

进化与异化: AI 合成主播的言说困境

高贵武 赵行知

摘要:言说,在现代汉语中泛指一般谈论、说话或指说出的话,其本身既有口语交流和信息传递的意义,又有宣教和宣传的意味。在传统媒体当中,主播的价值即主要在于以一种人格化和类人际传播的方式实现媒体向受众更好地言说。AI(人工智能)技术本是用来改善这种传播的技术工具,但同时也带来了人的主体性危机。从“异化”的理论视角,通过对10位不同程度触达和接受AI合成主播产品的受众进行深度访谈,可深入研究受众感知AI合成主播的“言说”过程,探析其可能遇到的三重困境,以批判性地考察AI合成主播的传播效能。

关键词:AI合成主播;主体性;交往理性;异化

中图分类号:G206 **文献标志码:**A **文章编号:**2096-5443(2023)02-0005-12

项目基金:中国人民大学“双一流”建设项目“中国人民大学马克思主义新闻观研究中心课题”(RMXY2021C021)

一、问题的缘起:作为言说工具的AI技术

言说在现代汉语中泛指一般谈论、说话或指说出的话^[1]。尽管有哲学家认为“不可言说之域并不存在^[2]”,有意无意地忽略言说作为沟通行为的属性,但无论宗教研究还是哲学研究,都将言说的命题归纳为“可言说/不可言说”的二分法中^[3],将言说与否作为划分默会知识与广义知识的标准^[4]。言说与言语和语言不同:在索绪尔看来,语言是一个超个体的、自给自足的符号系统,是一系列的社会规约和一套使言语活动得以成为可能的结构原则,而言语指的是个人对语言的具体使用,它是多变的、个性化的^[5];拉康也注意到了言说在建构主体层面的重要意义,在他看来,语言是先于主体而存在的^[6],是支配主体无意识结构的象征法则,而言说则是主体实现自我建构的某种行为,是一种在与他人建立关系的基础上确立自身存在的行为。

《春秋谷梁传·僖公二十二年》中也提到:“人之所以为人者,言也;人而不能言,何以为人?”揭示了言说作为人之内在规定性的本质,甚至将言说视为人之为人的根本凭据之一^[7]。由此看来,言说既是人的本质需求,也是人在社会化过程中的公共交往需求。而“言以载道”“言而有信”等论述又在一定程度上暗示言说的内容普遍具备公共性,或存在着以信誉为基础的价值标准。因此,来自人格化传播主体的言说,不仅有助于满足受众在虚拟交往中的言说(被言说)欲望,还能通过言说传递出权威、有说服力的信息,体现出媒体自身的人性化。而无论是真人主播,还是数字时代的AI合成主播,媒体所追求的就是力图还原本真的交往环境,使受众在媒介使用中感受到交往的愉悦,感知到自身的主体性。

言说既意味着口语传播层面的诉说与对话,也蕴含着公共层面的互通互信、社会与心理层面上的“心领神会”等精神交往含义,而人对言说的技术工具使用往往是从能够诉说和对话开始。人工智能(AI)之所以能够作为言说的技术工具,主要取决于以下三方面的可能性。其一,AI合成主播具备言说的社会条件。AI合成主播是真人主播的技术进化,在满足真人主播的基础功能前提下提升效

率、降低成本。当前主流媒体融合转型大力推广 AI 合成主播,赋予其高效传递信息的社会期待,这为 AI 合成主播承担公共言说的使命提供了社会现实基础;其二,人工智能技术具备言说的离身可能。言说行为与主体之间的互动是通过知觉来感知的^[8],正因为人活在知觉中,人的主体性才能够通过言说而得以存在。离身认知(disembodied cognition)的认知科学思潮,强调认知在功能上是能够脱离人的身体而独立存在的,换言之,身心是可以二元的,不需要肉身也能够感受和知觉,这就为 AI 合成主播拥有言说的可能性奠定了基础;其三, AI 合成主播具备服务于言说的技术偏向。言说本身虽然具备主观性,但也受到语言、思维、社会规则等的制约。认知语言学家认为,人们描述特定现象的生活无法绝对客观,总是从特定的视角观察对象。AI 合成主播背后的人工智能技术采用认知计算与算法学习的方式演进,当人工智能在“与人交互”与“向人学习”中逐渐适应既有交往规则时, AI 合成主播与人便存在规则的通约性,从而搭建了像人一样言说的现实基础。

尽管认知科学与认知语言学存在不同的范式之争,“身心能否二元”的争议也一直影响着人们对人工智能技术的伦理评判,但随着业界不断发展, AI 言说现象的出现引发的对于“言说工具”问题的讨论已是不争的事实。AI 技术不仅能提升主播的工作效率,在播报的准确性、全时性、侧重用户服务、丰富用户体验等方面亦为主持传播带来了更好的效能和体验。在电影《西蒙妮》中,人们曾描绘过 AI 合成主播的理想形态:“她是一个由计算机虚拟合成的、高度逼真的三维动画人物,一言一行都与真人一样,并可以完成一切表演、播报,且不会有任何绯闻。”^[9]当下, AI 合成主播早已触及早期人们对它的完美想象,而随着 AI 合成主播的不断演进,其在技术、理念与用途等方面也在不断走向纵深。

AI 合成主播蓬勃发展的实践背后同样存在着不少盲区与隐忧。关于 AI 合成主播未来如何发展的研究目前主要分布在用途展望和伦理反思两个方向,其中大多涉及具身性^[10]、媒介性^[11]、人格化以及智能化等问题。当 AI 合成主播和虚拟偶像的界限越来越模糊时,不少学者提出需要带着资本和平台的视角去看待背后的利益分成、粉丝经济与饭圈治理等伦理议题,期待创造一种技术主体的伦理自觉和专业性的自律。^[12]而在 AI 合成主播的传播效率与受众感知层面,现有研究则主要集中于新技术与社会行动者的双重视角上,针对感知新颖性与受众接受度、可信度等指标的测量服务于 AI 合成主播的形象、性格建构^[13],鲜少对受众的主观意愿进行更具人文色彩的研究。而对 AI 合成主播而言,不论是作为言说工具的技术属性,还是密切连接关系经济的类人际传播特质,都不能不引发技术进化对言说能力的想象与期待,但这样的技术演进同样带来了情感空洞、语言沟通失效等负面影响。原本用来改善传播效能的人工智能合成技术却带来了人的主体性危机,存在着“异化”的可能性。这些都需要从受众的视角去体察人机之间的关系与互动。

