

诗歌平仄与语义加工的认知神经机制*

张晶晶¹ 时颖¹ 邓善文¹ 李佳斌² 陈庆荣^{1,3}

(¹ 南京师范大学心理学院; ² 南京师范大学文学院; ³ 江苏省高校哲学社会科学实验室
——南京师范大学青少年教育与智能支持实验室, 南京 210097)

摘要 音乐是诗的生命。对于中国古代诗歌而言,平仄模式是诗歌音乐性的一种典型表达方式。本研究选取五言律诗作为实验材料,同时操纵诗歌的平仄模式和语义特征,采用 EEG 技术考察中国古代诗歌阅读过程中平仄感知和诗意理解的认知神经机制。行为结果显示,平仄的主效应以及平仄与语义之间的交互作用显著。ERP 结果进一步发现,诗歌阅读过程中存在平仄效应,表现为早期的 P200 效应,以及之后的 N400 效应和晚期的 LPC 效应;不仅如此,平仄感知和诗意理解在中晚期相互调控和整合,表现为在 N400 和 LPC 上存在交互作用。同时,为了揭示神经活动与平仄和语义特征的关系,对 EEG 数据进行了解码分析,发现基于 EEG 数据训练得到的神经网络能够对不同类型诗句进行有效分类。实验结果一致表明,人们在诗歌阅读过程中期待字音本身的和谐以及音与义的协调,这种期待影响着诗歌早期的语音表征和随后的音义整合。

关键词 诗歌, 音乐, 平仄, 语义, EEG

分类号 B842

1 引言

音乐是诗的生命(朱光潜, 2012)。在不同文化中,诗歌与音乐在起源时都是合二为一的艺术形式,诗常可歌,歌常伴乐(Aristotle, 1968; Nietzsche, 2012)。随着艺术分化,诗歌朝着文字意义的方向发展,转入有词而无调时期。在词调并立之时,诗的音乐在调上见出,词既离调之后,诗的音乐要在文字本身见出。以中国古代格律诗为例,为了体现语言的韵律,格律诗在字数、声调、韵脚等方面都有许多讲究(王力, 2012)。

汉语作为一种声调语言,平仄规则是中国古代格律诗独特的一种韵律特征。古代汉语有平声、上声、去声、入声四个声调,所谓“平”指的是平声(包括如今的阴平、阳平),“仄”指的是上去入三声。五言和七言格律诗平仄规则,指的是单句内平仄交替(如“仄仄平平仄”),联句间平仄相对(如出句“平平仄仄平”对仗入句“仄仄平平仄”);此外,后联出

句第二字与前联对句第二字平仄相粘,避免声调的单调重复(王力, 2012)。这两类声调有规律的更替起伏,让无音乐的诗歌听起来产生“吟唱”的感觉,从而赋予诗歌内在的“音乐性”。例如,“人间四月芳菲尽,山寺桃花始盛开”,如果把第二句改为“山寺桃花始盛放”,意义虽然没有明显变动,却立刻觉得不顺口不顺耳了。此外,四声还存在音质上的差别,即平仄的功用还在于构成字音本身的和谐以及字音与意义之间的协调,例如形容马跑时多用铿锵疾促的字音,形容水流宜用轻快圆滑的字音(朱光潜, 2012)。韩愈《听颖师弹琴歌》的前四句,“昵呢儿女语,恩怨相尔汝;划然变轩昂,勇士赴敌场”,即能看出四声调质能够产生和谐的印象,并且使得诗的音义并行。

20 世纪 80 年代兴起的认知诗学,开始关注人们如何感知诗歌中的韵律特征(张叶鸿, 2015; Tsur, 1983)。Jakobson (1960)在对西语诗歌考察的基础上提出了诗意性功能理论(poetic function theory),即

收稿日期: 2025-01-25

* 国家社会科学基金重大项目(21&ZD288)。

通信作者: 陈庆荣, E-mail: jscqr80@sina.com

诗歌文本使得读者有意识地知觉诗歌语言结构所具有的文本语言模式,特别是其中的韵律特征。已有诗歌阅读的实验研究主要围绕印欧语诗歌的典型韵律特征如重音、节奏展开。这些研究发现,阅读诗歌相比阅读散文,除了激活阅读相关区域,的确激活了更多与处理节奏和音乐相关的脑区,比如左侧和右侧颞上回(Zeman et al., 2013)。并且,随着诗歌难度增加,人们会分配更多的注意资源用来加工诗歌的形式特征,包括某些韵律元素(Yaron, 2002)。鉴于人们对于诗歌韵律特征的敏感性,诗歌韵律异常(轻重模式异常和押韵违反)会造成明显的阅读干扰,引起对于局部文本的重新阅读(Beck & Konieczny, 2020)。

一些研究者结合汉语特点,考察了汉语诗歌中停顿、押韵等韵律特征加工。停顿、押韵与平仄共同构成了古代汉语诗歌的典型韵律特征(王力, 2012; 朱光潜, 2012)。相较而言,平仄是诗歌全局性的声调组织规则,押韵是句末韵部呼应的回响机制,停顿在朗读层面形成语流切分与节拍分组。Li 和 Yang (2010)以七言绝句为实验材料,考察了古代汉语诗歌中韵律层级边界的加工,结果发现不同层级边界均能诱发中止正漂移成分 CPS (closure positive shift),表明被试能够感知古代汉语诗歌中不同位置的停顿。针对押韵的眼动研究(陈庆荣, 杨亦鸣, 2017)发现押韵违反影响早期眼动指标(如首次注视时间、凝视时间)和晚期眼动指标(如总注视时间、总注视次数),电生理研究(Chen et al., 2016)也发现,押韵违反诗句相比押韵合理诗句诱发波幅更大的 P200 和 N400。

张政华等人(2020)以七言绝句为实验材料,同时关注了诗歌中押韵与声调的加工,结果发现声调和押韵违反体现在 P200、N400 和 LPC 成分上,且受到被试音乐经验(张政华等, 2020)和实验任务(张政华等, 2021)的调节。这两项研究虽然关注了汉语诗歌中声调的加工,但是实验材料没有选择严格符合平仄规则的诗句,也没有考虑诗句中古音和今音的变化,因此并不是专门对平仄规则的考察。其次,在操纵诗句押韵和声调的同时,没有控制诗句的语义合理性程度,导致不同实验条件的诗句除了韵律模式不同,还可能存在着意义的差别,从而难以厘清实验效应究竟是什么因素导致的。

除了诗歌韵律特征,认知诗学同样关注诗歌意义的理解。根据特殊体裁假说(genre-specific hypothesis) (Hanauer, 1998)和传统文学理论(Culler,

1975),文本的阅读和欣赏模式明显受到体裁特征的影响,使得不同外在特征的文本体裁存在明显的加工差异。对于诗歌而言,读者会将注意投向特定的文本语言模式,特别是诗歌的韵律特征,使得阅读速度更慢,对语义的理解难度更高(Hanauer, 1998)。此外, Hanauer (1998)认为“声音的相似性常被视作意义上的相似性”,而中国古代诗歌的韵律特征使得诗歌在声音上具有明显的相似性。基于这两种观点,可见诗歌中的语义加工会受到韵律因素的调节。

韵律与语义关系的理论争论主要围绕加工进程和加工模式的独立性与互动性展开。一方面,传统观点认为语义加工在理解中占据主导,韵律只是辅助性特征。例如,模块化理论(the modularity of mind)(Fodor, 1983)认为,韵律加工是一个相对独立的封装模块,其输出被传递至更高阶的中央系统,以辅助语义构建。因此,韵律加工只发生在早期的感知阶段,与语义之间的互动是单向而有限的,语义信息不会影响韵律加工。另一方面,三阶段神经认知模型(Friederici, 2002, 2012)和互动激活模型(McClelland & Rumelhart, 1981)则主张韵律加工并非模块化,而是与语义加工存在实时的双向互动。这些模型强调,韵律与语义是并行加工,并非仅限于早期阶段;韵律线索能动态调节语义加工,反之亦然。

