

汉语四字成语语体特征再考察

— 以《现代汉语词典》语体标记为中心 —

张 梁*

<目次>

- | | |
|----------------|----------------|
| I. 引言 | IV. 带标成语的认可率调查 |
| II. 词汇语体的标记类型 | V. 结语 |
| III. 语体相关的其他标记 | |

<韩文摘要>

사자성어의 문체적 특성을 고찰하기 위해, 본 연구는 『현대한어사전』 애플리케이션 버전의 ‘성어’ 항목을 연구 대상으로 삼아 4,244개의 ‘사자’ 성어에 부여된 문체 표지를 면밀히 분석하였다. 그 결과, 명확한 문체 표지가 부착된 항목은 총 21개였으며, 그외에 ‘~儿’이 포함된 항목은 8개, ‘~的’ 표기가 있는 항목은 15개로 나타났다. 또한 경성(輕聲) 음절을 포함한 사자성어는 총 32개가 수록되어 있는데, 이들 역시 비교적 뚜렷한 구어체적 특성을 지닌 것으로 볼 수 있다. 이상의 세 가지 유형을 바탕으로, 본 연구는 총 80개의 사자성어를 대상으로 성어 인정율을 조사하였다. 그 결과, 문어체 표지가 부여된 성어가 가장 높은 수용도를 보이는 것으로 나타났다. 아울러 본 논문은 사자격(四字格)을 기반으로 한 다양한 성어의 원형적 특성에 대해서도 추가로 논의하였다. 비록 현재 중국어 학계에서

* 江原国立大学 中语中文学科 讲师

‘성어’에 대한 정의가 아직 확정되지 않았지만, 본 연구는 성어를 원형 이론의 관점에서 이해해야 한다고 본다. 즉, 하나의 관용 표현이 성어의 특성을 많이 충족할수록 모어 화자에게 ‘성어’로 인식될 가능성이 높아지며, 『현대한어사전』 역시 이러한 관점을 적절히 참고할 필요가 있을 것으로 보인다.

1. 引言

‘成语’的概念一直以来都是汉语词汇学界的一个热点问题, 对于一个词条是否属于成语范畴的界定问题曾引发大规模讨论。如刘叔新(1984)、王吉辉(1995)、徐耀民(1997)、刘正江(2006)、田辉(2009)、涂志成·谢莉(2009)等相关文章一直以来层出不穷。汉语成语如‘形式上为四字格’, ‘内部结构有对称性’, ‘语义具有双层性’, ‘来源具有典故性’, ‘音韵特点平起仄收’, ‘韵律节奏2+2’等诸多特点既鲜明但又不绝对。因此, 不少学者如施春宏(2002)、刘智芳(2004)、唐雪凝·许浩(2010)、张梁(2025)等开始利用认知语言学中‘原型范畴理论’来进一步认识成语这一概念。

其中关于成语的语体倾向, 相关研究则数量较少, 大抵是由于‘书面语体’特征十分显著, 不像‘语义双层性’等特点那么有争议性。但就语体理论来说, 成语也并非全然没有讨论价值, 如唐雪凝·许浩(2010)认为‘书面性’和‘有典性’等一样, 是成语的一般属性, 却不如形式上的‘四字格’强属性。张正生(Zhang: 2017, 2018)则从语体角度认为汉语成语等词汇单位的‘文言性’和‘书面性’并不重合。其他讨论汉语成语语体的研究如陈振艳(2016)、李宏玲(2021)等基本和Zhang类似, 主要立足于语体理论, 力图区分成语和类固定短语、惯用语之间的界限, 但很少给出具体的数据统计。除此之外, 学界鲜有从成语本体角度出发进行的探讨, 因此本文试图从辞书的元标记角度出发, 来考察成语的语体特征。

至于研究范围,本文选择的辞书是2016年出版的《现代汉语词典》第7版(以下简称《现汉》)。作为中国第一部现代汉语规范型词典,《现汉》由中国社科院语言研究所编纂,并由商务印书馆出版,其权威性不言而喻。此外,随着计算机技术的发展,《现汉》第7版更是在2019年推出了融媒体电子版,其中在‘分类词汇’板块中列有‘成语’4342条(其中四字成语4244条),相当于提供了一个现成的成语语料库,有利于展开定量研究,这是纸质版不具备的。本文就以电子版¹⁾中的四字成语为研究对象,进行成语语体特征的考察。

II. 词汇语体的标记类型

依据《现汉》‘凡例’第4.2条中的说明,一般条目中关于词汇的语体元标记共有〈口〉、〈方〉、〈书〉、〈古〉四种,分别表示该条目在口语、方言、书面上的文言词语、古代的提示性用法。这些在辞书中同样可以视为成语的语体元标记。《现汉》电子版中共收录四字成语4244条,本文整理后进行了穷尽式统计,发现其中除了〈古〉之外,其他三种标记皆有出现,本章将就这三种语体标记分别展开讨论。

1. 口语体标记

在4244个四字成语中,出现口语体标记〈口〉的成语共有9个词条,本文按其拼音顺序罗列在下〈表1〉中。

此外根据《现汉》凡例第3.3条,记载有关口语异读的情况:“口语中有比较通行的不同读音,在注音后面的括号内注明口语中的读音。如【一会儿】

1) 出现电子版与纸质版不一致的地方,本文以纸质版为准。其中仅涉及【避重就轻】一词,在电子版有标记〈书〉(电子版检索日:2025-12-15),而纸质版未标注,故以纸质版(见《现汉》75页)为准。

yīhuìr(口语中也读yīhuìr)；【作料】zuò·liao(口语中多读zuó·liao)。”成语中也同样存在‘口语’异读的情况：其中‘一本正经’与‘正儿八经’和‘正经八百’语义相近，且其中‘经’的读音在口语中都读作上声，但‘一本正经’没有出现口语体标记〈口〉，因此本文特意将其列入作为前两者的参照。

〈表 1〉带有口语体标记的四字成语

	词条	页数	释义
1	大天白日	245	〈口〉白天(强调)：~的，你怎么走迷了路！
2	点头哈腰	292	〈口〉形容恭顺或过分客气。
3	顶头上司	306	〈口〉指直接领导自己的人或机构。
4	黑灯瞎火	532	〈口〉形容黑暗，没有灯光：楼道里~的，下楼时注意点儿。 也说黑灯下火。
5	黑更半夜	532	(~的) 〈口〉指深夜。
6	见钱眼开	639	〈口〉见到钱财，眼睛就睁得大大的。形容人非常贪财。
7	溜须拍马	834	〈口〉指谄媚奉承。
8	正儿八经	1671	(口语中也读zhèng·er-bājīng)(~的) 〈口〉很正经：他是个~的庄稼人 咱们~地请他来吃顿饭。
9	正经八百	1672	(异-正经八摆) (口语中也读zhèngjīng-bābǎi) 〈口〉正儿八经。
-	一本正经	1532	(口语中也读yīběn-zhèngjīng) 形容很规矩，很庄重。

