

不同语言加工中的语音预测效应：回顾与展望

沈艺虹

天津师范大学外国语学院，天津

收稿日期：2025年9月15日；录用日期：2025年10月9日；发布日期：2025年10月24日

摘要

在言语交流的过程中，人们在听到某些内容之前，常常会根据情景和常识预先激活与之相关的语义、句法和语音信息。句法和语义信息的预测在许多研究中已经被证实，但关于语音信息的预测情况还未形成统一的结论。因此，本文旨在梳理近年来学界对于语音信息预测的研究情况。本文介绍了母语背景以及二语背景下语音信息预测的相关实验研究，并分析语音信息预测过程中可能存在的机制。本文还对未来研究提出建议：1) 重视汉语特色，应更多关注汉语音段以及声调在预测过程中的作用；2) 关注二语研究，重视语音信息在二语中的表现；3) 探究语音信息预测背后存在的机制，拓展以西方语言为基础的理论模型，推动本土理论国际化。

关键词

语音预测，二语，眼动，视觉情景范式

Phonological Prediction Effects in Different Language Processing: Review and Outlook

Yihong Shen

College of Foreign Languages, Tianjin Normal University, Tianjin

Received: September 15, 2025; accepted: October 9, 2025; published: October 24, 2025

Abstract

During verbal communication, people often predict semantic, syntactic, and phonological information related to upcoming content based on contextual and general knowledge. While the prediction of syntactic and semantic information has been widely demonstrated in numerous studies, the extent and nature of phonological prediction remain subjects of ongoing debate. This article aims to review recent research on the prediction of phonological information in language processing. It presents empirical studies conducted in both first-language (L1) and second-language (L2) contexts and

examines the potential mechanisms underlying phonological prediction. Furthermore, the article proposes several directions for future research: 1) placing greater emphasis on the distinctive features of Chinese, particularly the role of segments and tones in the prediction process; 2) increasing attention to L2 studies to better understand how phonological prediction operates in a second language; and 3) investigating the mechanisms underpinning phonological prediction, expanding theoretical models primarily based on Western languages, and promoting the internationalization of local theoretical frameworks.

Keywords

Phonological Prediction, Second Language, Eye-Tracking, Visual World Paradigm

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在语言加工过程中，人们往往会借助情境或常识，在尚未听到具体内容之前，就预先激活与之相关的语义、句法和语音等多层面的语言信息。比如，当人们在边观看含有蛋糕、帽子、船等事物的图片边听句子时，当听到“He wants to eat...”，但尚未听到“蛋糕”时，人们便会提前看向相应图片，这表示他们根据语义信息预测了言语的内容。当前心理学界认为，句法和语义信息的提前激活是语言加工过程中的一个普遍现象[1]。但是关于语音信息是否会被提前激活，产生语音预测效应这一问题，学界尚未形成一致的结论。

在母语环境中，不少研究已经证实，母语者能够在语言理解过程中激活预测词汇的语音表征，这些研究主要聚焦于英语母语者[2][3]。相比之下，二语学习者在二语理解过程中可利用的资源较少。尽管如此，已有研究表明，二语学习者可以对句法信息和语义信息进行一定程度的预测[4][5]，但对于语音信息的预激活情况，学界尚未有统一的结论。鉴于此，本文主要回顾过去近年来，利用 ERP 和眼动技术对语音信息预期的研究进程，总结现有研究的结果，并通过实验结果探讨语音信息预测的机制，在此基础上，提出对未来语音信息预测研究的建议。

