

语形视角下的《现汉》非四字成语探析

——兼论成语的原型特征

张梁*

目 录

1. 前言
2. 《现汉》中成语的定义与数据统计
3. 非四字成语的形式结构特征
 - 1) 双体成语
 - 2) 单体成语
4. 非四字成语的语义特征
5. 非四字成语的语体特征
6. 结语

中文摘要

为探讨成语这一概念的原型特征，并区分非典型成语与俗语、惯用语等概念的范畴，本文以《现代汉语词典》第7版认定的98个非四字成语为研究对象，对成语语形进行了详细探讨。在非四字成语中以双四字格的语形占优势，其次为五音节和六音节单体语形。双体成语中对称结构占主导地位，不仅是字数上的对称，结构和语义上也呈现相应明显对称状态。单体非四字成语的注音形式更接近于句子而非“语汇”，其中五音节成语多有相对应的四字形态。在语义方面，具有双层性的非四字成语比例明显高于四字成语。因此，本文认为“成语”的原型特征除了最重要的“四字形式”之外，结构对称性、语义双层性、语体书面化、来源典故性等这几方面也是成语在不同方面的原型特征。

* 강원대학교 중어중문학과 강사

关键词：现代汉语词典、成语、语形、俗语、熟语、原型

1. 前言

成语是汉语中较为特殊的一类词汇单位，也是汉语词汇学中的一个热点。前人学者已经从语音、结构、语义等各个方面对成语进行了大量研究。比较有代表性的专著类研究成果有马国凡（1978）的《成语》，向光忠（1982）的《成语概说》，刘洁修（1985）的《成语》，倪宝元·姚鹏慈（1990）的《成语九章》，莫彭龄（2001）的《汉语成语与汉文化》等等。各大出版社的多种成语词典也在一直不断更新再版，这也从侧面证明了成语在汉语词汇中较为特殊的地位。此外，其他学者如王勤·武占坤（1959）、刘叔新（1982）、温端政（2005）、周荐（2014）等学者也都在各自的词汇学著作中谈及成语的定义与各种特点。其中“四字”可谓是成语在形式上最为显赫的、有别于其他词汇单位的一个特征，这是各家学者的一个共识，如上海辞书出版社（2007）的新一版《中国成语大辞典（辞海版）》中就只收录四字成语，将非四字成语排除在外。因此，四字格或四音节的成语也可以称之为“典型”成语，但多数学者也并不断言“四字”就是汉语“成语”的必要条件。

不可否认的是，在大多数成语辞典中，或多或少都收录有一些非四字的条目，如商务印书馆2003年出版的《汉语成语小词典》和2013年出版的《新华成语大词典》，华语教学出版社（2023）《中华成语大词典》等。包括汉语学界的权威辞书《现代汉语词典》（以下简称《现汉》，本文特指2016年出版的第7版）也将不少非四字条目认定为“成语”。相对来说，海内外学者对非四字成语的研究成果就较为薄弱。近些年来有任会芹（2011）、王德毅·施向东（2016）、石清洋（2018）、王利·曹瑶（2021）等学者对这些非四字条目进行研究，但多集中于硕士论文和期刊论文，鲜见博士论文与专著。而且各家学者对这类成语的指称熟语未有统一（有学者称“非典型成语”，本文统称为“非四字成语”），可见相

对于较为典型的四字成语来说，学者们在这方面的研究成果都尚未完善，目前尚难以区分与“俗语”等其他熟语单位类型的界限。

本文将立足于最新的《现汉》，对其中收录的“成语”条目进行穷尽式统计，并以同为商务印书馆出版的《成语大词典》（2020版）为参照，利用认知语言学中的原型范畴理论，将研究重点放在处于边缘地带的非四字成语上，试图为“成语”这一概念梳理出一些原型特征，希望能为后继研究者提供些许参考。

2. 《现汉》中成语的定义与数据统计

本文将《现汉》电子融媒体版本¹⁾中被认定为“成语”的全部4342个条目进行了汇总统计，其中四字条目有4244个，占总量的97.74%。非四字条目共98个，仅占2.26%。按字数与音节统计，本文将98个“非四字成语”已穷尽式列出，其中需要略作说明的是词条“一时半会儿”中“儿”为儿化韵，《现汉》将其注音为四音节²⁾“yīshí-bànhuìr”，因此本文列入非四字成语中的“其他”单位，与其余4244个四字成语不作一类处理。由此，《现汉》中4342个成语分布本文整理如表-1所示：

〈表-1〉《现汉》中的成语数量

音节数	数量	条目
三音节	3	吹法螺、莫须有、破天荒；
四音节	4244	略；
五音节	20	病急乱投医、不期然而然、大旱望云霓、大鱼吃小鱼、何乐而不为、疾风知劲草、庐山真面目、鸣鼓而攻之、蚍蜉撼大树、十八般武艺、十八层地狱、十万八千里、铁杵磨成针、图穷匕首见、无所措手足、薰莸不同器、一鼻孔出气、一言以蔽之、欲速则不达、坐山观虎斗；
六音节	19	百闻不如一见、毕其功于一役、不知天高地厚、反其道而行之、风马牛不相及、过屠门而大嚼、化干戈为玉帛、九牛二虎之力、