在上表中值得注意的是，‘黑更半夜’和‘正儿八经’两个词条还在释义中标有(~的)。另外，‘大天白日’‘黑灯瞎火’和‘正儿八经’三个词条给出了例句，全部都是‘~de’字段的形式出现在句中。让人不免联想到‘的’与口语体词汇的关联性。

此外，‘正经八百’以‘正儿八经’释义，可以视为后者的副条，同样标有口

语体标记〈口〉。但是‘黑灯瞎火’的副条‘黑灯下火’²⁾却未进行标记, 是否为标记失误值得商榷。

2. 方言体标记

《现汉》电子版中出现方言体标记〈方〉的成语条目共有8个, 数量与口语体标记不相上下, 本文罗列如下〈表 2〉所示:

〈表 2〉带有方言体标记的四字成语

	词条	页数	释义
1	阿猫阿狗	1	〈方〉泛指某类人或随便什么人(含轻蔑意)。
2	不尴不尬	107	〈方〉左右为难, 不好处理。
3	吹灯拔蜡	206	〈方〉比喻人死亡或垮台。
4	归里包堆	489	〈方〉总计; 拢共: 家里~就我和老伴儿两个人。
5	呼么喝六	548	② 〈方〉形容盛气凌人的样子。 ³⁾
6	灰头土脸	578	(~儿)〈方〉 ① 满头满脸沾上尘土的样子: 他扬完了场, 闹了个~。 ② 形容神情懊丧或消沉: 你高高兴兴地去了, 可别弄得~地回来。
7	家长里短	623	(~儿)〈方〉指家庭日常生活琐事: 谈谈~儿。
8	一头雾水	1538	〈方〉形容摸不着头脑, 糊里糊涂。

由上〈表 2〉可见, ‘呼么喝六’和‘灰头土脸’都有两个义项, 其中‘呼么喝六’的方言体标记只适用于第2个义项。但‘灰头土脸’适用于全部两个义项。此外, ‘呼么喝六’和‘家长里短’两个词条还有儿化音标记(~儿), 可见方言类词汇显著的口语性特点。

汉语方言支系众多, 鉴于汉语的书写系统自秦统一后的‘书同文’特点, 虽

2) 见《现汉》532页【黑灯下火】: 黑灯瞎火。

3) 义项1: ① 掷色子时的喊声(幺、六是色子的点子), 泛指赌博喧哗声。

然方言词汇在语体范畴中并不直接体现某类语体特征，但现代汉语毕竟是北京语音为标准音并以典型的现代白话文‘著作’为语法规范的共同语，而《现汉》又作为一部规范型词典，因此结合以上8个词条来看，本文认为《现汉》中所标记有〈方〉的词汇或多或少都体现出了口语体倾向，从上表中的两个口语儿化音便可见一斑。

3. 书面语体标记

众所周知，大部分汉语成语具有显著的书面语体特征，但《现汉》中出现语体标记〈书〉的仅有4个条目，具体见下〈表3〉所示：

〈表 3〉带有书面语体标记的四字成语

	词条	页数	释义
1	首善之区	1206	〈书〉最好的地方，指首都。
2	刎颈之交	1374	〈书〉指同生死共患难的朋友。也说刎颈交。
3	伊于胡底	1541	〈书〉到什么地步为止(对不好的现象表示感叹)。
4	饕餮不继	1578	〈书〉指吃了上顿没有下顿(饕餮：早饭和晚饭)。

以上4个条目中，‘刎颈之交’有一个三字副条‘刎颈交’⁴⁾，虽然并不是在电子版收录‘成语’板块内，但是依然标记有书面语体特征〈书〉。结合2.1节中‘黑灯下火’的情况，《现汉》中对主副词条的语体标记可能各自不同。另外值得注意的一点是，《现汉》正文共1760页，4个词条都出现在辞书的后1/3。那么问题在于前2/3部分是否完全不存在值得标记为书面语体特征的成语？甚至能否推论《现汉》由不同责任人分段负责？这些问题都有待进一步商榷与验证。

综合以上三个表的条目来看，4244个四字成语中总共只有21个条目有语

4) 见《现汉》1374页。

体标记。从比例来看不到0.5%，几乎可以忽略不计。但是就如凡例第3.3条所说的，有些语体特征并不体现在标记中，也可能体现在具体的释义之中。

Ⅲ. 语体相关的其他标记

如前一章2.1和2.2节中所述，释义部分中的(～儿)和(～的)在语用中与口语体特征关系密切。而后本文发现在《现汉》凡例中的第2.3条和2.4条中也有具体说明，因此本章将(～儿)和(～的)也视为相关语体标记，分为两节与轻声问题一并探讨。

1. 儿化音相关语体标记

《现汉》第7版凡例中涉及儿化音的条例主要有2.3和3.8两条，其中第3.8条是关于儿化音的注音形式，而‘字形和词形’中的2.3条则详细说明了儿化音和语体及义项有关的标注方法：

书面上有时儿化有时不儿化，口语里必须儿化的词，自成条目，如【今儿】、【小孩儿】。书面上一般不儿化，但口语里一般都儿化，在释义前加“(～儿)”，如【米粒】条。释义不止一项的，如口语里一般都儿化，就把“(～儿)”放在注音之后，义项①之前，如【模样】条。如只有个别义项儿化，就把“(～儿)”放在相应义项号码之后，如：【零碎】①……②(～儿)。如单字儿化与非儿化意义相同，分别用于不同的搭配中，也在释义前加“(～儿)”，通过举例表示在有些搭配中儿化，在另一些搭配中不儿化，如“胆”(～儿)胆量：～怯|～大心细|～小如鼠|壮壮～儿。

儿化音(也称儿化韵)是由于词根语素与后缀‘儿’合并为一个音节而产生的，再加之许多汉语方言都有类似的音变现象⁵⁾，因此儿化音属于汉语中典型

的口语性标志。依据以上第2.3条凡例，本文将带有‘(～儿)’标记的四字成语条目和义项全部列出，发现确有根据义项而不同的儿化音条目，如下〈表4〉所示，共有8个条目：

