

大语言模型字幕英汉翻译 译文质量评估

——以文心一言和智谱清言为例

徐晓婕, 李华东

上海海事大学外国语学院, 上海

收稿日期: 2025年12月5日; 录用日期: 2025年12月30日; 发布日期: 2026年1月12日

摘要

随着全球流媒体平台对多语言字幕需求的激增, 大语言模型(LLMs)成为字幕翻译的重要工具。然而, LLMs在文化适配、口语表达及语境把握等方面的表现尚待系统评估, 现有自动化指标(如BLEU)难以反映多模态语境下的真实可接受性。本研究融合翻译目的论与图灵测试“自然等效性”理念, 构建三维英译汉质量评估模型, 从文化适应性、语句流畅性、翻译准确性三个维度对字幕质量进行量化评估。研究选取6类影视题材的9段字幕, 对比人工译本、文心一言(ERNIE Bot)和智谱清言(Zhipu Qingyan)的译文表现, 并结合201份问卷与统计检验进行实证分析。结果表明, 文心一言在三个核心维度上综合表现最优, 尤其在语句流畅性上具有显著优势; 受访者对AI翻译的核心诉求集中于提升准确性与语境理解。本研究为多模态适配的翻译质量评估提供了可操作框架与实证依据。

关键词

大语言模型, 字幕翻译, 文心一言, 智谱清言, 质量评估

Evaluating the Quality of English-to-Chinese Subtitle Translation by Large Language Models

—A Case Study of ERNIE Bot and Zhipu Qingyan

Xiaojie Xu, Huadong Li

College of Foreign Languages, Shanghai Maritime University, Shanghai

Received: December 5, 2025; accepted: December 30, 2025; published: January 12, 2026

Abstract

With the surging demand for multilingual subtitles on global streaming platforms, Large Language Models (LLMs) have emerged as pivotal tools for subtitle translation. However, the performance of LLMs in cultural adaptation, colloquial expression, and context comprehension remains to be systematically evaluated, as existing automated metrics (e.g., BLEU) fail to adequately reflect authentic acceptability within multimodal contexts. Integrating Skopos Theory with the concept of “Natural Equivalence” from the Turing Test, this study constructs a “Three-Dimensional English-to-Chinese Translation Quality Assessment Model” (3D-ECTQA Model) to quantitatively evaluate subtitle quality across three dimensions: Cultural Adaptability, Linguistic Fluency, and Translation Accuracy. The study selects nine subtitle segments from six genres of film and television, comparing the performance of human translation against ERNIE Bot and Zhipu Qingyan. Empirical analysis is conducted through 201 questionnaires and statistical tests. The results indicate that ERNIE Bot demonstrates the best overall performance across the three core dimensions, exhibiting a significant advantage in linguistic fluency. Furthermore, respondents’ core demands for AI translation center on enhancing accuracy and context understanding. This study provides an operable framework and empirical evidence for translation quality assessment adapted to multimodal contexts.

Keywords

Large Language Models (LLMs), Subtitle Translation, ERNIE Bot, Zhipu Qingyan, Quality Assessment

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

当今, 随着全球流媒体平台的蓬勃发展, 字幕翻译已超越单纯的语言转换, 成为跨文化传播与视听内容分发的核心环节。近年来, 大语言模型(Large Language Models, LLMs)的突破性进展为翻译领域带来了新的机遇, 其在大规模语料预训练与生成能力上的优势, 展现出实现高质量与高效率翻译的显著潜力。然而, 字幕翻译作为一种受时空制约的特殊翻译类型, 对语境理解、口语化表达及文化适配性具有高度依赖性, LLMs 在此细分领域尚待进一步实证探讨。

在现有的机器翻译质量评估体系中, 研究者多依赖 BLEU、ROUGE 等自动化指标。这些指标虽然在计算上具有便捷性, 但其底层逻辑仅能衡量译文与参考文本之间的表面词汇匹配度, 无法深入语义层面有效评估译文在多模态语境下的真实质量。特别是在字幕翻译场景中, 自动化指标难以捕捉口语化表达的流畅度、幽默元素的传递以及文化负载词的处理策略等核心要素。尽管已有部分研究尝试构建新的评估框架, 但往往局限于单一的语料类型或缺乏多维度的量化标准。因此, 在 LLMs 逐渐介入字幕生产流程的背景下, 亟需建立一套科学、系统且符合视听翻译特性的质量评估体系。

