

神经振荡视角下的第二语言听觉感知与理解

朱诗瑶 陆灵犀

北京语言大学心理与认知科学学院（生命健康研究院），北京

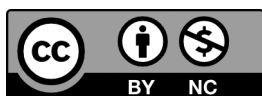
摘要 | 对第二语言（L2）学习者而言，在听觉言语加工中达到与母语者相当的速度与准确性始终是一项重要挑战。近年来，随着脑电图（EEG）和脑磁图（MEG）等具有高时间分辨率的神经电生理技术的发展，从神经振荡与神经夹带的视角研究第二语言加工过程，为揭示其神经机制提供了新的理论框架。本文系统梳理了近年来关于二语学习者听觉言语感知与理解中神经振荡的研究进展，重点总结不同振荡频段在言语加工层级中的功能作用，以及语言经验与环境因素对神经振荡模式的调节作用。现有研究表明，不同频段的神经振荡能够表征言语加工中从音素、音节到短语等不同时间尺度的信息；然而，关于母语（L1）与二语皮层追踪强度之间的关系，目前尚未形成一致结论，其差异可能受到实验范式、分析方法及被试语言经验的共同影响。此外，双语背景，尤其是二语熟练度，在安静与噪声环境下均显著调节二语听觉加工中的神经振荡模式。作为双语语言使用中的常见现象，语言转换通常引入更高的认知控制需求，其神经表征主要体现在 δ 和 θ 频段振荡功率的调节变化上。本文在此基础上综合讨论当前研究中的主要共识与分歧，并提出未来有必要结合复杂听觉情境与纵向设计，从神经振荡层面进一步深化对二语听觉言语加工机制的理解。

关键词 | 二语习得；言语加工；神经振荡；脑电/磁图

Copyright © 2026 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

<https://creativecommons.org/licenses/by-nc/4.0/>



1 引言

在当今经济与文化高度交织的全球化背景下，掌握多种语言已成为重要的社会与认知能力。在中国，有超过4亿人将英语作为第二语言（Second Language, L2）（Wei et al., 2012）。然而，大量研究表明，二语学习者在语言加工的速度和准确性方面通常难以达到母语者水平（Juffs & Harrington, 1995），这一差距在存在噪声干扰的复杂听觉环境中尤为突出，二语学习者在此类场景下表现出更为明显的语言加工劣势（Cooke et al., 2008）。

在传统神经语言学研究中，与语言加工相关的事件相关电位（Event-related Potential, ERP）分析在过去几十年中占据主导地位（Steinhauer et al., 2008; Luck & Kappenman, 2011）。近年来，随着脑磁图（Magnetoencephalography, MEG）等技术的发展，具有高时空分辨率的神经记录手段逐渐被引入语言研究领域，基于事件相关磁场（Event-related Magnetic field, ERF）的研究也开始为语言加工的神经机制提供新的证据（盛经纬, 高家红, 2021）。

与传统 ERP/ERF 研究主要关注语言诱发神经反应在

基金项目：北京市社会科学基金项目（21YJC010）。

通讯作者：陆灵犀，北京语言大学心理与认知科学学院（生命健康研究院）研究员，研究方向：语言认知神经科学。

文章引用：朱诗瑶, 陆灵犀. (2026). 神经振荡视角下的第二语言听觉感知与理解. *中国心理学前沿*, 8(1), 1-7.

<https://doi.org/10.35534/pc.0801001>

时间维度上的变化不同,神经振荡(Neural Oscillation)研究聚焦于大脑神经元群体活动的节律性、同步性与持续性特征(Buzsáki et al., 2012)。当大脑内部的神经节律受到外部刺激驱动时,神经活动的相位和振幅会与外界刺激的准节奏结构发生耦合,从而产生神经夹带(Neural Entrainment)现象(Obleser & Kayser, 2019)。在连续语音输入的情境下,神经夹带通常被视为有效言语理解的重要神经标志,其强度与听者的理解表现密切相关(Vanthornhout et al., 2018)。因此,从神经振荡与神经夹带的视角出发,为理解二语言语加工的神经机制提供了一个新的研究框架。

