TOWARDS A PROSODY MODEL OF ATTIC TRAGIC POETRY: FROM LOGOS TO MOUSIKÉ

Anastasia Georgaki
Department of Music
University of Athens, Greece
georgaki@music.uoa.gr

Stelios Psaroudakēs
Department of Music
University of Athens, Greece
spsaroud@music.uoa.gr

Martin Carlé
Institute of Music and Media Studies
Humboldt-University, Berlin
martin.carle@culture.hu-berlin.de

Panagiotis Tzevelekos
Department of Informatics and Telecommunications, University of Athens, Greece
taktzev@di.uoa.gr

ABSTRACT

Recently there has been increasing interest of scientists for the performance of singing or reciting voices of the past in utilising analysis-synthesis methods. In the domain of Ancient Greek musicology indeed, where we find the roots of the occidental music, the main research has been done mostly by scholars of classical Greek literature. However, there is still a vast territory for research in audio performances to be carried out with the help of new digital technologies.

In this paper, we will present an attempt to decode a recited text of Ancient Greek tragedy and render it into sound. At the first paragraph of this article we underline the origin of music arising from the melodicity of speech in Ancient Greek tragedy. In the second paragraph, we describe the methodology we have used in order to analyse the voice of S. Psaroudakēs, himself professor of Ancient Greek music, by an open source prosodic feature extraction tool based on *Praat*. We give a description of the prosodic analysis, implementation details and discuss its feature extension capabilities as well. Last, we refer to the difference between the Ancient and Modern Greek phonological system, the application of this research in music and further development.

1. INTRODUCTION

Our sources on the pronunciation of ancient Hellenic speech (at least, classical Attic) mention two main “prosodies”:

1) syllabic duration (long, short), and
2) musical accents (acute, grave, circumflex).
3) To this phonetic picture modern scholarship adds another prosody: stress accent.

Based on this information, an attempt has been made by Professor S. Psaroudakēs, to reconstruct the sound (intonation) of poetic recitation in an extract from the parodos of Aeschylus’ *Agamemnon* (ll.40-46) [1].

The present paper describes and analyses the recorded vocal contours. It concentrates especially on the pitch contour of an expert voice in order to explore the melodicity of Ancient Greek poetry according to an analysis of the theoretical sources and possible alternatives or modifications of their application in performance. Particularly with regard to the insight of the pre-semiotic linguist Wilhelm von Humboldt which he gained as he was *metrically* translating "Aeschylos Agamemnon" we like to quote: "[...]; one always thinks to find everything in the mental domain. Though it's not the place to elaborate on this here; it always appeared to me that it's predominantly the way in which letters link to syllables and syllables combine to words in language, and how in speech again those words link up with one another according to timing and tone which defines the intellectual, yes that it actually designates no less than the moral and political fate of a nation." [20, p. 136]

2. THE PROSODY OF THE TRAGIC POETRY IN ATTIC THEATER

As our main concern is to analyse the interpretation of a text by the performance of an expert in the field, like professor Psaroudakēs, this makes up just the beginning of a new research that can lead us to elaborate on the relationship between music and logos [5]. What was the pitch range of the voice while reciting the tragic text? Which were the timbers according to the meaning of the words, the dynamics, the rhythm?

According to scholars [17, 10] the solo singing voice was particularly associated with Greek tragedy. Early tragic actors’ roles may have consisted almost entirely of singing. The actor of fifth-century tragedy had to sing in a variety of metres in rapid succession and to negotiate the delicate transitions between them: the shift between recitative and lyrics was regarded as particularly emotive [19]. Anapaestic and lyric verses repeatedly alternated with iambic trimeters, and these were spoken.

[Dekaton men etos tod' ẹpẹẹ Priamọọ megas antidikos
Menelaaos anax ẹẹd' Agamemnọọn,
ditʿronuu Diotʿen kai diskẹẹptruu
timẹẹs okʿüron zdeugos Atrẹẹdaan,
stolon Argẹẹọọn kʿilionautẹẹn
tẹẹsd' apo kʿọọraas ẹẹran, stratiọọtin arọọgẹẹn,]

Figure 1. the recorded text: parodos of Aeschylus' *Agamemnon* (ll.40-46).

