

多模态视角下纪录片《舌尖上的中国》字幕翻译研究

彭 婷

长沙理工大学外国语学院, 湖南 长沙

收稿日期: 2026年8月8日; 录用日期: 2026年9月1日; 发布日期: 2026年9月15日

摘 要

纪录片《舌尖上的中国》不仅是一部展示中华美食文化的影视作品, 更是传播中国传统文化与生活哲学的重要载体。本文以多模态话语分析理论为基础, 从文化层面、语境层面、内容层面和表达层面四个角度, 系统探讨该纪录片在英文字幕翻译过程中所面临的挑战与应对策略。通过对具体字幕实例的分析, 揭示字幕翻译如何在多模态语境中实现意义的重构与跨文化传播。

关键词

多模态话语分析, 字幕翻译, 舌尖上的中国, 跨文化传播

A Study of Subtitle Translation in the Documentary *A Bite of China* from a Multimodal Perspective

Ting Peng

School of Foreign Languages, Changsha University of Science and Technology

Received: August 8, 2026; accepted: September 1, 2026; published: September 15, 2026

Abstract

The documentary *A Bite of China* is not only a visual presentation of Chinese culinary culture but also a vital medium for conveying traditional Chinese culture and life philosophy. Based on the theory of multimodal discourse analysis, this paper systematically explores the challenges and strategies involved in the English subtitle translation of the documentary from four perspectives: cultural, contextual, content, and expressive levels. Through the analysis of specific subtitle examples, the

study reveals how meaning is reconstructed and cross-cultural communication is achieved within a multimodal context.

Keywords

Multimodal Discourse Analysis, Subtitle Translation, *A Bite of China*, Cross-Cultural Communication

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

随着全球化与数字媒介的发展,中国纪录片逐渐走向国际舞台,成为传播中国文化、讲好中国故事的重要手段。其中,纪录片《舌尖上的中国》以其深厚的文化底蕴、细腻的人文关怀和精美的视觉呈现,受到了国内外观众的广泛关注。该片通过对中华饮食文化的描绘,展现了中国人对食物的热爱、对土地的敬畏以及对亲情与传统的珍视。然而,在跨文化传播过程中,语言并非信息传递的唯一渠道,图像、声音、音乐、动作等非语言符号同样在意义建构中起着不可忽视的作用。这就要求字幕翻译不仅关注语言的转换,更需综合考虑多种模态之间的协同作用。

多模态话语分析(Multimodal Discourse Analysis, MDA)作为一种新兴的跨学科研究方法,强调语言与其他模态资源(如图像、声音、动作等)之间的互动关系。冯德正、张德禄指出,多模态话语中的意义并非仅由语言单独建构,而是语言、图像、声音等多种符号资源共同作用的结果[1]。在纪录片这一以视听为主要传播手段的语篇类型中,多模态理论为字幕翻译提供了新的视角和分析框架,有助于译者深入理解原语信息的多层次意义,从而更有效地实现跨文化传播。

目前,国内关于《舌尖上的中国》的字幕翻译研究多聚焦于语言层面的翻译策略,如归化与异化、意译与直译的运用等[2],尚缺乏从多模态话语角度出发的系统探讨。因此,本文尝试以多模态理论为指导,围绕文化、语境、内容与表达四个层面,分析《舌尖上的中国》中典型字幕翻译案例,探讨其在多模态互动语境下的翻译特征与策略,旨在为纪录片字幕翻译提供理论支撑和实践借鉴,推动中国文化更好地“走出去”。

2. 多模态理论框架

多模态话语分析是在系统功能语言学和社会符号学基础上发展起来的研究路径,强调意义并非仅由语言建构,而是由语言、图像、声音、动作等多种符号资源共同完成。Halliday 的系统功能语言学为多模态意义研究提供了功能语言学基础[3],Kress 等学者则从社会符号学角度进一步强调不同模态具有各自的意义潜势,并通过相互协同参与整体意义建构[4]。在此基础上,张德禄[5]构建了多模态话语分析综合理论框架,从文化、语境、意义、形式和媒体等层面对多模态意义建构进行系统分析,在具体研究中可进一步从文化、语境、内容和表达四个方面考察不同模态之间的关系。