1) 即2019年上线的APP，本文利用的是iOS系统2.0.34版。

2) 四字成语中有“正儿八经”一条目的“儿”与之情况类似，但《现汉》注音为“zhèng·er·bājīng”，并未写作儿化韵形式，因此依旧视作四字成语处理。

		可望而不可即、杀鸡焉用牛刀、神不知鬼不觉、生米煮成熟饭、水至清则无鱼、万变不离其宗、无所不用其极、五十步笑百步、勿谓言之不预、迅雷不及掩耳、有一搭无一搭；
七音节	8	此地无银三百两、解铃还须系铃人、冒天下之大不韪、赔了夫人又折兵、牵一发而动全身、行百里者半九十、一失足成千古恨、长江后浪推前浪；
双四字	42	百尺竿头，更进一步、百足之虫，死而不僵、兵来将挡，水来土掩、不入虎穴，焉得虎子、差之毫厘，谬以千里、成事不足，败事有余、成也萧何，败也萧何、城门失火，殃及池鱼、道高一尺，魔高一丈、放下屠刀，立地成佛、己所不欲，勿施于人、近朱者赤，近墨者黑、流水不腐，户枢不蠹、明修栈道，暗度陈仓、皮之不存，毛将焉附、千里之堤，溃于蚁穴、千里之行，始于足下、前事不忘，后事之师、庆父不死，鲁难未已、取之不尽，用之不竭、人同此心，心同此理、人为刀俎，我为鱼肉、仁者见仁，智者见智、如临深渊，如履薄冰、三天打鱼，两天晒网、失之东隅，收之桑榆、失之毫厘，谬以千里、螳螂捕蝉，黄雀在后、桃李不言，下自成蹊、万事俱备，只欠东风、为渊驱鱼，为丛驱雀、无源之水，无本之木、项庄舞剑，意在沛公、星星之火，可以燎原、言者无罪，闻者足戒、一人得道，鸡犬升天、一言既出，驷马难追、以其昏昏，使人昭昭、以眼还眼，以牙还牙、以子之矛，攻子之盾、欲加之罪，何患无辞、鹬蚌相争，渔人得利；
双三字	4	彼一时，此一时、吃一堑，长一智、此一时，彼一时、满招损，谦受益；
其他	2	一时半会儿、冰冻三尺，非一日之寒

从上表-1的“词长”统计可以看出，《现汉》中的“非四字成语”中以“双四字格”为最主要的形式，达42个，是第二名五音节形式的两倍有余。其次是五音节（20个）与六音节形式（19个），二者数量相差无几。前三者合计占非典型成语总量的83%左右。再次是七音节形式（8个），除此之外的其余所有形式合计仅有9个条目。

还需值得注意的是非典型成语中的“双体现象”，除了最主要的“双四字格”，还有4个“双三字格”和1个不对称双体成语，即“冰冻三尺，非一日之寒”。此外六音节成语中还有几个似乎也可以视为“双三字格”，如“神不知鬼不觉”和“有一搭无一搭”。与此相对，“双三字格”中的条目似乎也可以“并合”视为六音节条目，因此有必要从成语的结构和语义等方面作进一步分析。

总之，从表-1可知，在“非典型成语”的范畴中，各词长形式并非是平均分

布的，这意味着“成语”这一概念有其自身的原型特征。比如《现汉》本身对“成语”一词的定义为：

【成语】人们长期以来习用的、简洁精辟的定型词组或短句。汉语的成语大多由四个字组成，一般都有出处。有些成语从字面上不难理解，如“小题大做、后来居上”等。有些成语必须知道来源或典故才能懂得意思，如“朝三暮四、杯弓蛇影”等。³⁾

可见《现汉》对“成语”的特点主要提到四个方面：(1) 来源特点：长期习用、有出处；(2) 结构特点：定型词组；(3) 形式特点：四个字；(4) 语义特点：有双层性，也可只有字面义。但是关于成语的句法组织、语音韵律、语体语用等特征却不曾言及。

不过各家学者在研究以四字形式为代表的成语时，认为成语还有以下多个特征，如语义结构的对称性（刘振前1998）、音韵结构的对称性（刘振前·邢梅萍2003）、结构类型中并列结构占主导（洪波2003）、古雅的语义色彩（周荐2014）等等。那么在脱离了“四字”这个最大的语形特点之后，处于成语范畴边缘地带的各类非四字成语是否依然具有以上这些特征就是本文主要的研究目的。特别是以《现汉》为代表的各类规范型词典，对于语言学类的词条，不止要看它如何定义，更要看它的定义是否符合其实际的收录情况。

3. 非四字成语的形式结构特征

从上文中表-1中可见非四字成语的形式非常多样化，因此无法完全像四字成语一样从“并列、主谓、动宾、述补、偏正”等基本结构特征入手。而且五种基本结构是建立在双音节词汇基础之上的，用于分析四语素的成语本就有些掣肘，如“问道于盲”“不期而遇”等成语就很难套用到这几种基本结构中去，更何况

3) 摘自《现汉》第7版，166页。

是语素更多的非四字成语。因此本文采用就简原则，以注音形式为主来研究非四字成语的形式特征，这主要适用于“双四字格”。原因在于《现汉》中对四字成语的注音成语的形式有着非常严格的标记原则，四字成分可以前后二二两分的成语基本以“2-2”标记⁴⁾。而“问道于盲”“不期而遇”等难以两分的成语注音则连写为“4”式。这两种注音形式在4244个四字成语中占97.6%左右⁵⁾。

