

# How does Qwen-Audio’s performance on the MMSU benchmark compare to Whisper-Large-V3 and OpenPangu-7B-MLA in ze

Assignee Research

June 10, 2026

## Abstract

Zero-shot cross-lingual knowledge transfer enables the multilingual pretrained language model (mPLM), finetuned on a task in one language, make predictions for this task in other languages. While being broadly studied for natural language understanding tasks, the described setting is understudied for generation. Previous works notice a frequent problem of generation in a wrong language and propose approaches to address it, usually using mT5 as a backbone model. In this work, we test alternative mPLMs, such as mBART and NLLB-200, considering full finetuning and parameter-efficient finetuning wi

## 1 Introduction

This paper examines: Empirical study of pretrained multilingual language models for zero-shot cross-lingual knowledge transfer in generation. Research question: How does Qwen-Audio’s performance on the MMSU benchmark compare to Whisper-Large-V3 and OpenPangu-7B-MLA in zero-shot cross-lingual transfer learning when evaluated on unseen low-resource languages?.

## 2 Methodology

Systematic literature search across multiple databases yielded 9 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 6.5/10.

## 3 Results

9 papers retrieved. 10 claims extracted; 5 independently verified. Quality review score: 6.5/10.

## 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.

## 5 Extracted Claims

Claim	Verified	Confidence
mBART with adapters performs similarly to mT5 of the same size in zero-shot cross-lingual knowledge transfer for generat	✓	0.38
NLLB-200 can be competitive in some cases for zero-shot cross-lingual knowledge transfer in generation tasks.	✓	0.37
Tuning the learning rate used for finetuning helps to alleviate the problem of generation in the wrong language.	✓	0.37
With too small or too large learning rates, the performance of mT5-base and mBART is affected.	×	0.05
The validation curves for full finetuning of mT5-base and mBART with various learning rates are presented in Figure 1 fo	×	0.07
Full set of plots demonstrating the effect of learning rate in each task-model-adaptation method-language combination is	×	0.06
The problem of generation in a wrong language is frequent in zero-shot cross-lingual knowledge transfer for generation t	✓	0.35
Previous works propose approaches to address the problem of generation in a wrong language, usually using mT5 as a backb	✓	0.34
Translating test input examples into the source language, generating outputs in the source language, and translating the	×	0.12
The drawbacks of translation-based approaches include high computational cost, lack of high-quality translation models f	×	0.12

## References

- <http://arxiv.org/abs/2310.09917v3>
- <http://arxiv.org/abs/1909.07342v1>
- <http://arxiv.org/abs/2308.10783v2>