

# Scaling Effects on Semantic-Paralinguistic Disentanglement in Multimodal SpeechLLMs

Assignee Research

June 9, 2026

## Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: What is the impact of model size scaling on the ability of multimodal SpeechLLMs to disintegrate semantic meaning from paralinguistic features in the MMSU benchmark, measured by F1-score improvements. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: MMSU: A Massive Multi-task Spoken Language Understanding and Reasoning Benchmark. Research question: What is the impact of model size scaling on the ability of multimodal SpeechLLMs to disintegrate semantic meaning from paralinguistic features in the MMSU benchmark, measured by F1-score improvements?.

## 2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.2/10.

## 3 Results

12 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 2.2/10.

## 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

## References

- <http://arxiv.org/abs/2506.04779v3>
- <http://arxiv.org/abs/1806.08685v2>
- <http://arxiv.org/abs/2207.08179v1>