2. 语言加工中的语音预测效应

2.1. 母语感知过程中的语音预测效应研究

许多研究已经表明，在母语背景下，被试者可以预测即将到来的语音信息。最早进行研究的是 DeLong 团队，他们利用 Event-Related Potential (ERP) 技术发现[2]，在阅读理解过程中，人们可以预测高度可预测词汇的语音特征。实验中，被试阅读一些句子，这些句子强烈暗示某个特定名词(例如，The day was breezy so the boy went outside to fly..., 暗示风筝)，随后出现可预期的名词短语(a kite)或非预期但合理的名词短语(an airplane)。非预期名词引发的 N400 (N400 是一个与语义加工密切相关的 ERP 成分，在目标刺激出现后 400 毫秒左右最为明显。一般认为 N400 与语义信息的提取或整合的难度有关，难度越大，N400 波幅越大)比预期名词更大。这表明母语者可以根据可预测词的开头语音信息，从而选择正确的冠词搭配。然而，后续脑电研究也指出，母语者对语音信息的预测会受到句末名词可预测性的调节，因此语音预测效应——即在语言理解过程中听者基于语境预先激活后续词汇语音信息的现象——并不总是稳定出现。此外，目前尚不明确该效应究竟是源于对语音信息本身的预测，还是对冠词与名词搭配关系的预期[6]。

因此,一些研究通过相关违例范式来探究语音预测效应。具体实验操作是,将句子 *The student is going to the library to borrow a book* 中,有高预期度的目标词 *book* 替换成语音相关词(*hook*)和无关词(*sofa*)以此观察被试在面对各个词汇时的 ERP 反应。结果表明语音相关词比其他无关词诱发了更小的 N400 波幅,即由于目标词和语音相关词之间存在语音上的重叠,这种预先激活的语音信息降低了语音干扰词出现时的加工难度。因此,可得出结论,在母语理解的过程中,词汇的语音信息是可以被提前激活的[3] [7]。

ERP 技术在实验过程中并不能检测预测的时间进程,即预测具体何时开始。因此,学者们开始采用眼动技术中的视觉情景范式来探究语音预测效应。与 ERP 技术相比,眼动技术能够直观地展示语音预测的过程。在视觉情景范式中,被试一边听有高预期度的句子,一边看电脑上的图片,这些图片包括目标图、语音干扰图和无关图。那么那些优先注视语音干扰图的被试,必定是预先激活了可预测词的形式(因为该语音形式相关词与可预测词或语境在语义等其他方面毫无关联)。Ito 在其 2018 中的研究发现[8],英语母语者在听到目标词前,对语音干扰图关注度显著高于无关图,这表明母语者能够根据语境提前激活目标词的语音信息。后续的研究也发现了母语者存在语音预测效应[9]-[11]。但也有研究并未发现母语背景下对语音信息的预测[5] [12] [13]。之前的研究多集中在英语这类拼音文字,这可能是语音预测效应不稳定的原因,与此不同的是,以汉语为语料的研究发现了明显的语音预测效应,这为语音预测效应提供了更加有力的证明。

以往的研究多是集中于英语等拼音文字,这些语言的语音信息重叠主要集中在音段部分,而汉字的音节除了辅音(onset)、元音(rime)外,还有声调,因此汉语中存在许多同音异义字,这让目标字与竞争字之间语音重叠度最大化,这可能是汉语中发现语音预测效应较为稳定的原因。Shen 在研究中使用双字组刺激[14],考察声调相似性和语境可预测度对声调在语言理解中的作用,发现只有在高度可预测的语境下,与目标词(象限 *xiang4xian4*)声调相似性高的词语(相册 *xiang4ce4*)会引起语音预测效应,但因实验中语音干扰词的声调和音段信息都与目标词相同,所以母语者是否能预测声调本身这一点还并不清楚。随后,Xu 运用了三个实验探究声调、音段对语音预测的效应[15],结果表明干扰词的声调和音段信息与目标词都相同时(即同音异义字,如,椒/*jiao1*/和胶/*jiao1*/)最容易引起语音预测效应,其次是仅音段信息相同,但并未发现声调单独引起语音预测效应,这表明音段信息可以独立于声调信息产生预测效应,声调则不然。Zhao 发现了同样的结果[16],声调相同并不会引起语音竞争效应,这表明作为语音信息一部分的声调信息,可被预测的程度最小。还有研究进一步探索认知处理系统对语音预测的影响,Li [17]在实验中设置不同的图片预览时间,但结果都呈现出稳定的语音预测效应,因此作者认为语音预测是视觉语境驱动的自动化过程,而非高阶认知推理。Li [18]进一步探究了工作记忆对语音预测,结果表明个人工作记忆容量的高低对语音预测能力没有显著影响,但工作记忆容量高的组别语音预测持续时间更长,这表明工作记忆容量并未决定语音预测是否发生,而是影响了预测表征的稳定性或持续性。这些研究结果为理解汉语中语音预测机制提供了新的视角,并揭示了音段信息在语言预测中的重要作用,也为其他声调语言的研究提供了新的研究思路,进一步探究声调在语言理解预期中的作用。同时一些学者也关注到了认知能力在语音预测中可能发挥作用,但目前的研究成果还不多,未能形成有力的证据。

