

Speech is a peculiarly human characteristic. No other animal produces the gamut of sounds that we do, and we use this to measure maturity: an ‘infant’ is one who does not speak. To program a computer to synthesize speech, then, is to teach a machine what it is to be human.

Text-to-Speech is essential for visually-impaired computer users. The voices are designed to be clear, and they can also be pedagogically useful in language labs, as a model of pronunciation and speech style. Poeta ex Machina is designed for interactive learning, and it is the first working attempt to synthesize the recitation of Latin poetry. It uses Perseus’ morphological analysis to find the potential scansion of a word, disambiguates each anepithet with a meter-independent scansion engine, produces a list of phonemes and associated pitch points, and calls MBROLA to create an mp3.

MBROLA is a popular phoneme-to-speech diphone synthesizer that has dozens of diphone databases, databases of recorded speech spliced at the midpoints of each phoneme to compensate for coarticulation (cf the difference of *l* in kale/cauliflower/cold). Allen’s *Vox Latina* references the pronunciation of Italian more than any other language besides English, and MBROLA’s Italian voice #2 is the default.

The default pronunciation is Classical, from Allen’s *Vox Latina*, with an Ecclesiastical option. The prosody of Latin poetry (the pitch structure of sentences and lines) is poorly attested. Poeta ex Machina derives its fundamental pitch contour from an internal model of lung air volume. A perversely long sentence with only sparse punctuation will cause a continuous reader to start at a high pitch and eventually descend as the lungs run out of air to a much lower pitch until the sentence provides a comma for a breath, when the pitch rises again and falls again until the lowest point at the end of the sentence. Daitz, rightly, has noted that some punctuation marks, at times, could not denote a pause, just grammar and pitch: between an elision, or after a short syllable, for example. The lung air volume approximation is adequate for most poetry: there is more punctuation than in prose, and many more parses of a sentence. (For example, Catullus 5.4 *soles occidere et redire possunt* makes a complete thought without the last word: *soles* can be a verb or a noun.) A grammar-based Romance accent, averaging European prosody and intonation, is in planning.

The scansion engine relies on Perseus’ morphological analysis, and the quality of scansion depends on the accuracy of the dictionary and the declension/conjugation algorithms, which are adequate. Corpus-based improvements to this are in development.

Papers, source code, and a live web application are available at [www.poetaexmachina.net](http://www.poetaexmachina.net). Thanks to Professors Joseph Pucci and Gregory Crane for feedback and encouragement.