〈表4〉带有‘～儿’标记的四字成语

	词条	页数	释义
1	挤眉弄眼	613	(～儿)用眉眼示意：几个人都对他～，叫他别去。
2	零七八碎	832	②(～儿)零散没系统的事情或没有大用的东西：整天忙些个～儿 桌上放着好些他喜欢的～儿。 ⁶⁾
3	十字路口	1181	(～儿)两条路纵横交叉的地方，比喻在重大问题上需要对去向做出选择的境地：处在人生的～，他将何去何从呢？
4	细皮嫩肉	1409	(～儿)形容人的皮肤细嫩：这姑娘长得～的。
5	一年到头	1536	(～儿)从年初到年底；整年：～不得闲。
6	有头有脸	1590	(～儿)形容有名誉，有威信：他在村里是个～的，说话很有分量。
7	针头线脑	1662	(～儿)缝纫用的针线等物，多比喻零碎细小的东西。
8	转弯抹角	1721	(～儿) ①沿着弯弯曲曲的路走：汽车～开进了村子。 ②形容路弯弯曲曲：这条路～的，可难走了。 ③形容说话、做事不直截了当：有什么意见就痛快说，别这么～的。

正文由上〈表4〉可见，‘转弯抹角’的三个义项都带有儿化音口语体特征，但‘零七八碎’的两个义项中，只有第二个义项带有儿化音。另外〈表3〉中的‘灰头土脸’和‘家长里短’两个条目也带有儿化音标记，其实也可以归入上表。

值得注意的是，以上条目中的10个义项中9个义项都有出现了例句，比例

5) 北京大学中文系现代汉语教研室，《现代汉语(增订本)》，北京：商务印书馆，2012，97-101页。

6) 义项-1见本文〈表5〉的词条7。

之高, 甚于〈表 1〉〈表 2〉中带有〈口〉和〈方〉的条目, 与〈表 3〉的书面语体条目形成了鲜明的对比。此外, 其中的4个例句中都出现了‘~的’字段, 也就是说儿化音条目与‘的’字的共现频率相当之高。由此也可见‘的’在一定程度上可以视为口语体标记, 这在《现汉》凡例第2.4条中可以得到佐证。

2. ‘的’字相关语体标记

同样是在《现汉》第7版凡例的‘字形和词形’部分中, 第2.4条对带有‘的’字标记的相关语体标记做了如下说明:

状态形容词或某些重叠式在口语中经常带“的”或“儿的”, 条目中一般不加“的”或“儿的”, 只在释义前面加“(~的)”或“(~儿的)”, 如:【白花】báihuāhua(~的)……, 【乖乖】guāiguāi(~儿的)……。

从以上表述中可以得知: 儿话音与‘的’字的共现频率如此之高并不是偶然。因此, 按照第2.4条的说明, 本文将出现有‘(~儿的)’的标记归入‘(~的)’字标记, 两者合计共15个条目, 如下〈表 5〉所示:

〈表 5〉带有‘~的’标记的四字成语

	词条	页数	释义
1	半生不熟	37	②(~的)不熟习; 不熟练: 他试着用~的英语跟外宾交谈。 ⁷⁾
2	本乡本土	62	(~的)家乡; 本地: 菜都是~的, 请尝尝 都是~的, 在外边彼此多照应点儿。
3	怪模怪样	477	(~儿的)形态奇怪: 她这身打扮土不土, 洋不洋, ~的。
4	怪声怪气	477	(~的)(说话、唱歌等)声音、语气或腔调古怪难听。也说怪声怪调。
5	慌手慌脚	572	(~的)形容做事慌张忙乱。

7) 义项1: ①食物没完全加工熟; 果实没有完全成熟: 肉煮得~的, 没法儿吃 | 葡萄~的, 太酸。

	词条	页数	释义
6	火烧火燎	594	(~的)形容身上热得难受或心中十分焦灼。
7	零七八碎	832	①(~的)零碎而杂乱：~的东西放满了一屋子 被~的事儿缠住了，走不开。 ⁸⁾
8	闷声闷气	890	(~的)形容声音低沉，不响亮：他感冒了，说话有些~的。
9	奶声奶气	934	(~的)形容小孩子稚嫩的声音和腔调。
10	蔫头耷脑	951	(~的)形容耷拉着脑袋，没精打采的样子：他干了一天的活儿，累得~的，连句话也不愿意说◇地里的茄秧被晒得~的。
11	蹑手蹑脚	957	(~的)形容走路时脚步放得很轻：他~地走进了病房。
12	平头正脸	1008	(~儿的)形容相貌端正。
13	头疼脑热	1320	(~的)指一般的小病：~的，有两天就能好。
14	阴阳怪气	1560	(~的)(性格，言行等)乖僻，跟一般的不同：他说话~的，没法儿跟他打交道◇天气老是这样~的，不晴也不雨。
15	油嘴滑舌	1586	(~的)形容说话油滑。

从上表可见，‘零七八碎’的第一个义项带有(~的)字标记，而第二个义项则带有(~儿)化标记(见<表 4>)。另外与儿化音共现的(~的)字标记条目有‘怪模怪样’和‘平头正脸’两个，<表 4>的例句中还有‘细皮嫩肉’‘有头有脸’‘转弯抹角’三个。除此以外，<表 1>中的‘黑更半夜’‘正经八经’同样带有(~的)字标记。可见在《现汉》中把‘~的’和‘~儿’一同视为口语体标记是不会有问题的。

此外，<表 5>中的‘怪声怪气’带有一个副条——‘怪声怪调’⁹⁾，但和‘黑灯下火’一样，‘怪声怪调’未像‘正经八百’一样有语体标记。结合<表 1>至<表 5>，虽然有些主、副词条的标记前后不一，但数量并不太多。

总体来算，《现汉》电子版收录的4244个四字成语中，带有语体标记的

8) 义项-2见本文<表 4>的词条2。

9) 见《现汉》477页【怪声怪调】：怪声怪气。

仅有43个条目。且其中只有4个是书面语体标记,另外39个都可视为口语体标记,这一现象显然与成语语体的传统认知大相径庭。但如果以标记性理论(Theory of markness)来看,就会发现这并不能说是意外。

标记性理论在语言学中通常用于描述同一语言系统内部一组对立范畴之间的结构关系,即在成对或成组的范畴对立中,某一项作为默认项、分布更宽、语义/语用限制更少,往往体现为形式上更简约,而另一项则带有额外限制、功能更特殊,因而表现为有标记项。虽说不能将词典是否加注标签直接等同于‘有无标记’,但本文认为词典编纂者的说明策略与用法提示需求(如对方言、口语等使用域的提示)等‘元标记’可以视为标记性理论在辞书语言体系中的‘投射’,两者在一定程度上呈映射关系。