1) 双体成语

已知《现汉》非四字成语中以“双四字格”为最主要形式，其数量占据全部非四字成语的43%左右。本文就可以依靠注音形式来研究双四字格内部的形式结构特征，并以此参照其他非四字成语的形式特点。

“双四字格”成语等“双体成语”的最显著特征便是以逗号区分前后两“节”，以数量较少的4个“双三字格”成语来看，“彼一时，此一时”“吃一堑，长一智”“此一时，彼一时”“满招损，谦受益”无一例外都是“111, 111”的注音形式，且前后“节”长度一致，结构相同，这就类似于四字成语中的并列结构，如“花红-柳绿”“碍手-碍脚”等。

但42个“双四字格”成语的情况相对较为复杂。本文按照前后节的注音形式是否一致分列如下表-2。

〈表-2〉《现汉》中的双四字格成语注音形式

结构分类	注音形式	条目
前后一致 (35)	112, 112	以其昏昏，使人昭昭
	4, 4	百足之虫，死而不僵、己所不欲，勿施于人、千里之堤，溃于蚁穴、千里之行，始于足下；(不对称) 成也萧何，败也萧何、道高一尺，魔高一丈、近朱者赤，近墨者黑、人同此心，心同此理、人为刀俎，我为鱼肉、皮之不存，毛将焉附、

4) 如果是词汇化程度不高的条目则一般以“2 2”形式标记。

5) 据本文对四字成语注音形式的初步统计，这两种注音形式分别有3468个、674个，合计4142个，第三位的“2 2”形式仅有70个，其余形式合计也仅有32个。

		如临深渊, 如履薄冰、无源之水, 无本之木、以子之矛, 攻子之盾; 城门失火, 殃及池鱼、放下屠刀, 立地成佛、万事俱备, 只欠东风、 项庄舞剑, 意在沛公;(不对称)
	2-2, 2-2	兵来将挡, 水来土掩、流水不腐, 户枢不蠹、明修栈道, 暗度陈仓、 庆父不死, 鲁难未已、取之不尽, 用之不竭、仁者见仁, 智者见智、 三天打鱼, 两天晒网、失之东隅, 收之桑榆、为渊驱鱼, 为丛驱雀、 以眼还眼, 以牙还牙、螳螂捕蝉, 黄雀在后、言者无罪, 闻者足戒、 一人得道, 鸡犬升天、一言既出, 驷马难追、鹬蚌相争, 渔人得利、 不入虎穴, 焉得虎子、成事不足, 败事有余;
前后不一 (7)	2-2, 4	百尺竿头, 更进一步、桃李不言, 下自成蹊;(不对称) 差之毫厘, 谬以千里、前事不忘, 后事之师、失之毫厘, 谬以千里;
	4, 2-2	星星之火, 可以燎原、欲加之罪, 何患无辞;(不对称)

由表-2可见, 大部分双四字格成语的前后节注音形式是一致的, 共有35个, 占有所有双四字格的83%。其中与四字类似的是单节注音形式, 其中“2-2”形式占绝大多数, 其次是连写的“4”式。即使是前后节注音形式不一致的7个词条中, 也是“4”与“2-2”形式的不对应。此外, 仅有“以其昏昏, 使其昭昭”的单节注音不是连写形式, 占总数的2.4%左右, 这与四字成语中的其他注音形式占比极其接近(见脚注5), 可能不仅仅是简单的巧合。

有了明确的注音形式, 就可以借助其来进一步探讨结构方面的对称性。注音形式的一致性可以作为结构对称性的参照, 但两者并非完全对应。从上表-2的条目来看, 双四字格的结构对称性可以大致分为“对称”和“不对称”两种, 其中“不对称”的双四字格已在表中标示。但在“对称”形式中还可以按照语素的词性、虚字的位置等要素是否完全对应分为“严式对称”和“宽式对称”。这与刘振前·邢梅萍(2000)文中对四字格成语按照语义结构的分类方法非常类似。

比如“严式对称”除了上文中提到的“以其昏昏, 使其昭昭”之外, “4, 4”形式中有“成也萧何, 败也萧何、道高一尺, 魔高一丈、近朱者赤, 近墨者黑、人同此心, 心同此理、人为刀俎, 我为鱼肉、如临深渊, 如履薄冰、无源之水, 无本之木、以子之矛, 攻子之盾”8个。“2-2, 2-2”形式中有“兵来将挡, 水来土掩、流水不腐, 户枢不蠹、明修栈道, 暗度陈仓、庆父不死, 鲁难未已、取之不尽, 用之不竭、仁者见仁, 智者见智、三天打鱼, 两天晒网、失之东隅, 收

