Recovering the Vowel Behind Schwa: Orthography-Guided /ə/ Resolution in English-to-Georgian Transliteration through Grapheme–Phoneme Alignment

Authors:Giorgi Lekiashvili, Caucasus University
Tinatin Bolkvadze, Tbilisi State University
Nata Chkhaidze, Ivane Javakhishvili University, Tbilisi, Georgia
Abstract ID: 462
Event: Linguistics 2026 Paris
Subject: 19. Computational linguistics
Presenter Name: Giorgi Lekiashvili

Abstract

English schwa creates a specific transliteration problem for Georgian. English can neutralize several vowel qualities in unstressed position as /ə/, whereas Standard Georgian has a five-vowel system /a, e, i, o, u/ and no phonemic schwa category to receive that reduction directly. A phoneme-only converter must therefore choose a full Georgian vowel after the English pronunciation has already discarded part of the contrast that would help make that choice. This study proposes an orthography-guided solution in which pronunciation first identifies a schwa token and a customized global sequence-alignment procedure then recovers the grapheme or graphemic chunk associated with that token. The implementation adapts Needleman–Wunsch dynamic programming to grapheme–phoneme alignment through longest-match grapheme tokenization, linguistic compatibility scores, context-sensitive gap costs, and one-grapheme-to-multiple-phoneme macro transitions. The recovered spelling feature conditions the Georgian output: in the evaluated core classes, <a>, <e>, <i>, <o>, and <u> aligned with unstressed schwa lead respectively to ა, ე, ი, ო, and უ, while <ou> is handled as a multigraph class. Structural rules may override vowel restoration where an English schwa is better suppressed in Georgian. A controlled 30-item diagnostic test, balanced across six spelling classes, recovered the pre-specified schwa source in all 30 target tokens. In a second ablation, removing spelling-alignment information changed 20 of 30 complete outputs, specifically the <e>, <i>, <o>, and <u> classes; <a> and <ou> converged with the system's generic AH/AX → ა fallback. These results are functional validation rather than an estimate of unrestricted accuracy. The contribution is a transparent computational treatment of schwa as an information-recovery problem at the interface of phonetics, phonology, orthography, and transliteration, not a claim that English spelling reveals a speaker's underlying phonological vowel.