Distance measurements of manipulations of rhythmic patterns in expressive music performance and in speech prosody

Fernando Gualda
Sonic Arts Research Centre
Queen's University – UK

ABSTRACT

Based on the hypothesis that expression can be regarded as manipulations of prosodic and acoustic features of a given sentence (or musical score), this research suggests an unified approach to comparing utterances and musical performances.

Possibly, the most distinctive differences in music and in speech are, respectively, the quantization of pitch (notes) and of timbre (phonemes). Emphasis (stress) and rhythm are prosodic features that belong to both domains.

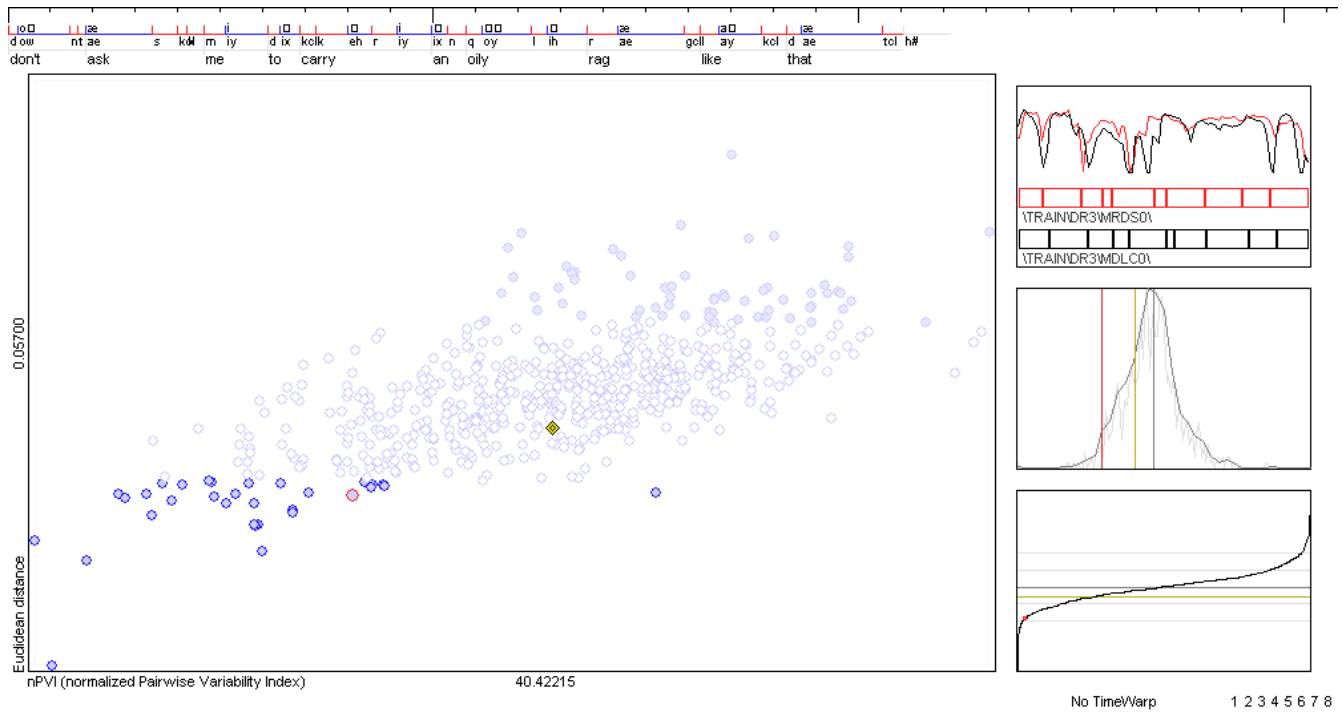
This research compares the normalized Pairwise Variability Index (nPVI), the *Inegáles* k-value of the KTH Rule System for Musical Performance, using Euclidean and Mahalanobis distances between rhythmic patterns of either notes or speech units extracted from audio recordings in the following studies:

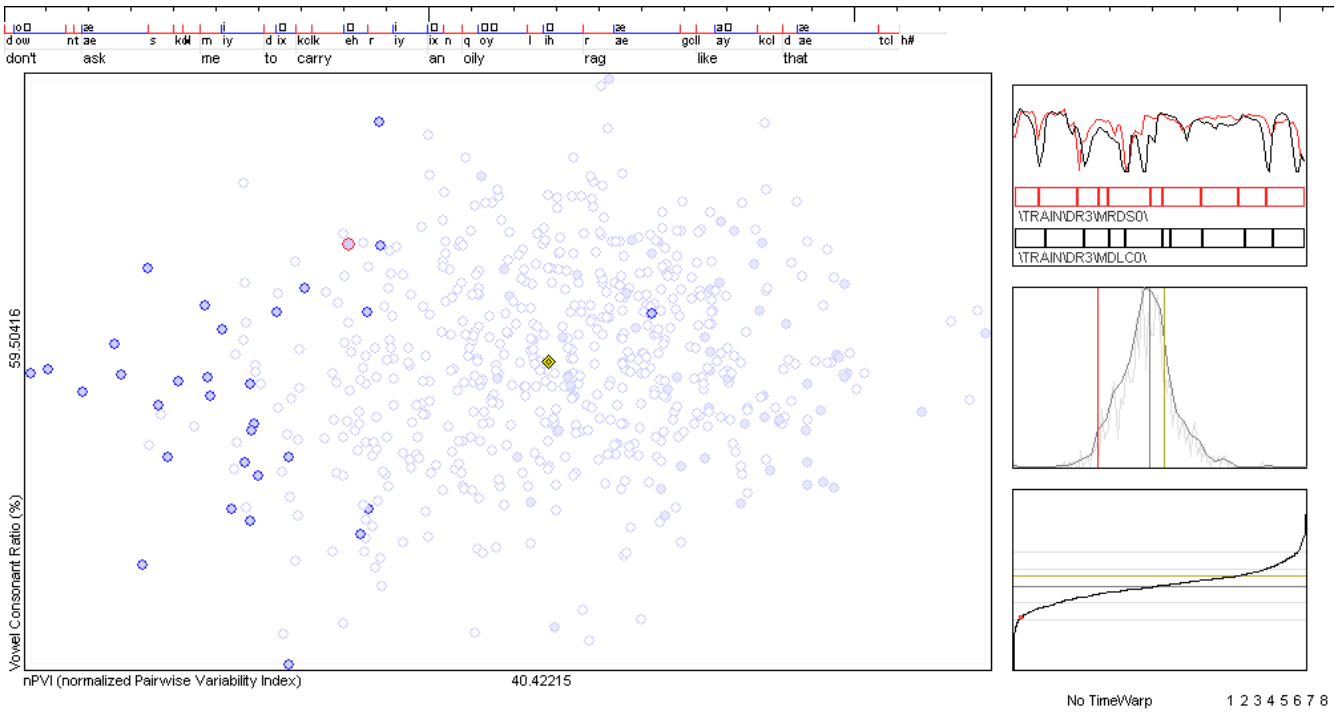
1) TIMIT – 628 utterances of the same sentence

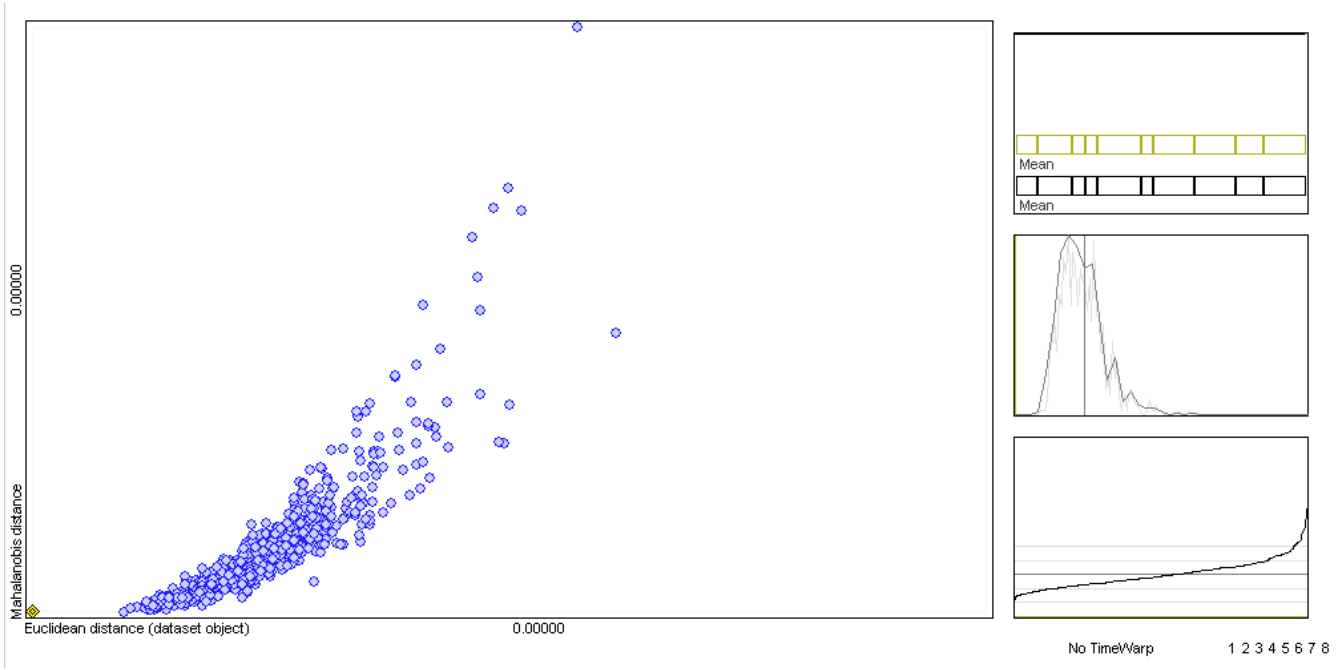
2) Niobe – 40 recordings of performances of the same musical phrase

