语音相似性调节听觉词汇语义加工的脑电研究

耿立波
江苏师范大学
语言科学与艺术学院
江苏徐州221009
libogeng@163.com

赵悦，陈茜
江苏师范大学
语言能力协同创新中心
江苏徐州221009
3585394513@qq.com
85079953@qq.com

摘要

语音与语义的交互加工机制是理解语言认知过程的核心问题之一。以往研究多集中于词汇层面的线性处理路径，而对音节内部语音片段在语义加工中的作用关注不足。为探索语音信息在词汇语义加工中的调节机制，并为心理加工模型建构提供实证依据，本研究采用事件相关电位（ERP）技术，结合听觉启动范式，考察汉语双音节词尾音（第二音节韵母）相似性对语义加工的影响。实验操控词对的尾音相似性（相同/不同）与语义关系（相关/无关），通过语义判断任务测量被试的行为反应与脑电指标。结果发现：（1）尾音相似词对在晚期N400时间窗内诱发更大的负波幅，提示尾音信息对语义加工过程具有显著调节作用；（2）语义启动效应在尾音不同条件下显著，而在尾音相同条件下消失，显示语音信息可影响语义加工的时间进程与效应强度。研究表明，在听觉词汇加工中，语音片段的结构特征（如尾音相似性）不仅被高度感知，而且会通过调节语义预激活和整合过程参与语义建构。这些发现支持语音-语义交互模型的构想，揭示了语言加工过程中低层语音输入对高层语义处理的动态影响，为听觉词汇识别的认知心理模型建构提供了重要证据。

关键词： 语音片段；听觉词汇加工；双音节词；事件相关电位（ERPs）

An ERP Study on the Modulation of Auditory Lexical-Semantic Processing by Phonological Similarity

Libo Geng
Jiangsu Normal University
School of Linguistic
Sciences and Arts
Xuzhou 221009, Jiangsu, China
libogeng@163.com

Yue Zhao, Qian Chen
Jiangsu Normal University
Collaborative Innovation
Center for Language Ability
Xuzhou 221009, Jiangsu, China
3585394513@qq.com
85079953@qq.com

Abstract

The interaction between phonological and semantic processing is central to understanding language comprehension. Traditional models often assume a modular and linear progression from phonological recognition to semantic activation. However, emerging evidence suggests a more dynamic interplay between lower-level phonological cues and higher-level semantic integration. To explore how sub-syllabic phonological features modulate semantic processing and to inform cognitive model development, the present study employed event-related potentials (ERP) in a semantic judgment task using an auditory priming paradigm. Focusing on disyllabic words in Mandarin Chinese, we manipulated the phonological similarity of the word-final segment (i.e., the rhyme of the second syllable) and the semantic relatedness between prime-target word pairs. Behavioral and ERP results revealed that (1) word pairs with similar final segments

第二十四届中国计算语言学大会论文集，第769页-第782页，济南，中国，2025年8月11日至14日。

(c) 2025 中国中文信息学会计算语言学专业委员会

elicited enhanced negativity in the late N400 time window, indicating a modulation of semantic processing by phonological similarity; (2) semantic priming effects were significant when word-final segments were dissimilar, but disappeared when they were phonologically similar. These findings suggest that native speakers are highly sensitive to word-final phonological information, which influences the timing and magnitude of semantic processing. The presence of similar sounds may reduce cognitive effort by preactivating phonological representations, but may also induce illusory semantic coherence, weakening sensitivity to semantic violations. The study supports an interactive phonological-semantic processing model in auditory word recognition and provides neurocognitive evidence for the role of phonological cues in semantic construction.

Keywords: phonological segments , auditory word processing , disyllabic words , Event-Related Potentials (ERPs).

1

语言理解是一个多层次、多通道的信息加工过程，语音输入与语义建构之间的交互机制构成了语言认知研究的核心议题。传统观点基于模块化语言加工模型，认为语音识别与语义建构属于串行加工过程，彼此相对独立（Marslen-Wilson,1978）。但越来越多的研究显示，语音与语义之间并非完全隔离，而是在加工过程中表现出显著的交互作用（McClelland,1986）。如McClelland 和Elman（1986）提出的轨迹模型（the TRACE model），是交互式模型的典型代表; 邻域激活模型（The Neighborhood Activation Model, NAM）（Luce, 1986）也是比较有影响力的听觉词语加工模型，这类模型都认为大脑对语音、语义、句法等各个维度的信息是进行着同步分析的，并得到了研究证实。比如Knap (2020)在词汇决策任务中控制词汇的邻域密度PND,发现高PND的词因发音相似词多，可能加速语音系统向语义系统的信息传递，从而缩短了词汇决策时间。

近年来，计算机语言模型的发展，如Transformer架构的预训练模型通过自注意力机制（Self-Attention）模拟了语音-语义的并行交互过程，其神经网络层中语音特征与语义表征的动态融合，为语言认知的“交互激活”假说提供了计算证据（Devlin et al., 2019），例如，Feng, C等（2021）通过高时间分辨率的ERP技术和创新的实验设计，首次在汉语中证实了口语产生中语义与语音的并行加工机制，这与BERT等模型中语音嵌入（Phonetic Embedding）与语义向量（Semantic Vector）的联合优化策略高度契合；Goldstein等（2022）研究发现人类大脑与自回归深度语言模型如GPT-2，在处理自然语言时共享核心计算原理。此类跨领域发现表明，语言认知的神经机制可能为改进计算机语言模型提供理论约束，而计算模型的预测能力也可反向验证语言学假设的普适性（Radford et al., 2019）。这些实证研究与计算模型的协同发展，正在推动语言认知研究从“模块化”向“动态交互系统”范式转变。

在自然语境中，语音以时间线性方式呈现，信息接收者无法像阅读书面文本那样即时获取全貌。这种特性促使人脑在部分语音输入的基础上启动预测与整合机制，实现对词汇和意义的提前建构（Marslen-Wilson, 1987）。Yonatan Bisk and Ari Holtzman（2020）基于神经网络的模型，通过大规模语料库训练，学习在部分输入（如句子开头）时激活多可能性预测，并通过上下文概率分布选择最可能的词汇序列。这种机制实际上模拟了人类语言理解的预测性。ERP研究进一步证实，在词汇识别尚未完成时，语义激活已可被观测（Kutas and Federmeier, 2000）。Caucheteux, C（2023）等认为大脑的预测是分层次组织的，不同区域预测的时间范围和层次结构不同，为理解语言处理模型和人工智能提供了新的视角，并强调了预测编码理论在语言处理中的重要性，这表明语义加工可能早于甚至反过来影响语音识别。

汉语作为一种以双音节词为主的孤立语，其词汇单位内部结构简明清晰，尤其适合用于探索音节内部语音片段在词汇识别与语义加工中的作用。已有研究多关注首音（词首声母）对词汇获取的预测功能（Anne Cutler，1990），Federmeier, K. D.and Kutas, M. (2001) 通过事件相关电位（ERP）发现，当首音与预期词汇不符时，会引发更大的N400脑电成分，表明首音在