技术进步是否真能带来更加有效的言说?受众对与 AI 合成主播之间的言说到底存在何种期待?为了审视和回答这些问题,本文尝试引入“异化”理论视角,批判性地看待 AI 合成主播的传播效能。具体到操作上,笔者采用深度访谈法对 15 位不同程度触达和接受 AI 合成主播产品的受众进行访谈(访谈信息如表 1),研究者总计对访谈对象进行了 15~30 小时的线上深度访谈(其中每位访谈对象的访谈时长分别约为 1~2 小时),共获得了近十万字的访谈材料。通过分析这些材料,受众感知“言说”的过程以及 AI 合成主播在言说主体、心理感知及交流观念、伦理等方面的困境开始逐渐浮出水面。

表 1 受众访谈信息表与筛选依据

被访者类型	访谈编号	备注	类型划分依据
X	X1	工作经历中曾参与开发某款 AI 新闻主播	相关领域从业者。由于工作之便深度接触 AI 合成主播的生产逻辑与生成过程。
	X2	某头部科技互联网公司数字人开发者	
	X3	某头部科技互联网公司数字人产品经理	

续表

被访者类型	访谈编号	备注	类型划分依据
Y	Y1	对 AI 合成主播有个人性别偏好取向	主要听众(观众)。有听 AI 合成主播播新闻的习惯或者观看 AI 合成主播主持节目的习惯。
	Y2	关注了较多虚拟偶像,某虚拟偶像团体的粉丝	
	Y3	曾有过主播工作经历,对 AI 合成主播有观看偏好	
	Y4	有过为 AI 配音的工作经历,是 3 年以上的 AI 合成主播受众	
	Y5	学习播音主持专业,为某 AI 主播角色做动捕经验,关注 AI 合成主播频度高	
	Y6	对 AI 合成主播有观看偏好,高校学生	
Z	Z1	高校学生,新闻传媒类专业,长期关注 AI 合成主播	咨询关注者。长期关注或者对人工智能主播感兴趣,对 AI 合成主播不同阶段的发展有所关注等。
	Z2	高校学生,财务管理类专业,长期关注 AI 合成主播	
	Z3	自媒体从业者,关注虚拟偶像或者 AI 数字人	
	Z4	法律行业从业者,长期关注 AI 合成主播,并撰写过相关分析文章	
	Z5	残障人士,对 AI 合成主播有关关注偏好,有使用各种数字产品经历	
	Z6	高校学生,播音主持专业,长期关注 AI 合成主播,以往兼职工作岗位曾被 AI 取代	

二、主体性缺失:信息生产中的主体困惑

主播通常不仅是信息生产的主体,更是信息传播的人格化中介。传统意义上的新闻主播,尽管主要完成的是播报任务,但其专业素养取决于能否在纷乱的信息中找寻新闻价值、通过二次处理稿件,在言说过程中整合信息、输出内容。在此基础上,新闻主播还需在播报中融入口语表达的艺术,发挥视听媒介的优势,在表情、肢体、神态等非语言符号中丰富报道内容。因此,发生在主播身上的言说既是一种信息传递行为,更是一种信息再生产行为。

依托于语音合成和音源库的技术基础,AI 合成主播已经实现了语音声纹的个性化和拟人化,在动作捕捉与 3D 计算机动画人物形象等技术的驱动下,AI 合成主播也越来越具备像人一样言说的条件。微软 2014 年建立的情感计算框架“微软小冰”甚至在升级中发展出了主播之外的社会化角色,类似于少女歌手、诗人等^[14]。在技术赋能之下,AI 合成主播与人言说的主体愈加丰富,能够再生产的信息类型也不再仅是狭义的新闻报道。然而,AI 合成主播虽然可以做到 24 小时不间断地播报信息,将信息传递的效能大大提升,但主播角色的艺术性和创造力也在这个过程中被科技感与新奇性淹没。口语传播中的创造力和艺术性并非技术进化的目的,而这正可能是被技术进化所隐藏的受众需求。作为口语传播的主体,AI 合成主播角色难以自主地理解播报内容,只能带来信息维度的增量;而 AI 合成主播作为新闻生产-传播-反馈链条中重要的主体,主播的角色由人转换为人工智能机器,言说的过程陷入技术黑箱。人们对 AI 作为交流主体的主体性困惑始终存续在人机交往之中,不但受众(人)存在“正在与我说话的主体是谁?”的困惑,技术进化还让受众本身难以辨别“事实”“意见”与“编程规则”,消解了情感互动层面真实人的主体性,带来信息生产行为中受众主体的规训。

人工智能技术虽让 AI 主播拥有更类人的皮肤,更逼真的声纹,更敏感的情绪识别,但成为主播

的 AI 如何言说? 被看做主播的 AI 能否完成像真人主播一样的言说使命? 这样的追问伴随着技术进化显得更加重要。在 AI 合成主播不断迭代的过程中, 元宇宙、虚拟偶像、粉丝经济等热点与之不断接合, 呈现出被消费主义意识形态和技术术语霸权裹挟的形态。当 AI 合成主播成为一种物化的沟通符号, 在具备虚拟感的播报和现实感的新闻之间, 在技术物与感性人之间, 在言说的过程中人们寻求参与科技和游戏的仪式性互动之间, 模糊了呈现事实和技术表演之间的界限, 将真正实现沟通与共情的可能转换为技术猎奇和偶像崇拜。

(一) 仪式感对信息量的侵蚀与吞没

相比于真人主播在大众媒介中的言说, AI 合成主播给了受众更多参与的新奇感和触及场景, 崭新的对话形象、高交互的交流方式、丰富多样的交往环境为受众带来了前所未有的言说仪式感(Z4, Z5, Z6), 有了这样新鲜的体验, 访谈对象提出, 受众往往很少会出于纯粹的信息需求选择 AI 合成主播作品(产品)(Y5, Y6)。AI 合成主播技术原本是为了满足信息的全时无间断播报, 给受众提供更及时、多元和全面的信息增量, 但在现实中这种信息量不但没有增加, 反而遭到技术进步造就的仪式感的吞没, 逐渐远离了最初的目标, 这不能不说是口语传播中 AI 合成主播语音技术的一大局限。而在技术局限客观存在的情形下, 区别于人的技术主体又会演化出更多超越主播的角色, 也会引发受众更多区别于对以往真人主播的差异性期待。