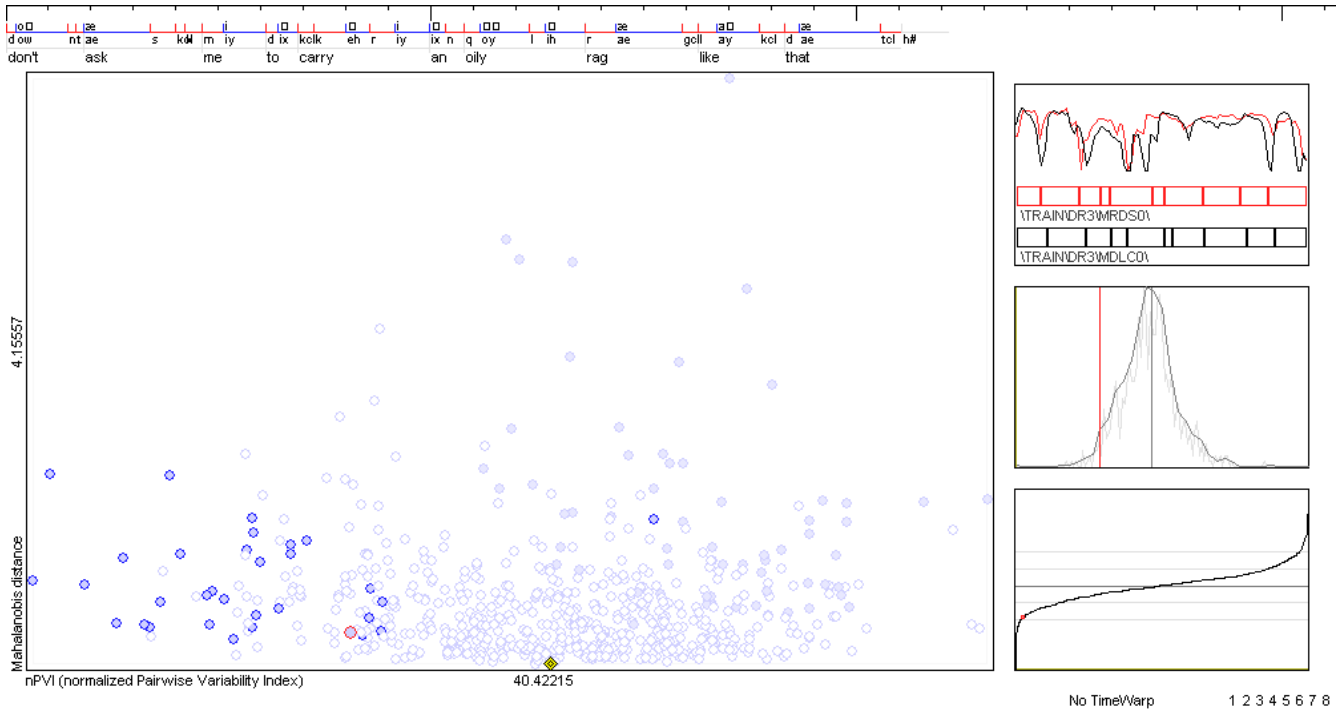
Conceptually, the normalized Pairwise Variability Index (nPVI) and the *Inegáles* k-value of the KTH Rule System for Musical Performance are correlated. The nPVI (Low, Grabe, and Nolan 2000) has been used to classify both utterances and musical themes (Patel 2008). In the *Inegáles* rule (Friberg et al. 2006) longer-shorter patterns (swing) are introduced in the rendition of musical scores containing notes with same durations.

The implementation of those methods led to the development of new software application, which can be demonstrated during the EMUS conference.


don't ask me to carry an oily rag like that
0.05700
Euclidean distance
nPVI (normalized Pairwise Variability Index)
40.42215
\TRAIN\DR3\MRDS0\
\TRAIN\DR3\MDLC0\
No TimeWarp
1 2 3 4 5 6 7 8


don't ask me to carry an oily rag like that
59.50416
Vowel Consonant Ratio (%)
nPVI (normalized Pairwise Variability Index)
40.42215
\TRAIN\DR3\MRDS0\
\TRAIN\DR3\MDLC0\
No TimeWarp
1 2 3 4 5 6 7 8


0.00000
Mahalanobis distance
Euclidean distance (dataset object)
0.00000
Mean
Mean
No TimeWarp
1 2 3 4 5 6 7 8


d ow nt ae s kd m iy d ix kcl k eh r iy ix n q oy l ih r ae gcl ay kcl d ae tcl h#
don't ask me to carry an oily rag like that
4.15557
Mahalanobis distance
nPVI (normalized Pairwise Variability Index)
40.42215
\TRAIN\DR3\MRDS0\
\TRAIN\DR3\MDLC0\
No TimeWarp
1 2 3 4 5 6 7 8