基于此,本文系统梳理了近年来关于第二语言听觉言语感知与理解中的神经振荡研究,重点总结影响大脑加工二语听觉言语过程的关键因素及其潜在神经基础。具体而言,本文将围绕以下几个方面展开综述:(1)不同振荡频段及其在言语加工中的功能作用;(2)大脑对母语(First Language, L1)与二语的皮层追踪强度关系;(3)影响二语听觉感知与理解中神经振荡现象的主要因素;(4)语言转换过程中的神经振荡特征。最后,对当前研究在神经机制层面存在的不足进行总结,并提出未来研究方向。

2 不同振荡频段在言语加工中的作用

在利用脑电图或脑磁图研究言语加工时,研究者通常首先记录大脑在处理连续语音流过程中的动态神经信号,并通过傅立叶变换将时域信号转换为频域信号,从而考察不同频段的神经振荡活动。现有研究主要关注五个经典频段: δ (1–4 Hz)、 θ (4–8 Hz)、 α (8–12 Hz)、 β (13–30 Hz)和 γ (30–150 Hz)频段。语言加工的不同方面可以通过不同频段的神经振荡活动模式加以表征(Bastiaansen et al., 2012; Ding & Simon, 2014)。 δ 频段振荡通常与单词加工、句法结构以及韵律线索的编码密切相关(Meyer et al., 2018; Doelling et al., 2014; Rimmele et al., 2021); θ 振荡被认为参与词汇语义的提取过程(Bastiaansen et al., 2003); β 振荡主要与句法整合和结构加工有关,而 γ 振荡则更多与语义结构的构建和精细语音信息处理相关(Bastiaansen & Hagoort, 2006; Bastiaansen et al., 2010)。

在听觉言语加工过程中,不同频段的神经振荡并非独立运作,而是通过跨频段协同编码言语刺激中不同时间尺度的层级信息,从而实现连续语音的整合和理解(Kayser et al., 2015)。神经夹带研究进一步揭示了这种多时间尺度加工机制:在自然言语流中,不同频带的神经活动倾向于与不同语言单位的时间结构发生同步。具体而言, γ 频段的神经活动往往与音素层面的快速变化相对应(Giraud & Poeppel, 2012; Zion Columbic et al., 2013; Nourski et al., 2015; Meyer, 2018), θ 频段的振荡更容易夹带音节呈现的节律(Luo & Poeppel, 2007;

Peelle et al., 2013; Meyer, 2018), 而 δ 频段的活动则倾向于追踪语音流中更宽时间窗的声学包络调制,与短语或更高层级的结构相关(Gross et al., 2013; Meyer, 2018)。

语音的皮层追踪(Cortical Tracking)通常指时间语音包络与皮层神经活动之间的相位耦合关系。对母语者而言,大脑通常能够在多个时间尺度上较为高效地追踪不同层级的抽象语言结构,如单词、短语和句子(Ding et al., 2016)。Ding等人(2016)采用频率标记范式,将汉语语音刺激中的单词、短语和句子结构分别设置为4 Hz、2 Hz和1 Hz的目标频率,结果发现汉语母语者的神经信号能够稳定地追踪这些频率成分,而不具备汉语知识的听者则无法追踪短语和句子层级的频率。随后,Getz等人(2018)在人工语言学习范式中进一步证明,当成人被试通过学习掌握了刺激中类似语言的短语结构分布线索后,其神经活动同样能够表现出对短语频率的追踪。由于刺激中并不存在明确的声学短语边界,这种皮层追踪必然反映了对抽象句法结构的内部表征。上述研究共同表明,成功的言语理解不仅依赖于声学线索的处理,还依赖于对抽象语言结构的构建,而这一过程可以通过周期性神经振荡及其层级夹带特性进行表征。

3 母语与二语皮层追踪强度的比较

已有研究对母语与二语皮层追踪强度之间的关系给出了并不完全一致的结论。Song等人(2018, 2020)发现,尽管二语学习者在整体言语理解准确性上低于母语者,但其对语音包络的神经响应反而更强。这一结果被解释为二语学习者在加工二语时需要投入更多的注意资源,以补偿其在语音感知和语言理解方面的不足。类似地,Reetzke等人(2021)也报告,与母语者相比,二语学习者表现出较弱的言语理解能力,但同时伴随增强的语音包络神经追踪,即更强的皮层追踪并不必然反映更高效的言语理解,而可能更多体现了对声学信号的增强注意或补偿性加工机制。

