

中文语体构式及其资源构建初探*

黄国敬 饶高琦

北京语言大学汉语国际教育研究院

ellenh1001@163.com raogaoqi@blcu.edu.cn

摘要: 语体与构式的研究一直受到学界的关注,二者结合而成的语体构式是否存在?是否能够构建相关的资源?本文通过厘清相关研究,论证了语体构式的存在,语体构式是表达语体功能的规约化的形式意义对,并对语体构式的内涵进行了描述。进一步尝试构建了相同语义下的平行语体构式资源,建立了资源构建规范并进行了实践,最终得到了5,233条构式,初步对这些构式数据进行了分析。最后,探讨了语体构式在语言研究、汉语教学、自然语言处理以及翻译与本体化等领域的可能应用空间。

关键词: 语体构式, 语体资源, 语体, 构式

A Preliminary Study on Chinese Register Construction and Its Resource Construction

Abstract: The study of register and construction has been receiving a lot of attention from the academic community. Does a combined register construction exist? Is it capable to construct relevant resources? In this paper, we clarify the relevant studies, argue for the existence of register construction, which is conventionalized form-meaning pairs, and express the features of the register construction. Further attempts are made to construct the resource of parallel register constructions under the same semantic meaning, then we established a specification for the resource and conducted the practice. 5,233 constructions are finally obtained, and the data of these constructions are preliminary analysed. Finally, the applications of register construction in the fields of language research, Chinese language teaching, natural language processing, and localized translation are discussed.

Keywords: register construction, register resource, register, construction

1. 引言

语体一直是学者受到广泛关注的问题,是语言为了完成不同的功能、适应不同的语域而形成的相对恒定的表达模式,其体现在语言生活的方方面面。目前语体研究众多,但语体资源实践则相对较少,并且较多关注在语体语料库构建和语体词词典编纂。但语体差异并不只是体现在词语级别的差异,而构式的内涵更为丰富,可以借以描写语体差异。所谓构式,就是具体语言系统中的规约化的形式意义对¹,在广义的构式观下,语体中能够形成规约化的形式—功能配对的内容自然也能够纳入构式。因此本文进一步提出了一种平行语体构式,即在同一语义下,由表达不同语体功能的构式所组成的一组平行构式,最终得以构建以平行语体构式为主体,辅之以实例集合的平行语体构式资源。该资源不仅能够服务于语言本体的研究,促进语言教学,还可为自然语言处理及生成提供基础资源。同时,这也是对构式语法的一种尝试性实践。

本文介绍了语体和构式的相关研究,建立了语体构式这一概念,并阐述了其内涵和特征。以此为基础构建了平行语体构式资源,描述了构建该资源的规范及实践,探讨了语体构式及其资源的理论和应用价值。

*本文系教育部人文社科青年项目“清末以来汉语报刊词汇使用计量研究(20YJC740050)”阶段成果

¹ 参见施春宏. 构式的观念:逻辑结构和理论张力[J]. 东北师大学报(哲学社会科学版),2016,(04):17-23.

2. 相关研究

2.1 语体研究

汉语语体的研究是关于汉语的基础性研究，近些年来也广泛受到学界关注。语体作为一种语言现象，与语言的使用息息相关，最能体现语言的交际功能，因此语体研究极为重要，需要持续的关注。目前语体的研究已经取得了一些成果，主要有语体理论研究、语体实证研究两方面。

语体理论研究成果众多。在语体的定义和分类方面，唐松波（1984）、李熙宗（2006）对以往的各种语体相关概念进行了区分。语体的划分主要有二分法、三分法以及多维度划分法三种。而语体语法的研究是语体理论研究中的重点和热点。吕叔湘（1977）、朱德熙（1987）、胡明扬（1993）先后指出汉语书面语驳杂不纯，从研究方法上提出应当区分不同语体进行语法研究。随着功能语法的发展，陶红印（1999）提出“以语体为中心的语法研究具有重大理论意义”，张伯江（2007）论证了语体意识在语法研究中的重要性。近些年来关于语体语法理论问题更是取得了一系列成果，冯胜利、施春宏（2018）提出语体语法的概念，并建立了语体语法的基础理论，揭示了语体的本质，构建出了汉语的语体系统，阐明了语体语法的原理和机制，确立了语体语法的单位系统，并且讨论了语体语法的使用范围与价值。

语体的实证研究也是语体研究的另一个热点。许多学者关注了语体的量化研究，丁金国（2009）利用语料库，建构了语体标记集；刘海涛（2007）、刘丙丽（2012、2013）、张聪（2018）、黄伟（2017、2018）和邵沁清（2021）等从词语和句子的各个角度，提取了不同的语言特征，进行了统计和分析，从而说明不同语言特征在语体分类的有效性，分析了口语和书面语之间的差异。另一类语体的实证研究则是关注语体的各类资源，主要有语料库和词典两类。北京语言大学 BCC 语料库包含文学、报刊、对话、古汉语等多领域；北京大学 CCL 语料库中也包含了法律、文学、报刊、口语等多个类型；冯胜利等（2017）构建了由叙事文等 6 类文体组成的 12 万字左右的语体语料库，从语法、韵律、语体信息三方面进行标注。《现代汉语词典》标注<口><方><书>，分别表示口语、方言、书面上的文言词语，闵家骥等（1991）编著有《汉语方言常用词词典》，施光亨等（2012）编著有《汉语口语词词典》，冯胜利等（2020）编著有《汉语八百对单双音节对应词词典》等。