纪录片字幕翻译具有明显的多模态特征。字幕并非独立的语言文本,而是与画面、人物动作、旁白、音乐和环境音等共同参与意义传递。同时,字幕还受到屏幕空间、呈现时间和观众阅读速度等限制[6]。因此,本文以多模态话语分析理论为基础,从文化、语境、内容和表达四个层面对《舌尖上的中国》英文字幕进行分析,重点考察语言模态与视觉、听觉模态之间的协同关系,以及这种关系对字幕信息选择、

意义重构和跨文化传播的影响。

(1) 文化层面

文化层面是多模态交际展开的重要基础，主要涉及意识形态和体裁结构潜势。前者体现特定文化中的价值观念、思维方式与生活习惯，后者则影响这些文化意义在具体语篇中的组织和表达。在纪录片中，文化意义并非仅依靠语言传递，还可以通过饮食、节庆仪式、人物关系、地域环境等视觉和听觉符号共同建构[7]。

在字幕翻译中，文化层面的关键在于协调文化信息的保留与目标语观众的理解。译者需要结合画面、声音等模态判断文化信息是否已经得到呈现，并据此采取保留、解释或适度调整等方式。例如，“年夜饭”“腊八粥”等文化负载词既具有鲜明的中国文化特征，又可能给目标语观众造成理解障碍，因此译文既要保持其文化识别度，也应借助必要的解释或模态互补降低理解难度，从而实现文化意义的有效传递。

(2) 语境层面

语境层面关注交际活动发生的具体情境，包括话语内容、参与者关系、交际方式以及时间、空间等因素。多模态语篇中的语境并非单纯依赖语言形成，而是由画面、声音、人物动作和语言等多种符号共同建构。对于纪录片字幕翻译而言，译者需要结合人物身份、场景变化、行为方式及视听信息判断字幕的具体意义，避免脱离语境进行孤立的语言转换。例如，在节庆、家庭聚会或地域文化场景中，同一词语可能因画面和人物关系而具有不同的情感色彩，因此字幕表达需要与具体情景保持协调。

(3) 内容层面

内容层面主要关注多模态语篇中意义的建构、信息的组织以及不同模态之间的关系。张德禄[5]指出，不同模态之间既可能形成互补关系，也可能存在非互补关系。在纪录片中，语言、图像、音乐和环境声音通常共同承担信息传递功能，因此字幕无需机械复制源语中的全部信息。当画面已经清楚呈现人物动作、地点或食物形态时，译文可以适当压缩重复信息；而当视觉信息不足以帮助目标语观众理解文化含义时，则可通过语言补偿增强信息的完整性。由此，字幕翻译实际上是根据不同模态的信息分布进行选择 and 重新组织的过程。在字幕翻译中，译者应综合分析不同模态间的关系，尤其关注互补关系下的信息分布情况。字幕作为语言模态的体现，其内容应根据其他模态所呈现的信息进行取舍与补充。例如，若画面已直观呈现某个信息(如地点或动作)，则字幕可省略冗余内容；反之，在画面信息不足时，字幕应适当增补语言说明，以弥补视觉模态的空白，从而提升观众的整体理解效率。

(4) 表达层面

表达层面主要考察意义通过语言与非语言符号实现的具体方式，包括字幕语言、语音语调、人物动作、画面构图、音乐及环境声音等。对于影视字幕而言，表达方式还受到屏幕空间、字幕持续时间和观众阅读速度等因素的限制[6]。因此，译者不仅需要考虑词汇和句式是否准确，还应关注译文的长度、节奏和语言风格能否与画面及声音相协调。在节奏较快的场景中，字幕通常需要适当压缩以保证可读性；而在情感性较强的叙事中，则应在有限篇幅内尽可能保留原语的情感色彩和语言风格，从而实现语言模态与其他视听模态之间的协调。

3. 《舌尖上的中国》的多模态语篇特征及其翻译制约

《舌尖上的中国》以中国饮食文化为叙事核心，通过语言、图像、声音等多种模态的协同作用，将食物制作、地域生活与人物情感联系起来。因此，其字幕翻译并不是对解说词的孤立转换，而需要根据不同模态之间的信息分布重新组织语言意义。译者既要利用画面和声音提供的辅助信息，又要考虑字幕的时空限制，在信息完整性、文化传递和观众接受之间进行选择。