言说最核心的特点在于主体运用语言符号与非语言符号融合创生出无比丰富的视听感受。人工智能、大数据和媒介融合为技术基底的 AI 合成主播主要通过人脸合成与语音合成技术“像人一样”地生产和传播信息。在人的生理结构中, 音高、音强、音色等特质不仅来自基因, 更来自原始时期人的口语交际和社交认同, 因此更具社会性。而 AI 语音合成最多只能在算力逻辑中穷尽“排列组合”模仿出不同的声纹, 却始终无法在社会意义上比拟真实的人声。“在断句以及语调上面, 我觉得听起来还是单调, 而且挺不自然的。时间久了也许会疲劳”(Y1), 而“听感的疲劳”(Y4, Z3, X3)一定程度上表明了语音合成技术运用数字化处理试图达到真人主播吐字归声的言说效果实际上收效甚微。

与之相反, 更多选择 AI 合成主播作品(产品)的受众却是被游走在虚拟与现实之间超真实的声感与观感(Z2, Z4, Z5, Z6)所吸引。在本研究中, 针对有真人原型、无真人原型以及纯虚拟形象的三类主播, 许多访谈对象选择了纯虚拟形象的 AI 合成主播。“有时我会专门关注 AI 合成主播的声音和表情, 当他纯粹是一个全新的虚拟人物的时候, 我不再去想是真是假, 而是开始喜欢它本身, 这和虚拟偶像也是一样的。”(Y3)早在互联网兴起初期, 美国、日本等不同国家开始试水 AI 合成主播。早期被一概称之为“虚拟主播”, 其使用场景更多地存在于游戏和动漫领域, 较为知名的有穿越火线的“灵狐”、B 站的虚拟歌手“洛天依”等。可见受众对 AI 合成主播的喜爱很大程度上并非将其看作是“主播”角色, 更多是认知为“偶像”“动漫人物”等角色。由此, 受众逐渐将与其交往和消遣的需求替代了更大程度上获取信息的需求。即便如此, 受众对 AI 合成主播的功能期待仍然是混杂的, “我认为未来 AI 主播能发挥的作用可能还是在信息层面”(X1, X2, Z1), 当被消费主义与饭圈文化所定义的“主播”成为重构言说的方式时, 受众就有可能将技术的模仿看作是一种真实, 将娱乐和情感消费看作是信息增量。

(二) 游戏性对传播交流感的替代

相比于真人主播, AI 合成主播会在受众对待言说的心态上注入更强的游戏属性, 从而忽略最初沟通中本应注重的交流感。在人际传播中, 不论是微笑、点头还是一颦一笑都为不同的言说场景留下注脚, 这意味着交流感对信息的传达和情感的共鸣有着重要的作用。交流感也成为传统播音员主持人在科班学习过程中的目标之一, 它常常可作为交流双方默契感的表征。AI 合成主播为增强交流感在面部合成与外形塑造技术上不断进化, 为了增强熟悉的交流感, 便采用了撒贝宁、康辉等大众熟悉的主持人为原型, 为了高度还原人际交往中的交流感, 便使用了大量的面部动态捕捉技术、3D 全

息技术。

卢卡奇认为,商品逻辑掩盖了“人与人关系这一商品本质属性的每一处痕迹”,致使“人们之间的关系带上了物的特性,并获得了‘虚幻的客观性’”^[15]。这提示人们除了关注技术进化的效果,还要关注提供技术的源头。AI 合成主播引得许多互联网技术公司纷纷下场借助算法、大数据等方式提升“主播”的产品质量。技术进化的内在动因与外在动力则使得 AI 合成主播成为各大技术公司与媒体竞标的“数字人”产品。当 AI 合成主播成为“人格化工具”时,受众并不能在其言说中感知到 AI 合成主播背后的产品属性,也无法意识到一次观看和交互能够生产多大的商业价值。当前受众在 AI 技术和元宇宙生态的影响下,改变了与 AI 合成主播言说和沟通的心态,正如访谈对象 X2 提到:“目前的 AI 合成主播交互的方式好像比较单一,有时候互动的乐趣来自像和 Siri 说话一样与他做游戏,看看能不能有趣地回答我的无厘头问题。”用游戏的心态替代本该存在于言说中的交流感,这样的转换隐蔽地发生在 AI 合成主播的技术演进中,让受众在“戏弄”“戏谑”中获得猎奇的情绪满足,而在游戏之后又实际上陷入了“对空言说”的欺骗性沟通之中(Y2, Y3)。就像科幻电影《她》中的情景,男主角沉浸在一场与机器人的爱情中异常痛苦,而机器人却只将他视作编程背后的无数终端之一。

(三)表演性对信息可信度的削减

相比于真人主播,新闻信息生产中的 AI 合成主播功能应用其实十分有限,甚至有时 AI 合成技术只停留于表演层面,用于彰显主播所在平台技术先进、理念超前。当表演作用强于实际操作功用时,AI 合成主播也极易沦为装饰性的技术物存在。一般而言,一种媒介就是一种工具或技术。媒介物是人在具身化的媒介实践中积累的交往经验和知识的体现,是各种与沟通、交往相关的技术的物质性载体^[16]。因此 AI 合成主播不仅具身体现为新旧技术的演变、迭代,还应以身体为构造尺度,在言说过程中调节、中介各种交往关系。但在虚拟平台上,“AI 合成主播”成为行业风口和发展热点后,难以避免被当作是来吸引更多的平台用户的“技术噱头”,并未起到深层的交往功用。访谈对象 X1 提到:“有一些动漫人物会也自称是用了 AI 技术,会在前面直接写 AI 合成主播,但实际上很多主播明显是一个动漫人物,所以这个能不能叫 AI 就会有一点存疑。”事实证明,以人工智能技术和 AI 合成主播为噱头“挂羊头,卖狗肉”的行为(X2, X3)其实不在少数。由于 AI 存在技术黑箱,受众的真假辨别能力非常有限,AI 合成主播的言说能力和作用广受质疑,也与其背后商品逻辑下隐蔽的符码力量紧密相关。