从《现汉》的体系来看,有标记的口语体成语只是少数,无标记的4201个成语才是占据绝对优势的主流。“无标记性是指与所有语言中发现的一般倾向相一致的特性,有标记性则是违背这种一般倾向的特性,即例外特性(相对共相)”¹⁰⁾。虽然标记性关注的是系统内部的对立与默认关系,而非简单的数量优势,但一般来说,无标记项的分布范围要比有标记项的大¹¹⁾。因此4个标有〈书〉的条目大概率才是‘意外’,如前文2.3节所述,存在标记失误的可能:即《现汉》后1/3部分对成语或有标记冗余的情况。

3. 轻声词与可轻读词的标注问题

除了前文所说的语体标记问题之外,轻声词和可轻读词的标注也与成语的语体特征也密切相关。可轻读词就是词条中某音节是否读作轻声都可以,并不作强制要求的词条。《现汉》一般用在标注原调的音节前加‘·’来表示可轻读音节¹²⁾。“一般来说,口语中的常用词才有读轻声音节的。”¹³⁾因此,标

10) 戴维·克里斯特尔,《语言学词典》,北京:商务印书馆,2000,216页。

11) 沈家焯,《不对称和标记论》,北京:商务印书馆,2018,24页。

12) 见《现汉》凡例第3.5条:“一般轻读、间或重读的字,注音上标调号,注音前再加圆点……”。表示读作轻声或原调皆可。以本文中【话里有话】一词为例,《现

有轻声的条目也可以视为具有较为明显的口语体特征。本文认为可视之为隐性语体标记。

同样从《现汉》电子版的4244个四字成语来看，本文发现注音作轻声或可轻读词标记的条目共有32个。其中可轻读词仅有‘话里有话’和‘果不其然’2个。由于数量较少，本文为行文简洁，一般按照‘轻声字(词)’的表达形式来叙述。在此按照轻声字在成语中出现的位置，制〈表6〉如下：

〈表 6〉含有(可)轻声字的四字成语

类型	注音形式	数量	例词
B位轻声	2-2	4	话里有话(可)、假模假式、狮子搏兔、正儿八经
	4	4	果不其然(可)、归里包堆、唧唧喳喳、死气白赖
D位轻声	2 2	12	安民告示、笔墨官司、顶头上司、东郭先生、断线风筝、狗皮膏药、江湖骗子、清水衙门、水磨工夫、绣花枕头、眼皮底下、一般见识
	2-2	9	不可收拾、不识抬举、好好先生、火烧眉毛、酒肉朋友、没头苍蝇、如意算盘、一门心思、依样葫芦
	4	2	满不在乎、无头告示
	112	1	过河卒子

从上〈表6〉可见，轻声字在四字成语中出现的最大特点当属只出现在B位和D位，即只有第二字和第四字才可能出现轻读现象。两者比例为1:3(8:24)，D位出现的频率又明显高于B位。其次，可轻读字只出现在B位。

此外，值得注意的一点是D位轻声字的这些‘双音节’轻声词在进入成语之前多数已有轻声现象，本文仅列举《现汉》里收录且明确标注为轻声的词(括号内为《现汉》第7版收录的页数)：

“收拾”(1201)、“膏药”(435)、“骗子”(999)、“抬举”(1263)、“先生”-2次

汉》标记为‘huà·lǐ-yōuhuà’。

13) 黄伯荣·廖序东，《现代汉语》，北京，高等教育出版社，2002，106页。

(1417)、“眉毛”(886)、“朋友”(968)、“苍蝇”(127)、“算盘”(1251)、“心思”(1456)、“葫芦”(551)；“告示”-2次(436)、“官司”(481)、“上司”(1147)、“风筝”(392)、“衙门”(1500)、“工夫”(448)、“枕头”(1664)、“底下”(282)、“见识”(639)；“在乎”(1629)；“卒子”(未收录)。

可见D位出现轻声的24个‘成语’中，除了‘卒子’未被《现汉》收录以外，其他都是既有的轻声词。但B位轻声字只有‘狮子’一词属于此类现象，也只有‘狮子搏兔’的[AB]位是以双音词形式进入四字格的。很显然，B位轻声和D位轻声的性质并不一样。B位轻声的成语符合四字格其中一种[2013](即‘中轻轻重’)的重音模式(冯胜利1997，朱赛萍2015，张梁2023等)；D位轻声的成语虽然大部分可视为‘2+2’结构，但既不遵循四字格主流的[0213](即‘轻中轻重’)模式，又不符合[2013]的模式，可以想见其在母语者意识中难以成为‘典型’的成语。

如前文所述，口语中的常用词才有读轻声音节的。因此这些双音节轻声词在成为成语的构成成分之后，形成的四字成语依然具有明显的口语体特征。比如其中‘顶头上司’‘归里包堆’‘正儿八经’三个条目已经带有〈口〉或〈方〉的语体标记。由此分析，在语音上表现出来的轻声形式也可视为隐性的口语体标记。

如此合计，即便加上〈表6〉中新增的29个轻声词，口语体倾向的成语也一共只有68个，在4244个四字成语中只占1.6%左右，属于可以被忽略的数量规模。那么这些口语体倾向明显的四字条目是否真如《现汉》收录的那样具备‘成语’资格？恐怕还需要进一步探讨与验证。

IV. 带标成语的认可率调查

基于多数成语带有书面语体特征的既有认知，以上〈表1〉至〈表6〉的条目——特别是具有口语体标记的条目是否具有成语资格；以及缘何被《现汉》

收录并认定为成语，成了必须要面对的问题。因此本文面向汉语母语者展开了问卷调查与相关讨论。

1. 问卷设计与数据统计

调查条目中，除了第3章结尾处提到的68个口语性成语和〈表3〉中的4个书面语体标记之外，还加入了2010年后以微博平台¹⁴⁾为主的中文社区兴起的6个高频‘网络新成语’¹⁵⁾作为填充语料¹⁶⁾，分别为‘不明觉厉’‘累觉不爱’‘十动然拒’‘喜大普奔’‘指鼠为鸭’和‘提灯定损’。其中‘指鼠为鸭’‘提灯定损’是诞生于近两年(2023年6月和2024年3月)有明确出处的社会性事件；而另外4个条目在2014年《中国语言生活状况报告》中的‘网络四字格’(227-228页)盘点中已经被明确提及，均形成于2011~2013年间，已有10余年历史。两者共同特点是不含轻读音节，既有利于和〈表6〉中的轻声成语条目进行对比，又可以从侧面一定程度了解成语的‘久远性’特征。

