# FlashSpeech Zero-Shot Speaker Adaptation and Word Error Rate in Emotional Speech Datasets

Assignee Research

June 6, 2026

**Abstract**

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: What is the impact of FlashSpeech's zero-shot speaker adaptation on word error rate degradation when evaluated on out-of-domain emotional speech datasets like CREMA-D. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Speaker Adaptation Using Spectro-Temporal Deep Features for Dysarthric and Elderly Speech Recognition. Research question: What is the impact of FlashSpeech's zero-shot speaker adaptation on word error rate degradation when evaluated on out-of-domain emotional speech datasets like CREMA-D?.

## 2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.7/10.

## 3 Results

11 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.7/10.

## 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/2305.01759v1
- http://arxiv.org/abs/2508.02112v1
- http://arxiv.org/abs/2202.10290v3