之桑榆、为渊驱鱼，为丛驱雀、以眼还眼，以牙还牙”至少10个。

“宽式对称”也不在少数，如“皮之不存，毛将焉附、成事不足，败事有余、言者无罪，闻者足戒、一人得道，鸡犬升天、不入虎穴，焉得虎子、螳螂捕蝉，黄雀在后、一言既出，驷马难追、鹬蚌相争，渔人得利”等8个。此外值得注意的是，在前后节注音形式不一致的双四字格中也存在宽式对称，如“差之毫厘，谬以千里、前事不忘，后事之师、失之毫厘，谬以千里”3个。这主要是由于虚语素在四字格中的位置所致：“以”在四字格第二字时一般四字格注音需要全部连写；“之”字在第二字时常作为指示代词可分写（如“差之毫厘”）、但在第三字时一般为助词，则需要连写（如“后事之师”）。但从词性角度来说，用“以”对“之”，以虚对虚，视为对称并无不妥。

如此算来，不对称的双四字格仅12个左右，其余30个“对称”式双四字格都可以视为广义上的“并列结构”，占71%左右的绝对优势。周荐（2014）在讨论双四字格时也早已发现对称关系占主导地位的现象⁶⁾。这与四字成语中并列结构的强势地位不谋而合⁷⁾。

由此可见，双四字格成语与四字成语相比，除了字数（长度）翻倍以外，在内部结构、前后关系等方面有着极大的相似性，可谓是“双倍版”的变体成语。而双三字格成语在形式的“对称性”方面也保持着较为明显的成语特色。最后“冰冻三尺，非一日之寒”的结构更接近于“主谓关系”，主谓关系在周荐（2014）看来，于双四字格中也属于不对称关系。由于不对称双体成语数量仅有1条，可视为孤例，在此不作展开讨论。

2) 单体成语

本节来看一下单体成语（为行文简洁，本节所述的“单体成语”不包括典型

6) 据周荐对《成语词典》（1983）和《中国成语大辞典》（1987）的调查显示：两本辞书共88个双四字格单位中，有82个存在联合的对称关系。（见《汉语词汇结构论》298页）

7) 如张梁（2023）在其统计的186个现代成语中，并列结构以107个占据首位。（见《现代汉语四字格韵律构词法研究》122页，表-98）

的四字成语)的形式特点。单体成语在没有符号标记切分前后节的情况下,与四字成语的差别看起来就是由字数多寡造成的“语形”长度差异。但如果进一步分析就会发现并非如此,就注音形式来看可一目了然。本文将51条非四字单体成语的注音形式按“语形”长度罗列如下表-3所示:

(表-3)非四字单体成语的注音形式

音节数	注音种类	注音形式(条目数量)
五音节	6种(20条)	11111(9);1112(1);1121(1); 122(1);212(6);221(2);
六音节	6种(19条)	111111(12);11112(1);12111(1); 1212(2);21111(2);222(1);
七音节	8种(8条)	1111111(1);121111(1);2212(1);21121(1); 111112(1);12112(1);12121(1);11212(1);
其他	3种(4条)	2-2(1);12(2);3(1)

由表-3可见,四音节以上的单体成语有一个最显著的特征,即没有任何连写方式⁸⁾。这也就意味着虽有47个单体条目被以“成语”身份收录在《现汉》之中,但也未被完全承认其“词化”后“词”的身份,与大量“2-2”式注音和完全连写的“4”式注音成语依然有着巨大的鸿沟。可见“词长”因素对一个“语汇”单位是否为“词”或“语”依然有着重要的影响力。

如果纯粹从注音形式来看,这些成语都是按照“一词一空”的方式注音,甚至就五音节和六音节成语而言,“一字一空”成了数量最多的注音形式。从注音形式的种类来看,字数越多,注音形式的数量也就越多,以七音节成语来看,更是每个条目都有自己的注音形式,与一般句子的注音方式来说并无二致。说明这些单体成语中,并没有固定或明显占优势的结构形式关系。以七字成语为例,一级句法划分就可分为以下五种:

2/5:此地无银三百两、解铃还须系铃人;

1/6:冒天下之大不韪;

8) “一时半会儿(2-2)”和“莫须有(3)”虽为连写,但属于四音节及以下条目。

4/3：賠了夫人又折兵、行百里者半九十、长江后浪推前浪；
 3/4：一失足成千古恨；
 3/1/3：牵一发而动全身；

在这种形式差异下，这些单体成语为何能被《现汉》认为是“成语”？他们在形式上是否与四字成语越接近就越容易被认可为是“成语”？本文认为有必要对与四字成语最接近的五字成语做进一步分析。相较于其他成语来说，20个五音节成语中大多都有其相对应的四字成语形式，而且多被成语词典或网络辞书收录其中，本文将其对应形式整理如下表-4所示：

〈表-4〉五音节形式与其四字成语对应形式

	五音节条目	四字成语	四字形式收录情况 ⁹⁾
1	不期然而然	不期而然	《现汉》110页
2	大旱望云霓	大旱云霓/大旱望云	百度教育-汉语
3	何乐而不为	何乐不为	《成语大词典》454页
4	疾风知劲草	疾风劲草	《现汉》610页
5	庐山真面目	庐山真面	《现汉》846页
6	鸣鼓而攻之	鸣鼓而攻/鸣鼓攻之	《成语大词典》739页
7	蚍蜉撼大树	蚍蜉撼树	《成语大词典》790页
8	铁杵磨成针	铁杵成针/磨杵成针	百度教育-汉语
9	图穷匕首见	图穷匕见	《现汉》1325页
10	无所措手足	手足无措	《现汉》1204页
11	薰莸不同器	薰莸异器	《现汉》1492页
12	一言以蔽之	一言蔽之/一言以蔽	《成语大词典》1301页
13	欲速则不达	欲速不达	《成语大词典》1371页

