

基于自适应学习需求的聋人视频无障碍字幕设计*



曹宇星^{1,2} 吴鹏泽¹ 林晓凡¹ 贾天龙²

(1. 华南师范大学 教育信息技术学院, 广东广州 510631;

2. 绥化学院 文学与传媒学院, 黑龙江绥化 152000)

摘要: 视频生动形象的直观画面深受聋人喜爱, 但由于聋人的听力受损, 使得聋人对视频的信息处理面临诸多困难, 无法有效接收、理解视频内容。为此, 文章以聋人的自适应学习需求为导向, 确定了包括数据处理、用户定位、字幕匹配和无障碍字幕呈现四个环节的聋人视频无障碍字幕设计流程。同时, 文章介绍了两个典型的聋人自适应学习场景: 面向聋人的交互式视频资源平台和人工智能技术支持的聋童识字阅读, 并分别提出了合适的无障碍字幕设计方案。进行基于自适应学习需求的聋人视频无障碍字幕设计, 不仅有利于提升聋人的学习兴趣并有效、完整地接收视频信息, 而且在一定程度上促进了聋人的教育公平。

关键词: 自适应学习; 聋人视频; 无障碍字幕

【中图分类号】G40-057 【文献标识码】A 【论文编号】1009—8097(2019)07—0099—07 【DOI】10.3969/j.issn.1009-8097.2019.07.015

《国务院关于加快推进残疾人小康进程的意见》高度重视残疾人的民生改善, 大力弘扬人道主义思想、扶残助残的中华民族传统美德和残疾人“平等、参与、共享、融合”的现代文明理念, 营造理解、尊重、关心、帮助残疾人的社会环境^[1]。在此背景下, 中国残疾人联合会、教育部、国家语言文字工作委员会、国家新闻出版广电总局联合采取了在视频作品中加配手语、字幕、文字提示等一系列措施, 以实现互联网和移动互联网的无障碍信息服务。

字幕是辅助聋人深入理解视频信息的重要方式。国外学者对于听障类视频字幕的研究较早, 从字幕设计到字幕组织都已有明确的规范, 如 Jelinek Lewis 等^[2]认为视频一定要具备隐藏式字幕 (Closed Captioning, CC) 的功能设计, 即除了对白字幕, 还应包括场景声音和配乐的解释信息, 以利于聋人学习并深化其对视频的理解。我国学者沈玉林^[3]认为在服务聋人语言理解方面, 视频中的字幕设计弥补了手语无法表达的概念和因手语表达比口语速度慢而造成的信息量“丢失”等不足; 而及时更新字幕标题的提示内容, 还有利于提高聋人收看节目的理解能力和书面表达能力。基于此, 本研究拟聚焦于聋人视频无障碍字幕设计, 并针对聋人的自适应学习需求进行具体的流程设计和适应场景分析, 以期实现聋人有效、完整地接收视频信息的目的。

一 聋人的自适应学习需求与聋人视频无障碍字幕设计

对聋人而言, 字幕缺失或字幕使用不当都会造成其认知发展中关键信息的缺失。这就需要研究者针对聋人在信息处理加工方式上存在的缺陷采取补偿性措施, 形成能满足聋人自适应学习需求的视频无障碍字幕设计方案。

1 聋人的自适应学习需求

失去听力, 意味着聋人失去了一个与社会进行有效沟通的渠道; 伴随着听力障碍, 聋人在基础语言能力、记忆能力、认知偏好、注意力分配能力等方面也会与听力正常者存在明显的差异^[4]。究其原因, 在于聋人的信息加工方式与听力正常者有所不同: 听力正常者可以很快明确自己想要传达的信息内容, 确定自己传达信息的方式对信息进行编码, 接收完整信息进行解码并

有所反馈；而聋人由于自身听力障碍导致信息传递与接收存在障碍，且在其信息加工的任意一个环节都有可能造成对信息的歪曲或误解。具体来说，信息处理各环节的角色活动和聋人面临的困难如表 1 所示。