从视觉模态来看，纪录片大量使用食材特写、烹饪过程、人物活动及地域景观等镜头。画面不仅承

担审美功能,也直接提供人物、地点、动作和食物形态等信息[8]。这种视觉信息可以减轻字幕的语言负担,使部分已经由画面明确呈现的内容得到压缩或省略。然而,视觉模态虽然能够展示食物“是什么”以及人们“怎样做”,却未必能够独立解释食物背后的节令传统、家庭伦理和文化象征。因此,译者仍需判断哪些文化意义能够由画面补偿,哪些信息需要通过字幕进一步显化。字幕的删减由此并非简单的语言缩短,而是一种建立在模态信息重新分配基础上的意义选择。

从听觉模态来看,旁白、人物话语、背景音乐和环境声音共同参与影片情感氛围的建构。尤其是解说词较为舒缓、含蓄的语调以及音乐节奏的变化,可以赋予语言超出字面意义的情感色彩。对于目标语观众而言,这些非语言声音能够部分保留原片的情绪基调,但字幕本身难以完整复制汉语语音中的节奏、语气和修辞效果。因此,当具有抒情色彩的中文解说被压缩为简洁的英文字幕时,虽然信息传递可能更加直接,却也可能造成情感浓度和语言意象的弱化。这说明听觉模态既能够对字幕形成补偿,也可能凸显源语声音与目标语文字之间无法完全对等的问题。

从语言模态来看,《舌尖上的中国》的解说词兼具知识介绍和情感叙事功能,其中包含大量饮食文化词汇、地域表达及具有文学色彩的语句。此类语言在汉语文化语境中往往能够借助观众已有的文化知识形成联想,而目标语观众未必具有相同的认知背景。因此,字幕翻译不能仅以字面对应作为评价标准,而需要考察译文是否重新建立了目标观众理解原语意义所需要的语境。必要的解释和显化能够降低文化理解障碍,但过度解释又可能削弱原有表达的含蓄性与文化陌生感;相反,过多保留源语形式虽然能够突出中国文化特征,却可能增加目标观众的理解负担。

由此可见,《舌尖上的中国》的字幕翻译实际上是语言、视觉和听觉信息共同制约下的多模态再建构过程。译者并非被动地“服务”原文,而是在字幕时空限制和目标观众认知差异的影响下,对原语信息进行保留、压缩、补偿和重组。这些处理既可能提高字幕的可读性和跨文化传播效果,也可能造成部分文化内涵、修辞意象和情感意味的损失。因此,对该纪录片字幕翻译的评价不能仅停留在译文是否准确或是否具有“中国特色”,还应进一步考察不同翻译选择对目标语意义产生了何种改变,以及这种改变是否有助于语言模态与其他视听模态之间形成有效协同。

4. 多模态话语分析在纪录片中的具体体现

《舌尖上的中国》是一部视觉、听觉与语言多模态信息融合的典范纪录片,其字幕翻译不仅要传达语言信息,更要体现文化意涵与视觉配合的默契。本节将从前述四个层面分析具体字幕翻译策略。

(1) 文化层面的字幕翻译

纪录片中的饮食文化不仅体现在食物本身,更蕴含着家庭观念、节庆仪式、地域信仰等深层文化意涵。文化层面字幕翻译的核心在于实现跨文化理解与认同,避免“文化缺席”或“误读”。

例 1:

原文:“除夕夜,一家人围坐在一起吃年夜饭。”

字幕翻译:“On Chinese New Year's Eve, the whole family gathers for a reunion dinner.”

分析:

该句中的“年夜饭”具有浓厚的中国文化内涵,是春节最重要的家庭仪式之一。若译为“New Year's Eve dinner”可能会让外国观众误以为是西方跨年晚宴,缺乏文化区分度。因此翻译时添加“中国”限定词并使用“reunion dinner”强调团圆含义,较好地保留了文化特性。

例 2:

原文:“腊八粥是冬季的传统节令食物。”

字幕翻译: “Laba porridge is a traditional Chinese dish enjoyed on the eighth day of the twelfth lunar month.”

