

# 《心理学报》审稿意见与作者回应

题目：言语的情绪韵律和情绪语义对听觉去信息掩蔽的作用

作者：郑茜 张亭亭 李量 范宁 杨志刚

## 第一轮

### 审稿专家 1 意见：

本文探究了情绪韵律和情绪语义对听觉掩蔽的不同影响及影响的机制，选题有价值，实验设计研究，数据分析详实。

**意见 1：**引言中详细写明了掩蔽有“能量掩蔽”和“信息掩蔽”，但直接定性“言语的韵律和语义中所包含的情绪信息的去掩蔽作用很可能是作用于信息掩蔽而非能量掩蔽”，缺乏可靠证据。实际上，听觉掩蔽范式是很难排除“能量掩蔽”的，如引言所述，Dupuis 和 Pichora-Fuller 证明了不同的情绪声音具有显著的声学特征差异，而能量掩蔽主要是目标与掩蔽声音在时间和频率上的重叠所导致的，所以掩蔽音与目标音之间声学特征的相似性必然会对能量掩蔽的效果造成影响。

**回应：**感谢审稿专家的意见！

我们同意审稿专家的意见，本研究中情绪韵律和语义信息（特别是情绪韵律信息）的去掩蔽作用很难完全排除能量掩蔽成分。但由于能量掩蔽发生在听觉外周系统，几乎不受认知、注意等高级加工的调节，具有很强的不可逆性(Wu et al., 2007; Yang et al., 2007)。以往研究表明，类似于知觉空间分离(Freyman et al., 1999, 2001)、内容或者声音启动(Freyman et al., 2003; Yang et al., 2007)等线索对于能量掩蔽的释放都有一定的效应，但一是程度相对于信息掩蔽而言更小（约为 1dB 左右），二是比较稳定。所以我们认为情绪韵律的去掩蔽作用主要还是作用于信息掩蔽而非能量掩蔽。原则上，我们认为通过分析识别曲线下不同部分的面积可以给出每种条件下能量掩蔽和信息掩蔽量的一个估计值，但我们暂时未掌握这种分析所需要的数学工具，所以在本研究中并未尝试这种方法。

**意见 2：**需要就掩蔽刺激进行更为详细地交代。实验使用“两个说话人音频叠加的混合声音”的正序语句和“时间逆转语句”为掩蔽刺激，“两个说话人音频叠加的混合声音”是否还能够听出语言文字？是否会成为不可识别内容的噪音？如果不具备可识别的信息内容，则不会引起信息掩蔽。而经过倒放的“时间逆转语句”，其必然是不可识别的非言语的噪音，引起的只会是能量掩蔽，而非信息掩蔽。

**回应：**感谢审稿专家的意见！

“两个说话人音频叠加的混合声音”的正序语句，仍能听出语言文字信息，但已经比较困难了。经过倒放的“时间逆转语句”，长期频谱信息与正序语句相似，因此二者产生的能量掩蔽几乎相等。虽然经过倒放的“时间逆转语句”与正序语句具有相反的时间模式，不具有任何可懂度的言语，不包含认知因素，但其保留了人类嗓音的基本特征，具有谐波和共振峰结构，会占用听者的注意资源，从而会导致知觉信息掩蔽。此前也有研究表明无任何可懂内容的类言语 (speech-like) 合成的声音可以产生信息掩蔽 (如Chen et al., 2012 *Informational masking of speech produced by speech-like sounds without linguistic content*)。并且，本文的一个重要观点即是，信息掩蔽中除了因掩蔽声内容所导致的认知上的干扰之外，掩蔽声音流本身作为一个听觉事件本身就会吸引听者的加工资源 (主要是注意资源)，因而，即便是完全不具备可懂度的声音，两个以上的声音流叠加在一起的时候也会导致被试在加工上的困难，具体体现在，听者需要在若干声音流中间确定并追踪目标声音流，而这种追踪过程是需要耗费知觉加工资源的，因而，也会造成额外的掩蔽。并且，与纯粹的能量掩蔽不同，这种因声音流数目的增加而不是因语义内容上的干扰所导致的掩蔽也能在很大程度上被知觉空间分离所释放。因而，我们认为它是一种与因掩蔽声语义内容对中枢加工资源的竞争所导致的 (认知上的) 信息掩蔽不同的信息掩蔽类型。

**意见 3:** 引言最后提到“主观空间分离范式利用优先效应能够使听者主观感知到目标声和掩蔽声来自不同的位置(虽然实际上目标声和掩蔽声播放位置相同)，其主要释放信息掩蔽，对能量掩蔽影响不大(Freyman et al., 1999, 2001)。”作者此处所引用的文献为英语环境下的去掩蔽探究。根据吴艳红(2005)的研究，在汉语环境下，主观空间分离可以引起去能量掩蔽，也可以引起去信息掩蔽，两种去掩蔽作用均会发生，此处论断有误。 [1]吴艳红,李文瑞,陈婧,王纯,屈宏伟,吴玺宏,李量.主观空间分离下的汉语信息掩蔽效应[J].声学学报, 2005(05): 462-467. DOI: 10.15949/j.cnki.0371-0025.2005.05.013.

