

# 大语言模型辅助的功能语篇分析： 理论、方法与实践\*

北京师范大学 于 晖 宋金戈 王 乐

**摘 要** 语言智能技术的发展推动外语教学与研究进入“大数据”时代, ChatGPT 的出现更是将自然语言生成与处理的语言学理论探究推至新高度。但是, 鲜有研究将 ChatGPT 的“预训练式深度学习能力”系统应用于功能语篇分析。鉴于此, 该研究基于系统功能语言学视阈, 尝试建构 ChatGPT 辅助的功能语篇分析工作模型, 并对《荒野的呼唤》进行生态话语分析操作演示, 旨在为大规模语料的功能分析与处理提供参考。

**关键词** 大语言模型 人工智能 语篇分析 语言功能 生态话语

**中图分类号** H319.3 **文献标识码** A **文章编号** 1001-5795(2024)05-0043-0009

**DOI** 10.20139/j.issn.1001-5795.20240507

## 1 引言

语言功能指语言在使用中所实现的目的和作用, 是语言系统的本质属性( Halliday, 1978), 因此基于语言功能的分析和应用成为外语教学与研究的重要路径。但由于语言功能的语境依赖性和语篇分析的主观性, 大规模文本的功能分析与处理始终是一个难题( Matthiessen *et al.*, 2022)。2022 年末, 美国人工智能实验室 OpenAI 推出大语言模型( Large Language Model, LLM) ChatGPT。一经发布, ChatGPT 便在全球激起对科技和学术伦理问题的广泛关注和探讨( Graf & Bernardi, 2023; 令小雄等, 2023), 也因其强大的信息处理和语言生成能力, 在计算机与智能科学、教育学、政治学、人类学和医学等领域引发“狂飙式”影响( 胡加圣、戚亚娟, 2023)。

目前, 语言学领域主要从理论和应用视角对 ChatGPT 展开研究。一方面, ChatGPT 通过庞大的模型参数和训练数据而涌现出强大的自然语言理解和生成能力( Bommasani *et al.*, 2022), 甚至挑战和驳斥某些当代主流语言学理论和方法( Hanson, 2023; Piantadosi, 2023), 但尚未有语言学理论能够对其底层原则给出合理且充分的解释( 袁毓林, 2023)。因此, 不同的语言学分支和流派从其语言生成机制

( Chomsky *et al.*, 2023; 冯志伟、张灯柯, 2023)、多语分析能力( 刘海涛、亓达, 2024; 秦洪武、周霞, 2024)、哲学思辨( Wang *et al.*, 2023; 张杨, 2023) 和理论启示( Ratten & Jones, 2023; 杨敏、王亚文, 2023) 等角度展开探讨。另一方面, 由于其卓越的语义理解和常识推理能力, 以及易操作的对话式人机互动模式( OpenAI, 2022), 依托 ChatGPT 的外语教研与改革( Kohnke *et al.*, 2023; 胡壮麟, 2023)、对外汉语教学( Cai, 2023; 刘利等, 2023)、机辅翻译( Sahari *et al.*, 2023; 耿芳、胡健, 2023)、性能测试( Li *et al.*, 2023; 赵雪等, 2023)、语言治理( Ferrari *et al.*, 2023; 饶高琦等, 2023) 等应用研究丰富, 但却鲜有研究涉及功能语篇分析。鉴于此, 本研究基于系统功能语言学理论和生态话语分析方法, 尝试建构 ChatGPT 辅助的功能语篇分析工作模型, 以此促进基于大语言模型的语言功能分析与实践。

## 2 理论基础

### 2.1 ChatGPT 的工作原理和技术解构

GPT 主要依靠预训练模型( pretrained model) 和转换器模型( transformer model), 融合语言处理、机器学习和深度学习等认知计算技术, 从而实现对话式人工智能生成内容( AI-Generated Content, AIGC)。

转换器模型是 GPT 发展演进的技术基础,其架构在大多数神经序列转导模型“编码—解码”机制的基础上,可以避免循环和卷积,进而显著促进并行计算效能,提升文本翻译质量(Vaswani *et al.*, 2017)。基于转换器架构的深层神经网络,ChatGPT 得以初步具备强大的序列任务处理能力和语义特征提取能力,知识储存能力显著提升(苏祺、杨佳野, 2023)。

同时, ChatGPT 具备敢于质疑、承认错误和无知以及大幅度提升对用户意图的理解能力等类人性特征,所依靠的主要算法是基于人类反馈的强化学习(reinforcement learning from human feedback, RLHF)机制,这是一种通过训练和微调大语言模型以指导其自身智能系统行为的方法。训练过程可分为三个步骤(OpenAI, 2022): (1) 收集演示数据并制定监督策略。该步骤旨在基于所收集的人工标注数据,通过监督式微调(supervised fine-tuning, SFT)预训练一个语言模型。(2) 收集比较数据并训练奖励模型(reward model)。奖励模型的建构旨在评估初始语言模型的输出并反馈奖励信号。人类标注者将对基于提示学习(prompt learning)方法所得的多对“提示—回答”数据,从而确定根据人类感知的梯度序列。(3) 针对奖励模型使用近端策略优化(proximal policy optimization, PPO)强化学习算法优化策略。该步骤旨在创建一个反馈循环以训练和微调初始语言模型副本的部分或全部参数。在该过程中,第二阶段的奖励模型将不断评估所生成的回答并为其打分。根据迭代过程中不同回答的奖励分数,于人类而言更为可取的回答将会被优先输出。总而言之,预训练模型是 ChatGPT 得以支持多种自然语言理解和生成的基础,提示学习方法是促进任务高效且正确完成的中枢,而基于人类反馈的强化学习机制则使这个大语言模型具备类人的思维方式和价值取向。

