

地方新闻节目字幕翻译新探索: 人机协作下的多语言处理挑战与智能化解决方案

吴梦婷

四川外国语大学成都学院, 四川 成都 611844

摘要: 本文聚焦于人机协作模式下新闻节目的中英字幕翻译, 特别是针对广东地区电视新闻的翻译挑战。研究揭示了多语言输入(普通话与粤语)及文化语境复杂性对翻译的影响, 评估了现有语音转文字技术和 AI 翻译工具的局限性。提出多语言检测、智能上下文理解和文化敏感算法等技术手段以增强翻译准确性和适应性。同时, 强调了人工干预在确保翻译质量、文化贴切性和新闻传播效果中的关键作用。最终, 展示了人机协作模式在新闻字幕翻译领域的潜力和优化效果。

关键词: 字幕翻译; 人机协作; 多语言处理

Subtitle Translation for Local News Programs: Multilingual Challenges and AI Solutions in Human-AI Collaboration

Wu, Mengting

Chengdu Institute Sichuan International Studies University, Chengdu, Sichuan, 611844, China

Abstract: This paper examines Chinese-to-English subtitle translation for news programs in a human-AI collaboration framework, focusing on the challenges faced in translating TV news from Guangdong. The study highlights the impact of multilingual inputs (Mandarin and Cantonese) and the complexities of cultural context on translation accuracy. It also evaluates the limitations of current speech-to-text technologies and AI translation tools. Technological advancements, such as multilingual detection, intelligent context recognition, and culturally sensitive algorithms, are proposed to improve translation precision and adaptability. Additionally, the paper emphasizes the essential role of human intervention in ensuring translation quality, cultural relevance, and the effectiveness of news dissemination. The study ultimately demonstrates the potential of optimized human-AI collaboration models for news subtitle translation.

Keywords: Subtitle translation; Human-AI collaboration; Multilingualism

DOI: 10.62639/sspsstr54.20240105

引言

随着全球数字化交流的日益加深, 字幕翻译的准确性与时效性在新闻传播中显得尤为关键。在广东这一多语言环境, 普通话与粤语并存, 为字幕翻译带来了独特的挑战。人机协作模式, 作为提升翻译效率与质量的有效途径, 虽已展现其潜力, 但面对混合语言输入及文化细微差异的复杂性时, 仍显不足。

当前的人工智能工具在处理方言、成语及文化特定元素时往往力不从心, 导致翻译结果不尽人意。这一现状凸显了人工干预在修正翻译错误、确保文化贴切性方面的重要性。

本研究特选定广东卫视 DV 现场节目于 2024 年 6 月 8 日播出的高考专题报道作为研究对象, 深入分析多语言环境下的翻译难题, 并审视现有 AI 技术的局限性。进而, 探索多语言检测、文化敏感性算法等技术创新, 以期在人机协作框架下, 找到优化翻译质量的更佳路径。同时, 本文也强调了人类专业知识在翻译过程中的核心价值, 为

新闻字幕翻译的未来发展提供了参考。

一、案例研究: 粤语与普通话新闻字幕的中英翻译

本研究的研究对象是广东新闻节目中有关报道了 2024 年高考首日的情况, 时常约 15 分钟, 节目聚焦于各考场外的紧张与期待氛围。节目采访了考生、家长、教师和工作人员, 体现了他们在这重要考试日的不同情绪和准备工作。家长们穿着旗袍、手举向日葵为考生送上祝福, 表现出对孩子考试成功的美好期望。节目还介绍了志愿者为考生提供帮助、警方保障交通安全等一系列措施, 确保高考顺利进行。

节目语言结合了正式报道与非正式访谈, 采访中充满口语化表达, 真实反映了参与者的情感和紧张心情。节目中普通话与粤语交替使用, 特别是在家长和考生的访谈中, 充分体现了地方语言的独特性和文化背景。像“旗开得胜”和“一举夺魁”这样的文化符号和成语被频繁使用, 反映了高考的传统和文化内涵。

(稿件编号: SSTR-24-5-F001)

