

多模态话语分析视角下外宣字幕翻译研究 ——以电影《叶问》为例

张千龙

北京语言大学英语和高级翻译学院, 北京

收稿日期: 2025年4月21日; 录用日期: 2025年5月29日; 发布日期: 2025年6月11日

摘要

全球化新浪潮背景下, 中国电影产业迎来国际传播的历史性机遇。同时, 我国电影工业体系日趋成熟, 不断产出高质佳作(如《流浪地球》《哪吒》系列等)。然而, 从跨文化传播的维度来看, 当前字幕外译依然存在着质量困境, 海外广泛传播的头部国产影片, 依然存在诸多错译、漏译甚至文化误读等问题, 严重影响了我国影视作品的国际接受度与传播效果。究其原因, 在于传统字幕翻译多拘泥于文本, 且忽视对受众接受度的反馈研究。因此, 通过多模态话语分析方法, 围绕文化语境、内容特征与表达策略等四个层面深入解构视听文本, 开创构建“受众反馈驱动”的翻译模式, 能为提升中国影视作品对外翻译的质量与文化传播力提供新的理论参考与实践方向, 助力中国优秀文化“走出去”。

关键词

字幕翻译, 多模态话语分析, 跨文化传播

A Study on Subtitling for International Communication from a Multimodal Discourse Analysis Perspective —A Case Study of the Film “Ip Man”

Qianlong Zhang

School of English and International Studies, Beijing Language and Culture University, Beijing

Received: Apr. 21st, 2025; accepted: May 29th, 2025; published: Jun. 11th, 2025

Abstract

Against the backdrop of the new wave of globalization, China's film industry is embracing a historic

opportunity for international communication. Meanwhile, China's film industrial system has become increasingly mature, continuously producing high-quality masterpieces (such as "The Wandering Earth" and "Ne Zha" series). However, from the perspective of cross-cultural communication, the current external translation of subtitles still faces quality dilemmas. In the leading domestic films widely spread overseas, there are still many issues such as mistranslations, omissions, and even cultural misinterpretations, which seriously affect the international acceptance and communication effectiveness of Chinese film and television works. The root cause lies in the fact that traditional subtitle translation has been overly constrained on the text while neglecting feedback research on audience acceptance. Therefore, by adopting multimodal discourse analysis, deeply deconstructing audiovisual texts from four dimensions including cultural context, content characteristics, and expression strategies, and establishing a "feedback-driven translation model" centered on audience needs, this study can provide new theoretical references and practical directions for improving the quality of external translation of Chinese film and television works and their cultural communication capabilities, thereby helping excellent Chinese culture to "go global".

Keywords

Subtitle Translation, Multimodal Discourse Analysis, Cross-Cultural Communication

Copyright © 2025 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

全球化进程的加速与数字技术的革新，使跨文化传播成为国家软实力竞争的核心场域。中国故事的对外传播，既是文化价值的输出，也是国际话语权的建构。在此驱动下，影视、短视频等视听媒介凭借多模态符号的强表现力，成为外宣实践的重要载体。然而，视听翻译的复杂性远超传统文本翻译，如何通过翻译模式的优化，实现从“语言转换”到“意义共享”的跨越，成为提升中国话语国际传播效能的关键命题。

当前字幕翻译研究面临极大挑战。传统研究长期受困于“文本中心主义”，忽视多模态协同作用与受众的主体性，导致研究无法突破传统翻译研究的束缚，同时与受众的认知需求相脱节。为此，本研究以多模态话语分析为理论框架，以受众反馈为检验，探索外宣翻译模式的革新路径。前者从文化、语境、内容与表达四层面解构视听文本的意义；后者通过受众反馈的分析，探寻不同语言符号的理解障碍与情感共鸣的驱动因素。二者的协同旨在突破传统翻译的线性逻辑，构建“受众反馈驱动”的正循环，助力中国故事从“消极单向输出”迈向“积极双向共鸣”，为全球化语境下的文化传播开辟新路径。

2. 理论背景

多模态话语分析(Multimodal Discourse Analysis)是系统功能语言学与符号学深度融合的产物，其核心在于揭示语言、图像、声音、动作等模态在意义建构中的协同机制[1]。该理论认为，人类交际的本质是多元符号系统的动态互动，意义的生成不仅依赖于单一模态的线性表达，更需要通过跨模态的互补、强化或竞争构建整体意义。在视听翻译领域，多模态话语分析突破了传统“文本中心主义”的局限，强调字幕翻译需与镜头语言、符号象征、音乐情感等形成跨模态适配，从而在目标语境中重构完整的文化意义。