表 1 信息处理各环节的角色活动和聋人面临的困难

信息处理环节	角色活动	聋人面临的困难
信息发送	发送人确定想传达的内容	聋人无法确定如何传达思想和情感
信息编码	发送人选择发送信息的方法（口语、书面、非语言或视觉方式）	聋人无法选择正确的方法，无法通过语言表达自己的感情；信息编码有障碍
转移信息	信息被有效发出	聋人讲话困难，口吃，存在发音障碍
接收信息	接收人有效接收信息	聋人因看不到手势或无法完全听到声音，导致内容接收有障碍
信息解码	信息被有效理解	聋人因词汇量、语境感知、语气信息等缺失，无法有效处理信息而出现误解
反馈	接收人做出有效反应	聋人不能做出正确反应，反馈不及时
回应	自行发送信息，使整个循环继续	聋人对信息无反应或需要充分时间回应

和普通学习者一样，聋人也有学习知识、有效获取信息的权利。但由于聋人在信息处理环节面临各种困难，故聋人学习者的学习之路注定充满艰辛。随着自适应技术及其相应推荐算法的不断进步，学习者个性化的学习渠道得以拓宽，这在认知评估、知识追踪、学习诊断等方面为聋人带来了便利。聋人的自适应学习需求是指依托网络，通过与系统实时对话，系统自动采集聋人的各种学习行为数据，对聋人学习者的听力损失程度、语言基础水平、学习风格、理解程度等进行智能分析，由此得到的符合聋人认知特征、学习进度的个性化学习需要。

2 聋人视频无障碍字幕设计

在视频资源中，字幕作为聋人的主要信息来源，其设计对于聋人开展自适应学习十分重要，其原因在于：如果缺少字幕或字幕设计不当，便无法实现有效的信息传播；手语可表达的词汇量较少，语序与文字表述很难一一对应，故需要引入恰当的字幕进行解释；聋人的文字语义理解能力较弱，很难从长句中快速提取关键词^[5]，故字幕需要进行合理设计。总的来说，聋人视频无障碍字幕设计是指通过深度挖掘聋人在学习过程中对知识和技能的掌握程度，依据对聋人的认知诊断结果和语言发展水平而进行的字幕匹配设计，旨在实现聋人对视频资源的无障碍学习和理解，在一定程度上满足聋人的自适应学习需求。

二 基于自适应学习需求的聋人视频无障碍字幕设计流程

针对聋人的自适应学习需求，为了实现聋人视频无障碍字幕设计的规范化和普及化，本研究根据字幕生成环节及其自适应动态修正的内容需要，制定了基于自适应学习需求的聋人视频无障碍字幕设计流程，如图 1 所示。

1 数据处理

数据处理是指聋人在观看视频的过程中对收集的数据进行认知风格辨别、听障程度检测（包括声强、频率和助听器佩戴情况等）、学习能力测试和学习行为分析。有效的数据处理可以实现对聋人学习者的精准定位，有利于为其提供适切的自适应学习服务。数据处理的目的是，通过

测验或记录聋人的学习行为（如字幕浏览速度、字幕识别情况、字幕类型风格选择），来判断聋人的感知觉特征（包括注意力、记忆力等）及其对相关知识点的掌握情况。当然，聋人对视频内容的理解不仅与字幕变换的速度和风格有关，而且依赖于聋人的学习能力及其已有的学习基础^[6]。因此，还需采集聋人学习者在视频观看过程中标注的难点、兴趣点，并进一步由自适应学习系统将数据进行聚类，形成聋人学习者的知识网络。

