的色彩”，使用面广，且具有向固定短语转化的特质。作者还观察到一类特殊的现象，如“水落油出”[[4]](#footnote-4)，他认为这种仿造的成语必须依附被仿造的成语，才能显示出它的修辞效果，然而并未深究，有些遗憾。

（齐沪扬，2001）进一步强调类固定短语“结构模式固定”和“意义单一”两个特点，其中结构模式固定是指“类固定短语是是替换了成语中的一些组成成分后构成的”。作者根据替换成分的多少分成了替换两个成分和替换三个成分两类，在前一类中又进一步分出“保留数词”、“保留相同语素”和“保留同义或反义语素”三类，基本都属于“对举”类型。

（李传军，2008）更系统地考察了类固定短语。他沿用齐沪扬的按替换成分分类方式，把类固定短语分为偶为内嵌型（A…C…）、奇位内嵌型（…B…D）、首尾内嵌型（…BC…）、一字支撑型（……C…，…B……）和组块内嵌型（其他格式）。他采用定量方法，“例不十，不立格”，提炼出300个格式。他还把类固定短语的能产方式总结为“类推”和“扩展”，与前人结论类似。

“待嵌格式”派

（周荐，1999）观察到《现代汉语词典》中存在这样一类特殊的词条，像“不…不…”、“千…万…”，他在该文的附注中首次采用“待嵌格式”来命名这一类词条，而《现代汉语词典》本身并未对这一类词条做出更多的解释。（周荐，2001）认为，之所以设定此类条目，是因为其在口语中出现频繁，却不足以设定条目，以“待嵌格式”以纲带目，目的是引导人们用某种架构去类推创造出新的固定语。他认为，四字格是待嵌格式最适宜选择的格式，而且《现代汉语词典》中绝大多数格式都是“一三字显，二四字隐”。他认为，“一三字隐，二四字显”的格式同样存在，《现代汉语词典》中没有收录，应该是无法为之在音序序列中找到一个合适的位置的缘故，属于技术处理上的困难，他将这两类待嵌格式统称为交替显现[[5]](#footnote-5)待嵌格式。此外，他还举出《现代汉语词典》中收录的“自相……”一例，说明这类邻接显现的格式远不如交替显现的格式的能产性强。

（范喜梅，2009）重点分析了由成对方位词构成的待嵌格式。她认为，这些待嵌格式内部结构虽大多并列，却要遵循一种约定俗成的语序，如“上”在“下”前，“左”在“右”前。作者还指出，在并列关系之外，还存在转折、因果、目的等其他关系。此外，作者还为方位待嵌格式拟出了6个释义模板。

（孟祥英，2010）对待嵌格式进行了系统研究，她整理出240个固定字与填充字交替出现的待嵌格式，并根据固定字的词性组合、意义异同分为多个小类。作者指出，固定部分多属于基本词汇范畴，且多具丰厚的文化蕴涵（如“三”表多）。嵌入部分则与固定部分的词性、词义相关，有时候可以嵌入多种词性。作者很好地总结、理清了前人的相关研究，可以算是集大成的作品。

其他术语体系下的研究

（张拱贵，1985）中首次提出了语汇格式的概念。这篇文章是作者为朱林清《现代汉语“格式”初探》作的序，他指出，朱作中分析的格式都是语法格式，这些格式包含固定词语，一般充作关联词组构成短语，且使用上带有某种习惯性，不能用一般语法规则解释。他认为，朱林清对格式的研究是启发性的，但“‘格式’不仅语法里有，语汇里（成语、惯用语、谚语等）更多”。他举出自己于1963年发表的《成语中的成套格式》一文，说明此类格式在成语中的大量存在。

（冯胜利，1997）从韵律角度讨论了四字格式的成因，他提出了格套化理论，认为某些四字格式的形成源于双音节词的格套化，即将双音节词拆开，中间填入某种具有音缀性质的填充词而成。他认为，双音节词以[[＿A][＿B]]形式拆分居多，而填充词用久用熟以后，便能形成[X＿Y＿]式的格套，从而产生更多的词。

（李宇明，1999）提出了词语模的概念：具有批量产生新词语的框架，其中“模标”指不变的成分，而“模槽”指空位，例如“X盲”中，“盲”是模标，“X”是模槽。他还强调，词语模不仅仅是一种创造词语的模子，也是一种思维模式，模标具有一种“类”功能，例如人们习惯于把有目的、有计划、影响面广、意义重大的工作称为“X工程”，“X工程”便成了一个思维的模式。而他认为，词语模主要分前空型和后空型，中空型的词语模很少，如“吃X饭”。