通过上述对语体研究的梳理，可以发现语体理论研究众多，但对其进行落地实践则较少，十分有必要对现有的理论进行验证。而实现语体理论的应用，则需要将语体引入到具体的语言形式，可以表现为词和词组，但这还远远不够，而构式相对而言则是一个更好的承载单位。同时，语体的实证研究中，专门的语体语料库资源极少，且规模不大，并且运用词典中存储的一些词语是远远不够的，语言单位需要革新，语料库也要随之更新，因此建立语体构式资源则是十分有必要的了。

2.2 构式语法

构式语法自 20 世纪 80 年代兴起，后得到迅速发展。构式语法为语言研究提供了一种新的视角和新的方法，对于汉语研究也是有着重要的意义，可以用来解释一些先前不好解释、或先前想不到去解释的语言现象。近些年构式语法研究持续受到关注，主要集中在理论研究和应用研究两方面。

美国著名认知语言学家 Goldberg（1995）对于构式语法的理论建构影响重大，她提出了影响最大的构式定义，提出了构式语法分析法，而 Fillmore、Kay&O' Connor（1988）对关联词“let alone”的句法、语义、语用特征做了精细的刻画，引起了人们对习语性构式的广泛关注。严辰松（2006）对构式语法的系列理论进行了整合，较为详细地阐释了什么是构式语法和构式，指出了构式语法的理论和实践价值及其存在的问题。陆俭明（2008）也肯定了构式语法理论的贡献，同时指出这种理论还很不完善。而构式语法理论经过发展在不断完善，Goldberg（2006）提出了构式概念的修正版，认为只要有足够的使用频率，即便是完全可预测的结构型式，也是构式。近些年国内学

者也在持续推进构式理论的研究，陆俭明（2016）从构式如何产生、构式理论与组块理论的结合以及构式与语言信息结构的关系出发提出了三点思考，施春宏（2022）从构式语法的基本理念和底层逻辑出发，对构式研究中重大而基础的相关议题作出了结构化、系统性和一致性的阐释，对构式理论进行了梳理。而在第三届构式语法研究高峰论坛中，展示了构式语法在理论与应用、共时与历时、汉语与外语等领域的前沿成果，可见构式语法研究众多，发展较快。

关于构式语法的应用，施春宏（2017）、郝曦（2022）先后进行了探讨。构式语法在语言教学和语言习得、语言知识库建设、辞书编纂、自然语言处理这几大领域中都能够得到应用，并且还以翻译为例探索了构式语法在其他领域的应用空间。而对于构式资源的实践则是少之又少，冯胜利、施春宏（2015）专著中的“三一语法”，与构式语法一样，所持的教学语法观是大语法观，即将语法和词汇视为一个连续统，甚至对两者并不做出严格的区分，而将所有语言单位都看作作为形式和意义/功能配对体的构式；陆俭明先生承担了教育部人文社会科学重点研究项目“现代汉语构式知识库建设及其应用研究”，已经构建了1,114条的现代汉语构式知识库，将为构式资源库的建设和开发以及构式理论知识的挖掘提供重要的参照。

构式语法尽管研究历史不长，但是已经产出了丰富的成果，已经形成了一套比较完善的理论，但是应用方面还存在一定的局限，主要只关注新型构式的产生和有限的既有构式形成机制等问题。而更广泛的语言生活中的问题，还有待与构式理论相结合进行探讨。本文则关注了语体问题，语体关注语言的场景信息，与语言使用联系密切，体现语体信息的语言结构还有待于进一步观察。构式唯有与语体信息相结合，才能够切实地进入到语言使用的现实场景中。

3. 语体构式

3.1 语体构式的存在

将语体与构式结合，产出语体构式。这是一个构想，那么语体构式是否有理论或实践依据？以下从两方面来对此进行论证。

第一，语体在语言中表现得形式是多种多样的，那么构式能否作为语体的承载体呢？

语体的表现形式主要包括了语体词、语体句、语体辞式和语体章法等一整套体系，而语体词是对言语成品进行语体切分的最后结果，是对言语成品进行语体分析的最小单位。那么语体词作为最为直观表现的单位，是能够进行形式化描写的，易于构建语体资源进行存储和调用。黄国敬（2022）一文基于《大词林》，完成了“语体词标注”“语体（词）链条标注”和“平行构式标注”三个任务，建立了以语体词为基础的语体分类资源。但是，仅仅构建语体词的资源是远远不够的。那么，以语体词为基础进一步构建构式资源是有其必要性和可行性的。构式指的是形式意义对，构式不仅仅是语言中不规则的习语、熟语等，也包括抽象的句型，甚至语素（含词缀）和词。语言中各种规约化的“形式—意义/功能”结合体都是构式，构式存在于语言的各个层面，内涵十分丰富。而语体构式可以既包括语体词的实体单位，也包括抽象的图式单位。例如，“这NV的”结构便是现代汉语口语中常用的一种构式，它的基本语义功能是表达说话人对当前语境中的某种事物或某种事态进行感叹性评价。可见，语体构式可以承载丰富的语言信息，并且有许多的实践和证明，是可以进行进一步的观察和研究的。