不能否认,在当下媒体传播中 AI 主播仍无法摆脱其作为一种“物”的命运,是一种被观看、被消费的对象,也是一种被人们使用的工具^[17]。2021 年发布的《广播电视和网络视听“十四五”科技发展规划》中虽提出要加快推进制播体系技术升级,推动虚拟主播广泛应用于新闻播报、天气预报、综艺科教等节目生产,创新节目形态,提高制播效率和智能化水平^[18],在政策的倡导下,许多地方县级融媒体和基层传媒平台也都配备了 AI 合成主播,并将其融入当地的内容生产流程中,但 AI 合成主播在语音播读新闻和客户端新媒体同步等方面仍存在诸多局限。从受众的角度看,AI 合成主播作为主播的角色功能正在被进一步弱化,除了各地的技术配备与使用场景并未完全达标之外,受众的接受欲望和信息记忆也都处于较低水准。访谈者 Z2 表示:“我觉得好像它是一个装饰性的,实用价值好像不是非常大,也没有感觉到因为搭载了这个技术,增强了什么样的互动性。”另外,大部分 AI 合成主播的应用都是录播好之后用合成技术转译,在本该保证的时效性上也大打折扣。这一定程度上表明了 AI 合成主播在新闻生产中的应用迎合了以物品为载体的“心理—文化”需求满足的符码消费,不仅停留在对物的片面占有和消耗,而且盲从了荫蔽在符码含义镜像后的价值判断^[19]。当 AI 合成主播不再发挥言说的核心功能,而是作为一种技术配备和科技符号存在于信息生产流程中时,始终作为言说主体的人就需要理性面对技术勾画的未来图景,避免陷于技术异化的风险中。

三、交流的无奈:人机互动中的伦理(观念)困境

在哈贝马斯的理论视野中,会话的顺利进行是一种理想的状态,而实际的情况是会话总会受到

“系统”的制约^[20]。这种制约主要体现在会话中的主体之间总是存在不平等的关系,而言说的导向也总是被外在的因素所影响,这种具备日常生活情境会话的困境与互动中忽略认知水平、文化共鸣与认识志趣等重要因素不无关系。尽管 AI 合成主播的研发团队一直在使用技术进化其情感感知力与交互能力,试图逐步改善 AI 合成主播与人之间的“言说”难题,如 2022 年科大讯飞基于 AI 虚拟人交互平台底座研发了“AI 虚拟人直播系统”,在上百个虚拟形象和上千个音库基础上提供多模感知、多维表达、情感贯穿、自主定制等服务,让虚拟主播的形象更真实、情感更灵动、交互更智能^[21],但即便如此,在 AI 合成主播与人言说的情境中,深受技术设置和智能编排的“系统”制约,处于人格化传播中的载体(即 AI 合成主播)带给受众的仍然是互动的假象。或千篇一律或性格各异的人格设计实际上是技术带来的伪个性化,受众事实上无法感知到真实而稳定的人格,也无法在“不平等”的人机互动中发现像真人主播一样的人格魅力,更难以体会到言说作为“交往行动”中的社会性。更为重要的是,在人机互动的过程中,许多受访者表示被“系统”侵蚀的会话还会反过来改造人的认知、爱好、行为倾向等,在会话的虚假契合中形成“交往异化”(Y2, Y3, Y5, Z2)。AI 合成主播作为技术主体会存在系统设定的“自我主义倾向”,频繁地以机器逻辑感知人为会话;而受众作为实在主体,也难以避免会在工具理性的思维定式中学着“像机器一样思考”(Z1, Z3),忽略会话与言说中可能出现的创新意识,受访者认为频繁陷入“被编程的自我意识”中是不被察觉的,“一旦后期反应过来,会有细思极恐的感觉”(Z4)。

人机之间的交往伦理与交流观念一直以来都处于变动之中,传播的“传递观”与传播的“撒播观”一直以来是人机交流观念的讨论核心。人机交流中的不平等,不仅会影响到人机交流的结果,它为人所培养的思维和行为习惯,特别是以自我为中心的惯性,也会为人际传播带来障碍^[22]。以下从交往感知、人格感受和情感位移三方面,分析受众效果层面透露出了哪些人-机交往中不稳定的观念,在观念与实践的变化中开启探讨人机伦理问题的大门。

(一)从技术好奇到技术恐慌的交往感知

AI 合成主播带给受众的交往感知往往呈现为两种截然不同的态度,分别是技术新奇感和技术恐慌感,二者不仅存在于不同受众接触 AI 合成主播的各个时期,也会在一定程度上相互转化(如图 1)。国际学者对 AI 合成技术的公信力问题保持了高度关注^[23],在技术带来的新奇之余,也有学者将人机互动的信任感问题纳入新闻伦理的考量中^[24],这表明人们普遍认为 AI 合成等技术对受众的交往感知影响是复杂的。新技术融入交往场景在初期一般都会引发受众的狂热追捧,而当技术发展到一定程度,受众对新鲜事物的热情逐渐消退,就能够感受到技术对人主体性的压迫感,从而开始产生技术恐慌。技术主义为了消除不确定性,以帮助人们获取安全感为借口大规模入侵社会到一定程度,则会忽视人的现实存在,制造出诸多人学空场,从而引发普遍的道德恐慌^[25]。正如被访谈者 Y3 提到:“我觉得 AI 合成主播的技术如果发展到特别像人,我会感到非常恐慌,一方面是难以分辨是不是真人,另一方面这也会让我在真实世界中感到迷茫。”技术新奇和技术恐慌的感知转换同样也体现在接触 AI 合成主播的程度上,在一般情况下,接触程度较高的受众会更更多地在意交往过程中的有效性和愉悦感,而接触相关资讯但不常与 AI 合成主播产品进行交往互动的受众则只会将注意力停留在认知新事物层面。面对 AI 合成主播的技术恐慌,机器人专家莫里曾在 1970 年警告指出,机器人不应该与真人相似,因为这样的机器人可能落入“恐怖谷”,当机器人几乎像人类一样行动时但却未能获得逼真的外表,人们对类似人类的机器人的反应会随着与人相近突然从认同转变为厌恶和惧怕^[26]。AI 合成主播显然正随着技术的进化而逐渐贴合“恐怖谷”效应的预测。因此,在交往实践中,作为言说客体的人依旧会被言说内容之外的技术因素和系统因素影响交往感知,甚至这样的交往感知还会深度介入交往的过程,让人们从关系中抽身而出思考“系统”的问题,从而脱离交往与言说的实质。

