

Doi: doi.org/10.70693/rwsk.v1i2.670

AI 主播与真人主播：一项语言审美视角的探析

牛勇博¹

(¹ 马来西亚博特拉大学, 雪兰莪, 马来西亚 43400)

摘要: 本研究从语言审美的角度分析了人工智能主播与真人主播的表现,并研究了不同语言体系(英语、汉语和日语)的语音美学特征。研究表明,真人主播在情感表达、语音韵律和人文价值方面更具优势,而 AI 主播尽管通过语音合成技术不断改进,但在审美方面仍存在限制。真人主播能够通过声音创造更具文化深度的表达,而 AI 主播难以再现这种美学体验,因为语言审美在汉语中不仅表现在声调和节奏上,而且还强调“意境”、“情致”和“韵味”等诗性特征。此外,不同语言体系的美学标准影响了观众对 AI 主播的接受度。例如,英语强调重音和节奏,而日语则强调音节均衡和敬语体系的韵律特征。研究表明,为了更好地满足不同语言用户的审美需求,人工智能主播未来的发展应结合汉语语言的美学特征。

关键词: AI 主播、真人主播、语言审美、语音美学、跨文化比较

一、中国美学中的声音美感与语言审美

有声语言审美是中国美学中一个重要而复杂的议题,涉及语音、韵律、声音的节奏、声调、音韵美、文辞音律等方面。在中国美学中,语言审美不仅仅是语义的传达,更包含了声音的韵律美和表达方式的艺术性。中国美学认为声音的美感对语言审美至关重要。“任何艺术形式都不能只满足于语言学层面的分析或呈现,而应该纳入审美交往的维度,考察创作者与受众之间的交往关系。表达者并非单方面的单项输出,而是在和受众的互动中,才能真正使人体体会到声音的美妙之处。”^[1]

音色的婉转、柔美或清越、铿锵能够直接影响人们对语言的感知和接受程度。在古典诗词或传统戏曲中,语音的优美往往被视为重要的审美标准,如平仄、押韵等都与声音美感有关。中国美学强调表达方式的艺术性,语言审美不仅包括语音的美感,还有表达时的声调、语气、停顿等细微之处,这些都能为语言增添审美的艺术感。“从中国美学的角度看,有声语言的审美是一个综合性的概念,涉及声音的美感、音韵的和谐、文辞的音律以及表达方式的艺术性。这些因素共同构成了语言美学,使人们在交流中更加愉悦、感性,并且更好地传达情感和意境。在官方政治、人士文化和市民生活三大层面的语言传播活动共同影响下,吸收传统文化精髓,结合古代社会语言传播实际需要的汉民族语言传统美学体系日渐形成,有声语言艺术活动的内在美学精神和审美范畴也日趋明确,典范美、人文美与艺术美是其主要审美范畴。”^[2]文章中明确给出了三种美的具体解释。““典范美”所指出的不仅是有声语言活动在礼乐仪式和社会传播中所保有的规范之美。实则更强调在有声语言垂范下所形成的社会文明之美。”^[2]“有声语言的“人文美”指的是对有声语言的思想、情感、人文价值等层面的美学考量。”^[2]““艺术美”的审美范畴是在美育性的审美评价标准的基础上进一步明确的有关有声语言活动在艺术表现与艺术精神方面的美学命题。”^[2]汉语有声语言表达尤为重视语言“意境”、“情致”、“韵味”等诗性功能的实现,从而领悟“言外之旨”,达到有声语言美学价值的提升。^[3]

换言之,在有声语言的审美过程中主要包含了一下的几个重要纬度。首先,受众对声音的审美离不开的是音色。音色的优美是有声语言审美的重要方面。张颂教授提出,有声语言的审美空间应具备规范性和艺术性,以为受众带来深刻而高雅的审美体验。^[4]在有声语言作品的审美过程中,规范性与艺术性相辅相成,缺一不可。一部优秀的有声语言作品,首要之务便是确保语言的规范性,这不仅关乎作品的传播效果,更

[作者简介]