现有实证研究进一步检验了这些理论观点。在印欧语系(如英语或德语)中,研究主要操纵重音位置(Rothermich et al., 2012; Rothermich & Kotz, 2013)和音节时长(Magne et al., 2007),而在汉语中,则针对短语节奏(如“种蒜”[1+1] vs. “种植蒜”[2+1]节奏; Luo & Zhou, 2010)或句子节奏(如规律二字词背景 vs. 不规则模式; Li et al., 2019)进行操控。这些研究发现,韵律违反往往诱发 N100(Li et al., 2019)、N400(Luo & Zhou, 2010; Magne et al., 2007)或者晚期正成分如 LPC (Luo & Zhou, 2010; Magne et al., 2007; Rothermich et al., 2012),并且韵律和语义在 N400 和 LPC 成分上存在相互影响(Li et al., 2019; Luo & Zhou, 2010; Rothermich et al., 2012)。其中,早期 ERP 效应主要反映对韵律违反的识别与检测,而晚期正成分则更多涉及韵律结构的整合与再分析。

尽管现有研究结果总体倾向于支持三阶段神经认知模型和互动激活模型,但是仍然存在一些局限性。首先,重音、音节时长、短语和句子节奏等

主要涉及声音的强度和时长,是“时间导向”的韵律特征(如形成节拍和分组),“音高导向”的韵律特征(如声调和语调)需要进一步科学研究(Hirst & Di Cristo, 1998; Ladd, 2008)。其次,对于词汇层面的重音或音节时长操纵,往往与词汇语义高度耦合,韵律变化可能使词汇听起来不自然甚至不像真实词语,从而直接或者间接干扰词汇识别和语义整合过程(Rothermich et al., 2012)。此外,对于句子节奏规律性的操纵,比如印欧语系中的轻重模式(Rothermich et al., 2012; Rothermich & Kotz, 2013)和汉语中的双字词模式(Li et al., 2019),并非语言固有的、自然规则,人们可能难以感知或有效整合这些模式,从而难以揭示韵律与语义加工的真实关系。

基于此,本研究以古代格律诗为实验材料,同时操纵诗歌的平仄和语义,采用 EEG 技术考察诗歌阅读中韵律和诗意加工的认知神经机制。研究的创新点体现在以下三个方面。首先,声调主要依赖基频(F0)模式,属于“音高导向”的韵律特征,本研究率先探讨声调组合与语义加工之间的相互影响,为理解韵律与语义关系的理论争论提供了重要补充。其次,对于诗歌平仄规则的操纵,比如“山寺桃花始盛开/山寺桃花始盛放”,虽然平仄模式发生了变化,但是词义并不改变,因此提供了更加干净的韵律操作。同时,古诗词是真实的文化语料,体现了声调在实际语言使用中的自然模式,比实验室构建的句子更具生态效度,减少人工操纵的偏差。

我们假设在早期阶段存在平仄效应,表现为平仄不合理诗句相比合理诗句诱发波幅更大的 N100 或者 P200。在随后的 N400 和 LPC 成分上,诗歌平仄感知和语义理解之间存在相互作用。N100 和 P200 是由视觉刺激或者听觉刺激诱发的早期 ERP 成分,对刺激的声学特征敏感(Näätänen & Picton, 1987),反映了将刺激输入与存储在记忆中的知觉表征进行匹配的过程(即“知觉匹配”, Luck & Hillyard, 1994)。N400 是语义加工的典型 ERP 指标(Brown & Hagoort, 1993; Hagoort et al., 2004; Kutas & Federmeier, 2011; Lau et al., 2008; Nieuwland et al., 2020),一些阅读或者言语理解研究发现 N400 也与语音加工有关,可能反映了人们对语音或者超音段信息的预期(Chen et al., 2016; Coch et al., 2005; Rothermich et al., 2012; Rugg, 1984; Zhang et al., 2023)。LPC 是在刺激呈现后 500~1000 ms 出现的具有广泛分布或中央-顶叶分布的晚期正成分,最初被认为反映对于句法结构的再分析(Friederici,

1995; Hagoort et al., 1999; Osterhout & Holcomb, 1992)或者反映句法整合难度(Kaan et al., 2000; Kaan & Swaab, 2003),后续研究发现 LPC 也与语义或者韵律的整合与再分析有关(Brouwer et al., 2012, 2017)。同时对 EEG 数据进行解码分析,如果基于 EEG 数据训练得到的神经网络能够对不同诗句类型进行有效分类,则进一步印证平仄和诗意是中国古代诗歌加工的重要特征。

2 方法

2.1 被试

选取古代文学相关专业的学生作为实验被试。为了保证被试拥有较为丰富的诗歌经验,请被试完成《古代诗歌知识测验测评问卷》(陈庆荣, 窦芬, 2021),选取测验分数在 60 分以上者。使用 G*Power 3.1 计算所需样本量,为了在显著性水平 0.05 下检测中等效应($f = 0.25$)并达到 95% 的统计检验力(Cohen, 1988),至少需要 36 名被试。为了防止存在行为正确率过低或者脑电伪迹过多而需要删除的被试,以及遗漏可能存在的较小实验效应,最终 48 名被试(女生 24 名,男生 24 名)参与实验。这些被试的年龄范围在 18~28 岁之间,平均年龄 21.25 ± 2.29 岁。诗歌经验得分在 60~98 分之间,平均得分 80.83 分($SD = 11.89$)。

被试均为右利手,母语为汉语,裸眼视力或矫正视力正常,无神经系统疾病或精神病史。所有被试均自愿参加实验,实验之前签署知情同意书,实验结束之后获得一定报酬。研究得到南京师范大学生物医学伦理委员会(伦理批准号: NNU202410030)的批准。

2.2 实验设计与材料

采用 2 (语义合理、语义不合理) \times 2 (平仄合理、平仄不合理) 两因素被试内实验设计,因变量为被试按键反应的正确率和阅读诗句中关键字词的 EEG 信号。

从全唐诗中选取 100 联诗句作为实验材料,这些诗句严格符合平仄规则,并且古音和今音在平仄上未发生变化(罗竹风, 1993)。为了排除押韵可能产生的影响,所选联句两句之间均不押韵。由于这些诗句语义合理并且符合平仄规则,因此构成第一种实验条件(semantic + tone +, 简称 S+T+)。由于实验中要求被试对诗句语义进行判断,因此需要选取难度较低的诗句,以保证被试能够理解诗句的意义。同时,为了避免被试基于熟悉性产生预期从而影响诗句的正常加工,所以选取不熟悉的诗句作为材

料。10 名不参加正式实验的古代文学相关专业学生对诗句难度和熟悉度进行了评定。难度评定采用 5 点量表, 其中“1”代表非常简单, “5”代表非常难。结果显示, 所选诗句的平均难度为 3 ($SD = 0.29$), 因此所选诗句难度并不高。熟悉性评定也采用 5 点量表, 其中“1”代表非常不熟悉, “5”代表非常熟悉。结果显示, 所选诗句的平均熟悉度为 2.36 ($SD = 0.37$), 因此这些诗句的熟悉度较低, 符合实验要求。

对联句中第二句诗句的最后一个字进行修改, 形成另外三种实验条件, 分别是语义合理但平仄不合理(S+T-), 语义不合理但平仄合理(S-T+), 语义不合理并且平仄不合理(S-T-)。四种实验条件下关键字的字频不存在显著差异($p = 0.10$)。实验材料举例如表 1 所示。

表 1 不同平仄合理性和语义合理性的古代诗句

实验条件	诗句样例
S+T+	竹凉/侵/卧/内/, /野月/满/庭/隅/。
S+T-	竹凉/侵/卧/内/, /野月/满/庭/院/。
S-T+	竹凉/侵/卧/内/, /野月/满/庭/啡/。
S-T-	竹凉/侵/卧/内/, /野月/满/庭/肺/。

注: 加粗的字为关键字。分段呈现诗句, 每两段之间以“/”分隔。

材料编写完成之后, 对诗句语义合理性进行了评定。26 名不参加正式实验的古代文学相关专业学生对 4 种实验条件下诗句的语义合理性进行评定, 每个被试只评定一组诗句的其中一个版本。采用 5 点量表进行评定, 其中“1”代表语义非常不合理, “5”代表语义非常合理。4 种条件语义合理性评分的均分和标准差分别是: S+T+ ($M = 4.52, SD = 0.73$), S+T- ($M = 4.40, SD = 0.72$), S-T+ ($M = 2.08, SD = 1.28$), S-T- ($M = 2.09, SD = 1.22$)。线性混合效应模型(Linear Mixed Model, LMM)分析显示, 语义合理性主效应显著($F = 674.00, p < 0.001$), 语义合理性与平仄合理性交互作用不显著($F = 1.36, p = 0.25$),