以上成语中还有的条目不止一个四字对应形式，如“铁杵磨成针”——“铁杵成针/磨杵成针”，说明其“成语化”¹⁰⁾倾向明显。除了以上13个词条外，“坐山观虎斗”和“病急乱投医”两个条目也能看见四字形式的实际语料，以下两则均选自北京大学CCL语料库：

9) 四字形式有两种以上的情况，收录出处标记为前者。其中网络辞书“百度教育-汉语”的检索日期为2025年9月10日。

10) 亦有学者称此过程为“雅化”或“四字化”等。

哈哈，开夏老师坐山观虎呵，赶紧孵个争端解决器吧

——2010s\201x\网络语料\微博网友

原因在哪儿？一是示威，担心别人抢了江山，通过挖祖坟，打击对手的政治信心，如武则天的掘坟焚尸；再就是病急乱医，大厦将倾，江山将失，也只得通过死马当活马医了。

——2000s\2009\倪方六《中国人盗墓史》

如果包含以上这两个条目，那么20个五音节条目中有75%都可以变化为四字形式。相较于六音节和七音节成语，五音节成语由于字数与四字成语的“语形”长度最接近，确实更加接近“成语”这一概念的典型形式特征。由此可见，把五音节成语视为“加字版”的变体成语也未尝不可。但是由于五字的单数“词长”，必然不易形成前后对称的并列结构，因此导致的五音节成语认可度不高，也是客观上不争的事实。

本章的分析显示，双四字格在结构上最接近于典型成语的并列特征，而五音节成语则在“词长”上更接近于典型成语的形式特征。因此相较于六音节（单体）和七音节成语而言，这两者数量更多，显然也更接近于“成语”原型范畴的语形特征。那么其他形式的条目又是缘何被《现汉》认定为是属于成语的范畴呢？比如三音节条目的词长与典型的四字成语也是只差一个音节，但在数量上却远逊于五音节条目。对于这些问题，下文将试图从语义和语体等角度继续分析非四字成语的特点。

4. 非四字成语的语义特征

在典型的四字成语中，多有学者提及成语作为熟语范畴下位成员的语义双层性。但如《现汉》自己对成语的定义与说明，语义双层性相较于形式结构等方面并不是成语的优势特征，仅从字面意义可以理解的成语也不在少数。根据周荐（2014）对《中国成语大辞典》的统计，有双层义和无双层义的“四字语”数

量“难分轩轻”，而相比较归于熟语的“三字语”则是基本是都具有双层义¹¹⁾。

那么《现汉》中非四字成语的语义双层性又是何种情况？本文通过《现汉》词条释义中的双层义提示词可以认为该词条具有明确的双层义，如【百尺竿头，更进一步】“比喻学问、成绩等达到了很高的程度以后仍继续努力。”其中“比喻”就是一个非常典型的双层义提示词，与词条的字面义差距甚远。除了“比喻”以外，“指”“形容”“表示”虽然侧重点各有不同，但都可以视作这一类双层义提示词。本文将释义中含有这些关键词的条目数据统计如下表-5所示：

〈表-5〉非四字成语释义中的提示词统计

提示词	数量	备注(参看主条的条目)
比喻	49	2 (解铃还须系铃人；杀鸡焉用牛刀)
指	21	2 (此一时，彼一时；仁者见仁，智者见智)
形容	10	-
表示	6	1 (有一搭无一搭)
/	13	-

从上表可见，在98个¹²⁾非四字成语中，不含任何提示词的有的仅有13个。而含有提示词的条目总数占比为86.7%左右。其中有5个条目虽然作为副条出现在《现汉》中，但参看其主条的释义也含有提示词，本章也算入在内。

相较与周荐(2014)所言有双层义和无双层义的四字语数量“难分轩轻”，非四字成语86.7%的比例明显偏高。这与《现汉》所认定“熟语”中的数据相当接近，特别是五音节及以上熟语中双层义占90.4%的比例¹³⁾。

但是说到《现汉》所认定的“熟语”和“成语”，两者在来源上有着较大区别。其中有明确历史出处的条目数据差异明显：非四字成语中明确标有来源出处的共有38个，占比接近40%。但《现汉》中全部451个“熟语”中仅有14个五音节以上条目有来源，即使范围只限定于五音节以上的148个“熟语”，其占比也不

11) 按《汉语惯用语词典》(1985)统计，有90%的都有比喻义。《汉语词汇结构论》322页。

12) 其中“百足之虫，死而不僵”(《现汉》29页)的释义出现“指”和“比喻”连个提示词，表中分别计入，因此表-5总数合计为99。

13) 据张梁(2025)〈汉语惯用语原型再议——以《现汉》第7版为中心〉中〈表-12〉数据

到10%。说明《现汉》对成语的收录确实与其定义中的“一般都有出处”相对应。如果没有出处，类似语形的条目则易被划分为其他类型的熟语。

标记有历史出处意味着该熟语拥有较长的习用历史，那么语义在使用过程中发生引申变化也是可以预见的结果。这和惯用语的“短命”特点¹⁴⁾完全相反，那么两者在双层义产生的来源上可能并不一致。还有不少人认为成为“惯用语”的条件之一就是其内容上的比喻性¹⁵⁾，因此其意义双层性可以说是“与生俱来”的。相对来说不是所有成语都具有“天生”的语义双层性。