牛勇博(1989—),男,云南昆明人,马来西亚博特拉大学在读博士研究生。

是审美体验得以成立的基础。语音、语法、词汇的规范,既是准确传递信息的前提,也是艺术性展现的基石。正如中国自古以来的哲学理念所言——“无规矩不成方圆”,在有声语言艺术中,规范之“规”塑造了表达的边界,而艺术之“圆”则赋予其韵律与美感。唯有在规范之中寻求突破,于严谨之上追求灵动,方能使有声语言作品既具传播价值,又不失艺术魅力。

其次,万物都是在律动与变化之中,唯有旋律美和节奏美的有声语言才具有审美价值。曾志华和卢彬的研究指出,播读者的语言形式应体现旋律美和节奏美,以增强有声读物的艺术感染力。^[5]

人类的听觉器官进化使我们能够体验声音的美感。这种感性体验是有声语言审美的基础,通过声音的抑扬顿挫、轻重缓急等特征,引发听众的情感共鸣。^[2]审美活动本质上是主客一体的,两者相互作用,彼此影响。审美主体的感知、情感与思考,往往受到自身的成长背景、生活经历、文化积淀等多重因素的塑造,而审美客体的存在形式、呈现方式也在无形中影响着主体的认知与体验。朱光潜先生在《谈美》一书中,以“一棵古松的三种态度”为例,深入剖析了人类对同一客体的不同理解方式:实用态度关注其材质与功能,科学态度探究其生长规律,而审美态度则超越了功利和理性,单纯欣赏其形态之美、风骨之韵。这一理论不仅揭示了审美的多重层面,也强调了审美活动的主观能动性。换言之,不同的阅历、认知与文化背景,会使人们在面对同一客体时产生截然不同的审美体验,而这正是审美活动的独特魅力所在。

最后,有声语言能够通过声音的高低、快慢等形式,传达特定的意象和情感。这种声音意象使听众在听觉上产生联想和共鸣,增强了语言的表现力和感染力。^[6]受众通过声音的高低、快慢、节奏变化,不同的语音特征可以传达特定的情感和画面,例如,低沉缓慢的语调带来庄重或悲伤的氛围,而高亢急促的声音则易让人感受到激动或紧张。这种声音意象能够激发听众的想象力,使其在听觉上产生联想,与语言内容建立深层次的情感共鸣,从而增强语言的表现力和感染力,使表达更加生动、富有魅力。

综上所述,有声语言审美在中国美学中扮演着关键角色,涉及多个方面。不仅仅是语义传达,还包括声音的韵律美和表达方式的艺术性。这种审美不仅使语言更具魅力,还加深了人与人之间的情感交流。在中国传统文化和现代社会语言传播的影响下,有声语言艺术活动的内在美学精神和审美范畴逐渐清晰,典范美、人文美和艺术美成为其主要审美范畴。通过强调语言的诗性功能,特别是语言意境、情致、韵味等方面的实现,有声语言美学的价值得以提升。

二、真人主播与 AI 主播在语言艺术美中的比较

审美活动是人类独有的一种思想和内心的活动,它是由外化到内化的过程,在这个过程中审美一半依赖于审美主体,另一半依赖于客观的现实世界。中国美学大师朱光潜先生说:“美,一半在你,一半在我。”也就是这个意思。“我”是审美的主体,“你”即客观现实。既然审美活动有一半在于主体另一半在于客体,就需要主体发挥主观能动性来完成审美活动或者必须具有审美意识,从这个层面上来说 AI 主播不能完成审美活动,也不具备审美意识。

真人主播通过情感表达、思想传递和人文价值观的呈现,能够以更自然、真实的方式传递他们的情感和思想,建立起深厚的情感联系。他们独具个性、丰富的经验和深厚的文化背景,这使得他们在广播电视领域更为引人注目。成为一名卓越的播音员和主持人需要经过专业的训练和长期的学习,这一过程从大学阶段的语音发声训练开始,逐步深入到播音创作基础等专业课程。毕业后,他们需要在实际工作中不断磨练自己,提高专业水平。这个过程不仅仅是技能的培养,更是一段宝贵的人生财富的积累。这些经历和经验为播音员和主持人奠定了良好的基础,使他们能够更好地参与人文审美活动。通过不断润色加工自己,他们能够更好地展现出艺术美和艺术精神,使得他们的语言表达更富有深度和广度。这样的专业背景和经历不仅提高了他们在广播电视领域的专业水平,也丰富了他们在人文美层面的表达能力。