表明语义不合理诗句的评分显著低于语义合理诗句, 并且诗句的语义合理性不受平仄合理性的影响, 因此诗句的语义操纵符合实验设计的要求。同时, 平仄合理性主效应不显著($F = 0.90, p = 0.34$), 即平仄合理与不合理诗句在语义合理性上没有显著差异, 因此正式实验中得到的平仄效应不会受到语义因素的混淆。

2.3 实验程序

实验在隔音、电屏蔽实验室开展。实验过程中, 屏幕中央首先呈现注视点“+” 1000 ms, 接着分段呈现诗句。每段呈现 600 ms, 段与段之间间隔 400 ms, 但是每句最后一个字与逗号或者句号之间的间隔为 600 ms。单独呈现标点符号是为了减少句末效应对关键字处 EEG 信号的影响(Just & Carpenter, 1980; Luo & Zhou, 2010)。句子呈现结束之后, 出现 1000 ms 空屏, 接着屏幕中呈现 4 个选项, 分别为: 平仄合理语义合理、平仄不合理语义合理、平仄合理语义不合理、平仄不合理语义不合理。要求被试根据自己对诗句的理解, 用鼠标点击选项做出判断(陈庆荣, 杨亦鸣, 2017)。选项位置经过了平衡, 使得每个选项出现在四宫格任意位置上的概率均为 25%。选项的呈现顺序也在被试间进行了平衡, 一半被试的选项是语义在前平仄在后, 例如“语义合理, 平仄合理”, 另外一半被试的选项是平仄在前语义在后, 例如“平仄合理, 语义合理”。被试做出判断之后, 出现 1500 ms 空屏, 接着呈现下一个试次。实验流程示意图如图 1 所示。

正式实验共有 4 套实验材料, 每个被试完成其中一套材料, 使得同一组诗句只给被试呈现一次。每套材料包含 100 联实验诗句, 每种实验条件各 25 联。为了避免被试产生反应偏向, 另外选取 40 联七言律诗作为填充材料放入每套材料中。因此, 给每个被试呈现 140 联诗句, 诗句呈现顺序进行伪随机排列, 使得同一实验条件的诗句不会连续出现

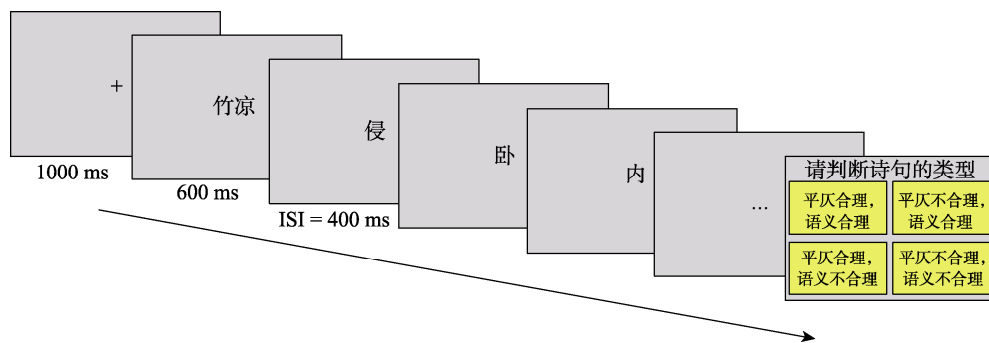


图 1 实验流程示意图

两次以上。每个实验分成两个组块, 每个组块包含 70 联诗句。组块顺序在被试之间进行平衡, 两个组块之间让被试进行短暂休息。在正式实验之前, 被试进行 10 个试次的练习实验, 帮助被试熟悉实验流程和实验任务。整个实验(包括准备时间)大约持续 2 小时。

2.4 数据记录与分析

2.4.1 行为数据分析

正确率分析采用逻辑线性混合效应模型(Logistic Linear Mixed Effects Model, LLMEM)进行。其中被试的按键反应(0/1)为因变量, 固定效应为语义合理性、平仄合理性, 随机效应为收敛状态下被试的最大斜率和截距以及项目的最大斜率和截距(Barr et al., 2013)。由于我们在编制材料时没有同时匹配四个实验条件关键字的笔画数, 因此将笔画数作为一个协变量纳入了模型: `model <- glmer(response ~ semantic * tone + stroke + (1 + semantic * tone | sub) + (1 | item), family = binomial, control = glmerControl(optimizer = 'bobyqa'), data)`。模型中, response 代表被试的按键反应, semantic 代表语义合理性, tone 代表平仄合理性, stroke 代表笔画数, subject 代表被试, item 代表诗句材料。如果存在交互作用, 采用 emmeans 包进行简单效应分析(Lenth & Lenth, 2018)。

2.4.2 脑电数据记录与预分析

采用美国 NeuroScan 公司生产的 64 导脑电记录系统与符合国际标准 10~20 记录系统的 Ag/AgCl 电极帽收集脑电数据。在线带通滤波器参数设置为 0.05~400 Hz, 在线采样频率为 1000 Hz。采集脑电信号时, 在双眼外侧安置电极记录水平眼电(HEOG), 在左眼上下位置安置电极记录垂直眼电(VEOG)。所有电极均以左耳乳突进行在线参考。采集过程中控制所有电极阻抗保持在 5 K Ω 以下。

采用 Python 的 MNE 软件包对 EEG 数据进行离线分析(Gramfort et al., 2013)。首先, 将以左右两耳乳突平均电位为基准进行离线重参考。随后, 采用 0.1~30 Hz 的数字带通滤波器对数据进行滤波。接着, 利用独立成分分析进行伪迹矫正, 结合人工检查, 剔除眼电、肌电等对脑电数据的影响。以关键字出现之前 200 ms 到出现之后 1000 ms 进行分段, 并用关键字出现之前 200 ms 的平均振幅进行基线校正。此外, 剔除在任意导联上波幅超过 $\pm 75 \mu\text{V}$ 的所有分段。平均而言, S+T+条件保留 24.40 (0.96)个试次, S+T-条件保留 24.67 (0.78)个试次, S-T+条件保留 24.33 (1.19)个试次, S-T-条件

保留 24.48 (0.71)个试次。

2.4.3 ERP 分析

通过观察 ERP 成分波形图, 确定 ERP 的分析窗口为 N100 (80~120 ms), P200 (150~250 ms), N400 (250~450 ms), LPC (500~700 ms)。ERP 分析采用线性混合效应模型(LMM)进行, LMM 相较于传统方差分析提供了更稳健的参数估计(Brauer & Curtin, 2018; Winter, 2013)。LMM 分析中, 包含 4 个因素, 分别是语义合理性(合理、不合理)、平仄合理性(合理、不合理)、半球(左、右)以及区域(前部、中部、后部)。半球和区域结合起来, 形成 6 个兴趣区, 每个兴趣区包含 5 个或者 6 个电极, 分别是: 左前区域(F1, F3, F5, FC1, FC3, FC5)、左中区域(C1, C3, C5, CP1, CP3, CP5)、左后区域(P1, P3, P5, PO3, PO7)、右前区域(F2, F4, F6, FC2, FC4, FC6)、右中区域(C2, C4, C6, CP2, CP4, CP6)、右后区域(P2, P4, P6, PO4, PO8)。为了控制被试个体差异和诗句材料差异对分析结果的影响, 将被试的随机截距以及在语义合理性和平仄合理性上的随机斜率、诗句材料的随机截距纳入模型, 同时将笔画数作为一个协变量纳入。最终构建模型 `model <- lmer(amplitude ~ semantic * tone * region * hemisphere + stroke + (1 + semantic * tone | subject) + (1 | item), control = lmerControl(optimizer = 'bobyqa'), data)`, 其中 amplitude 代表目标时间窗口内平均振幅, semantic 代表语义合理性, tone 代表平仄合理性, region 代表区域, hemisphere 代表半球, stroke 代表笔画数, subject 代表被试, item 代表诗句材料。如果与实验条件的交互作用达到显著, 采用 emmeans 包进行简单效应分析。进行多重比较时, 采用 Tukey 方法对 p 值进行校正。