第二，构式存在多种类型，范围极广，那么是否存在“语体构式”呢？

《语言学名词》中将“语体”定义为语言的社会功能变体，认为它是为适应不同社会活动领域中的交际需要所形成的、具有一定功能风格特点的语文表达体式。目前学界对于语体的定义仍未达成共识，但大多数学者都认同语体在交际功能上的属性，因此可以认为语体是具有语用功能的。而构式最经典的定义是：所谓构式，就是指这样的形式—意义对，它在形式或意义方面所具有的某些特征不能从其组成成分或业已建立的其他构式中完全预测出来。从这一定义出发，早期构式语法还认为，“构式意义既是语义信息，也包含焦点、话题、语体风格等语用意义，所有这些与构式的关系都是约定俗成的，是构式本身所具有的表达功能。因此，即便是可用语法规则推

理得出的句式，如果其语用意义特殊，它也同样属于构式。”该观点是能够完全印证语体构式是可以存在的，并且有独特的价值。

经过二十多年的发展，施春宏在经典定义的基础上，认为“构式指的是规约化的形式意义对。”可以看出将规约化作为构式的主要特征，而“不可预测性、非完全预测性”只是构式之所以成为构式的充分条件而非必要条件。构式语法的知识观，强调的是知识的固化过程及其规约性。那么，相对于词语句子，篇章级的语体中具有规约化的形式意义对也是能够纳入构式范围的。施春宏也多次提出：“只要是规约化的形义配对体，就具有了构式性，就能获得构式地位。语素、词、习语、语块、句式、语篇乃至语体、文体等，都具有形义结合的依存性和特定性，都是通过固化和规约化而形成的交际单位，因此都是构式。”这一论述也与上文中语法规则推理得出的句式而语用意义特殊仍然可以成为构式，是一脉相承的。

3.2 定义与内涵

语体构式，可以简单理解为，以表达语体功能为落点的构式，它带有某种语体特征。以下从语体构式的特点进一步论述。

第一，语体构式作为构式的一种，满足构式的基本特征。语体构式是带有语体特征的规约化的形式意义对，一方面具有规约化的特征，是人们意识固化的可以接受的形式意义对，而构式的表现不是临时性的，而是经过人们意识固化过程后可以认知的。另一方面，构式是形式意义对，既需要有一定的形式表现，这可以是语素、词、习语、语块、句式等等，也可以用其他的形式化描写。也需要有其内在意义，这个意义可以是多重的，例如对于词语来说，既包括词语的本义，也包括词语的色彩意义（其中就包括语体意义）。

第二，语体构式强调构式的语体特征。正如前文论述，构式的内涵是包罗万象的。语体构式作为构式的一种，最为突出的特征便是它的语体特征。一方面，语体特征是可以透过母语者的语感直观感受得出的，母语者自然能够分辨出该构式是否适用于某一语体中。另一方面，语言学研究者也进一步总结出语体的语言学表现，如典型口语词与书面词有其特殊的表现特征，语体实现的语法手段包括有语言的句型变换、句法移位以及韵律特征等等。如词语对“妈妈-母亲”，两词的形式和意义都是固定的，同时表达了明显的语体功能意义，还是典型的口语词和书面语词，便可以算做语体构式。

第三，语体构式有其丰富的内涵。首先，语体构式并不是凭空而出，是多方面共同作用的结果。语体构式并不是只含有语体特征，它是语法、词汇、语义、韵律等多方面共同作用的结果，需要借助词汇的形式、词法句法手段、韵律特征来表现出来，并且包含有多重的语义。其次，语体构式体现的语体特征是丰富的。语体的分类是二分法、三分法以及多维度分法，使用不同的分类方法，可以得出不同的语体构式，如使用二分法，便可得出非正式语体构式、正式语体构式。最后，语体构式表现的形式是多种多样的，既可以是词语、短语、习语一类确定的实体单位，也可以是包含有变项的词语、短语组合而成的半实体半图式单位，还可以包括包含变项的抽象句式而形成的图式单位。如非正式语体构式“编+np[表示作品]”，正式语体构式“编写+np[表示作品]”²，其中包含有词语、符合语法规则、表现出了韵律特征，除本身语义外，还包含语体意义，是多种因素共同发挥作用而形成的，是一种半实体半图式单位。