由此可以推断：语义上的双层性和历史来源有出处这两个特点应也当是成语的特征之二，虽然语义双层性并不是所有成语天生自带的，但是随着长期习用而产生双层义的条目就更有可能被汉语母语者认可为成语。想必这也就是“成语”名称来源的原因之一。那么进一步推测：如果一个条目在语形上无法符合“四字”要求，但在结构对称性、语义双层性、来源有出处等方面能够满足条件，那么也是有可能被认定为成语。比较典型的如以下几个条目：

【莫须有】“宋朝奸臣秦桧诬陷岳飞谋反……后用来**表示**凭空捏造。”
(《现汉》922页)

【庐山真面目】“宋代苏轼《题西林壁》诗……后用‘庐山真面目’**借指**事物的真相或人的本来面目。”(《现汉》846页)

【五十步笑百步】“战国时候……(见于《孟子·梁惠王上》)**比喻**自己跟别人有同样的缺点或错误，只是程度上轻一些，却讥笑别人。”(《现汉》1389页)

【赔了夫人又折兵】“《三国演义》里说……后用来**比喻**想占便宜，没有占到便宜，反而遭受损失。”(《现汉》984页)

以上四个成语虽然在语形和结构对称性上没有符合成语的原型特征，但来源有出处，语义也有双层性，因此被《现汉》认定为成语。但从母语者的语感

14) 崔希亮(2024)“惯用语的时代色彩突出……流行的东西总是短命的。”——《汉语熟语与中国人文世界(修订本)》第204页。

15) 周荐(2014)“……不少人看来，一个单位要被认作是惯用语须满足两个条件：形式上的三字格，内容上的比喻性”——《汉语词汇结构论》第268页。

来说，说它们是成语多少有些勉强。

5. 非四字成语的语体特征

四字成语的书面语体特征已有多位学者进行过论述，其原因多认为语义典雅的单位更喜音步平稳的节律（冯胜利1997、周荐2014、朱赛萍2015等）。那么非四字成语在丧失平稳的节律后是否依然具有这样的语体特征就非常值得再次探讨。以第四章最后提到的“赔了夫人又折兵”为例，该条目既没有相对应的四字形式，也没有偶字数成语的平稳节律，其成语的资格必然会受到质疑。本章将试图通过《现汉》中的语体标记来看这些非四字成语的语体特征。

在98个非四字条目中，只有一个“一时半会儿”有几化音，也只有该词条具有“口语中也读yīshí-bànhuǐr”的语体性解释，而且其他所有词条都没有<口>和<方>这样的语体标记。此外，从注音标记来看，有轻声字的也只有“赔了夫人又折兵”一个条目。这说明绝大部分非四字成语确实不具有口语体的特征，这和《现汉》的“熟语”条目确实存在较大语体差异。据张梁（2025）<表-11>的统计，五音节及以上的熟语口语体标记占比为23.6%，大大高于非四字成语，如：

【八字没一撇】<口>比喻事情还没有眉目。（《现汉》17页）

【板板六十四】<方>形容不知变通或不能通融。（《现汉》34页）

但是与此同时必须正视的事实是：即使从语体角度看，《现汉》所划定的非四字“成语”和“熟语”的界限也并非完全清晰，最具代表性的当属“赔了夫人又折兵”，既属于4342个“成语”之一，又属于451个“熟语”之一。想必是该条目既符合熟语的口语化特征，又具有成语“语义双层性、来源有出处”的特点。此外比较有争议的还有以下几对例子，本文制表-6所示，并辅以《成语大词典》的收录情况做参照。

〈表-6〉《现汉》中形式相近的成语和熟语条目

《现汉》成语	《现汉》熟语	《成语大词典》收录与否
吃一堑，长一智	不经一事，不长一智	收录 (p.192) / 收录 (p.113)
道高一尺，魔高一丈	魔高一尺，道高一丈	收录 (p.272) / 无
杀鸡焉用牛刀	割鸡焉用牛刀	无 / 收录 (p.394)
生米煮成熟饭	生米做成熟饭	收录 (p.956) / 无
有一搭无一搭	有一搭没一搭	无 / 无

如果单纯从语素角度认为“无”“杀”“煮”比“没”“割”“做”更具书面语体特征，似乎略显勉强。从上表-6可见《成语大词典》收录“割鸡焉用牛刀”而非“杀鸡焉用牛刀”，也未收录“有一搭无一搭”为成语。但《现汉》确实有意识将“熟语”视为更具口语性的条目，用以作为“成语”的语体对立面，而非语言学界普遍将“熟语”视为“成语”“惯用语”“歇后语”等形式的上位概念。因此仅从《现汉》融媒体版划定的“成语”和“熟语”的界限和归属来看，并不能简单地从语体、语义上将所有词条区分清楚。