©2025 中国计算语言学大会
根据《Creative Commons Attribution 4.0 International License》许可出版

词汇获取的早期阶段已被用于预测。但对尾音（词尾韵母）在加工中的调节角色仍缺乏系统考察，尾音在语义层面上的影响还是认知机制建模中的一大空白。

主流的神经语言模型主要有Friederici的模型、交互激活框架等。Friederici （2004, 2011）的神经语言学模型将语言处理分为三个层级——早期句法整合、语义整合、句法-语义接口，模型强调前颞叶的句法优先性，但可能低估了后部颞顶区在声调语言韵律-语义绑定中的作用，汉语尾音或韵母可能通过pSTG/pIPL的语义整合机制影响词汇选择。例如，词尾韵母的声学特征可能触发语音表征的精细调节，从而优化语义预测误差。交互激活框架较好解释了首音的早期预测功能（如Cutler, 1990），首音声母激活候选词汇节点（Friederici, 2012）实现快速竞争抑制，但该框架难以解释汉语尾音的语义调节作用：尾音的韵母信息在交互激活模型中通常被视为“晚期特征”，其语义贡献未被明确显示，应允许韵母的声学特征直接驱动语义预测，如“-ian”韵母激活“尖/千”等高频词汇（王露等，2018）。

此外，从韵律研究的角度来看，尾音所携带的超音段信息（如音高、音长）在语流感知中可能具有语义指引作用，Mark Liberman and Alan Prince（1977）较早将尾音和韵律结合研究，认为尾音的时长、强度和音高变化是韵律边界的重要标记，帮助听者划分话语单元。除此之外，尾音也可能通过诱发声音相似性的启动效应，引发认知上的语义协调错觉，从而影响语言系统对语义偏离的警觉性（Hyman,2007）。这些现象提示我们，语音信息本身可能与意义的建构过程发生复杂而动态的互动。

为进一步阐明语音与语义之间的交互关系，并推动心理语言加工模型向动态耦合系统发展，本研究聚焦汉语双音节词，通过操控词尾韵母的相似性与整词语义的相关性，设计听觉语义判断任务，并结合事件相关电位（ERP）技术，系统考察尾音相似性在语义加工中的调节作用。我们关注：（1）尾音信息是否影响语义加工的效应强度与时间进程？（2）语音-语义交互在神经时序上体现为何种加工阶段？（3）这种现象如何丰富我们对词汇识别中语音与语义相互作用机制的理解?

通过ERP成分分析，尤其聚焦N400与晚期成分（如LPC）的调节模式，我们希望揭示尾音信息在语义建构过程中的心理机制，与现有模型相比，本研究的理论更有助于深化我们对语言理解中语音与语义相互塑造关系的认识，并为构建更具解释力的听觉词汇识别模型提供理论与实证基础。

2 实验设计

2.1 实验被试

由于现代汉语存在着多种方言，不同方言之间对音位的具体发音和分类方式存在着各类差异，为避免被试的方言背景对语音加工可能产生的潜在影响，我们对被试进行了筛选，只保留更接近普通话语音的北方方言区的被试。具体来说，被试的范围被限制在东北官话、冀鲁官话、中原官话区域内，主要包括黑龙江、吉林、辽宁（辽东半岛除外）、河北、山东（胶东半岛除外）、江苏徐州等区域，且筛除了具有长期旅居史的被试。基于上述基本要求，本实验共招募了39 名被试，其中男生15 人，女生24 人。根据Cohen (2013)的研究,理想的统计检验力和效应量均需高于0.8。使用Faul 等人2009 年开发的GPower 软件3.1 版本（http://www.gpower.hhu.de/）对所需的被试样本量进行测算，本研究选取的被试数量符合要求。所有被试的年龄在18-26 岁之间，视力正常或矫正视力正常，无脑外伤和神经系统疾病史，无家族精神病史。在实验前，每位被试都完成了爱丁堡利手测试（Oldfield，1971），确保均为右利手，所有被试均自愿参加实验，并在实验前签署实验知情同意书和实验协议，填写相关信息并确认无误后参与实验，在实验结束后给予每位被试适当的报酬。

2.2 实验材料

本研究从现代汉语词典（第六版）中选取了160 个普通话中常用的双音节名词作为目标刺激，按照词语的词尾韵母是否相同、整词语义是否相关这两个变量为每个目标词匹配了一个启动词。根据汉语音节结构的相关理论，音节的音段信息包括声母和韵母两个部分，声母在前，韵母紧随其后。因此，我们将尾音相同定义为双音节词第二个音节的韵母相同，这与Zou等人（2019）的研究内容定义基本一致。值得注意的是，词尾韵母一致即可视为尾音相同，声调差异并没有影响尾音的一致性，如”小偷”和”扒手”作为一组语义相关的词，虽然词尾的偷[tou]和手[ou]声调并不一致，但是由于韵母ou一致，依然视为尾音相同。因此，启动—目标词对之间

(c) 2025 中国中文信息学会计算语言学专业委员会

的关系可分为四类：尾音相同语义相关（R+S+）、尾音相同语义无关（R+S-）、尾音不同语义相关（R-S+）、尾音不同语义无关（R-S-）。每种类型的刺激有40 对。

条件	R+S+	R+S-	R-S+	R-S-
例 1	比赛—奖牌	口腔—工匠	蟋蟀—昆虫	书包—苦瓜
例 2	国宝—熊猫	酸奶—火柴	粉笔—黑板	货币—胶囊

图 1: 实验语料举例

为避免不同词性的词语之间可能存在的加工差异，实验中使用的所有材料均为名词词语，少量兼类词也以名词词义为主。为保证材料的准确和精细，在语料筛选过程中，含有轻声音节、零声母音节的词语被剔除。除尾音、首音以外的其他片段相同等情况被严格剔除，如大地（词语内部两个音节首音相同）、栏杆（词语内部两个音节尾音相同）、花瓶—盆景（目标词首音同启动词非首音位置重叠）等。此外，汉语中存在大量同音字，我们剔除了可能因同音而产生歧义理解的词语，并确保启动词和目标词中不存在相同音节。

为保证所选的启动词和目标词能被快速加工和理解，我们邀请了54 名大学在读学生对实验材料的熟悉度进行五度测评，其中“1”代表“非常不熟悉”，“5”代表“非常熟悉”。实验成员对每个词语的熟悉度得分进行均值统计后，将熟悉度低于3.5 的词语筛除。测评人员还对词对之间的语义相关性进行了评分，同样采用五度量表的形式，其中“1”代表非常不相关，“5”代表“非常相关”。54 位测评人员对每组词对的语义相关度评分，并进行均值统计。对于R+S+和R-S+两种语义相关词对，只选取平均得分为3.5以上的词对，对于R+S-和R-S-两种语义不相关词对，只选取平均得分为2.5 以下的词对。四种条件下的词对语义相关度测评结果如图。

类型	启动词熟悉度	目标词熟悉度	语义相关度
R+S+	4.81 (0.11)	4.78 (0.13)	4.34 (0.28)
R+S-	4.86 (0.08)	4.81 (0.09)	1.46 (0.21)
R-S+	4.80 (0.18)	4.79 (0.21)	4.56 (0.25)
R-S-	4.83 (0.11)	4.77 (0.12)	1.40 (0.23)