为此, 本研究创新性地融合了功能派翻译理论中的目的论(Skopos Theory)与图灵测试中的自然等效性(Natural Equivalence)理念, 构建了三维英译汉字幕质量评估模型(Three-Dimensional English-to-Chinese Translation Quality Assessment Model, 简称 3D-ECTQA Model)。该模型试图突破传统评估的局限, 将抽象的翻译理论原则操作化为三个可量化的核心维度: 文化适应性、语句流畅性与翻译准确性。本研究旨在通过这一多维框架, 检验 LLMs 生成的字幕译文是否已具备接近人工译本的自然度与可接受性, 即观

众在观看时是否无法察觉机器痕迹,从而达到图灵测试所定义的自然等效状态。

本研究选取了涵盖多类型题材的经典影视作品作为语料来源,对比分析了传统人工译本与两款主流国产大语言模型——文心一言(ERNIE Bot)和智谱清言(Zhipu Qingyan)的译文表现。通过收集 201 份有效问卷的用户反馈数据,本研究不仅从受众感知视角为 LLMs 翻译提供了实证依据,丰富了现有的评估视角,也为未来多模态语境下的机器翻译优化提供了理论参考与实践路径。本研究将官方及字幕组发布的经过人工审校的定稿译本视为人工译本,作为质量评估的参照基准,以此对比大语言模型在无人工干预情况下的原始生成质量。

2. 文献综述

目的论(Skopos Theory)是功能派翻译理论的重要组成部分,由德国翻译学者汉斯·弗米尔(Hans Josef Vermeer)于 20 世纪 70 年代提出,标志着翻译研究范式从忠实源文向服务目标的根本转变。Skopos 一词源于希腊语,意为目的、目标或意图。Vermeer 强调,翻译的本质是一种跨语言的目的性交际行为,其核心原则为“翻译行为的目的决定翻译策略与过程”[1]。目的论提出三项基本原则:(1) 目的原则(Skopos Rule),即翻译策略和形式完全由译文想要实现的交际目的决定,不同类型的文本应服务于不同的目的;(2) 连贯原则(Coherence Rule),即译文必须自身连贯(intratextual coherence),在目标文化和语言中被受众理解为有意义且符合情境;(3) 忠实原则(Fidelity Rule),即译文应尽量贴近原文,但这种忠实并非绝对,而是需与目的和连贯原则相协调。这种多层面的指导原则为译者在面对影视、多媒体等具有时间、空间与视觉限制的字幕翻译场景时提供了操作弹性[2]。

在字幕翻译中,目的论的价值尤为突出。字幕不仅涉及语言转换,还需协调画面节奏、视觉空间和文化背景,传统字对字的忠实观往往难以适应。译者需灵活意译、替换文化意象,并注重译文风格,使观众在极短时间内理解内容。然而,目的论的原则偏重宏观理念和译者决策,难以直接量化评估字幕质量:它指出要连贯,但未说明如何连贯;强调需适应文化,却未提供衡量标准。因此,为将目的论理念转化为可操作的质量评估,本文提出三维英汉翻译质量评估模型,通过三个可观察、可度量维度,将目的的实现、语句连贯、语言忠实等抽象问题转化为观众可评分、研究者可统计的实证指标。这一模型既继承了目的论功能主义精神,又填补了其量化评估的空白。

与此同时,图灵测试(Turing Test)由英国计算机科学家艾伦·麦席森·图灵(Alan Mathison Turing)提出,用以检验机器是否具备智能[3][4]。其核心思路是:若测试者无法区分机器与人的对话表现,则认为机器具有智能。这一逻辑引申出自然等效性(Natural Equivalence)概念,即机器生成文本能否以自然、连贯的语言行为融入人类交际,使人类难以察觉其非人类身份。随着大语言模型的发展,这一概念获得新的学术意义,若机器生成的字幕在文化适应性、语句流畅性与翻译准确性上与人类译文无明显差异,即观众无法察觉机器痕迹,则可认为其具备一定程度的自然等效性。