Besides some important external evidence, tragic poetry offers internal clues to the way in which the voice was being used; in iambics people constantly use such verbs as *legein* and *phrazein* in reference to their own speech and that of their interlocutors, whereas the semantic range referring to lyric utterance, - which includes *melpein* and *aidein*, is quite different.[1] [10]

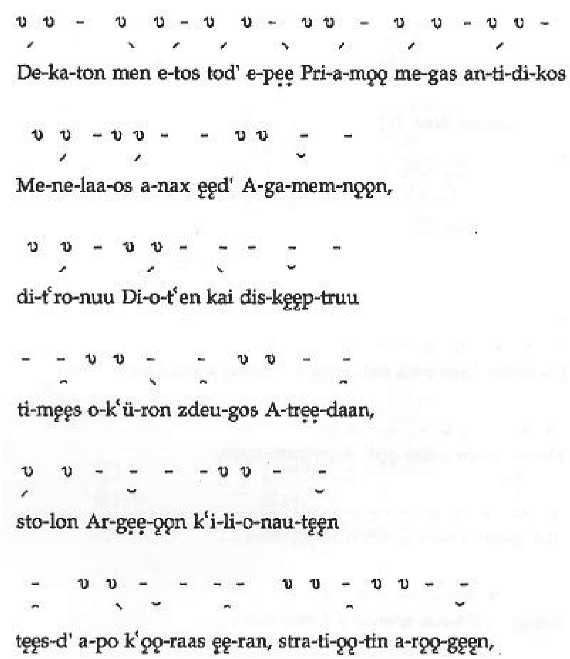

Figure 2. The tonal movement of the voice upon the analysis of Psaroudakēs [1].

Tragedy thus offered the dramatist a palette of vocal techniques with which to paint his sound pictures, and certain patterns can be discerned in the way that he handled them. Our main concern is to develop this palette of vocal techniques in order to trace new audio ways of performing the tragic text and thus make the connection of Tragedy to the etymology of *tragoudi* which in modern Greek means song.[2]

2. METHODOLOGY FOR THE PROSODIC MODELLING

Intonation modelling for speech synthesis is now one of the big issues facing not only speech synthesis systems but also music synthesis systems.

Figure 3. The rhythmic interpretation of the text upon the analysis of Psaroudakēs [1]

[1]assets.cambridge.org/97805216/51400/excerpt/9780521651400_excerpt.pdf

[2] Theatrical singers are attested from tragedies of Thespis in the sixth century BC to the Byzantine theatres in which Theodora performed in the sixth century AD, when the word 'tragedy' gave rise to what is still the word for 'song' in the Greek language (*tragoudi*).

One of the basic questions that derive from ancient Greek text is the intonation of speech according to the text and the emotions. Many evolutionary theories suggest that musical behaviour evolved in conjunction with, or as an adaptation of, vocal communication [7].

In music, pitch and temporal relations define musical tunes, which retain their identities across transformations in pitch level and tempo. In speech, pitch variation provides an important source of semantic and emotional information. In addition, temporal properties help listeners to determine boundaries between words and phrases. Typically, descending pitch contours and syllables or notes of long duration mark ends of phrases in speech [11] and in music [8].

The problem of intonation modelling for speech synthesis is summed up by the following quote regarding segmental effects on pitch: We have some basic intuitive ideas about what natural pitch should sound like, but we just don't understand enough to know how the pitch associated with a specific segment, in a specific syllable with a specific accent, in a specific word in a specific phrase with a specific phrase type, in a specific context, spoken by a specific speaker, should behave.

We have followed the next steps in order to extract the prosodic contours of the recited voice of Psaroudakēs according to the last research results of pronunciation [3] utilising the open source environment of *Praat* in order to investigate and compare the special interpretation and the diagrams that professor Psaroudakēs has designed by hand concerning the rhythm, the accent and the intonation. [1].

1.) Record the text by the special interpretation of S. Psaroudakēs according to the Allen's pronunciation system.
2.) Implement the prosodic feature extraction tools in order to get several features that describe the performative model of the prosody.

3. IMPLEMENTATION

For the implementation of the prosodic feature extraction tool we have chosen *Praat's* programmable scripting language [9].

An important reason to use *Praat* as our platform is that it provides an existing suite of high quality speech analysis routines, such as pitch tracking.

Several different types of features are computed based on the stylised pitch contour.

