




英汉程式语 心理表征对比研究

贾冠杰 王 云 李更春◎著

 科学出版社

国家社会科学规划基金项目“英汉程式语心理表征模式对比研究”（12BYY009）资助出版

英汉程式语 心理表征对比研究

贾冠杰 王 云 李更春◎著

科学出版社

北京

内 容 简 介

本书基于心理语言学理论,结合具体语言实际,对英语和汉语程式语心理表征的异同进行了详细认真的对比分析研究,是国内相关学科少有的专著,吸收了国内外心理学、心理语言学等学科的最新研究成果,理论联系实际。全书内容包括相关的综述类研究和实证研究两大部分,有助于揭示英汉程式语心理表征异同的深层根源,具有重要的理论和实践价值。

本书适合从事心理语言学研究的学者、大学英语教师和学生阅读使用。

图书在版编目(CIP)数据

英汉程式语心理表征对比研究 / 贾冠杰, 王云, 李更春著. —北京: 科学出版社, 2019.10

ISBN 978-7-03-062650-9

I. ①英... II. ①贾... ②王... ③李... III. ①心理语言学-对比研究-英语、汉语 IV. ①H0-05

中国版本图书馆 CIP 数据核字(2019)第 231276 号

责任编辑: 常春娥 / 责任校对: 贾娜娜

责任印制: 李 彤 / 封面设计: 润一文化

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京虎彩文化传播有限公司 印刷

科学出版社发行 各地新华书店经销

*

2019 年 10 月第 一 版 开本: 720×1000 B5

2019 年 10 月第一次印刷 印张: 11 3/4

字数: 197 000

定价: 88.00 元

(如有印装质量问题, 我社负责调换)

前 言

人类语言交际在很大程度上使用程式语，在口语和书面语的交流中，程式语起着举足轻重的作用。目前，国内外对程式语的定义极不统一，可以说是智者见智、百花齐放。较有代表性的英语程式语术语有 formulaic sequences、formulaic language、lexical bundles、lexical phrases、prefabricated chunks 等。相关的汉语术语除了程式语外，还有公式（化）语（言）、程式化序列、词块、（预制）语块、词汇短语、（词汇）组块、套语、多词单位、惯用语、固定用语、习语、复现组合、两个词/词以上的词串、言语程式、词的联合、词汇化句干等。之所以有如此多的术语，主要是研究者从不同的视角对程式语进行研究。程式语的相关术语之多，说明程式语是一个开放性的术语，研究者在具体的研究中，可以根据自己研究材料的需要去定义，把自己的研究定义在具体的研究框架之内。程式语覆盖面很广，具体研究只能涉及其中很小的一部分，因此，本书仅仅涉及程式语的一个很小的点，其中英汉程式语对比研究只涉及心理动词程式语。

从国内外相关研究来看，大都是以本族语者或二语习得者为研究对象，缺乏对英汉程式语心理表征的对比研究，因此，本书的研究就显得十分必要和及时。本书运用多种心理语言学实验方法，研究英汉程式语的心理表征，通过对比研究，在一定程度上揭示了英语本族语者、汉语本族语者和外语学习者程式语的意识水平、习得程度及其语言加工策略的选择模式等，有助于揭示英汉程式语心理表征异同的深层根源，其研究结果为相关理论研究、外语学习和外语教学提供了有价值的参考依据。

本书吸收了国内外心理学、心理语言学、应用语言学等学科的最新研究成果，查阅了国内外大量相关文献，深入调查分析，积极开展实证研究，理论联系实际。本书充分利用笔者在英国一年多的学习机会，对 60

位英语本族语者进行了多个实证研究，取得了相关的研究数据。同时，其他几位合作者对不同数量的中国英语学习者开展了多项实证研究，为英汉程式语开展比较研究提供了数据支持。在笔者的共同努力下，经过几年的深入研究，取得了比较理想的研究成果。

本书共 10 章，除第 1 章概述外，主要包括两大部分：第一部分包括五章，重点分析讨论相关的研究，如程式语加工研究、程式语加工的研究方法、心理表征研究、心理词汇研究等；第二部分包括四章，主要通过实证研究展开分析讨论，如英汉心理动词程式语心理表征对比研究、英汉心理动词程式语语音表征对比研究、英汉心理动词程式语语符逆序表征对比研究、英语程式语心理表征模式研究等。