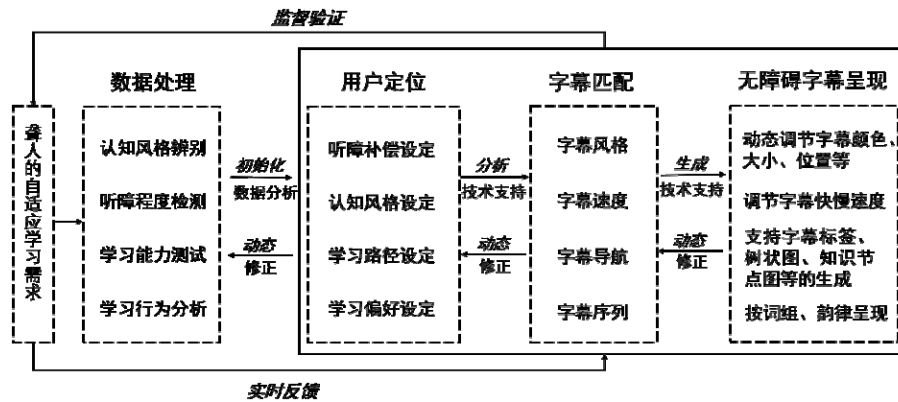


图 1 基于自适应学习需求的聋人视频无障碍字幕的设计流程

2 用户定位

用户定位主要根据数据处理环节的数据分析结果，对聋人学习者从听障补偿设定、认知风格设定、学习路径设定和学习偏好设定等四个方面进行定位。以听障补偿设定为例，本研究采用世界卫生组织（WHO-1997）日内瓦会议推荐听力障碍分级标准和残疾人残疾分类分级听力残疾评定标准，分别对聋人学习者的听力障碍、听力残疾进行等级划分，如表 2 所示。对聋人学习者听力障碍和听力残疾的分级，有利于为每位聋人学习者提供适合其自适应学习需求的无障碍字幕服务。如依据 WHO-1997 标准，26~40 分贝听力损失为轻度障碍，可采用隐藏式字幕设计，以符合聋人的听力康复需求；61~80 分贝听力损失为重度障碍，聋人进行理解、交流等活动重度受限，可选用基于 SETT（Student-Environment-Task-Tool）框架的无障碍字幕设计^[8]，以解决聋人断句困难、专有名词难以理解的问题；81 以上分贝听力损失为极重度障碍，这种情况下除了考虑使用字幕以外的辅助（如手语翻译），还需考虑建立聋人多模态语料库。

表 2 听力障碍和听力残疾的分级列表^[7]

世界卫生组织（WHO-1997 年） 日内瓦会议推荐听力障碍分级标准			残疾人残疾分类分级 听力残疾评定标准（GB/T 2634-2010）		
类别	级别	听力损失程度 (dB HL)	类别	级别	听力损失程度 (dB HL)
听力障碍	极重度	81	听力残疾	一级	> 90
	重度	61 ~ 80		二级	81 ~ 90
	中度	41 ~ 60		三级	61 ~ 80
	轻度	26 ~ 40		四级	41 ~ 60

3 字幕匹配

字幕具有工具性和文化性的双重特性，可以应用自适应引擎技术，从字幕风格、字幕速度、字幕导航、字幕序列等四个方面进行适应性匹配。

(1) 字幕风格

字幕风格可以根据用户的注册信息（如年龄、性别、学习背景、兴趣爱好等）来定制。字幕的含义既可以通过颜色、字体、字号等的设计，来表示发声者的不同（如图2所示），也可以通过清晰的字幕设计来解释手语（如图3所示）。为了更好地满足聋人的自适应学习需求，面向聋人的视频无障碍字幕风格预设应重点考虑：字幕自适应呈现样式，如提供低龄聋童的卡通版字幕样式风格或成年聋人简洁明了的风格等选择；字幕自适应位置定位，如字幕通常设置在屏幕底部，而自适应字幕可以将字幕放于发声处附近，通过不同的位置和“「」”等符号来辅助理解复杂语境；字幕自适应颜色搭配，如字幕背景色可自主识别视频背景的颜色并呈现强对比效果。



图2 不同的颜色、字体和字号，表示不同的发声者



图3 解释手语的字幕与背景对比清晰

(2) 字幕速度

字幕速度可以根据不同用户的听力监测、认知能力、学习行为等综合数据分析结果进行不同的设定，也可以根据用户的听障等级和前期测试结果选择无字幕、有字幕或手语版展示。此外，由于聋人记忆力较弱，可通过设置“字幕标签”重复回放已选的视频片段，以帮助聋人增强记忆效果，并自主掌控学习进度。