第四，语体构式具有组合性，主要体现在内在和外在两方面。内在的组合性指的是，实体单位、半实体半图式单位以及图式单位的语体构式都是组合而成的，是由常项或常项和变项组合而成，即语体构式是由其他单位组合而成的。外在的组合性指的是，语体构式的构件是由更低一级的单位组合而来的，语体构式也将与其他单位组合形成句子或篇章等更大的单位。即组成语体构式的词语、短语等都是由更低一级的实体单位组合而来的，并且语体构式并不是语言运用的最终形式，它仍然需要与其他单位组合才能够使用在具体的语言环境中。如语体构式“造房子”和“建造房屋”是由语体词“造”“建造”“房子”“房屋”作为常项组合而成，

² “np”表示名词性短语。构式举例如“编歌”“编话剧”，“编写歌曲”“编写话剧”等。

第五，语体构式通常并不完全具有递归性。组合的递归性指的是语法结构中的某个单位可以不断地被一个同功能的结构去替换和扩展，结果使基本结构中的项扩展能很长的多层套合结构，但作用等于原来的那个项。而对于语体构式，其中的构件的替换并不是能够不断产生的，是有限的，而且每产生一个新的构件的替换都将作为一个新的语体构式。语体构式的基本不能扩展，若特殊情况的扩展基本局限于某些副词或者虚词单位，也是有限的，也将作为新的语体构式存在。

第六，语体构式表现为一种线性序列。传统的语法结构经常使用树状结构，但语体构式基本不需要再进行内部的层层切分，它是由变项或变项与常项组合而成，内部没有层层嵌套。其表现形式也是各项单位排列而成的线性序列。

值得注意的是，语体构式与常规的短语结构语法组合相比，最重要的特征是有语体功能表达；并且相对于一般的动宾结构等，语体构式的表达则会更加具体；同时，语体构式是包含带有语体特征的常规短语结构语法组合，仍然包含确定的词语和短语，也包含有变项且不按照常规语法规则形成的短语结构。语体构式与语体词相比，语体构式是包括语体词的，语体构式可以由语体词构成，可以是单独的语体词，语体词与语体词组合，也可以不完全由语体词构成，语体词与词类组合成语体构式，还可以是完全不由语体词构成，这样语体构式也是可能存在的。在我们的实践和观察中，大部分的语体构式都是包含有语体词的。

4. 平行语体构式资源构建

通过对语体构式的探索，可以进一步尝试构建语体构式资源。本文则选择构建了平行语体构式资源，做出这项选择的原因主要是因为，一方面语体构式资源是一个广泛的资源，先从某一个角度出发尝试构建是符合基本的方法论要求的；另一方面，虽然语体构式资源可以是多种类型的，而构建同义的平行语体构式资源具有可行性和明确的适用场景：首先，某种语体并非单独存在的，是与其他语体对立共同存在的。那么构式若要表达相同语义，在不同语体中也是有不同的语体构式的；其次，在操作性上，产出某一语体的构式，对应产出该语义下的其他语体构式是自然而然可能发生的，并且语义的限制提供很好的条件；最后，相同语义的平行语体构式资源可以很好地服务于语言教学和语言研究等领域。

本研究的平行语体构式举例如下：“打”后接名词性短语，在名词性短语表示织物时，“打”可以转化为“织”。本研究中将“打+np”作为构式，当“np”指代“织物”时，与“织+np”组成平行构式。

4.1 数据来源

本文的目标是构建相同语义的平行语体构式资源，目前学界相关的直接资源较少，因此最终选取了语言生活中的实际语料来进行尝试。数据主要来源于两部分，一部分是从语体语料库中逐步统计出的口语词，另一部分通过语言模型生成的语体语句抓取语体构式。

借用了已经构建的语体语料库³，将口语与书面语的分词结果取交集，将口语词的结果与上述交集再取差集，选取频率大于5的词语，经过人工筛选后，初步得到了口语词词表。通过搜索语体资源，进一步补充了已有词典的方言词和口语词，最终得到了8,418个口语词，以这部分口语词为基础，进行了语体构式的构建。

语句的语料来自于人民日报，将部分《人民日报》作为训练数据，借用生成改写模型进行训练，试图通过改写提高正式程度，再选用另一部分数据运用转写模型自动生成4,000个语句。

4.2 资源构建规范与实践

构建资源总的原则在于：（1）保证构式成立，即该构式能够是规约化的形式-意义对；（2）是表达语体功能的语体构式，满足其上述特征；（3）完成平行的语体构式，即表达同一语义的语体构式是对照出现的。

³ 参见 Qinqing Tai and Gaoqi Rao. 2021. 汉语语体特征的计量与分类研究(A study on the measurement and classification of Chinese stylistic features). In Proceedings of the 20th Chinese National Conference on Computational Linguistics, pages 398 - 412, Huhhot, China. Chinese Information Processing Society of China.

