

Responsible NLP Checklist

Paper title: *Zero-Shot Context-Aware ASR for Diverse Arabic Varieties*

Authors: *Bashar Talafha, Amin Abu Alhassan, Muhammad Abdul-Mageed*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- the authors indicated that the question does not apply to their work
- the authors did not respond to the checkbox question

For background on the checklist and guidance provided to the authors, see the [Responsible NLP Checklist](#) page at ACL Rolling Review.

A. Questions mandatory for all submissions.

- A1. Did you describe the limitations of your work?

This paper has a Limitations section.

- A2. Did you discuss any potential risks of your work?

This work focuses on test-time decoding strategies for automatic speech recognition (prompting, prefixing, and proxy-guided selection) and does not introduce new models, datasets, or deployment scenarios that raise distinct societal or ethical risks beyond those already well documented for ASR systems.

B. Did you use or create scientific artifacts? (e.g. code, datasets, models)

- B4. Did you discuss the steps taken to check whether the data that was collected/used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect/anonymize it?

This work does not involve new data collection. All experiments use existing ASR benchmarks and previously curated evaluation sets.

- B6. Did you report relevant statistics like the number of examples, details of train/test/dev splits, etc. for the data that you used/created?

Relevant dataset statistics, including dataset descriptions and evaluation splits, are reported in Appendix A.1 (Datasets) and referenced throughout Section 4 (Experiments).

C. Did you run computational experiments?

- C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?

The experimental setup, model configurations, decoding settings, and hyperparameters are described in Section 4 (Experiments) and Appendix A.3 (Model Settings).

- C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?

Descriptive statistics and summary results (e.g., average WER/CER across datasets, beam-size trends, and oracle vs. top-1 comparisons) are reported in Section 4, Table-1 and Figures 23, with it made explicit when averages or single-run results are used.

The Responsible NLP Checklist used at ACL Rolling Review is adopted from NAACL 2022, with the addition of ACL 2023 question on AI writing assistance and further refinements based on ARR practice. ACL 2026 used a subset of ARR checklist form.

D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?

D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?

This paper did not involve human participants or annotators.

D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?

This paper did not involve human participants or annotators.

D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating (e.g., did your instructions explain how the data would be used)?

This study did not involve new data collection or direct interaction with human subjects; all data used comes from existing datasets with consent handled by the original dataset creators.

D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?

This paper did not involve new data collection or human subjects

E. Did you use AI assistants (e.g., ChatGPT, Copilot) in your research, coding, or writing?

E1. If you used AI assistants, did you include information about their use?

AI assistants were used only for light editing and language refinement during manuscript preparation, as English is not our first language, and did not contribute to the research, experiments, or scientific content.