

不同类型噪声环境下言语理解的脑机制研究*

耿立波

江苏师范大学
语言科学与艺术学院
/江苏 徐州
libogeng@163.com

薛紫炫

陕西师范大学
文学院
/陕西 西安
1484688752@qq.com

杨亦鸣¹

江苏师范大学
语言科学与艺术学院
/江苏 徐州
yangym@jsnu.edu.cn

摘要

文章使用 ERP 技术, 对比分析了安静、白噪声、汉语噪声、英语噪声四种听觉条件下, 听力正常的汉语母语者加工汉语句子的情况, 以探究信息掩蔽条件下语义加工的神经机制。研究发现不同噪声条件下诱发的 N100、N400、LPC 等 ERPs 成分具有不同的波形表现, 据此本文得出以下结论: 首先, 在语音掩蔽条件下, 对于难度较大的语义加工, 目标语音与掩蔽噪声在知觉层面的相似性并非主要影响因素, 而掩蔽噪声语义内容上的可懂度发挥着更关键的作用。其次, 当言语噪声为听者极其熟悉或完全陌生的语言, 其对语义加工的掩蔽干扰较小, 而当掩蔽噪声为听者接触过的语言但不是母语或主要语言, 其掩蔽效应可能更强。最后, 不熟悉的言语噪声中所包含的出现频率较小但能够被听者理解的语义内容, 与听者的预期相冲突, 引发听者的注意转移, 这些语义信息被传输至听觉中枢神经, 占用了原本用于目标刺激的认知资源, 从而增强了信息掩蔽的效果。

关键词: 信息掩蔽; 语义加工; 噪声类型; 言语可懂度; 事件相关电位

Brain Mechanism of Speech Comprehension in Different Noise Conditions

Abstract

By means of ERP, this paper compares the processing of Chinese sentences in quiet condition, white noise condition, Chinese noise condition and English noise condition, in order to explore the neural mechanism of semantic processing under information masking condition. It is found that the waveforms of N400, LPC and other ERPs induced by different noise conditions are different, based on which the following conclusions are drawn. Firstly, under speech masking condition, the perceptual similarity between target speech and masker is not the main factor for difficult semantic processing, while the intelligibility of semantic content of masker plays a more critical role. Secondly, the masking effect on semantic processing is smaller when the language is a very familiar or completely unfamiliar language, while the masking effect may be stronger when the masking noise is the language the listener has been exposed to but not the native language. Finally, The semantic content contained in the unfamiliar speech noise that appears less frequently but can be understood by the listener conflicts with the expectation of the listener, causing the attention transfer. These semantic information is transmitted to the

* 国家重点基础研究发展计划 (973)(批准号: 2014CB340502)、国家社科基金青年项目 (批准号: 16CYY021) 资助

¹ 通信作者. E-mail: yangym@jsnu.edu.cn

auditory central nerve, occupying the cognitive resources originally used for target stimulation, thus enhancing information masking intelligibility.

Keyword: Information masking, Semantic processing, Noise type, Speech intelligibility, Event-related potential

1 引言

所谓“噪声”，是指一切妨碍人们正常休息、学习和工作的声音。噪声无处不在，人们的口头言语交际行为往往会受到噪声的干扰，这一现象被称为“噪声掩蔽(noise masking)”。噪声掩蔽研究对人工智能的自然语言处理技术的发展有着积极的推动作用。以机器语音识别为例，2015年后，由于端到端技术兴起，语音识别系统的性能大幅度提升，2017年微软在 Switchboard 上词错误率降至 5.1%，在语音识别的精确性上首次超越了人类。很显然这是在一定限定条件下的结果，如果增设噪声使训练环境与真实的语言使用环境失配，即使是目前最先进的计算机模型，其语音识别性能也将急剧下降，根本无法与人类的听觉系统媲美。若要突破这一技术障碍，必须了解人脑在噪声环境下的言语加工机制，然后在此基础上赋予机器以人脑的功能。

噪声掩蔽的研究始于能量掩蔽，这种掩蔽效应发生在听觉外周，被认为是目标声音与背景噪声在时间和频率上重叠的结果(Miller,1947)。这一阶段的研究多以噪音(noise)²、音调(tone)等纯音作为背景噪声，关注频谱结构、强度、连续性等声学特征。随着“有信息内容”的语音掩蔽的使用，产生了能量掩蔽理论和模型不能涵盖的掩蔽效应，即信息掩蔽效应。信息掩蔽发生在听觉中枢神经，涉及更高级的加工层面(Kidd et al.,2008)。信息掩蔽的性质一直存在争议，越来越多的证据表明，掩蔽语音的语义内容是影响信息掩蔽的关键因素(许家铭等, 2020)。这一问题如果得到解决，将会为人工智能应用技术创造非常重要的理论价值，毕竟机器要真正地理解人类的语言，其重点是要理解语义，尤其是噪声条件下的语义理解。

2 信息掩蔽效应

2.1 信息掩蔽的诱因

混叠语音掩蔽的最新研究发现，信息掩蔽包含知觉与认知两个水平，在知觉水平层面，背景噪声与目标声音的相似度占主导作用，而在认知水平，背景噪声的可懂度与信息掩蔽的可懂度密切相关(杨志刚, 2008,2014)。

在知觉水平，目标声音与掩蔽噪声的相似性是影响音系掩蔽的重要因素。一般认为，目标声音与背景噪声在知觉特征、频率、呈现方式(单耳或双耳)等方面相似均会增强信息掩蔽(Brungart,2001;Kidd et al.,2008)。如 Brungart 等 (2001)发现目标声音与掩蔽噪声的发音者为同性的条件下产生的信息掩蔽强于目标声音与掩蔽声音的发音者为异性的条件。

前人关于相似性的判定多限于声学特征范围，随着“有信息内容”的言语噪声的使用，相似性的解释似乎难以适用。语言作为一个符号系统，其下包含词汇、语义、语法、语音等子系统，相似性的判定难以给出客观化的标准。而且，在语音掩蔽条件下，与语义加工相关的可懂度、熟悉度等因素似乎更为关键。