## 2.2 语言的结构、系统和功能

语言的本质和本体定位是语言研究的核心问题,但不同的语言学研究范式对其认识不一。系统功能语言学认为语言的本质是一个通过语义选择形成的动态系统网络,属于社会符号系统(Halliday, 1978),而语言研究的理论建构应围绕语言的结构、系统和功能展开(Halliday, 1994)。换言之,系统功能语言学虽主张系统第一性,即纵聚合关系是语言的首要特征,但并未忽视结构对于语言的重要作用,而是从功能出发解释语法,强调系统和结构之间或意义和形式之间

的张力和互补关系(Halliday & Matthiessen, 2014)。

功能语言学通过层次化层级(hierarchy of stratification)和实例化连续体(cline of instantiation)将语言形式的内部解释(语言结构构成的词汇语法规则)和外部解释(语言使用的目的和交际功能)相联系。一方面,系统功能语言学主张语境和语言系统之间是层次化体现关系,即语境可由语言系统体现。语言系统包括语义层、词汇语法层和音系层。因此,词汇语法变异可体现语义变异,进而引发语域变异。而语域作为由语境的一组特定值所决定的“语言功能变体”(Halliday & Matthiessen, 2014),也会对语言选择具有限制作用。另一方面,语言系统和文本之间也可通过实例化相贯通。实例化连续体意指从文本实例到文本类型再到语言系统潜势这样一个连续体,体现了意义不断概括提炼的过程(Halliday & Matthiessen, 2014)。系统位于潜势极,文本位于实例极,文本可被视为语言系统潜势的体现。同时,词汇语法系统及其功能的确定需要凭借它们在结构中的相对位置关系。例如,述位系统及其功能需要依靠小句中的主位进行定义。这也使得结构和功能之间是“一对多或多对一的张力系统”(刘向东、王博, 2021)。故而,任何自然语言文本的理解在本质上都源于语义,而语义则是由纵聚合关系决定的语境内选择,且语言文本的功能可通过观察和参考其结构或形式表征以及所使用的词汇语法资源进行分析。

由于系统功能语言学是以问题为导向的理论体系,其对于语言的结构、系统和功能之间相互关系的理解可满足语篇分析所需,而其对于语言系统本质的认识使其理论架构“在自然语言处理领域的应用更具灵活性和适用性”(黄国文、张培佳, 2020)。另外,语言学理论和语言智能之间存在天然共性,应是相互培养的关系(胡开宝、尚文博, 2022),语言智能技术可以助力基于语言学理论的语篇分析实践,以语言学理论作为支撑也可促进以数据驱动的语言智能发展(韩礼德等, 2011)。因此,基于系统功能语言学理论解构自然语言的资源,并借助 ChatGPT 的算法原理辅助分析其功能,或可弥补人工分析主观性较强和机器分析语境理解能力欠佳等缺点。

## 3 ChatGPT 辅助的功能语篇分析工作模型建构

功能语篇分析是评估趋向的、有步骤的“解释性活动”(explanatory activity),即“解释语篇如何以及为

何具有所表现出的价值”(Halliday & Hasan, 1976)。因为语篇分析不仅是简单的“说明性活动”(interpretive activity),而应是基于语篇形式特征的有理据、由浅入深的过程(黄国文, 2001),故此黄国文(2002)指出功能语篇分析可分为观察、说明、描述、解释和评估这五个步骤。王勇、孙亚迪(2021)基于读者理解语篇的逻辑顺序,并借鉴 Fairclough (1995) 和黄国文(2002)等研究,提出功能语篇分析的步骤和顺序可表现为“观察—描写—说明—解释”。综合以上研究,本研究沿袭 Halliday (1994) 所主张的“评估是位于解释之上的更高层次”,将功能语篇分析的步骤和顺序调整为“观察—描写—说明—解释—评估”。

观察既需要判断语篇是否存在被分析的价值,也要聚焦并检视语篇的语音、词汇语法和功能结构等方面,发现对于建构语篇某一方面意义而言至关重要的形式特征。描写旨在理论刻画形式特征,即对语篇进行词汇语法分析。词汇语法分析是“较低层次的语篇分析”,可为语言意义和功能的理解提供基础(Halliday, 1994)。说明指对语篇的形式特征进行语义功能阐述,是关联语篇形式特征和意义特征的关键。描写和说明不是“随想式、个人经验的”,而应是基于一定理论框架的系统剖析(黄国文, 2002)。解释是基于说明的进一步追问,是对语义特征的深度诠释。这一步骤往往需要借助语篇体裁和语境因素等一般规律,解释语篇表征的个别现象。评估的目的则是考察语篇在哪些方面是成功的,而在哪些方面是有所欠缺的。这个过程不仅需要斟酌语篇意义建构的各个方面,还需要深度探讨语言和语境之间的各种关系(Halliday, 1994)。由此可见,解释和评估是更为抽象的高维分析。因此,囿于 ChatGPT 语言分析的可操作性范畴,并基于功能语篇分析的步骤和顺序,本研究尝试建构 ChatGPT 辅助的功能语篇分析工作模型,详见图 1。

