

# 从口语与书面语角度 看GPT的自然语言处理弊端

◎王 宁 张世传

(河南理工大学文法学院 河南 焦作 454000)

**【摘要】**人工智能语言存在表达生硬、分析固化及话轮转换滞后等问题,原因在于其自然语言处理机制对自然语言理解不足,过度依赖书面语料进行分析。在不改变语料类型的前提下,关键在于理解并优化书面语与口语之间的转换。本文结合语言学与人工智能理论,探讨了文字和语言的关系,分析了人工智能语言与人类语言的差异,并给出相应建议,即人工智能应重视语境分析与应用,基于数据库解析的“最优解”未必适应灵活的话语空间,同时,对话中上下文联系也应得到关注。

**【关键词】**语言;文字;人工智能;ChatGPT

**【中图分类号】**H021

**【文献标识码】**A

**【文章编号】**2096-8264(2024)22-0134-03

**【DOI】**10.20024/j.cnki.CN42-1911/I.2024.22.041

## 一、引言

人工智能的语言生成能力受限多种因素,如语料规模和类型。书面语因其易总结应用,被AI广泛采用,而口语则因难以模仿,较少被采用。文字作为书面语的载体,在口语表达中有局限。当前,对如何利用文字表达语言的理解已深化。本文探讨将“文字生动表示语言”的理论应用于“AI语言生成问题”的可行性。

语言障碍已成为21世纪全球化发展的制约因素。因此,让计算机理解自然语言,实现自动化处理和个性化服务,是国际前沿研究,也是全球共同目标。<sup>[1]</sup>在ChatGPT广泛应用下,袁毓林(2023)提出语言学研究应转向以数据/计算密集型为主要构造对象的第四、第五范式,为语言学发展指明方向。

有学者认为,人工智能与人类有本质区别,前者学习,后者创造。然而,这与创造AI的初衷不符。尽管ChatGPT改善了以往AI语言的产力、话轮死板、用词单一、句法简单等问题,但如何使其更接近人类语言仍待解决。原因在于语言和文字的本质。

## 二、书面语系统与口语系统的区别

### (一)书面语系统

书面语是文字书写的产物,用于记录语言。文字的出现就是为了记录语言,于省吾提到,文字起源于原始社会,初衷是简单记录事物,防止遗忘。随着阶级形成和国家出现,生产力提升,人事繁杂,文字逐渐用于记录语言。<sup>[2]</sup>文字从语言独立成符号后,相较于语言,延长了思维时间,并经历了简化过程,导致书面语与口语差异。

人工智能语言代码的输入与输出的机制构想是摹仿人

脑对语言的输入与输出进行架构的,人所使用的是人脑,人工智能所使用的是自然语言处理机制,这是言语产出是否能符合语言的根源。文字是表达于书面的通过大脑加工后的语言代码,包括表音文字、表意文字和符号等类型,这些存在于视觉、听觉等感知下的文字与语言之间的变换是人工智能需要关注的重要一环。我们要清楚地认识到,人工智能的代码输入和输出机制与人类的语言解析和产出并不完全一致(虽然我们是这样预想的),而是与文字的表述模式有一定的相关性。因此,当回答“人工智能所发出的语言如何才能更像人的语言”这个问题时,关键之处在于如何使得文字的表述更接近以至于成为“真实的”语言。

### (二)口语系统

口语是语言的表达形式,人类依赖语言进行思考和交流。言语是心理现象,是人类运用语言的行为。起初,言语无固定规则,但随着社会共同语言的出现,受到制约,但仍保持自由生成性,为语言更新提供动力。大脑是言语接收与发出的中枢<sup>[3]</sup>,语言信息处理依赖词汇和语法。

在普通语言学中,存在两个基本的观点:第一,语音是语言的本质,文字是语言表达的辅助工具。第二,语言的存在先于语法,语法是对语言规律的滞后总结与归纳<sup>[4]</sup>。对于第一种观点,文字是大脑组织语言思维后的视觉映射,传递视觉图像和字义;语言则是思维组织输出为听觉的声响,传递语音序列和语义。这种差异导致语言受语境、心理等因素影响,文字受书写、表达方式等影响。

第二种观点中,“滞后”包括历时和共时的双向滞后。语法对新生语言的控制力逐渐减弱,如古代汉语语法不适用于现代汉语。严格遵循语法规律创造语言,忽略语境、用

词和语气,会导致听众难以接受,如人工智能回答问题时过于死板。

### (三)人工智能语言分析

人工智能语言追求准确性,因此其表达更为正式,易于识别与人类语言的差异。人类语言反映思维,听话者可判断说话者状况;而人工智能语言依赖于程序等,仅反映不完美的语法系统,难以实现人机语言融合。如某人言:“我该得的那个,你付都多了,我该得的应该得,这里拿多掉了。”虽语法错误,但人类能理解。这是因为非规范语言具有真实感,听话者可根据说话者身份、心理等进行判断。而人工智能缺乏这种能力。此外,这种语言难以文字表达,因文字需更多大脑加工。杨立昆(2022)认为,语言仅承载人类知识一小部分,大部分知识和动物知识非语言(非象征符号的)。<sup>[5]</sup>但人工智能语言具有强制正确性,输入输出固定,表达受限。能否将人类语言转移至人工智能?可从汉语独特性入手。