(3) 字幕导航

字幕导航是指根据聋人的学习风格测试和行为分析，来确定其学习路径和偏好设定。以知识树结构进行字幕分类导航，可以通过呈现含有关键字幕的视频片段，来明确自己所学的内容在整个知识体系中的位置，从而避免信息迷茫。

(4) 字幕序列

字幕序列即一组连续的、有相关联系的，并按照一定顺序排列的字幕，通过设计可以揭示字幕之间的关系或某种隐性存在的规律。根据学习者的认知水平和当前学习能力，冉若曦等^[9]采用 SETT 框架，以中文分词的技术方式提供字幕辅助方案。所谓中文分词，是指对字幕文本的韵律切分，即标注字幕文本的专有名词，辨别其词组序列的出场形式。如“大熊猫”是一个整体词语，而不能以“大”、“熊”、“猫”分别展示，其作用是强化聋人对概念的整体认知。

4 无障碍字幕呈现

基于对用户定位和字幕匹配模型的判断，聋人收看的视频端提供了可视化地呈现无障碍字幕的功能：在字幕风格方面，体现为对字幕颜色、字体大小及其位置定位等的动态调整；在字幕速度方面，体现为对字幕显示速度快慢的自适应调节；在字幕导航方面，体现为字幕标签和知识节点图的可视化；在字幕序列方面，体现为支持以聋人语言训练的词组、韵律等方式呈现。

三 基于自适应学习需求的聋人无障碍字幕适用场景

在人工智能、人机交互、智能识别等技术的支持下，聋人获取知识的途径不断拓宽。在此背景下，本研究拟选取两个典型的聋人自适应学习场景，尝试提出相适的无障碍字幕设计方案。

1 场景一：面向聋人的交互式视频资源平台

联合国教育、科学及文化组织（United Nations Educational, Scientific and Cultural Organization, UNESCO）曾强调，要鼓励将信息技术、通信技术和交互技术应用于推进残疾人信息无障碍资源建设^[10]。如希腊 Pandora 数字图书馆（官网地址：<http://multimedia-library.prosvasimo.gr/home>），它是一个免费、开源的面向聋人的交互式视频资源平台（下文简称“平台”）。目前，该平台拥有超过 23,000 个手语视频，其交互特性体现在可以通过数字标签的方式进行搜索与呈现，聋人可参与视频资源的管理和归类^[11]。据此，本研究拟提出适应于面向聋人的交互式视频资源平台的无障碍字幕设计方案：

聋人交互式视频资源中字幕风格的自适应设计。由于视频资源数量庞大，视频字幕的添加与更改的范围较大。因此，平台要为用户预设标准字幕设计端口，并可根据聋人年龄、听障程度将字幕的主要风格类型进行标准化处理。之后，基于用户对手语视频类别、内容、关键信息设置（如字体、大小、颜色）等的选择，平台相应地推荐无障碍字幕。

聋人交互式视频资源中字幕速度的自适应设计。字幕的速度变换与聋人的听力损失程度、已有语言基础、观看技巧和教师的教学引导有关。字幕呈现速度过快，聋人会因为无法完整地获取信息而导致传播无效；字幕呈现速度过慢，则容易让聋人产生视觉疲劳和心理疲劳。所以，面对聋人这一特殊群体，平台应提供可自适应调节的字幕速度：允许视频根据聋人的自适应学习需要进行速度调节或定制，并可随着视频内容的更新而实时调整手语、字幕等；允许聋人下载视频进行反复回放操作，以提升聋人的记忆效果。

聋人交互式视频资源中字幕导航的自适应设计。基于自适应学习的聋人无障碍字幕设计具有快速导航、可编辑、易操作等特征，以适应不同聋人的学习需要。目前，聋人交互式视频资源库已实现对视频字幕的高级搜索，且与搜索信息一致的视频字幕内容（如单词、短语等）将自动“标黄”，如图 4 所示。建议平台通过聋人的视频标注行为或搜索行为的数据分析进行内容汇聚，并以时间序列或知识地图的形式呈现，以帮助聋人快速了解知识的整体框架和自己所学内容在其中所处的位置，以更好地辅助聋人开展自适应学习。

