

基于眼动视觉情境范式的听理解加工研究

郑巧斐

金华职业技术大学师范学院, 浙江 金华

收稿日期: 2025年3月4日; 录用日期: 2025年4月15日; 发布日期: 2025年4月29日

摘要

视觉情境范式的兴起以及视听交互理论的发展, 为研究者提供了通过眼动模式分析听觉输入加工的新途径。目前, 研究者利用视觉情境范式探究了语言加工的语音、词汇和句法层面。这些研究为中国外语听力习得的研究提供了坚实的基础。中国外语教学与习得领域的研究者应勇于采用现代心理学和神经科学技术, 如眼动追踪技术, 针对中国外语学习者及汉语特点, 在听力习得和听力教学领域展开研究。

关键词

眼动, 视觉情境, 听理解, 习得

Research on Auditory Comprehension Processing Based on the Visual-World Paradigm

Qiaofei Zheng

Normal College of Jinhua University of Vocational Technology, Jinhua Zhejiang

Received: Mar. 4th, 2025; accepted: Apr. 15th, 2025; published: Apr. 29th, 2025

Abstract

The rise of the visual-world paradigm and the development of audio-visual interaction theories have provided researchers with a new avenue for analyzing auditory input processing through eye-tracking. Currently, researchers have utilized the visual-world paradigm to explore language processing at the levels of phonetics, vocabulary, and syntax. These studies have laid a solid foundation for research on the acquisition of listening acquisition in foreign languages in China. It is imperative that researchers in the field of Chinese foreign language teaching and acquisition embrace modern psychological and neuroscientific research technologies, such as eye-tracking, to conduct research tailored

to Chinese foreign language learners and the unique characteristics of the Chinese language in the areas of listening acquisition and instruction.

Keywords

Eye Movement, Visual-World, Auditory Comprehension, Acquisition

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

无论是在交流中，还是在语言学习中，听理解都起着非常重要的作用。与其他语言技能相比，听理解具有的内隐性和瞬时性特征使其成为教学和研究中的难点。不过，不管是交流、语言学习、听力测试，还是教学，人们在听的时候眼睛总是需要看某些地方，眼动追踪就是运用眼动仪来确定人的眼睛朝哪里看的过程。大量的研究已经证实，目光注视的位置通常与所关注和所思考的事物有关，特别是有目的地看某物的时候(Bojko, 2019: p. 9) [1]，这就是“眼 - 脑假说”(eye-mind hypothesis)。根据“眼 - 脑假说”，眼睛的注视过程反映的是大脑自动运作的认知加工过程，因此，我们可以通过眼动仪记录人们在听过程中眼睛的注视模式，来探究听语言的理解。

2. 基于眼动追踪的听理解加工研究背景

2.1. 眼动追踪追踪什么

眼动追踪在语言研究中一般用来研究阅读加工，但该技术除了可以呈现书面文本之外，还能呈现图像、场景、音频和视频等，这些刺激结合运用就可以用来探索听理解加工。在听理解加工的研究中，变量通过音频播放，眼动对视觉输入的注视就会锁定在一个相应的语言变量(字、词、句法结构等)上。屏幕上的不同区域被定义为兴趣区，研究通常报告首次注视点、首次注视时间、总注视时间、眼跳距离等指标。首次注视点的位置意味着最先认知加工的内容；首次注视时间能有效反映词汇通达的早期阶段特征，对多个言语特征反应敏感；总注视时间(total viewing time)是落在兴趣区的所有注视点的时间总和，反映的是较慢、较长时间的认知加工。当与恰当的控制条件相匹配时，较长的时间注视反映了更努力的加工，或更高的任务要求(Aline, 2020: p. 88) [2]。眼睛回视说明需要从前面提取信息帮助后面进行认知加工；眼跳距离大，说明在眼跳前的注视中所获得的信息相对较多，注视的内容难度低，可预测性高。在不同的注视任务下，眼动指标也会有比较大的差异。

2.2. 眼动视觉情境范式的兴起

眼动追踪的研究范式有多种。1974 年，Cooper 开创性地把视觉图片与听觉输入结合起来，设计出了“视觉情境范式 visual-world paradigm”。在实验中，他发现当人们听到“狮子”这个词的开头或整词时，会注视狮子的图片，当听到“非洲”这个词时会注视“狮子、斑马、蛇”，而不会看“狗”。当听到一个词的首字母发音时，大约有 55% 的概率，注视会转向与该信息相关的词；在听完一个词的第一个 5 秒之后，大约有 44% 的概率注视会从该词移向其他。实验证明，在边听边看的过程中，眼动模式和词汇表征及其语义激活存在显著的关联。该实验的逻辑是：被试对所听句子的加工与眼睛对刺激物的注视是同步进行的，即当被试听到某个词语时，被试的眼睛会注视视觉情景中的图片或实物中相关的物体(闫国利、