结合目的论与自然等效性,3D-ECTQA 模型在先前研究中确立了三大核心维度,文化适应性、语句流畅性、翻译准确性,分别对应目的论的目的原则、连贯原则和忠实原则,同时体现自然等效性对译文自然性和可接受性的要求。

随着大语言模型(Large Language Models, LLMs)的发展,字幕翻译迈入了智能化阶段。以 ChatGPT 为代表的模型,在多语翻译任务中表现出卓越的语义理解与文本生成能力[5]。LLMs 对口语化表达的精准捕捉与情感语义的传递能力,使其在影视字幕翻译中表现优异;在新闻类视频中,LLMs 依托强大的事实核对能力与信息整合能力,能够在较短时间内实现高准确度的内容转换与时效性表达;同时,凭借扎实的跨学科知识储备,LLMs 在教育类视频中能够较准确地传递知识概念。LLMs 翻译依赖大规模语料的预训练与指令微调机制,通过编码、语义建模与解码等流程,在语义建模、上下文理解与词汇生成方面相

较于传统机器翻译(MT)展现出更强的泛化能力[6]。其应用已覆盖新闻、商务、社科等领域,并在低资源语言翻译及多语言指令响应等任务中表现优异[7]-[9]。

然而,学界也指出其表现仍受到视频类型差异与任务复杂度的影响。首先,LLMs 的视频类型适配性仍显不足。研究表明,语言模型在叙事类或影视类视频中通常表现更佳,能够较好地捕捉情节逻辑与口语化特征[10],其次,现有字幕翻译研究多侧重于译文与原文之间的表层对应关系,而对译文的接受度和观众体验关注不足。最后,评价体系尚不统一,不同研究往往采用自建指标或主观评分,缺乏系统的量化标准,难以对模型输出的字幕质量进行客观比较[11]。

综上所述,现有研究在理论与实践层面仍存在以下三方面空白:1) 评估维度单一,传统自动化翻译质量指标多聚焦词面匹配,难以涵盖字幕翻译中的文化适应、语境协同与口语自然性等关键要素;2) 尽管 LLMs 在通用文本翻译中表现优异,现有评估框架缺乏对字幕翻译细粒度场景(如影视语料、口语化表达)的系统分析;3) 缺失用户感知角度与等效性验证逻辑,目前对大语言模型字幕译文质量的判断多基于专业评分或静态参考译文比对,缺乏观众是否能辨别机器与人类译文的自然等效性视角。因此本研究从目的论与自然等效性双重视角出发,构建三维字幕英译汉质量评估模型,并选取经典影视作品与多类型字幕版本为研究样本,结合观众反馈数据,实证探究大语言模型字幕英译汉是否已具备接近人工译文的自然性与可接受性。

3. 研究设计

3.1. 语料与对象

首先选定《肖申克的救赎》《莎翁情缘》《名利场》《火星救援》《老友记》《摩登家庭》六部经典影视作品的九段官方英文字幕为样本。对比三个版本译文:人工译本(基准)、文心一言(V3.3.0)和智谱清言(ChatGLM4)。向 LLMs 输入统一 Prompt,要求遵循文化适应性、语句流畅性及翻译准确性三项标准(见表 1)。

Table 1. The evaluation dimensions and standards of 3D-ECTQA model
表 1. 3D-ECTQA 模型评估维度与标准说明

维度名称	核心定义	评估标准
文化适应性	译文是否有效对接目标语言文化背景,能否恰当转换原文中的文化负载词汇、隐喻与社会语境。	(1) 是否准确传递源语文化信息 (2) 是否使用目标语常见表达 (3) 是否避免生硬直译或文化隔阂
语句流畅性	译文在目标语言中是否符合自然表达习惯,包括语法结构、节奏与语用习惯。	(1) 句法结构是否自然 (2) 语序是否符合中文习惯 (3) 语言是否口语化、符合语境
翻译准确性	译文在语义内容、术语使用及逻辑推演上是否准确重现原文信息,避免误译或遗漏。	(1) 是否存在信息增删误 (2) 是否准确还原人名、术语、事实 (3) 逻辑是否清晰一致

3.2. 方法与工具

通过“问卷星”收集 201 份有效问卷。问卷分为两部分:一是用户行为与技术感知调查;二是基于

3D-ECTQA 模型的译文盲评。采用 Wilcoxon Signed-Rank 检验和 Cliff's Delta 效应量分析数据, 以验证 LLMs 译文与人工译本的差异显著性。