然而,并非所有研究均支持上述观点。Zinszer等人(2022)的研究发现,在噪声条件下,二语学习者对母语的皮层追踪显著增强,而对二语的追踪在噪声与非噪声条件之间并未表现出显著差异,且整体上弱于对母语的追踪。这一结果与Song等人(2018, 2020)及Reetzke等人(2021)的发现形成对照。研究结果的不一致可能部分源于实验设计的差异:Zinszer等人(2022)采用被试内设计,在同一批被试中比较母语与二语的加工过程;而Song等人(2018)采用被试间设计,其二语学习者长期沉浸在二语环境中,整体二语水平可能高于被试内设计中所招募的学习者。不同的语言经验背景和熟练度水平,可能直接影响大脑对噪声的敏感性。

此外,研究方法的差异也可能是导致结果分歧的重要因素。与Zinszer等人(2022)主要从神经振荡和相

位同步的角度考察皮层追踪不同, Song等人(2018)和 Reetzke等人(2021)采用多元时间响应函数方法量化神经信号对连续语音特征的响应。不同分析框架对神经数据中时间和频率信息的侧重不同, 可能导致对追踪强度这一指标的解释存在差异。综上, 当前关于, 母语与二语皮层追踪强度关系的研究尚未形成一致结论。皮层追踪强度究竟反映的是语言理解效率, 还是注意资源分配与补偿性加工, 仍需在将来研究中结合更精细的实验设计和多方法分析加以澄清。

4 影响二语听觉理解与感知的神经振荡调节因素

综合现有研究, 双语背景, 尤其是二语熟练度, 是影响二语听觉理解与感知过程中神经振荡模式的核心因素之一。此外, 在复杂听觉场景中, 噪声的存在及其类型也会显著改变学习者对二语的加工方式, 并在神经振荡层面体现出来。基于此, 本节从噪声类型和语言熟练度两个方面, 系统梳理影响二语理解与感知中神经振荡现象的主要因素。

4.1 噪声类型

在复杂的声学环境下倾听和理解连续语音流是日常交流中的常见情境。为在多声源环境中选择性地关注目标语音, 听觉系统需要从复杂的声学输入中提取关键特征, 以实现目标语音与干扰信号的分离(Bregman, 1994)。这一过程显著增加了语音感知的认知负荷(Mattys et al., 2012), 且对二语学习者的影响尤为明显(Meador et al., 2000)。

大量行为研究表明, 噪声类型(Cooke et al., 2008; Brouwer et al., 2012; Kilman et al., 2014)以及语境信息(Bradlow & Bent, 2002; Shimizu et al., 2002; Aydelott & Bates, 2004; Aydelott et al., 2012; Van Engen, 2010; Brouwer et al., 2012)都会显著影响二语学习者在噪声中的语音感知表现(Broersma & Scharenborg, 2010; Cooke et al., 2010)。Kilman等人(2014)比较了能量掩蔽与信息掩蔽对二语言语加工的影响, 发现相较于能量掩蔽, 信息掩蔽条件下学习者表现显著更差, 表明背景噪声中可理解信息的干扰对二语加工尤为不利。

在神经层面, Song等人(2020)进一步考察了不同噪声类型(单一说话者噪声与多人混合噪声)对母语者和二语学习者言语感知的影响。结果显示, 母语者在单一说话者噪声条件下表现更佳, 而二语学习者未表现出类似优势, 并且整体上更容易受到噪声的不利影响。这些发现提示, 噪声类型不仅影响二语学习者的行为表现, 也会通过调节神经振荡模式改变其对言语信息的加工方式。

4.2 语言熟练度

已有研究表明, 二语熟练度会显著调节二语学习者

在安静环境下的神经振荡模式。Lizarazu等人(2021)利用脑磁图技术考察熟练度对二语皮层追踪的影响, 发现随着二语学习经验的增加, δ 和 θ 频段的语音皮层追踪显著增强。鉴于 δ 和 θ 频段分别对应短语和音节层级的信息处理, 该结果表明, 语言熟练度的提高有助于增强学习者对音节结构的预测能力, 从而促进听觉皮层中持续的 θ 频段振荡; 换言之, 学习者越精通目标语言, 其对即将到来的音节信息的预测越精确, 相关频段的振荡功率也越强。