董卿在中国是一位家喻户晓的著名节目主持人,她的主持风格不但大方亲切,更重要的是她的语言表达可以让人有一种“腹有诗书气自华”的感觉。腹有诗书气自华就反映出了董卿有一种内在的文学与人文氛围,蕴含着丰富的思想和情感。腹有诗书表示个体内部拥有丰富的文学知识和诗意情怀。这反映了对文学的热爱和追求,文学氛围赋予了这句话一种深邃的文学意境。气自华强调个体内在的气质和修养。这里的“华”指的是文明、高尚的品质,说明个体通过诗书的积累和学习,使得自身的气质变得丰富和高尚,具有一种自我充实的美学特征。殊不知董卿平时就是一个爱阅读享受阅读的人,她喜欢阅读文学哲学方面的书籍,正是她的这一生活习惯才让董卿的语言具有很明显的人文性、艺术性。董卿不论是在中央电视台的《春节联欢晚会》还是在《朗读者》节目中的语言表达都体现出来一种内在的自信和自豪感。这是因为拥有诗书知识和高尚气质的人在言行间透露出一种自信,这种自信也是一种美学上的迷人特征。一位优秀的播音员或者是主持人,都必须具备文学知识和气质修养,以及对自信和自豪感的表达,这种表达展现了一种内在充实和高尚的美学特征,使得语言表达更加丰富而深沉具有人文之美。

然而,对于 AI 主播而言,“人文美”在其语言表达中存在一些限制。尽管一些 AI 系统能够模拟人类语

音和情感,但它们缺乏真人主播深度理解和情感共鸣的能力。文字对 AI 主播来说仅仅是一种简单的识别功能,缺乏人类的“思考”和对文字的“感受”能力,这在审美活动中是至关重要的。正如前文所述,理想的审美状态需要主客体的统一,而 AI 缺乏这种主观能动性,无法完成自主阅读和真实经历百态。AI 缺少了人生的真实经历和积累,因此无法像真人主播一样完成审美活动。这也意味着它无法达到“腹有诗书气自华”的表达水平,因为缺少了与人类相似的深度文学知识和高尚情感氛围。在缺乏真实人生体验的情况下, AI 主播无法通过语言表达展现出丰富的人文价值观和情感共鸣,这成为其在人文美领域受限的根本原因。因此,尽管 AI 技术不断发展,但在人文美的层面上,它仍然无法取代真人主播所具有的深刻理解、情感共鸣和个体独特性。

真人主播的语言表达艺术美在于他们的表达方式、语言运用和声音的音调变化等方面。通过巧妙运用这些艺术手段,他们能够更好地激发观众的审美感知,使得语言传递更具有艺术性。播音员主持人的语言表达技巧主要体现在外部技巧方面,包括停连、重音、语气和节奏。这些看似独立的要素实际上构成了一个有机的整体系统。例如,停连的位置与重音有关,同时也与节奏和语气紧密相连;重音强调的方式也与语气和节奏密切相关。这种复杂的系统使得真人主播在语言表达中能够展现出更高水平的艺术美。相比之下, AI 主播的艺术美受限于算法和训练数据,可能在模仿自然语言和情感表达方面存在一些局限。虽然技术不断进步,前文中提到的 AI 主播“海蓝”的新闻评论播音展现了相当高水准,无论是语言表达的“内外部”技巧还是语音语调都表现出了很高的水平,但在一些细致和高度个性化的艺术表达方面仍有改进的空间。虽然 AI 主播在一些方面取得了显著的进展,但在真人主播所展现的高度个性化和复杂的艺术表达方面,仍需不断努力和改进。这也强调了真人主播在艺术美领域的独特价值和无法替代的地位。

真人主播能够在语言活动中展现更多的个性、创造力和独特的艺术精神。他们可以灵活运用语言和表达方式,使得节目更具有创造性和独创性。AI 主播相对缺乏真人主播的创造性和灵活性,他们的语言合成也依赖于真人主播的语言数据库。但随着技术的发展,未来可能会有更具创造性的 AI 语音模型。