2.2. 二语感知过程中的语音预测效应研究

预期对于非母语者来说可能更为关键,因为预激活的语义、语音等信息能在一定程度上减轻外语交流过程中的认知负荷,从而降低外语理解的难度。外语句子理解之所以比母语句子理解更为困难,很可能是因为在外语语境下人们的预期效率更低[19]。与母语者相比,二语者在二语理解的过程中受到的影响因素更多,预测加工过程更为复杂。

有关二语背景下非目标语的母语激活效应的研究存在较大分歧,双语词汇通达模型中的非选择性加

工观则认为, 双语词汇通达总是伴随非目标语言的同步激活[20] [21]。据此, 母语对译词的语音表征也应被激活。Ito 探讨了日语 - 英语二语者和英语母语者在视觉情景范式下测试语音预测效应是否存在, 研究结果表明, 相较于母语者, 日语母语 - 英语二语者不能激活二语的语音表征, 并不存在二语预测效应[8]。但徐晓东探究了汉语 - 英语双语者在理解英语句子时能否激活母语中的音段和声调信息[22], 结果表明高水平的二语者在二语理解过程中能够激活母语语音信息, 存在母语语音激活效应。Amos 选择了二语水平较高的同传人员作为被试, 但在实验中也并未发现二语者预先激活了母语语音信息[23]。

在探讨双语者的语言处理过程中, 关于是否会激活非目标语言的问题, 目前学界形成了两种不同的观点。选择性加工观点主张, 在双语词汇的通达过程中, 只有目标语言会被激活, 而非目标语言则不会被触发[24]。根据这一理论, 当双语者在预测二语词汇时, 其母语中的对应词汇的语音信息不应被同时激活, 因此, 语音预测效应应当仅限于二语层面[8] [11] [22] [25]。

相对地, 非选择性加工观点认为, 在双语词汇通达的过程中, 非目标语言总是会伴随目标语言一起被激活[22] [26] [27]。基于这一理论, 在二语处理过程中, 母语中的对应词汇的语音表征也应当被激活。这两种观点提供了不同的视角来解释双语者在语言加工中的行为模式, 对于理解双语者的语言处理机制具有重要意义。

现有研究结果并不一致, 因此尚不能确定在二语感知过程中, 是否会激活母语语音信息, 支持何种理论, 私以为, 这受到二语学习者学习水平、学习能力、认知策略、年龄等多种因素的影响[11] [28]。

3. 语言理解中预测行为的加工策略

3.1. 产出预期机制(Prediction-by-Production Mechanism)