Pereira等人(2022)探讨了双语背景对二语理解中振荡动力学的影响, 发现早期双语者与晚期双语者在 α 和 θ 频段振荡功率上并未表现出显著差异。然而, 这一结果并不能直接否定熟练度对神经振荡的调节作用: 若缺乏单语者或二语初学者作为对照组, 则难以区分两种可能性: (1) 不同发展阶段的二语学习者在神经振荡层面确实不存在差异; (2) 神经振荡差异仅在二语熟练度达到某一特定阈值后才会显现。因此, 关于熟练度如何调控二语学习者不同频段神经振荡的具体机制, 目前仍存在争议。这种不一致既可能源于所关注的振荡频段不同, 也可能在同一频段中因研究设计差异而呈现不同结果。

在嘈杂的声学环境下, 二语学习者理解二语的困难主要来源于其在词汇和句法信息获取方面的不足(Hahne & Friederici, 2001)。这一困难可能反映了噪声和熟练度对不同层级语言结构解析的不同干扰方式: 噪声主要影响音节层面的神经夹带, 而语言熟练度则更显著地调节短语层面的神经夹带(Blanco-Elorrieta et al., 2020)。Blanco-Elorrieta等人(2020)使用脑磁图记录不同熟练度汉英双语者在多种噪声水平下对音节和短语结构的皮层追踪情况。结果发现, 对音节层级的追踪主要受噪声强度影响, 并随噪声水平升高而下降; 相比之下, 对短语层级的追踪则在很大程度上依赖于语言熟练度, 熟练度较高的学习者表现出更强的短语层级皮层夹带。然而, Zinszer等人(2022)报告了不同的结果: 在无噪声条件下, 随着熟练度提高, 对二语的皮层追踪显著增强; 而在噪声条件下, 熟练度对二语追踪强度的影响并不显著。针对熟练度与二语皮层追踪关系的不一致结果, 目前存在两种可能解释。第一种观点认为, 在安静条件下, 二语追踪强度随熟练度提升而增强, 当二语学习者接近母语水平时, 对母语与二语的皮层追踪均可能显著增强。第二种观点则认为, 熟练度与二语追踪之间可能呈倒U型关系, 而现有研究可能仅捕捉到了该曲线的左侧阶段。上述解释仍有待通过更系统的纵向研究和更精细的熟练度划分加以验证。

除熟练度外, 二语习得年龄也可能影响神经振荡模式。Wöstmann等人(2017)指出, α 频段功率与自上而下的注意控制密切相关, 且 α 功率的增强通常与语言理解所需努力程度呈正相关。在此基础上, Grant等人

(2022)利用脑电图研究习得年龄对二语学习者在噪声中言语理解的影响,发现习得年龄越晚, α 频段振荡功率越低。然而,由于较晚的习得年龄通常伴随较低的语言熟练度,目前尚无法明确区分 α 功率下降究竟源于习得年龄本身,还是语言熟练度降低的结果。

5 语言转换中的神经振荡现象

语言转换(Language Switching)是指多语者在交流过程中交替使用两种或两种以上语言的现象。在实际语言使用中,二语学习者往往需要在母语与二语之间灵活切换,以实现有效表达。相较于单一语言的持续使用,语言转换通常伴随着更高的错误率、更长的反应时以及更显著的神经反应,反映出其较高的认知加工需求(Grainger & Beauvillain, 1987)。

以往的行为研究主要从产出和理解两个层面考察语言转换过程。例如,一些研究关注双语者在口头产出中如何在不同语言之间切换(Jackson et al., 2001; Gollan et al., 2014),一些则系统探讨了影响语言转换效应的关键因素,包括语言熟练度(Costa et al., 2004)以及语言转换的方向性与非对称性(Peeters et al., 2014)。此外,也有研究提出,在言语理解过程中,双语者同样需要不断在两种语言系统之间进行选择,并抑制来自非目标语言的干扰(Costa & Santesteban, 2004)。

近年来,神经振荡研究为理解语言转换的神经机制提供了新的视角,相关工作开始关注语言转换如何影响连续言语的感知与理解(Xie et al., 2019; Pérez et al., 2019)。Xie等人(2019)通过分析诱发振荡探讨语言转换的认知控制机制,发现言语理解与语言产出在神经振荡模式上表现出一定的相似性,主要体现为 δ 和 θ 频段的调节作用。具体而言,与二语到母语的转换相比,母语到二语转换后 θ 频段振荡功率显著增强,表明从优势语言切换到非优势语言需要更多的认知控制资源。