普遍研究认为，能被听者理解的言语噪声会产生更强的掩蔽效果，也就是说，在目标语音为听者母语的前提下，以听者的母语或熟悉的语言作背景噪声，信息掩蔽的强度最大，当掩蔽语音为听者接触过的语言但不是母语或主要语言，其掩蔽效果小于那些更容易被识别出来的掩蔽噪声(Brouwer et al., 2012)。

Van Engen(2007)为探究语音掩蔽的不利影响是来自声学特征还是掩蔽噪声的语义内容，要求英语母语者分别在包含两个说话人的英语噪声、包含六个说话人的英语噪声、包含两个说话人的普通话噪声、包含六个说话人的普通话噪声的条件下，对带有母语口音的英语句子进行理解。

² “噪音”区别于“噪声”，前者为物理学概念，指发声体做无规则振动时发出的声音。

结果发现, 具有语言内容的掩蔽对听者的影响不同于非语言掩蔽, 且在包含两个说话人的掩蔽条件下, 母语为英语的听者受到英语掩蔽的负面影响要大于汉语听者。Calandruccio(2013)的实验设置了英语、荷兰语和普通话可懂度不同的三种掩蔽噪声, 以英语单语者在句子识别任务中的成绩为依据, 结果显示英语掩蔽、荷兰语掩蔽、普通话掩蔽条件下目标句子的识别表现呈梯度提升。Calandruccio 将这一结果更多地归于目标语音与背景噪声的相似度上的差异, 但无法排除掩蔽噪声的语义信息对听者的干扰。许博渐(2015)考察言语可懂度对信息掩蔽的影响, 要求被试复述目标刺激(汉语无意义语句), 同时设置了可懂度不同的三类掩蔽噪声, 即英语无意义句、汉语无意义句、汉语有意义句。结果显示以汉语有意义句为掩蔽噪声条件下的复述正确率显著低于以汉语无意义句为掩蔽噪声的条件, 同时显著低于以英语无意义句为掩蔽噪声的条件。

以上研究均在一定程度上支持了这样一种观点, 即当掩蔽噪声为听者的母语或熟悉的语言时, 掩蔽噪声的语义内容成为影响信息掩蔽的关键因素, 掩蔽强度可能随噪声可懂度的增强而增大。不过这些研究都未能清楚地解释, 在混叠语音掩蔽条件下, 背景噪声对目标语音的掩蔽效应是源于掩蔽噪声在语言上的干扰, 还是源于目标语音和掩蔽噪声在声学或语音上的相似。而且最新的研究认为, 以听者熟悉的语言和不熟悉的语言作为掩蔽, 掩蔽效果相对较小, 而当掩蔽语音为听者接触过的语言但不是母语或主要语言, 掩蔽的情况将变得更加复杂, 掩蔽效果与背景噪音的可懂之间不可能仅是简单的正比关系。

除相似性、可懂度外, 背景噪声的不确定性也是影响信息掩蔽的重要因素。所谓的不确定性是指在同一次实验的不同试次间的刺激模式的随机性, 如随机改变掩蔽声音的频率、时长或起始时间以及空间位置等。Neff 与 Green(1987)的实验发现缩短掩蔽声音与目标声音起始时间的差异, 听者感知目标声音的表现会随之提高; Kidd 等(1998)指出通过随机改变掩蔽声音的频谱, 可提高特定目标纯音的阈限; Allen 等(2011)等以言语信号为材料, 发现掩蔽声音位置的随机化会使听者对目标音节的识别阈限增加。

与相似性的研究情况类似, 针对背景噪声的可懂度对信息掩蔽的影响, 前人研究多以合成纯音作为刺激, 将不确定性限制在声学特征的范围。依据最新研究成果, 掩蔽语音在语义内容上的不确定性刺激也可能诱发信息掩蔽, 而如何判定语义内容上的不确定刺激, 这是一个值得探究的问题。

2.2 信息掩蔽的神经机制

一个主要的观点认为在信息掩蔽条件下, 言语噪音占用了目标语音的知觉与认知资源。如过程干扰假说(interference-by-process hypothesis)认为目标语音与言语噪声认知加工过程的冲突或重叠导致了信息掩蔽效应(Oksama et al., 2016)。用于听觉加工的认知资源的数量相对固定, 如果进入听觉系统的有效语音流越多, 那么被占用的认知资源也就更多, 目标语音流能获得的资源相对变少。

Nakai(2005)要求被试听一名女性朗读故事, 不同的 block 中使用不同的背景噪声, 包括没有噪声、一名男性说话者产出的语音(DV)、与目标刺激同一名女性说话者产出的语音(SV)。与没有噪声的基线条件对比, DV 条件下双侧 STG 激活, 这在一定程度上说明中枢皮层听觉区域既处理了目标语音又处理了噪声。在 SV 条件下, 双侧颞叶、前额叶和顶叶区域的激活明显增强, 且与 DV 条件相比, SV 条件在补充前语言区(pre-SMA)、右顶叶和双侧前额叶区域的激活更强烈, 研究认为这些区域的激活是为弥补 SV 条件下噪声与目标语音对中枢资源的激烈竞争, 也就是说, SV 条件下的信息掩蔽效应更强烈。

Gutschalk 等(2008)要求听者在时间和频率都随机的音调噪声中检测有规律重复的目标音调, 脑磁图结果显示在早期时间窗内, 无论被试是否感知到了音调信号, 反应都存在且波形强烈, 但到后期时间窗(50-250 msec), 被试的反应则与目标检测之间存在显著相关, 这表明听觉皮层在后期处理阶段参与了需要更多认知处理的任务。Scott(2013)指出噪声条件下的听觉言语加工之所以会产生额外的前额叶、顶叶和扣带皮层的活动, 主要在于听者对噪声的感知占用了额

外的认知控制网络。

此外，注意在语音掩蔽中也发挥着重要的作用。首先，听觉客体的形成与选择都依赖于选择性注意，在复杂的听觉条件下，听者必须从不同的声源中选择目标声音或客体去加工，这一过程往往由内源性注意和外源性注意共同决定，前者取决于观察者的目标，是自上而下目标驱动的过程，后者受刺激的显著性影响，是自下而上刺激驱动的过程。相关 ERP 证据表明，在多人谈话的场景下，听者需要耗费几秒钟以在语流中集成连续音调的过程中集中注意力。