基于原有的理论体系，张德禄提出了多模态话语分析的理论框架，该框架以层级化结构整合文化、语境、内容与表达四个核心维度，通过系统性关联揭示多模态交际的运作机制[2]。文化层面作为框架的

深层根基，为多模态交际提供意义生成的土壤，其通过意识形态与体裁的双重作用，制约着语境、内容及表达层面符号的选择与组合。语境层面聚焦交际发生的即时情境，由话语范围、话语基调及话语方式等要素动态交互，也会受到文化层面的宏观制约。内容层面主要划分为两个维度，话语意义通过概念意义、人际意义与谋篇意义的结合，将抽象语境要素转化为具体的语义单元，其实现方式受制于语境的三个要素；形式维度则依托视觉、听觉、触觉等模态系统，将话语意义转化为可感知的符号形式。表达层面作为符号的载体，整合语言与非语言符号，确保多模态意义的精准外化。这一层级结构为多模态话语的生成与解析提供了从文化到符号的完整研究路径。基于上述理论框架，本文以《叶问》为分析对象，从文化层面、语境层面、内容层面和表达层面，进行系统地多模态文本分析，以期提升字幕外译质量开拓新路径。

3. 《叶问》翻译的多模态话语与受众反馈分析

中国影视作品的跨文化传播面临一个现实挑战：文化精品的创作周期漫长，例如《哪吒之魔童闹海》耗时五年才完成制作。这种“慢生产”与社交媒体时代信息传播的“快节奏”需求存在矛盾。因此，推动中国电影“走出去”的关键策略之一，在于盘活现有经典资源，通过二次传播扩大国际影响力。这些作品已通过本土市场验证，但如何突破文化距离、触达海外受众是一个难题。调研显示，当前最普及的海外传播方式是通过平台推荐、意见领袖的推广与二次传播[3]。其中，电影 reaction(观看反应视频)在一定程度上突破了版权对海外电影宣发的限制，又因其独特的即时性和互动性，为研究跨文化接受提供了鲜活素材。社交媒体平台的即时反馈机制，也使得这类内容能够快速积累受众讨论。

选择《叶问》作为研究案例，则有以下原因。首先，动作类型的受众基础与明星资本效应。根据《2021年度中国电影国际传播调研报告》[4]，华语动作片以30%的偏好度位居海外网络受众最受欢迎类型榜首。本片以李小龙老师叶问为原型，李小龙与主演甄子丹作为文化符号所具备的附加价值也有效降低了跨文化传播的认知门槛。其次，本土化创作与外宣价值的平衡。相较于《卧虎藏龙》等为迎合西方市场定制的“特供片”，《叶问》系列是面向中国观众的本土创作，对其翻译策略的讨论更贴近“文化外宣”的核心诉求——如何让原生于本土语境的作品被海外观众理解并共情。第三，文戏与武戏的协同深化。影片不仅以武打动作作为卖点，更通过家国情怀、民族气节等主题刻画(如影评人周黎明所言“以拳脚写史诗”)，使其在武打片类型中兼具思想深度，为翻译中的文化诠释提供了优质样本。

本研究通过多模态话语分析与受众研究的互补性构建翻译研究的分析框架。多模态话语分析从文化、语境、内容、表达四层面解构文本的多模态符号系统，受众研究则聚焦于受众的能动性解读(本文选取了数十位百万粉丝量或浏览量的 YouTube 自媒体创作者，即意见领袖，作为观察对象)。二者结合形成“生产-传播-接受”的闭环分析：一方面，多模态话语分析揭示意义构建的内在逻辑；另一方面，受众反馈能够反映翻译策略和实践的效果。后续案例分析将基于此框架，从模态分析、受众反馈到策略调适逐层推进，以期探索出字幕外译新模式。

3.1. 文化和语境层面

文化层面以意识形态与语篇体裁为核心。视听翻译虽受时空与语境制约，但借助多模态协同，能够有效调和文化差异。情景语境包括时空、场景、人物关系等要素。视听翻译依托多模态互补，通过对语境的深刻认识，实现跨文化交际的有效传递，确保信息在目标语境中的自然流畅与情感共鸣[5]。