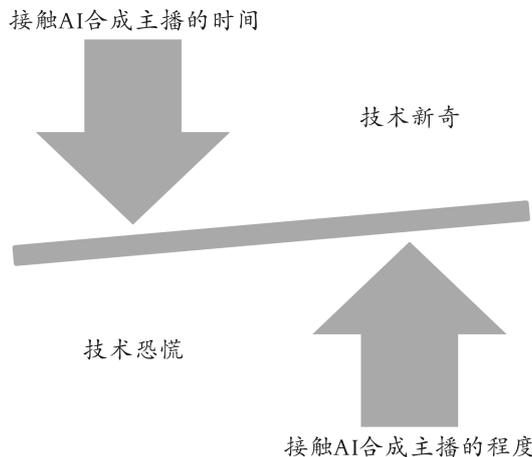


图 1 受众接触 AI 合成主播的交往感知变化示意图

(二) 从难辨真伪到不辨真伪的人格感受

一般而言, AI 合成主播的虚拟人格主要沿着两条思路进化:一是逐步通过模仿实现拟人化,二是通过技术能力和编程设计实现个性化。这其中都包含着受众对 AI 合成主播不可避免的“真/假”判断,在很长一段时间内更是人们关注 AI 合成主播的焦点。当关于 AI 合成主播的报道和宣传集中于如何像真人时,受众的期待不再存在于希望看到“以假乱真”的 AI 主播,而是转向了彻底区分技术背后的真假二元对立,产生“不必辨认真伪”的人格期待。此外,在视觉影像所建构的消费世界中,现代性审美精神也会通过确立标准与规范的方式对身体进行压抑与控制^[27]。被访谈者 Z2 谈到:“除了 AI 合成主播模拟真人的脸之外,未来还有可能再慢慢模仿身材,再模仿说话,就是我觉得我们在把掌控自己身体的权利让步给了一种新的技术,我觉得让我对于我们的主体性会产生怀疑。”

事实上,无论是 AI 合成主播的拟人化(虚拟)人格还是其个性化的人格设定背后体现的仍是受众在观点与情感层面的不同的人格期待,前者是通过附着在真人原型上获得性格与情感的“再生”,后者则是通过技术形成自己独特的个性,拥有解决问题的能力与智慧。目前 AI 合成主播的离身认知、具身认知、认知交互及人机交互还是初阶段,其主体性、创造性、艺术性仍然有限,类似局限亦体现在初阶人工智能新闻传播的采集、生产、分发、接收、反馈中^[28]。在“认知可以被计算”的大前提下,人工智能脱离人的身体而存在观念意识的设想目前仍面临着很大挑战,在人机互动的交往实践中 AI 合成主播技术上无法实现身心二元,互动的过程中主体与客体之间也不能同时脱离身体与灵魂孤立存在。被访者 Y3 提到:“即便 AI 合成主播说出了一个观点,我都会想这个观点真正的输出者到底是他还是他背后的人,去追根溯源,因为我无论如何都没有办法把一个人工智能当作一种思想的源头。”可见,作为言说客体的人始终期待交往的另一面是具备独立思考能力和智慧的人格载体,而 AI 合成主播却无法实现这样的人格期待,即便不执着于“真假”,受众仍然无法脱离寻根溯源的本能去关注人机交往中的内容与信息。

(三) 从互动言说到技术性陪伴的情感位移

在智能机器人能否作为“第六媒介”的讨论中,有学者指出人机传播与人机交往存在本质性的区别,人机传播和人机互动强调机器与人之间的信息流动性,而人机交往既包含前者,也囊括机器与人之间交往的“主体能动性”^[29]。当前人工智能的技术发展仍处于人机传播与人机互动的层面,但与 AI 合成主播言说却成为人机交往设想中的理想情境之一。这是因为侧重于“人-机”双向交往行为的 AI 合成主播,不仅能够具备以往数字媒介所能提供的定制服务功能,还能将大众传播与人际传播、人内传播结合,通过人机交往充分唤起人类全方位、多感官、高镜像的沉浸体验、具身认知和主体意识。

即便技术发展仍然在路上,但互动性的言说已不再是受众对 AI 合成主播唯一的期待和想象。当前的 AI 合成主播虽然无法完全在思想、认知和情绪上与真人产生共鸣,但却能够以一种稳定的技术式存在给予老年群体、空巢小孩、残障儿童等群体提供陪伴的情感价值(Z4,Z5)。例如当前人机交互领域中对情感计算的相关研究,就是以人类情感为基点,利用计算机仿真技术与方法将人类情感进行量化分析并进行适当反馈^[30],探究其技术肌理如何促进形成与人的良好关系。AI 合成主播或许更适用于未来人们之间长期的情感陪伴与亲密关系的建立,毕竟,它可通过精准的情感计算与识别,对用户进行有效的情感补偿,与用户建立专属的情感陪伴,使之成为用户信任的、依赖的、有影响力的智能主体^[31]。随着未来技术的不断演化,人与 AI 的互动方式、交流途径会愈加丰富,“作为治疗(陪伴)的 AI 主播”(Y2)“作为游戏 NPC 的 AI 主播”(Z1)以及更加工具服务类的交流联系会更加紧密,但考虑到人工智能带来的伦理安全风险和社会后果,仍有学者持续地展开讨论,以人文主义为基点来限制人-机之间更深度情感关系用来规避可能出现的负面影响。

四、简单化场景:口语“灵韵”的消逝

灵韵是本雅明在《机械复制时代的艺术作品》一书中提出的概念,他认为,传统的艺术品是不可复制的,这种独一无二性使原作本身具有了一种神秘的“灵光”,即“灵韵”^[32],人们在这些艺术杰作面前,只能顶礼膜拜,聚精会神地欣赏。原本的机械复制技术打破了艺术品的灵韵,而今天 AI 合成主播的出现更像是不断复制真人主播形象出现的虚拟人,在视听领域一定意义上也打破了原本真人主播既有距离感又有此时此地本真性的口语灵韵。