4.2.1 资源构建规范

第一，主要将语体分为正式语体与非正式语体，这也意味着语体构式是成对出现的。崔希亮（2020）提出“正式与非正式是两端清晰中间模糊的一对概念，我们要探寻正式语体和非正式语体的分界线，只能从两端入手。”⁴我们认同这种说法，因此选取具有两端语体特征明确的非正式语体构式和正式语体构式是更具有操作性的。

第二，语体构式语义一致。正如之前所述，构建同义的平行语体构式资源具有可行性和明确的适用场景，本文的目标是构建相同语义的平行语体构式，因此语体构式的语义是基本一致的，即互相转换使用是基本可以接受的。

第三，以词、短语和句式作为构式的基础。语体构式有实体单位、半实体半图式单位、图式单位，但构式的形式仍然需要以词语、句式等为载体，并且语体构式的建立初步可以借鉴传统语法研究已经比较成熟的词语、短语和句式，争议更少，更容易为人接受。

第四，以词类和短语功能类作为其中的变项进行填充。仅仅以词语或短语构成的实体单位有限，并且效率不高。那么，借用词类表达可以组成半实体半图式单位和图式单位，可以极大地扩充资源。

第五，语义辅助解释平行语体构式。词类限制已经是相对有效的手段了，但仍有一部分情况需要更多的限制。为了保证平行语体构式成立，可以对词类进一步限制，即语义，语义的范围相比词类较难进行形式化描述，因此语义解释会配以实例集合，从而使得构式清晰准确。

第六，可使用韵律特征来辅助构式。冯胜利在韵律语法领域研究颇深，正式语体与非正式语体韵律格式不同，因此单双音节的选择也可以辅助语体构式的构建。

第七，例句作为实例便于理解及验证。每一个语体构式都需要有对应的例句，可以对构式进行验证，也丰富了资源本身。

4.2.2 平行语体构式资源实践

基于不同来源的数据，资源构建一方面从口语词本身出发，寻找典型例句（例句主要来源于BCC或CCL语料库），寻找对应的书面语词，进而构建语体构式；另一方面直接从已有语句出发，在已有语句的基础上，运用已有的语体词资源，建立以语体词为基础的构式，验证构式，进行修改或者限制，再次验证，直至无误。总的实践步骤是从例句出发，提取或创作典型语体构式，以典型语体词或短语为基点确定对应另一语体的构式，进行并完善形式化描写，语体构式和平行语体构式构建完成均需要验证，验证不成立则需要修改直至成立，无法修改则跳过，最终完成完整的平行语体构式。

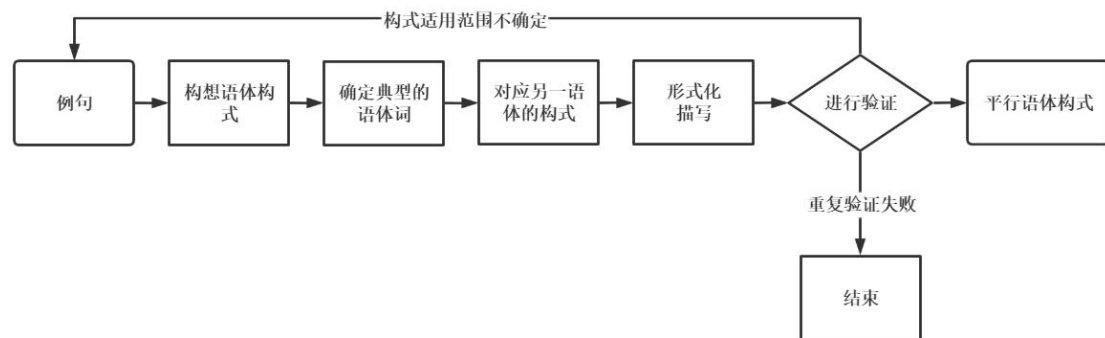


图 1 平行语体构式资源实践流程图

第一，语体构式本身要成立。从例句出发提炼出构式，这与构式的规约化特征相关，构式使用中的知识是逐步概括出来的知识，而例句甚至多个例句则便于提取其中的语体特征且更加可靠。构式经过验证是必不可少的。还需要进行形式化描写，而词性，语序，韵律都可作为辅助手段，

⁴ 参见崔希亮.正式语体和非正式语体的分野[J].汉语学报,2020(02):16-27.

具体可表现为词性符号、连接符号、句首句尾标记符、单双音节符等等。

第二还要满足其平行构式成立。选取典型的语体词后，则需要确定词语的义项，这主要依靠《现代汉语词典》，还可参考百度汉语、汉语大词典等进行补充。进一步可以采用词语或短语级别的替换，便于直接生成对应另一语体的构式。由于相同语义的要求，那么词性限制不足则需以语义解释或语义实例集合进一步限制。

实际操作中，一般会选用例句中的语体词以及口语语体作为起点，主要是可供参考的学术理论资源和已有的语体词基础资源比较丰富，基于语体词的资源更容易构建，标注员也便于操作，也可以提高例句验证的有效性。当然在语言生活中可能还存在很多不由语体词构建的语体构式资源，留待以后的工作中来完成。另外，需要标注员有较好的语言学基础，对于语言的敏感能够更好地辅助完成标注任务。