本片一场戏中(图1)，廖师傅慕名拜访叶问，以抱拳作揖的武术礼仪向叶问请教，背景音包含鸟鸣与角色衣物摩擦的自然环境声。对白部分，廖师傅以“切磋武艺”表明来意，字幕将其译为“practice”。值得注意在此前武馆晨练场景中，“练功”一词同样被译为“practice”，两个不同文化概念(竞技性武艺

交流与日常训练)在字幕中共享了同一词汇符号。视觉模态上,廖师傅表情谦和,肢体动作(如抱拳礼、微倾的身体姿态)与叶问的礼貌回应构成互动,听觉模态中环境音暗示场景的公开性与非对抗性,与尔后比武时场景的封闭性形成了对比。



Figure 1. Master Liao visits Ip man
图 1. 廖师傅拜访叶问

翻译上,“practice”指代的概念造成了一定的困惑,媒体创作者@ Reel Reject (以下意见领袖均简称为@XX)误将廖师傅的切磋请求理解为“拜师”,认为其恭敬夸赞叶问武功修为后提出“practice”是寻求建立师徒关系。这一误读暴露了翻译在多模态协同中的问题:文本模态方面,译者未区分“切磋”与“练功”的文化差异(前者强调对等竞技,后者指向个人训练),而视觉模态中的抱拳礼、表情互动等东方武术礼仪符号未能通过字幕得到解释。当“practice”同时覆盖两种场景时,目标受众仅能依赖自身文化经验填补语义空缺,这就导致语义发生偏离,将“切磋”误解为“拜师”。

下面在金山找率众踢馆的场景中(见图 2),人物行为与场景符号呈现显著差异。金山找虽也做出抱拳作揖的武术礼仪,但其站姿、神态与中气十足的对白语气,配合身后四名师弟双臂环抱,显示保守和敌对的肢体语言,共同营造了危机氛围。背景环境从廖叶切磋时的庭院鸟鸣转为武馆街的喧闹人声,暗示事件的公开性与群体围观压力。但是译者仍将“切磋”译为“practice”,但此时金山找对佛山武师产生的威胁感、背景音乐中隐约的鼓点节奏所共同构建出的情景语境,与廖叶切磋时非对抗、自然的语境形成了鲜明对比。



Figure 2. Jin Shan Zhao challenges masters
图 2. 金山找踢馆

观众通过前一场武戏已建立“practice”与“比武”的初步关联,在此场景中结合多模态符号(如人物设计的攻击性、环境冲突暗示)主动修正了对“practice”的理解。@Lost in Nostalgia 指出:“这根本不是‘practice’,而是擂台挑战”——这一反馈印证了多模态系统的动态补偿机制:当文本模态的翻译(两场“切磋”均译作 practice)阻碍语义精确传递时,视觉模态的对抗性符号(挑衅表情和肢体压迫)、听觉模态

的环境提示(喧闹声和鼓点)与观众不断累积的观影经验形成交互,进而主动将“practice”从“日常训练”引申为“公开挑战”的语境含义。这种跨模态的意义重构,既暴露字幕翻译的无注性导致的其在文化概念区分上的受限,也凸显多模态系统的动态补偿能力。

前两场戏分别代表着佛山门派竞争,南北地域矛盾。随着电影来到下一幕,日本侵华的时代背景已经成为毋庸置疑的主旋律。由此可见,第三场“切磋”戏的核心矛盾源于历史语境的根本性转变。场景中(见图3),作为日军翻译的李钊代表日军将领三浦,邀请中国人“比武会友”。这里的视觉模态呈现矛盾性:李钊西装革履、语气温和的姿态与身后荷枪实弹的日本士兵形成鲜明对照。听觉模态上,低沉悬疑的旋律与李钊斯文友善的对白产生反差,营造出紧张的戏剧性。译者捕捉到了语境转换对这里的“切磋”有了一定程度的颠覆,将“切磋”改译为“fight with”,而非延续之前的“practice”,试图通过程度升级凸显日军以比武会友为名实施恐怖压迫的意图。



Figure 3. Li Zhao's invitation

图3. 李钊的邀请

“fight with”兼具“与某人打架”和“与某人并肩作战”两层含义,陌生语境下容易产生歧义,进而导致跨文化误读。例如@Reel Reject 误以为日本人寻找会武功的中国人为日本效力,因此将报名比武的中国人视为与日军同流合污(traitor)。这一误判源于译者翻译策略的转换失当:前两场戏通过多模态协同已使观众接受“practice”是作为“对抗性比武”的一种委婉表达。第三场突然改用“fight with”,虽可以看到译者试图区别特殊语境的能动性,却割裂了观众已积累的符号解读经验。这也证明了仅通过主观判断,而调整策略,对某处进行片面的修改无法满足视听翻译的整体需要。