在言语理解领域,Pérez等人(2019)考察了二语学习者在处理两种已知语言混合语音流时的神经振荡特征,发现 δ 频段的神经夹带能力整体下降,但在随后进行的记忆回忆任务中并未观察到显著的行为差异。此外,语言转换方向对 θ 频段振荡功率的影响呈现出非对称性。这些结果表明,尽管语言转换未必直接削弱对长时间语音信息的记忆与提取,但其引入的额外认知负荷能够在神经振荡层面被清晰地捕捉。

总体而言,相较于单一语言的持续使用,语言转换在言语理解和产出过程中均会显著增加认知负荷,其神经表现主要体现在 δ 和 θ 频段振荡功率的调节上。同时,语言转换的方向性也是影响神经振荡模式的重要因素。这些发现表明,神经振荡指标能够为理解双语语言控制与语言转换机制提供关键线索。

6 总结与展望

综上所述,现有研究表明,从神经振荡的视角考察第二语言听觉言语加工,有助于揭示大脑在不同时间尺度上整合声学信息与语言结构的动态机制。不同振荡频段在言语加工层级中的功能分工,为理解二语学习者如何解析连续语音流提供了重要线索。然而,在这一框架下,关于二语加工中神经振荡模式的具体表征仍存在较大不确定性,尤其是在皮层追踪强度、跨语言差异及其调节机制方面,不同研究结果尚未形成统一解释。总体来看,二语听觉加工并非单一因素驱动,而是受到语言经验、环境约束与任务需求等多重因素的共同调控,这些因素通过影响不同频段神经振荡的协同作用,塑造了二语学习者对言语信息的感知与理解方式。

总体而言,从神经振荡的视角探讨二语听觉言语感知与理解的神经机制仍处于起步阶段,现有研究主要集中在识别影响神经振荡的关键因素。未来研究有必要在此基础上进一步深化。首先,应更加系统地考察语言熟练度如何调控不同频段的神经振荡,尤其是在复杂听觉环境中,以明确熟练度与神经振荡之间的动态关系。其次,鉴于行为研究已充分表明噪声类型,特别是不同掩蔽类型,对二语学习者噪声中言语识别具有显著影响,未来有必要将这一变量引入神经振荡研究框架,系统比较不同噪声类型对二语学习者神经振荡模式的影响。通过结合行为表现与神经振荡指标,有望更全面地揭示二语听觉言语加工的神经机制。

参考文献

- [1] 盛经纬,高家红.(2021).脑磁图仪的前世今生与未来.《物理》,50(7),463-469.
- [2] Aydelott J & Bates E. (2004). Effects of acoustic distortion and semantic context on lexical access. *Language and Cognitive Processes*, 19, 29-56.
- [3] Aydelott J, Baer-Henney D, Trzaskowski M, Leech R & Dick F. (2012). Sentence comprehension in competing speech: Dichotic sentence word priming reveals hemispheric differences in auditory semantic processing. *Language and Cognitive Processes*, 27(7/8), 1108-1144.
- [4] Bastiaansen M, Van Berkum J J & Hagoort P. (2003). Syntactic processing modulates the θ rhythm of the human EEG. *NeuroImage*, 17, 1479-1492.
- [5] Bastiaansen M & Hagoort P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in Brain Research*, 159, 179-196.
- [6] Bastiaansen M, Magyari L & Hagoort P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *Journal of Cognitive Neuroscience*, 22, 1333-1344.