4.3 资源分析

经过两轮的标注和质检，最终得到了 5,233 对平行语体构式。以下将从资源的表现，描述和分析资源的基本概况。

数据来源	口语词	自动生成语句
数量	8418	4000
产出构式	4755	478
百分比	56.5%	12.0%

表 1 不同来源数据产出平行语体构式的表现

第一，从不同来源的数据来看资源的生成，如表 1 所示。可以看出，不同的数据来源产出平行语体构式的比例都不是非常高，这也说明这项任务是非常有难度的，要经过重重筛选才能完成。而从口语词出发最终产出的结果是远远高于语句产生的结果的，主要是因为自动生成的语句本身质量一般，并且《人民日报》本身的正式程度已经比较高，而且从语句中筛选典型语体构式也是有一定难度的。

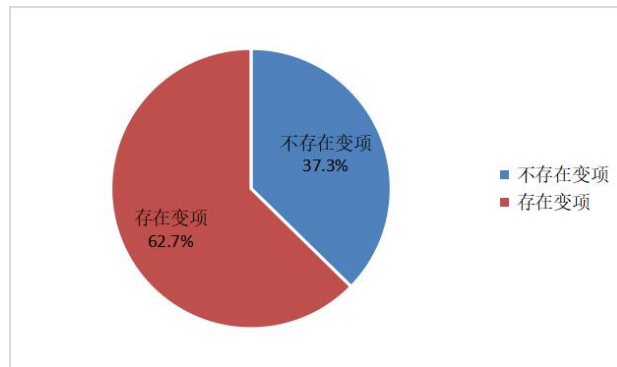


图 2 平行语体构式是否存在变项

第二，从是否有变项来看资源的情况。如图 1 所示，产出的 5,233 对平行语体构式中，其中 1954 对构式不存在变项的，占比 37.3%，这些构式中都是词语或者短语，如：要不是=>若非，得了绝症=>身患绝症；其中 3,279 对构式是存在变项的，占比 62.7%，这些构式中常项与变项用“+”连接，常项则是词语或短语，变项则是词类以及需要的语义限制，如：np+心里不是滋味=>np+心里五味杂陈。

	存在语义限制	不存在语义限制	总计
变项构式对数量	744	2535	3279
占变项的百分比	22.7%	77.3%	100%

表 2 存在语义限制变项的语体构式对

进一步统计发现，在存在变项的 3,279 对平行语体构式中，有语义限制的变项构式对有 744 个，占比 22.7%，如：有+np=>患有+np，np=疾病相关，np=胃病/心脏病/慢性病……；不需语义限制的构式对有 2,535 个，占比 77.3%。需要语义限制的构式对存在，但比例并不高，这是因为在实际标注过程发现，常项与变项进行组合时，符合语法规则的同时，一般也会自动限制了语义，如：请+np+帮忙=>请+np+协助，“np”自动就带有“与人相关”的语义限制，不再需要人工总结。

	存在韵律限制	不存在韵律限制	总计
变项构式对数量	75	3204	3279
占变项的百分比	2.3%	97.7%	100%

表 3 存在韵律限制变项的语体构式对

另外，还对含有韵律限制的变项构式对进行了统计。在存在变项的 3,279 对平行语体构式中，有韵律限制的变项构式对有 75 个，占比 2.3%，如：a{1}+喇叭叭=>非常+a{1}；不需语义限制的构式对有 3,204 个，占比 97.7%。数量极少的原因在于构式对中的音节需要统一，这样的对应更加严格有效，更易于转化。而因音节不对应无法构成构式对时，可以将典型实例存储为常项。如：怪自己=>自责。对比而言，在存在变项的构式对中，语义限制的作用还是略微大于韵律限制的作用。韵律特征更适用于仅存在常项的构式直对。

构件数量	1	2	3	4	≥5
非正式语体构式	1965	1993	916	258	101
累计频率	37.6%	75.6%	93.1%	98.3%	100%
正式语体构式	1965	2097	876	233	62
累计频率	37.6%	77.6%	94.4%	98.8%	100%

表 4 不同语体构式构件数量分布

第三，从构式由几部分构件组成来看不同语体的构式表现。仅存在一个常项的语体构式，构件数量为 1，如：不大会儿、不够意思等，由一个常项和一个变项组成的语体构式，数量为 2，由一个常项和两个变项组成的语体构式，数量为 3，依此了推。如：要多+a+有多+a，有 4 个构件，十分+a，有 2 个构件。通过统计，如表 3 所示，可以发现构件数量为 3 时，覆盖的不同语体构式都已超过 90%，而构件数量为 4 时，覆盖的不同语体构式则已超过 98%。可以发现，在现代汉语中，当一个构式的构件在 3 个或 4 个时，基本就可以表达出一定的语体信息了。另外，通过加权计算，非正式语体构式的构件数量平均约为 1.97，正式语体构式的构件数量平均约为 1.92，二者相差不多，说明语体差异与构件数量的多少关系并不直接。