**回应:** 感谢审稿专家的意见!

我们在论文中的某些表述可能让审稿专家以为我们认为知觉空间分离对能掩蔽掩蔽没有什么影响，这可能是一个误会，实际上，知觉空间分离确实对能量掩蔽是有一定效果的，不仅吴艳红老师 2005 年的汉语研究得到了这一结果，Freyman et al., 1999, 2001 的研究中也有发现知觉空间分离对于能量掩蔽有一定的影响，只不过影响程度很小 (约 1dB 左右的释放量)，且在不同研究中都比较恒定。吴艳红 (2005) 的研究中没有直接计算识别阈限差值，但是在 Wu et al., (2011) 一篇直接比较知觉空间分离对汉语普通话和英文掩蔽释放作用的研究中，发现在能量掩蔽条件下空间分离对汉语和英语的掩蔽释放效果是非常接近的 (分别为 1.3 和 1.2dB 左右)。因为此前确实有很多研究表明知觉空间分离对能量掩蔽释放效果的这种恒定性，我们在本研究中也没有包含能量掩蔽条件，而是将重点直接放在信息掩蔽上。

但原则上总体释放量中应该包含一些能量掩蔽成分。但因为知觉空间分离对能量掩蔽的相对恒定性，我们假定不同情绪线索条件下因知觉空间分离所导致的能量掩蔽释放量是相对恒定的，因而不同情绪线索所导致的掩蔽释放量仍在很大程度上主要反映其对信息掩蔽的作用。

**意见 4:** 导言主要在论述情绪信息的去掩蔽作用，对操控的其他自变量，特别是信噪比，没有加以论述；导言的呈现逻辑不够清晰，有的段落只介绍一个研究，而且最终没有关联到研究问题上去；没有阐述为什么要韵律情绪和语义情绪，或是说明找出二者差异对于我们的生活有何帮助；没有写明研究假设。

**回应:** 感谢审稿专家的意见！

第一，针对专家所提出的关于信噪比变量未作深入讨论的意见，我们是出于文章整体结构的考虑。因为本文的最核心的目标是要考察情绪韵律和语义线索的去掩蔽作用，我们主要通过计算有知觉空间分离和没有知觉空间分离下的识别率差值来作为被释放的信息掩蔽的指标（严格说来其中可能包含有少量能量掩蔽，见上一条回复意见）来分析，设置不同的信噪比水平可以帮助我们获取在不同掩蔽强度下的掩蔽释放情况，更主要地，通过对不同信噪比条件下识别率的参数拟合，我们可以得到一个比较稳定的识别阈限作为考察不同线索掩蔽释放作用的主要指标。而关于信噪比对于能量掩蔽和信息掩蔽成分的影响，此前已经有很多研究证明能量掩蔽对信噪比更敏感，信息掩蔽不太敏感。为更好地聚焦于核心问题，论文写作过程中我们就没有对信噪比变量展开深入论述。

第二，关于导言的呈现逻辑不够清晰，有的段落只介绍一个研究，而且最终没有关联到研究问题上去这一问题。

感谢审稿专家的宝贵意见，初稿中的引言有些地方确实写得不够好，问题提出链条不够清晰，修改稿中我们对此做了针对性的修改，调整了部分段落的内容，增加引用了一些新的文献，见引言中标示蓝色字体的部分。

第三，关于没有阐述为什么要研究韵律情绪和语义情绪，或是说明找出二者差异对于我们的生活有何帮助问题。

我们确实未在引中阐述研究韵律情绪和语义情绪的理论意义和实际意义。补充内容如下：

“从实际意义来看，在现实的交流场景中，不论是情绪韵律还是情绪语义，都是在交流中经常会出现的泛语言元素，了解这元素在交流中的作用和具体机制，对于提升日常交流的效能，以至于开发更为有效的人工耳蜗算法都有重要的启示。”，见论文引言部分后面蓝色字体部分。