现综合以上分析,提出两种调适本例翻译策略的方案。第一种方案采用统一化翻译策略,将三处“切磋”均译为“spar with”,“spar”能够表示非正式比武较量。术语的一致性依托多模态系统的平衡能力和观众对冲突升级的共性理解,不仅可以避免因“一词多义”导致认知混乱,还不会引起跨文化误读。第二种方案则实施差异化补偿策略,根据场景意识形态的特殊性分级调整:廖叶切磋保留“spar with”,能够体现“非正式”的隐含规则(异化翻译);金山找踢馆转为“challenge”(归化翻译),明示踢馆文化的公开性;日军胁迫场景则译为“duel”(补偿翻译),补足画面中枪械等符号信息。该方案以词汇分级降低海外观众的理解负担。无论选择何种路径,均需立足动态化、全局化的分析框架,了解受众心理的认知路径,再进行翻译改善。

3.2. 内容和表达层面

内容层面以概念、人际与谋篇意义为枢纽。视听翻译虽不同于一般文本翻译,其通过大量的非语言符号,动态补偿文本阐释的连贯性,构建整体话语的意义。表达层面由语言媒体和非语言媒体构成。视听翻译借助视听符号的多模态形式重构话语意义,强化多模态文本的接受效果[6]。

开场旁白对于佛山的背景介绍以繁体呈现(见图 4), 同时在视觉上通过字体的残破设计强化年代感。背景音使用二胡等传统乐器演奏, 听觉模态展现侠义的文化基调。文本模态中(详见表 1), 说明性文本信息量大, 视觉上占据整体画面的大量比例, 没有足够空间提供给需要空间坐标跟踪的硬字幕, 又受限于字幕瞬时性(8.7 秒时长), 综合考虑译者依旧选择将字幕固定于底部区域, 并调小了字号。同时, 译者对原文进行缩略处理, 例如删除“尚武成风”等修饰性表述, 保留核心信息(如“佛山, 武术之乡”)。字幕将“南派武术”译为“Southern China martial arts”, 未补充文化注释, 仅传递地理方位概念。



Figure 4. Opening scene of the film
图 4. 开场动画

Table 1. Statistical analysis of information in three sets of subtitles
表 1. 三组字幕的信息统计

	区间	字符数	总持续时间	每秒字符数
开场动画	01:45~01:54	228	00:00:08.743	26.1
过场动画	34:57~35:06	223	00:00:08.512	26.2
片尾动画	01:42:07~01:43:16	567	00:00:49.910	11.4

通过观察可知, 多数观察对象已经建立了对佛山城市形象的认识, 但依然存在一些消极反馈。@TBR Schmitt 因字幕信息密度过高暂停影片进行详读, 并对“Southern China”的译法产生困惑, 误以为中国的南北方存在类似朝鲜与韩国的政治分裂, 没有认识到源文本实际表达的武术流派的地域差异。这一误解表明, 针对该文本的直译导致观众仅能依赖自身认知框架来填补语义空缺。

译者使用的缩略策略(如删除“尚武成风”等修饰词)在控制字幕信息密度、提升跨文化可读性上已取得基础成效。那么针对“南派武术”等文化负载词, 在不影响剧情理解的情况下, 能否进一步缩略为“佛山是中国武术的主要发源地”, 既可规避“Southern China”可能引发的政治误读, 又能通过简化术语减轻观众认知负荷, 使字幕更适配时空限制下的效率需求。但需注意, 调适后的这一策略是否合理, 需于实际测试中重新观察受众对南北流派冲突的情节接受度是否提高。

第二幕过场动画通过视觉、听觉与文本模态的叠加表现战争背景(见图 5)。视觉模态以日军战机轰炸镜头引出说明性文字; 听觉模态通过炮火和机枪扫射的背景声强化硝烟感。文本模态中直译了源语中加强情感的语言符号(如“卢沟桥事变爆发!”保留感叹号), 字体延续年代感设计以契合历史基调, 字幕区域未对高密度展现战争惨烈的程度词(如“民不聊生”)进行缩略, 略显拥挤。海外观众基于二战的共性记忆可捕捉场景基调, 把握时代背景。但对日军侵华缺乏基本认知, 甚至@Reel Reject 在此前一直误认为是中国侵略日本。这种错误认知反映出我国外宣传播的不足。



