

Reconstructing Performed Homer: Metrically Constrained Speech Synthesis for Ancient Greek Epic

Federico Di Pasqua^{1,*}, Markus Hafner¹, Pierfrancesco Melucci², Paolo Merialdo³ and Riccardo Torlone³

¹University of Zurich, Zurich, Switzerland

²Sapienza University of Rome, Rome, Italy

³Roma Tre University, Rome, Italy

Abstract

In antiquity, Greek and Latin poetry derived much of its musicality from quantitative rhythm: the patterned alternation of long and short syllables. Readers today can scan Homeric metre, but they rarely perceive syllable quantity as a living rhythmic feature. Homeric epic is therefore often encountered as a kind of mute score, preserving its formal structure but not its acoustic force. This paper presents a computational framework for producing plausible, philologically grounded reconstructions of Homeric recitation by combining metrical annotation, controllable speech synthesis, and cross-lingual analysis of Arabic quantitative verse. Arabic is used not as a proxy for ancient Greek performance, but as a comparative source for modelling how quantitative rhythm is realized acoustically in recorded recitations of Classical Arabic poetry. The proposed pipeline converts Homeric hexameter into syllable-level prosodic scaffolds, estimates timing parameters from Arabic recitation, maps those parameters onto language-independent prosodic targets, and generates metrically constrained Homeric Greek audio whose assumptions are explicit and evaluable. Rather than claiming to recover Homeric performance in an absolute sense, the framework makes hypotheses about duration, caesura, and phrasing audible and available for scholarly assessment. The contribution is twofold: it provides digital classics with a new instrument for studying performed metre, and it offers AI research a culturally specific test case for interpretable speech generation across historically distant quantitative poetic traditions.

Keywords

artificial intelligence for cultural heritage, Homer, ancient Greek, speech synthesis, prosody, quantitative metre, digital classics, cross-lingual speech synthesis, cross-lingual prosody transfer, Arabic poetry

1. Introduction

Contribution type and topics. Project paper on language technologies, generative AI, and intangible cultural heritage.

Ancient Greek epic has reached modern scholarship as text, but it was not originally experienced as silent script. Homeric verse belonged to a performed poetic culture in which rhythm, phrasing, formulaic diction, and voiced delivery were constitutive features of meaning [1, 2, 3]. The words survive; the metre survives; many of the formal structures that organized performance survive. What has largely disappeared is the acoustic experience that made those structures perceptible.

The loss is especially acute because Greek and Latin poetry were built on quantitative metre. In dactylic hexameter, rhythm depends on the patterned sequence and variation of long and short syllables. This differs from the stress-based and accentual patterns through which many modern readers perceive poetic rhythm. Homeric metre can therefore be analysed and scanned, but its rhythmic organization is not easily heard. The result resembles reading the score and libretto of an opera without hearing it performed: the words and formal structure remain, but a central dimension of their effect is lost.

Reconstructing Performed Homer begins from this problem. The project asks whether artificial intelligence can convert the surviving formal structure of Homeric verse into audible, historically informed

4th Italian Workshop on Artificial Intelligence for Cultural Heritage (IAI4CH 2026, <https://ai4ch.di.unito.it/>), co-located with the 24th International Conference of the Italian Association for Artificial Intelligence (AIIA 2026). 6-9 October 2026, Perugia, Italy

*Corresponding author.

✉ federico.dipasqua@uzh.ch (F. Di Pasqua)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

recitation. Its purpose is not to reproduce “Homer as he sounded,” but to test whether philological and metrical analysis can be expressed as explicit prosodic constraints. These constraints can then guide the generation of plausible renderings of Homeric hexameter whose assumptions are documented, varied systematically, and subjected to scholarly and perceptual evaluation.

The project’s central comparative move is the acoustic analysis of quantitatively metered Classical Arabic poetry. Contemporary recitations provide recorded material in which the durational realization of quantitative rhythm can be studied. Arabic is not treated as a historical surrogate for Greek performance, nor is genealogical continuity between the traditions assumed. The choice rests on quantitative metre and recorded recitation rather than linguistic relatedness; Indo-European traditions such as Vedic Sanskrit remain relevant alternatives.