刘中富(2016)明确主张“成语是带有层次性的开放系统”，本文持类似观点，如果以上这些‘新成语’能够经受住时间的考验，在汉语历史中保持生命力，那么若干年后也有可能进入到成语的范畴。最后本文还选择了‘守株待兔’一词作为典型成语¹⁷⁾的代表性条目，和〈表1〉中的‘一本正经’一起，同作为参照条目。由此共形成80个条目作为问卷范围，并随机打乱作答顺序。

由于问卷条目数量众多，为减少被试者的作答负担，本文选择利用个人微信渠道小范围点对点发放作答，以提高问卷的有效率，并且不收集被试者性别、年龄等社会信息。虽然大部分成语语素为常用字，但为尽量避免条目

14) 检索时间范围为2010年1月1日~2025年12月31日(因新浪微博于2009年8月上线运营，本文取整年度计算)。

15) 条目来源参考2010年~2025年的《中国语言生活状况报告》。

16) 填充语料在本问卷中为干扰项，如这6个不入典的干扰项全部被选为成语，则视为无效问卷处理。

17) ‘守株待兔’符合前言中所述成语的多数标准：形式上为四字格、内部结构有对称性、语义具有双层性等，详见本文〈表8〉与4.2节有关‘原型范畴理论’的论述。

中的生僻字等因素影响被试者,对作答者要求为大学专科以上学历母语者,仅进行‘是’或‘否’的二选一调查¹⁸⁾。问卷收集时间为2025年12月19日至2026年2月19日。获得样本共97份。

在剔除无效问卷¹⁹⁾后,共获得有效样本88份。按照本文前述的标记,共分为‘口、方、书、儿、的、轻B、轻D、新’8组,其中‘守株待兔’和‘一本正经’视为带有书面语体特征的条目,以确保每一类都至少有5个以上条目,形成聚集规模。问卷调查结果的基本数据如下〈表7〉所示:

〈表7〉成语语体类别的认可率

标记	数量	平均认可率	条目(成语认可率-由低到高排列)
口	9	50.00%	大天白日(18.18%)顶头上司(22.73%)黑更半夜(28.41%) 正经八百(43.18%)正儿八经(63.64%)黑灯瞎火(64.77%) 溜须拍马(65.91%)点头哈腰(69.32%)见钱眼开(73.86%)
方	8	50.57%	不尴不尬(20.45%)阿猫阿狗(22.73%)呼么喝六(39.77%) 归里包堆(44.32%)吹灯拔蜡(63.64%)家长里短(63.64%) 一头雾水(69.32%)灰头土脸(80.68%)
书	6	77.65%	伊于胡底(51.14%)首善之区(57.95%)饕餮不继(77.27%) 刎颈之交(88.64%)守株待兔(95.45%)一本正经(95.45%)
儿	8	51.29%	十字路口(15.91%)一年到头(17.05%)有头有脸(44.32%) 针头线脑(50.00%)零七八碎(59.09%)细皮嫩肉(68.18%) 转弯抹角(70.45%)挤眉弄眼(85.23%)
的	14	56.25%	慌手慌脚(26.14%)本乡本土(29.55%)头疼脑热(31.82%) 奶声奶气(42.05%)平头正脸(46.59%)怪声怪气(47.73%) 怪模怪样(52.27%)闷声闷气(61.36%)半生不熟(64.77%) 火烧火燎(65.91%)阴阳怪气(72.73%)蔫头耷脑(77.27%) 蹑手蹑脚(81.82%)油嘴滑舌(87.50%)
轻B	6	54.73%	死气白赖(47.73%)唧唧喳喳(50.00%)假模假式(52.27%) 狮子搏兔(52.27%)话里有话(53.41%)果不其然(72.73%)

18) 问卷引导语为“您是否认为下列80个四字条目属于成语?请根据直觉语感进行判断即可。本次问卷仅面向汉语母语者,感谢配合!”

19) 作答时间短于60秒的;长于500秒的;重复作答的情况等(另见脚注16)本文视为无效问卷,共9份。

标记	数量	平均 认可率	条目(成语认可率-由低到高排列)
轻D	23	48.37%	江湖骗子(21.59%)好好先生(29.55%)眼皮底下(30.68%) 一般见识(32.95%)没头苍蝇(36.36%)东郭先生(38.64%) 断线风筝(42.05%)依样葫芦(42.05%)安民告示(43.18%) 酒肉朋友(43.18%)无头告示(44.32%)满不在乎(46.59%) 不可收拾(47.73%)过河卒子(50.00%)水磨工夫(50.00%) 一门心思(51.14%)笔墨官司(54.55%)绣花枕头(59.09%) 狗皮膏药(60.23%)清水衙门(60.23%)如意算盘(71.59%) 火烧眉毛(73.86%)不识抬举(82.95%)
新	6	49.24%	累觉不爱(38.64%)十动然拒(38.64%)指鼠为鸭(46.59%) 喜大普奔(48.86%)提灯定损(54.55%)不明觉厉(68.18%)

根据统计结果,全部80个成语条目的平均认可率为53.18%;但由于书面语体类条目数量较少,难以反映各语体类全貌。因此8组‘成语’可以大致按语体标记分为书面语类(77.65%)、口语类(51.37%)和暂无标记的网络新成语类(49.24%)这三大类。

由〈表7〉可见,最突出的无疑是带有标记〈书〉的一组,以77.65%高居第一,比第二‘的’字组高出20%以上。即使去掉‘守株待兔’和‘一本正经’两个作为典型成语的参照条目,其余4个成语的平均认可率也高达68.75%。其中‘伊于胡底’(51.14%)该条目在书面语条目中认可率最低,但也与口语类的总体平均认可率(51.37%)相当。

此外,虽然近年来的网络新成语与其他几组相比尚未有语体标记,但认可率(49.24%)却与另外几组口语类的数据相当接近,与两组轻声成语(共29个)的平均认可率相比也几乎持平(49.69%)。其内部差异也并不明显:‘指鼠为鸭’和‘提灯定损’两个相对较新条目的平均认可率与4个2011~2013年间形成的较早条目仅相差不到2%。

其余6组口语类的成语认可率都在48~57%之间,由高到低排序分别为‘的-B-儿-方-口-D’,首尾差异之小不足10%,尚且不如一二名的〈书〉字组与‘的’字组。可见这几组距离母语者心目中典型的‘成语’范畴还有相当一段距