本书的出版得到了国家社会科学规划基金项目的资助，科学出版社常春娥等多位老师给予了精心指导和大力支持。在完成国家社会科学规划基金项目和编写本书的过程中，很多同事、朋友、学生等都给予了笔者真诚帮助。同时，本书参考了一些国内外学者的研究成果，在此一并对以上单位和个人表示衷心感谢。

由于时间仓促，加之笔者水平所限，书中可能存在不妥之处，敬请广大读者批评指正。

贾冠杰

2019 年 1 月

目 录

前言

第 1 章 概述	1
1.1 相关研究梳理	1
1.2 主要概念	4
第 2 章 程式语加工研究	15
2.1 引言	15
2.2 程式语加工研究的进程	16
2.3 程式语加工研究的主要内容	18
2.4 程式语加工研究存在的问题	22
2.5 程式语加工研究未来的研究方向	23
2.6 小结	25
第 3 章 研究方法	26
3.1 程式语加工的实验方法	26
3.2 第二语言心理语言学研究方法	35
第 4 章 心理表征	41
4.1 引言	41
4.2 心理表征的定义	41
4.3 心理表征研究的发展	43
4.4 心理表征模型	47
4.5 小结	52
第 5 章 心理词汇	53

5.1	引言	53
5.2	心理词汇研究的发展	54
5.3	心理词汇模式	56
5.4	<i>The Mental Lexicon</i> 杂志十年研究分析	65
5.5	心理词汇研究存在的问题	69
5.6	未来研究方向	70
5.7	小结	72
第 6 章	多语心理词汇模式及其应用	73
6.1	引言	73
6.2	单语心理词汇模式	74
6.3	多语心理词汇模式	75
6.4	多语心理词汇模式及其应用	82
6.5	小结	84
第 7 章	英汉心理动词程式语心理表征对比	86
7.1	引言	86
7.2	研究设计	88
7.3	结果与讨论	92
7.4	小结	100
第 8 章	英汉心理动词程式语语音表征对比	102
8.1	引言	102
8.2	相关研究	103
8.3	研究设计	105
8.4	结果与讨论	108
8.5	小结	115
第 9 章	英汉心理动词程式语语符逆序表征对比	117
9.1	引言	117
9.2	研究背景	117

9.3 研究设计	119
9.4 研究结果	123
9.5 讨论	127
9.6 小结	130
第 10 章 英语程式语心理表征模式	132
10.1 引言	132
10.2 实验一：自定步速逐词阅读实验	133
10.3 实验二：自定步速逐句阅读实验	138
10.4 我国英语学习者程式语心理加工与表征模型	141
10.5 小结	144
参考文献	145
附录	177

第1章

概 述

关于程式语的研究，近些年来引起了国内外众多学者的高度重视，本章重点讨论相关研究，两个重要的术语“程式语”和“心理表征”以及本书的结构。

1.1 相关研究梳理

语言学习本质上是学习和理解程式语 (Ellis, 2012)。程式语在我们的生活中无处不在，在口语和书面语的交流中，程式语起着举足轻重的作用。早在 1916 年，Ferdinand de Saussure 就谈到了粘着法构词 (agglutinations)，即两个或多个语言符号通过无意的融合 (fusion) 而成为一个非分析性单位。1925 年，Harold Palmer 指出了多词或“词汇”组块在语言学习中的重要性。在心理学领域，Miller (1956) 在一篇关于短时记忆容量的 7 ± 2 制约的经典论文中提出了“组块化” (chunking) 这种语言加工策略。此外，语料库语言学研究使人们认识到：语言具有双重属性，即可分析性 (analyticity) 和程式性 (formulaicity)。它既是一个以语法为基础的、可分析的、相对封闭的系统，又是一个以记忆为基础的、程式化的、相对开放的系统 (Skehan, 1998, 转引自丁言仁、戚焱, 2005: 49)。在语言习得研究领域，学者们也研究了相同的现象。这种单词组合既是第一语言习得又是第二语言习得的重要特征。然而，许多关于程式语的研究要么是基于语料库的描述性研究，要么是基于语言习得的应用性研究。前者侧重于对程式语加以识别，对其在语篇中的用法进行描述，后者主要研究语言初学者 (母语习得者或二语学习