5. 应用空间

语体构式以及平行语体构式资源是有广泛的应用空间的，可以应用于语言本体研究、汉语教学、自然语言处理以及语言服务等领域。宏观的观念及具体的资源

第一，语体构式可以为语言研究提供新的视角。本研究阐述了语体构式的概念及其特征，丰富了构式语法体系，并且在构式语法理论的指导下，构建了平行语体构式资源。而且，语体构式的观念，为语言本体研究提供了新视角，可以用来尝试解释一些不易解释的语言现象。

第二，语体构式能够辅助汉语教学。在语文教学当中，培养学生的口头表达能力与书面表达能力至关重要。运用语体构式的观念进行教学，能够凸显语体的重要性，促进学生语体意识的培养，形式意义相结合的构式也便于教师教学，提高教学质量。同时平行语体资源提供了在不同语境中不同的表达方式，帮助教师进行查询使用，并且可在教学过程中不断完善

资源。在国际中文教育中，如今国际中文教育与职业教育相结合是国际中文教育的重要方向和任务。而职业带有语体属性，如商务与农业的语言教学自然不同，广泛的语体构式不仅在教学中将形式与意义相结合，而且特别重视在不同语境中语言的使用。平行语体构式资源则提供了丰富的语言教学资源，具体的语言知识资源可为教学提供极大的便利。

第三，语体构式与自然语言处理密切相关。语体构式为不同语体的语言知识提供了形式化描写的方式，扩大了语言知识形式化的范围，能够为自然语言处理提供新的解决方式。平行语体构式资源可以作为数据资源直接应用于众多的自然语言处理场景当中，如自动校对、语体改写、语言润色等工作均需要丰富的语体资源，并且形式化的描写也能够极大地提高任务的效率和准确性。

第四，语体构式还可以与语言服务相结合。应急语言服务具有重要作用，其针对的服务场景类型多样，面向不同的服务场景，则需使用不同的语体，语体构式可发挥重要作用。如公告的内容需要使用正式的语体构式，语言抚慰则要以非正式的语体构式为主。当需要在不同的场景中表达相同语义时，平行语体构式资源即可直接使用，极大地提高语言的传播效率，可及时提供便捷准确的应急语言服务。翻译与本地化中必然会面对语体问题，这些问题不仅仅存在于词语层面，而是构式级别的。语体构式是本地化翻译需纳入思考的，不同的产品和市场需要使用不同的语体构式，才能传达更加准确的信息，提高翻译的质量。

6. 总结

本研究提出了语体构式的概念，认为语体构式表达语体功能的规约化的形式意义对，其具有构式的基本特征，凸显语体特征，内涵丰富，具有组合性的特征，通常不具有递归性，表现为一种线性序列。并且初步完成了平行语体构式的资源构建任务，描述了资源构建的规范和实践过程，得到了 5,233 对平行语体构式，对数据进行了统计和分析。之后，探索了语体构式的应用空间，语体构式及其资源可在语言研究、汉语教学、自然语言处理以及语言服务等领域有所助力。

今后语体构式仍然需要更加深入的研究。语体构式的内涵需要进一步探索，发掘完全不含语体词的语体构式，继续夯实语体构式的理论基础。当以语体构式的视角思考语言问题时，可对语体构式与其他语言本体现象有新的认识与思考，并且语体构式需要紧跟语体与构式理论的发展，用语体构式的视角进行解构和研究。语体构式资源也需要做更多的扩充。平行语体资源还可以进一步补充，扩大数据的规模并且不断提高数据的质量。还需要构建更多类型的语体构式资源，充分利用资源提供更广阔优质的服务。

参 考 文 献

- [1] Fillmore,C.,P.Kay&M.O'Connor. Regularity and idiomaticity in grammatical conditions: the case of LET ALONE[J].Language64(1988):501-538.
- [2] Goldberg,A.E. Constructions:A Construction Grammar Approach to Argument Structure [M]:Chicago:Chicago University Press,1995.
- [3] Goldberg,Adele E. Constructions at Work: The Nature of Generalization in Language. Oxford: Oxford University Press,2006.
- [4] Guojing Huang, Liwei Zhou, Gaoqi Rao, and Jiaojiao Zang. 2022. 基于《同义词词林》的中文语体分类资源构建(Construction of Chinese register classification resources based on "Tongyici Cilin"). In Proceedings of the 21st Chinese National Conference on Computational Linguistics, pages 431-443, Nanchang, China. Chinese Information Processing Society of China.
- [5] Qinqing,Tai and Gaoqi,Rao.2021.汉语语体特征的计量与分类研究(A study on the measurement and classification of Chinese stylistic features).In Proceedings of the 20th Chinese National Conference on Computational Linguistics, pages 398-412, Huhhot, China. Chinese Information Processing Society of China.
- [6] 崔希亮.正式语体和非正式语体的分野[J].汉语学报,2020(02):16-27.
- [7] 丁金国.基于语料库的语体风格研究——兼论量化与质化的关系[J].烟台大学学报(哲学社会科学

版),2009,22(02):114-120.