The project uses these measurements to estimate candidate timing parameters, including long-to-short duration ratios, phrase-final lengthening, pause distributions, and local rate variation. The parameters are not transferred directly as Arabic surface prosody. Instead, they are normalized, represented as plausible ranges, and mapped onto language-independent controls for a duration-aware speech-synthesis system. The resulting Homeric Greek audio is generated through the interaction of three explicitly separated components: a philologically reviewed prosodic scaffold, a comparative parameter-estimation layer, and a controllable synthesis model.

The investigation is organized around three questions. First, can syllable quantity, metrical position, and phrase structure be represented as explicit synthesis parameters that a speech model realizes accurately and consistently? Second, do parameters estimated from Arabic recitation improve the audibility or controllability of quantitative rhythm relative to manually selected ratios and unconstrained text-to-speech? Third, which controlled realizations help specialists and non-specialists perceive metrical and phrase-level organization without making the output unintelligible or mechanically timed?

Generated audio is understood here as an executable model of a philological hypothesis. A decision about the strength of durational contrast, the treatment of a metrically long position, or the acoustic realization of a caesura can be heard, measured, compared with alternatives, and revised. The synthesis does not provide independent historical evidence for the assumptions encoded in it; its value lies in making their acoustic consequences available for examination. Negative results are therefore informative: if Arabic-informed controls provide no advantage over simpler baselines, or if metrically regular timing obscures linguistic phrasing, the experiment identifies a limit in the proposed model.

The contribution is methodological rather than merely demonstrative. For digital classics, the framework turns scansion from a descriptive notation into an audible research instrument. For speech technology, it defines a demanding generation task in which symbolic structure, externally controlled duration, cross-lingual evidence, and evaluation by domain specialists must be integrated in the absence of recorded Homeric Greek performance. For cultural-heritage research more broadly, it proposes a transparent way to investigate not a recoverable ancient artefact, but a partially lost mode of perception: the experience of quantitative rhythm as sound.

2. Research context

Research on Homeric oral tradition has established that formulaic diction, metre, and performance are interdependent rather than separable features of the poems. Recent work has refined this account through the study of formularity, dialect, and poetic creativity [2, 3]. Reconstruction-oriented scholarship has likewise shown that ancient sound can be modelled without disguising conjecture as certainty. Danek and Hagel’s *Homeric Singing*, for example, makes interpretive choices audible and open to criticism [4]. The present project builds on this experimental principle through explicit annotation, comparative timing estimates, and controlled synthesis.

Computational work supplies much of the required formal representation. SEDES models position in Greek hexameter, while related studies analyse caesurae and metrical shapes at scale [5, 6, 7, 8]. Dactylo similarly demonstrates that Greek hexameter can be profiled computationally [9]. These resources treat metre as structured data; the present question is how that data must be enriched before it can guide

acoustic realization.

Speech synthesis provides possible realization engines. Poetry-specific text-to-speech has demonstrated explicit control of literary reading [10], and multilingual systems using phonological or articulatory representations have addressed low-resource and zero-shot settings [11, 12]. Duration-aware and expressive architectures offer further control over timing and delivery [13, 14, 15]. Cross-lingual prosody transfer establishes a technical precedent for representing and mapping phrase-level features across languages [16, 17], while recent experiments on Latin verse explore stress-guided generation for ancient languages [18, 19]. The proposed contribution is to combine Homeric metrical structure, comparative timing measurements, and explicit duration control within one evaluable framework.

Classical Arabic text-to-speech resources demonstrate that speech technology can address the relevant language variety [20], but they do not by themselves provide evidence about poetic performance. In particular, an audiobook corpus cannot substitute for a documented corpus of metrical recitation. The project must therefore distinguish technical resources for Arabic synthesis or alignment from the poetic recordings used to estimate duration. This distinction also limits what existing cross-lingual transfer studies can establish: they motivate the representation of prosodic controls, not the historical or linguistic validity of transferring a particular ratio to Homeric Greek.

3. Proposed framework

The framework has three linked components: a Homeric prosodic scaffold, an Arabic parameter-estimation layer, and a controllable synthesis layer. Evaluation feeds back into all three by identifying which annotations, measurements, or controls fail to produce stable and perceptible effects.