由此看来，最直观有效的方法似乎依旧是从“语形”角度出发，比如就“惯用语”来说，其原型特征中最重要的当属“语形”特征，即“词长”为三音节（苏向丽2008、周荐2014、崔希亮2024等）。那么相对来说“成语”这一概念如果说要有“语形”上的原型特征，那必然是四字格/四音节形式，因此“双四字格”和“五音节”作为与成语语形特征最接近的两种形式，其数量在非四字成语中占据前二也就不足为怪了。

那么就语形归属来说，最后的问题就落在一些五、六、七字的“语汇”单位上。这些单体条目似乎没有具体概念的归宿，这种略显尴尬的局面本文认为主要来自于以下几个原因：

1) 从“词汇”角度看，现代汉语词汇的主要词长（也可以说“词形”）以双音节和单音节为主，同时也有不少三音节与四音节词。但四音节以上的单位就很难在汉语母语者心里层面具备“词感”。以汉语外向型词表来看，一般都不收录五音节及以上词条，最具代表性的如《等级标准》¹⁶⁾和《HSK词汇大纲》¹⁷⁾。

16) 全称为《国际中文教育中文水平等级标准》（2021）北京语言大学出版社。

2) 从“语汇”角度看, 熟语作为上位概念, 在划分其下位的几个并列概念时, 较为公认的形式有成语、惯用语、俗语、谚语、歇后语五种¹⁸⁾。其中“成语”的四字形式、“惯用语”的三字形式、“谚语”的对举结构、“歇后语”的破折号和谜语双关, 都是直观的形式标准。唯独“俗语”的形式标准并不明晰。崔希亮 (2024) 在《汉语熟语与中国人文世界 (修订版)》一书中, 对熟语的分类介绍唯独缺少“俗语”, 由此可见一斑。

3) 如果把这些五、六、七音节的单体“语汇”单位姑且都称之为“俗语”, 那么从“俗语”本身形式的多样性看, 是否有必要再按五、六、七音节继续划分为三类? 从本文研究的非四字成语来看, 超过8个音节的就有被断为“双体”结构的趋势, 这显然与人类的短时记忆限度 (7±2) 有关¹⁹⁾。可以大胆预测, 一个“俗语”如果超过九个音节, 那么其连“语汇”的资格都将难以维持, 更不用说算是哪一类熟语了。

综上, 以非四字成语为主, 包括《现汉》“熟语”中的五音节及以上单位, 它们在“语形”上没有唯一共同特征, 在数量上也没有哪一类占绝对优势, 即使将其都称之为“俗语”, 这一概念似乎也很难找到统一的形式特征作为归属。在大部分条目都有双层语义的情况下, 如果硬要以语体特征来划归为非典型的“成语”或典型的“俗语”, 似乎有些牵强。

因此, 研究这些“语汇”单位, 还是需要回到其本身的内部结构、语义特点和表现内容等方面入手。如果在这些方面都无甚区别, 那么再根据“语形”将其统称为“俗语”也不为过。

17) 孔子学院/国家汉办《国际汉语教学通用词汇大纲》(2014年版) 北京语言大学出版社。

18) “格言”“警句”“名言”等形式即使被视为“熟语”单位, 也一般被归入“其他”, 甚至不属于“熟语”, 从其“言、句”的命名方式就可以判断其形式极有可能超过“语汇”的范畴。

19) 金立鑫 (2007) 《语言研究方法导论》, 第395页。

6. 结语

本文以《现汉》第7版电子融媒体版本认定的4342个“成语”为研究对象，从非四字成语的角度对成语的语形进行了详细探讨。在《现汉》划定的98个非四字成语中，单体成语共51个，双体成语有47个，两者数量相差无几。

双体成语的主要成员为42个双四字格，以超过总数40%的比例占第一，此外还有双三字格4个。双体成语具有明显的结构特征，不仅字数上前后一致，在结构和语义上也呈现相应对称状态。具有对称结构的双体成语占70%以上。这与四字成语中并列结构占绝对优势相呼应，两者在内部结构上有着高度一致性，可以推测前后对称（包括音步、语义、结构）是成语的原型特征之一。

单体非四字成语中以五字和六字形式为主，七字和三字成语居后。这些单体成语的注音形式更接近于句子而非“语汇”。其中五字成语多有相对应的四字形态，说明其更接近于成语四字语形的原型特征，两者容易发生转化。但六字和七字成语的语形更接近于俗语。就成语的语形数量作为原型特征而言，本文以《现汉》数据制如下图-1所示（其中“一时半会儿”在图中作五字处理）：

图-1 《现汉》成语数据统计

三字 (3)	四字 (4244)	五字 (21)	六字 (19)	七字 (8)
	双四字 (42)			
	双三字 (4)			
	四, 五字 (1)			

在语义方面，具有双层性的非四字成语比例明显高于四字成语，这极有可能是非四字单位在语形上不符合原型特征但仍然被认定为是成语的理由之一。此外习用历史较长是成语区别于惯用语的另一个原因。

因此，本文认为“成语”原型特征除了最重要的“四字形式”之外，结构对称性、语义双层性、语体书面化、来源典故性这几点也是成语在不同方面的原型

特征。但这几者之间具体权重如何，还有待进一步验证。

由于非四字成语在成语范畴中处于边缘地带，加之在语形和语义特点上与俗语都非常接近，虽然二者在语体上确实存在较大差异，但仅凭语体特征难以将两者的界限划清。因此将来有必要进一步探讨成语和俗语的原型特征来确定这些“语汇”的归属，这也有助于研究各熟语形式之间相互转化的条件和动因。