2.4.4 基于 CNN-LSTM 的 EEG 分类分析

ERP 分析是一种编码分析, 可以了解大脑如何处理诗歌的平仄和语义维度。同时, 我们对 EEG 数据进行了解码分析, 通过解码 EEG 信号来识别不同类型的诗句。解码分析通过多变量建模和对 EEG 信号中细粒度信息的充分利用, 可以捕捉信号中的局部模式和复杂特征, 更好地反映大脑的整体活动模式(Cox & Savoy, 2003; Kriegeskorte & Douglas, 2019; Tong & Pratte, 2012)。这种双向分析分别从不同的角度帮助我们理解大脑的功能和机制, 可以为分析大脑如何加工平仄规则和诗歌语义提供强有力的证据(Kriegeskorte & Douglas, 2019)。解码分析中, 采用卷积-长短期记忆联合网络(CNN-LSTM)

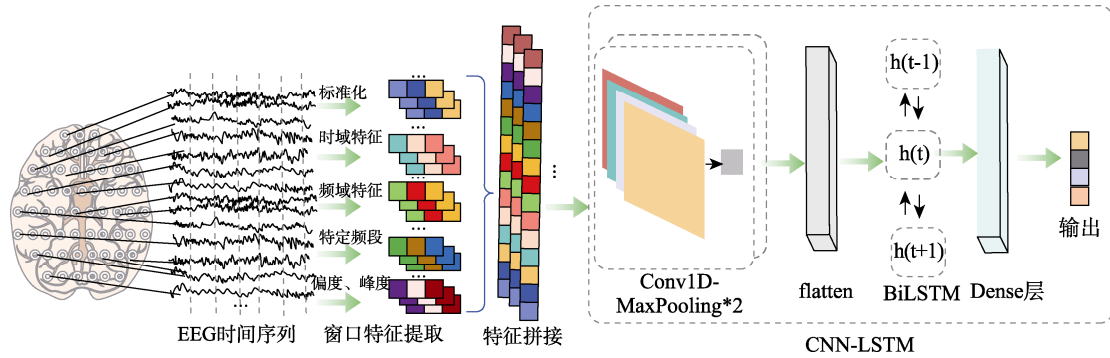


图 2 基于 CNN-LSTM 的 EEG 分类分析流程图

注：彩图见电子版，下同

来通过 EEG 数据对不同诗句类型进行分类 (Bashivan et al., 2015; Tabar & Halici, 2017), 具体流程如图 2 所示。如果在测试数据集上较为可靠地分类出四种诗句类型, 那么 EEG 信号一定包含解码语义和平仄规则的信息。

参考 ERP 分析, 从原始信号中提取 150~700 ms 作为分析窗口(例如 Borra & Magosso, 2021; Li et al., 2024; Tian et al., 2024)。随后采用 5 折交叉验证的方法, 将整个数据集划分为 5 个互斥子集, 每轮选取其中 4 折作为训练集、1 折作为验证集进行模型训练和评估, 循环 5 次后对全部测试结果取平均作为模型的综合评价。数据预处理阶段使用滑动窗口(窗口时长 0.1 秒, 重叠率 85%)切分信号, 对每个通道进行 z-score 标准化。在进行特征提取时, 时域特征包含标准差、均值、极值、峰峰值、偏度、峰度来反映信号幅度波动特性与分布形态。频域特征使用 Welch 法计算频域能量分布。此外, 获取不同频段(delta、theta、alpha、beta 和 gamma)的平均功率刻画信号的频谱信息。接着采用谱熵(spectral entropy)量化信号复杂度。最终, 对以上 5 类特征进行了特征拼接作为模型

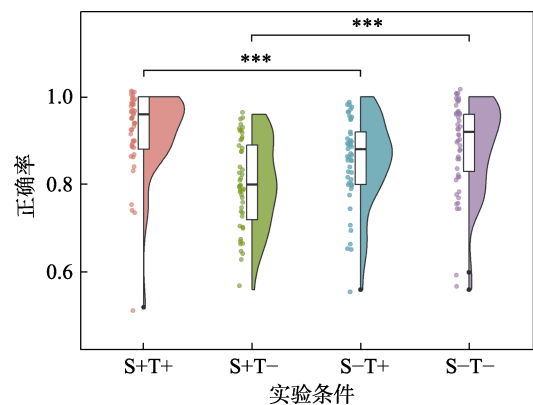
的分类输入。在此基础上, 本文采用 CNN-LSTM 混合网络, 以融合局部特征提取与时序动态建模能力。空间特征提取部分采用 CNN 架构, 以 60 通道原始 EEG 信号为输入。第一卷积层使用 64 个核宽为 5 的一维卷积核, 后接 BatchNorm1d 和 ReLU 激活函数; 第二卷积层将通道数扩展至 128, 保持相同的卷积核宽度。两层卷积后均采用 2×2 最大池化操作, 使时间维度从 240 逐步降至 120 和 60。逐层递增的通道数强化了对复杂空间特征的提取能力, 而渐进式下采样在压缩时域信息的同时保留了对细微时间变化的敏感性。

时序建模部分由两层双向 LSTM 构成, 每层包含 128 个隐藏单元。双向结构能够同时捕获前向和后向的时序信息, 输出 256 维特征向量, 层间引入 Dropout 以缓解过拟合风险。随后, 两层全连接网络依次将特征维度从 256 降至 128, 再映射至 4 类输出。中间层配置 BatchNorm1d、ReLU 激活函数及 0.2 的 Dropout 率, 以提升训练稳定性和模型泛化能力。最终通过 Softmax 层输出类别概率分布。

3 结果

3.1 行为数据分析结果

如图 3 所示, 被试在 4 种实验条件下按键反应的平均正确率及标准差分别是: S+T+ ($M = 92.25\%$, $SD = 0.09$), S+T- ($M = 80.42\%$, $SD = 0.10$), S-T+ ($M = 85.42\%$, $SD = 0.10$), S-T- ($M = 88.75\%$, $SD = 0.10$)。逻辑线性混合模型(Logistic Linear-Mixed model)分析显示, 平仄合理性主效应显著($\chi^2(1) = 7.01$, $p = 0.008$), 平仄合理诗句的正确率显著高于不合理诗句。平仄合理性与语义合理性的交互作用显著, $\chi^2(1) = 52.45$, $p < 0.001$ 。简单效应分析显示,

图 3 四种实验条件下的正确率。*** $p < 0.001$ 。

当平仄合理时, 语义不合理条件相比合理条件正确率显著更低($OR = 0.37, SE = 0.09, z = -4.04, p < 0.001, 95\% CI = [0.31, 0.45]$); 当平仄不合理时, 语义不合理条件相比合理条件正确率显著更高($OR = 2.48, SE = 0.53, z = 4.28, p < 0.001, 95\% CI = [0.89, 6.94]$)。

将所有被试错误的试次放在一起(如图 4 所示), 4 个实验条件下的具体错误试次数量分别为: S+T+ 条件 93 个, S+T- 条件 236 个, S-T+ 条件 176 个, S-T- 条件 135 个。从图中可以看出, 发生错误判断最多的两种类型, 是将 S+T- 类型的诗句判断为 S-T- 类型(179 个试次), 以及将 S-T+ 类型的诗句判断为 S+T+ 类型(143 个试次)。

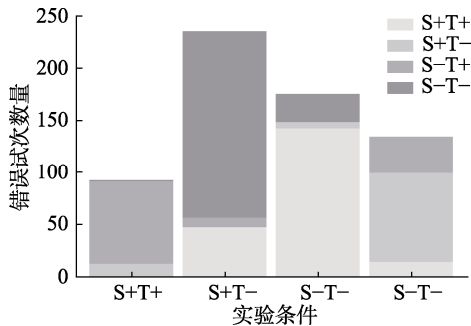


图 4 四种实验条件下所有被试错误试次分布图

3.2 ERP 分析结果

被试在四种实验条件下关键字处的总平均 ERP 如图 5 所示。在 N100 (80~120 ms) 时间窗口, 没有发现任何显著的主效应或者交互效应($F_s < 1.07, p_s > 0.34$)。在 P200 (150~250 ms) 时间窗口, 平仄合理性与区域交互作用显著, $F(2, 27851.3) = 3.71, p = 0.024, \eta_p^2 = 0.0003$ 。简单效应检验显示(如图 6a 所示): 在头皮前部区域($\beta = 0.51, SE = 0.22, z = 2.26, p = 0.02, 95\% CI = [0.07, 0.95]$), 平仄不合理诗句相比合理诗句诱发更大的 P200, 在头皮中部区域边缘显著($\beta = 0.39, SE = 0.22, z = 1.75, p = 0.080, 95\% CI = [-0.05, 0.83]$), 在头皮后部区域两种类型诗句之间没有显著差异($p = 0.88$)。