3.1. Homeric annotation as prosodic scaffold

The project will build a reviewed pilot corpus of lines selected from the Homeric epics and represented at syllable level. Each record will include normalized text, a reconstructed phonological representation, syllable boundaries, phonological quantity, metrical position, foot structure, caesural candidates, phrase segmentation, and uncertainty. These categories must remain distinct. Phonological quantity is a property of linguistic form; a metrical position describes a syllable’s function within hexameter; acoustic duration is one possible realization influenced by segmental composition, context, and phrasing.

The phonological layer requires an explicit textual and historical policy. The corpus documentation will identify the critical edition or editions and the pronunciation model, and record how the system handles the traditional, dialectally mixed poetic language of Homeric epic, including digamma, elision, correction, hiatus, synizesis, and pitch accent. The selected critical text will serve as a documented textual basis, not as a recovered Homeric *Urtext*. Where a form admits more than one defensible analysis, each alternative will receive a stable identifier and confidence note. This prevents decisions made for synthesis convenience from becoming invisible claims about the language.

The scaffold combines discrete annotations with numerical controls. Its discrete layer records the line’s linguistic and metrical analysis. Its numerical layer specifies duration multipliers, pause ranges, and phrase-final adjustments relative to a model’s contextual baseline. Values may be assigned by rule, estimated from Arabic recordings, or varied systematically. This separation permits the same philological annotation to generate several acoustic hypotheses without changing the underlying text.

Duration targets will also preserve the internal structure of syllables. A syllable may be phonologically heavy because it contains a long vowel or diphthong or because it is closed by a consonant; a metrically long position may also arise through context-dependent phenomena. Uniformly lengthening every segment would not model these distinctions explicitly. The numerical layer will therefore permit phone-level allocation within a syllable while retaining syllable-level targets for metrical comparison.

Caesurae illustrate the need for this intermediate representation. A caesura is not automatically an audible silence: it may be realized through a pause, final lengthening, pitch reset, or a weaker

grouping effect. The scaffold therefore marks caesural locations and possible realization types rather than inserting fixed pauses. Uncertain scansion or segmentations are retained as alternatives instead of being silently resolved during preprocessing.

3.2. Arabic quantitative verse as comparative source

The project will assemble a small pilot corpus from existing recitations; candidate sources include the datasets described by Al-Shaibani et al. [21], subject to checks of speaker metadata, reuse terms, and suitability for phonetic alignment. Analysis will address long-to-short duration ratios at syllable level and pause placement, phrase-final lengthening, and local rate variation at higher levels. Corpus selection and annotation will record the poem, metre, speaker and speaker background, such as whether the performer is a professional reciter, actor, or scholar, as well as performance style, language variety, licensing, and recording conditions.

Measured duration cannot be attributed to quantity alone. Segment identity, syllable closure, stress, phrase position, speaking rate, and recitation style may all affect the result. The estimation layer will therefore model these factors where the available data permits and report uncertainty rather than presenting one Arabic ratio as a universal constant. The comparative layer concerns timing; Greek pitch accent will be specified separately within the pronunciation model.

Recordings will be aligned at phone and syllable level and checked manually on a representative subset. Estimates will be normalized by speaker and rate, stratified by relevant linguistic and metrical contexts, and reported as distributions or plausible ranges rather than a single point value. The pilot corpus should include more than one speaker and metre whenever licensing and available recordings permit. If the evidence is too sparse to separate quantity from performance style, that limitation will be carried into the Homeric Greek controls as an uncertainty range.

Arabic-informed controls are an optional refinement: the system can first be tested with manually selected ratios. The project will compare unconstrained synthesis, Homeric Greek metrical controls with manually selected ratios, and controls informed by Arabic measurements. This design determines whether the comparative layer adds value beyond explicit scansion alone and prevents Arabic performance from being treated as direct evidence for Greek historical practice.

3.3. Controllable speech synthesis

The synthesis system receives normalized Homeric text, reconstructed phonology, and the prosodic scaffold. A duration-aware multilingual architecture will serve as the initial realization engine, since the task requires phone- or syllable-level timing control and cannot be learned from native Homeric recordings. Candidate architectures include FastSpeech 2 and multilingual systems derived from articulatory-feature representations [13, 11, 12]. Model selection will depend on whether inference-time durations can be overridden reliably and whether the phoneme inventory can represent the chosen phonological reconstruction of Homeric Greek.