Figure 5. Transition scene (excerpt)

图 5. 过场动画(节选)

第一幕背景字幕的缩略策略(删除“尚武成风”等修饰词)虽契合技术性精简需求,但第二幕战争描述中涉及日军暴行的文本(如“尸横遍野”“烧杀抢掠”)不可因瞬时性的可读问题而简化,因其“无实意却重情感”的表述正是外宣翻译的核心载体。在视听翻译的时空限制与外宣翻译的本质诉求发生冲突时,应以国家形象构建与历史真实性的有效传播为优先考虑,只有这样,视听翻译方能突破文化壁垒,向世界传递真实的声音[7]。

影片在民众反抗日军,涌向叶问后结束。转场后来到了片尾动画部分。片尾采用类似幻灯片的蒙太奇形式,生动地体现出多模态叙事的结构特点。视觉模态上,通过真实历史照片的线性排列构建时间脉络(见图 6),照片内容从宏观历史过渡到人物微观叙事,构建出集体记忆与个人叙事的双重闭环。背景音以弦乐为主,强化了历史沧桑感与厚重感。文字模态上,说明性文字没有沿用陈旧感的设计,而改用简约的楷体,虽增强现代感却与老照片的怀旧质感形成风格割裂。字幕位置沿用对白字幕的一般形式,即固定于画面底部,并未与画面文字同步排版与字体风格。



Figure 6. Epilogue scene (excerpt)

图 6. 片尾动画(节选)

观众在观影结束时的情感状态显著影响了接受效果。在电影情绪高点结束后的短时间内,观众会处于叙事满足感与信息渴求度的峰值交汇期,往往对表达层面出现的纰漏或瑕疵有极强的包容。尽管存在字幕遮挡画面、字体风格不统一等问题,但受访者更关注内容信息本身而非呈现形式。观察对象的评论多聚焦历史(反法西斯抗战的艰辛)与文化(李小龙),无人批评字幕排版问题,这种反馈倾向揭示出观众在观影末期更倾向于情感认同而非理性审视。

在视听翻译的多模态语境下,字幕技术的优化存在独特的跨模态补偿机制。理论上,说明性文字需通过内嵌、空间定位捕捉、字幕特效等手段,实现字幕与其他模态的时空协同与风格协调[8]。但本片反

馈结果显示，当泛黄的历史影像与厚重弦乐共同构建时代氛围时，即便字幕未做旧化处理，观众仍能通过多模态符号补足对年代感的意识，这种感官补偿使得形式适配成为非强制需求。这揭示出视听翻译相较于传统文本翻译的稳定性优势：其质量存在跨模态容错区间，模态的局部缺陷可由其他模态进行补偿。这种多模态特性同时也为翻译创新提供了可能：在满足信息保真的前提下，译者可借力媒体特性开发超文本表达。例如魔幻题材中咒语字幕采用光效动态渲染，创造出纸质媒介无法承载的沉浸体验。模态协同的延展性，使得视听翻译始终处在受技术赋能与理论创新影响的最前沿，其本身也会随着技术的发展焕发出无限可能。

4. 受众反馈驱动的字幕外译模式

通过多模态话语各层面的分析，本文厘清了翻译在意义建构中的逻辑机制；而受众反馈则反映出翻译策略与实践的实际效果，有助于在明确逻辑基础上对翻译活动进行更为现实和有效的认知。二者共同为翻译策略的进一步优化提供了系统的借鉴路径。本文从模态分析、受众反馈到策略调适，构建了一个将优化后的翻译产品投入新一轮实践检验、不断迭代完善的良性循环模式。在此基础上，结合字幕外译的特点，明确外译受众，提出了一种由受众反馈驱动的外译模式(见图 7)：在前期准备阶段，译者需基于多模态话语分析，对源文本的视觉、听觉和语言等模态进行系统性解构，同时结合受众中心理论对目标受众的文化认知、语言习惯及观影心理进行预判分析。在翻译实践中，译者应巧妙利用多模态的补偿机制，选择恰当的翻译策略。翻译产品输出后，需整合受众反馈，评估翻译质量和产品接受度，进而优化翻译策略，提升产品质量[9]。