离。即便数据如此接近,依然可以整理出以下一些信息:

1) 方言体标记〈方〉的成语认可率与带有口语体标记〈口〉的成语最为接近,都在50%上下,相差不到1个百分点,可见方言性成语的口语体特征相当显著。

2) ‘儿’字组的成语认可率为51.28%,比‘的’字组(56.25%)的成语更接近带有口语体标记的〈口〉字组,也可以反过来证明‘的’字组成语相对来说更接近书面语体特征。‘的’字组排名第二便可见一斑。

3) ‘轻B’组的认可率为54.73%,比数量最多但认可率最低的‘轻D’组(48.37%)高出5%左右,可见D字最轻的四字格比B字最轻的[2013]式(即‘中轻轻重’)更难被母语者识别为‘成语’。因此,‘轻D’组虽然结构上和[0213]型(即‘轻中轻重’)成语类似,多为‘2+2’复合式,但轻声与重音的影响显然不可小觑。本次调查结果支持张梁(2023)的‘四字格轻声接受度假说’,即在四字成语中,B位轻声的可接受度大于D位轻声²⁰⁾。由此也可见B位轻声和D位轻声两者的音系性质并不一样,应区别看待。

2. 结果分析与相关讨论

除了上节中较为直观的数据之外,还可以通过一些比较极端的条目来分析统计结果:‘一本正经’以94.45%的认可率与‘守株待兔’并列第一,但与意义相近且同样出现口语异读的‘正经八百’(43.18%)和‘正儿八经’(63.64%)相比,认可率高出1/3~1/2,可见《现汉》的口语体标记确有区分语体的作用。

同‘守株待兔’这种典型成语相反的是,‘十字路口’(15.91%)‘一年到头’(17.05%)和‘大天白日’(18.18%)三个条目在问卷中的认可率最低,均不足20%,与‘一本正经’差距显著。本文也尝试利用‘原型范畴理论’(Prototype Theory)来说明不同‘成语’认可率差异巨大的原因。

20) 张梁,《현대 중국어四字格 운율 조어법 연구》, 한국외국어대학교 박사학위논문, 2023, 88쪽.

原型范畴理论与经典的范畴化理论(Categorization Theory)相对, 一般来说拥有以下几项假定:(a)实体是根据其自身属性来加以范畴化的, 通常是个连续统(continuum), 而非经典范畴理论中的两分型理论;(b)各范畴之间的边界是模糊的、不确定的;(c)同一范畴内的成员地位并不相等, 有较好的样本与较差的样本之分, 较好的成员处于集合的核心地位, 更为典型, 也更能代表范畴²¹⁾。早有前人学者²²⁾利用该理论来处理汉语词汇学中一些具体的界定问题。

问卷中的‘守株待兔’就是本文按照‘较好的样本’所选择的参照条目, 而‘十字路口’‘一年到头’和‘大天白日’从〈表 7〉中的数据上看, 显然就是范畴中‘较差的样本’, 处于‘成语’的边缘地带, 因此超过80%的被试者并不认可其具有成语的资格。但成语这一范畴显然并不只具有一种原型特征, 如‘十字路口’若按照早期范畴化理论, 用成语的语义双层性²³⁾作为‘区别性特征’²⁴⁾的话, 显然不能说其是处于边缘地带的‘较差的样本’。

故而本文将引言中所提及前人学者列出的诸多成语原型特征²⁵⁾汇总并整理为以下〈表 8〉所示11条, 并将这认可率最低的三个词条与最高的‘一本正经’和‘守株待兔’作适用性对比。虽然以下特征来源复杂, 涵盖形式、音系、语义、来源、语体与句法等不同层级, 但目前学界对成语的各特征权重比例仍未有定论, 因此在初期阶段出于可操作性的考量, 该表暂未进行加权操作, 难免不够客观, 今后随着研究深入有待进一步补充完善。

21) 如Lakoff(1987)、Taylor(1989)、廖秋忠(1991)等。

22) 如董秀芳(2011)利用动词的及物性这一原型范畴来解释动宾式双音词的语义限制问题。可见《词汇化——汉语双音词的衍生和发展》第170-172页。

23) 《现汉》释义中的明确提到其比喻性用法, 具体见1181页【十字路口】。

24) 先行研究如: 刘叔新(1984)、王吉辉(1995; 1998)、李行健(2001)、陈桂成(2005)、刘正江(2006)、涂志成·谢莉(2009)等持类似观点。

25) 包括但不限于以下文献: 崔希亮(1993)、刘振前·邢梅萍(2000, 2003)、刘智芳(2004)、唐雪凝·许浩(2010)、周荐(2014)、朱赛萍(2015)、刘中富(2016)等。

〈表 8〉成语原型特征适用性对比

	一本正经	守株待兔	十字路口	一年到头	大天白日
四字形式	+	+	+	+	+
节奏: 二二相承	+	+	+	+	+
音韵: 平起仄收 ²⁶⁾	-	+	-	-	+
重音模式: [0213]	+	+	±	±	+
结构对称性(骈体)	-	+	-	-	±
语义双层性	+	+	+	-	-
历史久远性 ²⁷⁾	±	+	-	-	-
来源典故性	±	+	-	-	-
语体书面性	+	+	-	-	-
句法叙述性	+	+	-	+	-
语素古雅性	±	±	-	-	-
	7.5	10.5	3.5	3.5	4.5

由上表可见, ‘守株待兔’基本符合前人学者提出的所有成语特点, ‘一本正经’则紧随其后。而‘十字路口’和‘一年到头’则多数不符合成语的原型特征, 也符合本次问卷排名垫底的调查结果, 因此《现汉》认可其为‘成语’显然有待商榷。

此外值得注意的是, 以上诸多特征中如‘二二相承’‘平起仄收’和‘重音模式’, 甚至‘结构对称性(骈体性)’²⁸⁾都是建立在‘四字形式’的基础之上的。因此本文认为, 将形式上的四字特征视为汉语‘成语’的最大特点也不为过。除此外的‘语义双层性’‘来源典故性’等特征, 包括本文考察的‘语体书面性’, 对成语原型的重要性排序和具体占比具体如何, 值得今后进一步深入研究。

26) 依据崔希亮(1993)《汉语四字格的平起仄收势——统计及分析》, 本文具体处理原则指四字成语的第二音节为平声, 第四音节为仄声的情况。

27) ‘历史久远性’和‘来源典故性’本文参照商务印书馆国际有限公司2013出版的《成语大词典》。该词典基本所有成语出处都标有年代。出处为清末民国时代的成语本文标为‘±’。