An initial feasibility study will compare candidate models on four requirements: phoneme coverage, duration and pitch control, stability when external timing is imposed, and reproducibility across speakers or random seeds. The chosen system may be used without target-language fine-tuning or adapted through multilingual and articulatory representations; it will not be described as conventionally “low-resource” if no Homeric speech data is used. Homeric Greek phones absent from the training inventory will require documented approximation or feature-based composition.

Controls will modify contextual baseline durations rather than assign one absolute value to every long or short syllable. Higher-level parameters will govern phrase grouping and possible caesural realizations. The implementation must also document the mapping of Greek sounds into the model’s inventory and the source of the synthetic voice.

The voice itself will be treated as a realization resource, not a reconstruction of ancient vocal identity. Speaker identity, perceived gender, and expressive style will be documented as properties of the synthesis system rather than treated as evidence for ancient performance. Where technically feasible,

the same scaffold will be rendered with more than one synthetic speaker. Agreement across voices would suggest that an effect comes from the metrical controls rather than the idiosyncrasies of one pretrained speaker; disagreement would identify another source of uncertainty.

Every generated file will be paired with a machine-readable manifest containing its text, phonological input, annotation version, target durations, phrase settings, comparative parameter source, model version, and random seed. Controlled variants can then differ in one assumption at a time. Synthesis functions here as an experimental instrument: it exposes the acoustic consequences of an analysis and permits those consequences to be measured and judged.

4. Evaluation, reuse, and limitations

4.1. Evaluation design

Evaluation combines controlled comparisons with technical measurements, assessment by specialists in Greek philology, Arabic poetics, and speech technology, and perceptual measures. The principal conditions are unconstrained text-to-speech, manually parameterized Homeric Greek metre, Arabic-informed timing, and variants with different caesural or phrase-level controls.

Technical evaluation will measure phone- and syllable-duration error against prescribed targets, realized long-to-short ratios, stability across repeated generations, intelligibility, and responsiveness to parameter changes. Evaluation materials will include held-out lines and controlled ablation tests, in which one component—quantity, caesural treatment, or Arabic-informed timing—is removed at a time to measure its contribution. The analysis will also test whether results remain stable across speech rate and synthetic voice.

Specialist evaluation will use rating scales, pairwise comparisons, and error annotation. Scholars of Greek language and metre will assess pronunciation, metrical legibility, and philological plausibility; specialists in Arabic poetics will evaluate corpus selection, measurement, and the limits claimed for comparison; speech researchers will assess robustness and controllability. Inter-rater agreement will be reported for specialist judgments.

Listener studies will test more than whether two outputs sound different. After minimal instruction, participants will be asked to discriminate metrically constrained output from rhythmically disrupted controls, identify recurring rhythmic organization, and judge phrase or caesural groupings. These evaluations distinguish acoustic compliance, scholarly plausibility, and perceptual effect; without ancient recordings, listener judgments of naturalness cannot establish historical authenticity. Study size, recruitment, scales, and statistical comparisons will be specified before the main evaluation rather than after listening to the outputs.

4.2. Reuse and sustainability

The project will release, where rights permit, a pilot Homeric corpus with linguistic and metrical annotation; documented mappings from annotations to prosodic controls; measurement procedures for Classical Arabic recitation and derived parameter ranges; reproducible inference configurations; and an interface for comparing reconstruction hypotheses. Durable annotations and manifests will be separated from model-specific code so that the scholarly data remains usable as architectures change.

Each release will include schemas, annotation guidelines, version history, licensing information, and example lines that connect the textual analysis to generated manifests. A benchmark designation will be used only once the project defines a stable dataset split, tasks, baselines, and metrics.

The same materials support teaching and research. Students can compare visual scansion with controlled audio, while researchers can render competing analyses under shared conditions. The interface will support both analytical comparison and direct listening, while clearly identifying every output as a reconstruction hypothesis. The method may later be adapted to Latin hexameter, Greek lyric, and other traditions whose formal structures survive more securely than their acoustic realization. Sustainability will rest on open, versioned annotations and documented transformations rather than on any

proprietary voice or transient model.

4.3. Responsible reconstruction

Synthetic speech can sound more authoritative than the evidence warrants. Outputs will therefore be labelled as reconstruction hypotheses and accompanied by the assumptions that produced them: textual basis, pronunciation, scansion, timing parameters, caesural treatment, comparative source, and model settings. The system cannot recover either an original Homeric text or the authentic sound of ancient performance: each output begins from a specified critical edition and phonological analysis and renders one documented set of acoustic assumptions.