根据一项研究显示,观众对 AI 主播和真人主播的语音审美存在差异。一项研究采用控制实验法,将相同内容的新闻由 AI 主播和真人主播分别播报,并对观众的记忆效果进行比较。结果发现, AI 主播的播报效果在观众的新闻记忆方面弱于真人主播。此外,设置社会化线索的主播(如真人主播)比未设置社会化线索的主播(如 AI 主播)更能增强观众的新闻记忆效果。^[7]同时,研究还让观众分别观看了 AI 主播和真人主播播报相同新闻内容的视频。结果显示,观众对真人主播的播报有更好的记忆效果和接受度。这表明,尽管 AI 主播技术在不断进步,但在情感表达、社会化线索等方面,真人主播仍具有优势。^[8]这表明,审美涉及语音和人与人之间的情感互动。根据朱光潜的《谈美》,“美不单存在于对象之中,而是在观赏者的心灵活动中得到实现。” AI 主播缺乏真实的情感体验,因此缺乏审美价值,而真人主播的表达能够激发观众的情感共鸣,使信息传递更加生动。例如,在《朗读者》节目中,董卿的语言表达具有韵律感和层次变化,使观众沉浸其中。然而, AI 主播的语言表达,即使是准确流畅的,也难以营造相同的审美体验。

从语言的艺术性来看,真人主播在语音的旋律美和节奏美上具有明显优势。中国传统语言美学强调音韵和谐、抑扬顿挫,语言不仅是信息的载体,更是一种艺术表现形式。^[8]真人主播能够根据内容调整节奏、重音和语调,使语言富有层次和变化,增强其艺术感染力。例如,央视新闻主播康辉的播报节奏严谨、语调富有变化,使新闻更具表现力。而 AI 主播虽然可以模仿自然语音,但在节奏控制和情感递进方面仍存在模式化的问题,难以达到真人主播的艺术表现力。^[9]

研究表明,观众对真人主播的播报有更好的接受度和记忆效果。^[8]这可能与拟人化效应有关,即人类更容易对具有情感、个性和互动能力的对象产生亲近感。此外,沉浸理论(Immersion Theory)也解释了真人主播更容易让观众沉浸在语言表达中,而 AI 主播由于情感层次不足,可能会导致观众产生“审美疲劳”^[10]。例如,在《新闻联播》中,主持人的表达方式结合了节奏变化、语调高低、语气变化等手段,使语言更加生动,增强观众的审美体验。而 AI 主播由于语言变化不丰富且不能根据文本内容进行创作,长时间聆听可能会使观众产生单调感,降低观看体验。

总体而言,真人主播在人文美、艺术美和艺术精神方面具有更大的优势,因为他们能够借助个人经验、情感表达和创造性思维,使得语言活动更具深度和广度。然而,随着 AI 技术的不断发展, AI 主播在某些场景下可能会越来越接近真人主播的表现水平。

三、跨文化视角下不同语言的审美范式

语言作为文化的载体,其美学特性受到文化背景的影响。在人工智能技术迅猛发展的今天,人工智能主播已成为一种新兴的传播方式,全球范围内被广泛使用。然而,观众对 AI 主播的接受度和审美评价在不同语言和文化背景下有所不同。英语、日语等语言体系对语音艺术美的理解,以及人工智能主播在不同文化环境中的表现差异,分析文化背景对观众接受度的影响。

英语作为强调重音和节奏的语言,其审美特征主要集中在重音和节奏上。在英语中,重音的位置对于表达意义和情感至关重要。人们认为,准确的重音和自然的节奏是英语语音美的关键。研究表明,英语的节奏关注重轻音节的交替,这种重音的存在对英语的语音美起着重要作用。^[11]此外,重音的正确使用对于传达英语

的语义和情感具有重要作用。^[12]

此外,英语强调音质的清晰度和音调的变化,以传达不同的情感和态度。例如,语调的升降可以表示疑问、肯定或讽刺等多种语气。研究指出,英语的音高变化对语调的表达具有重要影响。^[13]同时,音质的清晰度对于英语语音的审美也至关重要。^[14]

相反,汉语是声调语言,其审美特征主要体现在声调的使用上。汉语的每个音节都有特定的声调,这些声调对表达词汇意义和情感至关重要。研究表明,汉语的节奏依赖于不同声调的组合,这与英语的重音节奏形成鲜明对比。^[15]此外,汉语的声调变化对于表达不同的词义和情感具有重要作用。