其次，背景噪声会造成听者的注意转移。Nakai(2005)的研究显示，当信息掩蔽效应增加时，非听觉区域的活动会增加，这可能反映了大脑为应对在噪声条件下注意力需求的增加而进行了补偿性活动。听觉交际场景的变化(如目标说话者的变换)需要重新分离或形成听觉客体，即要进行注意力的重新聚焦，这便需要投入额外的切换成本，结果会导致听者言语知觉的下降，表现为反应时间更长、错误率更高(Bohan et al., 2018)。

神经科学研究发现，部分与注意相关的脑区在语音掩蔽条件下也会激活。如听觉掩蔽的 PET 及 fMRI 研究有一个一致的发现，当听者对目标声音进行选择性的反映，随着语流数量增加，左前颞叶激活最显著，主要位于双边颞上回(superior temporal gyrus, STG)/颞上沟(superior temporal sulcus, STS)，这一区域的激活表明注意选择加工参与了这一过程。Scott 等(2013)发现与复杂的非语音基线(如旋转语音)相比，听者在语音噪声条件下的选择性反应沿着 STS 向前延伸，就偏侧性而言，这些反应的振幅在左半球更高、更稳固，这表明左侧颞叶具有一种与选择注意相关的特殊的功能，而先前研究已经证明，注意力控制主要激活左半球的相关脑区，包括额下回(inferior frontal gyrus, IFG)、上顶叶(superior parietal lobe, SPL)和顶内沟(intraparietal sulcus, IPS)(吕思慧等, 2020)。

2.3 小结

前人研究仍然存在一些问题：首先，注意与语音掩蔽的关系仍有待深入探究，现有的关于注意的理论多来自视觉研究，听觉注意的证据相对较少，难免出现偏颇；其次，前人研究侧重考察背景噪声的声学特征，对其语义内容的关注较弱，而且对掩蔽条件下句子及语篇等语言单位的加工不足，这也导致噪声掩蔽研究在语言学方面的意义稍显薄弱；再次，侧重于能量掩蔽与信息掩蔽的对比，缺乏信息掩蔽内部的考察，即对不同条件下的信息掩蔽效应的对比，因此对信息掩蔽效应的认识不全面；最后，多数实验对掩蔽语音的设置两极化，如有意义汉语语句或无意义汉语语句，可懂度极高或极低，缺少过渡性的可懂度适中的掩蔽。

总之，前人研究较一致地认为信息掩蔽涉及认知加工，必然会消耗认知与注意资源，但很难确定导致资源竞争的诱因。根据前人一般观点，不论是从相似性还是从可懂度出发，对于汉语母语者，当目标刺激为汉语句子时，汉语噪声的掩蔽效应应该比英语噪声大，因为汉语噪声与目标汉语句子更相似，可懂度也最高。据此，本研究试图验证这一观点，并考察在不同背景噪声条件下，听觉言语加工尤其是语义加工在神经机制上的差异。

3 研究方法

本研究采用 ERP 听觉实验，语义合理性判断任务，句子类型(语义合理句/语义违反句)×听觉条件(安静/白噪声/汉语噪声/英语噪声)被试内两因素混合设计。

3.1 实验被试

本实验共招募 15 名汉语母语者，年龄在 23-28 岁之间。被试均为右利手，视力或矫正视力均在 1.0 以上，双侧听力正常，无耳科疾病、脑外伤或神经系统病史。正式实验前对被试的英语水平和注意力水平进行测评，所有参与实验的被试的综合英语水平均属于 intermediate level³、注意力测评分数均达标(殷恒婵, 2003)。

³ 参照《Oxford University Press and University of Cambridge Local Examinations Syndicate》英语水平量表。

3.2 实验语料

1) 噪声语料

汉语噪声由四名普通话标准的汉语母语者录制；英语噪声则由四名英语母语者录制，为平衡性别因素的影响，每种言语噪声的产出者均为两男两女。在隔音的录音室，每位录音者大声朗读一篇指定的文章选段⁴。所有语音材料均被转换为 WAV 格式，并使用 Cool Edit 2.1 及 Praat 语音编辑软件分别对汉语噪声和英语噪声进行剪切及参数设置，尺度音强(scale intensity)设置为 68dB SPL，采样率设置为 44100Hz。

2) 目标语料

实验采用听觉语义判断任务，语料设计参考张华等(1990)，共包含 160 个语义合理句和 160 个语义违反句。所有语料均为包含 5 个关键词的简单句，句法结构为“修饰语+主语+谓语+修饰语+宾语”或“主语+状语+谓语+修饰语+宾语”。语义违反句通过调换语义合理句的句末词获得。经过语料接受度调查，语义合理句的平均可接受度达到 80%及以上，语义违反句的平均语义违反性达到 80%及以上。目标语料由一名普通话标准的汉语母语者(女性，年龄 23 岁)在隔音的录音室录制，采样频率设置为 44100Hz、分辨率设置为 16bit。目标语料示例见表 1：

条件	例子
语义合理句	无情的病魔吞噬了年轻的生命
语义违反句	无情的病魔吞噬了年轻的手势

表 1.语义判断实验语料

目标语料与噪声语料叠加，目标语料呈现水平为 65 dB SPL，噪声语料呈现水平为 68 dB SPL，信噪比为-3 dB。

3.3 实验程序

电脑屏幕中央呈现实验指导语，被试阅读后按任意键开始实验。正式实验开始，电脑屏幕中央首先出现一个 1000ms 的“+”，同时插入式耳机播放噪声，1000ms 后播放目标句，目标句时长约为 3000-5000ms，目标句呈现完毕，屏幕中央出现一个“???”，即开始按键判断，语义合理按“1”键，语义违反按“3”键，被试按键后“???”即刻消失。

判断屏的最大时长为 2000ms，若被试未能在规定时间内做出判断，判断屏呈现结束后自动切换，每个 trial 之间有一个 400-600ms 的随机间隔。实验共分为 4 个 block，每个 block 随机对应一种听觉条件，每个 block 内的语料随机呈现，前后两个 block 之间有短暂休息，但噪声不终止。具体内容见图 1：