者)产出的程式语。

在国内,关于程式语普遍问题的研究和一般心理表征的研究有一些相关成果,但是,对程式语心理表征的研究非常少。李红和缪道蓉(2009)对程式语(他们称之为“公式化语言”)的心理加工与表征进行过研究,他们研究了公式化语言频率及语言水平是否影响公式化语言在中国英语学习者二语心理词库中的整体表征和提取,其研究结果显示公式化语言频率对整体加工有一定的作用。但是,由于李红和缪道蓉的文章是一篇期刊论文,有很多需要研究的问题该篇论文并未涉及。另外,通过对“中国知网”近20年(1999—2018年)相关文章的搜索,只有3篇关于“英汉程式语心理表征模式对比研究”的相关文章:“英汉心理动词短语心理表征对比研究——来自直线性序列式语法判断范式的证据”(王云、贾冠杰,2017),“英汉心理动词双字/词语符逆序表征对比研究——来自掩蔽启动范式的证据”(王云、刘新芳,2018),“英汉心理动词程式语语音表征对比研究——来自视觉掩蔽启动范式的证据”(王云、刘新芳,2018)。在我国26年(1993—2018年)的国家社科项目里,笔者获批的项目是唯一的关于“英汉程式语心理表征模式对比研究”的项目。

国外对程式语心理表征进行研究的时间和相关研究成果的出现都早于我国,如Sosa和MacFarlane(2002)运用单词监控范式研究了受试对不同频率的短语中功能词“of”的反应时间;Schmitt、Grandage和Adolphs(2004)使用两种方法研究了复现词(recurrent clusters)的心理表征问题,即首先通过语料库分析识别出若干复现词,然后将它们嵌入某种心理语言学听写任务中,以研究它们是否以整体形式进行存储;Underwood、Schmitt和Galpin(2004)使用眼动跟踪方法研究了受试在阅读过程中是如何加工程式语的;Conklin和Schmitt(2008)通过比较本族语者与非本族语者对程式语和与其匹配的非程式语的阅读时间,研究了程式语可能具有的加工优势;Siyanova和Schmitt(2008)进行的系列研究探讨了二语学习者对形-名搭配的产出与加工;Tremblay(2009)研究了从语料库中提取的四词和五词程式语是否被受试以整体形式进行存储和加工;Ellis、Simpson-Vlach和Maynard(2008)则运用语料库语言学、心理语言学和语

言教学方法对程式语进行了研究。Elke (2015) 研究了基于课堂教学的外显性词汇教学对 24 名二语学者识别、线索输出及运用学术程式语的影响, 结果显示学生们收到了明显的效果。Hoa (2016) 采用了以交际和意义为中心的三种不同的关注形式教学的方法 (form-focused instruction) 对中国英语学习者运用英语程式语情况进行了研究; 结果显示, 学生经过三种任务的训练, 他们的多项测试成绩都有了明显进步。D. Sidtis 和 J. J. Sidtis (2018) 基于右脑皮层下过程, 研究了程式语的情感性, 证据支持了语言的双过程模式, 右脑皮层下系统调节了程式语。Rafieyan (2018) 通过对 45 名日本英语学习者的程式语知识和他们语言熟悉程度的关系进行了研究, 本研究有助于目标程式语在外语课堂教学中的作用。

纵观国内外研究现状, 虽有一些相关研究成果, 但是, 到 2018 年为止, 还没有看到对英汉程式语心理表征模式进行对比的相关研究。

从国内外相关研究来看, 大都是以本族语者或二语习得者为研究对象, 缺乏对英汉程式语心理表征模式的对比研究, 因此, 本书的研究就显得十分必要和及时。本书运用多种心理语言学实验方法研究英汉程式语的心理表征模式, 通过对比研究, 可以在一定程度上揭示英语本族语、汉语本族语和外语学习者程式语的意识水平、习得程度及其语言加工策略的选择模式, 有助于揭示英汉程式语心理表征差异的深层根源, 其研究结果可以为相关理论研究、外语学习和外语教学提供很有价值的参考依据。英汉程式语的心理表征模式既有相同点, 也有不同点; 程式语的频率对英汉程式语心理表征模式会产生一定的影响; 不同水平学习者程式语心理表征模式的不同并不一定由他们的语言水平来决定, 而是可能受多种因素影响。

本书重点研究英汉程式语心理表征模式的相同点和不同点, 分析其存在的原因。由于研究学习者的不同语言的心理表征是个十分复杂的心理过程, 通常被称为对“黑匣子”的研究, 因此, 需要谨慎设计和实施多种实验, 并避免各种可能因素的干扰。

基于以上研究的目的, 本书力图回答以下四个研究问题:

(1) 英汉程式语的心理表征模式是否存在差异? 如果是, 存在什么样的差异?

(2) 不同英语水平的中国英语学习者程式语的心理表征模式与英语本族语学习者程式语心理表征模式有什么不同?