## 三、如何用书面语料表现口语

当模仿物缺少必要现象,即失去了“灵魂”。常说汉语缺少形态变化,语序和虚词至关重要。人工智能的自然语言处理不仅需要理解语序、虚词带来的差异,更要理解在某些情况下语序与虚词的改变在意义表达不会产生任何影响这一人类语言理解特征。

### (一)语序

语序是一种语法手段,但却不一定会改变意义的表达。例如下面的两个句子:

(1)你淋着雨没有/雨淋着你没有

(2)肉末夹烧饼/烧饼夹肉末

例(1)同义,例(2)异义。例(1)前句,在实际语言环境中的使用更倾向于询问“你”的感受,身上是否被雨打湿;后句则更倾向于对“雨”效果的询问,雨有没有被你躲开。例(2)前句给人以肉包裹着烧饼,肉多的心理体验;后句则是烧饼为主,夹缝中有肉末,让人觉得肉少。

问ChatGPT两者区别,AI回答:两种表达方式主语和语序不同,但含义相同,并以系统中所保存的“着”表示“现在”的依据作出这两句话是在询问“正在”发生的事情。实际中,“着”在此句中仅起结构助词作用,对语义无影响。

在进一步对ChatGPT进行询问的回答中更为明显:

问:这两种表达方式都只能表示“正在”吗?

答:对的,这两种表达方式通常用来询问对方当前的状态或情况。它们都含有“正在”的意思,即询问是否在某个特定的时间点或时间段内正在被雨淋着。请注意,这些表达方式强调的是持续性的动作或状态,而不是雨是否曾经淋过你或你曾经淋过雨。

人工智能仅仅以“着”表示现在,不表示过去或未来’这一程序代码为依据,对于语境与事实置若罔闻,在实际语言环境中,若是下雨后问对方是否淋雨了,也会以这种方式进行询问。询问发出的时间是现在,对象也是现在,而问

询的事物是过去。

ChatGPT对于例(2)的回答更是让人啼笑皆非:

问:“肉末夹烧饼”和“烧饼夹肉末”有什么区别?

答:二者实质相同,仅表述顺序不同。通常可互换,无明显区别。前者指切开烧饼,夹入肉末,再合上;后者是将肉末置于烧饼上,对折或卷起,使肉末被包裹。简言之,二者皆描述同一种食物组合,顺序差异而已。

首先这段话的逻辑就出现了语义矛盾。首句末句皆言明是同样的食物,同样的组合,只是顺序不同,但描述时却用不同的制作工艺技巧进行叙述,说前句是“夹有肉末的烧饼”,是半包裹状态,后句是“肉末被烧饼包裹”,是全包裹状态,因此得到的食物并不同。事实上,在实际语言环境中这两种表达描述的是同一种食物,只是给人的感觉不同。前一种表述给人以肉多的心理体会,即肉包裹着烧饼;后句给人以肉少的感觉,即烧饼包裹着肉。食材的前后顺序给人以食物数量多少的心理体会,就如配料表是以含量多少进行排序,排名靠前则含量更多,靠后则较少。

总之,对于语序带来的影响应当被准确对待,细致划分,了解到这种表述的社会性是一个很重要的方面。

### (二)虚词

虚词在现代汉语中很重要,但有时删减不影响句意。如:

(3)买不起别买/要是买不起就别买

(4)洗干净净收着/洗得干干净净收着

例(3)和例(4)前后句意思相同,虚词的省略并无语义上的改变,尤其在口语中更常见。两句在自然口语对话中并没有显著的区别,GPT却过于敏感:

问:“买不起别买”和“要是买不起就别买”有区别吗?

答:两者微小差别。“买不起别买”简洁,强调行为建议,若买不起,别买。未提条件。“要是买不起就别买”较复杂,先给条件(买不起),再建议不买。语气较灵活,可寻求替代品或改变消费习惯。核心意思相同,皆建议经济不允许时不购买。第二句提供条件,给人更多思考空间。

人工智能难以理解“意合法”,虽前句省略了“要是……就”,但当我们对其进行解读时会自动加上这种关联,口语中转换无碍。但问人工智能,由于后句增加了关联结构,其解读就必然会有所增加以解读关联词的出现,这种解读是毫无必要的,更是不符合人之解读的。

例(4)后句的“得”是句子中表示结构关系的虚词,但在口语中也可以不说出来,人工智能却把注意力放在了“得”上:

问:“洗干净净收着”和“洗得干干净净收着”有区别吗?