白学军, 2018: p. 116) [3]。但遗憾的是, 该研究范式直到上世纪 90 年代才再次被 Tanenhaus 和他的同事们运用于语言学的研究。

Tanenhaus 等[4]在 1995 年运用这个范式进行了听理解中视觉信息与语言信息整合的研究, 文章发在 “Science” 上, 引起了广泛关注。在该研究中, Tanenhaus 等通过给与被试有歧义的听觉语言指令刺激来观察被试在受到歧义指令时的眼睛运用指标。研究发现, 即使是在语言加工的最初阶段, 视觉环境也会影响单个词的识别, 并介导句法加工。个体对听觉指令的处理是渐进的, 在听到指令中的相关词后, 眼睛会立即跳视物体, 因此, 眼部运动指标提供了洞察伴随语言理解的心理过程的证据。此外, 视觉语境影响了单个词内临时歧义的解决, 视觉语境所提供的信息会影响听觉指令的句法处理。Tanenhaus 等(1995) [4]指出, 在定义明确的任务下, 眼动可以用来观察在自然条件下听理解背后的快速心理过程, 而且, 这一范式可以扩展到从言语识别到合作问题解决过程中的对话互动等问题。

2.3. 眼动追踪可用于听理解加工研究的理论依据

视听交互作用理论

听觉信息与视觉信息加工模块并不是独立工作的, 两个模块的信息加工存在复杂的交互作用。眼动视觉情境范式依赖眼动追踪技术实时监控被试在理解听觉语言时对视觉输入的注视, 将听觉语言刺激和视觉刺激进行了跨通道整合, 以此考察听理解的心理加工过程。其基本理论假设是, 眼球运动能够很好地反映语言理解状态, 视觉中的眼球运动和相关的听觉输入存在紧密的锁时关系, 被试能够根据听觉输入的信息预测视觉中即将出现的相关物体或文字, 因此在听觉输入相关目标信息之前, 通常能够将视觉注意聚焦到关键目标上(Kamide *et al.*, 2003 [5]; Altmann & Kamide, 2007 [6])。相应地, 所输入的视觉信息也能够帮助被试预测或筛选可能出现的听觉信息。

3. 基于眼动追踪的听理解加工研究

3.1. 语音层面的研究

1976 年, McGurk 和 MacDonald 发现了“麦格克效应”(the McGurk effect), 他们给被试观看一段视频, 视频中的“ga”被发成了“ba”, 被试却认为他们听到的是“da”, 这表明了音节层面听觉与视觉之间的交互作用。人们在听到目标词时, 随着词头、词腹、词尾的依次展开, 与之发音相同的竞争词依次被激活(Allopenna *et al.*, 1998 [7]; Dahan *et al.*, 2001 [8])。例如, 当听到研究人员发出“拿起烧杯 beaker”这个指令时, 图片中的竞争物 beetle、speaker 会被依次激活, 但无关物 carriage 则不会被注视。而与双音节词相比, 单音节词的押韵竞争效果不太可能出现(Simmons & Magnuson, 2018) [9]。此外, 说话者的声音分贝也会影响听理解中目标词的识别, 相对于日常声音(60 dBA), 较轻柔的声音(50 dBA)会增加词头、词腹相同的竞争词的激活, 非常柔和的语音(40 dBA)则会减缓该类激活, 而增加词尾相同的竞争词被激活(Hendrickson *et al.*, 2020) [10]。这些研究都支持了言语识别模型 TRACE (trace model of speech perception)。

3.2. 词汇层面的研究

3.2.1. 词汇识别的研究

袁周敏等(2020) [11]通过梳理 1934 年到 2020 年间国外语言学学科眼动研究的现状时发现, 研究者多数采用了眼动视觉情境范式来研究影响词汇识别加工中的关键因素。Tanenhaus 和 Spivey-Knowlton (1996) [12]让被试按照音频指令观看、选择甚至移动目标物体, 来考察被试对视觉区域所展示的指示物的注视时间和方式。Huettig 和 Altmann (2005) [13]利用视觉情境范式, 探究了听觉输入对视觉词汇识别的影响, 该研究表明, 在听到目标词(如, dog)时, 被试偏向于注视类别相关的图片(如, pig)或主题相关的图片(如,

bone), 而不是无关的干扰因素。这些发现与语义记忆的局部模型(localist models of semantic memory)一致, 在该模型中, 概念是由语义网络中的节点表征的, 具有相似含义的概念在语义网络中彼此连接。当口语词或物体识别发生时, 参照节点的激活扩散到相关节点, 即语义相关概念, 从而支持了视觉情境任务中语言介导的眼球运动(Chow *et al.*, 2017) [14]。