该模式突破传统文本翻译的线性结构，将多模态文本分析与受众接受心理研究相结合，充分体现了受众在翻译过程中的核心地位，能够不断优化翻译产品与服务，有效避免与市场需求脱节，从而确保外宣翻译的传播效力。

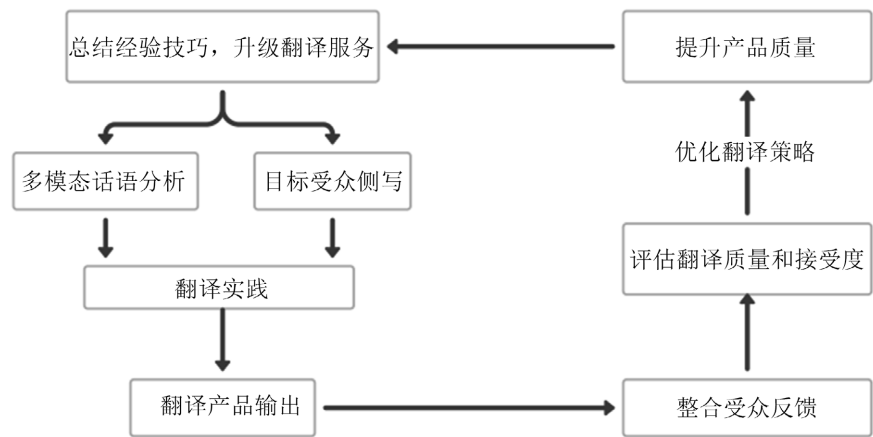


Figure 7. Audience feedback-driven translation model
图 7. 受众反馈驱动的翻译模式

5. 结束语

在新一轮全球化浪潮中，中国故事的对外传播面临诸多挑战。作为关键的评价机制，受众反馈为翻译服务的优化提供了实证依据。本研究发现，将受众的真实需求系统性地融入翻译实践，并通过多模态视听符号的深度分析重构语境，不仅有助于提升翻译文本的接受度，更能促进传播从单向的“文化输出”向双向互动的“价值共享”转变。这一基于反馈机制的翻译路径，不仅增强了外宣作品的传播效能，也为构建

国际传播的话语新格局提供了方法论支撑。反馈驱动的翻译模式具有良性循环的潜力，可持续推动翻译产品迭代优化。未来研究可进一步拓展多语种、跨文化语境下的受众调研，系统构建适用于不同语境的反馈收集与分析体系；同时，探索人工智能辅助下的多模态翻译策略，也将为规范化、智能化的外译实践提供新路径。通过这些深入的理论探索与技术实践，全球受众将更全面、真实地理解和感知中华文化。

参考文献

- [1] 武建国, 李育静. 多模态语境重构与中国影视文化的传播——以影片《我和我的祖国》字幕翻译为例[J]. 山东外语教学, 2024, 45(2): 11-21.
- [2] 张佳优. 多模态话语分析视角下纪录片字幕翻译研究——以《我在故宫修文物》英译版为例[J]. 今传媒(学术版), 2020, 28(2): 116-118.
- [3] 黄会林, 郭星儿, 蒋正邦. 中国电影海外传播平台传播效果研究——2023 年度中国电影国际传播调研报告[J]. 现代传播(中国传媒大学学报), 2024, 46(1): 39-48.
- [4] 黄会林, 黄昕亚, 祁雪晶. 中国电影海外网络受众接受度的实证研究——2021 年度中国电影国际传播调研报告[J]. 现代传播(中国传媒大学学报), 2022, 44(1): 74-81.
- [5] Jewitt, C., Bezemer, J. and O'Halloran, K. (2016) *Introducing Multimodality*. Routledge.
<https://doi.org/10.4324/9781315638027>
- [6] 辛红娟, 陈可欣. 多模态话语分析视角下外宣纪录片翻译研究——以《四季中国》为例[J]. 对外传播, 2020(2): 54-56.
- [7] 李燕, 麻争旗. 中国电影海外传播的字幕困境及其外译理念的革新[J]. 电影评介, 2021(7): 20-25.
- [8] 俞可平. 认知语境下的中文纪录片字幕英译与制作的效度研究——以《港珠澳大桥》记录电影为例[J]. 翻译研究与教学, 2021(2): 23-28.
- [9] Suojanen, T., Koskinen, K. and Tuominen, T. (2015) *User-Centered Translation*. Routledge, 143-147.
<https://doi.org/10.4324/9781315753508>