⁴ 汉语噪声文本均选自《散文海外版》；英语噪声文本均选自《English Digest》。经测评，所有文章不存在过时或有争议的内容。

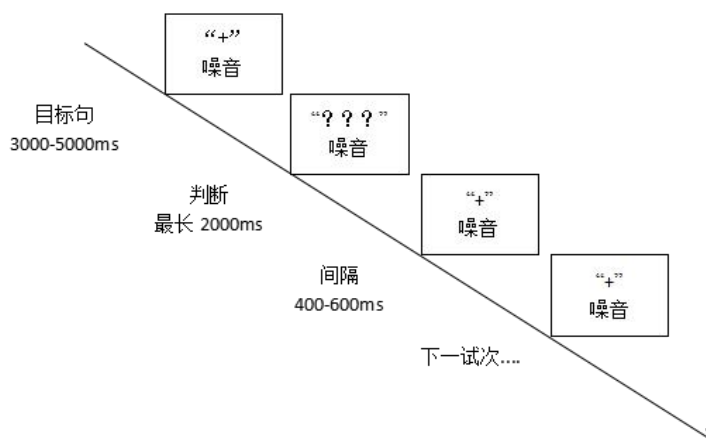


图 1.语义判断实验流程

3.4 脑电记录

实验采用 E-prime1.0 编程，Neuroscan 64 导 Ag/AgCl 电极帽记录脑电，电极按国际 10-20 系统放置。在线记录脑电时，以鼻尖为参考，接地点在大脑中线 FZ 和 FPz 的中点，左眼上下垂直位置记录垂直眼电，双眼外侧眼角水平位置记录水平眼电。实验中所有电极与头皮之间的电阻降至 $5k\Omega$ 以下。在线采样率为 1000Hz，滤波带通为 0.05-100Hz。

3.5 数据处理与分析

使用 Neuroscan 4.3 对脑电数据进行离线分析处理。首先剔除波形杂乱或被污染的脑电并去眼动，接着进行脑电分段，分析时程为刺激呈现前 100ms 至刺激呈现后 1000ms，随后进行第一次基线校正，设置 30Hz(24dB/oct)的低通滤波，然后进行去伪迹处理、第二次基线校正，最后对安静、白噪声、汉语噪声与英语噪声，语义合理与语义违反八种条件下的脑电数据进行分类叠加平均。

将所有有效被试的脑电进行组平均叠加处理，根据采集到的 ERPs 总平均波形图，主要诱发了 N100、N400、LPC 成分，选取的时间窗为 100-250ms、350-500ms、700-850ms，选取的电极点包括左前脑区、左后脑区、右前脑区、右后脑区、中间脑区(CZ、C1、C3、C2、C4、PZ、P1、P3、P2、P4)。使用 SPSS 22.0 软件进行重复测量方差分析(repeated measures analysis of variance, RANOVA)，并使用 Greenhouse-Geisser 法对不符合球形假设的检验进行校正。对涉及到实验条件且具有统计学意义的交互作用进行简单效应分析，采用 Bonferroni 方法进行多重比较校正，显著性水平为 $p < 0.05$ 。

4 实验结果

4.1 行为学数据结果

为保证被试在实验过程中认真加工实验语料，本实验设置了句子类型判断任务，实验程序软件记录了听觉语义判断实验的反应时(reaction time)和正确率(accuracy)，见表 2：

听觉条件	句子类型	正确率		反应时	
		均值	标准差	均值	标准差
安静	合理	0.92344	0.049555	408.49375	133.430323
	违反	0.93594	0.049974	415.55156	119.330137
白噪声	合理	0.82344	0.09936	481.71875	140.190289
	违反	0.72656	0.138885	616.71094	194.423566
汉语噪声	合理	0.77656	0.092407	555.79375	165.713697
	违反	0.56094	0.207559	632.6625	205.893059
英语噪声	合理	0.68125	0.177365	563.34531	166.978417
	违反	0.45937	0.166051	646.3	234.94188

表 2.行为学数据

对正确率进行重复测量分析发现，听觉条件的主效应显著($F(2, 44)=127.430, P=0.000$)。句子类型的主效应显著($F(1, 22)=6.659, P=0.017$)，语义合理句的正确率显著高于语义违反句。听觉条件和句子类型存在交互作用($F(2, 44)=6.340, P=0.004$)。

对反应时进行重复测量分析发现，听觉条件的主效应显著($F(2, 44)=33.818, P=0.000$)。句子类型的主效应显著($F(1, 22)=15.193, P=0.001$)。句子类型和听觉条件存在交互作用($F(2, 44)=9.935, P=0.000$)，对此进行简单效应分析。