- [8] 房旭.近十多年汉语语体语法研究述评[J].语言研究集刊,2022(01):104-116+401-402.
- [9] 冯胜利,施春宏.论语体语法的基本原理、单位层级和语体系统[J].世界汉语教学,2018,32(03):302-325.
- [10] 冯胜利,施春宏.三一语法:结构·功能·语境--初中级汉语语法点教学指南[M].北京大学出版社,2015.
- [11] 冯胜利,王永娜,王丽娟.汉语八百对单双音节对应词词典[M].Phoenix Tree Publishing Inc,2020.
- [12] 冯胜利,王永娜.语体标注对语体语法和叙事、论说体的考察与发现[J].汉语应用语言学研究,2017.
- [13] 冯胜利.汉语语体语法概论[M].北京语言大学出版社,2018.
- [14] 郝墩,施春宏.构式语法的应用空间探析[J].语言文字应用,2022,(01):17-30.
- [15] 胡明扬.语体和语法[J].汉语学习,1993,(02):1-4.
- [16] 黄伟,刘海涛.词频分布参数可以细分汉语新闻语体吗?[J].语言教学与研究,2017(04):79-86.
- [17] 黄伟.基于词频的现代汉语语体计量研究.In 计量语言学研究进展[M].浙江大学出版社,2018.
- [18] 李熙宗.关于语体的定义问题[J].复旦学报(社会科学版),2005(03):176-186+196.
- [19] 刘丙丽,牛雅娴,刘海涛.汉语词类句法功能的语体差异研究[J].语言教学与研究,2013(05):97-104.
- [20] 刘丙丽,牛雅娴,刘海涛.基于依存句法标注树库的汉语语体差异研究[J].语言文字应用,2012,(04):134-142.
- [21] 刘海涛.汉语句法网络的复杂性研究[J].复杂系统与复杂性科学,2007,(04):38-44.
- [22] 陆俭明.对构式理论的三点思考[J].外国语(上海外国语大学学报),2016,39(02):2-10.
- [23] 陆俭明.构式语法理论的价值与局限[J].南京师范大学文学院学报,2008,(01):142-151.
- [24] 吕叔湘.通过对比研究语法[J].语言教学与研究,1992(02):223-231.
- [25] 闵家骥.汉语方言常用词词典[M].浙江教育出版社,1991.
- [26] 钱一华.汉语语体研究回顾与前瞻[J].现代语文,2022(02):77-83.
- [27] 施春宏,蔡淑美.构式语法研究的理论问题论析[J].外语教学与研究,2022,54(05):643-655+798.
- [28] 施春宏.构式的观念:逻辑结构和理论张力[J].东北师大学报(哲学社会科学版),2016,(04):17-23.
- [29] 施春宏.构式三观:构式语法的基本理念[J].东北师大学报(哲学社会科学版),2021,(04):1-15.
- [30] 施春宏.构式语法的理论路径和应用空间[J].汉语学报,2017(01):2-13+95.
- [31] 施春宏.语体何以作为语法[J].当代修辞学,2019(06):1-20.
- [32] 施光亨.汉语口语词词典[M].商务印书馆,2012.
- [33] 邵沁清.汉语语体特征的计量与分类研究[D].北京语言大学,2021
- [34] 唐松波.文体、语体、风格、修辞的相互关系[J].当代修辞学,1984,(02):32-33.
- [35] 唐雪凝,张金圈.表感叹性评价的“这 NV 的”构式分析[J].语言科学,2011,10(02):182-189.
- [36] 陶红印.试论语体分类的语法学意义[J].当代语言学,1999(03):15-24+61.
- [37] 王寅.构式语法研究[M].上海外语教育出版社:当代语言学丛书,2011.
- [38] 王永娜.现代汉语语法的语体研究之发展[J].烟台大学学报(哲学社会科学版),2021,34(06):101-114.
- [39] 严辰松.构式语法论要[J].解放军外国语学院学报,2006,(04):6-11.
- [40] 张伯江.语体差异和语法规律[J].修辞学习,2007(02):1-9.
- [41] 张聪,刘海涛.词频计量指标与汉语语体演化[J].外语教学,2018(02):19-24.
- [42] 曾毅平,李琳.当前语体与语言风格的阐释、描写、应用研究——第十届语体风格学学术研讨会综述[J].当代修辞学,2015,(02):94-95.
- [43] 赵博文,张伟丹,刘文秀.构式语法的前沿探索和理论交融——基于第三届构式语法研究高峰论坛的成果概述及相关思考[J].语言教学与研究,2022(05):108-112.
- [44] 周毅,王建华.九十年代以来报刊语言研究综述[J].浙江传媒学院学报,2006,(01):5-7.
- [45] 朱德熙.现代汉语语法研究的对象是什么[J].中国语文,1987(5).
- [46] 朱军.汉语语体语法研究综述[J].汉语学习,2012,(05):72-79.