图 2: 实验语料测评结果

在各筛选标准的限制下，部分条件的语料始终未能搜集齐全，为确保条件间的一致性，我们在每种条件中都选取了3 对词频、熟悉度和语义相关度得分都相对匹配的词对，在实验过程中这些12 个词对呈现了2 次，其他词对只呈现一次。为尽可能避免重复效应的可能影响，相同的词对被设置为在不同的list 呈现，且在block 之间有短暂休息。语料筛选完毕后，将所选词语录制为音频形式。所选录音人员为播音主持专业的硕士研究生，经过专业的普通话发音训练，普通话水平达到一级乙等水平。实验材料的音频录制工作在隔音良好的专业配音实验室进行，以减少环境噪音、回声等可能干扰。所有实验材料全部由一位录音人员在连续状态下录制，最大限度避免音色、音量、发音节奏等变化对被试的听觉加工过程产生影响。音频录制使用Adobe Audition 2022 软件进行，录制参数为22100 Hz，单声道，位深度为16 bit。录制完毕后将完整音频分段，每个词语形成单独的WAV 格式音频文件。使用Praat 软件对每个音频进行对齐，使各音频的发音位置为音频的起始位置。每个词语的发音时长约为1000ms 左右，对时长短于900ms 和时长长于1100ms 的音频进行了时长调整，使用Praat 软件对整个音频文件进行了等比例的时长微调，确保调整后的音频时长在900 1100ms 范围内，且听感流畅自然。在每位被试完成实验后，主试对音频质量等进行了回访，被试一致表示音频清晰自然，时长适中，音量

(c) 2025 中国中文信息学会计算语言学专业委员会

合适，能够快速反应出所读的词语是什么。在语料测评和录制阶段，共有55 名志愿者参与，他们在正式实验前接触过实验语料，因而没有参与到正式实验中。

类型	启动词熟悉度	目标词熟悉度	语义相关度
R+S+	4.81 (0.11)	4.78 (0.13)	4.34 (0.28)
R+S-	4.86 (0.08)	4.81 (0.09)	1.46 (0.21)
R-S+	4.80 (0.18)	4.79 (0.21)	4.56 (0.25)
R-S-	4.83 (0.11)	4.77 (0.12)	1.40 (0.23)

图 3: 实验语料音频时长统计结果

2.3 实验程序

本研究采用被试内设计，参照Thierry等（2007）的实验方法，但将刺激呈现的方式改为听觉呈现。给被试佩戴入耳式耳机，并在实验前调试音量、佩戴位置等，保证实验过程中声音清晰且倾听舒适。实验时被试手拿按键盒进行反应。被试需端坐在屏幕前保持适当距离，实验过程中避免频繁眨眼或是产生较大的身体动作。实验程序编制使用的是E-prime3.0 软件。程序分为3 个list，每个list 包含53 或54 个词对，在每个list 结束后让被试休息三分钟。实验刺激以伪随机方式呈现。在每个试次中，先呈现500ms 的“+”符号作为注视点，提醒被试集中注意力，然后呈现启动词，在语料录制和剪辑阶段，我们已经对启动词和目标词的时长进行了控制，因此每个听觉词汇呈现的时长相似，在听感上几乎无差异。启动词呈现结束后，经过500ms 的间隔时间开始呈现目标词。被试在听到目标词后，尽可能快速、准确地进行判断任务，即判断先后呈现的两个词语在意义上有没有相关性。相同的答案最多连续出现三次。被试完成判断后，立即进入下一个试次，超过2000ms 没有完成判断的试次将直接结束，并跳转到下一个试次。在每个试次内，屏幕上将一直呈现“+”符号，在被试进行判断后，“+” 消失，并在下一试次重新出现。具体实验流程见图4。在每个list 开头呈现指导语，被试阅读并明确实验任务后，进行简短的练习，练习语料不出现在正式实验中。为减少不同按键反应对语言认知加工过程产生影响，按键答案在不同被试间平衡。

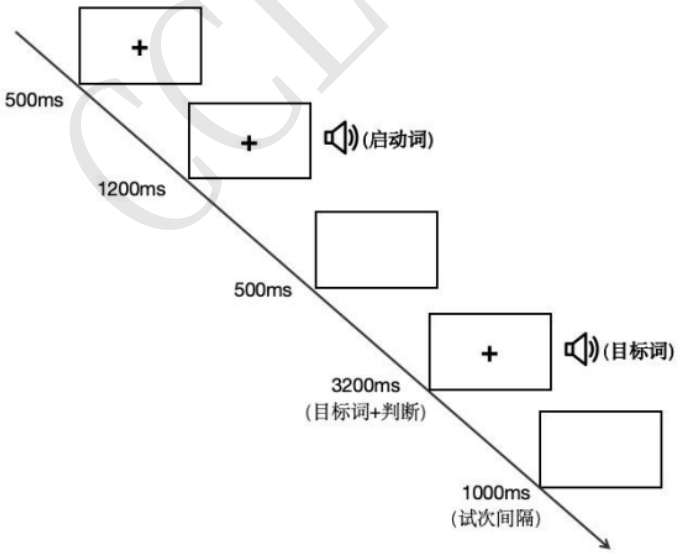

图 4: 实验流程图

2.4 脑电采集和记录

数据采集工作在标准的隔音实验室进行，实验人员使用事件相关电位技术记录被试任

(c) 2025 中国中文信息学会计算语言学专业委员会

务过程中的脑电情况，该技术具有安全无创、时间分辨率高等优势。该实验室使用的是美国Neuroscan 公司生产的ESI-256 脑电记录系统，以及该系统配套的64 导Ag/AgCl 电极帽来进行数据采集，电极位置按照国际10-20 系统来放置。实验程序采用E-prime 3.0 编写。在实验前，被试被要求仔细清洗头皮并吹干，正确佩戴帽子后，由主试向各电极位置注入导电膏。脑电的在线采样率为1000Hz，滤波带通为0.05—100Hz。由于材料以听觉形式呈现，在线参考使用了鼻尖位置。接地电极GND 位于大脑中线上FZ 和FPZ 两个电极点的中点。同时，我们在左眼瞳孔的垂直位置记录了垂直眼电，在两侧眼睛的外眼角位置记录了水平眼电。在整个实验过程中，头皮与电极之间的阻抗将被始终保持在5kΩ以下。