4.2 ERP 数据结果

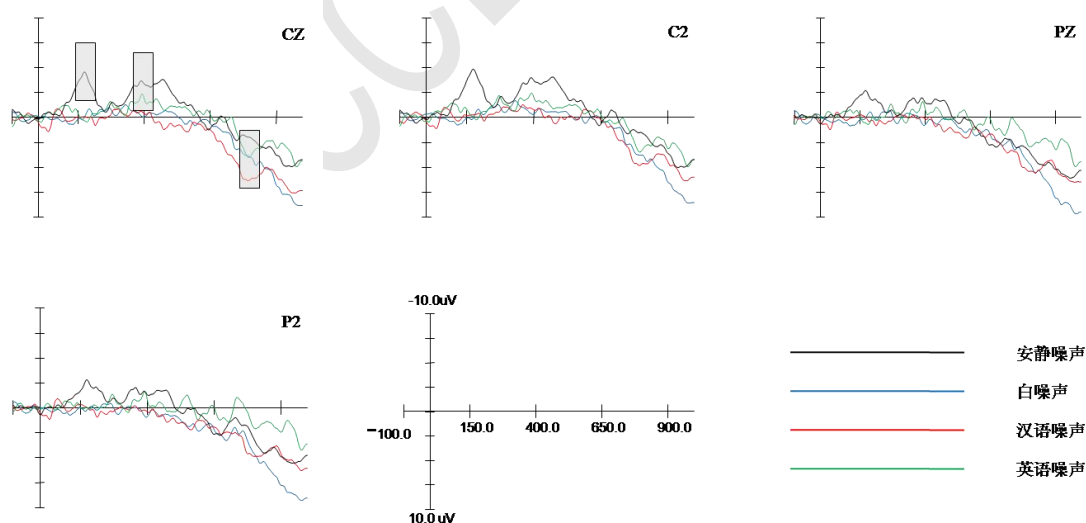


图 2.不同听觉条件下语义合理句的波形图

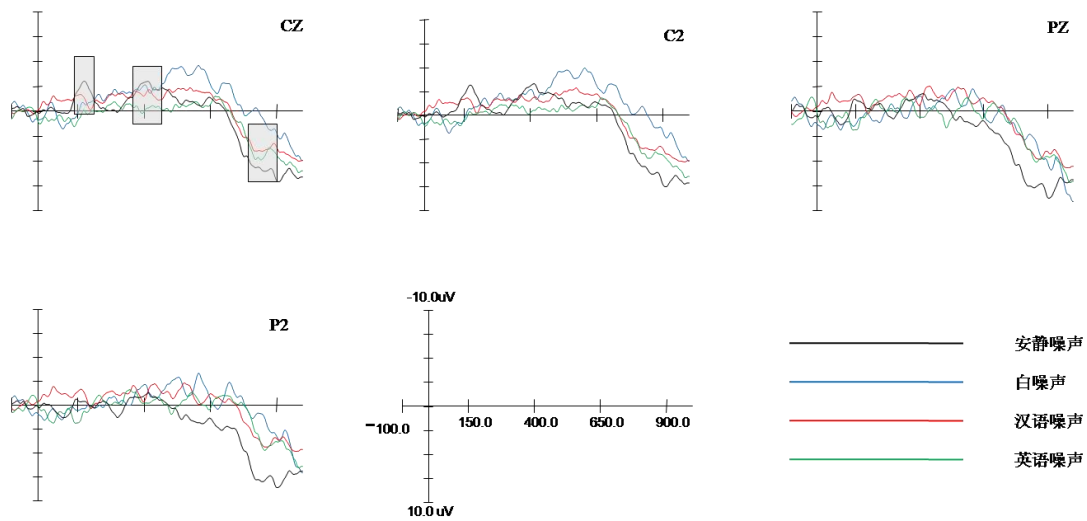


图 3.不同听觉条件下语义违反句的波形图

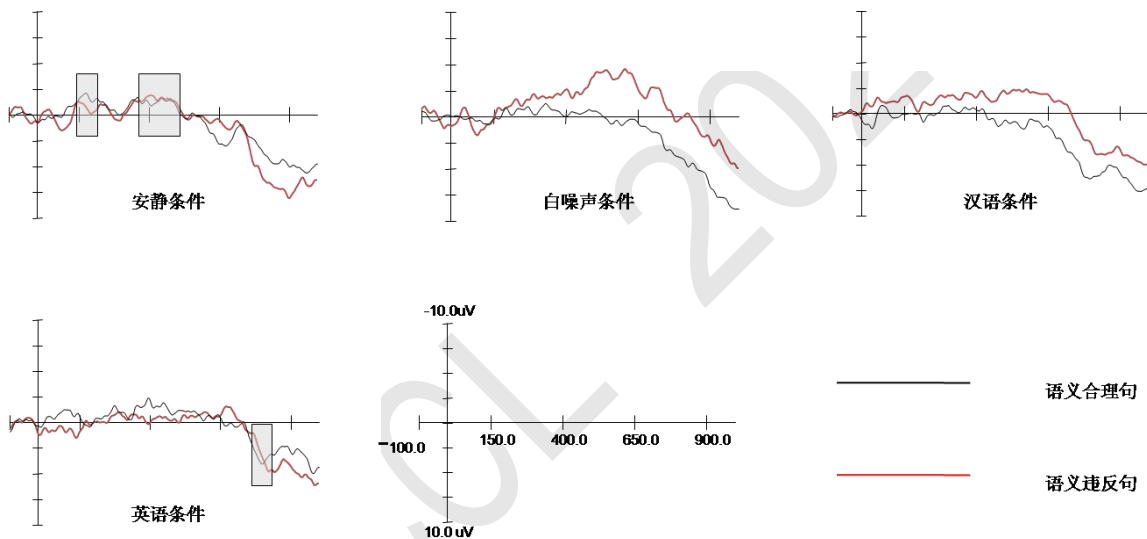


图 4.不同听觉条件下语义合理与语义不合理句的对比波形图

1) N100

就 N100 成分的波幅而言，脑区主效应显著($F(4, 60)= 8.721, P=0.000$)，左后脑区、右后脑区与中部脑区的波幅均大于左前脑区；右后脑区与中部脑区的波幅显著大于右前脑区。

噪声×脑区有交互作用($F(8, 120)=7.097, P=0.000$)，经简单效应分析，对于左前脑区，安静条件的波幅既显著大于汉语噪声条件又显著大于英语噪声条件，对于右前脑区，安静条件的波幅既显著大于汉语噪声条件又显著大于英语噪声条件；对于安静条件，左前脑区显著大于左后脑区、右后脑区与中部脑区的波幅；右前脑区显著大于左后脑区、右后脑区和中部脑区；中部脑区大于右后。

噪声×句子类型有交互作用($F(2, 30)=3.789, P=0.034$)，经简单效应分析，对于语义合理句，安静条件的波幅显著大于汉语噪声条件；对于语义不合理句，汉语噪声条件的波幅显著大于英

语噪声条件。

句子类型×脑区有交互作用($F(4, 60)=2.596, P=0.045$)，经简单效应分析，对于语义不合理句，左前脑区显著大于左后脑区、右后脑区和中部脑区，右前脑区显著大于右后脑区，中部脑区显著大于右后脑区。

2) N400

就 N400 成分的波幅而言，脑区主效应显著($F(4, 60)=4.342, P=0.004$)，左前脑区的波幅显著大于左后脑区，右前脑区的波幅显著大于右后脑区。