图 1 中虚线内操作需分析者人工进行,实线代表 ChatGPT 可辅助完成的操作。通过对目标语篇进行全面的观察和把握,分析者得以明确其意义建构的重要形式特征,并以此选择相应的功能语言学理论作为分析框架。其后,分析者需预先设定 ChatGPT 的对话角色,并通过知识“投喂”进行初步训练,使其熟悉所选理论的定义、假设、术语和基本应用等内容。经过理论预训练,任务文本可被输入依序进行描写和说明,即词汇语法分析和语义功能分析。两个步骤的分析都需要根据 ChatGPT 的反馈或回答进行调整,通过指令或提示反复聚焦并着重训练 ChatGPT,以提升每次回答的有效性和准确性,最终输出被理论解构的文本。但是,被输出的文本只是“半成品”,仍需经过分析者的解释和评估方可实现完整的功能分析。由于 ChatGPT 辅助的本质是分析者主导的人机互动协商过程(文秋芳、梁茂成, 2024),该工作模型的操作流程可被视为一个持续的、反复的动态循环。

#### 4 语篇功能实例分析

本文基于 ChatGPT 辅助的功能语篇分析工作模型进行实例分析,探讨如何针对大规模文本和不同语言功能进行策略性调整。生态话语分析和系统功能语法在理论渊源、理论基础和价值取向等方面均具有紧密相关性(Halliday, 1990; 辛志英、黄国文, 2013),且语篇的生态意义常伴随概念、人际和语篇意义的动态发展被建构和传递(何伟、张瑞杰, 2017)。因此,本研究基于系统功能语言学理论和生态话语分析方法,对 Jack London 所著的《荒野的呼唤》(The Call of the Wild)进行生态话语功能分析。

经审读和观察,寓言式小说《荒野的呼唤》整体以第三人称全知全能和第一人称有限叙事视角,讲述宠物犬 Buck 被迫回归自然的过程,经验元功能建构资源丰富多样,适用及物性系统理论分析。但是,鉴于及物

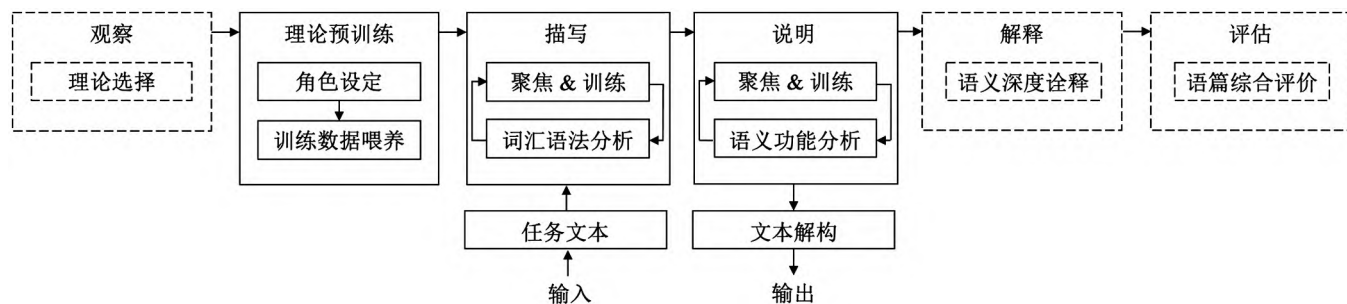


图 1 ChatGPT 辅助的功能语篇分析工作模型



性分析难以深度解释语义表征复杂的小句,且简单小句的及物性分析可辅以 UAM Corpus Tool 等语料库统计软件。本研究尝试基于及物—作格综合模式对目标文本进行生态性分析(于晖、王丽萍,2020)。该综合模式着眼于及物模式和作格模式的互补性,细化过程类型和参与者角色,可以深度解构语言建构经验意义的方式和过程,从而有助于揭示语篇深层生态意义和价值取向。演示借助 ChatGPT 4.0 模型,基本操作流程如下。

#### 4.1 理论预训练

在进行训练数据“投喂”前,为保证 ChatGPT 4.0 回答的质量和效率,往往需要为其预设一个学术身份。虽然 ChatGPT 4.0 的 Research & Analysis 模块已提供一些自带提示词的学术 GPT(如 Scholar GPT),但其数据资源较为宽泛且模糊,不足以满足本研究的分析所需。因此,手动角色预设仍是必要环节。通常情况下,角色预设指令或提示越详细越好,包括角色身份、知识能力、目标任务的内容描述和完成要求等。本工作模型旨在完成功能语篇分析,角色预设需要突显“具备一定系统功能语言学理论知识和实践能力的分析者或研究者”,因此指令可表述为: You are a researcher in the academic field of Systemic Functional Linguistics, possessing extensive research experience and a strong understanding of frontier research trends. Your primary task is to master the theoretical knowledge that I provide and apply it to discourse analysis according to my specifications.

训练数据“喂养”旨在使 ChatGPT 4.0 了解目标理论,并初步具备相应的语言功能分析能力。虽然通过联网和多模态升级,ChatGPT 4.0 的知识获取与整合能力明显提升,但大量且复杂的数据有时也令其回答出现分析错杂、逻辑混乱和结果呈现不合意等问题(Halaweh, 2023; 秦涛等, 2024)。因此,分析者需要注意所提供数据的准确性、完整性和内容一致性,并可通过对数据再加工以提升 ChatGPT 4.0 的理论理解深度和语义推理信度。在本研究中,分析者首先激发 ChatGPT 4.0 对系统功能视阈下及物性和作格性理论的学习,然后聚焦两者的并协互补关系以引出及物—作格综合模式,最后给予其相应的分析实例以促进输出的效度。在同 ChatGPT 4.0 的互动中,指令 1 为: Can you introduce transitive model and ergative model in Systemic Functional Linguistics in detail?

ChatGPT 4.0 迅速理解指令并完成任务,生成的回答理论理解正确,能够对及物模式的过程类型和参

与者角色、作格模式的作格句和中动句等概念进行较为全面的梳理总结。随即发布指令 2: Please explain the complementarity between transitive model and ergative model, and further integrate them with reference to the figure below.