在 N400 (250~450 ms) 时间窗口, 语义合理性的主效应显著, $F(1, 48.9) = 108.73, p < 0.001, \eta_p^2 = 0.69$, 语义不合理诗句相比合理诗句诱发更大的负波。语义合理性与平仄合理性的交互效应显著, $F(1, 47.9) = 7.18, p = 0.01, \eta_p^2 = 0.13$ 。简单效应检验显示(如图 6b 所示): 当平仄合理时, 语义不合理诗句相比合理诗句诱发更大的负波($\beta = -2.39, SE = 0.27, z = -8.86, p < 0.001, 95\% CI = [-2.92, -1.86]$); 当平仄不合理时, 语义合理性效应仍然存在, 但是明显

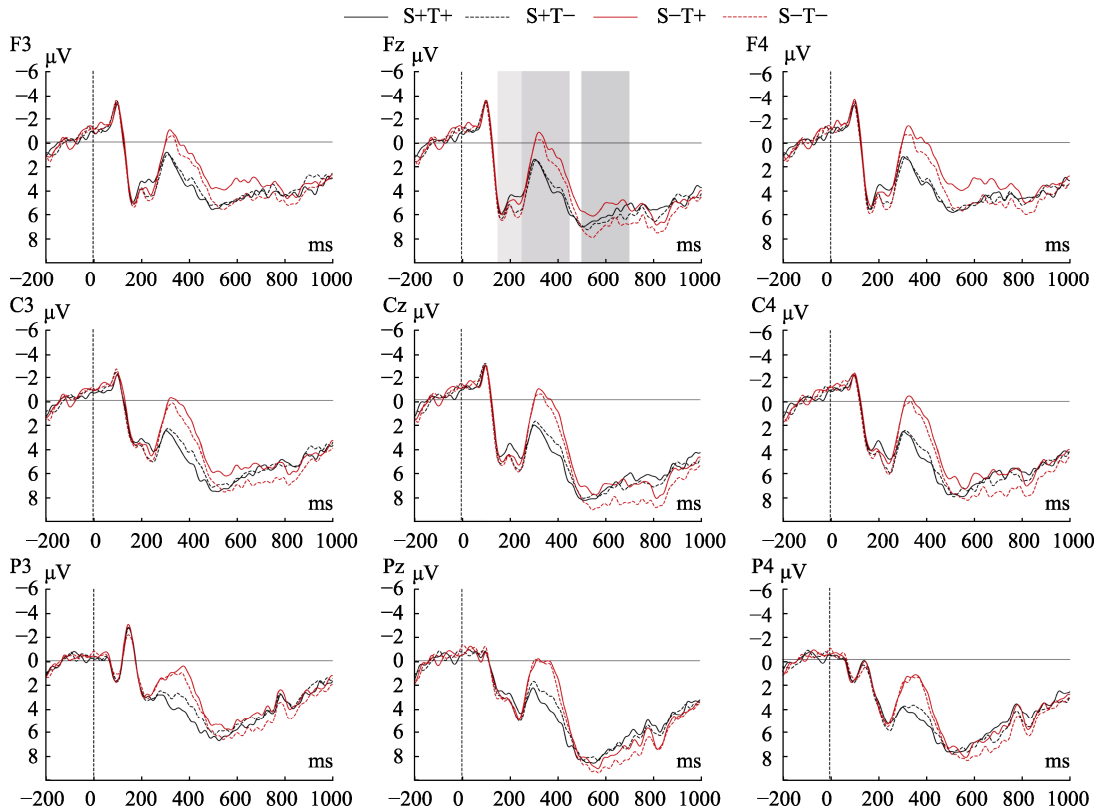


图 5 四种不同类型诗歌关键字诱发的平均波幅。

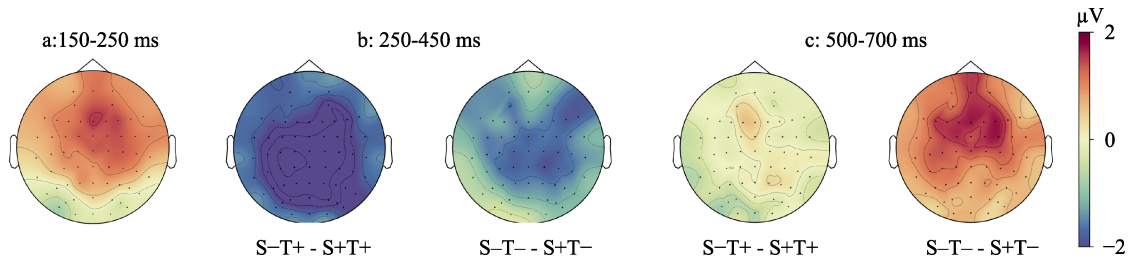


图 6 平仄合理性的主效应及其与语义合理性的交互效应。(a)平仄合理性效应(P200);(b)平仄合理性和语义合理性的交互作用(N400);(c)平仄合理性和语义合理性的交互作用(LPC)。

更弱($\beta = -1.32, SE = 0.27, z = -4.99, p < 0.001, 95\% CI = [-1.84, -0.80]$)。由此可见,合理的平仄规则能够促进诗歌意义加工。

在 LPC (500~700 ms)时间窗口,平仄合理性的主效应显著, $F(1, 48.9) = 4.04, p = 0.05, \eta_p^2 = 0.08$,平仄不合理诗句相比合理诗句诱发更大的 LPC。语义合理性与区域的交互作用显著, $F(2, 27852.4) = 4.47, p = 0.01, \eta_p^2 = 0.0003$,简单效应检验显示语义不合理相比语义合理在头皮前部区域诱发更大的负波($\beta = -0.64, SE = 0.33, z = -1.97, p = 0.05, 95\% CI = [-1.28, -0.002]$),在中部和后部语义效应不显著($ps > 0.64$)。语义合理性与平仄合理性的交互作用显著, $F(1, 48.3) = 9.61, p = 0.003, \eta_p^2 = 0.17$ 。简单效应检验显示,此时的平仄效应受到诗句语义合理性的调节(如图 6c 所示):当语义不合理时,平仄不合理诗句相比合理诗句诱发更大的 LPC ($\beta = 1.24, SE = 0.37, z = 3.35, p < 0.001, 95\% CI = [0.51, 1.96]$);当语义合理时,平仄合理性效应消失($p = 0.21$)。

3.3 基于卷积-长短期记忆联合网络(CNN-LSTM)的 EEG 分类结果

根据脑电信号,利用卷积神经网络对四种诗句类型进行分类,分类结果如图 7 所示。在图 7(1)中,精确率(precision)指在所有被模型预测为某一类别

的样本中,实际属于该类别的样本比例,反映了模型在正类预测上的准确性;召回率(recall)指在所有实际属于某一类别的样本中,被模型正确预测为该类别的样本比例,反映了模型对正类样本的覆盖能力;F1 分数(F1-Score)是精确率和召回率的调和平均值,用于综合评估模型的精确性和召回能力。这些指标均是多分类任务中常用的分类性能指标。在图 7(2)中,真阳性率(true positive)指实际正类中被正确识别为正类的比例,假阳性率(false positive)反映实际负类中被误分类为正类的比例;而 AUC 是 ROC 曲线下的面积,反映模型的整体分类能力,红色虚线表示随机分类器的参考线。

如图所示,各类别的 AUC 值在 0.95~0.97 之间,宏平均 AUC (Macro-Averaged AUC, 所有类别 AUC 的算数平均数)和微平均 AUC (Micro-Averaged AUC, 将所有类别的样本合并后计算 AUC 值)均达到 0.96,充分验证了模型对不同类型诗句的区分能力。对分类结果中的精确率、召回率、以及综合精确率和召回率指标的 F1 分数的置换检验结果发现,所有条件的真实分类结果均显著高于随机分类概率 25% ($ps < 0.001$),表明模型能够较为可靠地识别不同平仄和语义特征的诗句。