**意见 5:** 本文使用 G\*power 计算三因素被试内设计的被试量，据我所知，G\*power 不支持多于 1 个被试内或被试间变量的计算。在 Stack Exchange 上有人反映，其向 G\*power 的团队邮件咨询过，得到回复如下：“They informed me that the current version of G\*Power (3.1.9.2)

cannot conveniently do power analyses for repeated measures designs with more than one within-subject or between-subject factor. It is possible using the "Generic F test" option, but this is considerably more complicated.“ 来源: Repeated measures within factors settings for G\*Power power calculation 页面底部最后一个回答。

<https://stats.stackexchange.com/questions/59235/repeated-measures-within-factors-settings-for-gp-ower-power-calculation>

回应: 感谢审稿专家的意见! 我们在利用 G\*Power 计算样本量需求的过程中错误地使用了 G\*Power 的设定条件。在修改稿中我们删除了这部分内容。

意见 6: 情绪韵律实验中的目标语句为“由同一个年轻女性(C)以不同情绪韵律(中性、开心)读出并进行录音”, 不同情绪韵律的语音声学特征不一致, 会对能量掩蔽的效果造成不同影响。引言中也有提到 Lu(2018)在研究情绪线索对去掩蔽的影响时, 为了避免声学因素的影响, 采用了目标语句和情绪音的匹配, 本文既然提到了这个影响因素, 为何在实验中没有考虑?

回应: 感谢审稿专家的意见!

这确实是一个很重要而微妙的问题。本研究之所以采用情绪韵律和情绪语义线索, 是因为在现实的语言交流中, 发声韵律和语义内容是传递情绪信息的两种最为重要的方式, 其中情绪韵律是一种泛语言学线索, 而情绪语义是一种语言学线索。以往关于这两种线索在言语加工中的作用的研究很多 (如 Pell, M. D., et al. (2011). "Emotional speech processing: Disentangling the effects of prosody and semantic cues." *Cognition & Emotion* 25(5): 834-853; Schwartz, R., & Pell, M. D. (2012). Emotional speech processing at the intersection of prosody and semantics. *PloS one*, 7(10), e47279.等), 因此, 考察这两种线索的去掩蔽机制具有重要的生态学意义。但由此所带来的问题是不同情绪韵律的声音在声学特征上的差异确实可能会对能量掩蔽带来一定的影响。但鉴于本研究中主要采用的分析指标是不同因素所导致的知觉空间分离带来的去掩蔽量, 而之前的研究表明这种技术对于能量掩蔽的作用相对较小且较恒定, 因而, 我们认为就本研究的结果而言, 情绪韵律所导致去掩蔽作用主要反映的还是信息掩蔽上的差异。但审稿专家所提出的这方面的问题确实是一个值得考虑的问题。在今后的研究中我们可以考虑借鉴 Lu et al., (2018)的条件学习范式, 考察在更好地匹配声学因素条件下言语的情绪性的去掩蔽效果。

意见 7: 在使用心理测量函数拟合实验数据的部分中, 该处明显参考 Lu(2018)的文章, 为何没有做引用? 在 Lu(2018)的文章中, 等式左边只有  $y$ , 此时  $y$  作为因变量, 作者将等式右边改为  $P(y)$ , 此时  $y$  是自变量,  $P$  才是因变量, 而等式右边并没有作为自变量的  $y$ , 是一个错误且意义不明的计算式, 请作者进行修正。在对计算式中参数进行解释的部分, 说“ $\sigma$  是函数的斜率”。根据斜率的定义, 斜率, 是表示一条直线 (或曲线的切线) 关于 (横) 坐标轴

倾斜程度的量。只有直线才有固定的斜率参数，曲线的切线斜率是变化的，没有固定值， $\sigma$ 不能称之为斜率。请作者进行修改。

回应：感谢审稿专家的意见！

我们使用 logistic 心理测量函数对被试识别率的数据进行拟合是参考了以前的一篇研究，使用的软件是 Origin 8.0 软件。经核实，公式的规范形式中等号左侧确实是  $y$  而不是  $p(y)$ ，我们已经在正文中对这一错误进行了更正，修改了对相关参数含义的不准确表述，增加了对 Lu 等人(2018)文献的引用。

意见 8：对于掩蔽释放量的计算，为什么情绪韵律的掩蔽释放量=无主观空间分离条件下中性-开心？有主观空间分离条件下的中性-开心不也是情绪韵律单独导致的结果吗？同理，主观空间分离的掩蔽释放为什么只考虑中性韵律？如果综合考虑掩蔽释放量两种条件下的情绪韵律掩蔽释放量，会对实验结果产生影响吗？