英语的音调变化用于表达不同的语气和情感,尽管汉语以声调著称。例如,在英语中,升调通常用于疑问句和降调用于命令句或陈述句。汉语重音不仅用于强调信息,还可以表达说话者的情感和态度。例如,可以通过加重词的读音表达强调、惊讶或不满等情绪。

然后两者之间也存在着一些显著性的差异。在英语中,重音具有词汇辨义功能,即通过改变重音的位置,可以改变单词的词性或意义。例如,'record'作为名词时重音在第一音节,作为动词时重音在第二音节。而在汉语中,重音主要用于语用功能,如强调或对比,但不会改变词汇的基本意义。但是,汉语的声调是词汇意义的本质部分,不同的声调可以区分词汇意义。例如,'妈'(第一声)表示'母亲',而'骂'(第四声)表示'责骂'。相比之下,英语的音调变化主要用于表达句法功能或说话者的态度,但不会改变单词的基本意义。

日语的语音美学特征深受其语言结构和文化背景的影响,主要体现在音节均衡、声调和音高的变化以及丰富的敬语体系。首先,日语以开音节为主,音节结构相对简单,使得发音流畅自然,这种均衡性被认为是一种语言美感的重要体现。^[16]相比于英语和汉语较为复杂的音节结构,日语的音节相对规则,使得整体语音听起来更加柔和、有节奏感。此外,日语的声调变化较为微妙,音高的升降不仅影响词义,还会影响表达的情感和态度。其次,日语的声调变化微妙,音高的升降不仅影响词义,还影响表达的情感和态度,正确的音高控制被视为优美表达的标准。^[17]

除了音节和音高,日语的敬语体系也是其语音美学的一大特色。日语的表达方式受到社会等级、场合以及人际关系的影响,不同的敬语,如尊敬语、谦让语、郑重语等在语音表现上也有所不同。研究表明,敬语表达通常伴随着更加平缓的语调、较长的音节停顿和更高的声调,以表现出礼貌和尊敬的态度。^[17]这种礼貌用语的得体性不仅反映了日本文化中对人际关系的重视,也构成了日语语音美学的重要部分。因此,日语的语音美学不仅体现在语言的结构层面,也与日本社会文化紧密相关,使其在不同语言体系中展现出独特的韵律和审美特征。

文化背景影响语言的美学特性,因此人工智能主播在不同语言体系中的表现和观众接受度也不同。英语、汉语和日语是三大主要语言体系,每个语言的语音美学都不同。

英语强调重音和节奏。重音的正确使用影响语音的含义和情感表达。在英语语音表达中,音质的清晰度和音调的变化至关重要,因为它们可以传递不同的语气和态度。相反,汉语是一种声调语言,其语音美学主要体现在使用声调。不同声调组合会影响词汇意义和情感表达,而汉语重音更多用于语用功能,如强调或对比。

音节结构、声调变化和敬语体系影响了日语的语音美学。日语以开音节为主,使语言整体听起来流畅,同时微妙的音高变化影响表达的情感和态度。此外,敬语体系在语音表现上也有所体现:礼貌用语通常伴随着更平缓的语调和较长的音节停顿,这表明日本文化重视人际关系。

总而言之,英语、汉语和日语的语音美学差异会影响人工智能主播的表现以及他们在不同的文化背景下接受度。了解这些差异有助于改进人工智能语音技术,使其更适合不同语言用户的审美需求。

四、结语

本研究研究了人工智能主播与真人主播在语言审美方面的差异,并通过跨文化视角分析了不同语言体系的审美特征。研究表明,真人主播在语音美学方面具有天然优势,因为他们有丰富的情感表达、文化背景和个性化特征。人工智能主播的人文价值和艺术性仍然难以完全替代真人主播。此外,英语、汉语和日语的语音美学差异影响了AI主播在不同文化背景下接受度。未来,随着人工智能技术的发展,如何增强AI主播的情感表达与文化适应性,将成为提升其语言美学价值的重要方向。

参考文献:

- [1] 赵玉明. 中国广播电视通史[M]. 北京: 中国广播电视出版社, 2006: 196.
- [2] 祝捷, 钟威虎. 中华有声语言审美范畴溯源——基于一种媒介考古的分析[J]. 现代传播(中国传媒大学学报), 2023, 45(07): 93-101.
- [3] 耿艳丽. 谈有声语言表达提升美学价值的途径[J]. 开封大学学报, 2008(2): 2. DOI: 10.3969/j.issn.1008-343X.2008.02.019.