分析:

“腊八粥”直接音译为“Laba porridge”，保留了独特文化标识；后半句通过解释“the eighth day of the twelfth lunar month”补充背景知识，实现文化信息的补偿式传递，是一种典型的“归化 + 异化”结合策略。

(2) 语境层面的字幕翻译

语境层面的翻译强调字幕与人物关系、具体场景及视听氛围之间的协调。画面与声音不仅能够补充语言信息，也会影响观众对字幕情感和语义的理解。

例 3:

场景: 一位年迈的母亲在为子女准备饭菜，背景配以温柔音乐与昏黄灯光。

原文: “这是她每天最期待的时刻。”

字幕翻译: “This is the moment she looks forward to the most every day.”

分析:

译文基本保留了原句的情感意义，其中“the moment”与母亲准备饭菜的画面、柔和的音乐和灯光共同强化了这一日常行为的情感价值。这里的温情并非完全由字幕语言产生，而是语言、画面和声音共同建构的结果。因此，译文采用较为直接的表达反而避免了过度解释，使观众能够借助其他模态理解人物情感。

例 4:

场景: 介绍新疆地区饮食习惯，镜头扫过开阔草原与烤全羊场景。

原文: “这里的人热情奔放，食物也充满了力量。”

字幕翻译: “People here are full of passion, and so is their food.”

分析:

“full of passion”与人物活动及开阔的地域画面相呼应，“so is their food”则将人物性格与食物特征联系起来，使字幕保持了原文的修辞结构。但“力量”被处理为“passion”，实际上将“强烈、有力量感”的食物特征转化为“热情”，语义发生了一定偏移。虽然这种转换增强了字幕与人物形象的连贯性，却也弱化了原文对食物感官特征的描写，体现出语境协调与语义准确之间的权衡。

(3) 内容层面的字幕翻译

内容层面主要关注不同模态之间的信息分配。由于画面能够直接呈现部分动作和事物，字幕可以适当压缩重复信息；但译者的增补和重组也可能改变原文的信息重点。

例 5:

原文: “煎、炸、炖、煮，各种烹饪方式让食物拥有千变万化的口感。”

字幕翻译: “Frying, deep-frying, stewing, and boiling — each method creates a unique texture and flavor.”

分析:

译文保留了四种烹饪方式，并借助“each method creates...”重新组织信息，使字幕结构更加紧凑。值得注意的是，“千变万化”被简化为“unique”，而“口感”被扩展为“texture and flavor”，说明译者并未完全按照原文进行对应，而是根据英语表达习惯重新分配信息。结合烹饪过程的视觉画面，这种调整有助于观众快速理解，但也使原文强调的丰富变化程度有所减弱。

例 6:

画面中展示包饺子的过程, 未出现旁白。随后字幕显示:

字幕翻译: “Making dumplings is a cherished family ritual during Spring Festival.”

分析:

如果原片此处确实没有对应的语言信息, 那么这一字幕实际上超出了普通语言转换的范围, 更接近对视觉信息的解释性增补。译文将画面中的“包饺子”进一步解释为春节期间的“family ritual”, 能够帮助不了解中国文化的观众理解其家庭与节庆意义, 但同时也加入了画面本身未必能够完全确认的文化解释。因此, 这种增补体现了多模态再语境化的作用, 也说明译者在补偿文化信息时需要避免对视觉意义作过度阐释。

(4) 表达层面的字幕翻译

表达层面主要考察字幕的语言形式能否与影片节奏和视听风格相协调。由于字幕受到时间和屏幕空间限制, 译文通常需要在语言简洁性与原文风格再现之间进行平衡。

例 7:

原文: “滚烫的汤底, 加入牛骨慢炖数小时, 香气四溢。”

字幕翻译: “A boiling broth, simmered for hours with beef bones — rich and aromatic.”

分析:

译文采用较短的结构, 将制作过程压缩为“simmered for hours with beef bones”, 并以“rich and aromatic”概括汤底的感官特征, 与画面中的热气和食物特写形成呼应。这种压缩符合字幕快速阅读的需要, 但“rich”属于一定程度的意义增补, 而原文中的“滚烫”与“香气四溢”也被重新组织。因此, 该译法在保持整体感官效果的同时, 对原文信息进行了适度重构。

例 8:

原文: “一碗面, 不仅是一种食物, 更是家的味道。”

字幕翻译: “A bowl of noodles is more than food — it’s the taste of home.”