作者简介: 吴梦婷(2003-), 女, 汉族, 安徽人, 本科学历, 四川外国语大学成都学院 2021 级在读, 主要研究方向机器翻译及数字媒体。

二、文献综述：语音转文字与机器翻译技术的发展

本节综述了语音转文本技术的发展历程、技术原理、多语言处理能力，以及机器翻译在字幕翻译中的应用，同时探讨了面临的挑战和解决方案。

（一）语音转文字技术发展

随着机器学习和自然语言处理（Natural Language Processing, NLP）技术的飞速发展，语音到文本技术，也就是自动语音识别（Automatic Speech Recognition, ASR），近几十年来取得了显著进步。从最初依赖繁琐人工训练的基于规则算法，到如今运用深度学习模型——如递归神经网络（Recurrent Neural Network, RNN）和卷积神经网络（Convolutional Neural Network, CNN）——的现代系统，实现了对大规模数据集的自动化高效处理，极大地减少了人工干预的需求。

随着全球化进程的加快和语言资源的日益丰富，语音识别技术逐渐支持更多的语种和方言，包括普通话和广东话。国内的一些代表性研究成果，如讯飞听见 APP，以其卓越的实时语音转文字能力而闻名，支持多种语言和方言的转写。

（二）多语言输入的挑战

随着实时转录需求在媒体、教育和无障碍等领域的增长，语音转文本技术日益复杂。但一大难题是准确转录多语言及方言混用的场景，这在广东等地尤为常见，如普通话与粤语的切换。

最新多语言语音转文本系统，如谷歌的多语言 ASR 模型，能自动识别并切换语言，提升了多语言环境下的转录准确性。然而，面对快速的语言或方言切换，这些系统仍易出错，对新闻字幕等场景构成挑战，因为准确性关乎内容完整。

国内，讯飞听见 APP 以强大的实时语音转文字能力脱颖而出，支持多语言和方言，展现了中国技术的卓越。它不仅语音识别精准，还融入了边录边翻、智能编辑、文本翻译等功能，极大地加

速了记录流程，提升了工作效率。

（三）机器翻译在字幕翻译中的应用演变

从基于规则的机器翻译方法向神经机器翻译（Neural Machine Translation, NMT）的转变，显著增强了翻译的流畅度及上下文理解能力。然而，面对多语言输入的复杂场景与文化细微差别的精准捕捉，当前技术仍面临重大挑战。国际范围内，如 Vaswani 等人（2017）在转换器模型上的开创性工作，以及国内上海外国语大学王鹏程（2022）针对成语翻译的专项研究，均揭示了 AI 在处理富含文化深意与特殊语言现象时的局限性。

在字幕翻译领域，特别是在实时性要求高的媒体环境中，AI 工具的应用日益广泛。全球范围内（如 Tuominen 等人，2023）及中国本土（如李楚菡，2022）的研究均聚焦于 AI 在新闻广播、影视作品等字幕翻译中的效能，指出 AI 虽展现出巨大潜力，但在确保语言转换的精准度、适应特定语境下的习惯用语，以及满足实时性需求方面，仍需人工介入以提升翻译质量。

针对粤语与普通话等多语言输入环境下的 AI 语音识别与字幕生成、机器翻译研究，目前尚处于前沿探索阶段，国内外相关研究成果较为稀缺。

三、研究前期准备

在准备阶段，笔者首先对本研究中的原视频进行下载，原视频总计进长为 2 小时，与高考内容相关的仅有 15 分钟，笔者此因进行了相关的剪辑。同时笔者初步拟定人机合作方案，最开始的设计是利用国内外较为先进的实时字幕翻译工具对原视频直接进行实时字幕翻译。笔者选择的国内 AI 工具是科大的“讯飞听见字幕”，国外平台则选择的是 RASK AI，笔者期待这两款工具能提供原声字幕识别、自动打轴和自动翻译等一站式服务。但是，另人遗憾的是，两个平台的表现差强人意，“讯飞听见字幕”在目前没有粤语和普通化混合输入模块，且仅支持粤语到普通话翻译，不支持粤语到英语的翻译（如图 1 所示）。