- [7] Bastiaansen M, Mazaheri A & Jensen O. (2012). Beyond ERPs: Oscillatory neuronal dynamics. In *Oxford Handbook of Event-Related Potential Components* (pp. 31–49). Oxford University Press.
- [8] Blanco-Elorrieta E, Ding N, Pykkänen L & Poeppel D. (2020). Understanding requires tracking: Noise and knowledge interact in bilingual comprehension. *Journal of Cognitive Neuroscience*, 32(10), 1975–1983.
- [9] Bradlow A R & Bent T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America*, 112(1), 272–284.
- [10] Bregman A S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT Press.
- [11] Broersma M & Scharenborg O. (2010). Native and non-native listeners' perception of English consonants in different types of noise. *Speech Communication*, 52, 980–995.
- [12] Brouwer S, Van Engen K J, Calandruccio L & Bradlow A R. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. *Journal of the Acoustical Society of America*, 131(2), 1449–1464.
- [13] Buzsáki G, Anastassiou C A & Koch C. (2012). The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, 13(6), 407–420.
- [14] Cooke M, Garcia Lecumberri M L & Barker J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America*, 123(1), 414–427.
- [15] Cooke M, García Lecumberri M L, Scharenborg O & van Dommelen W. (2010). Language-independent processing in speech perception: A study of English consonant identification in 8 European languages. *Speech Communication*, 52, 954–967.
- [16] Costa A & Santesteban M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50, 491–511.
- [17] Ding N, Melloni L, Zhang H, Tian X & Poeppel D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 19(1), 158–164.
- [18] Ding N & Simon J Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience*, 8.
- [19] Doelling K B, Arnal L H, Ghizza O & Poeppel D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage*, 85, 761–768.
- [20] Getz, Ding N, Newport E L & Poeppel D. (2018). Cortical tracking of constituent structure in language acquisition. *Cognition*, 181, 135–140.
- [21] Giraud A L & Poeppel D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517.
- [22] Gollan T H, Kleinman D & Wierenga C E. (2014). What's easier: Doing what you want, or being told what to do? Cued versus voluntary language and task switching. *Journal of Experimental Psychology: General*, 143(6), 2167.
- [23] Grainger J & Beauvillain C. (1987). Language blocking and lexical access in bilinguals. *The Quarterly Journal of Experimental Psychology Section A*, 39(2), 295–319.
- [24] Grant A M, Kousaie S, Coulter K, Gilbert A C, Baum S R, Gracco V... & Phillips N A. (2022). Age of acquisition modulates alpha power during bilingual speech comprehension in noise. *Frontiers in Psychology*, 13, 865857.
- [25] Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P & Garrod S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology*, 11(12), e1001752.
- [26] Jackson G M, Swainson R, Cunnington R & Jackson S R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, 4(2), 169–178.
- [27] Juffs A & Harrington M. (1995). Parsing Effects in Second Language Sentence Processing: Subject and Object Asymmetries in wh-Extraction. *Studies in Second Language Acquisition*, 17(4), 483–516.
- [28] Hahne A & Friederici A D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, 4(2), 123–141.
- [29] Kayser C, Wilson C, Safaai H, Sakata S & Panzeri S. (2015). Rhythmic auditory cortex activity at multiple timescales shapes stimulus-response gain and background firing. *Journal of Neuroscience*, 35(20), 7750–7762.
- [30] Kilman L, Zekveld A, Hällgren M & Rönnberg J. (2014). The influence of non-native language proficiency on speech perception performance. *Frontiers in Psychology*, 5, 651.
- [31] Lizarazu M, Carreiras M, Bourguignon M, Zarraga A & Molinaro N. (2021). Language proficiency entails tuning cortical activity to second language speech. *Cerebral Cortex*, 31(8), 3820–3831.
- [32] Luck S J & Kappenman E S. (2011). *The Oxford handbook of event-related potential components*. Oxford University