(3) 程式语使用的频率是否会对英汉程式语心理表征模式产生影响? 如果是, 有何影响?

(4) 中国英语学习者不同的英语水平是否影响程式语在其心理词库中的表征模式? 其表征模式有何不同?

为了回答以上四个研究问题, 在对国内外相关研究进行认真翔实分析研究的基础上, 笔者开展了八个实证研究(见第7章至第10章), 这八个实证研究回答了这四个问题。

1.2 主要概念

人类语言交际在很大程度上使用固定表达, 这些固定表达经常以程式语形式被人记忆, 程式语是语言表征的基本水平, 其形式和意义具有可靠的关联(方秀才, 2013); 本书主要涉及三个最主要的概念, 即程式语、心理表征和心理词汇, 本节简单讨论这三个主要概念。

1.2.1 程式语

1. 程式语的术语

目前, 国内外对程式语的定义不统一。国外著名学者 Wray (2002: 9) 总结了如下 57 个与程式语相关的术语: Amalgams、automatic、chunks、Clichés、co-ordinate constructions、collocations、complex lexemes、composites、conventionalized forms、F[ixed] E[xpressions] including I[dioms]、fixed expressions、formulaic language、formulaic speech、formulas/formulae、fossilized forms、frozen metaphors、frozen phrases、gambits、gestalt、holistic、holophrases、idiomatic、idioms、irregular、lexical simplex、Lexical(ized)

phrases、lexicalized sentence stems、listemes、multiword items/units、multiword lexical phenomena、noncompositional、noncomputational、nonproductive、nonpropositional、petrifications、phrasemes、praxons、preassembled speech、precoded conventionalized routines、prefabricated routines and patterns、ready-made expressions、ready-made utterances、recurring utterances、rote、routine formulae、schemata、sentence builders、set phrases、semipreconstructed phrases that constitute single choices、stereotyped phrases、stable and familiar expressions with specialized subsenses、stereotypes、stock utterances、synthetic、unanalyzed chunks of speech、unanalyzed multiword chunks、units。

较有代表性的程式语术语有：formulaic sequences (Wray & Perkins, 2000)、formulaic language (Wray, 2002, 2004, 2008)、lexical bundles (Biber & Conrad, 1999)、lexical phrases (Nattinger & DeCarrico, 1992)、prefabricated chunks (Becker, 1975)等。之所以有如此多的术语，主要是因为研究者从不同的视角对程式语进行了研究(见表 1.1)。

表 1.1 程式语研究中相关的术语、定义及其角度*

研究者与年份	术语	定义	角度
Jespersen (1924)	formulas	重音、节奏和单词都不能改变的且作为整体习得的固定表达(转引自 Hickey, 1993: 28)	语言形式+语言习得
Lyons (1968)	ready-made utterances	本族语者以不可分析的整体形式学得并用于特定场合的表达(p.177)	语言习得+语言使用
Hakuta (1974)	prefabricated patterns	与某种可移动的成分,如可填入名词短语或动词短语(的空槽)一起运作的句子片段(pp.288-289)	语言形式
Wong-Fillmore (1976)	formulaic speech	那种对说话者来说完全或部分地作为不可分析的、固定的或自动的语言单位(p.295)	语言心理
Krashen & Scarcella (1978)	routines and patterns	routines: 整体记忆的话语或短语; patterns: 部分是创造的且部分是整体记忆的(p.283)	语言心理
Pawley & Syder (1983)	memorized sequences and lexicalized sentence stems	memorized sequences: 说话者或听话者能够有意识地组装或分析但在多数使用场合是作为整体或自动链接的形式而调用的单词串(p.205); lexicalized sentence stems: 与小句长度相当或更长的单位,其语法形式和词汇内容在很大程度上或完全是固定的;其固定成分形成了一种指称某种被该文化所认可的概念的标签,成为该语言中的一个术语(p.191)	语言形式+语言心理