3.2.2. 词汇语义加工的研究

Altmann 和 Kamide (1999) [15]首次运用视觉情境范式来研究语义预测的处理。在研究中, 显示器上显示“一个男孩、一个蛋糕、一个球、一辆玩具汽车、一辆玩具火车以及其他物体”, 当英语母语者听到 the boy will eat...时的目光转向蛋糕比听到 the boy will move...时要快。研究显示, 被试会看屏幕中尚未命名的物体, 这都表明他们会对即将到来的听觉信息进行预测。

Yee 和 Sedivy (2006) [16]研究发现, 语义关联效应的发生与语义激活的词汇项是否展示在图片中无关。眼动技术可以用来有效地测量在听的过程中词的意义(以及它的形式)变得活跃的时间过程, 可提供不同词群的词汇加工过程的详细时间进程信息。

Huettig 和 McQueen (2007) [17]通过四个实验, 研究了听 - 看过程中语音、视觉特征和语义的相互作用, 他们发现, 在听觉输入过程中, 如果被试有足够的时间来做视觉信息上的选择, 就会在早期表现出语音偏向, 在晚期会表现出语义和拼写偏向, 但是, 如果被试没有足够时间观看, 则不太会关注语音的干扰。因此, Huettig 和 McQueen (2007) [17]认为如果被试没有足够的时间进行视觉信息选择, 那么就没有足够的时间进行级联语音激活来匹配听觉输入, de Groot 等(2016) [18]对语义和视觉形式干扰物的研究也有类似发现。Huettig 和 McQueen (2007) [17]同时提出, 如果视觉信息只是打印的文字, 则搜索仅依赖于语音匹配, 作者认为, 这是因为在印刷的单词中, 语音信息与搜索最相关, 而且对被试来说, 完成一个由四个单词组成的基于语音的搜索可能比完成一个仅基于视觉特征的四张图片的搜索更容易。然而, 这个结论的提出显然只是基于表音文字的特点, 作为意音文字的汉语在视觉情境范式下的加工表现还有待考察。

汉语是有声调的语言, 而且存在大量的同音词¹, 人们需要对所听到的词进行辨析才能完成词汇通达的过程。采用视觉情境范式的眼动追踪研究发现, 在高限制性句子语境中, 声调在词汇激活阶段起到了重要作用, 而在低限制性句子语境中, 声调、字形在词汇激活、识别阶段都发生了作用(Rayner *et al.*, 1998 [19]; 王露等, 2018 [20])。因为高水平语音信息能够影响低水平信息的加工过程(Moss & Marslen-Wilson, 1993 [21]; Salverda *et al.*, 2014 [22]), 不过在高限制性语境中词频效应不显著, 而在低限制性语境中存在词频效应(卢张龙等, 2008 [23]; 任桂琴等, 2012 [24])。

Yip 和 Zhai (2018) [25]探究了语境效应对听力理解中双音节口语同音词加工过程的早期影响。实验中, 被试边听汉语同音词结尾的句子边看屏幕, 60 名汉语母语者被分成两组, 一组看呈现的文字, 一组看线描图。研究结果显示, 无论是看文字版本还是看图片版本, 目标词出现前 60 ms 就显示出较强的语境效应, 当被试听到一个多义汉语同音词时, 他们会自动快速地使用上下文进行选择(或预先选择)合适的意义。

3.2.3. 词汇识别的研究

研究人员通常从词汇激活(lexical activation)和竞争效应(competition effect)的角度进行分析。被试在认知目标词之前, 听觉输入的词会激活多个候选词, 它们会在被试的词库中竞争识别。视觉情境眼动追踪可以通过捕捉被试在屏幕上的搜索过程来揭示词汇识别中的微妙激活和竞争效应。

Chambers 等(2009) [26]进行了英法双语者的听力研究, 发现无论法语熟练程度如何, 英语母语者被试在法语听力中听到目标词时都会跳视音位相同但意义不同的英语竞争词。杨雯琴、李荣宝(2012) [27]采

¹ 周勇翔主编的《现代汉语同音词词典》共收录现代汉语中的同音词(包括部分三音节词、成语、熟语等)近 7000 组, 收词总数达 15,000 余条, 2009, 商务印书馆。

用了与 Chambers 等(2009) [26] 相同的范式, 在中英这两种不同类型的语言实验下, 在英语非熟练的中国英语学习者身上也观察到了明显的中英双语间的词汇竞争情况, 发现对中英双语间语音相似的竞争项的注视次数比例显著高于非相似项。这两个实验都体现了语音相似的母语词汇与二语词汇间的竞争, 支持了心理学中双语者(或多语者)的综合词库观点。从二语习得的角度看, 实验结果证实了二语习得中的母语词汇干扰现象, 无论二语学习者的二语水平如何, 两种语言的类型是否相似, 跨语际干扰都存在, 这也为二语(外语)教学与学习提供了重要参考。