4. 问卷结果分析：用户感知与技术痛点

4.1. 用户依赖与质量敏感度：从“辅助”到“刚需”

调查数据显示(见表 2), 外语影视内容的消费已具有广泛的受众基础, 84.73%的受访者表示会经常或偶尔观看外语影片。在这一高频消费场景下, 字幕表现出极强的刚性需求特征: 高达 96.06%的受访者表示观影时需要依赖字幕, 其中 66.5%为“总是需要”。这一数据充分印证了字幕已超越单纯的语言辅助功能, 成为跨语言观影的核心媒介。

Table 2. Core feedback from the questionnaire results

表 2. 问卷结果核心反馈

问题类别	关键结果	支持率/比例
1. 您通常观看外语影片的频率?	偶尔(48.77%) 经常(35.96%)	84.73%
2. 您观看外语影片时, 是否经常需要依赖字幕?	是, 总是需要(66.5%) 有时需要(29.56%)	96.06%
3. 您认为字幕翻译对您的观影体验有多大影响?	比较大, 但可以通过其他方式弥补(如查看原声版)(50.74%) 非常大, 直接影响对作品的理解和评价(30.54%)	81.28%
4. 您对目前影视作品字幕翻译的准确性满意吗?	比较满意(76.35%) 非常满意(7.39%)	83.74%
5. 在使用 AI 字幕翻译时, 以下哪些问题最让您困扰?	语境理解不足(83.25%) 翻译不准确(78.33%) 语音识别错误(58.13%)	前三困扰平均占比超 55%
6. 您是否注意到过 AI 字幕翻译的应用?	有时注意到(43.35%) 是, 经常注意到(27.59%) 很少注意到(22.66%)	70.94%
7. 您对 AI 字幕翻译持什么态度?	支持, 但认为仍需人工审核(68.97%) 非常支持, 认为能提高翻译效率和质量(12.81%) 不太支持, 担心影响翻译准确性(5.42%)	支持率: 81.78%
8. 您对 AI 字幕翻译的未来发展有何期待?	提高翻译准确性(57.34%) 提升语境感知力(20.98%)	此二需求总占比超 78%

然而, 用户对字幕质量的敏感度与当前供给质量之间存在显著落差。虽然有 81.28%的受访者认为字幕翻译质量对观影体验有“非常大”或“比较大”的影响, 直接关乎对作品深度的理解; 但在满意度评价上, 仅有 7.39%的受访者对现有字幕翻译准确性表示“非常满意”。这表明, 尽管现有的字幕翻译服务(含传统机翻)能满足基本的信息传递, 但尚未达到观众期待的“信达雅”高标准, 高质量的字幕翻译仍是市场上的稀缺资源。

4.2. AI 技术痛点与态度：信任赤字与深层语义挑战

针对 AI 字幕翻译技术的应用，受访者呈现出审慎接纳的态度特征。虽然 AI 技术的市场渗透率较高(70.94%的受访者曾注意到 AI 字幕)，且有 81.78%的受访者对 AI 应用持支持态度，但这种支持是有条件的，高达 68.97%的受访者强调 AI 翻译仍需“人工审核”。这反映出用户对 AI 独立生成内容的信任度不足，倾向于将 AI 视为提效工具而非完全替代者。

造成这种信任赤字的根本原因在于技术痛点的转移。问卷数据显示，受访者在使用 AI 字幕时面临的核心困扰并非基础的语音识别(占比 58.13%)，而是更高阶的语义理解问题：语境理解不足(83.25%)位居痛点之首，说明 AI 在处理多模态语境(如画面暗示、语气潜台词)时仍显笨拙；翻译不准确(78.33%)紧随其后，显示出 AI 在术语准确性和长难句逻辑上仍有待提升。

这一结果反映出当前 AI 字幕技术的瓶颈，即听写能力的短板已逐渐补齐，而深层理解与表达的精准度成为新的核心矛盾。这也与受访者对未来的期待高度一致，超过 78%的用户呼吁 AI 应重点提升翻译准确性与语境感知力。

5. 实证分析：译文质量多维对比

为进一步验证前文评估结果的可靠性与显著性，鉴于文心一言在 LLMs 组中的综合表现更突出，本研究选取文心一言与人工译本作为对比对象，逐维度深入剖析其在英译汉质量上的差异及统计学意义。