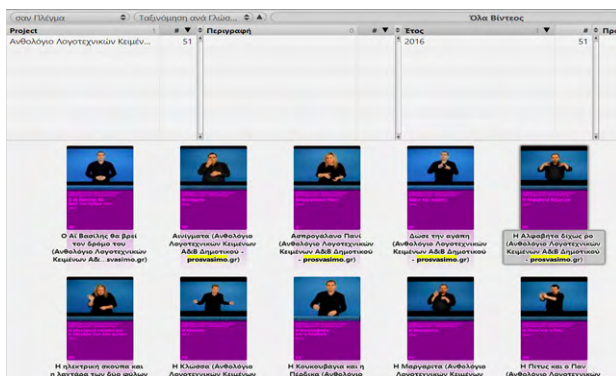


图 4 视频字幕的高级搜索与自动“标黄”

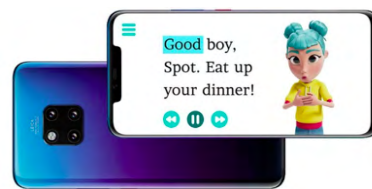


图 5 StorySign 字幕与手语实时对应

聋人交互式视频资源中字幕序列的自适应设计。建议平台依据字幕语义,进一步提升视频信息之间的关联度,为关键字幕匹配相关的拓展性资源。但是,这的确是一项极其艰巨的任务,需要网络行为数据分析师与专业手语翻译、学科领域专家共同合作完成。

2 场景二:人工智能技术支持的聋童识字阅读

世界各地都有很多聋童在努力尝试阅读,但由于听力障碍,很难将平时看到的唇语、手语与相应的文字对应起来,而长期对应匹配缺失,导致聋童的识字率普遍偏低。我国华为公司采用人工智能技术并结合动作捕捉技术,进行 StorySign 字幕与手语实时对应功能的研发,目前可实现将书籍上的部分文字实时转化为手语视频,并依据聋童的阅读速度将字幕自动“标亮”,让听障儿童可以像听力正常者一样享受美妙的故事^[12],如图 5 所示。此外,华为研发组还与欧洲聋人联盟、英国聋人协会等慈善机构合作,开发了能够提供与字幕相匹配的真实面部表情和准确手语的卡通人物形象,且视频内容选择范围贴近聋童的日常生活,从而降低了聋童的认知负荷,丰富了聋童的阅读资料。据此,本研究拟提出适应于人工智能技术支持的聋童识字阅读的无障碍字幕设计方案:

依据聋童的学习风格进行自适应多元化预设。在确保教育视频字幕科学、准确的前提下,进一步配合聋童的听障程度和记忆特点,匹配多样的字幕风格样式和文字出场节奏,充分给予聋童理解、消化知识点的时间,进而有效、完整地传达信息。

跨学科协同设计聋童视频无障碍字幕。从特殊儿童心理学、教育学等多学科角度出发,对聋童进行个性化字幕信息补偿,增加能吸引聋童注意力的卡通人物、贴近聋童生活的字幕内容等,在降低聋童认知负荷的同时,努力提升聋童的学习兴趣。

四 结语

字幕是语言的传播载体,而使用字幕是尊重聋人群体的表现。视频字幕的添加是聋人获取信息的基本需求,但由于不同聋人的听障程度、知识水平、学习需求等均存在差异,故应依托差异化学习理念,针对聋人的自适应学习需求,从字幕风格、字幕速度、字幕导航、字幕序列等方面进行聋人视频无障碍字幕设计。期待借助本研究,能吸引更多的有识之士参与进来,共同探讨聋人视频无障碍字幕的设计路径,探寻基于学习理论的聋人认知加工补偿方案和人工智能、大数据等技术支持下聋人自适应教育的实现方式^[13],努力让聋人获得更多平等学习的机会。