3.3. 句法层面的研究

Tanenhaus 等(1995) [4] 也是最早运用眼动追踪技术来研究听理解中句法歧义消解问题的研究, 在那之后, 听理解加工研究中句法层面的研究就日益丰富。

3.3.1. 句法启动的研究

句法启动又称结构启动, 指在几种可替换的结构中, 被试倾向于选择最近刚听到、看到或产出过的那个结构。自 Bock (1986) [28] 的经典句法启动实验以后, 句法启动逐渐成为认知语言学和心理语言学中考查习得与母语句法表征的常用行为工具(陈庆荣, 2012 [29]; 夏赛辉、汪朋, 2017 [30])。而不少学者的研究证实边听边看相结合的视觉情境范式是研究语言句法启动效应的有效途径。Scheepers 和 Crocker (2004) [31] 通过被试的听理解过程考察了德语主动句和被动句的认知过程, 他发现, 听觉输入通道的句子结构特点显著影响着被试对图片的注视轨迹。例如, 如果听力输入句(启动句)是 SVO 结构, 那么被试会更多地注视与受事相关的视觉信息。

Cleland 和 Pickering (2003) [32] 研究了名词短语结构的启动, 发现听者倾向于在对话中重复对方使用的结构, 而且句法结构的重复倾向发生在名词短语层面, 被试在准备名词短语等话语时, 会激活与首要名词相关的概念, 而较少激活与语义相关的概念。Melinguer 和 Cleland (2011) [33] 进一步研究了名词短语在句子中的位置对句法启动效果的影响。研究发现, 与句子结尾处的名词短语相比, 位于句首的名词短语会显著影响后续的句法决策。这项发现显示, 在听力理解过程中, 理解句子或结构的前半部分成分往往更重要, 相应地也需要付出更多的努力。

Arai 等(2007) [34] 研究了视觉输入对听觉理解的影响。该研究共有两个实验, 实验一, 被试大声朗读屏幕上的双宾语句(DO)或者介词宾语句(PO), 然后按键进入下一屏——图片, 图片显示 1000 ms 后, 被试会听到与图片相关的句子, 该句子与被试先前所看到的句子动词相同、结构相同, 但内容不同。实验二, 听觉输入的句子与视觉输入的句子只有结构相同, 动词和内容都不相同。研究发现, 实验一的条件下(句子结构和主要动词相同), 视觉输入的句子(DO/PO)能促进听觉句子(DO/PO)的理解; 实验二的条件下(句子结构相同, 主要动词和句子内容不同), 先前视觉输入的句子(DO/PO)对之后听觉句子(DO/PO)的理解毫无帮助。

然而, 不同于 Arai 等(2007) [34] 的研究, Thothathiri 和 Snedeker (2008) [35] 虽然也采用了相同的范式, 但两者的研究方法和结论有所不同。后者的研究是让被试听完指令后再表演出来, 发现被试听 DO 句的动词后比听到 PO 句的动词更显著表现出对主题的偏好, 他们的结论是在听理解中, 结构相同的句子即使其动词不同, 前句的出现依然有助于后句的加工。Thothathiri 认为之所以与 Arai 等(2007) [34] 得出的结论不同, 可能是分析时间窗口的差异造成的。

3.3.2. 否定句加工的研究

陈广耀等采用眼动视觉情境范式对部分汉语否定句的理解进行了系列研究。陈广耀等(2014) [36] 研究了汉语中类别型状态不确定的独立否定句(如, 裙子不是蓝色的)的听理解加工, 发现当被试听到“裙子不是蓝色的”句子时, 被试倾向于先注视“蓝色裙子”的图片, 再转向注视其他颜色裙子的图片, 也就是被

试在早期对“与事件否定状态一致的图片”注视概率更高，他们会先激活事件的否定状态，然后再进行激活扩散。之后，陈广耀等(2016) [37]又对结果不确定离散型否定句(如，卡通不是米老鼠)的听理解进行了实验，研究结果还是支持了否定句理解的两个阶段模拟加工理论。陈广耀等(2018) [38]进一步对比了强弱语义语境下的否定句加工。强语义语境，如“胳膊不是弯曲的”，备择选项为“伸直的胳膊”；弱语义语境，如“裙子不是蓝色的”，备择选项为“黑色的裙子”。该研究结果显示，否定加工是命题符号加工和经验模拟加工共同作用的结果，支持了符号相互依存假设。

尽管经过系列实验，不过，陈广耀等在研究后也指出，这几个实验材料所选句式相对单一，被试容易造成练习效应。此外，该实验所选事物概念比较具体，且没有生命，人们对于抽象概念的否定句的理解，主语为有生命的人或动物的否定句的加工是否也是如此有待进一步研究。