- [4] 张健. 语音的美学: 刘勰的声律论[J]. 北京大学学报(哲学社会科学版), 2024, 61(1): 96-108.
- [5] 范藻. 在“嗟叹”与“咏歌”之间 ——有声语言如何成为“艺术”的美学解析[J]. 中外艺术研究, 2021, (3): 45-53.
- [6] 曾志华, 卢彬. 中国有声读物播读评价体系构建研究[J]. 现代传播, 2023, 45(6): 123-130.
- [7] 王宇. 声乐艺术中的音色审美探析[J]. 公务员期刊, 2023, (5): 78-82.
- [8] 刘娜, 黎樟浩, 吴晔. AI 主播与真人主播的播报效果研究[J]. 青年记者, 2023(06): 58-61. DOI: 10.15997/j.cnki.qnjz.2023.06.019.
- [9] 冯子萱, 舒永芳. AI 主播 vs 真人主播: 一项基于受众感知的实验研究[J]. 中国新闻评论, 2024, 5(3): 16-25. DOI: 10.35534/cnr.0503002.
- [10] 张莎莎. 人工智能时代 AI 主播的伦理审视和风险规避[J]. 当代电视, 2022(5): 84-87.
- [11] 王珊. 人工智能语音在新闻播报中的应用与挑战[J]. 现代传媒, 2023, 41(4): 45-52.
- [12] 王芳. 英语语音的重音节奏特征及其在语言表达中的作用[J]. 外语教学与研究, 2023, 55(4): 78-85.
- [13] 李明. 英汉语音对比研究: 从重音到声调的语言审美差异[J]. 语言研究, 2022, 40(6): 102-110.
- [14] 张伟. 英语语音中的音质与音调变化及其情感表达[J]. 外国语, 2023, 46(2): 59-66.
- [15] 赵欣. 汉语声调的音高变化及其在语言美学中的作用[J]. 语言学论丛, 2024, 37(1): 88-95.
- [16] 陈晨. 跨语言语音学视角下的英汉节奏模式研究[J]. 语音学报, 2023, 12(3): 47-55.
- [17] 李双. 日语中的敬语研究[D]. 吉林: 吉林大学, 2011.
- [18] 崔亚蕾. 日语敬语在表现说话人“品位”方面的作用研究——以日剧中的人物对话为例[J]. 日本问题研究, 2016, 30(1): 71-80. DOI: 10.14156/j.cnki.rbwtj.2016.01.008.

AI anchors versus live anchors: an exploration of a linguistic aesthetic perspective

Niu Yongbo

(¹ Universiti Putra Malaysia, Selangor Malaysia 43400)

Abstract: This study analyzes the performance of AI anchors and real anchors from the perspective of linguistic aesthetics, and investigates the phonetic aesthetic characteristics of different language systems (English, Chinese and Japanese). The study shows that real anchors have more advantages in emotional expression, voice rhythm and humanistic values, while AI anchors still have limitations in aesthetics despite continuous improvement through voice synthesis technology. Real anchors can create more cultural depth of expression through sound, while AI anchors are difficult to reproduce this aesthetic experience, because language aesthetics in the Chinese language is not only expressed in the tone and rhythm, but also emphasizes the "mood", "sentiment" and "rhyme". The reason is that linguistic aesthetics in Chinese is not only expressed in tone and rhythm, but also emphasizes poetic features such as "mood", "sentiment" and "flavor". In addition, the aesthetic standards of different language systems affect the audience's acceptance of AI anchors. For example, English emphasizes stress and rhythm, while Japanese emphasizes syllabic balance and rhythmic features of the honorific system. The study suggests that in order to better meet the aesthetic needs of users of different languages, the future development of AI anchors should incorporate the aesthetic characteristics of the Chinese language.

Keywords: ai anchors, live anchors, linguistic aesthetics, phonetic aesthetics, cross-cultural comparisons