参考文献

- [1]国务院.国务院关于加快推进残疾人小康进程的意见[OL].
<http://www.gov.cn/zhengce/content/2015-02/05/content_9461.htm>
- [2]Jelinek Lewis M S, Jackson D W. Television literacy: Comprehension of program content using closed captions for the deaf[J]. The Journal of Deaf Studies and Deaf Education, 2001,(1):43-53.
- [3]沈玉林.从电视手语新闻谈对受众的服务[J].新闻窗,2013,(6):60-61.
- [4]Brusilovsky P. Adaptive hypermedia[J]. User Modeling and User-adapted Interaction, 2001,(1-2):87-110.
- [5]Hyde M, Zevenbergen R, Power D. Deaf and hard of hearing students' performance on arithmetic word problems[J]. American Annals of the Deaf, 2003,(1):56-64.

- [6]Burnham D, Leigh G, Noble W, et al. Parameters in television captioning for deaf and hard-of-hearing adults: Effects of caption rate versus text reduction on comprehension[J]. *The Journal of Deaf Studies and Deaf Education*, 2008,(3):392-404.
- [7]孙喜斌,刘志敏.残疾人残疾分类和分级《听力残疾标准》解读[J].*听力学及言语疾病杂志*,2015,(2):105-108.
- [8][9]冉若曦,李晗静,吕会华.基于 SETT 框架的无障碍字幕教学视频研究与实现[J].*北京联合大学学报*,2016,(3):88-92.
- [10]UNESCO. ICTs in education for people with disabilities: Review of innovative practice[OL].
<<https://unesdoc.unesco.org/ark:/48223/pf0000193655>>
- [11]Kourbetis V, Boukouras K, Gelastopoulou M. Multimodal accessibility for deaf students using interactive video, digital repository and hybrid books[A]. *International Conference on Universal Access in Human-Computer Interaction*[C]. Toronto: *Lecture Notes in Computer Science*, 2016:93-102.
- [12]太平洋电脑网.华为新应用借助 AI 帮听障儿童更好地阅读书籍[OL].
<<https://mobile.pconline.com.cn/1207/12071366.html>>
- [13]吴鹏泽,杨琳.学习理论视角的数字化无障碍学习发展与变迁[J].*中国电化教育*,2018,(12):136-141.

The Design of Video Barrier-free Subtitles for Deaf People based on Adaptive Learning Needs

CAO Yu-xing^{1,2} WU Peng-ze¹ LIN Xiao-fan¹ JIA Tian-long²

(1. School of Educational Information Technology, South China Normal University, Guangzhou, Guangdong, China 510631;

2. School of Literature and Media, Suihua College, Suihua, Heilongjiang, China 152000)

Abstract: The vivid intuitive pictures of video are deeply loved by deaf people. However, deaf people face many difficulties in the information processing of video and cannot receive and understand video content effectively, due to hearing impairment. Therefore, oriented to the adaptive learning needs of deaf people, this paper determined the design flow of video barrier-free subtitles for deaf people, which included data processing, user positioning, subtitle matching and barrier-free subtitles presentation. Meanwhile, two typical adaptive learning scenarios for deaf people were introduced: interactive video resource platform for deaf people and literacy reading for deaf children supported by artificial intelligence technology, and appropriate design programs of barrier-free subtitles were put forward, respectively. The design of video barrier-free subtitles for deaf people based on adaptive learning needs was conducive to improving the learning interests of deaf people, effectively and completely receiving video information, and promoting education equity to some extent for deaf people.

Keywords: adaptive learning; video for deaf people; barrier-free subtitle

*基金项目:本文为国家社会科学基金青年课题“面向听障儿童的优质教育视频资源的建设机制与效果研究”(项目编号:15CXW038)的阶段性研究成果。

作者简介:曹宇星,华南师范大学在读博士,绥化学院讲师,研究方向为特殊教育信息化、教育电视,邮箱为 cyxcom@qq.com。

收稿日期:2019年3月14日

编辑:小米