回应：感谢审稿专家的意见！

1. 本研究对于掩蔽释放量的计算方法是参照 Lu 等人(2018)的研究。Lu 等人使用主观空间分离范式考察经过情绪条件化的声音是否能够在信息掩蔽的条件下，提高对目标声的再认，起到去掩蔽的作用。在其研究中将无主观空间分离时，情绪学习前的识别阈限与情绪学习后的识别阈限二者之间的差值作为单独的情绪线索的掩蔽释放量。将在情绪学习前，无主观空间分离时的识别阈限与有主观空间分离时的识别阈限二者之间的差值作为单独的空间分离线索的掩蔽释放量。情绪和主观空间分离两种线索结合时的掩蔽释放量等于情绪学习前且无主观空间分离时的识别阈限减去情绪学习后且有主观空间分离时的识别阈限。

2. 在进行自检过程中，发现在实验 2a 中情绪语义的掩蔽释放量应为 0.11dB，而不是论文中写的-0.11dB。对实验 2 中相关内容进行修改：

“单因素方差分析表明，听觉线索类型效应显著， $F(2, 87) = 20.44, p < 0.001$ 。情绪语义的掩蔽释放量( $M = 0.11, 95\% CI: [-0.67, 0.88]$ )显著低于主观空间分离( $M = 2.84, 95\% CI: [2.05, 3.63], p < 0.001$ )和二者结合的掩蔽释放量( $M = 2.97, 95\% CI: [2.36, 3.58], p < 0.001$ )”。见论文实验 2a 结果部分蓝色字体内容。将图 10 也进行相应的修改。

“对于情绪语义线索，在知觉信息掩蔽下的掩蔽释放量与在知觉和认知信息掩蔽下的掩蔽释放量差异不显著， $t(55) = -1.48, p = 0.144, 95\% CI: [-1.64, 0.25]$ ”，见论文实验 2a 和 2b 结果比较部分。将图 14 也进行相应的修改。

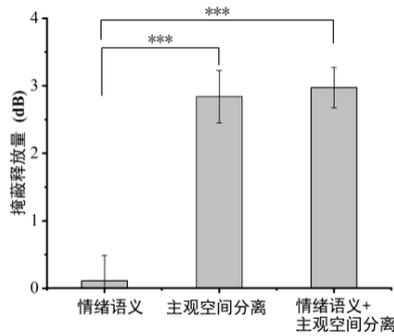


图 10 不同线索的掩蔽释放量

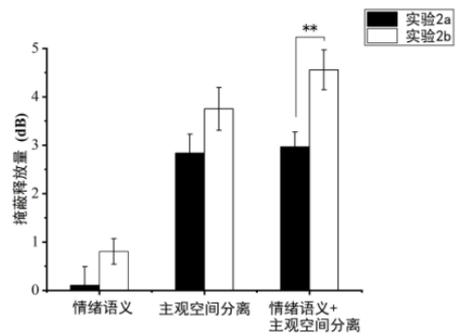


图 14 实验 2a 和实验 2b 去掩蔽量结果对比

3. 改变情绪韵律和情绪语义的掩蔽释放量的计算方法，综合考虑在有无主观空间分离两种条件下的情绪韵律和情绪语义掩蔽释放量。结果发现，情绪韵律线索在知觉信息掩蔽下的掩蔽释放量和在知觉、认知双重信息掩蔽下的掩蔽释放量差异不显著， $t(55) = -1.97, p = 0.054, 95\% \text{ CI}: [-1.72, 0.02]$ 。情绪语义线索在知觉信息掩蔽下的掩蔽释放量与在知觉和认知信息掩蔽下的掩蔽释放量差异不显著， $t(55) = -1.93, p = 0.059, 95\% \text{ CI}: [-1.40, 0.03]$ 。以上结果与本研究的计算方法结果一致。

意见 9: 在实验 2 中，所选的信噪比测量点与实验 1 不同，请说明理由。在实验 2 的设计中说明的信噪比设置为 0, -4, -8, -12dB，后续图表内容均与此不一致，何处正确？

回应: 感谢审稿专家的意见！这是我们初稿中的笔误所致，实验 2 中所选用的信噪比与实验 1 相同，均为 -8dB, -4dB, 0dB, 4dB。正文中对相应内容进行了修改。

意见 10: 在结果分析中，多次提到“由于方差不齐性，报告矫正后的结果”，请说明是根据何种方法进行的矫正。方差齐性是使用方差分析的前提条件，方差不齐建议使用 Welch anova 或 Brown-Forsythe anova 或非参数检验。