- Press.
- [33] Luo H & Poeppel D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54, 1001–1010.
- [34] Mattys S L, Davis M H, Bradlow A R & Scott S K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, 27(7/8), 953–978.
- [35] Meador D, Flege J E & Mackay I R A. (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition*, 3(1), 55–67.
- [36] Meyer L. (2018). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience*, 48(7), 2609–2621.
- [37] Nourski K V, Steinschneider M, Rhone A E, Oya H, Kawasaki H, Howard III M A & McMurray B. (2015). Sound identification in human auditory cortex: Differential contribution of local field potentials and high gamma power as revealed by direct intracranial recordings. *Brain and Language*, 148, 37–50.
- [38] Obleser J & Kayser C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, 23(11), 913–926.
- [39] Peeters D, Runnqvist E, Bertrand D & Grainger J. (2014). Asymmetrical switch costs in bilingual language production induced by reading words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(1), 284–292.
- [40] Peelle J E, Gross J & Davis M H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, 23(6), 1378–1387.
- [41] Pereira Soares S M, Prystauka Y, DeLuca V & Rothman J. (2022). Type of bilingualism conditions individual differences in the oscillatory dynamics of inhibitory control. *Frontiers in Human Neuroscience*, 16, 910910.
- [42] Pérez A & Duñabeitia J A. (2019). Speech perception in bilingual contexts: Neuropsychological impact of mixing languages at the inter-sentential level. *Journal of Neurolinguistics*, 51, 258–267.
- [43] Rimmele J M, Poeppel D & Ghitza O. (2021). Acoustically driven cortical δ oscillations underpin prosodic chunking. *eNeuro*, 8(4).
- [44] Reetzke R, Gnanateja G N & Chandrasekaran B. (2021). Neural Tracking of the Speech Envelope Is Differentially Modulated by Attention and Language Experience. *Brain and Language*, 213, 104891.
- [45] Shimizu T, Makishima K, Yoshida M & Yamagishi H. (2002). Effect of background noise on perception of English speech for Japanese listeners. *Auris Nasus Larynx*, 29(2), 121–125.
- [46] Song J & Iverson P. (2018). Listening effort during speech perception enhances auditory and lexical processing for non-native listeners and accents. *Cognition*, 179, 163–170.
- [47] Song J, Martin L & Iverson P. (2020). Auditory neural tracking and lexical processing of speech in noise: Masker type, spatial location, and language experience. *The Journal of the Acoustical Society of America*, 148(1), 253–264.
- [48] Steinhauer K, Connolly J F, Stemmer B & Whitaker H A. (2008). Event-related potentials in the study of language. In H A. Whitaker (Ed.), *Concise Encyclopedia of Brain and Language* (pp. 91–104). Elsevier.
- [49] Van Engen K J. (2010). Similarity and familiarity: Second language sentence recognition in first- and second-language multi-talker babble. *Speech Communication*, 52(11/12), 943–953.
- [50] Vanthornhout J, Decruy L, Wouters J, Simon J Z & Francart T. (2018). Speech intelligibility predicted from neural entrainment of the speech envelope. *JARO: Journal of the Association for Research in Otolaryngology*, 19(2), 181–191.
- [51] Wei R & Su J. (2012). The statistics of English in China. *English Today*, 28, 10–14.
- [52] Wöstmann M, Lim S J & Obleser J. (2017). The Human Neural Alpha Response to Speech is a Proxy of Attentional Control. *Cerebral Cortex*, 27(6), 3307–3317.
- [53] Xie N, Li B, Zhang M & Liu H. (2019). Role of top-down language control in bilingual production and comprehension: Evidence from induced oscillations. *International Journal of Bilingualism*, 23(5), 1041–1063.
- [54] Zinszer B D, Yuan Q, Zhang Z, Chandrasekaran B & Guo T. (2022). Continuous speech tracking in bilinguals reflects adaptation to both language and noise. *Brain and Language*, 230, 105128.
- [55] Zion Golumbic E M, Ding N, Bickel S, Lakatos P, Schevon C A, McKhann G M, Goodman R R, Emerson R, Mehta A D, Simon J Z, Poeppel D & Schroeder C E. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a ‘cocktail party’. *Neuron*, 77(5), 980–991.

Second Language Auditory Perception and Comprehension from the Perspective of Neural Oscillations

Zhu Shiyao Lu Lingxi

Cognitive Science and Allied Health School, Beijing Language and Culture University, Beijing

Abstract: For second language (L2) learners, achieving processing speed and accuracy in auditory speech comprehension comparable to those of native speakers remains a major challenge. In recent years, advances in high-temporal-resolution electrophysiological techniques, such as electroencephalography (EEG) and magnetoencephalography (MEG), have provided a new theoretical framework for investigating L2 processing from the perspectives of neural oscillations and neural entrainment. This review systematically synthesizes recent research on neural oscillatory activity underlying auditory speech perception and comprehension in L2 learners, with a particular focus on the functional roles of different oscillatory frequency bands across hierarchical levels of speech processing, as well as the modulatory effects of language experience and environmental factors on neural oscillatory patterns. Accumulating evidence suggests that neural oscillations at different frequency bands encode speech information at multiple temporal scales, ranging from phonemes and syllables to phrasal structures. However, the relationship between cortical tracking strength in the first language (L1) and L2 remains unresolved, with inconsistent findings likely arising from differences in experimental paradigms, analytical approaches, and participants' language experience. In addition, bilingual background, especially L2 proficiency, has been shown to significantly modulate neural oscillatory patterns during L2 auditory processing under both quiet and noisy listening conditions. As a common phenomenon in bilingual language use, language switching typically imposes increased demands on cognitive control, which are primarily reflected at the neural level by modulations in oscillatory power in the delta (δ) and theta (θ) frequency bands. Based on these findings, this review integrates converging and diverging evidence in the current literature and highlights the need for future studies to incorporate complex auditory environments and longitudinal designs, in order to further elucidate the neural mechanisms of L2 auditory speech processing from a neural oscillation perspective.

Key words: Second language acquisition; Language processing; Neural oscillations; M/EEG