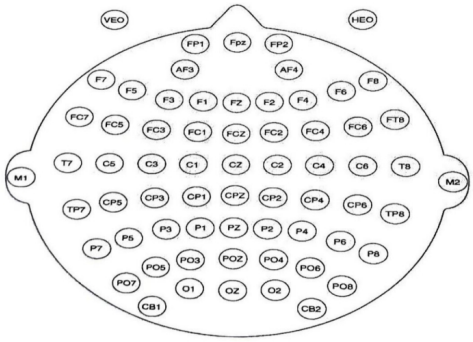

图 5: 导电极分布图

2.5 数据处理与分析

数据分析使用matlab 2017 和eeglab 2023.1 工具包进行。对分段采集的脑电数据进行拼接、融合后进行电极定位，去除VEO、HEO、M1、M2、CB1、CB2 等无用电极。按照0.1—30Hz 进行带通滤波，并进行48—52Hz 的凹陷滤波。将采样率降低为500Hz 后，手动去除漂移、伪迹，并对采集质量不佳的坏导进行差值计算。运行独立成分分析（independent component analysis，ICA）运算并据此去除眼电、肌电等干扰成分。在分段步骤中，记录和分析了每个试次目标词呈现-100 1000ms 时间窗内的数据，目标词呈现前100ms 的时间窗用于基线校正，并运用绝对阈值法剔除信号幅度超出±80μV 的信号。对每位有效被试的数据进行预处理后，使用eeglab 工具包中的erplab 插件分配实验相关的事件码，并导出单个被试每个条件多个试次下的脑电加工情况。将不同被试的脑电数据进行组水平叠加，绘制出被试在四种条件下的平均脑电波形图，根据波形图和实验需求选取合适的时间窗绘制脑地形图。最后，我们提取出了四种条件下所诱发的脑电波在感兴趣时间窗和电极点的幅值、潜伏期等量化数据。为考察大脑半球和区域对实验条件的影响，我们将大脑分为左、右半球，将所有脑区分为前部、中部、后部三个部分，由此将所有电极点划分为多个感兴趣区，即左前脑区（F1，F3，FC1，FC3），右前脑区（F2，F4，FC2，FC4），左中脑区（C1，C3，CP1，CP3），右中脑区（C2，C4，CP2，CP4），左后脑区（P1，P3，P5）和右后脑区（P2，P4，P6），以及位于头皮中线的前、中、后部电极点（FZ、CZ、PZ）我们计算出了这些感兴趣区内各个电极点的平均幅值。借助SPSS 26.0 软件，我们对收集到的行为数据和脑电数据进行了重复测量方差分析，使用Greenhouse-Geisser 法进行p 值校正，p 值低于0.05 的数值被定义为具有统计学意义的显著效应，并进一步进行简单效应分析，根据分析结果形成报告。

3 实验结果

3.1 行为数据结果

实验过程中，我们主要记录了被试在听到目标词后进行语义判断任务的反应时和正确率作为主要的行为数据指标。共160 个试次，每种条件下40 个试次。各条件下被试进行语义判断任务的行为数据见图6。对于四种条件下被试的平均反应时和正确率，我们进行了语音特征（尾音相同、尾音不同）× 语义特征（语义相关、语义无关）的两因素重复测量方差分析。对正确率的分析结果显示，语音特征主效应显著，F（1，35）=6.590，$p < 0.05$，η_p^2=0.158，

(c) 2025 中国中文信息学会计算语言学专业委员会

尾音不同条件下的正确率高于尾音相同条件。语义特征主效应显著，F（1，34）=35.744，$p < 0.001$，η_p^2=0.505，语义无关条件下的正确率高于语义相关条件。语音特征× 语义特征交互效应显著，F（1，35）=15.962，$p < 0.001$，η_p^2=0.313。简单效应分析显示，无论尾音是否相同，语义无关条件下的正确率都显著高于语义相关条件（尾音相同条件下$p < 0.001$，尾音不同条件下p=0.001）。只在语义相关条件下，尾音不同时的正确率显著高于尾音相同条件，p=0.001，语义无关条件下，正确率差异不显著，p=0.212。

类型	正确率（%）		反应时（ms）	
	平均值	标准差	平均值	标准差
R+S+	93	7	1288.95	107.09
R+S-	99	1	1298.93	94.10
R-S+	96	4	1191.38	106.90
R-S-	99	2	1290.21	91.39

图 6: 实验行为数据

对反应时数据的分析结果显示，语音特征的主效应显著，F（1，35）=31.052，$p < 0.001$，η_p^2=0.470，尾音相同条件下被试的反应时显著长于尾音不同条件。语义特征的主效应显著，F（1，35）=14.830，$p < 0.001$，η_p^2=0.298。语义无关条件下被试的反应时显著长于语义相关条件。语音特征× 语义特征交互效应显著，F（1，35）=6.527，$p < 0.05$，η_p^2=0.157。在语义相关条件下，被试对尾音相同词对的反应时显著长于尾音不同词对，$p < 0.001$。在语义无关条件下，反应时差异不显著，p=0.089。尾音相同条件下，反应时差异不显著，p=0.086。尾音不同条件下语义无关条件的反应时显著长于语义相关条件下。

3.2 ERP数据结果

我们记录了目标词呈现前100ms 至呈现后1000ms 的脑电反应数据。根据前人的研究结果结合实验结果绘制的波形图，选取了听觉脑电实验常关注的N100，P200，早期N400，及晚期N400 几个脑电成分进行重点分析。N100 成分的时间窗为80-120ms，P200 成分选定的时间窗为150-250ms，N400 成分选定的时间窗为300-500ms，晚期N400 成分选定的时间窗为500-700ms。针对各个选定的时间窗，我们将进行语音特征（2 水平：尾音相同、尾音不同）× 语义特征（2 水平：语义相关、语义无关）× 脑区（3 水平：前部、中部、后部）× 半球（2 水平：左半球、右半球）的多因素重复测量方差分析。

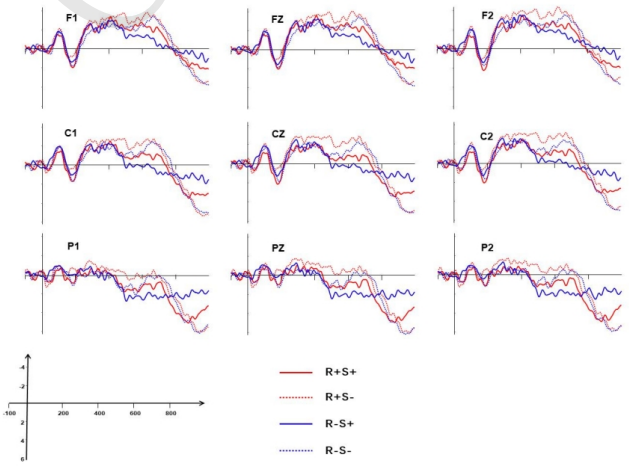

图 7: 四种条件下听觉加工的ERP 总波形图

(c) 2025 中国中文信息学会计算语言学专业委员会

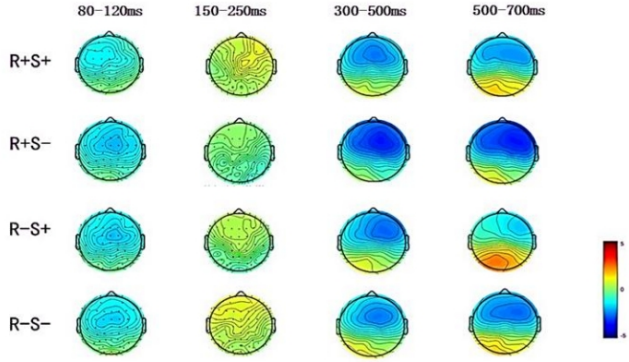

图 8: 四种条件下听觉加工的ERP 总地形图

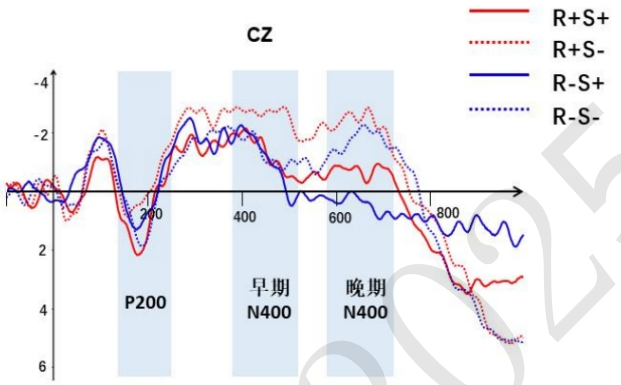

图 9: CZ 电极点的ERP 总波形图

3.2.1 尾音相似未对N100 波幅产生显著影响

在80-120ms 的时间窗下，我们进行了语音特征（2 水平：尾音相同、尾音不同）×语义特征（2 水平：语义相关、语义无关）× 半球（2 水平：左半球、右半球）× 脑区（3 水平：前部、中部、后部）的多因素重复测量方差分析。结果显示语音特征主效应不显著，F（1，34）=0.050，p = 0.824。语义特征主效应不显著，F（1，34）=0.241，p = 0.627。半球主效应不显著F（1，34）=0.209，p = 0.650。