回应: 感谢审稿专家的意见！

在进行重复测量方差分析时，因方差不齐性所采用的校正方法是 Greenhouse-Geisser 矫正，我们在正文中作了标明。

另外，由于方差不齐性，对实验 2b 中情绪语义、主观空间分离线索和二者结合时的掩蔽释放量进行的单因素方差分析改为 Welch 方差分析。具体地，将实验 2b 结果中“对其进行单因素方差分析。结果表明，听觉线索类型效应显著(由于方差不齐性，报告矫正后的结果)， $F(2, 78) = 26.67, p < 0.001$ 。”改为“对其进行单因素方差分析(由于方差不齐性，对其进行 Welch 方差分析)。结果表明， $F(2, 48.94) = 35.70, p < 0.001$ ”。见论文……

**意见 11:** 实验一中, 情绪韵律线索对两种播放形式的掩蔽刺激均具有去掩蔽作用, 即便对两种掩蔽刺激的去掩蔽量没有显著差异, 作者下结论说“可以克服掩蔽声音在知觉上造成的干扰, 但不具有去认知信息掩蔽的作用”也是基于阴性结果下推论。

**回应:** 感谢审稿专家的意见。

这里的表述确实有不够严谨的地方。实验一中将正序言语掩蔽和逆序言语掩蔽条件下的情绪韵律的去掩蔽效应对比, 没有发现显著的差异, 原文中基于此得出了情绪韵律“可以克服掩蔽声音在知觉上造成的干扰, 但不具有去认知信息掩蔽的作用”的结论。但实际上两种条件下情绪韵律线索的去掩蔽量之间是存在一定程度的差异的, 没有达到预设的0.05水平, 但可以认为是达到了边缘显著水平 ( $p = 0.086$ )。我们在正文中对此结果的表述和由之得到的结论都做了修改。删除了情绪韵律不具有去认知信息掩蔽作用的表述。

**意见 12:** 知觉掩蔽、认知掩蔽的对立是有争议的, 认知掩蔽中应该含有知觉掩蔽。

**回应:** 感谢审稿专家的意见!

关于信息掩蔽的成分问题是此领域中长期以来都存在争议的问题, 此前人们对信息掩蔽最大的共识就在它像一个行李箱词汇, 所有不能归因于听觉外周阶段时频重叠的掩蔽成分都被归因于信息掩蔽。但近年来已经有越来越多的研究者认为信息掩蔽中包含着发生机制不同的亚成分。但这些亚成分有多少种, 分别基于何种机制仍然存在问题。比如一种长期流行的观点认为信息掩蔽中包含由目标声、干扰声相似性所导致的掩蔽和目标不确定性所导致的掩蔽两种成分。但这种分类的一个关键问题在于相似性和不确定性的界定和它们之间的彼此关系很难在概念上加以澄清 (Lutfi, 1993)。Cooke, Lecumbern 和 Barker(2008)认为信息掩蔽由刺激成分的错误分配, 掩蔽声音所带来的注意竞争, 高认知负荷和可识别的言语所带来的干扰这四种元素组成。Mattys, Brooks 和 Cooke (2009) 也提出言语掩蔽效应主要由知觉和认知负载 (perceptual and cognitive loads) 两种成分导致, 他们具体指出信息掩蔽包括掩蔽声音所带来的注意竞争、可识别的言语所带来的干扰和高级认知负荷成分。但我们认为这里所提到的可识别的言语所带来的干扰和高级认知负荷成分之间的关系同样比较模糊, 难于做严格的概念区分。本文中将信息掩蔽分为知觉和认知两种亚成分的方法源于杨志刚等人(2014)的观点。当然, 在认知心理学中, 通常认为所有的信息加工过程都是“认知”过程。因而, 知觉也是一种认知过程。但这里对知觉掩蔽成分和认知掩蔽成分的界定有更为狭义的限定, 简单地说, 知觉掩蔽成分被认为是由不同的(可知觉为独立的)声音流之间的相互竞争所导致的, 与声音流中所包含的内容无关; 而认知掩蔽被认为是由掩蔽声中的可识别的内容信息对高级认知资源的竞争所导致的, (原则上)与声音流的数目无关。我们之所以采用这种概念框架主要的原因有两点, 一是两种信息掩蔽亚成分的界定简单且机制相对清晰; 二是这种概念框架也与认知心理学中关于注意的两大类模型相契合。在认知心理学中, 注意模型总体上可以分为两大类, 一类是以过滤器模型和衰减模型为代表的通道模型, 认为注意的选择是基