续表

研究者与年份	术语	定义	角度
Peters (1983)	speech formula	一种多词素短语或句子, 它要么通过社会协商 (social negotiation) 要么通过个体演变 (individual variation) 成为说话者词库中的一种整体的预制项目 (p.2)	语言形式+语言心理
Nattinger & DeCarrico (1992)	lexical phrases	形式与功能的复合体, 一种介于传统的词汇与句法两极之间的词汇—语法单位; 它们被视作某种整体, 因此与词汇相似, 但多数词汇短语由一个以上的单词构成, 且许多词汇短语能同时由句法规则派生而来, 正如其他句子一样。它们的使用受到语用能力原则的制约, 后者对它们进行选择并赋予它们特定的功能 (p.36)	语言形式+语言功能
Moon (1997)	multi-word items	一种由两个或更多的单词 (一个单词只是一种正字法单位) 组成……并形成的一种有意义且不可分割的词汇项目 (p.43)	语言形式+意义
Biber & Conrad (1999)	lexical bundles	由三个或更多的单词组成的、出现频率很高的序列, 这些单词具有统计上同现的趋势 (p.183)	语言形式+使用频率
Wray & Perkins (2000)	formulaic sequences	一个由单词或其他意义成分组成的连续或非连续的序列, 该序列是或看起来是预制的, 即在使用时是从记忆中整体提取或存储的而不是通过语言的语法生成或分析的 (p.1)	语言形式+语言心理
Wood (2002)	formulaic language	作为语块被产出和被记忆的而不是由单个词汇和规则生成的多词或多形式词串 (p.3)	语言形式+语言心理
Jiang & Nekrasova (2007)	formulaic sequences/formulas	以较高频率、作为短语或连贯的语义单位出现的多词表达 (p.433)	意义+使用频率
Wray (2008)	morpheme equivalent unit	一个词或词串。该词串可以是完整的, 也可以是含有空格的。该词或词串被当作一个语素一样来处理, (它的意义) 不依赖于其组成成分的组合。(p.12, 转引自徐泉, 2010: 76)	语言形式+语言心理

*见李更春, 2011: 22-24

国内相关研究对程式语的定义也各不相同, 使用的术语也多种多样, 其中包括公式(化)语(言)、程式(短)语、程式化序列、词块、(预制)语块、词汇短语、(词汇)组块、套语、多词单位、惯用语、固定用语、习语、复现组合、一个词/词以上的词串、言语程式、词的联合、词汇化句干等。每一位研究者都是根据自己具体的研究目的, 确定程式语的具体定义。

程式语的词汇单元超过一个单词的长度, 坚持这种初步特征是合理的。尽管有关频率和公式性的确切关系的看法不尽相同, 但近年来, 有关公式化语言的研究大幅度增加, 特别是计算机技术的发展使有效地搜索语

料库中重复出现的多字字符串成为可能 (Wray, 2008)。

Wray (2008: 8) 对程式语的本质提出了三条理论观点, 第一个观点是心理词库是异常形态的 (heteromorphic), 这一观点与人们如何加工程式语有关。第二个观点是词汇的内容是通过唯一需求分析 (needs only analysis) 决定的, 这和人们如何学习程式语相关。第三个观点是语素的等效单元 (morpheme equivalent unit) 能让说话者控制听话者, 这和人们用程序语干什么有关。

Wray (2008: 8) 在给程式语下定义时, 把程式语比喻为一个连续体, 这非常形象。连续统的一端是争论不休的定义, Wray 把这些争议归为三大类; 另一端是没有争议的习语 (idioms), 如 kick the bucket、rain cats and dogs。在这两个极端中间是其他的程式语形式。例如, ①同通常的意思完全背离的词语或 (和) 语法规则, 如 “By and large”; ②表达具有比喻意义, 但是比起习语, 其与字面意思相近些, 如 “set one’s store by”; ③一个或多个词有特殊的意思, 如 “do someone in”; ④一系列的惯用语, 如 “you’d better V” “not very well”; ⑤“半开放模型/结构”指的是一个完整短语或者句子, 一部分是固定的词语结构, 一部分是开放型的结构, 如 “the end of the N” “as result of NP” “the way in which”。

很多程式语的解释都把程式性规定在只有指示性符号的词语中, 这些词语包括单个单词组成的词语、由连字符组成的单词、作为整体发音和作为部分发音不同的词组及意思和单个单词的联系不大的词组。但是, 另一种观点认为, 各个元素项目首先具有程式性, 之后才有指示性的符号。

Wray (2008: 12) 对“词素对等单位”的定义如下: “一个单词或一个字符串, 无论是不完整的, 还是能够插入可变元素的, 只要它能如一个词素般加工, 不需依靠其他任何的成分来完成形式与意义的匹配, 就是词素对等单位。”该定义反映出词素对等单位的属性, 并暗指它在语言中具有惯用语材料的起源性。

Wray 和 Perkins (2000: 1) 首次提出 “formulaic sequences” 这一术语, 并将之定义为: “一个由单词或其他意义成分组成的连续或非连续的序列, 该序列是或看起来是预制的, 即在使用时是从记忆中整体提取或存