3.3.3. 复杂句加工的研究

现在，眼动视觉情境范式也开始在二语习得复杂句听理解加工研究中得到运用。Chun 和 Kaan (2019) [39]研究了中国高级英语学习者在处理含有歧义关系分句的复杂句时使用动词语义信息的预测行为。研究结果显示，与英语母语者相似，高水平英语学习者甚至在听到目标词之前就使用动词的语义属性对目标词进行了预期性观察，这与已有的二语语义线索预测效应研究结果相一致。

3.4. 其他非语言信息对听语言理解影响的研究

眼动追踪还可以用来考察其他非语言信息，如，动作、事件以及情景记忆等对听觉语言加工的影响。Chambers 等(2004) [40]的实验表明，物体的动允性(affordance)会影响句法的早期加工过程。Knoeferle 等(2005) [41]的实验也表明，一旦被试听到句中的动词，就会马上对视觉场景中的题元角色进行分配。Sussman 和 Sedivy (2003) [42]、Chambers 和 Juan (2008) [43]发现，被试对听觉通道的语言内容预期不仅基于物体的特点，还会受视觉场景情节记忆的影响。

4. 小结与展望

4.1. 眼动视觉情境范式生态有效

不同语言的实验研究表明，眼动视觉情境范式是研究听力理解中语音、词汇识别和句子理解即时认知过程的有效方法，不仅是研究成年人听-看语言加工的流行技术(Huettig *et al.*, 2011 [44])，也是研究二语(外语)习得的理想技术(van Hell & Tanner, 2012 [45]; Chamber & Cooke, 2009 [26]; Farris-Trimble & McMurray, 2013 [46])。

转向眼动追踪已经成为了目前第二语言习得研究的一大趋势，该趋势强调研究中并行数据的收集(Aline, 2020: p. 23) [2]。眼动视觉情境范式能让被试在不受干扰的自然条件下听录音、阅读文本、观看视频、选择判断、移动物体、面对面交流等操作，既能测量即时眼睛对视觉刺激的反应，又能同时测量来自听觉刺激的反应，能根据眼睛注视模式和眼睛注视时间对正在进行的语言加工提供精确的时间锁定和精细的测量，比如测试数据、问卷、有声思维能提供更多信息。其优势不仅在于理解言语是如何映射到视觉指示物上的，还在于理解听觉、视觉、注意、记忆、语言和任务目标是如何相互作用的，从而为研究一般认知而不仅仅是语言处理提供了丰富的前景(Sajin, 2016: p. 2) [47]。

4.2. 立足于二语听力理解的研究展望

4.2.1. 融合心理学理论与技术

语言研究的多学科交叉并不是今天才兴起的。自古希腊时期之后，我们对于语言的整体理论思考，包括主要研究方法的创制，是在学科之间借鉴、补充、融合之中一步一步地走到今天的。一部语言学理

论研究史就是一部语言交叉学科研究发展史(陈平, 2021) [48]。

眼动追踪技术源于心理学领域, 虽然近年来该技术在语言学领域的运用有日渐增长的趋势, 但大部分视觉情境范式的心语言学研究还是围绕心理学的主题进行。比如, 句法歧义加工、语音感知、语义、主题和视觉形态竞争效果、句子加工预测和语言加工的个体差异, 及心理语言学的其他主题(Sajin, 2016: p. 2) [47]。

鉴于眼动视觉情境范式在听理解研究中的优势, 二语习得研究者和语言学者有必要主动融合心理学的最新理论和眼动技术到二语言习得研究中, 扩展第二语言习得的研究主题。未来的研究可以进一步探讨以下具体问题。

二语学习者的认知机制: 通过眼动追踪技术, 研究二语学习者在不同语言任务中的认知加工路径, 揭示其与母语者的差异和相似性。

多模态语言加工: 结合语音、视觉、动作等多种模态, 研究二语学习者在复杂语言环境中的信息整合机制。

跨文化语境下的语言加工: 探讨不同文化背景下的二语学习者在语言加工中的文化适应性和认知策略。

4.2.2. 具体研究方向与建议

尽管眼动视觉情境范式为语言研究提供了新的视角, 但当前研究仍存在一些不足, 未来的研究可以在以下方面进行深入探索。

1) 二语学习者与母语者的对比研究

眼动模式的异同: 系统比较二语学习者和母语者在听理解过程中的眼动模式, 分析其差异是否源于语言特征、认知能力还是学习经历。例如, 研究汉语作为二语学习者在处理汉语否定句、复杂句时的眼动模式, 与母语者进行对比, 揭示其加工机制的差异。

语言特征与认知机制的关系: 探讨二语的语言特征如何影响二语学习者的认知机制。例如, 研究汉语声调对二语学习者词汇识别的影响, 以及不同语言类型(如表音文字与意音文字)在视觉情境范式下的加工表现。