于通道的；第二类是由 Waugh 和 Norman (1965) 所提出的中枢资源理论，认为注意的选择是基于外部的任务难度和可用的中枢资源两种因素。而实际上这两类模型彼此之间并不一定是非此即彼的关系，而是可以彼此融合的。本文中所采纳的信息掩蔽两种亚成分的观点即可看作是两类注意理论的一种整合。知觉掩蔽体现的是通道间的竞争，而认知掩蔽体现的是对可用的中枢资源的竞争。因而，在我们的概念框架中，知觉掩蔽和认知掩蔽是两种不同的掩蔽成分。但确实如审稿专家所言，即便是在我们所采用的理论框架中，这两种掩蔽成分间的关系也依然存在一些不甚清楚的地方。主要体现在两点，一是它们之间是否彼此独立的问题，二是两者在微观发生学上的时间顺序问题。但这两个问题我们目前仍没有明确的认识，需要今后的研究进一步探索。

**意见 13:** “言语的情绪语义能够优先获取听者更多的认知加工资源，具有去认知信息掩蔽的作用，但不具有去知觉信息掩蔽的作用。”也是基于阴性结果做推论。

**回应:** 感谢审稿专家的意见。这里得出的言语的情绪语义的去掩蔽作用主要基于实验 2a 中未发现情绪语义的主效应，而实验 2b 中发现情绪语义存在主效应的结果。在后面对正序和逆序言语掩蔽条件下情绪语义去掩蔽效应的直接对比确实没有得到显著的差异，这可能是因为在通过数据拟合计算识别阈限的过程中损失了部分原始识别率数据所导致的。但我们认为上基于识别率的结果得出上述结论是基本上是站得住脚的。

**意见 14:** 次要问题

**回应:** 感谢审稿专家的意见！这里审稿专家提出了 5 个次要问题：

第一，文中多次提到“完全被试内设计”、“一元方差分析”，建议用“被试内设计”、“单因素方差分析”。

我们同意审稿专家的意见。并按照审稿专家的意见进行相应地修改。

第二，文中多次将“简单效应分析”写为“简单简单效应分析”，请认真校对文中的笔误。

在这里我们是参照舒华老师的《心理学研究方法》一书将三因素交互作用显著之后的进一步分析称为简单简单效应分析。

第三，每个实验条件下只有 11 个 trials，数量是否不足？

实验试次我们主要考虑两个因素，一是整体材料数量，一是参考以前的类似研究。此前有研究（比如 Wu et al., 2011 的跨语言掩蔽研究）中用的是每种条件下 13 个 trials，本研究通过操纵词汇的情绪语义来控制句子的情绪信息。每个句子都是无意义的汉语语句，其主语、谓语和宾语情绪信息需一致。本研究所有的词汇材料均来自中文词汇数据库(CLD) <http://www.chineselexicaldatabase.com/>(Sun et al., 2018)。实验前由 35 名非实验被试对词汇(动词: 571, 名词: 1091)的效价和唤醒度进行 9 点评分。选择效价评分为 4.50-5.50, 唤醒度相对较低的动词(182 个, 评分  $\leq$  6.00)和名词(364 个, 评分  $\leq$  3.51)作为中性词构成目标语句。评

分为 5.80-9.00，唤醒度相对较高的动词(107 个，评分  $\geq 5.26$ )和名词(214 个，评分  $\geq 4.23$ )作为积极情绪词。本研究的实验均有 16 个实验条件。综合考虑，研究确定了每种条件下 11 个 trials，从结果来看，这种设置下我们可以得到每种条件下比较稳定的识别结果。

第四，在数据分析中提到“听觉线索类型效应显著”，请先定义听觉线索类型的含义。

文中我们谈到线索类型时主要指三种：情绪（韵律或语义）、知觉空间分离和这两者的结合。

第五，在实验 2b 的结果分析中，说“主观空间分离和信噪比交互作用显著， $F(3, 78) = 46.52, p < 0.001, \eta^2 = 0.641$ 。简单效应分析发现，不管信噪比的大小，无主观空间分离条件下的识别正确率均显著低于有主观空间分离条件下的识别正确率， $ps < 0.001$ ”。后续语句不是描述交互作用的语句，照如此描述，应是无交互作用。既然有交互作用，虽然在各信噪比下，无主观空间分离的识别正确率均低于有主观空间分离，但也应说明二者在不同信噪比下的差异程度是不一样的。

回应：这里在表述上确实容易引起误解，统计结果确实表明在信噪比的不同水平下，知觉空间分离与不分离之间在识别率上都有显著差异，但是差异量却有比较大的差异。这一点从图中也可以比较明确地看出。我们在正文中补充了效应量 Cohen's d 值，以解释交互效应显著的来源。

.....