图 1

而 RASK AI 的模块支持自动语言识别，也能实现粤语到英语的翻译，但是无法识别原视频中的普通话文本，这表明，目前这款 AI 也无法实现多语言环境的自动切换（如图 2 所示）。

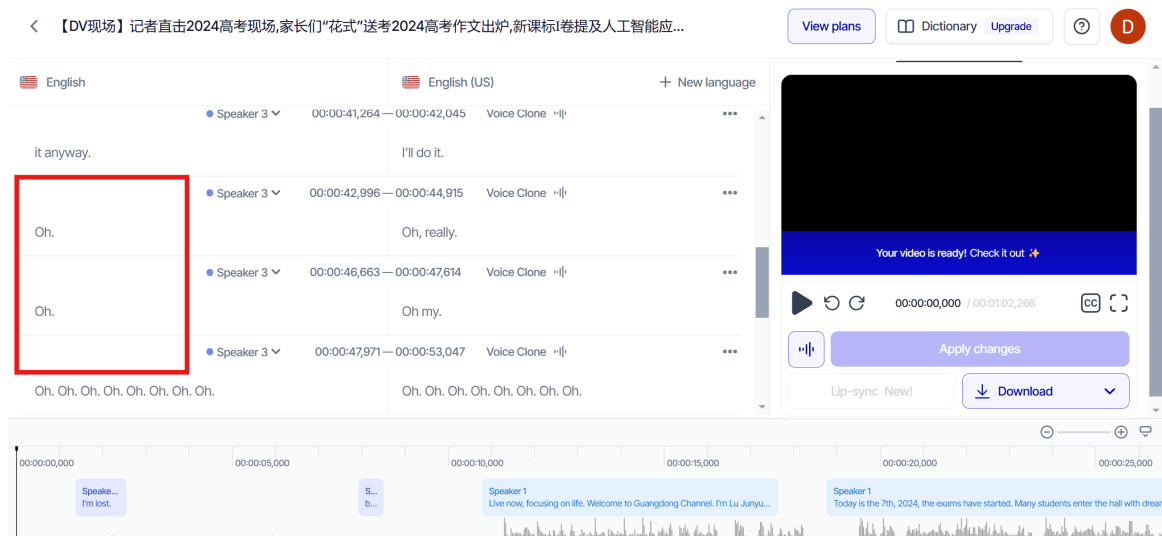


图 2

鉴于目前各 AI 功能局限性，笔者决定拆分人机合作步骤，首先是完成自动语音识别生成中文字幕文件，然后再进行机器翻译。

四、语音转文字的技术难点及解决方案

笔者在语音转文本工具（如“讯飞听见”）中，分别选择普通话和粤语两种输入，分两次进行转录。在这一过程中遭遇了极大挑战，错误频繁出现，迫使人不断介入，在继续翻译前对转录内容进行逐一校正，从而拖慢了整体进度。

（一）语音转文字技术难点

在混合语言环境下，转录工具常难以区分粤语与普通话，如将粤语中的“咗呢”（意为“做了”）错误转录为普通话中的“左呢”（意为“左”），造成上下文错乱。这种误读直接影响了翻译的连贯性，要求人工手动调整，增加了工作量并延长了翻译周期。

此外，工具在处理日常对话中的口语和口头禅时也显得力不从心。例如，“讲真”（粤语，意为“老实说”）常被误译为“将真”（普通话，但此词无实际意义），导致意思完全扭曲。若无人工审核，这些错误将直接反映在翻译结果中，误导受众，因此必须在翻译前进行细致的人工校正。

转录中的另一大挑战在于转录工具难以捕捉并准确转译广东文化特有的表达方式。比如，前文提到的“左呢”这类短语，这暴露了人工智能在理解地域性语言细节上的不足。这进一步凸显了提升语音转文本技术，特别是在方言识别与理解方面的迫切需求。

（二）语音转文字技术解决方案

为解决前文中探讨的转录难题，尤其是混合语言输入与口语的挑战，本研究可采纳多种技术策略来提升语音转文本工具的精确度。

首先，要优化语言检测，引入自适应多语言模型，该模型能在转录过程中自动检测并切换语言，有效减少错误，减少人工干预。其次，应用上下文感知模型，以准确理解说话者意图，特别是在处理非正式口语时，避免误转。另外，在包括南方粤语