参考文献

单行本：

- 成语大词典编委会，《成语大词典》，北京：商务印书馆，2020年。
- 崔希亮，《汉语熟语与中国人文世界（修订本）》，北京：商务印书馆，2024年。
- 符淮青，《现代汉语词汇（增订本）》，北京：北京大学出版社，2017年。
- 孔子学院·国家汉办，《国际汉语教学通用词汇大纲》，北京：北京语言大学出版社，2014年。
- 教育部（中外语言交流合作中心），《国际中文教育中文水平等级标准》，北京：北京语言大学出版社，2021年。
- 金立鑫，《语言研究方法导论》，上海：上海外语教育出版社，2007年。
- 刘洁修，《成语》，北京：商务印书馆，1985年。
- 马国凡，《成语》，呼和浩特：内蒙古人民出版社，1978年。
- 莫彭龄，《汉语成语与汉文化》，南京：江苏教育出版社，2001年。
- 倪宝元·姚鹏慈，《成语九章》，杭州：浙江教育出版社，1990年。
- 温端政，《汉语语汇学》，北京：商务印书馆，2005年。
- 向光忠，《成语概说》，武汉：湖北人民出版社，1982年。
- 周荐，《汉语词汇结构论》，北京：人民教育出版社，2014年。
- 中国社会科学院语言研究所词典编辑室，《现代汉语词典》，北京：商务印书馆，2016年。

期刊论文：

- 陈振艳，《成语和类固定短语的语体鉴别及语体动因》，《浙江树人大学学报》第16卷，第6期，2016年。
- 冯胜利，〈论语体的机制及其语法属性〉，《中国语文》第5期，2010年。
- 洪波，〈对外汉语成语教学探论〉，《中山大学学报论丛》第2期，2003年。
- 李振凌，〈俗语的性质和范围〉，《烟台职业学院学报》第4期，2006年。
- 刘振前·邢梅萍，〈汉语四字格成语语义结构的对称性与认知〉，《世界汉语教学》第1期，2000年。

苏向丽, 〈现代汉语惯用语的词汇化等级分析〉, 《语言教学与研究》第5期, 2008年。

王德毅·施向东, 〈八字格成语浅析〉, 《鸡西大学学报》第8期, 2016年。

王利·曹瑶, 〈《中国成语大辞典》非四字格成语探析〉, 《黄冈师范学院学报》第41卷, 第1期, 2021年。

王勤, 〈俗语的性质和范围——俗语论之一〉, 《湘潭大学学报》第14卷, 第2期, 1990年。

姚锡远, 〈“熟语”的种属地位及其定义域〉, 《汉字文化》第2期, 1998年。

姚鹏慈, 〈试论汉语成语的长度〉, 《广播电视大学学报(哲学社会科学版)》第1期, 1998年。

张梁, 〈汉语惯用语原型再议——以《现汉》第7版为中心〉, 《韩中言语文化研究》第76辑, 2025年。

学位论文:

刘振前, 《汉语成语的对称特征与认知》, 华东师范大学博士学位论文, 1998年。

石清洋, 《面向汉语国际教育的非四字成语研究》, 山东师范大学硕士学位论文, 2018年。

孙月, 《俗语在熟语中的定位及其特征》, 河北大学硕士学位论文, 2006年。

任会芹, 《成语辞典所收录的非四字条目研究》, 河北大学硕士学位论文, 2011年。

Abstract

An Analysis of Non-Four-Character *Chengyu* in the *Xiandai Hanyu Cidian* from a Morphophonological Perspective

Zhang Liang

This study aims to identify the prototypical features of *chengyu* (Chinese idioms) and to delineate their boundaries from closely related categories such as *suyu*. Taking 98 non-four-character *chengyu* listed in the seventh edition of the *Xiandai Hanyu Cidian (Modern Chinese Dictionary)* as the research corpus, the paper conducts a detailed examination of their morphophonological forms. The findings show that double four-character patterns constitute the majority of non-four-character *chengyu*, followed by monomorphic five- and six-syllable forms. Among disyllabic *chengyu*, symmetrical structures are predominant, exhibiting symmetry not only in character count but also in syntactic organization and semantic correspondence. Monomorphic non-four-character *chengyu* display prosodic patterns more akin to sentences than to lexical items, and many five-syllable idioms have corresponding four-character variants. In terms of semantics, non-four-character *chengyu* show a significantly higher tendency toward dual-layered meanings. Therefore, this paper argues that, in addition to the most essential feature—the “four-syllable form,” other aspects such as structural symmetry, semantic duality, formal stylistic register, and origin from classical allusions also constitute the prototypical characteristics of *chengyu* (Chinese idioms).

Key words : *Xiandai Hanyu Cidian*, four-character *chengyu* (idioms), morphophonological form, *suyu* (proverbs), *shuyu* (fixed expressions), prototype

투 고 일 : 2025. 10. 10. / 심 사 일 : 2025. 10. 15.~ 2025. 11. 15. / 게재확정일 : 2025. 11. 20.