脑区主效应显著，F（2，68）=32.543，p < 0.001，η_p^2=0.489。前部脑区诱发的波幅比中部更负，p < 0.05，前部脑区的波幅比后部脑区更负，p < 0.001，中部脑区的波幅比后部脑区更负，p < 0.001。语音特征× 半球×脑区交互效应显著，F（2，68）=3.459，p < 0.05，η_p^2=0.092。简单效应分析结果显示，在尾音不同条件下，左半球前部比后部显著更负，p < 0.001，中部脑区幅值比后部脑区更负，p < 0.001，但前部和中部脑区的幅值没有出现显著性的差异。在右半球，前部脑区幅值显著高于中部脑区，p < 0.001，前部脑区幅值显著高于后部脑区，p < 0.001，中部脑区幅值显著高于后部脑区，p < 0.001。在尾音相同条件下，在左半球，前部脑区幅值显著高于中部脑区，p < 0.05，前部脑区幅值显著高于后部脑区，p < 0.001，中部脑区幅值显著高于后部脑区，p < 0.001。在右半球，前部脑区幅值显著高于中部脑区，p < 0.05，前部脑区幅值显著高于后部脑区，p < 0.05，中部脑区幅值显著高于后部脑区，p < 0.001。

3.2.2 尾音相似未对P200 波幅产生显著影响

在150-250ms 的时间窗下，我们进行了语音特征（2 水平：尾音相同、尾音不同）× 语义特征（2 水平：语义相关、语义无关）× 脑区（3 水平：前部、中部、后部）× 半球（2 水平：左半球、右半球）的多因素重复测量方差分析。

(c) 2025 中国中文信息学会计算语言学专业委员会

语音特征主效应不显著，F（1，34）=0.258，p <= 0.614。语义特征主效应不显著，F（1，34）=0.219，p = 0.643。半球主效应不显著F（1，34）=0.021，p = 0.886。脑区主效应显著，F（2，68）=3.718，p < 0.05，η_p^2=0.099。语义特征× 半球交互效应显著，F（1，34）=5.648，p < 0.05，η_p^2=0.142，简单效应分析后未发现显著差异。语音特征× 半球× 脑区交互效应边缘显著，F（2，68）=3.114，p = 0.051p=0.051，η_p^2=0.084。尾音不同条件下，左半球前部脑区平均幅值显著高于中部脑区p < 0.05；在右半球中部脑区的平均幅值比后部脑区显著更高，p < 0.05。

3.2.3 尾音相似对早期N400 成分有影响

在300-500ms 的时间窗下，我们进行了语音特征（2 水平：尾音相同、尾音不同）× 语义特征（2 水平：语义相关、语义无关）× 脑区（3 水平：前部、中部、后部）× 半球（2 水平：左半球、右半球）的多因素重复测量方差分析，结果显示语音特征主效应不显著，F（1，34）=2.688，p = 0.110。语义特征主效应不显著，F（1，34）=1.491，p = 0.203。

半球主效应显著F（1，34）=10.849，p < 0.01，η_p^2=0.242，右半球产生的波幅更负。脑区主效应显著，F（2，68）=76.825，p < 0.001，η_p^2=0.693，前部脑区诱发的负波幅值显著高于中部脑区，p < 0.001，前部脑区幅值显著高于后部脑区，p < 0.001，中部脑区幅值显著高于后部脑区，p < 0.001。语音特征× 语义特征× 半球交互效应显著F（1，34）=4.275，p < 0.05，η_p^2=0.112。在尾音相同语义相关条件下，左右半球幅值没有出现显著差异。尾音相同语义无关条件下，半球间幅值差异显著，p = 0.001。尾音不同语义相关条件下，半球间幅值差异显著，p < 0.05。尾音不同语义无关条件下，半球间幅值差异显著，p < 0.05。在语义无关条件下，在左半球，尾音相同与不同条件诱发的波幅差异不显著，p = 0.077。在右半球，尾音相同与不同条件诱发的波幅差异边缘显著，p = 0.051，尾音相同条件诱发的负波波幅比尾音不同条件更大。而在语义相关条件下，两半球中尾音相同与不同条件诱发的波幅差异都不显著。在尾音相同条件下，右半球提取的脑电波中，语义无关条件产生的脑电波比语义相关条件下更负，p < 0.05。左半球提取的脑电波无显著差异。尾音不同条件下，无论左半球还是右半球，语义相关和无关条件的脑电波幅没有显著差异。

3.2.4 尾音相似对晚期N400 成分影响显著

在500-700ms 的时间窗下，我们进行了语音特征（2 水平：尾音相同、尾音不同）× 语义特征（2 水平：语义相关、语义无关）× 脑区（3 水平：前部、中部、后部）× 半球（2 水平：左半球、右半球）的多因素重复测量方差分析。

语音特征主效应显著，F（1，34）=8.072，p < 0.01，η_p^2=0.192。尾音相同条件下波幅比尾音不同条件的波幅更负。语义特征主效应显著，F（1，34）=14.510，p < 0.001，η_p^2=0.299。语义无关条件下波幅比语义相关条件更负。半球主效应显著，F（1，34）=15.347，p < 0.001，η_p^2=0.311，右半球诱发的波幅比左半球诱发的波幅负性更强。脑区主效应显著，F（2，68）=137.826，p < 0.001，η_p^2=0.802。前部脑区诱发的负波波幅比中部更高，p < 0.001，前部脑区诱发的幅值比后部脑区更高，p < 0.001，中部产生的负波比后部诱发的波幅更高，p < 0.001。

语音特征× 语义特征× 半球交互效应边缘显著，F（1，34）=3.339，p = 0.076。在尾音相同的情况下，在左半球，语义相关和无关条件诱发的负波波幅差异不显著，p = 0.063；在右半球差异显著，p < 0.05。在尾音不同调节下，在左半球，语义相关和无关条件诱发的负波波幅差异显著，p < 0.05。在右半球，语义相关和无关条件诱发的负波波幅差异均显著，p < 0.05。在语义相关的情况下，在左半球，尾音相同与不同条件诱发的波幅差异显著，p < 0.05，在右半球，尾音相同与不同条件诱发的波幅差异不显著，p = 0.059。

4 分析与讨论

4.1 行为数据分析

对行为数据的分析结果显示，在正确率方面，所有被试在四种条件下都达到了85%以上，表明被试认真完成了规定的语义判断任务。针对四种条件下正确率的方差分析结果显示语音特征主效应、语义特征主效应及其交互效应都产生了显著的差别。其中，尾音相同语义不相关条