Uncertainty will be visible at the point of listening, not confined to technical documentation. Users will be able to distinguish alternatives grounded in textual or metrical ambiguity from variants introduced only to explore acoustic parameters. Explanatory text will also separate confidence in an annotation from the naturalness of the generated voice, since a fluent output is not necessarily a better historical hypothesis.

The Arabic component requires both methodological and cultural care. Before corpus selection is finalized, the project will establish formal collaboration with a specialist in Classical Arabic poetics, who will review the selection, interpretation, and evaluation of the material. Should new recordings be required, they will use documented consent and reuse terms; existing material will be used only under compatible licences. Publications and interfaces will state that Arabic supplies comparative acoustic measurements, not evidence of historical identity or contact between the traditions.

Listener studies will follow applicable institutional requirements for research with human participants. Within the project team, Federico Di Pasqua serves as principal investigator and digital-humanities lead; Markus Hafner leads Homeric philology, including textual, phonological, and metrical policy, annotation guidelines, and expert validation; Pierfrancesco Melucci leads speech-model adaptation, controllable generation, and acoustic evaluation; and Paolo Merialdo and Riccardo Torlone jointly lead software and data engineering, including pipeline architecture, provenance, and the comparison interface.

5. Conclusion

Reconstructing Performed Homer proposes a transparent method for converting Homeric metrical analysis into controlled audio. Its contribution is not a claim to recover ancient performance, but a framework for rendering and testing specific hypotheses about quantitative timing and phrasing. By separating annotation, comparative parameter estimation, synthesis, and evaluation, the project makes each assumption inspectable and revisable. The project would offer a concrete example of how AI can help scholars reason through incomplete evidence by making the consequences of their interpretations audible. The same principle could inform other cultural-heritage reconstructions, provided that the links between surviving evidence, modelling assumptions, and generated outputs remain explicit and open to evaluation.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT for paraphrasing and rewording, improving writing style, grammar and spelling checks, and formatting assistance. The authors reviewed and edited all output and take full responsibility for the publication's content.

References

- [1] M. L. West, *The Making of the Iliad: Disquisition and Analytical Commentary*, Oxford University Press, Oxford, 2011. doi:10.1093/acprof:osobl/9780199590070.001.0001.