#### 审稿专家 2 意见：

该研究通过两个实验，采用主观空间分离范式，通过操纵掩蔽声类型的方式，分别探究言语的情绪韵律和情绪语义去信息掩蔽的机制。说明言语的情绪韵律能够克服掩蔽声音在知觉上造成的干扰，但不具有去认知信息掩蔽的作用。而言语的情绪语义能够优先获取听者更多的认知加工资源，具有去认知信息掩蔽的作用，但不具有去知觉信息掩蔽的作用。总的来说，实验设计较为合理，实验结果较为可靠。但是实验中还存在一些问题，需要进一步说明：

意见 1：作者说明在材料准备过程中，所有目标语句均由同一个年轻女性以不同情绪韵律读出并进行录音。然后由 20 个人对声音韵律的效价和唤醒度进行 5 点评分。不知是否包括实验 2 中的所有中性录音。如果包括的话，不同情绪语义对声音韵律的效价和唤醒度是否有影响？

回应：感谢审稿专家的意见！

实验 2 中的中性录音包括实验 1 中的中性录音。我们考察不同情绪语义是否对声音韵律的效价和唤醒度产生影响。结果发现，在感知声音韵律的效价上，情绪语义效应不显著， $t(212) = 1.53, p = 0.127, 95\% \text{ CI}: [-0.00, 0.03]$ 。在感知声音韵律的唤醒度上，情绪语义效应不显著  $t$

(212) = 0.46,  $p = 0.650$ , 95% CI: [-0.05, 0.08]。我们将这部分内容补充进了正文中，见实验 2a 方法部分蓝色字体内容。

**意见 2:** 有文献表明听者性别对情绪韵律的感知也会有影响的。而在研究 1a 与 2b 当中男女生的比例相差比较大，最好能加以说明。

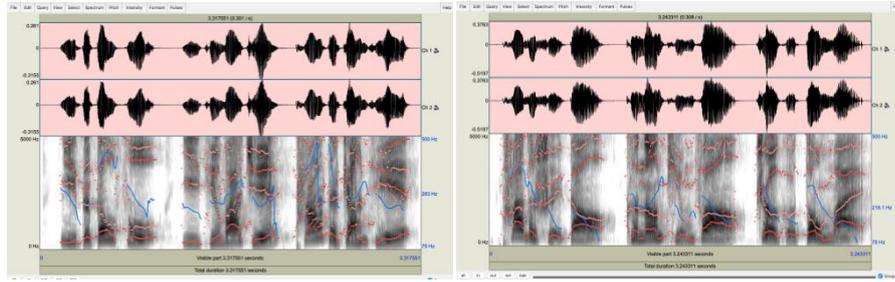
**回应:** 感谢审稿专家的意见！这是一个很有意思的问题，然而非常抱歉的是因为我们所在的实验室曾于 2021 年暑期时整体搬迁，学生用于记录被试信息的笔记本在搬家的过程中不慎遗失。故暂时没有办法补充分析这一因素的作用。但在今后的研究中这是一个很值得考虑的方向。

**意见 3:** 该实验为何选择高兴与中性相比？其他的情绪是否有相应的效果？

**回应:** 感谢审稿专家的意见！这个问题是我们在实验设计阶段曾重点考虑的一个问题。第一，此前曾有研究表明不同的情绪效价对于言语加工的作用不同，积极情绪促进言语加工，而消极情绪会抑制言语加工（Gordon, M. S. and M. Hibberts, 2011）。第二，更为实际的原因是，实验材料准备阶段所录制的消极情绪韵律的句子效果不能令人满意。由于本研究最重要的目标是要考察是否情绪韵律和情绪语义线索具有不同的去掩蔽机制，故从整体实验结构的角度出发，我们暂时舍弃了消极情绪韵律和消极情绪语义的条件。今后的研究中可以更为深入地考察不同情绪类型或不同情绪唤醒度因素在去掩蔽中的作用问题。

**意见 4:** 在“引言”中作者探讨了听者在噪音中对不同情绪下发出单词的再认行为表现不同，可能是由于情绪韵律具有明显的可辨别的声学特征造成的。但是在该实验中，作者却没有比较中性韵律与开心韵律的声学特征，特别是基频特征。若能补充一些声学方面的分析，将更加有说服力。