(c) 2025 中国中文信息学会计算语言学专业委员会

件下的正确率最高，尾音相同语义相关条件下的正确率最低。尾音相同条件下正确率略低于尾音不同条件，语义相关条件下的正确率低于不相关条件。交互效应表明不论尾音是否相同，语义无关词对的判断正确率都比语义相关词对更高。且在语义无关条件下，尾音是否相同没有对正确率产生影响，但在语义相关的词对中，尾音信息产生了影响。表明任务过程中，被试不仅对词汇意义进行了思考和匹配，而且对语音特征进行了一定的判断，但这种认知过程更多地集中在语义上。

从反应时来看，语音特征与语义特征的主效应和交互效应仍然显著。尾音相同语义不相关条件下的反应时最长，尾音不同语义相关条件下的反应时最短。尾音相同条件下反应时比尾音不同条件下的反应时更长，语义不相关条件下的反应时比语义相关条件下的反应时更长。结合正确率的结果可见，与语义相关词对相比，被试对语义无关词对进行判断的正确率更高，但同时反应时更长。这可能是任务特性导致的，由于本实验使用的是语义判断任务，对于显而易见相关的词对，被试可以快速而准确地作出判断，但不相关词对需要从多种词汇语义关系上进行思考和判断，被试的判断过程可能采取了放慢速度保证正确率的加工策略。同时，与尾音不同条件相比，尾音相同条件下的判断正确率更低，反应时更长，表明词对之间的尾音联系增加了理解和判断词汇意义的难度，使这一过程时间变长，也耗费了更多认知资源。由于行为数据是听和理解词语的完整过程相融合的结果，我们将结合脑电数据进一步观察和解析人们听和理解熟悉的双音节词的具体过程。

4.2 ERP数据分析

根据词对之间的音义关系，我们关注尾音信息和意义关系如何分别影响双音节词的听觉加工过程。我们选取了目标词出现前100ms 至出现后1000ms 的时间窗，对被试的语言加工理解过程进行了记录，根据绘制的波形图，结合各个时间窗内提取出的波幅具体数值，我们对任务诱发的四个脑电成分进行了分析。

4.2.1 语音早期对尾音相似的注意敏感度较低

在本实验中，被试在听觉刺激输入早期就产生了N100 成分，表明对语音刺激产生了注意，表明对语音刺激具有较强的敏感性。脑区分析显示，前部脑区产生的负波波幅最大，到中部和后部依次减弱。由于这些语音材料在录制和剪辑的过程中将音色、音量、时长、频率等控制在统一的范围内，在语音的典型特征上并没有明显的差异，而N100 成分受到声音的物理属性的影响比较大。因而在加工的早期阶段，这些刺激材料还不足以激发被试的敏感性，察觉到语料之间的差异。

波形图显示，在200ms 左右，四种条件的语音刺激都诱发了向下的正波。该成分也是语音早期加工的重要指标（Dorman, 1974），反映了早期注意、词汇语音、语音辨别等早期认知过程。Huang 等人（2014）认为P200反映了听觉模态下语音信息在短时记忆中的暂留，因而与跨模态研究相比，被试在听觉单通道模式下能更加直接地感知到启动词与目标词之间的语音相关性，从而减小认知加工的负荷，有助于听觉言语理解的认知过程。前人文献中报告的相对于控制条件，同音启动和尾音启动激发了更大的P200 成分的结果，表明P200 成分可能对语音相似性敏感（Kong et al., 2010）。且同音条件下还引发了语义激活，而尾音相同条件下，语音片段的重叠没能产生与完全相同的音节相同的效应，未能引发词对之间的语义激活。

然而，前人研究中使用的材料以单音节词为主，且在视觉模态下进行，为了避免正字法的可能影响，编创语料的过程中筛除了相同声符和义符的词对组合，以更纯粹地观察语音相似性的影响。但在我们的研究中，听觉模态的双音节词随着时间展开，在P200 成分的时间窗下，相同的尾音还没开始呈现，因而无法在150-250ms的时间窗内观察到语音片段相似性的影响。此外，在Huang等人（2014）的经典研究中，语义操控没能引发P200 的变化，表明该成分可能受到加工过程的影响比较小。这与我们实验发现的P200 效应是一致的。

已有多项研究（Huang et al., 2014; Huang et al., 2016; Chen et al., 2016）发现指示早期语音加工的P200 成分会受到语音相似性影响，语音不相似项目产生的P200 正波幅度比语音相似项目产生的波幅更大。在本实验中虽然出现了P200 成分，但相似的尾音信息并未对其波幅产生显著影响。对比这些研究的异同，我们认为双音节词音频材料的时长较长是造成语音启动效应与前人相反的主要原因。本研究中使用的双音节词音频时长为1000ms 左右，在P200 时间窗（150-250ms）关键的尾音信息还未出现，因而相同的尾音诱发的正成分很可能延迟出现，而大量文献已经证实300-500ms 的时间窗语义加工已经开始（Kutas et al., 1980;Kutas et al., 1984;

(c) 2025 中国中文信息学会计算语言学专业委员会

Kutas et al., 2000; Kutas et al., 2010），因此在早期N400 时间窗内可能有多个加工过程同时在进行，语音和语义等因素的影响叠加，共同影响着脑电波的波幅和地形分布。虽然早期语音加工的P200 成分会受到语音相似性影响的具体机制尚不明晰，但形同尾音引发的主效应至少提示我们末尾位置的语音信息引起了大脑的注意。

4.2.2 早期N400受尾音相似性的调节

N400 是经典的反映语义加工相关过程的脑电成分（Kutas et al., 1980; Kutas et al., 2011）。在400ms 左右，四种条件下都出现了向上的负波成分，我们提取了300-500ms 的平均波幅值进行的统计结果显示，语义不相关情况下，被试诱发的N400成分波幅显著大于语义相关的情况。该结果与众多经典的语义认知加工研究结果相一致，语义违反或不一致情况下，被试理解意义的认知资源消耗增加，反映了语义的不匹配造成词汇加工难度增加。由于实验中使用的语义判断任务中，所有语料都使用了相同词性的名词，且避免了启动词和目标词可以进行句法组合的情况，因此词对之间的语义关联比句子整合或短语整合的情况下更弱，因而语义不一致产生的违反效应也会比句子理解实验中意义整合违反产生的N400 效应更小。

尾音相同词对和尾音不同词对诱发的N400 成分波幅有显著差异，该成分在前部脑区最显著，右半球高于左半球，表明相同的尾音信息引发了听者的注意。大量使用ERP 技术（Rugg et al., 1984a, b; Praamstra et al., 1994; Dumay et al., 2001; Chen et al., 2016; Desroches et al., 2009; Liu et al., 2006）和眼动追踪技术（Zhao et al., 2011; Zou et al., 2022; Zou, 2017; Zou et al.,2019; Wang et al., 2021）的研究也证实了听者能够对尾音片段信息表现出较强的神经响应，本研究的结果与这些研究的结果相一致。