- [2] C. Bozzone, Formularity, in: *Homer's Living Language: Formularity, Dialect, and Creativity in Oral-Traditional Poetry*, Cambridge University Press, Cambridge, 2024, pp. 5–63. doi:10.1017/9781009067157.002.
- [3] B. McGillivray, M. A. Rodda, Computational valency lexica and Homeric formularity, *Journal of Greek Linguistics* 24 (2024) 264–289. doi:10.1163/15699846-02402003.
- [4] G. Danek, S. Hagel, Homeric singing: An approach to the original performance, Scholarly online publication, Austrian Academy of Sciences, 2002. URL: <https://www.oeaw.ac.at/kal/sh/index.htm>, accessed 13 September 2026.
- [5] S. A. Sansom, Sedes as style in Greek hexameter: A computational approach, *TAPA* 151 (2021) 439–467. doi:10.1353/apa.2021.0017.
- [6] S. A. Sansom, Breaking Hermann's bridge from Homer to Nonnus: Towards a stylometry of caesurae, *The Classical Quarterly* 75 (2025) 41–57. doi:10.1017/S0009838824000892.
- [7] S. A. Sansom, D. Fifield, SEDES: Metrical position in Greek hexameter, *Digital Humanities Quarterly* 17 (2023). URL: <https://dhq.digitalhumanities.org/vol/17/2/000675/000675.html>. doi:10.63744/h8amf4dwm329.
- [8] S. A. Sansom, D. Fifield, Epic rhythm: Metrical shapes in Greek hexameter, *Greek, Roman, and Byzantine Studies* 64 (2024) 350–377. URL: <https://grbs.library.duke.edu/index.php/grbs/article/view/17015>.
- [9] O. Schoisswohl, E. C. Papakitsos, Automated metric profiling and comparison of ancient Greek verse epics in hexameter, *Linguistik Online* 103 (2020) 159–177. doi:10.13092/lo.103.7192.
- [10] J. Koch, F. Lux, N. Schauffler, T. Bernhart, F. Dieterle, J. Kuhn, S. Richter, G. Viehhauser, N. T. Vu, PoeticTTS - controllable poetry reading for literary studies, in: *Interspeech 2022*, 2022, pp. 1223–1227. doi:10.21437/Interspeech.2022-10841.
- [11] F. Lux, N. T. Vu, Language-agnostic meta-learning for low-resource text-to-speech with articulatory features, in: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Association for Computational Linguistics, Dublin, Ireland, 2022, pp. 6858–6868. URL: <https://aclanthology.org/2022.acl-long.472/>. doi:10.18653/v1/2022.acl-long.472.
- [12] F. Lux, J. Koch, N. T. Vu, Low-resource multilingual and zero-shot multispeaker TTS, in: *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, Association for Computational Linguistics, Online only, 2022, pp. 741–751. URL: <https://aclanthology.org/2022.aacl-main.56/>. doi:10.18653/v1/2022.aacl-main.56.
- [13] Y. Ren, C. Hu, X. Tan, T. Qin, S. Zhao, Z. Zhao, T.-Y. Liu, FastSpeech 2: Fast and high-quality end-to-end text to speech, in: *International Conference on Learning Representations*, 2021. URL: <https://openreview.net/forum?id=piLPYqxtWuA>.
- [14] Y. A. Li, C. Han, V. S. Raghavan, G. Mischler, N. Mesgarani, StyleTTS 2: Towards human-level text-to-speech through style diffusion and adversarial training with large speech language models, in: *Advances in Neural Information Processing Systems*, volume 36, 2023, pp. 19594–19621. URL: https://proceedings.neurips.cc/paper_files/paper/2023/hash/3eaad2a0b62b5ed7a2e66c2188bb1449-Abstract-Conference.html. doi:10.52202/075280-0860.
- [15] K. Shen, Z. Ju, X. Tan, Y. Liu, Y. Leng, L. He, T. Qin, S. Zhao, J. Bian, NaturalSpeech 2: Latent diffusion models are natural and zero-shot speech and singing synthesizers, in: *The Twelfth International Conference on Learning Representations*, 2024. URL: <https://openreview.net/forum?id=Rc7dAwVL3v>.
- [16] J. Swiatkowski, D. Wang, M. Babianski, P. Lumban Tobing, R. Vipplerla, V. Pollet, Cross-lingual prosody transfer for expressive machine dubbing, in: *Interspeech 2023*, 2023, pp. 4838–4842. doi:10.21437/Interspeech.2023-437.
- [17] J. Swiatkowski, D. Wang, M. Babianski, G. Coccia, P. Lumban Tobing, R. Vipplerla, V. Klimkov, V. Pollet, Expressive machine dubbing through phrase-level cross-lingual prosody transfer, in: *Interspeech 2023*, 2023, pp. 5546–5550. doi:10.21437/Interspeech.2023-441.
- [18] M. Ciletti, Prompting the Muse: Generating prosodically-correct Latin speech with large lan-

- guage models, in: Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop), Association for Computational Linguistics, Vienna, Austria, 2025, pp. 740–745. URL: <https://aclanthology.org/2025.acl-srw.48/>. doi:10.18653/v1/2025.acl-srw.48.
- [19] M. Ciletti, Veras audire et reddere voces: A corpus of prosodically-correct Latin poetic audio from large-language-model TTS, in: Proceedings of the Eleventh Italian Conference on Computational Linguistics (CLiC-it 2025), volume 4112 of *CEUR Workshop Proceedings*, CEUR-WS.org, Cagliari, Italy, 2025, pp. 256–262. URL: <https://aclanthology.org/2025.clicit-1.27/>.
- [20] A. Kulkarni, A. Kulkarni, S. A. M. Shatnawi, H. Aldarmaki, ClArTTS: An open-source Classical Arabic text-to-speech corpus, in: Interspeech 2023, 2023, pp. 5511–5515. doi:10.21437/Interspeech.2023-2224.
- [21] M. S. Al-Shaibani, Z. Alyafeai, I. Ahmad, A. K. S. Al-Zahrani, Poem meter classification of recited Arabic poetry: Integrating high-resource systems for a low-resource task, 2025. doi:10.48550/arXiv.2504.12172, arXiv:2504.12172.