听觉条件×脑区有交互作用($F(8, 120)=4.424, P=0.000$)，经简单效应分析，对于安静条件，左前脑区的波幅显著大于左后脑区、右后脑区与中部脑区；右前脑区的波幅显著大于右后脑区。

听觉条件×句子类型有交互作用($F(2, 30)=7.009, P=0.003$)，经简单效应分析，对于汉语噪声条件，语义违反句的波幅显著大于语义合理句，对于英语噪声条件，语义合理句的波幅大于语义违反句；对于语义合理句，英语噪声条件的波幅显著大于汉语噪声。

3) LPC

就 LPC 成分的波幅而言，听觉条件×句子类型有交互作用($F(2, 30)=6.093, P=0.006$)，经简单效应分析，对于汉语噪声条件，语义合理句的波幅显著大于语义违反句，对于英语噪声条件，语义违反句的波幅显著大于语义合理句；对于语义合理句，汉语噪声条件的波幅显著大于英语噪声。

噪声×句子类型×脑区具有交互作用($F(8, 120)=4.839, P=0.000$)，经简单效应分析，对于汉语噪声条件，在左前脑区、左后脑区、右前脑区、右后脑区与中部脑区，语义合理句的波幅显著大于语义违反句；对于英语噪声条件，在左前脑区、左后脑区、右前脑区、右后脑区与中部脑区，语义违反句的波幅显著高于语义合理句。对于语义合理句，在左前脑区、左后脑区、右前脑区右后脑区和中部脑区，汉语噪声条件的波幅显著大于英语噪声条件。

5 分析与讨论

5.1 实验结果分析

根据不同听觉条件下语义合理句与语义违反句的对比波形图，本研究得出了一个有趣的实验结果，即在安静、白噪声、汉语噪声条件下，语义违反句诱发的 N400 波幅大于语义合理句，但英语噪声条件则相反，语义合理句诱发的 N400 波幅大于语义违反句。SPSS 重复测量结果验证了 ERPs 波形图上的波幅趋势，对于 N400 平均波幅值，听觉条件与句子类型具有交互作用，汉语噪声条件下语义违反句的波幅显著大于语义合理句，英语噪声条件下语义合理句的波幅大于语义违反句。此外，行为数据也支持了 ERP 结果，汉语噪声条件的语义合理性判断的正确率显著高于英语噪声条件下的正确率。

这一结果在一定程度上表明，在特定条件下，对于听力正常的成年汉语母语者而言，英语噪声对其加工目标语音的语义信息的干扰大于汉语噪声。一般研究认为，N400 与语义加工加工，且加工难度越大，需要的加工资源越多，N400 波幅越显著，因此，语义违反句诱发的 N400 波幅显著大于语义合理句。但在本实验中，英语噪声条件下，语义违反句诱发的 N400 波幅显著缩小，甚至小于语义合理句诱发的 N400 成分。这表明在英语噪声条件下，语义违反句的加工受到了更大的阻碍，这一阻碍在很大程度上来自英语噪声对目标句在注意及认知资源上的分割或占用，所剩资源难以维持加工难度较大的语义违反句。

神经网络中用于某一特定任务的心理认知资源的数量是有限的，如果那些包含语义信息的言语噪声进入听者的中枢神经系统，必然会分割原本用于目标语音的资源。语义合理句的加工难度相对较小，消耗的认知资源相对较少，剩余的认知资源足够用于对言语噪声的加工。相比之下，语义违反句的加工难度较大，需要的认知资源相对更多，认知资源难以同时负担目标语音与言语噪声的加工，在这种情况下，言语噪声便极可能夺取目标语音加工的认知资源。

由于难以保证在语义合理句与语义违反句条件下诱发的 N400 成分有无言语噪声的参与,本研究将各听觉条件下的语义合理句与语义违反句条件下诱发的 N400 波幅相减,得出了各听觉条件下语义合理句与语义违反句的 N400 波幅差值,然后对各听觉条件下语义合理句与语义违反句的 N400 波幅差值进行了重复测量分析。结果显示白噪声条件下的 N400 波幅差值与英语噪声条件下的波幅差值具有显著差异,与汉语噪声条件的波幅差值无差异。前人研究已经证明,白噪声产生能量掩蔽效应,言语噪声产生能量掩蔽与信息掩蔽的叠加效应,且能量掩蔽效应对语义加工的干扰作用小于信息掩蔽效应。再结合本实验结果,在很大程度上可以得出,英语噪声较汉语噪声对目标汉语句语义加工的信息掩蔽效应更强。

后期时间窗内诱发的 LPC 的波形趋势与 N400 的波形趋势相符。一般研究认为, LPC 正成分的形成与持续注意、记忆的形成、刺激的编码等过程相关,其波幅的大小取决于心理认知资源的分配情况(Azizian et al., 2007), LPC 正成分和分析评价过程的重要指标之一,反映了加工过程中认知资源的持续投入(Fields et al., 2016)。评价分析是个体根据自身需要和环境变化做出的资源分配的控制性加工过程,在这一过程中,那些得到充分分析的刺激将会诱发较大的 LPC 波幅。

根据本实验的 ERPs 波形图,在英语噪声条件下语义违反句诱发的 LPC 波幅显著大于语义合理句,汉语噪声条件则相反。SPSS 重复测量分析结果显示,白噪声与汉语噪声条件下语义合理句的波幅均显著大于语义违反句,而在英语噪声条件下,语义违反句的波幅显著大于语义合理句。这一结果很可能表明,在英语噪声条件下,语义违反句的语义加工受阻严重,为弥补早期语义加工的不足,个听者在后期仍然持续进行认知加工以达到补偿目的,因此在晚期时间窗诱发出更显著的 LPC 成分。

综上,根据本实验结果,得出如下假设,即在噪声掩蔽条件下,当汉语母语者对汉语句进行难度较大的听觉认知加工时,英语噪声的掩蔽效应强于汉语噪声,在这一过程中,有限的心理认知资源难以同时维持目标语音与英语噪声的语义加工。以汉语句为目标刺激,较之英语噪声,汉语噪声必然在声学特征、语音形式等方面与目标语音更为相似,很显然,本实验的结果并不能用相似性假说解释,下面本文将从掩蔽噪声的可懂度、不确定性角度进行分析。