- Range features: These features reflect the pitch range of a single word or a window preceding or following a word boundary. These include the minimum, maximum, mean, and last values of a specific region (i.e., within a word or window) relative to each word boundary. These features are also normalised by the baseline values, the top line values, and the pitch range using linear difference and log difference.
- Range features: These features reflect the pitch range of a single word or a window preceding or following a word boundary. These include the minimum, maximum, mean, and last values of a specific region (i.e., within a word or window) relative to each word boundary. These features are also normalised by the baseline values, the top line values, and the pitch range using linear difference and log difference.
- Movement features: These features measure the movement of the contour for the voiced regions of the word or window preceding and the word or window following a boundary. The minimum, maximum, mean, the first, and the last stylised values are computed and compared to that of the following word or window, using log difference and log ratio.
- Slope features: Pitch slope is generated from the stylised pitch values. The last slope value of the word preceding a boundary and the first slope value of the word following a boundary are computed. We also include the slope difference and dynamic patterns (i.e., falling, rising, and unvoiced) across a boundary as slope features, since a continuous trajectory is more likely to correlate with non-boundaries; whereas, a broken trajectory tends to indicate a boundary of some type.
- Energy features: The energy features are computed based on the intensity contour produced by *Praat*.

Similar to the features, a variety of energy related range features, movement features, and slope features are computed, using various normalisation methods.

- Other features: We add the gender type to our feature set. Currently the gender information is provided in a metadata file, rather than obtaining it via automatic gender detection.

	Duration Features	$\bar{F_0}$ Features	Energy Features
Word	√	√	√
Phone	√	×	×
Vowel	√	×	×
Rhyme	√	×	×
VUV	×	√	×
Raw Pitch	×	√	×
Stylized Pitch	×	√	×
Pitch Slope	×	√	×
Raw Energy	×	×	√
Stylized Energy	×	×	√
Energy Slope	×	×	√

Table 1. The use of raw files for extracting various features

Given a corpus with audio and time aligned words and phones as input, our tool first extracts a set of basic elements (e.g., raw pitch, stylised pitch, VUV) representing duration, and energy information, as is shown in Figure 5 (c). Then a set of duration statistics (e.g., the means and variances of pause duration, phone duration, and last rhyme duration), related statistics (e.g., the mean and variance of logarithmic values), and energy related statistics are calculated. Given the duration, , and energy information, as well as the statistics, it is straightforward to extract the prosodic features at each word boundary, according to the definition of features in and our tool documentation [12].

We describe below how to obtain and represent these basic elements in *Praat*. Table 1 summarises their use in the computation of the prosodic features.

Word and Phone Alignments: A forced alignment system 2 is used to determine the starting and ending times of words and phones. In our tool these alignments are represented in *TextGrid IntervalTiers*, as in Figure 4

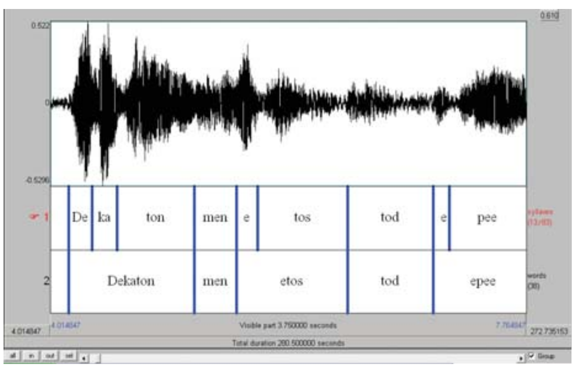

Figure 4. Sound and Textgrid Tiers of syllabus and words

Vowel and Rhyme: The starting and ending times of vowels and rhymes are determined from the phone alignments. As to rhyme, we only consider the last rhyme, which is defined to be the sequence of phones starting from the last vowel and covering all the remaining phones in a word. Vowels and rhymes are also represented in *TextGrid IntervalTiers.*

We rely on *Praat's* autocorrelation based pitch tracking algorithm to extract raw pitch values, using gender dependent pitch range.