答:“洗干净净收着”和“洗得干干净净收着”的区别在于前者强调结果,后者强调过程和结果。两者均表示洗净后收起,但“洗干净净收着”更简洁,注重结果,而

“洗得干干净净收着”更详细,强调过程和结果。根据场景和喜好选择即可。

回答单纯从虚词“得”入手进行分析,而汉语语法意合性显著,非典型规则不易把握。“得”的语法功能在本句中已经虚化,而回答反复从语法、语用角度解释“洗”在句子中的语法作用和语用表现,是语言规则的刻板表述,是被输入的语法代码的表述。这种解读并不适用于现实生活,在口语中人们经常使用省略“的、地、得”这种结构助词的表述方式,对于语义和表达并无影响。

总之,GPT的自然语言处理并没有意合的“程序”。在自然口语对话中,也经常有不理解对方意义的情况出现,但这种情况是少数,即使出现了,也能够通过多次的对话去解决这个“误会”。GPT不仅难以持续性对话,更是以高频词“误解”显现,这是不合适的。若将词序和虚词作为一种自然语言的特殊现象对人工智能进行培养,或许会出现对同一问题的不同形式而意义相同的回答,相对于同样的死板回答来说则更显得“人性化”。

### (三)语境表现

从根本上说,当前系统只能构建语言骨架,无法充分展示语言丰富性。婴儿表现的不流畅、错误的发音却可以被识别,这是因为其独特的音调、音高、音色及气力不足导致的断句、重音。言语障碍者虽然无法通过语言被他人理解,但却可以通过情绪感知、不断询问猜测到想要表达的意思,这是因为失语症患者仍然具有“人”的语言能力,通过声带震动和肢体语言表达。这些体现了语言魅力和“人”的语言特点,即使不使用规范的语言,也能表达意义,这与认知和文化共鸣有关,是共情和意义相通的关键。

总之,这些都无法通过大数据或基本方程获得。真正“语言”的真实构建不仅需要结构主义语言学对操作程序的讲究、转换成语法对结构模式及其背后心智过程的探索,以及功能语言学基于认知科学的研究,还需要经验观察,立足于“言语”事实,不能只关注局部,而忽视整体,只有两者相结合,才能实现人工智能自然语言的正确发展。

深度神经网络受生物学启发,瞳孔摄入像素,大脑皮层细胞发现边缘和方向,判断物体形状和类别。这种程序导致人工智能难以真正理解运用语言,生理判断虽似人类思维,实为简单思维模式,这不仅包括“看山不是山,看水不是水”这样的高度抽象思维逻辑,还包括无需思考的“习惯”<sup>①</sup>逻辑。

语言学有一个经典歧义问题,即:

#### (5)咬死了猎人的狗

说话人身处于“猎人”的尸体旁,眼睛观测到了一幅“狗”咬死“猎人”的场景,经过大脑的思维后,从口中对他人以语言的形式指出“这条狗是咬死了猎人的那条”这件事,接收信息者以听觉形式接收的信息在大脑中思索不会感觉有歧义或不解;但如果说话人以文字的形式进行书写,

接收信息者以视觉形式接收的信息经过大脑分析后却会得到两种不同的理解:狗被咬死了/猎人被咬死了,以致产生不知所云与不明其意的困惑感。而人工智能是以文字为基数,通过视觉或听觉方式发出信息的一种传输系统,其使用文字符号来表情达意,即使是以音响形式发出,也是以文字的思维逻辑原封不动地表述,这也就是为什么人工智能有时被戏称为“人工智障”,总是与人类的语言有着显而易见的差别感。

当今人工智能以深度学习为主要学习方法,这种学习确实会使机器获取大量语言范式并生成语言蓝本,但这种学习是由无数的人类语言组合而成的类似“共同语”的系统,其选择的解答也通常为“最优解”。但在不同语言环境中,这种所谓的“最优解”并非适合人类的思维感情。因此,对于语境的认知势必要成为人工智能在对话中语言选择的主要影响因素。

### 四、结语

书面语系统与口语系统存在现实差异,我们可以通过对语序、虚词等汉语自身特点在书面语、口语上的表现来解决这些差异问题。这就需要我们更加深入地对自然口语对话进行研究以及对自然语言处理的研究,找到更多的具有汉语特征的口语语料,与书面语语料进行对比分析,以实现一种数据密集型的研究范式,为语言学及人工智能语言提供进一步发展的空间。

### 注释:

①例如:中国人见面后常问“吃了吗”属于文化逻辑,熟练使用键盘的人可以不看键盘快速打字属于空间逻辑,人在看见美好事物的时候眼球会随物而动属于生理逻辑等。

### 参考文献:

- [1]中国科学技术协会主编,中国人工智能学会编著.人工智能学科路线图[M].北京:中国科学技术出版社,2022:57-58.
- [2]于省吾.关于古文字研究的若干问题[J].文物,1973,(02):32-35.
- [3]程琪龙.认知语言学概论——语言的神经认知基础[M].北京:外语教学与研究出版社,1999:37-38.
- [4]索绪尔.普通语言学教程[M].裴文译.上海:教育出版社,2002.
- [5]杨立昆.语言的有限性决定了AI永远无法比肩人类智能[DB/OL].[https://mp.weixin.qq.com/s/?\\_biz](https://mp.weixin.qq.com/s/?_biz),2022-8-26.

### 作者简介:

王宁,女,河南焦作人,硕士研究生,研究方向:汉语国际教育。

张世传,男,河南郑州人,硕士研究生,研究方向:汉语言文字学。