

《心理科学进展》审稿意见与作者回应

题目：微语调表情：检测情绪泄露的新框架

作者：申寻兵，丰婷婷，盛静，彭咏梅，刘仪辉，李雅方，陈振彩

第一轮

审稿人意见：

论文试图通过分析情绪泄露条件下的语音特征，考察是否存在微语调表情。理论上真实体验到的情绪可以通过面部泄露，也可以通过语音泄露出来，微语调表情在理论上具有存在的可能性。论文希望通过言语情绪识别和系列心理学实验，来为微语调表情的存在提供依据，丰富微表情的理论内涵。

以下几个问题需要作者进一步澄清：

1. 语音语调存在着地域、地区及语言本身特性等方面的差异，那么这些差异是否会影响到语调表情的分析？是否会影响根据这些语调情绪信息判断出是否存在欺骗？
2. 论文中提及了机器学习，但相关内容较少，建议进一步丰富相关内容，比如具体拟用哪些算法处理和学习哪些相应的特征。

意见 1: 1. 语音语调存在着地域、地区及语言本身特性等方面的差异，那么这些差异是否会影响到语调表情的分析？是否会影响根据这些语调情绪信息判断出是否存在欺骗？

回应: 诚如您所言，确实不同的地域、地区的语音语调及语言本身特性存在多方面的差异，为了排除这些因素对结果的影响，首次，我们在研究构想的视听表情数据库的构建工作中，尽可能多的选取有代表性的数据库，充分考虑文化背景的差异，现阶段主要包括英文和中文，后续将拓展到其他语言的语调表情数据库（在论文中对此进行了进一步说明，请见“3.1 研究 1：欺骗条件下视听表情数据库的构建”中标蓝部分的内容）。

其次，我们分析泄露条件下（欺骗条件）同一个体的存在泄露和不存在泄露的语调表情（如欺骗的“yes”和诚实的“yes”语音），排除个体差异（这一点已经在 3.2 研究 2 中的“5）游戏秀中欺骗与诚实语音特征比较分析”中有所说明）。

最后，我们将分析语调中普遍的基本情绪特征，正如面部表情一样，面部虽千差万别，但表达出来的基本表情具有普遍性。

因此，这些差异本身虽对分析可能存在影响（如识别上存在 in-group 优势，参见 Laukka, P., & Elenbaas, H. A. (2021). Cross-cultural emotion recognition and in-group advantage in vocal expression: A meta-analysis. *Emotion Review*, 13(1), 3-11.），但通过上述处理，尽可能的将影响降低到最低程度。更主要的，我们的目的是分析在言语中是否存在短暂的情绪泄露（比如可能在无情绪的语音中有短暂的某种情绪，如当众演讲中讲者努力保持镇定但泄露出来紧张、担忧），即便有这些差异存在，对于情绪泄露的分析应无影响（情绪泄露跟用什么语言说、是谁说等关系不大）。

同样，对于基于微语调表情的欺骗检测，我们分析的是欺骗时可能泄露的恐惧、高兴或内疚（主要包含基本情绪悲伤）等普遍性的情绪对应的声学特征，重点在分析情绪泄露和普遍的语音情绪特征；同时，分析同一个人的诚实和欺骗语音，从而排除个体差异的影响。

意见 2: 论文中提及了机器学习，但相关内容较少，建议进一步丰富相关内容，比如具体拟用哪些算法处理和学习哪些相应的特征。

回应: 感谢您的宝贵意见！我们在 3.2 研究 2 中的“3) 伴随微面部表情的语调表情特征分析”部分，增加了下述内容：

“对微语调表情特征的分析，除了使用 Praat 和 OpenSMILE 提取声学特征外，还将采用深度学习的方法来处理微语调表情的时间动态特征。通过将整段音频切分成短时序列来捕获微语调表情的动态特征；使用 Attention LSTM 来建模 (Xie, et.al, 2019) 找到微语调表情出现的时点。”

微语调表情具有短暂和动态的特性。针对动态特性，输入的数据不是整段音频提取到的特征，而是按短时序列来建模（比如，frame length: 25 ms, frame shift: 10 ms）；针对短暂的特性，我们将使用 Attention LSTM 来建模，因为持续时间短，如果采用一些要计算平均值的方法，将会很容易把反映微语调表情的声学特征值给平均掉，而普通 LSTM 对所有时间窗进行隐式平均，因此加入 Attention 机制，让模型学会“听哪里”，而不是平均“听完整段话”。总之，在后续的分析中，将选取可以提取高时间分辨率特征及可以测量具体微语调表情出现时点的深度学习算法。

此外，在“3.3 研究 3: 微语调表情在欺骗检测中的应用”部分，增加了下列内容：

“机器学习软件 WEKA(<https://www.cs.waikato.ac.nz/ml/weka/>) 及其深度学习工具包 WekaDeeplearning4j 将被用于对欺骗和诚实语音的分类。”

增加的相关内容，在修改稿中字体作了标蓝处理。再次感谢专家给出的中肯意见！

第二轮

编委意见:

这篇研究构想的研究内容具有一定的前沿性，且有潜在的应用价值，也符合研究构想栏目的基本要求，可以发表。不过，对文题稍有疑惑：用一个问句作为标题似乎不太符合惯例，也没有告诉读者整篇文章的主题。建议作者还是概括一下研究的主要内容和观点，提炼出一个陈述式的更符合研究内容的标题。供参考。

修改说明: 根据您的意见，将题目调整为：**微语调表情：检测情绪泄露的新框架。**