The raw pitch contour (Figure 5) is smoothed and the voiced/unvoiced regions are determined and stored in a *TextGrid IntervalTier. Praat's* pitch stylisation function is used to stylise raw values over each voiced region (Figure 6). Both raw values and stylised values are represented in *PitchTiers*. The pitch slope values are generated based on the stylised pitch contour, and are stored in a *TextGrid IntervalTier.*

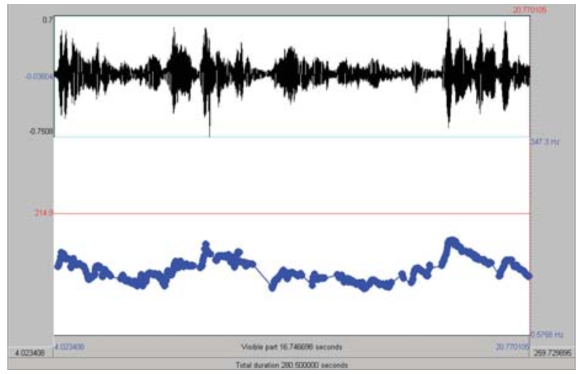

Figure 5. Sound and raw pitch countour (Pitch Tier)

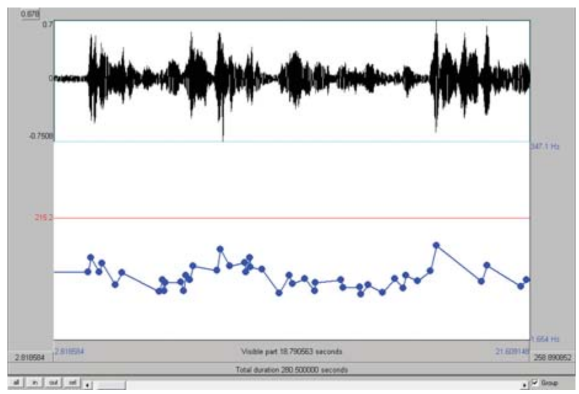

Figure 6. Sound and stylised pitch contour (Pitch Tier).

Energy: Intensity values are computed for each frame and stored in an IntensityTier (Figure 7). Since there is no intensity stylisation function in *Praat*, we choose to represent intensity values in a *PitchTier*, and apply the pitch stylisation function to stylise the intensity contour. Note that stylisation is performed on the entire intensity contour, in contrast to the pitch case, for which this applies only in voiced regions. The raw and stylised intensity values are stored in *PitchTiers*, and the slope values are stored in a TextGrid *IntervalTier.*

As we discussed above, the major advantage of building a prosodic model based on *Praat* is the capability of taking advantage of *Praat's* existing built-in speech analysis algorithms and other *Praat* scripts that have been written as extensions. In addition, because *Praat* is a public domain tool, there is the promise of future extensions to *Praat* functionality.

Although the features we have implemented have been used for a variety of event detection tasks, they are not necessarily equally effective for all tasks. Hence, it is important to have the flexibility to easily add new features to the system.

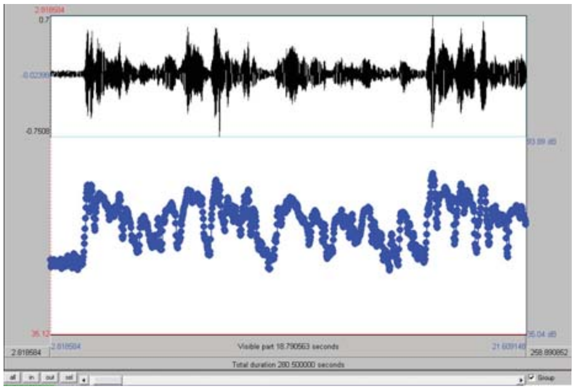

Figure 7. Sound and raw intensity contour (Intensity Tier).

4. FURTHER DEVELOPMENT

This research is only the starting point of understanding the audio aspects of Ancient Greek music.

In the future we hope to focus, on the extraction of rules that describe this kind of prosody and implement it in the model of Greek speech synthesis *Demosthenes* [18] in order to instruct the system with new information about ancient Greek poetry. This will be a further starting point to compare the phonological systems of Ancient[3] and Modern Greek concerning singing and the relation of the language.

On the other hand our research will be appreciated by directors and actors of Ancient Greek tragedy who seek for approaching the realistic way of interpreting the text [4, 8].

5. CONCLUSION

In this paper, we have attempted to extract especially the pitch contours as well as other features that characterise the interpretation of the tragic text prosody upon the special analysis of professor S. Psaroudakēs.

This research initiates a new domain: computational archeomusicology where computer-based analytical methods are used for the study of long-term musical behaviour and evolution of ancient Greek music, its systematisation and systems. The procedure of the prosodic feature extraction has been achieved through the environment extraction tools we built based on *Praat*.