认知能力的影响: 考察二语学习者的认知能力(如工作记忆、注意力控制)对其听理解加工的影响, 以及这些认知能力如何与语言加工相互作用。

2) 二语听力教学的有效性研究

教学方法的优化: 通过眼动追踪技术, 研究不同听力教学方法(如预测训练、语境化教学)对二语学习者听理解能力的提升效果。例如, 设计实验比较在有无语境支持的情况下, 二语学习者对复杂句的理解效果, 从而为听力教学提供科学依据。

测试与评估: 研究当前的二语听力测试是否能有效衡量学习者的语言知识和语言能力。例如, 通过眼动数据评估大学英语听力测试题目的难度和效度, 提出改进测试内容和形式的建议。

个性化教学: 基于眼动追踪数据, 分析不同水平二语学习者的听理解特征, 为个性化教学提供支持。例如, 根据学习者在眼动数据中表现出的词汇识别困难或句法加工问题, 设计针对性的教学方案。

3) 语言与非语言信息的整合研究

语境对语言加工的影响: 进一步研究语境(如视觉场景、动作信息)对二语学习者语言加工的影响机制。例如, 通过眼动实验研究视觉场景中的物体动允性如何影响二语学习者的句法加工。

跨模态信息的整合: 探索二语学习者如何整合来自不同模态(如语音、视觉、动作)的信息进行语言理解。例如, 研究在视听结合的条件下, 二语学习者如何利用视觉线索预测和验证听觉信息。

4) 汉语语言知识对二语学习的作用

汉语语言特征的加工优势：研究汉语的语言特征(如声调、语义丰富性)如何帮助二语学习者在听理解中进行预测和加工。例如，探讨汉语声调在二语词汇识别中的作用，以及如何通过汉语语言知识促进二语学习者的语义预测能力。

汉语背景下的二语教学策略：基于汉语语言特点，设计适合中国二语学习者的听力教学策略。例如，研究如何利用汉语的语义网络结构帮助二语学习者更好地理解和记忆词汇。

5) 技术与方法的创新

多模态数据的融合：结合眼动追踪、脑成像(如 EEG、fMRI)等技术，全面揭示二语学习者的语言加工机制。例如，通过眼动和脑电数据的同步采集，研究二语学习者在句法加工中的时间动态和神经基础。

大数据与人工智能的应用：利用大数据分析和人工智能技术，对眼动数据进行深度挖掘，发现潜在的语言加工规律。例如，通过机器学习算法分析大量眼动数据，预测二语学习者的语言水平和学习困难点。

通过以上具体的研究方向和建议，未来的研究可以更加系统地探索二语听力理解的机制，为二语教学和学习提供更具针对性的理论支持和实践指导。

基金项目

浙江省社会科学联合会研究性课题“基于眼动追踪的汉语比较范畴认知加工研究”，项目号为2024N105。

参考文献

- [1] Bojko, A. 眼动追踪: 用户体验优化操作指南[M]. 葛缨, 何吉波, 译. 北京: 人民邮电出版社, 2019: 9.
- [2] Aline, G. (2020) Eye Tracking in Second Language Acquisition and Bilingualism. Routledge, 88.
- [3] 闫国利, 白学军. 眼动分析技术的基础与运用[M]. 北京: 北京师范大学出版社, 2018: 11, 116.
- [4] Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M. and Sedivy, J.C. (1995) Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science*, **268**, 1632-1634. <https://doi.org/10.1126/science.7777863>
- [5] Kamide, Y., Scheepers, C. and Altmann, G.T.M. (2003) Integration of Syntactic and Semantic Information in Predictive Processing: Cross-Linguistic Evidence from German and English. *Journal of Psycholinguistic Research*, **32**, 37-55. <https://doi.org/10.1023/a:1021933015362>
- [6] Altmann, G.T.M. and Kamide, Y. (2007) The Real-Time Mediation of Visual Attention by Language and World Knowledge: Linking Anticipatory (and Other) Eye Movements to Linguistic Processing. *Journal of Memory and Language*, **57**, 502-518. <https://doi.org/10.1016/j.jml.2006.12.004>
- [7] Allopenna, P.D., Magnuson, J.S. and Tanenhaus, M.K. (1998) Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, **38**, 419-439. <https://doi.org/10.1006/jmla.1997.2558>
- [8] Dahan, D., Magnuson, J.S. and Tanenhaus, M.K. (2001) Time Course of Frequency Effects in Spoken-Word Recognition: Evidence from Eye Movements. *Cognitive Psychology*, **42**, 317-367. <https://doi.org/10.1006/cogp.2001.0750>
- [9] Simmons, E.S. and Magnuson, S. (2018) Word Length, Proportion of Overlap, and Phonological Competition in Spoken Word Recognition. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*. Madison, 25-28 July 2018, 1064-1069.
- [10] Hendrickson, K., Spinelli, J. and Walker, E. (2020) Cognitive Processes Underlying Spoken Word Recognition during Soft Speech. *Cognition*, **198**, Article 104196. <https://doi.org/10.1016/j.cognition.2020.104196>
- [11] 袁周敏, 易保树, 张砚妮. 国外语言学学科眼动研究: 现状与前瞻(1934-2020) [J]. 外语界, 2020(6): 46-53.
- [12] Tanenhaus, M.K. and Spivey-Knowlton, M.J. (1996) Eye-Tracking. *Language and Cognitive Processes*, **11**, 583-588. <https://doi.org/10.1080/016909696386971>
- [13] Huettig, F. and Altmann, G.T.M. (2005) Word Meaning and the Control of Eye Fixation: Semantic Competitor Effects and the Visual World Paradigm. *Cognition*, **96**, B23-B32. <https://doi.org/10.1016/j.cognition.2004.10.003>
- [14] Chow, J., Aimola Davies, A. and Plunkett, K. (2017) Spoken-Word Recognition in 2-Year-Olds: The Tug of War between