指令 2 中所提及的图片详见图 2(于晖、王丽萍, 2020)。ChatGPT 4.0 可自动识别图文语言并完成对目标语言的转换,因此图文语言可以不一致,但应说明输出要求。按照指令,ChatGPT 4.0 将及物模式和作格模式进行系统整合,并对不同的过程类型小句进行及物视角和作格视角的阐述和举例。随即发布指令 3: Now I will give you some examples of functional analysis based on the integration of transitive model and ergative model that we discussed previously. You need to learn and master the tagging approach used in these examples. The examples are as follows: 1) The glass [Medium] was broken [Process: material] by the cat [Agent]. 2) We [Behaver] can dance [Process: behavioural] without lobsters [Accompaniment: comitative].

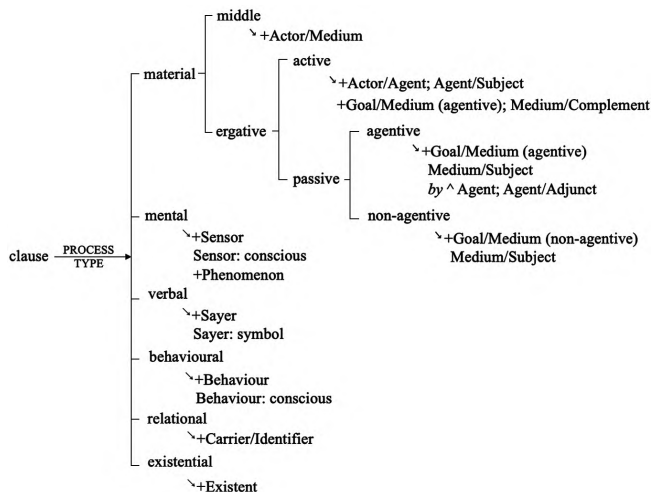


图2 及物—作格综合模式下过程类型及参与者角色

根据指令 3, ChatGPT 4.0 进一步解构所给例句,并总结标注方式,学习效果显著。此外,ChatGPT 4.0 还能举出新例句并进行分析,其回答如下: Applying the Approach to New Examples. Using this approach, you can analyze and tag new examples by following these steps. For instance: The team [Agent] completed [Process: material] the project [Goal] on time [Circumstance: time]. By mastering this tagging method, you can systematically analyze and integrate the transitive-

ty and ergative models to gain a deeper understanding of the functional roles in any given clauses. 若 ChatGPT 4.0 的回答并未直接生成新例分析,分析者可以发布指令(如“Can you give some examples using this tagging approach”)对其进一步追问。ChatGPT 4.0 的回答也符合预期。由此可见,ChatGPT 4.0 能够举一反三,且可以对分析做出较为专业的评价。

以上的指令可有序发布,也可进行简单整合。经实践,ChatGPT 4.0 虽可直接整合及物模式和作格模式,但回答不够完善且逻辑性较差。在给出图 2 所示的理论框架后,其回答则能够以图 2 为基础进行更为系统的阐述和总结。此外,若 ChatGPT 4.0 难以自主获取并学习目标理论和知识,或理论学习表现较差,可直接(如以网站链接的方式)为其提供目标理论的内容以助其顺利完成理论学习。由于 ChatGPT 4.0 尚未达到人类水平,分析者在同其互动过程中应尽量以具体、准确的语言清晰描述并传递信息或发布指令(文秋芳、梁茂成,2024),以保证其可以理解并顺利完成任务。

#### 4.2 描写: 词汇语法分析

在本研究中,描写阶段的任务是对目标文本《荒野的呼唤》进行词汇语法分析,即及物—作格分析。由于 ChatGPT 4.0 支持网盘和多种格式文本的上传,分析者可以选择任何一种形式向 ChatGPT 4.0 提供分析文本。本研究以 Word 文档格式为例,进行操作演示。指令 4 如下: Please analyze the paragraphs in this word document based on the previous tagging approach.

经过短时间的读档,ChatGPT 4.0 快速生成回答。经查阅,其不仅能够自动划分小句,还可以较为出色地完成及物—作格分析任务,且标注格式规范。以《荒野的呼唤》正文第一段为例,ChatGPT 4.0 的回答见图 3。

由图 3 可知,ChatGPT 4.0 虽然能够较为准确地分析目标语篇,但针对某些过程动词的及物性类型判断存在误差,如其将“BUCK did not read the newspapers”中的 read 判断为行为过程动词。在目标语境中,read 的及物性标注应更趋向为物质过程动词。此时,我们可以通过为 ChatGPT 4.0 提供相关例证分析,进一步促进其分析的信度。指令 5 如下: The transitivity analysis should be approached with caution. When the meaning of the verb leans toward “physiological and psychological” behavior, we recognize it as a behavioral process, such as “She is laughing”. When the meaning is more inclined to describe the occurrence of events, i. e. “doing & happening”, we should categorize it as a material process, such as “BUCK did not read the newspapers”. Verbs that are prone to incorrect analysis in transitivity include “taste”, “smell”, “see”, “hear”, etc. For example, “smell” can be identified as behavioral process verb in “She smelled the flowers”, but it becomes a relational process verb in “The soup smells delicious”. Therefore, please focus on the usage and context of the verbs so as to ensure accurate and convincing analysis.