前人研究中很多都报道了在400-500ms 出现了相同的尾音信息诱发的负波，尾音不同词对比尾音相同词对诱发更大的负向波幅（Dumay et al., 2001；Praamstra et al., 1994）。但大部分研究采用的实验任务是押韵判断任务，显性地测试押韵特征加工的过程。同时，也有很多研究报道了任务类型对加工过程的影响，即在显性的语音任务中，会出现语音N400 效应，（Perrin et al., 2003; Yoncheva et al., 2013; Yoncheva et al., 2014）。表明自上而下的加工过程调控着听觉加工的具体过程，在不同的任务需求下，语音加工者将注意力集中在词语的不同方面，反映在脑电波上即出现了N400 效应的差异。

在本实验中，我们使用的语义判断任务没有显性的语音加工需求，因此，即使词对之间的语音关联产生了启动效应，与语义启动的激活强度相比较，也会弱很多。在N400 时间窗内，左右两半球的波幅值开始产生显著性差异，且右半球波幅值显著高于左半球。这可能是因为右半球在加工语音特征和声调等韵律特征方面更为突出，相同尾音的存在使右半球在本实验的任务加工中更加活跃。

4.2.3 晚期N400 语义整合受尾音相似影响

晚期N400 成分通常出现在词汇语义加工的后期阶段，与语义相关的认知过程有关，如语义理解、词汇检索和语义记忆等。尾音相同的效应直到500-700ms 的时间窗内才开始显现，尾音相同词对产生的晚期N400 成分波幅比尾音不同词对诱发的晚期N400 波幅更大。因此我们认为，该成分是由语义加工过程诱发的。语音主效应表明在该时间窗内，被试对相同的尾音具有较强的敏感性。词对之间的语音关联可能使被试将更多注意力投注到语音层面上，导致用于语义判断的认知资源被分散，从而增加加工难度。该结果与行为结果也较为一致，尾音相同语义无关的情况下，被试进行判断任务的平均反应时最长，表明当语音的相关性与语义相关性产生冲突时，任务难度增加。

同时，语义主效应显著，语义无关词对诱发的晚期N400 波幅比语义相关词对更大。这符合大量文献中报告的语义违反增加认知负荷，导致N400 波幅增加的结果。与300-500ms 相比，晚期N400 成分可能更多地涉及了语义加工的过程，这一方面是因为听觉刺激材料随着时间而完整呈现，另一方面是因为语音加工比较早，而语义通达发生得比较晚，在500-700ms 的时间窗下，语义加工进入重要阶段，词义通达、记忆和进行比较判断任务大量消耗着认知资源。地形图上也反映了这一特点，在500-700ms 的全脑平均地形图上，脑区激活的程度比前期阶段和中期阶段更大，表明认知和思考过程主要发生在后期阶段，该过程大量耗费着大脑的认知资源。

尾音相同条件下的N400 波幅显著高于尾音不同条件$p < 0.05$，与Cutler（1990）首音效应方向相反。我们推测这可能与汉语韵律特征有关。汉语作为声调语言，尾音承载关键语义信息，而英语等重音语言中首音更为关键。当大脑判断一对词为押韵后，就会调用更多的认知资

(c) 2025 中国中文信息学会计算语言学专业委员会

源来对押韵这一特征进行分析，从而导致N400的波幅升高，这是语义判断和押韵判断叠加的结果，同时也是二者进行资源竞争的结果，因此在尾音相似的条件下语义启动效应会有所减弱。这是声调语言所独有的特点。汉语中声调具有范畴化作用，汉语中同一音节的不同声调可能对应完全不同的词义，如“ma”可表“妈”“麻”“马”“骂”。尾音的声调变化直接关联语义分化，导致大脑对尾音的声学特征，如基频曲线、时长更敏感，从而导致相同的尾音消耗了更多的认知资源。Maggu等（2021）通过声调语言粤语群体的绝对音高AP研究，揭示了大脑对声调语义的编码机制、尾音声学特征的神经敏感性，以及韵律模式对语义权重的强化作用。

5 结论与展望

本研究表明听者对尾音信息表现出强烈的神经响应，表明在汉语双音节词的听觉加工中，词尾的信息部分激活稳定和强烈，这可能意味着尾音信息在汉语词汇识别中通过预测编码机制发挥关键作用，值得注意的是，尾音信息的强神经响应可能与其在语音工作记忆（phonological working memory）中的存储功能相关——词尾作为韵律边界标记，可能通过增强语音表征的稳定性，促进后续语义提取。

尽管本实验采用高时间分辨率的ERP技术，利用高时间分辨率捕捉到人脑当中非常细微的变化，同时尽可能全面的排除额外因素的干扰，比如通过方言筛选控制了语言经验对语音加工的影响，但不得不承认，仍存在若干潜在混淆因素需进一步探讨,而这些因素都极有可能影响本次实验的结果。首先性别差异可能影响语言网络的神经表征（Lindell and Lum ,2008;Wang et al.,2011），但是我们的实验并没有将这一因素考虑在内，男女在尾音和语义的不同关系中应该会有细微的差别；其次音乐训练与阅读能力可能分别通过听觉敏感性和语义网络的强化间接调节语音处理，Maggu等（2021）认为音高和语调的感知能力是和词汇通达成正相关的，而在Knapp（2020）的实验当中又证实了词的邻域密度越高，词汇提取越快越准的特点，验证了语音信息在词汇加工早期阶段的重要性，这些都会影响被试的反应时和N400的波幅；最后样本量较小（N=39）也可能限制统计效力，扩大样本的容量才能更进一步的得出符合实际特点的结论，才能提高对特定群体认知策略差异的认识。尽管贝叶斯框架的初步模拟支持尾音的预测编码假说，但注意漂移与节奏效用的潜在干扰仍需通过实验设计优化排除。在以后的研究当中可以从以上几个方面来完善实验设计，从而更准确的揭示汉语母语者对尾音处理的神经机制。

一般认为大脑会在已经获得的语音信息基础上预测即将出现的语音内容，比如说Micha Heilbron等（2022）通过结合深度神经语言模型（GPT-2）与脑电或脑磁成像技术，揭示了语言理解中多层预测机制的存在,同时强调高层级语义预测对底层语音输入的“自上下”调控，但本研究通过尾音相似效应发现，底层感知线索如声调、韵母可直接触发语义预激活，表现为在尾音相似的条件下，语义相关的N400波幅降低，表明语音特征可直接参与语义竞争，这和“语音输入需经层级过滤”的假设不同，挑战了层级预测的框架。这可以为语言处理模型如预测模型提供新的证据，进一步深化模型框架，也为新的研究提供了新思路,促进汉语计算机语言模型的发展。

在汉语这类声调语言中，尾音的声调模式如升调或降调可以激活语义网络，王露等（2018）通过眼动实验研究汉语口语词汇识别中声调、语境的作用，发现被试对竞争词的整体匹配更关注，音段信息和声调在识别过程中均起作用。这表明声调信息参与了词汇识别，这与英语等重音语言的尾音处理机制存在显著差异。在本研究中，我们进一步揭示了汉语语音在人脑中的处理机制，尤其是补充证明了汉语母语者对尾音加工的特异性，而计算机语言模型的进一步改善依赖于人们对于语言处理的生理机制的探索已经得到证实：Tikochinski（2025）聚焦于大语言模型（LLMs）与人类大脑在语言语境整合机制上的差异，进一步表明计算机语言模型对于人脑处理机制的模拟的重要性。因此，深化对人脑对语音加工机制的认识有助于为改进计算机语言模型对人脑语言处理机制的模拟效能。

最后，本研究通过ERP技术揭示了尾音信息在汉语词汇加工中的核心地位，但其具体的神经机制仍需计算模型的进一步验证，或者使用FMRI来直接观察尾音相似与否的脑区激活状况，为尾音在汉语母语者当中的重要性提供神经机制的支持。也可应用跨语言与多模态的计算模拟，采用视觉和听觉两种不同感觉通道的刺激进行押韵任务，这将深化我们对声调语言韵律加工普适性与特异性的理解。

(c) 2025 中国中文信息学会计算语言学专业委员会

参考文献

Marslen-Wilson W D, Welsh A. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology*, 10(1):29-63.