Phonological and Semantic Activation. *Journal of Memory and Language*, **93**, 104-134.
<https://doi.org/10.1016/j.jml.2016.08.004>

[15] Altmann, G.T.M. and Kamide, Y. (1999) Incremental Interpretation at Verbs: Restricting the Domain of Subsequent Reference. *Cognition*, **73**, 247-264. [https://doi.org/10.1016/s0010-0277\(99\)00059-1](https://doi.org/10.1016/s0010-0277(99)00059-1)

[16] Yee, E. and Sedivy, J.C. (2006) Eye Movements to Pictures Reveal Transient Semantic Activation during Spoken Word Recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **32**, 1-14.
<https://doi.org/10.1037/0278-7393.32.1.1>

[17] Huettig, F. and McQueen, J.M. (2007) The Tug of War between Phonological, Semantic and Shape Information in Language-Mediated Visual Search. *Journal of Memory and Language*, **57**, 460-482. <https://doi.org/10.1016/j.jml.2007.02.001>

[18] de Groot, F., Huettig, F. and Olivers, C.N.L. (2016) When Meaning Matters: The Temporal Dynamics of Semantic Influences on Visual Attention. *Journal of Experimental Psychology: Human Perception and Performance*, **42**, 180-196.
<https://doi.org/10.1037/xhp0000102>

[19] Rayner, K., Fischer, M.H. and Pollatsek, A. (1998) Unspaced Text Interferes with Both Word Identification and Eye Movement Control. *Vision Research*, **38**, 1129-1144. [https://doi.org/10.1016/s0042-6989\(97\)00274-5](https://doi.org/10.1016/s0042-6989(97)00274-5)

[20] 王露, 万鹏, 任桂琴. 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[J]. 心理与行为研究, 2018, 16(2): 150-156.

[21] Moss, H.E. and Marslen-Wilson, W.D. (1993) Access to Word Meanings during Spoken Language Comprehension: Effects of Sentential Semantic Context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19**, 1254-1276. <https://doi.org/10.1037/0278-7393.19.6.1254>

[22] Salverda, A.P., Kleinschmidt, D. and Tanenhaus, M.K. (2014) Immediate Effects of Anticipatory Coarticulation in Spoken-Word Recognition. *Journal of Memory and Language*, **71**, 145-163. <https://doi.org/10.1016/j.jml.2013.11.002>

[23] 卢张龙, 白学军, 闫国利. 汉语词汇识别中词频和可预测性交互作用的眼动研究[J]. 心理研究, 2008, 1(4): 29-33.

[24] 任桂琴, 韩玉昌, 刘颖. 句子语境中汉语词汇识别的即时加工过程[J]. 心理科学进展, 2012, 20(4): 493-503.

[25] Yip, M.C.W. and Zhai, M. (2017) Context Effects and Spoken Word Recognition of Chinese: An Eye-Tracking Study. *Cognitive Science*, **42**, 1134-1153. <https://doi.org/10.1111/cogs.12570>

[26] Chambers, C.G. and Cooke, H. (2009) Lexical Competition during Second-Language Listening: Sentence Context, but Not Proficiency, Constrains Interference from the Native Lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **35**, 1029-1040. <https://doi.org/10.1037/a0015901>

[27] 杨雯琴, 李荣宝. 中国英语学习者母语加工中的二语竞争激活——来自眼动研究的证据[J]. 外语研究, 2012, 29(4): 34-39, 112.

[28] Bock, J.K. (1986) Syntactic Persistence in Language Production. *Cognitive Psychology*, **18**, 355-387.
[https://doi.org/10.1016/0010-0285\(86\)90004-6](https://doi.org/10.1016/0010-0285(86)90004-6)

[29] 陈庆荣. 句法启动研究的范式及其在语言理解中的争论[J]. 心理科学进展, 2012, 20(2): 208-218.