ChatGPT 4.0 纠正之前的错误,并对类似过程动词做出新的判断,如在“in his sleep [circumstance: location] he [Actor/Agent] heard [Process: mental] the faintest sound [Goal/Medium] and knew [Process: mental] whether it heralded peace or peril [Phenomenon/Range]”中,“hear”由行为过程被改判为心理过程。针对其他的分析偏差,本研究采取类似的方式不断进行聚焦和调整,ChatGPT 4.0 的分析准确性渐次提升。另外,本研究尝试逐步扩充 Word 文档的词数。但是,

Clause 1: BUCK [Behaver/Medium] did not read [Process: behavioral] the newspapers [Goal/Range].  
 Clause 2: he [Senser/Medium] would have known [Process: mental] that trouble was brewing [Phenomenon/Range], not alone for himself [Circumstance: purpose], but for every tide-water dog, strong of muscle and with warm, long hair, from Puget Sound to San Diego [Circumstance: accompaniment].  
 Clause 3: men [Actor/Agent] had found [Process: material] a yellow metal [Goal/Medium].  
 Clause 4: steamship and transportation companies [Actor/Agent] were booming [Process: material] the find [Goal/Medium].  
 Clause 5: thousands of men [Actor/Agent] were rushing [Process: material] into the Northland [Circumstance: location].  
 Clause 6: These men [Senser/Medium] wanted [Process: mental] dogs [Phenomenon/Range].  
 Clause 7: the dogs [Carrier/Medium] they wanted were [Process: relational: identifying] heavy dogs [Attribute], with strong muscles by which to toil [Circumstance: manner], and furry coats to protect them from the frost [Circumstance: purpose].

图 3 ChatGPT 4.0 及物—作格段落分析示例

随着词数的增加, ChatGPT 4.0 在处理语篇细节方面的表现愈加不理想, 尤其是上传文档中文本数据过多或内容较长时, ChatGPT 4.0 的分析准确性显著下降。为解决这一问题, 目前较为合适的方法是拆分文档内容, 多次上传分析。

#### 4.3 说明: 语义功能分析

本研究的说明阶段旨在基于描写阶段的及物—作格分析进一步揭示小句的生态意义或价值取向, 即通过生态话语分析判断小句属于生态有益性、生态模糊性还是生态破坏性。小句生态取向的判断标准或原则被称为“生态哲学观”, 即话语分析者“个人所拥有的世界观、生态伦理以及对事物和事件的识构”(黄国文、陈旸, 2016)。本研究所采取的生态哲学观强调生命平等、和谐共生, 主张生态成分自身存在及活动的合理性和自主性, 倡导自然与人类之间的积极互动、彼此爱护。

及物—作格综合模式可以细化参与者角色, 但小句生态取向的判断还需考虑过程动词的语义。换言之, 即便参与者的角色相同, 不同语义的过程动词也可能造成小句所传递的生态意义或价值截然不同。通过整合细化参与者角色(图2)和过程动词的生态语义分类, 于晖、王丽萍(2020)以生态成分及其活动所体现的自然形象为锚点, 建构生态话语及物性分析框架(见图4)。该框架以自然参与者为中心, 解构及物

性过程类型, 以此辅助小句生态取向分析。

对目标文本《荒野的呼唤》进行生态话语分析, 指令6如下: Please judge the ecological orientations of the paragraphs in this word document clause by clause. Ecological orientations within Ecolinguistics can be classified as beneficial, ambivalent and destructive according to analysts' ecosophy or ecological philosophy, i. e. researchers' ecological ethics. The ecosophy at present argues for the rationality and autonomy of the existence and activities of ecological components, advocates the positive interaction and mutual love between nature and mankind, and emphasizes the equality and harmonious coexistence of life. Your analysis may refer to the figure below.

指令6所提及的图片详见图4。指令6旨在激发 ChatGPT 4.0 对于生态话语分析方法的学习, 并将之应用于目标语篇的分析。分析者也可参考及物—作格分析的预训练步骤拆分指令, 如先激发其对于生态话语分析方法的学习, 再上传语篇进行生态取向分析。经实践, 两种方式达成的效果基本一致。读档结束后, ChatGPT 4.0 能够作出较高质量的回答, 且能够给出该种生态取向判断的原因和数量汇总。仍以《荒野的呼唤》正文第一段为例, ChatGPT 4.0 的回答见图5。