在统计方法上，首先通过描述性统计分析两类译本在三个维度上的得分分布特征，计算平均值(M)，其中 M1 和 M2 分别代表文心一言和人工译本的均值，以反映数据的集中趋势。其次，考虑到同一文本的评分在两个译本间存在配对关系，且评分数据为序数型变量、样本量较小(n=9)，难以满足正态分布假设，故采用 Wilcoxon Signed-Rank 检验比较两组数据中位数的差异。

为量化差异的实际效应强度，引入 Cliff's Delta 效应量(δ)作为补充指标。 δ 值介于[-1, 1]之间，绝对值越接近 1 表明差异越显著。根据效应大小，将其划分为可忽略(negligible)、小(small)、中(medium)和大(large)四个等级。此外，本研究还报告了 95%置信区间(CI)，以增强结果解释的稳健性。本次数据分析基于 Python 3.9.22 环境，调用 scipy.stats 与 cliffs_delta 等库完成统计检验，具体结果见表 3。

Table 3. Quality statistics and test results of three-dimensional English-Chinese translation of ERNIE Bot and artificial translation

表 3. 文心一言与人工译本三维英译汉质量统计与检验结果

维度	M1	M2	p 值	95% CI	效应量 δ	r 效应量强度
文化适应性	2.44	1.22	0.0742	[0.33, 1.89]	0.741	0.632
语句流畅性	2.56	1.22	0.0155	[0.56, 2.00]	0.815	0.829
翻译准确性	2.44	1.44	0.0547	[0.22, 1.67]	0.630	0.652

5.1. 文化适应性能力：文心一言 > 智谱清言 > 人工译本

在文化适应性维度，文心一言表现出最佳的语境贴合度。统计数据显示，文心一言的平均得分(M1=2.44)显著高于人工译本(M2=1.22)，均值差达 1.22。效应量分析进一步证实了这一优势(Cliff's Delta=0.741，属于 Large 级别)，且 95%置信区间[0.33, 1.89]未跨越 0，表明其优势具有统计学意义。

在处理含有隐喻的文化语境时，人工译本有时存在过度解读的倾向。例如原文“we'll meet at the checkout counter”，人工译本将其升华为“我们绝对会在人生的收银台见”。这种处理虽增加了文学色彩，但在影视快节奏语境中显得过于刻意且偏离原意。相比之下，文心一言译为“咱们收银台见。记住

哦, 收银台见”, 既保留了原文的口语重复强调, 又符合中国观众的认知习惯, 实现了更自然的文化归化处理。智谱清言的译文则相对中规中矩, 略显书面化。

5.2. 语句流畅性能力: 文心一言 > 智谱清言 > 人工译本

这是差异最为显著的维度。统计检验结果显示, 文心一言的得分均值($M1 = 2.56$)高于人工译本($M2 = 1.22$), Wilcoxon 检验 P 值为 $0.0155 (< 0.05)$, 达到了统计学显著水平。数据表明 LLMs 在生成符合中文句法规范和口语节奏的文本方面相对突出。

案例中, 面对 “I’m just gonna, kind of...implode” 这类极具口语色彩的表达, 各版本差异明显。人工译本直译为 “我就会内爆而死”, 虽然术语准确, 但在表达情感崩溃的语境下显得生硬机械。智谱清言保留了 “内爆” 一词, 虽加入了语气词 “嗯……”, 但整体语感仍显技术化。文心一言则灵活处理为 “我恐怕就会……彻底崩溃”。尽管 “崩溃” 弱化了物理意义, 但极佳地契合了人物当下的绝望情绪与中文口语习惯, 增强了字幕的带入感与幽默感。

5.3. 翻译准确性能力: 文心一言 > 智谱清言 > 人工译本

在翻译准确性维度, 文心一言($M1 = 2.44$)依然优于人工译本($M2 = 1.44$)。效应量 Cliff’s Delta 值为 0.630 (Large), 表明其在信息完整性与术语把控上的优势明显。