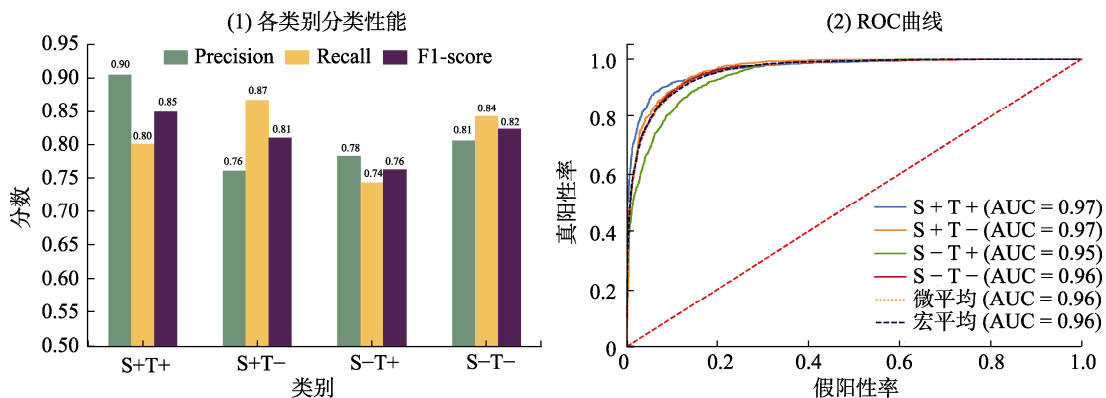


图 7 根据 EEG 数据对不同诗句类型进行分类的性能(1)和多类别分类的 ROC 曲线(2)。

4 讨论

诗歌是语言和音乐合成的一种艺术形式,“不能欣赏诗的音乐者对于诗的精微处恐终隔膜”(朱光潜, 2012)。本研究采用 EEG 技术,通过给诗歌经验者呈现严格符合平仄规则的五言绝句,同时操纵这些诗句的平仄和语义合理性,考察中国古代诗歌阅读过程中平仄和诗意理解的认知神经机制。行为结果显示,平仄合理性主效应显著,并且平仄合理性和语义合理性之间存在交互作用。ERP 结果进一步显示,在诗歌阅读过程中存在平仄效应,表现为早期的 P200 效应,并且平仄和语义在 N400 和 LPC 成分上存在交互作用。同时,基于 EEG 数据训练得到的神经网络能够有效分类四种类型实验诗句,进一步说明了平仄规则和诗歌意义是诗歌加工的重要特征。本研究结果表明,人们在阅读中国古代诗歌过程中形成了讲究平仄规则的特殊偏好,并且诗歌的这种音乐性功能与诗歌意义理解之间在中晚期存在相互作用。

平仄规则对诗歌意义加工的影响体现在行为结果上。当诗句平仄合理时,人们对于语义不合理诗句的判断难度更高,这与诗歌阅读中押韵特征的加工(Chen et al., 2016; 陈庆荣, 杨亦鸣, 2017)以及口语语篇理解中韵律加工(Rothermich et al., 2012)的研究结果一致。对错误试次模式进行分析发现,当诗句的平仄合理时,人们倾向于将语义不合理的诗句判断为语义合理。与此相对,诗句平仄不合理时,人们对于语义合理诗句判断的正确率更低,同样表明诗歌意义的理解受到诗歌韵律特征的影响,诗句平仄不合理时人们更加倾向于认为诗歌语义也不合理。

ERP 结果进一步揭示,诗歌阅读中存在平仄效应,这种效应在关键字呈现 200 ms 左右既已发生。平仄不合理诗句相比合理诗句诱发波幅更大的 P200 成分,是因为被试根据诗歌阅读经验产生自上而下的韵律预期。当呈现的诗句平仄不合理时,自下而上的刺激输入与被试记忆中的平仄表征无法匹配,预期错误带来的加工代价导致诱发波幅更大的 P200 成分(Chen et al., 2016)。尽管以往研究也发现包括 P200 在内的早期脑电成分反映形式预期加工(比如语音、正字法),但是 P200 反映形式预期究竟发生在语言加工的哪个阶段仍然存在争论。根据感觉假设(the sensory hypothesis),P200 上发现的韵律预期效应反映了韵律预期影响的是以基本感

知特征为输入的最底层、前词汇加工水平(Dambacher et al., 2009; Dikker et al., 2010; Dikker et al., 2009),而词汇识别假设(the word recognition hypothesis)则认为,韵律预期影响的是词汇加工阶段,在这一阶段字词的语音信息得到通达(Connolly & Phillips, 1994; Nieuwland, 2019)。P200 效应到底反映前词汇加工或是词汇加工需要后续研究进一步厘清,但是反映出基于平仄规则的韵律预期效应在诗句阅读的早期加工过程已经产生影响。

在随后的 N400 效应上,诗歌语义加工受到平仄合理性的调节。通过操纵德语句子的重音模式(Rothermich et al., 2012)或者汉语的短语节奏(Luo & Zhou, 2010),同样发现句子的韵律特征和句子语义在 N400 上存在交互作用,合理的韵律模式能够促进句子加工。在先前一项考察诗歌押韵和语义加工的研究中,也发现了类似的结果,即押韵模式的规则性对诗歌语义 N400 效应存在调节作用(Chen et al., 2016)。在诗歌阅读过程中存在明显的诗意性功能(poetic function),即诗歌文本使得读者更多关注诗歌的韵律特征(Hanauer, 1998)。诗意性功能的焦点是诗歌文字之间的内部联系,表现为诗歌具有的相似性、规则性和重复性等结构特征。平仄规则作为中国古代诗歌的典型韵律特征,具有明显的相似、相对和重复结构。根据内隐或者外显的学习经验,读者阅读古代格律诗会有意识地注意诗歌的平仄结构,读到平声时期待仄声的到来,读到仄声时又预期平声的复返。当出现的字音不符合诗歌平仄规则时,犹如步步上楼梯时猛然发现一级楼梯特别高或者特别低,会迅速分配更多的注意资源到诗歌韵律特征之上,从而导致此时语义效应减弱。

平仄规则和诗歌意义的交互影响持续到 LPC 时间窗口。平仄不合理条件相比合理条件诱发更大的 LPC,这与先前语音、韵律相关研究一致,例如语音或者声调违反会诱发波幅更大的 LPC(张政华等, 2021; Hu et al., 2012; Schirmer et al., 2005; Zhang et al., 2023)。同样,在对句子最后一个单词或者音乐最后一个音符的音高进行操纵时,音高违反也会诱发更大的 LPC(Magne et al., 2006; Schön et al., 2004)。这些研究结果共同表明,LPC 可能反映冲突的检测与解决(Bornkessel-Schlesewsky & Schlewsky, 2008; Kim & Osterhout, 2005),以及更一般的整合与再分析过程(Brouwer et al., 2012; Liu et al., 2010)。

Astésano 等人(2004)通过拼接技术对句子的语

调进行了操纵,例如将陈述句开头与疑问句结尾拼接,形成语调不合理条件。尽管这种韵律操纵与语用因素存在一定耦合(Rothermich et al., 2012),但其对句子整体音高模式的操控,与本研究的平仄规则具有较强的可比性。该项研究与本研究结果具有高度相似性:语调违反会诱发更大的晚期正成分,并且与句子语义存在交互作用,在语义不合理时韵律效应波幅更大。在晚期阶段,读者调动认知资源,尝试对不合理的诗句进行再次分析,在这个有意识的加工过程中,通过对相似性、重复性等韵律结构规则的分析,重构诗歌文本的非线性层级结构(Hanauer, 1998)。具体而言,当语义不合理时,不规则的平仄会加剧整体的不和谐,导致更大的 LPC 效应;而当语义合理时,诗歌的诗意缓冲了韵律瑕疵,从而使该效应消失。