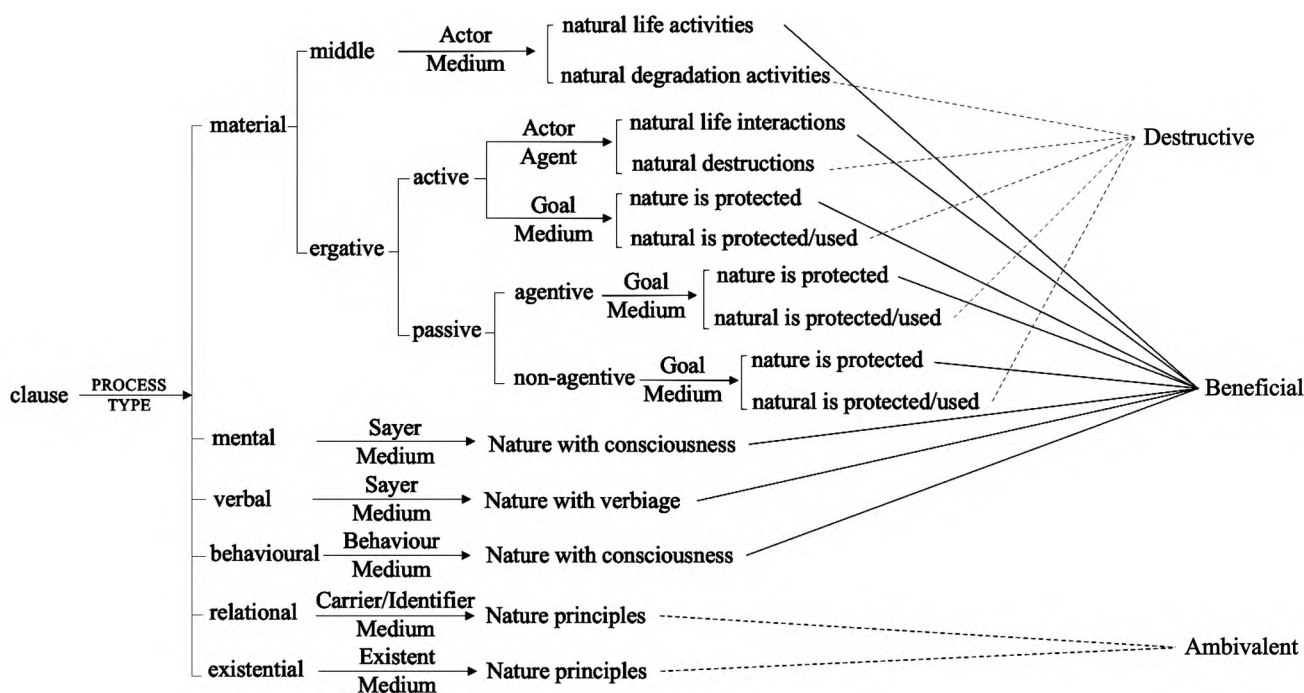


图4 生态话语及物性分析框架

Clause 1: BUCK [Actor/Agent] did not read [Process: material] the newspapers [Goal/Medium].  
 Orientation: Ambivalent  
 Reasoning: This clause describes Buck's action in relation with newspapers , with no direct ecological impact or engagement. [...]  
 Clause 3: men [Actor/Agent] had found [Process: material] a yellow metal [Goal/Medium].  
 Orientation: Destructive  
 Reasoning: The blooming of the gold find by companies indicates industrial exploitation , which is generally harmful to the environment. [...]  
 [Summary]  
 Beneficial Clauses: None  
 Ambivalent Clauses: 1 , 2 , 7  
 Destructive Clauses: 3 , 4 , 5 , 6

图 5 ChatGPT 4.0 生态取向段落分析示例

同及物一作格分析相似, ChatGPT 4.0 的生态话语分析能力仍在一定程度上受限于目标语篇的长度。另外, 其难以脱离语境直接识别拟人、隐喻等修辞手法, 如 ChatGPT 4.0 起初将“BUCK did not read the newspapers”判断为生态模糊性小句。但是, 当在指令中增加“Buck, as the hero of this story, is a dog”表述后, ChatGPT 4.0 将其改判为生态有益性小句, 并给出原因“The action of not reading the newspapers aligns with Buck's nature as a dog, emphasizing the autonomy and rationality of his existence”。由于小句的生态性分析是在及物一作格分析的基础上展开, ChatGPT 4.0 数据处理的工作负荷相对较小, 且经过多轮对话和调整, ChatGPT 4.0 辅助分析的准确性显著提升。通过人工对分析结果进行审查和矫正, 发现 ChatGPT 4.0 的准确率趋近 80%, 基本符合人机协同数据处理的二八分工规律(许家金, 2024)。基于人工修订后的分析结果, 本研究接着发布指令 7, 旨在通过统计 6 种过程类型及其生态取向的数量, 并采用卡方检验比较生态有益性和生态破坏性两组数据, 以呈现目标文本《荒野的呼唤》的整体词汇语法和生态语义表征。指令 7 如下: Please calculate the number of process types and their ecological orientations of the clauses in this word document. Can you present the data in one tabular format, with “Process Type” as the vertical header and “Ecological Orientation” as the horizontal header? Furthermore, can you perform a chi-square test on the two sets of data, beneficial orientations and destructive ones?

经过较长时间的理解和分析, ChatGPT 4.0 的回答见表 1 和表 2。

表 1 ChatGPT 4.0 对小句过程类型和生态取向的统计结果呈现

Process Type	Beneficial	Ambivalent	Destructive	Total
Material	2133	794	374	3301
Mental	291	57	19	367
Verbal	34	54	33	121
Behavioral	34	25	15	74
Relational	352	282	56	690
Existential	16	43	4	63
Total	2860	1255	501	4616

表 2 ChatGPT 4.0 的卡方检验结果呈现

Statistic	Value
Chi-Square Statistic	91.531
Degree of Freedom	5
p-value	$3.20 \times 10^{-18}$

由表 1 和表 2 可知, ChatGPT 4.0 能够按照要求进行数据统计, 但未能给出该种表征的深度解读。因此, 分析者仍需进行解释和评估, 进而完成对《荒野的呼唤》的功能语篇分析。 $p < 0.05$ , 这表明目标文本《荒野的呼唤》生态有益性语义表征与生态破坏性语义表征之间存在显著性差异, 整体呈现生态有益性取向。此外, 前文中的指令是笔者经多次调整后所得出的较具可行性的结果, 但在不同情况下, 其他分析者或仍需要根据 ChatGPT 4.0 回答中的具体偏差不断调整指令, 以提升其输出的质量。

在同 ChatGPT 4.0 的交互和协商过程中, 其表现出强大的思维和推理能力, 以及语义理解和语用能力, 这也使其在批量文本分析和处理效率方面表现出



色,可显著节省传统人工标注时间。但是,若无准确且清晰的指令,以及多轮的聚焦和调整,ChatGPT 4.0难以如期完成所给任务。总而言之,ChatGPT 4.0能够基于语言现象和理论学习发现和归纳语言规律,并根据用户要求对所提供的数据进行深度理解、推断和分析(秦洪武、周霞,2024)。用户也可通过质性表述、量化数据和图像等形式向 ChatGPT 4.0 提供信息或发布指令,多种信息形式的结合可以显著提升其学习和分析效率,促进其准确、高效地完成语言功能分析。