区等方言区域，使用方言特定模型提升转录准确性，确保地区特色表达的正确转录。此外，引入增强型声学模型，专注于声调变化，区分发音相近的词汇，减少混淆。最后，建立用户反馈机制，让工具从用户纠正中学习，持续优化性能。

综上所述，实施上述解决方案将使语音转文本工具在混合语言环境下达到更高水准的精确度，确保无论是正式还是非正式的语音输入都能得到准确转录。

五、AI 翻译的挑战与解决方案

在完成原文的详尽转录与精细校对流程之后，本研究进而启动了人机协同的翻译实践。在此过程中，以大语言模型 ChatGPT-4o（或类似高级版本）为代表的先进系统展现出了卓越的翻译性能。它支持粤语到英语，普通话到英语的翻译，且其输出的译文大多自然流畅，语境贴合度很好，如现场有家长表示“觉得应该重视吓，有啲仪式感，为个女女打气。”被译为“I think we should make this special—it’s important to cheer on my daughter.”，这个译本对“仪式感”和“打气”这些表达的处理自然贴切，完全没有直译的生硬感。另一个例子是主持人感叹高考不易时说：“关关难过关关过，前路漫漫易灿灿！”被译为“Every challenge will be overcome, and the path ahead will shine bright.”，此译文不仅精准捕捉了原文的深层含义，还巧妙地保持了韵脚的和谐，体现了高度的文化适应性。

然而，当前人工智能翻译领域仍面临若干关键挑战与待解问题，其核心聚焦于成语表达的精确传达、富含文化意蕴内容的恰当诠释以及复杂句式结构的精准解析。尽管如 ChatGPT-4o 这样的高级模型已展现出高度智能，但在处理细微语言差异、确保文化敏感性和准确性方面仍显不足，往往需要依赖人类专家的深度介入与知识辅助，以强化翻译结果的文化相关性和准确性。

（一）AI 翻译面临的挑战

以原视频中一位记者来到一个考点，形容考

点的热烈气氛时说到“有啲家长攞住‘旗开得胜’嘅小旗仔,亦有啲家长身穿旗袍,攞住向日葵,寓意旗开得胜、一举夺魁。”ChatGPT-4o将其中的成语“旗开得胜”与“一举夺魁”准确地译为“start off victorious”与“win in one attempt”,然而,由于缺乏对成语所承载的深厚文化底蕴及在高考激励语境中独特情感价值的深刻理解,其译文难以充分传达原成语的精髓与激励效果。

同理,有记者描述“而我身后亦都企咗两位同学,身着状元服,喺进入考场嘅必经之路为考生送上祝福。”还有家长介绍自设计的衣服时说“‘寓意系‘旗开得胜’。‘旗开得胜’,仲有锦鲤——高考锦鲤。”AI将“状元服”与“高考锦鲤”的直译(“top scorer outfit”与“Gaokao Koi fish”)忽视了它们作为中国特色文化符号所蕴含的丰富寓意与象征意义,可能导致非中国文化背景的受众在理解上产生偏差或困惑。

(二) AI 翻译技术解决方案

首先,引入语境学习模型可以为人工智能在成语表达及富含文化意义短语的阐释方面提供显著优势。通过深度分析并学习成语所处的具体语境,人工智能能够生成更多蕴含文化背景的精准确译。以高考语境中的成语“旗开得胜”为例,其翻译可被优化为“May you start smoothly and succeed”,此译法超越了字面直译“start off victorious”,更贴切地传达了原成语在高考场景中的祝愿与激励之意。

其次,开发文化敏感性算法是提升人工智能系统处理文化特定短语能力的关键步骤。此类算法能够识别并适应短语背后的文化价值,如在翻译“状元服”时,不仅识别其字面意义,更深入挖掘其作为中国传统文化中象征学术成就的服饰这一深层含义,从而给出如“traditional top scorer outfit, a symbol of academic excellence”的译文,增强了翻译的文化敏感性和深度。同理,对“高考锦鲤”的处理AI也能更加精细化,可以译为“there's also the koi fish, which represents good luck, especially for the Gaokao.”让英语受众更能理解其文化内涵。