产出预期机制理论提出, 人们在预测句子内容时, 实际上是通过语言产出系统来实现的[29]。具体而言, 在理解句子的过程中, 人们会不自觉地对当前可获取的内容进行一种隐性的模仿(covert imitation)。这种模仿意味着理解者将说话者的话语视作自己即将要表达的内容, 并利用自身的语言产出系统, 根据预测来补充未完全表达的部分。在这个过程中, 理解者不仅会综合考量语境所传递的信息, 还会同时考虑其他非语言因素, 例如与说话者共享的背景知识等。因此, 对于可预测的信息, 产出预期机制被认为是一种更为精确的预测方法。该机制还主张, 预期加工过程与语言产出过程相一致, 遵循从语义激活到句法激活, 最终到语音信息激活的层次顺序。由于语音信息的预激活位于预期加工过程的最后阶段, 因此它特别容易受到干扰。这个加工机制可以解释为什么许多研究即使在母语背景下, 也并未发现母语语音预测效应。其一是因为语音信息的预激活过程过于短暂和迅速, 有限的实验工具未能采集到相关数据[3]。其二是听者在实验过程中受到其他因素的影响, 如视觉图片、语言水平未能及时进行预测行为[22] [23]。

3.2. PACS 多机制预测系统

Huettig 提出[24], 语言预测不是由单一机制完成的, 而是至少依赖四种不同的机制, 他称之为 PACS 多机制预测系统(production-, association-, combinatorial-, simulation-based prediction), 这四种机制可以互补、互动, 甚至竞争:

1) **P (Production-based) 产出模拟**: 听者借助自身的语言产出系统, 在内心默默“复述”或补全说话人即将说出的词句, 形成从词形到语音的完整预激活; 该机制在需要个人产出或个人词汇量较高是发挥作用较大;

2) **A (Association-based) 联想激活**: 词汇间因语义、语音、视觉或经验共现而形成的联结自动扩散, 使相关词快速被预激活; 其强度取决于个人语言经历与识字量, 会在无语法约束的单纯听觉理解中占主

导;

3) C (Combinatorial)结构组合: 利用句法、语义角色等规则,对即将出现的句法范畴和词汇进行算法式预测;当句子框架高度受限或任务强调语法判断时,该机制贡献最大;

4) S (Simulation-based)情境模拟: 通过心理意象或事件知识在“脑内播放”话语描述的场景,从而激活与事件相关的物体、动作及其名称;语境越丰富、画面感越强,该机制越能向下级联触发具体词形或语音表征。

四条机制既可形成联结(情境模拟→联想扩散→结构组合→产出模拟),也可能相互冲突(联想词与语法要求不符),其实际权重由语境约束度、任务需求、认知资源与个体语言能力动态调节[10]。

目前研究因为实验任务不同、被试语言水平不同等因素结果不同,不能完全证实语言预测的确切机制。但语言理解中预测行为十分复杂,如果只关注单一的预测加工机制,并不能很好地解释这一过程,建议未来研究可以进一步探索多种加工机制之间的协同合作是如何发挥作用的。

4. 总结与展望

本文梳理了近年来采用 ERP 技术和眼动技术探究语言理解中语音信息预测的国内外研究,发现母语者在母语背景下可以产生语音预测效应,尤其在以汉语为语料的研究中更为明显。而二语者因为二语熟练度、认知策略等因素影响,并未能产生母语语音预测效应。产出预期加工机制可以解释部分语音预测效应,但仍需探讨背后更加复杂的机制,如 PACS 多机制预测系统,因此,为未来的研究提出几点建议:

第一,重视汉语特色,在母语感知的研究中,母语者在高限制的情境下会产生母语语音预测效应,但关于音段在预测效应中的作用,之前的研究主要集中于英语、意大利语等拼音文字。而汉语不同,它用声调区分意义,因此存在同音异义词,这让目标词和竞争词间的语音重叠度达到最大,汉语中的音段信息和声调在母语和二语感知中发挥的作用需要进一步的实证研究。第二,关注二语研究,检验二语感知过程中的语音预测效应的研究数量还较少,尚未形成一致的结论,需要更多的实验研究来检验这种效应是否存在,何时发挥作用以及如何发挥作用。第三,深入挖掘语音预测的认知机制,突破以印欧语系为核心的理论框架,将汉语等非印欧语言的声调、音段等证据纳入模型,打造跨语言普适、本土概念国际化的语言预测新范式。