5.1 信息掩蔽的机制探讨

实验后调查问卷提供了一个有趣且有意义的发现。部分被试指出对于他们而言,汉语噪声的可懂度极高,因此他们不仅能够轻松理解汉语噪声的内容,而且还能够较容易地从听觉上转移对汉语噪声的注意。但对于本实验使用的英语噪声的语义内容,他们只能理解部分单词或句子,所以当他们在一连串较难懂的英语语流中听到可以理解的英语单词或短语时,他们会变得“兴奋”,不受控制地想去加工更多的语义信息。

调查问卷的这一结果为我们提供了若干思路。首先,听者的注意力会受到噪声的干扰,尤其是包含可理解的语义信息的言语噪声,从而造成注意的转移与注意资源的竞争。一般认为, N100 成分与注意相关,注意力集中时,波幅值增加(Hillyard et al., 1998)。从本实验 ERPs 波形图可以看出,安静条件下诱发了极为显著的 N100 成分,但在噪声条件下, N100 波幅显著缩小。SPSS 重复测量的结果显示安静条件下 N100 的波幅显著大于噪声条件,尤其大于言语噪声条件,而言语噪声条件下 N100 的潜伏期大于安静条件和白噪声条件。这表明包含语义信息的言语噪声对听者注意的干扰显著大于无语义信息的纯音噪声,掩蔽噪声的语义内容在信息掩蔽中占重要作用。

其次,那些“成功”转移听者注意的语义信息被传输至听觉中枢进行加工,它们占用了原本用于目标语音加工的认知资源, ERPs 上表现为言语噪声条件下语义加工诱发的 N400 成分的波幅显著缩小。由于认知资源竞争导致语义加工受阻,为弥补前期语义加工的不足,听者在后期加工过程中对资源分配进行控制性加工,并对目标刺激持续投入心理认知资源以达到补偿的

目的。因此在晚期时间窗，言语噪声条件下的语义加工诱发出了更显著的 LPC 成分。

那么，一个关键问题是，在什么条件下或者哪些语义信息能够导致听者的注意转移呢？能够被听者理解当然是一个重要前提，此外，本研究推测包含意外性、独特性、新颖性的因素是另一个重要特征，这些因素能够更有效地转移听者的注意力，消耗注意资源，进而帮助部分可懂信息进入中枢神经系统完成高级加工。

Moray(1959)发现，当听者在嘈杂的听觉环境中听到了高度相关或熟悉的刺激(如自己的名字)，即使这一刺激出现在一个未被注意的听觉流中，也依然能够吸引听者的注意。Conway(2001)指出这个效应越强，听者越难选择性听到本应被注意的听觉流。因此，前人研究指出一些意外性、独特性因素的出现会增强外源性注意(Kaya et al., 2014)。本研究选择的英语噪声为包含四名说话人的有意义的英语语篇，虽未测评其可懂度的具体数值范围，但就被试英语水平⁵和语料难度而言，该英语掩蔽中既包含可理解的语义信息，又包含不可理解的语义信息，且可懂信息的比例小于不可懂信息。我们似乎可以假设，英语噪声中那些出现频率较小的可懂信息，便属于掩蔽噪声中的不确定因素，而在这一情况下，可懂度与不确定性之间似乎建立起了联系。

而掩蔽噪声中那些“出乎意料”但又可以被听者加工的信息，其出现与听者的预期相冲突，造成了注意的脱离与转移。偏差刺激干扰的相关研究认为偏差刺激的影响不仅因为偏差刺激的小概率或新颖性，还可能是认知预期冲突导致的受损行为。李毕琴等人(2014)发现无关声音刺激的意外变化会干扰任务表现，这背后的神经机制是无关声音刺激序列引发的预测机制，预测机制导致注意网络系统开启警觉功能。

这一假说似乎可以为本实验的结果提供一定的解释。在噪声掩蔽条件下，神经系统中预测机制首先会对即将执行的任务或加工做出预判，预判依据为目标任务的需要和掩蔽噪声的性质。目标任务的需要一般为加工目标声音、屏蔽掩蔽噪声，掩蔽噪声的性质则由掩蔽声音中的标准刺激或主要特征主导。预测机制依据这两点去激活注意网络，注意网络进入准备或警觉状态。当掩蔽噪声中的标准刺激或主要特征的出现时，注意定向速度加快，听者能较快且容易地将注意集中于目标声音，注意脱离和转移过程消失且损耗小；偏差刺激的出现则使得注意定向受到干扰，注意脱离和转移的损耗重新出现。

由不同类型噪声引发的注意脱离和注意转移的损耗不同。当言语噪声内容的可懂度极高或极低时，通过预测机制，听者高度专注于目标刺激，噪声掩蔽效应相对较小，但当言语噪声中既包含可懂信息又包含不可懂信息，且可懂信息的概率小于或等于不可懂信息时，可懂信息作为一种偏差刺激，其出现使得注意脱离和转移的损耗重新出现，听者同时也要增大认知资源的投入，信息掩蔽效应也随之增强。

这一结果在一定程度上表明，掩蔽声音的可懂度与信息掩蔽强度并不是简单的正比关系。以听者熟悉的语言和不熟悉的语言作为掩蔽，掩蔽效果相对较小，而那些既包含可懂信息又包含不可懂信息的言语噪声，其掩蔽强度相对更大。由此，本研究猜测掩蔽声音的可懂度与信息掩蔽强度之间极有可能是倒 U 型的曲线关系，即随着掩蔽语音可懂度的增大，掩蔽程度逐渐增强，当掩蔽语音中的可懂信息量与不可懂信息量相当时，掩蔽程度可能会达到峰顶，而后逐渐递减。

6 结论

本研究使用 ERP 技术，以听觉汉语句子作为目标刺激，设置安静、白噪声、汉语噪声和英语噪声四种听觉条件，考察了言语噪声条件下听觉句子语义加工的神经机制，较为深入地探究了信息掩蔽效应的性质及属性。