需要注意的是,在 Chen 等人(2016)的研究中,押韵违反并未诱发 LPC 效应。这种加工进程的差异可能源于两个因素:首先,韵母在词汇识别和语义限制方面具有更强的约束力和影响力,使得相较于声调,韵母加工相对更快(Hu et al., 2012; Li et al., 2014; Tong et al., 2008)。其次,押韵和平仄是诗歌中两种不同的韵律特征。押韵主要涉及字词尾部的音韵匹配,属于相对局部的语音模板检测,通常不需要进行大规模的结构重新整合。而平仄是建立在单句之内和联句之间的一组声调规则,属于更加整体的结构性规则,需要重新分析诗歌的韵律结构或进行修复,因而在晚期诱发 LPC。

同时,我们采用卷积-长短期记忆联合网络(CNN-LSTM)来通过 EEG 数据对不同诗句类型进行分类。这种解码分析是编码的逆过程,目标在于揭示编码的内容,有时被描述为“读取大脑”或者“破解神经代码”(Haynes & Rees, 2006; Tong & Pratte, 2012)。EEG 分类分析发现,通过 EEG 数据能够较为可靠地区分出 4 种诗句类型,说明诗歌的语义和平仄特征以某种被分类器利用的格式存在于 EEG 信号中(Kriegeskorte & Douglas, 2019)。因此,解码分析进一步揭示了大脑对诗歌的编码信息,说明平仄规则和诗歌意义是中国古代诗歌加工的重要特征。

本研究同时操纵中国古代诗歌的平仄模式和语义特征,采用 EEG 技术观察诗歌阅读中韵律和诗意的认知神经机制。结果发现,中国古代诗歌阅读中存在平仄效应,表现为平仄模式在早期阶段影响词汇语音表征,以及在中晚期阶段与诗歌意义的

相互制约和整合,最终实现诗歌的理解与欣赏过程。同时,基于 EEG 数据训练得到的神经网络模型能够对四个实验条件进行有效区分。由此可见,人们在诗歌阅读中偏好节奏和谐、错落有致的句子,进而期待声音与意义的和谐统一。实验结果验证了认知神经诗学模型(Neurocognitive Poetics)(Jacobs, 2015),即诗歌的前台特征(比如平仄、押韵)与后台特征(比如词汇和句子理解)之间存在相互作用,并在此基础上进一步揭示了诗歌阅读中韵律特征和诗歌意义交互的时间进程和模式。研究结果也揭示了声调组合这一音高导向的韵律特征与语义之间如何共同加工,支持三阶段神经认知模型和互动激活模型:在加工进程上,韵律效应贯穿语言加工的不同阶段,而非局限于早期阶段;在交互模式上,韵律与语义之间存在双向相互影响,体现出动态互动而非单向流动。

本研究主要关注了诗歌经验者的阅读过程,未来研究可以进一步考察更普遍的一般读者如何加工诗歌的平仄和语义,以及两组人群在加工方式上存在何种区别和联系。此外,根据双通路神经模型(Schwartz & Kotz, 2013),人脑中存在两条通路分别加工刺激的节奏维度和内容维度,这是大脑的一个基本功能。未来研究需要探讨诗歌韵律与诗歌意义各自的加工通路特征,以及这两条通路之间可能存在的相互作用模式。最后,鉴于诗歌与音乐在韵律结构和情感表达功能上都具有紧密的关联,后续研究可以系统分析诗歌与音乐在认知加工上的异同点,考察音乐经验对诗歌韵律和情感加工的潜在影响机制。

参 考 文 献

- Aristotle, A. (1968). *Poetics* (Vol. 9). Clarendon Press.
- Astésano, C., Besson, M., & Alter, K. (2004). Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research*, 18(2), 172–184.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bashivan, P., Rish, I., Yeasin, M., & Codella, N. (2015). Learning representations from EEG with deep recurrent-convolutional neural networks. *arXiv preprint arXiv:1511.06448*.
- Beck, J., & Konieczny, L. (2020). Rhythmic subvocalization: An eye-tracking study on silent poetry reading. *Journal of Eye Movement Research*, 13(3), 1–40.
- Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on “semantic P600” effects in language comprehension. *Brain Research Reviews*, 59(1),

- 55–73.
- Borra, D., & Magosso, E. (2021). Deep learning-based EEG analysis: Investigating P3 ERP components. *Journal of Integrative Neuroscience*, 20(4), 791–811.
- Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, 23(3), 389–411.
- Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science*, 41(Suppl 6), 1318–1352.
- Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, 1446, 127–143.
- Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5(1), 34–44.
- Chen, Q., & Dou, F. (2021). Constructing a Questionnaire to Test the Knowledge of Ancient Chinese Poetry: A preliminary result. *Journal of Nanjing Normal University (Social Science Edition)*, 3, 73–82.
- [陈庆荣, 窦芬. (2021). 古诗知识经验测评问卷的初步编制. *南京师大学报(社会科学版)*, 3, 73–82.]
- Chen, Q., & Yang, Y. (2017). Cognitive mechanisms in reading ancient Chinese poetry: Evidence from eye movements. *Social Sciences in China*, 3, 48–76+205.
- [陈庆荣, 杨亦鸣. (2017). 古诗阅读的认知机制: 来自眼动的证据. *中国社会科学*, (3), 48–76+205.]
- Chen, Q., Zhang, J., Xu, X., Scheepers, C., Yang, Y., & Tanenhaus, M. K. (2016). Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition*, 154, 11–21.
- Coch, D., Grossi, G., Skendzel, W., & Neville, H. (2005). ERP nonword rhyming effects in children and adults. *Journal of Cognitive Neuroscience*, 17(1), 168–182.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.
- Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, 6(3), 256–266.
- Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) “brain reading”: Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*, 19(2), 261–270.
- Culler, J. (1975). *Structuralist poetics: Structuralism, linguistics and the study of literature*. Routledge.
- Dambacher, M., Rolfs, M., Göllner, K., Kliegl, R., & Jacobs, A. M. (2009). Event-related potentials reveal rapid verification of predicted visual input. *PLoS One*, 4(3), e5047.
- Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, 21(5), 629–634.
- Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, 110(3), 293–321.
- Fodor, J. A. (1983). *The modularity of mind*. MIT Press.
- Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, 50(3), 259–281.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2), 78–84.
- Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16(5), 262–268.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... Hämäläinen, M. S. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7, 267.
- Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. M. Brown, & P. Hagoort (Eds.), *The neurocognition of language*, (pp. 273–316). <https://doi.org/10.1093/acprof:oso/9780198507932.003.0009>
- Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304(5669), 438–441.
- Hanauer, D. (1998). The genre-specific hypothesis of reading: Reading poetry and encyclopedic items. *Poetics*, 26(2), 63–80.
- Haynes, J., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523–534.
- Hirst, D., & Di Cristo, A. (Eds.). (1998). *Intonation systems: A survey of twenty languages*. Cambridge University Press.
- Hu, J., Gao, S., Ma, W., & Yao, D. (2012). Dissociation of tone and vowel processing in Mandarin idioms. *Psychophysiology*, 49(9), 1179–1190.
- Jacobs, A. M. (2015). Neurocognitive poetics: Methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, 9, 186.
- Jakobson, R. (1960). Linguistics and poetics. In T. Sebeok (Ed.), *Style in language* (pp. 350–377). Cambridge, MA: MIT Press.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354.
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15(2), 159–201.
- Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15(1), 98–110.
- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52(2), 205–225.
- Kriegeskorte, N., & Douglas, P. K. (2019). Interpreting encoding and decoding models. *Current Opinion in Neurobiology*, 55, 167–179.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62(1), 621–647.
- Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933.
- Lenth, R., & Lenth, M. R. (2018). Package ‘lsmeans’. *The American Statistician*, 34(4), 216–221.
- Li, S., Zhang, T., Yang, F., Li, X., Wang, Z., & Zhao, D.