参考文献

- [1] 王芙蓉,王敏. 语言加工中的预测行为研究综述[J]. 西安外国语大学学报, 2020, 28(3): 59-64.
- [2] DeLong, K.A., Urbach, T.P. and Kutas, M. (2005) Probabilistic Word Pre-Activation during Language Comprehension Inferred from Electrical Brain Activity. *Nature Neuroscience*, **8**, 1117-1121. <https://doi.org/10.1038/nn1504>
- [3] Ito, A., Corley, M., Pickering, M.J., Martin, A.E. and Nieuwland, M.S. (2016) Predicting Form and Meaning: Evidence from Brain Potentials. *Journal of Memory and Language*, **86**, 157-171. <https://doi.org/10.1016/j.jml.2015.10.007>
- [4] Foucart, A., Martin, C.D., Moreno, E.M. and Costa, A. (2014) Can Bilinguals See It Coming? Word Anticipation in L2 Sentence Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **40**, 1461-1469. <https://doi.org/10.1037/a0036756>
- [5] Ito, A., Dunn, M.S., III. and Pickering, M.J. (2017) Effects of Language Production on Prediction: Word vs. Picture Visual World Study. *IEICE Technical Report*, **117**, 31-36.
- [6] Martin, C.D., Thierry, G., Kuipers, J., Boutonnet, B., Foucart, A. and Costa, A. (2013) Bilinguals Reading in Their Second Language Do Not Predict Upcoming Words as Native Readers Do. *Journal of Memory and Language*, **69**, 574-588. <https://doi.org/10.1016/j.jml.2013.08.001>
- [7] Laszlo, S. and Federmeier, K.D. (2009) A Beautiful Day in the Neighborhood: An Event-Related Potential Study of Lexical Relationships and Prediction in Context. *Journal of Memory and Language*, **61**, 326-338. <https://doi.org/10.1016/j.jml.2009.06.004>
- [8] Ito, A., Pickering, M.J. and Corley, M. (2018) Investigating the Time-Course of Phonological Prediction in Native and