[30] 夏赛辉, 汪朋. 句法启动与二语复杂结构学习[J]. 现代外语, 2017, 40(1): 69-79.

[31] Scheepers, C. and Crocker, M.W. (2004) Constituent Order Priming from Reading to Listening: A Visual-World Study. In Carreiras, M. and Clifton, C., Eds., *The On-Line Study of Sentence Comprehension: Eye-Tracking, ERP, and Beyond*, Psychology Press, 167-187.

[32] Cleland, A. (2003) The Use of Lexical and Syntactic Information in Language Production: Evidence from the Priming of Noun-Phrase Structure. *Journal of Memory and Language*, **49**, 214-230.
[https://doi.org/10.1016/s0749-596x\(03\)00060-3](https://doi.org/10.1016/s0749-596x(03)00060-3)

[33] Melinger, A. and Cleland, A.A. (2011) The Influence of Sentential Position on Noun Phrase Structure Priming. *Quarterly Journal of Experimental Psychology*, **64**, 2211-2235. <https://doi.org/10.1080/17470218.2011.586709>

[34] Arai, M., van Gompel, R.P.G. and Scheepers, C. (2007) Priming Ditransitive Structures in Comprehension. *Cognitive Psychology*, **54**, 218-250. <https://doi.org/10.1016/j.cogpsych.2006.07.001>

[35] Thothathiri, M. and Snedeker, J. (2008) Give and Take: Syntactic Priming during Spoken Language Comprehension. *Cognition*, **108**, 51-68. <https://doi.org/10.1016/j.cognition.2007.12.012>

[36] 陈广耀, 张维, 陈庆, 等. 类别型状态不确定独立否定句的加工机制: 来自眼动实验的证据[J]. 心理学报, 2014, 46(10): 1426-1441.

[37] 陈广耀, 陈颖心, 邓玉梅, 等. 结果不确定离散型否定句加工机制: 锚激活与限制满足模型的修正与补充[J]. 心理科学, 2016, 39(5): 1064-1070.

[38] 陈广耀, 何先友, 刘涛. 强弱语义语境下的否定句加工机制[J]. 心理学报, 2018, 50(2): 186-196.

[39] Chun, E. and Kaan, E. (2019) L2 Prediction during Complex Sentence Processing. *Journal of Cultural Cognitive Science*,

3, 203-216. <https://doi.org/10.1007/s41809-019-00038-0>

[40] Chambers, C.G., Tanenhaus, M.K. and Magnuson, J.S. (2004) Actions and Affordances in Syntactic Ambiguity Resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **30**, 687-696. <https://doi.org/10.1037/0278-7393.30.3.687>

[41] Knoeferle, P., Crocker, M.W., Scheepers, C. and Pickering, M.J. (2005) The Influence of the Immediate Visual Context on Incremental Thematic Role-Assignment: Evidence from Eye-Movements in Depicted Events. *Cognition*, **95**, 95-127. <https://doi.org/10.1016/j.cognition.2004.03.002>

[42] Sussman, R.S. and Sedivy, J. (2003) The Time-Course of Processing Syntactic Dependencies: Evidence from Eye Movements. *Language and Cognitive Processes*, **18**, 143-163. <https://doi.org/10.1080/01690960143000498>

[43] Chambers, C.G. and Juan, V.S. (2008) Perception and Presupposition in Real-Time Language Comprehension: Insights from Anticipatory Processing. *Cognition*, **108**, 26-50. <https://doi.org/10.1016/j.cognition.2007.12.009>

[44] Huettig, F., Rommers, J. and Meyer, A.S. (2011) Using the Visual World Paradigm to Study Language Processing: A Review and Critical Evaluation. *Acta Psychologica*, **137**, 151-171. <https://doi.org/10.1016/j.actpsy.2010.11.003>

[45] van Hell, J.G. and Tanner, D. (2012) Second Language Proficiency and Cross-Language Lexical Activation. *Language Learning*, **62**, 148-171. <https://doi.org/10.1111/j.1467-9922.2012.00710.x>

[46] Farris-Trimble, A. and McMurray, B. (2013) Test-Retest Reliability of Eye Tracking in the Visual World Paradigm for the Study of Real-Time Spoken Word Recognition. *Journal of Speech, Language, and Hearing Research*, **56**, 1328-1345. [https://doi.org/10.1044/1092-4388\(2012/12-0145\)](https://doi.org/10.1044/1092-4388(2012/12-0145)

[47] Sajin, S.M. (2016) Language-Vision Interactions in the Visual World Paradigm: Examining the Effects of Display and Working Memory. Ph.D. Thesis, Binghamton University.

[48] 陈平. 语言交叉学科研究的理论与实践[J]. 语言战略研究, 2021, 6(1): 13-25.