Report on Musical Prosody

Kamlesh Lakshminarayanan. EE 675-Prof. Elaine Chew

It would be difficult to say whether music or speech came first. We appreciate music because of its correlation with speech at various levels. The paper by Caroline Palmer and Sean Hutchins on musical prosody discusses its forms and functions while drawing parallels from relevant speech literature. A performance is the effort of a composer who defines the score, its pitch and duration and of a performer who manipulates the pitch, amplitude and timbre. These variations define musical prosody. What is expressed in these manipulations is not definitive. Therefore we study how it is expressed. Music prosody is performer driven and obligatory. This implies that every musical performance focuses on a set of variables that define how we perceive the music. Performers tend to stick to the same tempo and dynamics while repeating a performance. They show a pattern in choosing prosody. Also, several elements of expressive performance like aesthetic, historic and contextual elements go into the selection of prosodic variables. The paper discusses whether all these factors are involved in helping us define general principles of musical expression.

To understand musical grammar, we consider linguistic prosody. The term meter refers to a specific rhythm determined by the number of beats and the time value assigned to each note in a measure of a verse or a musical piece. Phonological theory lays special emphasis on prosody and a section of it deals with metrical structure. Speech prosody distinguishes temporal based properties like rhythm from frequency based properties like intonation whose purpose is to indicate presence of emotional or pragmatic information. Temporal aspects of music and speech prosody share several similarities and differences. Example, there is a correlation between musical meter and linguistic stress but music and speech differ in beat regularity.

Musical prosody serves several functions like segmentation, prominence, coordination and emotional response. A phrase is generally segmented by prosodic cues which define boundaries by changes in acoustic variables. Phrase-final lengthening is an example that describes this concept. Segmentation also helps distinguishing musical voices. Thus, although prosody is not a necessary condition for segmentation it is a sufficient condition to influence the listener and indicate phrase structure. Prosody also indicates prominence of events. It is used to emphasize musical structure to a piece where melody, meter and phrasing are not completely defined in the score. It is done by lengthening, articulating or playing important beats relatively louder. Exaggeration and use of vibrato are other means of rendering prominence. Structural expectation also leads to systematic prosody. Collaborative music requires coordination and onset timings define synchronization between performers. Cohesion is achieved through a conductor or one of the performers play a lead role. Tempo plays an important role in a collaborative effort unlike speech where tempo is based on the individuals preferred rate. Music is also known to evoke a strong emotional response. Unlike in speech, music prosody does not necessarily convey the performers' emotional state. The fundamental frequency of a speaker indicates his/her emotion. Tone variations, articulation, tempo, timing and vibrato are some of the acoustic properties which when varied can be broadly used to classify the five primary emotions of sadness, happiness, anger, fear and surprise. Thus, prosody is one of the variables that help us define cohesion and emotion in music.

Several models relate the structure of music and its prosody. Preference rules define variation in the models output because the some degree of flexibility is exercised by the performer. Listeners prefer such model output over one generated by a computer. Inclusion of the composers' pulse into the structure was found to be more acceptable than otherwise. Models defining the tempo of the musical piece towards the phrase endings and at major structural breaks have also been proposed. Musical prosody can also be modeled with kinematic laws of physical motion. Example, timing characteristics of ensemble performances can be modeled by a cubic polynomial which connected portions of varying tempo in order to minimize drastic changes in acceleration. Findings show that tempo changes are often coupled with changes in loudness. So a model that connected the two was proposed.

Music prosody aids the listener in several ways. Acoustic properties that give structure to music helps the listener to perceive small units of the music and thereby forming a foundation for understanding higher-order relationships. A better appreciation of the structure creates expectations in the mind of the listener which in turn helps him in grasping the emotional content of the piece. Intonation and prosody are commonly used across cultures in communicating with infants through speech and music. Also, prosodic variation in pitch peaks, range and height help infants in better understanding. This means that greater exposure to music at early stages of child growth could provide a novel and effective learning environment for infants.

Stress-timed and syllable-timed language speaking composers lay different emphasis on amount of durational variation and inter-consonantal intervals. Studies reveal that the two domains share common neural mechanisms. Also, musically trained listeners are more receptive to fundamental frequency and emotional changes in speech. Future research could lead to higher correlation between the two domains.