从应用场景来看,AI 合成主播在游戏娱乐领域、信息传播领域和电商直播领域都有着丰富的发展空间,但在各个场景中的角色和定位却各有不同。在游戏娱乐领域中,AI 合成主播侧重于借助符合游戏场景化的形象来参与互动(Z1);在信息传播领域中,AI 合成主播侧重于运用语音制作和合成技术完成“语音播报”的功能,其人格化传播多受到媒体形象和媒体定位的制约;在电商直播中,商业和消费的场景有助于将 AI 合成主播塑造成商品品牌的“代言人”,借助 AI 合成主播自身的人格特质来吸引更多消费者的购买行为,如淘宝虚拟主播“小玉”等主要用途便是推销商品(Y3)。结合业界的实践来看,目前的 AI 合成主播更倾向于拆分功能特性,服务于特定的应用场景,将人在各个场景中的多元需求简单化。然而,属于人丰富而深厚的情感与情绪都应当在与主播的言说互动中存在,人并不是在消费的场景仅仅为了消费,也不是在游戏的场景仅仅沉迷于游戏,“我进入淘宝直播间,并不是为了买东西,有时候觉得与 AI 合成主播偶然相遇是件很有趣的事情,希望能多互动互动,但直播间中的 AI 主播好像只是想卖出商品。”(Z2)

马尔库塞在 1964 年出版的《单向度的人:发达工业社会意识形态研究》中曾提出“单向度的人”概念,意指个人在高度发达的工业社会中丧失了合理批判社会现实的能力^[33]。新媒介语境下,技术同样为人们建构了一个全新的“单向度”空间^[34]。人在各个言说场景中都有对社交的现实需求,而社交感很大程度来源于口语交往中的多义性与即时反馈。当 AI 合成主播过多陷入“单向度”的场景与功能时,就无法在言说中抓住灵光一闪般的口语“灵韵”。说到底,言说仍需依赖声音媒介,技术不应该以“进化”为名掩盖口语和社交在言说场景中所带来的可能性。技术进化可能对口语“灵韵”的侵蚀主要源于以下几个方面。

其一,技术设定可能导致社交能动性萎缩。在大部分 AI 合成主播使用场景中,不论是带有大众媒介属性的新闻播读,还是具备私人订制属性的聊天式主播,都仍处于独白型的叙事模式。言说与交往的意义在于不断地对话与沟通,在不断的双向反馈中让话题向纵深处发展。而 AI 合成主播的单向播读却不能获得预期中的反馈,即无法得到用户满意与否的反馈;AI 合成主播与受众的交互过程中同样也无法生成具备“主体间性”的叙事空间,AI 合成主播有既定的播读和反应文本,而这些都来自编程与技术设置。只有在主体的交互中,才能实现我与他人的相互构造,并在这个交互过程中

认识自我,强化个体的主体性^[35]。正如胡塞尔所言:“我所获得的并不单纯是在身体上与我对峙的、通过联想的结对而又回溯到我心理物理的此在的他人。”^[36]因此,从严格意义上来说,AI 合成主播与受众之间并不存在话语共同体的言说可能性,即不论 AI 合成主播和谁对话都可能拥有同样的应答“套路”,这实际上说明了 AI 合成主播目前与受众难以形成有效的社交关系。

其二,工具理性可能导致言说过程中的自我主义倾向。在哈贝马斯的交往理论中,交往理性有助于主体间的调和。在更多的应用场景中,AI 合成主播不仅是传递信息的“智能喇叭”,更需要面临被询问、被求助和被交谈的互动情境。当言说的场景中受众存在“观点、态度需求”时,AI 合成主播只能以由算法设定的工具逻辑来应答,无法随着受众的实际情绪变化而共情。甚至就连 AI 合成主播程序中的“共情力”也是对平均情绪的测算,具备技术先入为主的机器自我主义倾向,人机双方都存在不可通约的“自我主义倾向”,这背后透漏出的是人机之间不平等的言说地位。在人类与人工智能能力不对等的情况下,人类与人工智能将处于统治与奴役的主奴辩证法中^[37],这意味着当前普遍应用的弱人工智能只能成为被奴役的“他者”。如果理解到人与技术之间实际上存在彼此控制和掣肘的关系,就会发觉人其实难以与机器共情,机器也难以通过口语交流和言说真正地参与沟通。

其三,智能程式化一定程度上影响口语文化中的二度创作。对受众而言,口语文化的来源是声音媒介的各个侧面能够释放出能量的总和。在播音主持传播中,从原始文本的“一度创作”到新闻播读的“二度创作”之间有很大差异,这不仅仅是从原作者的思想片段转化为文字媒介,再将文字媒介背后的本意、引申义和联想义串联起来转化为声音媒介的过程,更是凝结不同阶段创作者的个人体验、经历沉淀与文化认知和个人风格。经过口语的评述与播读,受众不仅能够在播读神态、语音面貌、停连重音等细节背后打捞出原作者文字的“言外之意”,不同的二度创作者还能为原作者的文字拓宽想象空间,营造出不同的情景与语境。在这个纬度上,情感的生发并非一种程式化的机械处理,而是涌动着的不不断变化着的感情起伏^[38]。失去经过多重创作后的内容,口语文化的灵韵面临消逝的风险。AI 合成主播目前的发展尚未满足新闻播报的艺术性,AI 合成主播不断智能化的演进趋势也只能缩短文本所富含口语文化的意义链条,甚至有可能出现因为技术与人发声逻辑的不同而产生意义误读的情况。

其四,技术简化交往情境并可能形成固定单一的交流模式。复杂情境是人类生活的映射,人类的口语交往常常是在复杂情境中生发的。人改造工具的同时工具和环境也改造了我们。人在与 AI 合成主播互动的过程中会带入人类世界口语中的习惯词汇与交流偏好,而长期的人机互动也会逐渐影响人们的交流方式与交往习惯。智能化的语音交往往往具备言说的欺骗性,机器的功能设置也只能涉及有限的场景,难以覆盖人类会面临的所有情形,而人的口语交流行为具备随机性和多元性。AI 合成主播的互动是程序设定的虚拟反馈,其互动不会因观众的意见和情感而发生变化,只是根据内容的需要和程序的设定流露出合理的情绪^[39]。当 AI 合成主播以二进制的技术设置回应处于复杂情境中的人时,不但无法与人达成有效言说,而且还会影响受众,从而有可能将固定单一的交流对话模式代入到日常生活当中。