（黄燕璇，2011）沿用（邵敬敏，2008）提出的框式结构概念分析了方位对举四字结构的特点。她认为，这类四字结构中的方位词大多已虚化，虚化后才能形成“量多”这个结构义。此外，作者还提到成语的凝固性分两种，一种是来自典故的、一开始就凝固的结构，另一种则是经过繁衍、筛选、淘汰后形成的优势结构。因此诸如“东张西望”这样的凝固成语，从历史的角度也属于框式结构。

（韩立秋，2015）则用构式语法去分析古汉语的凝固结构，提出了许多前人不曾涉及到的凝固结构，如“不亦…乎”、“何其……”等。他强调语言一直处于一个动态过程中，“一旦临时结构所反映的认知经验得到了认同，并能够引起人们再次使用的兴趣时，凝固结构就很有可能在重复使用中发生变化”。他还指出，“随着使用频率的提高，各个成分的本源义就会被遗忘，人们对其内部形式就可能做出重新分析”，我们认为这是非常重要的能产机制。

1.3.2 前人研究的阶段性特点

从定性到定量

“定量分析”是近年来的大趋势，也是青年学者的优势所在。2000年之后，相关的研究基本都带有“语料库”特色，且无一例外地选用了北京大学中国语言学研究中心（CCL）在线语料库，这可能与该语料库提供丰富的查询手段、能很好地查找结构不连续的语法和语汇格式有关。

在定量分析上，数李传军的研究最为透彻。他提出了“型式的能产性与格式的能产性并不是简单的正相关关系”观点，从“型式——格式——语例”三层对涉及的类固定格式进行了定量描述，清晰地描述了这类短语的总体分布情况。

从现象到体系

阶段性的另一个体现是从“定义”到“释义”。早前，学者们多从不同的角度对所见到的语言现象进行了描述，因而产生了许多术语；而随着汉语语法研究的深入和国外先进理论的引入，青年学者们更倾向用各种语法理论框架去解释这些现象。

一方面，范喜梅、黄燕旋等把关注点从整体现象聚焦到某一类（方位对举格式）具体现象上，将这类实例丰富、语义抽象的格式描述得更加立体；另一方面，李传军、孟祥英等在前人基础上将类目进一步细化，力图建立起坚固的体系，从而能概括、解释更多的语言实例。

总而言之，随着研究的深入，人们对四字表达的认识逐渐由感性上升到理性，逐渐建立起一些共识，如语法性质更偏向词、语义透明而表现力丰富等，清楚地认识到它在汉语中的重要地位。

1.3.3 前人研究的不足

## 1.4研究方法

## 1.5论文结构安排

# 参考文献

# 致谢

# 北京大学学位论文原创性声明和使用授权说明

**原创性声明**

本人郑重声明： 所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文作者签名： 日期： 年 月 日

**学位论文使用授权说明**

（必须装订在提交学校图书馆的印刷本）

本人完全了解北京大学关于收集、保存、使用学位论文的规定，即：

* 按照学校要求提交学位论文的印刷本和电子版本；
* 学校有权保存学位论文的印刷本和电子版，并提供目录检索与阅览服务，在校园网上提供服务；
* 学校可以采用影印、缩印、数字化或其它复制手段保存论文；
* 因某种特殊原因需要延迟发布学位论文电子版，授权学校□一年/□两年/□三年以后，在校园网上全文发布。

（保密论文在解密后遵守此规定）

论文作者签名： 导师签名：

日期： 年 月 日

1. 当然，作为信息处理提高匹配效率和匹配准确率的手段，对于这种表达“见一个收一个”也无可非议。 [↑](#footnote-ref-1)
2. 即张斌，文炼为其笔名。 [↑](#footnote-ref-2)
3. 张国宪这篇文章并不针对四字表达进行讨论，但后人研究多以其文为参考，沿用“对举”概念。另外，“对举格式”只是四字表达中的一部分，不能概括四字表达全貌，这与其他派系不同。 [↑](#footnote-ref-3)
4. 例句如下：“龙二井有油和水的矛盾，这是它的特殊性。周队长说，要促使矛盾转化，就要捞水，把水捞干。我们想一不做，二不休，搞它个水落油出。” [↑](#footnote-ref-4)
5. 原文中称“两字交替显现，两字交替隐含”，我们简称为交替显现；同样地，后面的邻接显现，原文为“两字邻接显现，两字邻接隐含”。 [↑](#footnote-ref-5)