**回应:** 感谢审稿专家的意见，我们实验前确实没有对不同情绪韵律的声音的声学特征进行分析，收到审稿专家的意见后我们利用 Praat 软件提取了高兴和中性情绪韵律各 88 个句子的基频和共振峰信息（样例如附图 1 所示），之后进行统计分析，结果发现不同情绪韵律的声音在这些声学指标上确实存在显著的差异。具体的，分析表明，开心情绪韵律与中性韵律在 F0、F2、F3、F4 上差异均显著[F0:  $t(174) = 26.23, p < 0.001$ , 95% CI: [42.13, 48.98]; F2:  $t(174) = 2.64, p = 0.009$ , 95% CI: [16.45, 113.99]; F3:  $F(1, 160.06) = 38.06, p < 0.001$ ; F4:  $F(1, 161.12) = 7.26, p = 0.008$ ]。开心情绪韵律与中性韵律在 F1 上差异不显著,  $t(174) = 0.73, p = 0.467$ , 95% CI: [-12.25, 26.60]。这结果在一定程度上表明开心情绪韵律具有去掩蔽的作用是由于其 F0 较高。描述结果如附表 1 所示。但作者认为对于本研究的主题，即情绪韵律线索和情绪语义线索在去信息掩蔽方面有不同的机制而言，这一声学特征的分析并无直接的关系。因此我们打算将其作为正式内容写在正文中。



附图 1.中性与积极情绪韵律声音波形与时频图示意

附表 1：中性与积极情绪韵律声音的基频与共振峰参数

情绪韵律	F0(Hz)	F1(Hz)	F2(Hz)	F3(Hz)	F4(Hz)
中性	245.14 ±	692.51 ±	1862.42 ±	3275.04 ±	4241.15 ±
	10.42	62.70	184.09	72.42	54.66
积极	290.70 ±	699.69 ±	1927.64 ±	3194.82 ±	4214.93 ±
	12.53	67.77	140.86	98.16	73.09

此外，我们还对自检过程中发现的一些其他诸如措词，标点等小问题进行了系统修改，在此不一一列举。详见正文。再次感谢审稿专家的辛苦工作。

## 第二轮

**审稿人 1 意见：**作者已经详细地处理了一审中审稿人提出的意见，文章质量有大幅提升，建议在文稿进一步修改得更为简洁后可接收发表。

**审稿人 2 意见：**作者认真仔细地回答了评审的所有问题，并对不当之处进行了必要的说明与完善，使得研究结论更加妥当合理，学术解析更加严谨，应用价值也得到进一步体现。同意发表。

**编委意见：**此稿件同意发表

**主编终审意见：**

结论中，有“言语的情绪韵律能够优先吸引听者更多的注意，主要对知觉信息掩蔽起作用。言语的情绪语义能够优先获取听者更多的认知加工资源，具有去认知信息掩蔽的作用，而不具有去知觉信息掩蔽的作用。”陈述，其中的“情绪韵律能够优先吸引听者更多的注意”

和“情绪语义能够优先获取听者更多的认知加工资源”之间的关系不容易区分，“认知加工资源”相对比较宽泛，它可能包含了“注意”。建议明确“认知加工资源”的含义。

回应：感谢主编的宝贵意见，此处表述确实有容易混淆之处，我们在原文中相应部分将“认知加工资源”修改为“和内容加工相关的高级认知加工资源”，具体为，“言语的情绪语义能够优先获取听者更多**和内容加工相关的高级**认知加工资源，具有去认知信息掩蔽的作用，而不具有去知觉信息掩蔽的作用。”

除此外，我们通读全文，还发现有几处句子不太通顺之处，为方便读者理解，在不影响整体语义的情况下对句子进行了适当的调整。具体有如下几处（所标页码为本次所提交文稿中页码）：

1. 第 3 页第一段最后一句，“这一结果也为情绪韵律具有去听觉掩蔽的作用提供了神经机制证据。”修改为“这一结果也为情绪韵律具有去听觉掩蔽的作用提供了神经机制**方面**的证据。”
2. 第 3 页第三段最后一句，“这不仅说明了对听觉呈现的消极单词的语义加工不受注意的影响，而且也表明了其能够占用更多的资源”修改为“这说明不论是在注意还是在非注意条件下，听觉呈现的消极情绪词都会占用听者更多的加工资源。”
3. 第 4 页第一段第二行，“使其能够在混合的声音流中追随”修改为“使其能够在混合的声音流中被追随”。
4. 第 4 页第二段第四行，“而句法正确的无意义语句具有语义可懂度”修改为“而正序的无意义语句具有一定程度的语义可懂度”。
5. 第 8 页第一段第三行，“这可能是由于知觉条件相对困难的情况下”修改为“这可能是由于**在**知觉条件相对困难的情况下”。