五、讨论与启示:进化中的异化

2022年5月6日,海外 AI 合成主播 Vox Akuma 在 B 站进行了首次直播,播出时长 1.7 小时,营收达 111 万元人民币,付费人数近 4 万人^[40]。以 AI 合成主播为中心的虚拟主播产业如火如荼地发展不仅意味着主播这一大众媒体时代的人格化传播角色正在搭载人工智能技术的发展焕发出新的生机和活力,具有了更加丰富多元的言说功能,越来越多地介入人们的日常生活和大众文化之中;同时也意味着 AI 合成主播在技术上的再次进化,意味着 AI 合成主播从“形似”走向“神似”的技术演化路径。如前文所述,AI 主播技术进化的起点是语音、皮肤以及画像的精准模拟,而进化的方向则是塑造 AI 合成主播的虚拟人格和发掘其人格化魅力,即从赋能到“赋魂”^[41]。这也从一个侧面说明,在

技术的演化过程中, AI 合成主播同时承担着“技术性的主播”和“言说的工具”的两重角色,且二者间具备强烈的关系互动可能性。

不无遗憾的是,在信息的生产与传播技术的加持下,原本该处于重要地位的人格化传播主体——主播却不幸陷入了技术的黑箱,原本传受之间的交流与反馈空间也变得愈加狭小。智能时代的技术吞没了口语文化应有的亲密人际交往,以信息的通达、效率为有效传播标准的 AI 合成主播的出现和进化也忽视了受众对“主播”的人格化角色期待。口语言说的艺术与灵韵的消逝已在证明技术的进步可能会忽视和阻碍人在社交情境与口语表达中可释放的可能性。AI 合成主播实际上带给受众的是欺骗性的互动想象和通过技术编码批量复制的流水线“人设”。而当 AI 合成主播追求的“人格化”成为一个可供制造和消费的文化符码时, AI 合成主播也就陷入了一种由技术进化引发的言说异化的风险。

异化的概念来自德国古典哲学,黑格尔、费尔巴哈、马克思以及哈贝马斯对其都有阐释和发展。异化理论与主体有着紧密的逻辑关系^[42],常被用来阐释本来由主体产生和创造的物最终成为外在于主体,并与主体存在相对立力量的社会现象。马克思的研究则使异化的主体聚集于人。在马克思看来,人类社会在发展自身的同时,创造了各种人工物,然而在特定的社会条件下,人的活动及其产物并不一定都符合人的需要和目的,甚至完全背离了人的需要和目的,成为统治者、控制人、支配人的异己力量^[43]。

从 AI 主播的发展实践来看,技术虽然一定程度上实现了主播言说向着更加“有效”方向上的进化,并从某个侧面促进了真人主播在人格化方面的进化^[44],但同样也使主播失去了情感的真实联结,失掉了口语传播的创作力与想象力,甚至完全背离了人的需要和目的。因此, AI 合成主播所实现的只能是有限的言说,甚至是对空言说。人创造出 AI 主播是为了还原或模仿真人之间的交流,交流主体的主体性(属于人的主体性)是追求的目标之一,但却在模仿或还原的过程中被技术异化,反而丧失了真正属于人的主体性。最重要的是,这一过程隐蔽地发生在人机交流中,反而以人机高度互动为掩盖,使人容易忽视其中数字人格的“虚伪性”,陷入技术狂欢中。

在互联网实现的各种连接中,人与人的连接是核心^[45]。这似乎在启示人们首先应当在技术演化中嵌入人文价值,具体到 AI 主播中就是要透过 AI 合成主播的设计、编程和命名来真实地创造人格化传播技术主体,从交互方式到沟通效果上充分考虑受众的需求和感受,做到“以人为本”地建构人格;而在人机共生的过程中,则要平衡信息效率与情感连接的关系,发挥人工智能在新闻播报中的离身认知、具身认知及交互应用方面的效能,重视主持传播中的口语文化与表达艺术,从而实现 AI 合成主播在技术创新与人文精神上的协调和统一。

在 AI 合成主播所进行的言说与交往中,充斥着受众对理想人格、身份认同和技术互动的强烈诉求,也处处折射出受众面临技术飞速进步时在心态、期待和需求等各个方面的及时调试。技术的狂飙突进不仅为人类带来了更加便捷、高效的海量信息,更重塑和改变着人与技术交往的方式和形态。看似欣欣向荣的 AI 合成主播业态前后其实也潜藏着人的主体性和交往方式异化的风险,进化中的异化现象值得警醒,人机之间亦依然存在对空言说的窘境,为了避免让言说再次落入异化的陷阱,一切都需要回到技术进化的原点重新思考。

参考文献:

- [1] 孙维张. 佛源语词词典. 北京: 语文出版社, 2007: 305.
- [2] 彭喆, 梁雯. 勒·柯布西耶的不可言说的空间: 诺斯替主义与精神分析. 美术大观, 2022, 2: 126-132.
- [3] 张昭炜. 中国儒学的缄默维度特质. 世界宗教研究, 2018, 1: 15-31.
- [4] 陈晓平. 知识定义与默会知识——从“盖梯尔问题”谈起. 现代哲学, 2013, 6: 78-83.
- [5] 张延飞, 张绍杰. 再论索绪尔的语言本体论——与李文新先生商榷. 外语学刊, 2014, 1: 9-13.