28) 部分六字格、双三字格和双四字格等条目符合该特征。

需要补充的是，成语作为“人们长期以来习用的定型词组或短句”²⁹⁾，在句中一般处于词和句子的中间层，因此并不能通过成语数量直接反应语篇的语体特征。朱晓睿·骆健飞(2023)认为“成语的多少无法反映文本正式度，为弱指标，不具备区分正式度的作用”，本文持类似观点。本次研究的调查结果同样支持其文中的其他部分结论：

“儿化词、第二人称代词、必读轻声词、除‘的’以外的句末语气词以及语气副词为负向指标，即此类词越多，文本正式度越低，皆具备强制性，与正式语体相排斥。嵌偶韵律词、合偶双音词、古汉语虚词为正向指标，此类词越多，正式度越高，为强指标，但不具备强制性。”³⁰⁾

以上结论意味着如果一个四字单位中一旦带有儿化音(如本文〈表 4〉内条目)、轻声音节(如本文〈表 6〉内条目)等情况，就不易被母语者认可为‘成语’，即该四字单位的‘口语性/非正式性’较强。但四字条目中常常由合偶双音词³¹⁾(‘二二相承’平衡律³²⁾、古汉语虚词(如本文〈表 3〉内‘伊于胡底’的‘于’、‘刎颈之交’的‘之’)组成，那么就意味着其‘书面性’较强。也就是说‘成语’内部既有‘负项指标’，又有‘正向指标’，甚至同一个条目内部兼有正负两个指标，如‘果不其然’的‘不’[-可轻读音节]和‘然’[+古汉语虚词]，那么自然无法通过成语的多少反应文本的语体特征，只能视其为‘文本正式度的弱指标’。想必这就是‘书面性’虽为成语的原型特征，但远不及‘四字形式’属性更重要的原因³³⁾。

29) 见《现汉》166页【成语】。

30) 朱晓睿·骆健飞，〈汉语正式语体与非正式语体差异的量化研究〉，《汉语国际教育学报》第14辑，2023，123-124页。

31) 详情可见专著：王永娜，〈汉语合偶双音词〉，北京语言大学出版社，2015。

32) 具体可见：冯胜利·刘丽媛(2020)〈语体语法的生物原理与生成机制〉一文：“相对的两个节律单位如长短一致，则具正式性。”

33) 《现汉》电子版的‘成语’板块中，四字形式占97.74%(4244/4342)。

V. 结论

本文以《现代汉语词典》电子融媒体版本中的‘成语’板块为研究范围,详细考察了4244个四字成语的语体类‘元标记’。其中带有口语体标记〈口〉的条目共有9个;带有方言体标记〈方〉的条目共有8个;带有书面语体标记〈书〉的条目共有4个(详见〈表1〉至〈表3〉)。此外并无带有古语体标记〈古〉的条目,合计21个四字条目带有明确的语体标记。

除了《现汉》中明确出现的语体标记之外,本文还依照凡例中的第2.3条和第2.4条考察了可带有‘~儿’和‘~的’标记的条目。带有‘~儿’标记的条目共有8个;带有‘~的’标记的条目共有15个(详见〈表4〉和〈表5〉)。其中‘零七八碎’一词根据义项不同,第一义项带有‘~的’标记,而第二义项则带有‘~儿’标记。此外,在注音形式中标为轻声词或可轻读词的条目也可以视为具有较为明显的口语体特征。这类含轻声字的四字成语在《现汉》中共有32个条目,其中第二字B位轻声条目8个;第四字D位轻声条目24个。但‘项头上司’‘归里包堆’‘正儿八经’三个条目已经带有〈口〉或〈方〉的语体标记。

结合以上几种情况:有43个成语带有相关的语体标记,新增29个以轻声的注音形式体现的语体标记,共72个条目。在此基础上再加6个网络新成语和2个‘典型成语’的参照条目。本文以80个四字条目为问卷范围展开了成语认可率调查。结果显示带有书面语体特征标记的成语最受认可。新兴成语则与其他几类口语组的认可率相当接近。另外,几组口语性成语的比较中,‘~的’字组比‘~儿’字组更受认可;第二字读轻声的‘轻B’组比第四字读轻声的‘轻D’组更受认可。

以问卷调查的结果为基础,本文进一步讨论了有关成语原型的一些特征因素,其中‘书面性’倾向确实可以视为成语的原型特征之一,但不如形式上的‘四字格’更有约束力。虽然目前汉语学界对‘成语’的定义仍然未有定论,但只要一个条目符合越多的成语原型特征,那么就更容易被母语者认可为是典型的‘成语’。此外就《现汉》的收录范围来说,由于‘成语’内部本身就包含了一

部分‘口语性’较强的条目，虽然数量不多，但或多或少仍然影响着成语‘书面性’的原型特征。因此无法通过成语使用的多少直接反应语篇的文本正式度。

综上所述，基于词典学的角度，本文推测各类辞书对成语语体的标注和处理将来可能会有两个大方向，一个是将‘口语性’较强的非典型条目剔除出‘成语’范围。比如《现汉》今后可适当参照‘原型理论’来收录成语，不然即便是四字条目，可能依然会令读者产生诸如‘十字路口’和‘一年到头’为何是成语的困扰。想必这样处理的结果会使‘书面性’特征的重要度进一步上升，与文本正式度挂钩。另一个方向则是在承认‘口语性成语’的前提下，继续完善对成语的语体标记。比如将含轻声音节和带儿化音节的条目等都打上〈口〉语体标记，那么相对应的是否应当将含有古汉语虚词语素的条目全部都打上〈书〉或〈古〉语体标记也值得进一步思考。这种处理结果会让更多条目带有书面语体标记，远不止〈表 3〉中的四个。如带有‘之’字的四字成语在《现汉》电子版中就多达160条，光这一项就大大超过本文统计的‘口语性成语’总数68个。继而如何判断一个成语中是否含有古汉语语素(如‘饕餮不继’中的‘饕’‘餮’为实词语素)等可操作性问题就会接踵而至。因此‘书面性’虽为成语原型特征的‘一般属性’，但是否应该像〈表 3〉中的条目一样对其打上书面语体标记，本文持审慎意见。

〈参考文献〉

- 张梁, 《현대 중국어四字格 운율 조어법 연구》, 한국외국어대학교 박사 학위논문, 2023.
- 张梁, 〈语形视角下的《现汉》非四字成语探析——兼论成语的原型特征〉, 《한중언어문화연구》 제2집, 2025.
- 北京大学中文系现代汉语教研室, 《现代汉语(增订本)》, 北京:商务印书馆, 2012.