McClelland J L and Elman J L. 1986. The TRACE model of speech perception. *Cognitive psychology*, 18(1): 1-86.

Luce P A and Pisoni D B. 1998. Recognizing spoken words: The neighborhood activation model. *Ear and hearing*, 19(1): 1-36.

Luce P A and Pisoni D B. 1986. Neighborhoods of words in the mental lexicon. *Perception Psychophysics,*, 39(3), 155–158.

Knapp, D. L. 2020. Manipulating phonological density during a lexical decision task: An ERP investigation. University of South Alabama. ProQuest Number: 28026814.

Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 4171–4186.

Feng, C., Damian, M. F., and Qu, Q. 2021. Parallel processing of semantics and phonology in spoken production: Evidence from blocked cyclic picture naming and EEG. *Journal of Cognitive Neuroscience*, 33(4): 725–738.

Goldstein, A., Zada, Z., Buchnik, E., Schain, M., Price, A., Aubrey, B., ... Hasson, U. 2022. Shared computational principles for language processing in humans and deep language models. *Nature Neuroscience*, 25(3): 369–380.

Radford, A., Narasimhan, K., Salimans, T.,and Sutskever, I. 2022. Language models are unsupervised multitask learners. *OpenAI Blog.*,

Bisk, Y., Andreas, J., Holtzman, A., Bengio, Y., Chai, J., Thomason, J., ... and Turian, 2020. Experience grounds language. arXiv preprint arXiv:2004.10151.

Kutas M，and Federmeier K D. 2000. Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, 4(12): 463-470.

Caucheteux, C., Gramfort, A., and King, J.-R. 2023. Evidence of a predictive coding hierarchy in the human brain listening to speech. *Nature Human Behaviour*, 7: 430–441

Cutler, A. 1990. Exploiting prosodic probabilities in speech segmentation. *Nature In Recognizing Continuous Speech*, 106-121.

Federmeier, K. D., and Kutas, M. 2001. Meaning and Modality: Influences of Context, Semantic Memory Organization, and Perceptual Predictability on Picture Processing. *Nature Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1): 202 - 224.

Friederici, A. D. 2004. The neural basis ofsyntactic processes. *Nature NeuroImage*, 3ed,789–801.

Friederici, A. D. 2011. The neural basis of language processing: From structure to function. *Physiological Reviews*, 91(4): 1357–1392.

Pierrehumbert, J. 1980. *The phonology and phonetics of English intonation.*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Hyman, L. M. 2007. *How (not) to do phonological typology: The case of pitch-accent.*. UC Berkeley Phonology Lab Annual Report, 3(3): 654–685.

Rugg M D. 1984a Event-related potentials in phonological matching tasks. *Brain and language*, 23(2): 225-240.

Rugg M D. 1984b Event-related potentials and the phonological processing of words and non-words. *Neuropsychologia*, 22(4): 435-443

Dumay N, Benraiss A， Barriol B,et al. 2001 Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience*, 13(1): 121-143

(c) 2025 中国中文信息学会计算语言学专业委员会

Chen Q, Zhang J, Xu X, et al. 2016 Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition*, 154: 11-21

Chen W F, Chao P C, Chang Y N, et al. 2016 Effects of orthographic consistency and homophone density on Chinese spoken word recognition *Brain and Language*, 157: 51-62.

Desroches A S, Newman R L, Joanisse M F. 2009 Investigating the time course of spoken word recognition: Electrophysiological evidence for the influences of phonological similarity *Journal of Cognitive Neuroscience*, 21(10): 1893-1906.

Liu Y, Shu H,and Wei J. 2006 Spoken word recognition in context: Evidence from Chinese ERP analyses. *JBrain and language*, 96(1): 37-48.

Zhao J, Guo J, Zhou F, et al. 2011 Time course of Chinese monosyllabic spoken word recognition: Evidence from ERP analyses *Neuropsychologia*, 49(7): 1761-1770

Zou L, Chen Y. 2019. The roles of tonal and segmental information in spoken word recognition for L2 speakers: evidence from dutch learners of mandarin. In Proceedings of the 19th International Congress of Phonetic Sciences.

Zou T, Liu Y, and Zhong H. 2022 The roles of consonant, rime, and tone in mandarin spoken word recognition: an eye-tracking study. *Frontiers in Psychology*, 12: 740444.

Wang Y, Zang X, Zhang H, et al. 2021 The processing of the second syllable in recognizing Chinese disyllabic spoken words: Evidence from eye tracking. *Frontiers in Psychology*, 12: 681337.

Maggu, A. R., Lau, J. C. Y., Waye, M. M. Y., and Wong, P. C. M. 2021 Combination of absolute pitch and tone language experience enhances lexical tone perception. *Scientific Reports*, 11(1): 1485.

Huang X, Yang J C, Zhang Q, et al. 2014 The time course of spoken word recognition in Mandarin Chinese: A unimodal ERP study. *Neuropsychologia*, 63: 165-174.

Huang X, Yang J C, Chang R, et al. 2016 Task modulation of disyllabic spoken word recognition in Mandarin Chinese: A unimodal ERP study. *Scientific reports*, 6(1): 25916.

Chen Q, Zhang J, Xu X, et al. 2016 Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition*, 154: 11-21.

Kutas M,and Hillyard S A. 1980 Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427): 203-205.

Kutas M,and 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.

Lindell, A. K., and Lum, J. A. G. 2008 Priming vs. rhyming: Orthographic and phonological representations in the left and right hemispheres. *Brain and Cognition*, 68(2): 193–203.

Heilbron, M., Armeni, K., Schoffelen, J.-M., Hagoort, P., and de Lange, F. P. 2022 A hierarchy of linguistic predictions during natural language comprehension. *Proceedings of the National Academy of Sciences*, 119(32): e2201968119.

王露，万鹏，任桂琴. 2018 汉语口语词汇识别中声调、语境的作用：来自眼动研究的证据. 心理与行为研究, 16 (2): 150–156.

Tikochinski, R., Goldstein, A., Meiri, Y., Hasson, U., and Reichart, R. 2025 Incremental accumulation of linguistic context in artificial and biological neural networks *nature communications*, 16(1), 803.

孙文梅and 刘海伦. 2013 汉语双字词语音关联对错误记忆的影响. 心理科学, 36 (6): 1388–1392.

Mitra, P., and Coch, D. 2019 An ERP study of cross-modal rhyming: Influences of phonology and orthography. *Psychophysiology*, 56(4): e13311.

(c) 2025 中国中文信息学会计算语言学专业委员会