- [6] 黄汉平. 拉康的主体理论与欲望学说. 文学评论, 2010, 3: 194-199.
- [7] 胡百精. 共识与秩序: 中国传播思想史. 北京: 中国人民大学出版社, 2022: 57.
- [8] 吴新民. 梅洛·庞蒂现象学的奠基观念及其当下启示. 渭南师范学院学报, 2018, 11: 54-61.
- [9] 相芯科技观察: AI虚拟主播简史. 雪球科技共话互联, 2019-05-29. [2022-11-15] <https://xueqiu.com/1779398252/127449120>.
- [10] 郭建中. 身体、遮蔽与新中区: 对AI合成主播技术具身的反思. 现代传播(中国传媒大学学报), 2022, 1: 98-103+105.
- [11] 余文, 吴飞. 从AI合成主播看新闻传播的融合创新发展. 传媒, 2022, 4: 31-33.
- [12] 刘霞, 陈昌凤. 3D+AI主播: 伦理演进与价值引领. 青年记者, 2021, 5: 60-61.
- [13] 王忆希, 吴福仲, 王峥. 人工智能新闻主播何以被接受?: 新技术与社会行动者的双重视角. 全球传媒学刊, 2021, 8: 86-102.
- [14] 喻国明, 耿晓梦. 试论人工智能时代虚拟偶像的技术赋能与拟象解构. 上海交通大学学报(哲学社会科学版), 2020, 28: 23-30.
- [15] George Lukács. History and Class Consciousness. MIT Press. Marcuse, 1971: 83; 91.
- [16] 杜丹. 共生、转译与交互: 探索媒介物的中介化. 国际新闻界, 2020, 5: 18-34.
- [17] 胡丁琿. 国内AI主播的演变历程、发展困境与未来出路. 中国传媒科技, 2021, 6: 26-28.
- [18] 国家广播电视总局. 广播电视和网络视听“十四五”科技发展规划. 2021-10-08. [2022-08-04] http://www.nrta.gov.cn/art/2021/10/8/art_113_58120.html.
- [19] 范和生, 刘凯强. 符码消费镜像中的心理异化及实践逻辑——兼论鲍德里亚的消费社会学. 宁夏社会科学, 2016, 2: 104-109.
- [20] 徐苗苗, 李晓敏. 哈贝马斯语言哲学对“交往异化”问题的分析与解决. 黑龙江社会科学, 2022, 3: 16-20+128.
- [21] 吴长锋. AI虚拟主播“上新”性格与角色可按需定制. 科技日报, 2022-05-18(002).
- [22] 彭兰. 人机传播与交流的未来. 湖南师范大学社会科学学报, 2022, 5: 12-22.
- [23] H. Van der Kaa, E. Kraemer. Journalist Versus News Consumer: The Perceived Credibility of Machine Written News// The Computation+Journalism Symposium. New York: Columbia University, 2014: 24-25.
- [24] T. Montal, Z. Reich. I, Robot. You, Journalist. Who is the Author? Authorship, Bylines, and Full Disclosure in Automated Journalism. Digital Journalism, 2017, 5(7): 829-849.
- [25] 张洪根. 论技术主义时代的道德恐慌策略及其悖论. 自然辩证法研究, 2004, 7: 70-73+88.
- [26] 朱体正. 仿人机器人的法律风险及其规制——兼评《民法典人格权编(草案二次审议稿)》第799条第一款. 福建师范大学学报(哲学社会科学版), 2019, 4: 117-128.
- [27] 荆莹莹. 偶像养成类真人秀节目中身体景观形塑与权力话语运作. 当代电视, 2021, 8: 72-76.
- [28] 於春. 传播中的离身与具身: 人工智能新闻主播的认知交互. 国际新闻界, 2020, 5: 35-50.
- [29] 林升梁, 叶立. 人机·交往·重塑: 作为“第六媒介”的智能机器人. 新闻与传播研究, 2019, 10: 87-104+128.
- [30] 付心仪, 薛程, 李希等. 基于姿态的情感计算综述. 计算机辅助设计与图形学学报, 2020, 7: 1052-1061.
- [31] 崔洁, 童清艳. 解构与重构: “人格化”虚拟AI新闻主播再思考. 电视研究, 2022, 2: 62-64.
- [32] 鲍里斯·格洛伊斯. 艺术力. 杜可柯, 胡新宇译. 长春: 吉林出版集团股份有限公司, 2016: 75.
- [33] 马尔库塞. 单向度的人. 张峰, 吕世平译. 重庆: 重庆出版社, 1998: 31.
- [34] 窦艳伟, 林进桃. 论新媒体语境下“单向度”社交生活的建构与呈现. 文化与传播, 2021, 6: 5-8.
- [35] 韩敏, 赵海明. 智能时代身体主体性的颠覆与重构——兼论人类与人工智能的主体间性. 西南民族大学学报(人文社科版), 2020, 5: 56-63.
- [36] 埃德蒙德·胡塞尔, 克劳斯·黑尔德. 生活世界现象学. 黑尔德编//倪梁康, 张廷国译. 上海: 上海译文出版社, 2002: 194.
- [37] 黑格尔. 精神现象学(上卷). 贺麟, 王玖兴译. 北京: 商务印书馆, 1997: 127-128.
- [38] 张颂. 中国播音学. 北京: 北京广播学院出版社, 1994: 34-35.
- [39] 孔令强. 模仿、创新与新闻黑箱——对“AI合成主播”的技术反思. 传媒, 2020, 17: 47-49.
- [40] 张核, 甲子光年. 虚拟人是来爱你的, 还是来挣钱的?. 澎湃新闻, 2022-01-26. [2022-07-15] <https://www.thepa->

per.cn/newsDetail_forward_16444632.

- [41] 杨名宜,喻国明. 赋能与“赋魂”:数字虚拟人的个性化建构. 编辑之友,2022,9:44-46.
- [42] 周瑞辉,王忠桥. 马克思异化理论及其实践向度. 北方论丛,2016,1:108-112.
- [43] 魏巍. 刍议科技异化与科技生态化//科学发展·生态文明——天津市社会科学界第九届学术年会优秀论文集(上)2013:125.
- [44] 高贵武,杨航. 现实虚拟:技术发展与主持传播的人格进化,中国主持传播研究,2019,2:3-15.

Evolution and Alienation: the Articulation Dilemma of AI Synthetic Anchors

Gao Guiwu, Zhao Xingzhi (Renmin University of China)

Abstract: Articulation has the meanings of oral expression and message transmission, as well as preaching and propaganda in Buddhism. Artificial intelligence (AI) synthesis technology is a technical tool to improve communication, but it causes a crisis of human subjectivity. This paper introduces alienation theory's perspective and conducts in-depth interviews with ten audiences who have reached and received AI synthetic anchor products at various levels, and analyzes the triple dilemma that AI synthetic anchors may face by deeply studying the process of audience perception of Articulation to critique the communication effectiveness of AI synthetic anchors.

Key words: AI synthetic anchor; subjectivity; communicative rationality; alienation

■ 收稿日期: 2022-08-05

■ 作者单位: 高贵武, 中国人民大学新闻学院; 北京 100872

赵行知, 中国人民大学新闻学院

■ 责任编辑: 刘金波