此外,针对成语等比喻性语言的专门识别模块训练,对于提升人工智能翻译成语的能力至关重要。此类模块能够精确捕捉成语的内在含义与情感色彩,如将“一举夺魁”译为“Seize the premier position in a single endeavor!”,既保留了原成语的激励效果,又符合目标语言的表达习惯。

跨语言迁移学习技术的应用,则进一步拓宽了人工智能处理多语言及混合语言输入的能力边界。通过融合普通话与广东话等方言的训练数据,人工智能模型能够更有效地适应语言变异与混合语境,确保翻译结果的准确性与适应性。

最后,构建基于特定文化范例的增强型多语言模型,是整合多种语言翻译能力、精准传达文化内涵的有效途径。以高等教育相关短语为训练素材的模型,能够深刻理解并准确翻译如“旗开得胜”等富含文化内涵的短语,既保留其字面意义,又深刻揭示其背后的文化意蕴与象征价值。

(三) 人机结合解决方案

在技术持续进步的驱动下,翻译的准确性虽已显著提升,然而,确保翻译内容的文化贴切性与情感深度仍离不开人类译者的深度参与。特别是对于富含文化意蕴的成语表达,如“天公作美”,人工智能的直译往往难以捕捉其精髓,需依据目标语言的表达习惯与自然度进行细致调整,此过程高度依赖于人类译者的专业判断与语言艺术。

此外,翻译流程中实时引入的人工反馈机制,不仅为人工智能算法的持续优化提供了宝贵的数据支持,还通过直接指出翻译中的问题与不足,为算法迭代指明了方向。这种即时反馈循环,促进了人工智能与人类智慧的紧密融合,共同推动了字幕翻译过程向高效与文化适应性并重的方向发展。

综上所述,尽管人工智能工具在翻译任务中已能承担大量工作,但在追求翻译的准确性、细致度及文化契合度方面,人类译者的角色依然无可替代。因此,构建一种将人工智能高效处理能力与人类译者深厚文化底蕴相结合的混合翻译模式,成为提升字幕翻译质量,尤其是在高考等多元文化、多语言环境下实现快速且精准翻译的有效途径。

六、结论

本文研究了中英字幕翻译的挑战和解决方案,尤其是在经常使用普通话和粤语的多语言环境中。虽然人工智能驱动的语音转文本和翻译工具在快速高效地处理大量数据方面取得了长足进步,但它们在处理多语言切换、复杂语言结构和文化细微差别内容方面仍有困难。这些局限性凸显了人工干预的必要性,以确保翻译既符合上下文的准确性,又具有文化敏感性。

本研究认为人机协作结合了人工智能技术和人类专业知识的优势,目前,仍然是字幕翻译最有效的方法。人工智能擅长初始转录和基本翻译等任务,而人类译员则通过处理人工智能经常忽略的微妙之处,在完善输出方面发挥着至关重要的作用。这种合作不仅能提高翻译速度,还能确保高质量和文化共鸣的成果。

随着人工智能技术的不断发展,其适应复杂语言输入和文化背景的能力也将不断提高。然而,在监督翻译过程和确保最终输出保持准确性和文化相关性方面,人类的专业知识仍将至关重要。字幕翻译的未来将取决于如何优化人类专业知识与人工智能创新之间的平衡,从而在多语言环境中实现效率与质量的双赢。

参考文献:

- [1] Vaswani, Ashish, et al. “Attention Is All You Need.” *Advances in Neural Information Processing Systems*, vol. 30, MIT Press, 2017, pp. 5998-6008.
- [2] 王鹏程. “‘机器翻译+译后编辑’在汉英翻译项目中的优势与局限.” 上海外国语大学, 2022.
- [3] Tuominen, Tiina, et al. “Exploring the gaps in linguistic accessibility of media: The potential of automated subtitling as a solution.” *The Journal of Specialised Translation* 39 (2023): 77-98.
- [4] 李楚茵. “数字化时代的志愿翻译行为研究.” 北京外国语大学, 2022.