储的而不是通过语言的语法生成或分析的。”此后，Wray (2002: 9) 对上述定义进行了修正，将“意义”(meaning)一词从该定义中删去。该定义是目前为止涵盖面最广的定义之一，它既包括那种高度习用的且不可改变的单词串(如 by and large)，又包括那种语义上透明、句法上灵活且含有可填入开放词类的空槽的单词串(如 NP be-TENSE sorry to keep-TENSE you waiting) (Wray & Perkins, 2000: 1)。该定义从心理语言学角度指明了程式语的本质特征，即“整体提取或存储”。也就是说，某语言形式是否为程式语既不取决于其形式或意义的完整性，也不取决于其出现的频率，而是取决于语言使用者是否以整体形式对其加以对待。Wray (2002, 2008) 就用“formulaic language”替代“formulaic sequences”。从现有的相关研究成果来看，目前，我国进行相关研究的学者更多采用“程式语”这个术语。另外，程式语覆盖面更广，更能代表类似的相关研究。因此，本书采用“程式语”(formulaic language)这一术语。

从表 1.1 可以看出，学者们尝试从不同角度对程式语加以描述或界定。Wray (2002) 对程式语的定义可操作性不强，且没有考虑程式语的频率、语义或功能因素；Biber 等对词束的界定虽然可操作性较强，但稍显主观且无法排除无意义的单词组合；Becker (1975) 对预制语块的定义显然只是印象性的，无助于程式语的判别。总之，由于程式语研究起步较晚，加之研究者的背景各异，到目前为止，该领域中的术语和定义问题还没有得到彻底解决。笔者认为，程式语难以定义的原因首先在于其多样性，它们在长度、使用目的、固定性与整体存储的程度等方面存在差异。就长度来说，它们可长可短，长的如“You can lead a horse to water but you can't make him drink.”，短的如“Oh no!”，也可以介于两者之间。就使用目的来说，它们通常被用来实现不同的目的。例如，①表达一种语义信息或观点(如 Blood is thicker than water.)；②实现一种语用功能(如 Just looking.)；③以某种确切的、可理解的方式传达具体的信息。在固定性方面，这些序列可以是完全固定的(如 King and Queen)，或具有若干空槽，其中可填入合适的单词或单词串(如 ___ made it plain that ___)。在整体存储的程度上，有些程式语可能只以整体形式存储，有些程式语可能

具有双重存储模式（即整体存储+单个的词和句法），有些程式语的存储模式还可能会因人而异。其次在于程式语是一种多面性现象（a multi-faceted phenomenon），不能仅从少数几个方面而应该从形式、意义、功能、心理表征和社会文化特性等方面对其进行全面的描述与界定。在综合国内外已有定义的基础上，笔者将程式语重新定义为：某言语社团以整体形式辨识、存储和提取，连续或非连续、具有较完整的结构与意义/功能或参与语篇构建、使用频率较高的多词单位。该定义既涵盖了传统语法所认可的程式化表达（如习语、谚语、短语/词组以及其他类别的固定表达），又涵盖了通过语料库驱动方法识别的多词单位或词束。

2. 程式语的分类

从已有的相关文献来看，研究者从不同的角度对语言中的程式性现象进行了分类：第一类基于语言的形式，第二类基于语言的意义，第三类则基于语言的功能。这些不同的分类从一个侧面反映了程式性现象的多面性：它是一种集语法、词汇、功能和语境为一身的特殊的、复杂的语言现象。由于研究者的视角不同，分类的数量就存在很大差异，下面是国内外学者们从不同视角对程式语进行的分类。

二分法。Cowie（1988）根据语言功能将程式语分为两类：一类是具有语篇和社会功能的公式语；另一类是“复合体”（composites），即“作为句子中的成分（如宾语、补语或附加语等）并对其指称或命题意义具有贡献的词汇构件”。

三分法。国内外一些专家基于语言意义对程式语进行了分类。例如，Gibbs 和 Nayak（1989：133-134）依据语义可分解程度，将习语分为三类：可分解习语（如 pop the question、break the ice）、不正常可分解习语（如 carry a torch、spill the beans）和不可分解习语（如 hew the fat、shoot the breeze）。Cacciari 和 Glucksberg（1991，1994）根据语义透明度，将习语分为非透明的、透明的与准隐喻的。Fernando（1996）也依据意义标准将习语分为纯习语（pure idiom）、半习语和字面习语（literal idiom）。