以科幻题材中的术语翻译为例, 原文 “unused hydrazine”, 人工译本误译为 “联氨限制着”, 导致核心信息严重失真且语句不通。智谱清言译为 “未使用的肼”, 术语最为严谨, 但 “肼” 字对于普通大众存在认知门槛。文心一言则译为 “未使用的联氨燃料”。该译文虽有小瑕疵 (将 MDV 误译为月球驾驶舱), 但通过增补 “燃料” 一词进行了语义显化, 有效降低了普通观众的认知负荷, 体现了其根据受众背景灵活调整翻译策略的能力。

6. 结论与展望

本研究基于翻译目的论与自然等效性理论, 构建了 3D-ECTQA 字幕质量评估模型, 并通过问卷调查与统计检验, 对文心一言、智谱清言及人工译本进行了多维对比。研究结果表明: (1) LLMs 在字幕翻译中的 “类人” 潜力: 在所选语料下, 大语言模型 (尤其是文心一言) 在语句流畅性方面表现出一定优势, 并在部分口语化表达处理上接近或略优于人工译本, 显示出较高的自然等效性潜力; (2) 差异化表现: 文心一言在文化适应与语境重构上展现出一定灵活性, 倾向于语义显化; 智谱清言在术语严谨性上较为稳定, 但在口语节奏处理上存在一定局限; (3) 用户需求的差异: 尽管 AI 的语言生成能力有所提升, 但用户对字幕的核心需求已从单纯的听写准确延伸至语境感知与情感共鸣, 这仍是当前 AI 需要进一步改进的方向。

需要指出的是, 本研究仅选取了 9 个影视片段作为语料, 虽涵盖多种题材, 但难以全面代表所有复杂的视听语境 (如双关语、方言等)。此外, 研究主要基于文本输入, 尚未充分考虑视频画面和音频语调对翻译决策的潜在影响。未来研究可在更大规模、更多样化的语料上进行验证, 并探索多模态信息在字幕翻译评估中的作用, 以进一步丰富和优化 3D-ECTQA 模型的适用性。

综上, 本研究所提出的 3D-ECTQA 模型为字幕质量评价提供了可操作框架, 并为人机协作模式下字幕生产与微调提供初步数据参考, 但结论应限于本研究语料和条件。

参考文献

- [1] Reiss, K., Nord, C. and Vermeer, H.J. (2014) Towards a General Theory of Translational Action: Skopos Theory Explained. Routledge.
- [2] 高红. 中国电影字幕翻译之 “切” 的原则[J]. 上海翻译, 2015(2): 28-33.

- [3] Jacquet, B., Jamet, F. and Baratgin, J. (2021) On the Pragmatics of the Turing Test. 2021 *International Conference on Information and Digital Technologies (IDT)*, Zilina, 22-24 June 2021, 123-130.
<https://doi.org/10.1109/idt52577.2021.9497570>
- [4] Turing, A.M. (2021) Computing Machinery and Intelligence (1950). *Mind*, **59**, 33-60.
- [5] Achiam, J., Adler, S., Agarwal, S., *et al.* (2023) GPT-4 Technical Report.
- [6] 赵鑫, 窦志成, 文继荣. 大语言模型时代下的信息检索研究发展趋势[J]. 中国科学基金, 2023, 37(5): 786-792.
- [7] 孙光耀, 赵志泉, 沈思, 等. 基于大语言模型的人文社会科学汉英机器翻译研究[J]. 数据分析与知识发现, 2025, 9(4): 32-45.
- [8] 侯钰涛, 阿布都克力木·阿布力孜, 史亚庆, 等. 面向“一带一路”的低资源语言机器翻译研究[J]. 计算机工程, 2024, 50(4): 332-341.
- [9] 赵衍, 张慧, 杨祎辰. 大语言模型在文本翻译中的质量比较研究——以《繁花》翻译为例[J]. 外语电化教学, 2024(4): 60-66+109.
- [10] Shen, S. and Garg, A. (2025) Adapting Large Language Models for Movie Domain with Narrative Understanding Tasks. *Proceedings of the 29th Conference on Computational Natural Language Learning*, Vienna, 31 July-1 August 2025, 187-200. <https://doi.org/10.18653/v1/2025.conll-1.13>
- [11] 雷静, 李明明. 机翻字幕质量评估的模型构建研究——以“人人译视界”和“讯飞听见字幕”为例[J]. 大连民族大学学报, 2023, 25(2): 176-182.