The perception of music, according to Longuet-Higgins, involves the “interplay” of physical sounds with a listener’s knowledge of musical conventions. In a linguistic sense this is an interaction between sonic events and the grammatical context of those events. Longuet-Higgins describes the grammatical context of Western music as grounded in two fundamental “frames of reference,” and he creates conceptual structures in order to represent them visually. The first is a tonal frame, comprised of major and minor scale degrees built from the same tonic; and the second is a rhythmic frame, comprised of tempo and meter. The visual representations Longuet-Higgins constructs are designed to aid our understanding of how we interpret sonic events, and thereby cognitively transform them into what we recognize as music.

As a means of demonstration Longuet-Higgins developed a program that transcribes melodies into a script analogous to musical notation. The program begins with the assumption that the first note is the tonic, and that the first two notes establish the tempo. It then makes corrections to these frames as “distortions” occur. The limitations of the program are not only technological, but also result from of the narrow parameters in the “frames of references.” The program analyzes the tonal frame first and then proceeds to the rhythmic frame, and the implication that these can be treated as independent processes is an evident concern to Longuet-Higgins. Furthermore, the program is limited to monophonic melodies that are strictly composed as such. In other word, parts form a polyphonic melody cannot be analyzed because many of the important events delineating the tonality and rhythm are not found in each part. Finally, although the program is able to capture rhythmic articulations such as staccato, it is not capable of accounting for rhythmic ornaments (trills, mordents, etc.) and treats them instead as singular events.

The major contributions of the Longuet-Higgins’ lecture are concepts of “tolerance” in the rhythmic frame and “sharpness-remoteness” in the tonal frame. Both concepts serve
as criteria for how we interpret musical events. Tolerance is a temporal constant (typically around a tenth of a second). In order for a listener to perceive an event as a distinct rhythmic entity, its onset must differ from the metrical beat where it occurs by more than the Tolerance. This allows the listener to separate out rhythmic events into a tiered structure, and to predict the rhythmic development. The Sharpness-Remoteness scale is equally critical for the perception of tonality. Sharpness and Remoteness are measured separately along a linear path of notes in harmonic space, which move away from the center by an interval of a P5\(^{th}\). Sharpness is closely related to key (e.g., if ‘c’ = 0: ‘g’ has 1 sharp and a Sharpness of 1; ‘f’ has 1 flat and Sharpness of –1, etc.). Remoteness is the difference of sharpness, or the number of P5\(^{th}\)s, between any 2 notes (e.g., in the case of ‘f’ to ‘g’, the Remoteness = 2). Typically, intervals with Remoteness less than 6 are diatonic intervals, and those greater than 6 are chromatic.

Question: L-H claims that syncopation is more “confusing perceptually.” How can this be verified? Isn’t this notion simply based on the assumption that we actually “hear” what is being visually represented?