## 5 结语

虽然 ChatGPT 仅表现出人类智能的初级水平(Koubaa *et al.*, 2023; 陆俭明, 2024),但在自然语言理解、生成与处理方面所展现的优越性能可在某种程度上弥补语言研究中存在的技术和数据局限性,进而为促进针对语言功能的文本批量分析提供参考。因此,本研究从系统功能语言学视角探讨了 ChatGPT 辅助功能语篇分析的可能性,并尝试建构工作模型,同时基于生态话语及物性分析理论框架演示其操作流程。研究表明,ChatGPT 辅助的功能语篇分析具备一定的可行性和较强的可发掘潜力。但是,ChatGPT 以英语作为高资源语言(OpenAI, 2022),使得其针对汉语语言功能分析的有效性仍需进一步验证。在今后的研究中,以 ChatGPT 为代表的大语言模型和语言学理论之间的互补关系及其发展应用或是未来研究的方向。□

## 参 考 文 献

- [1] Bommasani, R. *et al.* On the opportunities and risks of foundation models [EB/OL]. Retrieved from <https://doi.org/10.48550/arXiv.2108.07258>, 2022.
- [2] Cai, W. ChatGPT can be powerful tool for language learning [EB/OL]. Retrieved from <https://universityaffairs.ca/career-advice/career-advice-article/chatgpt-can-be-powerful-tool-for-language-learning/>, 2023.
- [3] Chomsky, N., Roberts, I. & J. Watmull. Noam Chomsky: The false promise of ChatGPT [EB/OL]. Retrieved from <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html?searchResultPosition=1>, 2023.
- [4] Fairclough, N. *Critical Discourse Analysis: The Critical Study of Language* [M]. London: Longman, 1995.
- [5] Ferrari, F., van Dijk, J. & A. van den Bosch. Observe, inspect, modify: Three conditions for generative AI governance [EB/OL]. Retrieved from <https://doi.org/10.1177/14614448231214811>, 2023.
- [6] Graf, A. & R. E. Bernardi. ChatGPT in research: Balancing ethics, transparency and advancement [J]. *Neuroscience*, 2023(8).
- [7] Halaweh, M. ChatGPT in education: Strategies for responsible implementation [J]. *Contemporary Educational Technology*, 2023(2).
- [8] Halliday, M. A. K. & R. Hasan. *Cohesion in English* [M]. London: Longman, 1976.
- [9] Halliday, M. A. K. *Language as Social Semiotic: The Social Interpretation of Language and Meaning* [M]. London: Edward Arnold, 1978.
- [10] Halliday, M. A. K. New ways of meaning: The challenge to applied linguistics [J]. *Journal of Applied Linguistics*, 1990(6).
- [11] Halliday, M. A. K. *An Introduction to Functional Grammar* (2nd edn.) [M]. London: Edward Arnold, 1994.
- [12] Halliday, M. A. K. & C. M. I. M. Matthiessen. *Halliday's Introduction to Functional Grammar* (4th edn.) [M]. London and New York: Routledge, 2014.
- [13] Hanson, S. J. Dispatching "poverty of the stimulus": Did chatbot just learn context free grammar from positive examples? [EB/OL]. Retrieved from <https://lingbuzz.net/lingbuzz/007284>, 2023.
- [14] Kohnke, L., Moorhouse, B. L. & D. Zou. ChatGPT for language teaching and learning [J]. *RELC Journal*, 2023(2).
- [15] Koubaa, A. *et al.* Humans are still better than ChatGPT: Case of the IEEEExtreme competition [J]. *Heliyon*, 2023(11).
- [16] Li, L. H. *et al.* Evaluation on ChatGPT for Chinese language understanding [J]. *Data Intelligence*, 2023(4).
- [17] Matthiessen, C. M. I. M. *et al.* *Systemic Functional Insights on Language and Linguistics* [M]. Singapore: Springer, 2022.
- [18] OpenAI. Introducing ChatGPT [EB/OL]. Retrieved from <https://openai.com/blog/chatgpt#OpenAI>, 2022.
- [19] Piantadosi, S. T. Modern language models refute Chomsky's approach to language [EB/OL]. Retrieved from <https://>