- Non-Native Speakers of English: A Visual World Eye-Tracking Study. *Journal of Memory and Language*, **98**, 1-11. <https://doi.org/10.1016/j.jml.2017.09.002>
- [9] Ito, A. (2019) Prediction of Orthographic Information during Listening Comprehension: A Printed-Word Visual World Study. *Quarterly Journal of Experimental Psychology*, **72**, 2584-2596. <https://doi.org/10.1177/1747021819851394>
- [10] Kukona, A. (2020) Lexical Constraints on the Prediction of Form: Insights from the Visual World Paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **46**, 2153-2162. <https://doi.org/10.1037/xlm0000935>
- [11] 吴晓钢, 朱珩. 德语母语与二语语音预测效应的比较研究[J]. 现代外语, 2022, 45(6): 833-843.
- [12] Ito, A. and Sakai, H. (2021) Everyday Language Exposure Shapes Prediction of Specific Words in Listening Comprehension: A Visual World Eye-Tracking Study. *Frontiers in Psychology*, **12**, Article ID: 607474. <https://doi.org/10.3389/fpsyg.2021.607474>
- [13] Angulo, Chavira, A., Castellón-Flores, A., López-Santillán, H. and Arias-Trejo, N. (2023) Phono-Semantic Prediction during Language Comprehension: Effects of Working Memory. *Proceedings of the Annual Meeting of the Cognitive Science Society*, **45**, 1266-1272.
- [14] Shen, W., Hyönä, J., Wang, Y., Hou, M. and Zhao, J. (2020) The Role of Tonal Information during Spoken-Word Recognition in Chinese: Evidence from a Printed-Word Eye-Tracking Study. *Memory & Cognition*, **49**, 181-192. <https://doi.org/10.3758/s13421-020-01070-0>
- [15] Xu, X., Ji, C., Li, T. and Pickering, M.J. (2024) The Prediction of Segmental and Tonal Information in Mandarin Chinese: An Eye-Tracking Investigation. *Language, Cognition and Neuroscience*, **40**, 56-70. <https://doi.org/10.1080/23273798.2024.2395549>
- [16] Zhao, Z., Ding, J., Wang, J., Chen, Y. and Li, X. (2023) The Flexibility and Representational Nature of Phonological Prediction in Listening Comprehension: Evidence from the Visual World Paradigm. *Language and Cognition*, **16**, 481-504. <https://doi.org/10.1017/langcog.2023.38>
- [17] Li, X., Li, X. and Qu, Q. (2022) Predicting Phonology in Language Comprehension: Evidence from the Visual World Eye-Tracking Task in Mandarin Chinese. *Journal of Experimental Psychology: Human Perception and Performance*, **48**, 531-547. <https://doi.org/10.1037/xhp0000999>
- [18] Li, X. and Qu, Q. (2023) Verbal Working Memory Capacity Modulates Semantic and Phonological Prediction in Spoken Comprehension. *Psychonomic Bulletin & Review*, **31**, 249-258. <https://doi.org/10.3758/s13423-023-02348-5>
- [19] Ito, A. and Pickering, M.J. (2021) Automaticity and Prediction in Non-Native Language Comprehension. In: Grüter, T. and Kaan, E., Eds., *Bilingual Processing and Acquisition*, John Benjamins Publishing Company, 25-46. <https://doi.org/10.1075/bpa.12.02ito>
- [20] Lagrou, E., Hartsuiker, R.J. and Duyck, W. (2011) Knowledge of a Second Language Influences Auditory Word Recognition in the Native Language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **37**, 952-965. <https://doi.org/10.1037/a0023217>
- [21] Mishra, R.K. and Singh, N. (2013) Language Non-Selective Activation of Orthography during Spoken Word Processing in Hindi-English Sequential Bilinguals: An Eye Tracking Visual World Study. *Reading and Writing*, **27**, 129-151. <https://doi.org/10.1007/s11145-013-9436-5>
- [22] 徐晓东, 朱宏丽. 二语语境下母语语音信息的预激活——基于眼动视觉情境范式的研究[J]. 外语教学与研究, 2024, 56(4): 569-580+640.
- [23] Amos, R.M., Seeber, K.G. and Pickering, M.J. (2022) Prediction during Simultaneous Interpreting: Evidence from the Visual-World Paradigm. *Cognition*, **220**, Article No. 104987. <https://doi.org/10.1016/j.cognition.2021.104987>
- [24] Costa, A. (2009) Lexical Access in Bilingual Production. In: Kroll, J.F. and De Groot, A.M.B., Eds., *Handbook of Bilingualism*, Oxford University Press, 308-325. <https://doi.org/10.1093/oso/9780195151770.003.0018>
- [25] Amos, R.M., Hartsuiker, R.J., Seeber, K.G. and Pickering, M.J. (2023) Purposeful Listening in Challenging Conditions: A Study of Prediction during Consecutive Interpreting in Noise. *PLOS ONE*, **18**, e0288960. <https://doi.org/10.1371/journal.pone.0288960>
- [26] Costa, A., Caramazza, A. and Sebastian-Galles, N. (2000) The Cognate Facilitation Effect: Implications for Models of Lexical Access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **26**, 1283-1296. <https://doi.org/10.1037/0278-7393.26.5.1283>
- [27] Hoshino, N. and Kroll, J.F. (2008) Cognate Effects in Picture Naming: Does Cross-Language Activation Survive a Change of Script? *Cognition*, **106**, 501-511. <https://doi.org/10.1016/j.cognition.2007.02.001>
- [28] Huettig, F. (2015) Four Central Questions about Prediction in Language Processing. *Brain Research*, **1626**, 118-135. <https://doi.org/10.1016/j.brainres.2015.02.014>
- [29] Pickering, M.J. and Garrod, S. (2013) An Integrated Theory of Language Production and Comprehension. *Behavioral and Brain Sciences*, **36**, 329-347. <https://doi.org/10.1017/s0140525x12001495>