分析:

译文以“more than food”对应“不仅是一种食物”, 并保留“the taste of home”这一核心隐喻, 使食物与家庭情感之间的联系得到延续。破折号形成的停顿也与原文的转折节奏较为一致。在背景音乐和人物画面的配合下, “home”能够进一步激活亲情与归属感, 因此这一译法较好地利用了语言与视听模态之间的情感互补。不过, “家的味道”所包含的中国家庭饮食经验难以完全由“the taste of home”复制, 其文化内涵仍存在一定程度的泛化。

综合以上四个层面的分析可以发现, 《舌尖上的中国》的字幕翻译并不是对原语信息的机械转换。视觉和听觉模态可以补偿字幕中被压缩或省略的信息, 而字幕本身也会通过增补、转换和重新组织改变部分意义。因而, 对字幕翻译效果的评价既要考察不同模态之间是否形成有效协同, 也应关注这种协同过程中产生的信息强化、弱化与文化意义变化。

5. 结语

《舌尖上的中国》的字幕翻译并非单纯的语言转换, 而是在语言、画面、声音等多种模态共同作用下进行的意义重构。本文以多模态话语分析理论为基础, 从文化、语境、内容和表达四个层面对相关字幕案例进行了分析。研究发现, 视觉和听觉模态能够为字幕提供重要的语境补偿, 使译者可以对部分重

复信息进行压缩或省略；与此同时，对于文化负载词、情感表达及具有特定文化内涵的信息，字幕又需要通过保留、解释或适度增补帮助目标语观众建立必要的理解语境。

从本文所分析的案例来看，多模态字幕翻译并不存在绝对理想的转换方式。信息的压缩能够提高字幕的简洁性和可读性，却可能造成文化意涵或修辞效果的弱化；解释性增补有助于降低目标观众的理解障碍，但也可能增加阅读负担，甚至改变原文的信息重心。因此，对字幕翻译效果的评价不能仅停留在语言是否准确、表达是否流畅，还应关注译文在与其他模态协同过程中产生的信息强化、弱化及意义变化。译者需要根据具体画面、声音、字幕时空限制以及目标观众的文化认知，对原语信息进行有选择地保留与重组。

对于中国纪录片的国际传播而言，这意味着字幕翻译不应一味追求文化信息的完整解释，也不宜为了目标语的自然流畅而过度削弱原有文化特征。当画面能够承担解释功能时，字幕可以适当简化语言信息；而当文化意义无法通过视觉或听觉模态直接获得时，则应保留必要的文化标识并进行适度补偿。只有根据不同模态的信息功能和目标观众的接受需求动态调整翻译策略，才能在字幕有限的表达空间中兼顾可理解性与文化特征，使中国饮食文化及其背后的生活观念获得更有效的跨文化传递。

参考文献

- [1] 冯德正, 张德禄. 多模态语篇分析的进展与前沿[J]. 当代语言学, 2014(1): 88-99.
- [2] 陈楠楠. 多模态话语分析视角下影视字幕翻译[J]. 文学教育(上), 2020(8): 134-135.
- [3] Halliday, M.A.K. and Matthiessen, C.M.I.M. (2014) *Halliday's Introduction to Functional Grammar*. 4th Edition, Routledge.
- [4] Kress, G. (2010) *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge.
- [5] 张德禄. 多模态话语分析综合理论框架探索[J]. 中国外语, 2009, 6(1): 24-30.
- [6] Díaz Cintas, J. and Remael, A. (2021) *Subtitling: Concepts and Practices*. Routledge.
- [7] Dash, A.K., Patnaik, P. and Suar, D. (2016) A Multimodal Discourse Analysis of Glocalization and Cultural Identity in Three Indian TV Commercials. *Discourse & Communication*, 10, 209-234. <https://doi.org/10.1177/1750481315623892>
- [8] Moerdisuroso, I. (2017) Social Semiotics and Visual Grammar: A Contemporary Approach to Visual Text Research. *International Journal of Creative and Arts Studies*, 1, 80-91. <https://doi.org/10.24821/ijcas.v1i1.1574>