- lingbuzz. net/lingbuzz/007180 , 2023.
- [20] Ratten , V. & P. Jones. Generative artificial intelligence ( ChatGPT ) : Implications for management educators [J]. *The International Journal of Management Education* , 2023 ( 3 ) .
- [21] Sahari , Y. , Al-Kadi , A. M. T. & J. K. M. Ali. A cross sectional study of ChatGPT in translation: Magnitude of use , attitudes , and uncertainties [ J ]. *Journal of Psycholinguistic Research* , 2023( 6 ) .
- [22] Vaswani , A. *et al.* Attention is all you need [EB/OL]. Retrieved from <https://arxiv.org/abs/1706.03762> , 2017.
- [23] Wang , F. Y. *et al.* What does ChatGPT say: The DAO from algorithmic intelligence to linguistic intelligence [J]. *IEEE/CAA Journal of Automatica Sinica* , 2023( 3 ) .
- [24] 冯志伟 , 张灯柯. GPT 与语言研究 [J]. 外语电化教学 , 2023( 2 ) .
- [25] 耿芳 , 胡健. 人工智能辅助译后编辑新方向——基于 ChatGPT 的翻译实例研究 [J]. 中国外语 , 2023( 3 ) .
- [26] 韩礼德 , 姜望琪 , 付毓玲. 篇章、语篇、信息——系统功能语言学视角 [J]. 北京大学学报( 哲学社会科学版 ) , 2011( 1 ) .
- [27] 何伟 , 张瑞杰. 生态话语分析模式建构 [J]. 中国外语 , 2017( 5 ) .
- [28] 胡加圣 , 戚亚娟. ChatGPT 时代的中国外语教育: 求变与应变 [J]. 外语电化教学 , 2023( 1 ) .
- [29] 胡开宝 , 尚文博. 语言学与语言智能 [J]. 华东师范大学学报( 哲学社会科学版 ) , 2022( 2 ) .
- [30] 胡壮麟. ChatGPT 谈外语教学 [J]. 中国外语 , 2023( 3 ) .
- [31] 黄国文. 功能语篇分析纵横谈 [J]. 外语与外语教学 , 2001( 12 ) .
- [32] 黄国文. 功能语篇分析面面观 [J]. 国外外语教学 , 2002 ( 4 ) .
- [33] 黄国文 , 陈旸. 生态哲学与话语的生态分析 [J]. 外国语文 , 2016( 6 ) .
- [34] 黄国文 , 张培佳. 系统功能语言学的性质、特点及发展 [J]. 现代外语 , 2020( 5 ) .
- [35] 令小雄 , 王鼎民 , 袁健. ChatGPT 爆火后关于科技伦理及学术伦理的冷思考 [J]. 新疆师范大学学报( 哲学社会科学版 ) , 2023( 4 ) .
- [36] 刘海涛 , 亓达. 大语言模型的语用能力探索——从整体评估到反语分析 [J]. 现代外语 , 2024( 4 ) .
- [37] 刘利等. “ChatGPT 来了: 国际中文教育的新机遇与新挑战”大家谈( 上 ) [J]. 语言教学与研究 , 2023( 3 ) .
- [38] 刘向东 , 王博. 语言系统的分形与功能变异 [J]. 外语教学理论与实践 , 2021( 2 ) .
- [39] 陆俭明. 人工智能语言模型 ChatGPT 不会影响人类的修辞知识生产范式 [J]. 当代修辞学 , 2024( 2 ) .
- [40] 秦洪武 , 周霞. 大语言模型与语言对比研究 [J]. 外语教学与研究 , 2024( 2 ) .
- [41] 秦涛等. ChatGPT 的工作原理、关键技术及未来发展趋势 [J]. 西安交通大学学报 , 2024( 1 ) .
- [42] 饶高琦 , 胡星雨 , 易子琳. 语言资源视角下的大规模语言模型治理 [J]. 语言战略研究 , 2023( 4 ) .
- [43] 苏祺 , 杨佳野. 语言智能的演进及其在新文科中的应用探析 [J]. 中国外语 , 2023( 3 ) .
- [44] 王勇 , 孙亚迪. 语篇分析中的解释 [J]. 中国外语 , 2021 ( 5 ) .
- [45] 文秋芳 , 梁茂成. 人机互动协商能力: ChatGPT 与外语教育 [J]. 外语教学与研究 , 2024( 2 ) .
- [46] 辛志英 , 黄国文. 系统功能语言学与生态话语分析 [J]. 外语教学 , 2013( 3 ) .
- [47] 许家金. 话语研究的人机协同创新模式 [R]. 贵阳: 第十九届中国英汉语比较研究会英汉语篇分析学术研讨会 , 2024.
- [48] 杨敏 , 王亚文. ChatGPT 的“理解”与“意义”: 论其生成语言背后的形式、功能与立场 [J]. 中国外语 , 2023( 3 ) .
- [49] 于晖 , 王丽萍. 生态话语及物性分析模式探究——以教育语篇为例 [J]. 外语与外语教学 , 2020( 6 ) .
- [50] 袁毓林. 超越聊天机器人, 走向通用人工智能——ChatGPT 的成功之道及其对语言学的启示 [J]. 当代语言学 , 2023( 5 ) .
- [51] 张杨. ChatGPT 与意象感知问题: 再论塞尔的人工智能语言批判 [J]. 外语学刊 , 2023( 6 ) .
- [52] 赵雪等. 面向语言文学领域的大语言模型性能评测研究 [J]. 外语电化教学 , 2023( 6 ) .

基金项目: 本文为国家社科基金一般项目“中英学术论文写作中的思辨对比研究”(项目编号: 19BYY233)、国家社科基金青年项目“话语权建构视角下跨学科语篇多声系统研究”(项目编号: 24CYY065)和北京师范大学外国语言文学学院课程思政项目“语篇分析课程思政建设”(项目编号: KCSZ2022007)的阶段性成果。

收稿日期: 2024-04

通讯地址: 100875 北京市 北京师范大学外国语言文学学院

# LLM-Aided Functional Discourse Analysis: Theory , Method and Practice

YU Hui ,SONG Jinge & WANG Le

( School of Foreign Languages and Literature , Beijing Normal University , Beijing 100875 , China)

**Abstract:** The development of linguistic intelligence technology has propelled foreign language teaching and research into the era of “big data” , with the emergence of large language models like ChatGPT pushing the exploration of linguistic theories in natural language generation and processing to new heights. However , there is a scarcity of studies applying its “pre-trained deep learning capability” systematically to functional discourse analysis. In view of this , this research constructs a ChatGPT-aided

model for functional discourse analysis from a systemic functional perspective , and demonstrates the operation of ecological discourse analysis with *The Call of the Wild* , aiming to provide guidance for the functional analysis and processing of large-scale corpora.

**Key words:** Large Language Model; Artificial Intelligence; Discourse Analysis; Language Function; Ecological Discourse