[3] In modern Greek, especially in singing, we find great differences in the accentual system. Ancient Greek is generally referred to as a language with melodic accent [3] using register tone (as opposed to e.g. Mandarin Chinese which uses contour tone), varying the pitch of the voice and thereby changing the lexical or the grammatical meaning of a word [16, p. 253]. Modern Greek with its dynamic accent focuses on using stress, i.e., using more air and muscular energy, thereby producing a mixture of increased loudness, pitch and quantity [16, pp.249–250]. This difference manifested itself in the Modern Greek writing system, in that all accentual signs that had been introduced by Aristophanes of Byzantium around the third century BCE, except for the acute, the οξεία, were dropped from the official orthography in the reforms of the 1970s and 80s.

6. REFERENCES

[1] Psaroudakes S. "Meters in ancient Greek tragedy «Μέτρα στην αρχαία ελληνική τραγωδία»", *Μούσα 3:77*, Athens, 1996, pp. 77-92.

[2] Allen, W. S. *Accent and rhythm, Prosodic features in Latin and Greek: a study in theory and reconstruction.* Cambridge University Press, Cambridge, 1993.

[3] Allen, W. S. *Vox Graeca. The pronunciation of Classical Greek.* Cambridge University Press, Cambridge, 1987.

[4] Georgaki, A. "Allegories et mutations vocales au sein de la Yannis Christou (1926-1971) et Michael Adamis(1929)", in: Giordano Ferrari, H. (ed.) *la voix dans la dramaturgie contemporaine Grecque*, Gallimard, Paris, 2008.

[5] Lohmann, J. *Mousiké et Logos.* Ed.T.E.R, Paris, 1989.

[6] Loraux, N. "La voix endeuillée. Un essai sur la tragédie", MIT Press, Cambridge, MA, 1999.

[7] Dissanayake, "Antecedents of the temporal arts in early mother-infant interaction", in: Wallin N., Merker B. & Brown S. (eds.), *The Origins of Music*. Cambridge, MA: MIT Press, 1992 pp. 389-410.

[8] Narmour, E. *The analysis and cognition of basicmelodic structures*, University of Chicago Press, Chicago, 1990.

[9] Boersma P. & Weeninck, D. "Praat, a system for doing phonetics by computer", T*echnical Report 132, Inst. of Phonetic Sc.*, University of Amsterdam, 1996.

[10] Easterling P & Hall E. (Eds.) "The art of the actor", *Greek and Roman Actors: Aspects of an Ancient Professionm*, in: www.cambridge.org Press, pp. 370-380.

[11] P. Price, Ostendorf, M., Shattuck-Hufnagel, S. & Fong, C. "The Use of Prosody in Syntactic Disambiguation", *J. of the Acoust. Society of America 90*, 6, 1991, pp. 2956-2970.

[12] Huang Z., Chen L., & Harper M. "Purdue Prosodic Feature Extraction Toolkit on Praat", *Zeitschrift für Phonetik 16*, Diesterweg, 2006, pp. 293-326.

[13] Devine, A. M. & Stephens, L. D. *The prosody of Greek speech,* Oxford University Press, New York, 1994.

[14] Xydas, G. Spiliotopoulos, D. & Kouroupetroglou, G. "Prosody Prediction from Linguistically Enriched Documents Based on a Machine Learning Algorithm", *Proceedings of the 6th International Conference of Greek Linguistics (6th ICGL)*, Rethymno, Greece, September 18-21 2003.

[15] Devine, A. M. & Stephens, L. D. *The prosody of Greek speech,* Oxford University Press, New York, 1994.

[16] Ladefoged, P. *A course in phonetics. Forth Worth:* Harcourt Brace, 1993.

[17] Pöhlmann, E. & West, M. L. Documents of Ancient Greek Music. Clarendon, Oxford Press, Oxford, 2001.

[18] Xydas, G. & Kouroupetroglou, G. "The DEMOSTHENES Speech Composer", *Proceedings of the 4th ISCA Tutorial and Workshop on Speech Synthesis*, SSW4, Perthshire, Scotland, September 2001, pp. 167-172

[19] Aristotle, Problems, Book 19, 6.

[20] Aeschylos Agamemnon, metrical translation by Wilhelm von Humboldt, bei Gerhard Fleischer dem Jüngeren, Leipzig 1816, introduction, p. 136