- (2024). A dynamic multi-scale convolution model for face recognition using event-related potentials. *Sensors*, 24(13), 4368.
- Li, W., Wang, L., & Yang, Y. (2014). Chinese tone and vowel processing exhibits distinctive temporal characteristics: An electrophysiological perspective from classical Chinese poem processing. *PLOS ONE*, 9(1), e85683.
- Li, W., & Yang, Y. (2010). Perception of Chinese poem and its electrophysiological effects. *Neuroscience*, 168(3), 757–768.
- Li, X., Shao, X., Xia, J., & Xu, X. (2019). The cognitive and neural oscillatory mechanisms underlying the facilitating effect of rhythm regularity on speech comprehension. *Journal of Neurolinguistics*, 49, 155–167.
- Liu, Y., Li, P., Shu, H., Zhang, Q., & Chen, L. (2010). Structure and meaning in Chinese: An ERP study of idioms. *Journal of Neurolinguistics*, 23(6), 615–630.
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31(3), 291–308.
- Luo, Y., & Zhou, X. (2010). ERP evidence for the online processing of rhythmic pattern during Chinese sentence reading. *NeuroImage*, 49(3), 2836–2849.
- Luo, Z. (1993). *Chinese Dictionary Compendium*. Shanghai: Publishing House of Chinese Dictionary Compendium.
- [罗竹风. (1993). *汉语大词典*. 上海: 汉语大词典出版社.]
- Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinot, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: Behavioral and electrophysiological evidence. *Cerebral Cortex*, 17(11), 2659–2668.
- Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18(2), 199–211.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375–407.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, 24(4), 375–425.
- Nietzsche, F. (2012). *The birth of tragedy*. Courier Corporation.
- Nieuwland, M. S. (2019). Do ‘early’ brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience & Biobehavioral Reviews*, 96, 367–400.
- Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., ... Von Grebmer Zu Wolfsturn, S. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society B*, 375(1791), 20180522.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31(6), 785–806.
- Rothermich, K., & Kotz, S. A. (2013). Predictions in speech comprehension: fMRI evidence on the meter-semantic interface. *NeuroImage*, 70, 89–100.
- Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia*, 50(2), 232–244.
- Rugg, M. D. (1984). Event-related potentials in phonological matching tasks. *Brain and Language*, 23(2), 225–240.
- Schirmer, A., Tang, S. L., Penney, T. B., Gunter, T. C., & Chen, H. C. (2005). Brain responses to segmentally and tonally induced semantic violations in Cantonese. *Journal of Cognitive Neuroscience*, 17(1), 1–12.
- Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41(3), 341–349.
- Schwartz, M., & Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience & Biobehavioral Reviews*, 37(10), 2587–2596.
- Tabar, Y. R., & Halici, U. (2017). A novel deep learning approach for classification of EEG motor imagery signals. *Journal of Neural Engineering*, 14(1), 016003.
- Tian, X., Zhu, L., Zhang, M., Wang, S., Lu, Y., Xu, X., Jia W., Zheng, Y., & Song, S. (2024). Social anxiety prediction based on ERP features: A deep learning approach. *Journal of Affective Disorders*, 367, 545–553.
- Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology*, 63(1), 483–509.
- Tong, Y., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in mandarin Chinese. *Language and Cognitive Processes*, 23(5), 689–708.
- Tsur, R. (1983). *What is cognitive poetics?* Katz Research Institute for Hebrew Literature, Tel Aviv University.
- Wang, L. (2012). *The metrics of Chinese poetry*. Beijing: Zhonghua Book Company.
- [王力. (2012). *诗词格律*. 北京: 中华书局.]
- Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *arXiv preprint arXiv:1308.5499*.
- Yaron, I. (2002). Processing of obscure poetic texts: Mechanisms of selection. *Journal of Literary Semantics*, 31(2), 133–170.
- Zeman, A., Milton, F., Smith, A., & Rylance, R. (2013). By heart: *What is cognitive poetics?* An fMRI study of brain activation by poetry and prose. *Journal of Consciousness Studies*, 20(9–10), 132–158.
- Zhang, Y. H. (2015). Cognitive poetics and interdisciplinary research on literary comprehension. *Journal of Tsinghua University (Philosophy and Social Sciences)*, 30(2), 139–147+190.
- [张叶鸿. (2015). 认知诗学与跨学科文学理解研究. *清华大学学报(哲学社会科学版)*, 30(2), 139–147+190.]
- Zhang, Z., Han, M., Zhang, F., & Li, W. (2020). Musical training improves rhythm integrative processing of classical Chinese poem. *Acta Psychologica Sinica*, 52(7), 847–860.
- [张政华, 韩梅, 张放, 李卫君. (2020). 音乐训练促进诗句韵律整合加工的神经过程. *心理学报*, 52(7), 847–860.]
- Zhang, Z., Zhang, H., Sommer, W., Yang, X., Wei, Z., & Li, W. (2023). Musical training alters neural processing of tones and vowels in classic Chinese poems. *Brain and Cognition*, 166, 105952.
- Zhang, Z., Zhang, H., Wang, B., Zheng, Z., Zhao, L., & Li, W. (2021). The processing of the tone and the vowel in poems under different tasks: Evidence from ERPs. *Studies of Psychology and Behavior*, 19(6), 728–735.
- [张政华, 张航, 王兵, 郑子龙, 赵黎明, 李卫君. (2021). 不同任务下声调和韵母在诗句中的加工过程——来自 ERP 的证据. *心理与行为研究*, 19(6), 728–735.]
- Zhu, G. Q. (2012). *On Poetry*. Beijing: Zhonghua Book Company.
- [朱光潜. (2012). *诗论*. 北京: 中华书局.]

Cognitive neural mechanisms of tonal patterns and semantic processing in poetry

ZHANG Jingjing¹, SHI Ying¹, DENG Shanwen¹, LI Jiabin², CHEN Qingrong^{1,3}

(¹ School of Psychology, Nanjing Normal University, Nanjing 210097, China)

(² School of Chinese Language and Literature, Nanjing Normal University, Nanjing 210097, China)

(³ Adolescent Education and Intelligence Support Lab of Nanjing Normal University, Laboratory of Philosophy and Social Sciences at Universities in Jiangsu Province, Nanjing 210097, China)

Abstract

Music has long been considered the soul of poetry, a reflection of their shared origins across diverse cultural contexts. Historically, poetry and music were inseparable, with poems frequently performed through songs. As the arts evolved, however, poetry gradually shifted toward a focus on literal meaning, shedding its musical accompaniment and entering an era of “words without tunes.” Yet the musical essence of poetry endures, embedded in the rhythmic, tonal, and melodic qualities of language itself. In classic Chinese poetry, tonal patterns serve as a defining prosodic feature, shaping both the rhythmic flow and the musical character of a poem.

The present study investigated the neural mechanisms underlying tonal patterns in classic Chinese poetry and their interaction with semantic processing. Forty-eight participants with extensive experience reading classic Chinese poems took part in an electroencephalography (EEG) experiment. They were presented with four versions of classic Chinese poems in a pseudo-random order: (1) semantically congruous with regular tonal patterns (S+T+), (2) semantically congruous with irregular tonal patterns (S+T-), (3) semantically incongruous with regular tonal patterns (S-T+), and (4) semantically incongruous with irregular tonal patterns (S-T-). Each participant viewed 100 experimental stimuli and additional 40 filler stimuli. After reading each poem, participants rated the reasonableness of both its semantics and tonal patterns.

Throughout the experiment, both behavioral responses and EEG data were recorded. Accuracy rates and event-related potentials (ERP) were analyzed using linear mixed-effects models via the *lmerTest* and *emmeans* packages in the R, and EEG preprocessing was conducted using the MNE software package in Python. Accuracy results showed that participants performed best on the S+T+ condition, whereas congruous poems with irregular tonal patterns (S+T-) yielded the lowest accuracy.

ERP analyses revealed that tonal patterns exerted a continuous influence on both early and late stages of poetry reading. Specifically, in the P200 window, an interaction between tonal patterns and scalp region indicated that irregular tonal patterns elicited larger P200 amplitudes at anterior and central sites. In the N400 time window, a significant interaction between semantics and tonal patterns emerged: semantically incongruous lines produced a more pronounced negative component than congruous lines when tonal patterns were regular, whereas this semantic difference diminished under irregular patterns. Furthermore, in the LPC time window, irregular tonal patterns evoked larger positivities for semantically incongruous poems, while semantic congruity eliminated the tonal pattern effect altogether. Finally, deep learning models trained on the EEG data reliably distinguished among the four experimental conditions, indicating robust neural signatures associated with the combined tonal-semantic processing.

In summary, the current findings underscore the significant impact of tonal patterns on the reading of classic Chinese poetry. Irregular patterns modulated phonological representation at early stages and constrained semantic comprehension at later stages. Supporting the neurocognitive poetics model (NCPM) of literary reading, these results shed light on how prosodic elements and semantic meaning dynamically interact over time during poetic processing.

Keywords poetry, music, tonal patterns, semantics, EEG