- 陈桂成,〈成语界说与成语词典立目〉,《辞书研究》第1期,2005.
- 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6期,2016.
- 成语大词典编委会,《成语大词典》,北京:商务印书馆,2020.
- 崔希亮,〈汉语四字格的平起仄收势——统计及分析〉,《当代修辞学》第1期,1993.
- 崔希亮,〈正式语体和非正式语体的分野〉,《汉语学报》第2期,2020.
- 戴维·克里斯特尔,《语言学词典》,北京:商务印书馆,2000.
- 董秀芳,《词汇化——汉语双音词的衍生和发展》,北京:商务印书馆,2011.
- 冯胜利,〈论汉语书面正式语体的特征与教学〉,《世界汉语教学》第4期,2006.
- 冯胜利,〈论语体的机制及其语法属性〉,《中国语文》第5期,2010.
- 冯胜利·刘丽媛,〈语体语法的生物原理与生成机制〉,《民俗典籍文字研究》第26辑,2020.
- 韩硕,〈汉语成语界定浅议〉,《语文学刊》第16期,2015.
- 韩越,〈汉语成语特征分析〉,《湘潭师范学院学报》第4期,1999.
- 黄伯荣·廖序东,《现代汉语》,北京:高等教育出版社,2002.
- 教育部语言信息管理司,《中国语言生活状况报告(2014)》,北京:商务印书馆,2014.
- 李宏玲,《基于韵律和语体理论的汉语成语和惯用语对比研究》,天津大学硕士学位论文,2021.
- 李行健,〈成语规范问题〉,《辞书研究》第2期,2001.
- 廖秋忠,〈《语言的范畴化:语言学理论中的典型》评介〉,《国外语言学》第4期,1991.
- 刘静静,〈成语、惯用语界定标准之我见——由“表意双层性”引发的思考〉,《南华大学学报(社会科学版)》第1期,2016.

- 刘叔新,《词汇学和词典学问题研究》,天津:天津人民出版社,1984.
- 刘叔新,《汉语描写词汇学》,北京:商务印书馆,2005.
- 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.
- 刘振前·邢梅萍,〈四字格成语的音韵对称与认知〉,《语言教学与研究》第3期,2003.
- 刘正江,〈意义的双层性是汉语成语的“区别性特征”〉,《新疆财经学院学报》第1期,2006.
- 刘智芳,〈从原型理论看成语、惯用语的划分〉,《池州师专学报》第1期,2004.
- 刘中富,〈成语的界定与成语的层次性〉,《山东师范大学学报(人文社会科学版)》第4期,2016.
- 莫彭龄,〈关于成语定义的再探讨〉,《常州工业技术学院学报(社会科学版)》第1期,1999.
- 施春宏,〈词义的认知模式与词义的性质及构成——兼谈成语的性质〉,《辞书研究》第6期,2002.
- 沈家煊,《不对称和标记论》,北京:商务印书馆,2018.
- 唐雪凝·许浩,〈认知视野下的成语再认识——从经典理论到原型理论〉,《东岳论丛》第10期,2010.
- 陶红印,〈试论语体分类的语法学意义〉,《当代语言学》第3期,1999.
- 田辉,〈对“成语”界定的分析与探求〉,《楚雄师范学院学报》第10期,2009.
- 涂志成·谢莉,〈汉语成语界说新论〉,《理论界》第9期,2009.
- 徐耀民,〈成语的划界、定型和释义问题〉,《中国语文》第1期,1997.
- 王吉辉,〈成语的范围界定及其意义的双层性〉,《南开学报(哲学社会科学)》第6期,1995.
- 王吉辉,〈意义的双层性及其在成语、惯用语划分中的具体运用〉,《南开学报(哲学社会科学)》第4期,1998.

- 王永娜, 《汉语合偶双音词》, 北京: 北京语言大学出版社, 2015.
- 姚鹏慈, 〈关于成语语感与成语度的思考〉, 《广播电视大学学报(哲学社会科学版)》第2期, 内蒙古广播电视大学, 2002.
- 张伯江, 〈语体差异和语法规律〉, 《修辞学习》第2期, 2007.
- 中国社会科学院语言研究所词典编辑室, 《现代汉语词典(第7版)》, 北京: 商务印书馆, 2016.
- 周荐, 《汉语词汇结构论》, 北京: 人民教育出版社, 2014.
- 朱赛萍, 《汉语的四字格》, 北京: 北京语言大学出版社, 2015.
- 朱晓睿·骆健飞, 〈汉语正式语体与非正式语体差异的量化研究〉, 《汉语国际教育学报》第14辑, 2023.
- Lakoff, G., *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press, 1987.
- Taylor, J. R., *Linguistic categorization: Prototypes in linguistic theory*, Oxford: Oxford University Press, 1989.
- Zhang, *Dimensions of Variation in Written Chinese*, London: Routledge (Taylor & Francis Group), 2017.
- Zhang, “Mapping Stylistic Variation with Correspondence Analysis”, *Journal of Technology and Chinese Language Teaching*, Volume 9 no. 2, 2018.

<Abstract>

To investigate the stylistic characteristics of Chengyu, this study takes the “成语” section of the electronic multimedia edition of the *Xiandai Hanyu Cidian* (Modern Chinese Dictionary) as its research scope and conducts a detailed examination of the stylistic labels assigned to 4,244

Chengyu. Among these, 21 entries carry explicit stylistic labels. In addition, there are 8 entries marked with “-r”(~儿), 15 entries marked with “-de”(~的), and 32 Chengyu containing neutral-tone syllables; these Chengyu can also be regarded as exhibiting relatively salient colloquial features. On the basis of these, This study conducted an Chengyu recognition survey covering 80 four-character entries. The results show that Chengyu labeled as belonging to the written style receive the highest level of acceptance. Furthermore, the study discusses a number of prototypical features of Chengyu based on the four-character pattern. Although there is still no consensus in Chinese linguistics on the definition of “成语” this study argues that Chengyu should be approached from the perspective of prototype theory: the more idiomatic features a fixed expression possesses, the more readily it is recognized as an Chengyu by native speakers, and the *Xiandai Hanyu Cidian* should take this perspective into consideration to an appropriate extent.

Key Words : 成语(Four-Character Idioms), 标记(marker), 四字格(Chinese Four-Character Structure), 语体(stylistic register), 现代汉语词典(*Xiandai Hanyu Cidian*)