本实验结果显示，在行为数据上，英语噪声条件下听者执行语义判断任务的正确率显著低于汉语噪声。ERP 数据上，英语噪声条件下语义违反句诱发的 N400 波幅显著小于语义合理句，

⁵ 根据《Oxford University Press and University of Cambridge Local Examinations Syndicate》的测评分数，所有被试的综合英语水平均属于 intermediate level，平均分约为 38.267。

汉语噪声条件下则相反。白噪声条件下语义合理句与语义违反句 N400 波幅差值与英语噪声条件下的 N400 波幅差值存在显著差异,与汉语噪声条件相比无差异。英语噪声条件下语义违反句诱发的 LPC 波幅显著大于语义合理句,汉语噪声条件则相反。这在一定程度上表明,对于听力正常的汉语母语者,英语噪声对其进行汉语句语义加工的掩蔽效果大于汉语噪声。

本研究推测,在语音掩蔽条件下,对于难度较大的语义加工,目标语音与掩蔽噪声在知觉层面的相似性并非主要影响因素,而掩蔽噪声语义内容上的可懂度发挥着更关键的作用。听者极其熟悉或完全陌生的言语噪声对语义加工的掩蔽干扰较小。而当掩蔽噪声为听者接触过的语言但不是母语或主要语言,在一定条件下其掩蔽效果反而更大。言语噪声中出现频率较小但能够被听者理解的语义内容,其出现与听者预期相冲突,更容易引发听者的注意转移,这些语义信息被传输至听觉中枢神经,占用了原本用于目标刺激的认知资源,从而增强了信息掩蔽的效果。

参考文献

- Miller, George A. The masking of speech[J]. *Psychological Bulletin*, 1947, 44(2):105.
- Kidd Gerald and Mason Christine R and Arbogast Tanya L. Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns[J]. *The Journal of the Acoustical Society of America*, 2002, 111(3): 1367-76.
- 杨志刚. 听觉信息掩蔽的亚成分: 知觉信息掩蔽和认知信息掩蔽[D]. 北京:北京大学,2008.
- 杨志刚, 张亭亭, 宋耀武,等. 听觉信息掩蔽的亚成分:基于行为和脑成像研究的证明:[J]. *心理学进展*, 2014.
- Brungart D S et al. Informational and energetic masking effects in the perception of multiple simultaneous talkers[J]. *The Journal of the Acoustical Society of America*, 2001, 110(5 Pt 1) : 2527-38.
- Brouwer S , Engen K V , Calandruccio L , et al. Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content[J]. *The Journal of the Acoustical Society of America*, 2012, 131(2):1449-1464.
- Engen K V , Bradlow A R . Sentence recognition in native- and foreign-language multi-talker background noise[J]. *Journal of the Acoustical Society of America*, 2007, 121(1):519.
- Calandruccio L , Brouwer S , Engen K V , et al. Masking release due to linguistic and phonetic dissimilarity between the target and masker speech[J]. *American Journal of Audiology*, 2013, 22(1):157-164.
- 许博渐. 言语可懂度与知觉相似性对于信息掩蔽的影响[D].河北大学,2015.
- Neff D L, Green D M. Masking produced by spectral uncertainty with multicomponent maskers[J]. *Perception & Psychophysics*, 1987; 41(5):409-415.
- Kidd G , Mason C R , Rohtla T L , et al. Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns[J]. *Journal of the Acoustical Society of America*, 1998, 104(1):422-431.
- Allen Kachina et al. Masker location uncertainty reveals evidence for suppression of maskers in two-talker contexts[J]. *The Journal of the Acoustical Society of America*, 2011, 130(4) : 2043-53.
- Oksama L, Jie L, J Hyönä. Identity information of multiple moving objects is extracted in a serial manner during multiple identity tracking: An eye-tracking study[J]. *Journal of Vision*, 2016; 16(12):1263.
- Nakai T, Kato C, Matsuo K. An fMRI study to investigate auditory attention: a model of the cocktail party phenomenon[J]. *Magnetic Resonance in Medical Sciences An Official Journal of Japan*

- Society of Magnetic Resonance in Medicine, 2005; 4(2):75.
- Gutschalk A, Michey C, Oxenham A J. Neural Correlates of Auditory Perceptual Awareness under Informational Masking[J]. PLoS Biology, 2008; 6(6).
- Scott S K, Mcgettigan C. The neural processing of masked speech[J]. Hearing Research, 2013; 303(9):58-66.
- Bohan D, Chen C S, Long Y H et al. Neural mechanisms for selectively tuning into the target speaker in a naturalistic noisy situation[J]. Nature Communications,2018; 9(1):R1024-R1027.
- 吕思慧, 张璐, 钟舒明等. 基于虚拟现实注意力训练对抑郁发作患者认知功能干预的研究[J]. 中华精神科杂志, 2020;53(5):384-391.
- 殷恒婵.青少年注意力测验与评价指标的研究[J]. 中国体育科技,2003;39(3):51-53.
- 张华, 曹克利.汉语最低听觉功能测试的设计及初步应用[J]. 中华耳鼻咽喉头颈外科杂志, 1990; 25(002): 79-82.
- Azizian A, Polich J. Evidence for attentional gradient in the serial position memory curve from event-related potentials[J]. Cognitive Neuroence Journal of, 2007,19(12):2071-2081.
- Fields E C, Kuperberg G R. Supplementary Material.2016.
- Hillyard S A , Vogel E K , Luck S J . Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 1998,353(1373), 1257-1270.
- Moray, N. Attention in dichotic listening: Affective cues and the influence of instructions.Quarterly Journal of Experimental Psychology, 1959,11, 56–60.
- Conway A , Cowan N , Bunting M F . The cocktail party phenomenon revisited: the importance of working memory capacity[J]. Psychonomic Bulletin & Review, 2001, 8(2):331-335.
- Merve K E , Mounya E . Investigating bottom-up auditory attention[J]. Frontiers in Human Neuroscience, 2014, 8.
- 李毕琴, 袁子璇, 胡竹菁. 偏差声音刺激对注意网络系统的影响[C]. 中国心理学会.第十七届全国心理学学术会议论文摘要集. 中国心理学会, 2014;3.