Cognitive Models of Tonal Tension

Eileen Wagner[*]
eileen.wagner@student.uva.nl

April 25, 2014

Tension plays a central role in music perception: the anticipation of relaxation or release and the ensuing resolution creates the ebb and flow of a listening experience. It can be seen as part of musical expectancy (Schmuckler, 1989; Huron, 2006) and has been linked to musical affect (Krumhansl, 1997, 1996; Koelsch, 2012; Rohrmeier & Koelsch, 2012). It may also influence aspects of musical performance (Palmer, 1996). Despite its emergent nature, tension has been judged consistently by participants (Krumhansl, 1996), even taking into account familiarity with the piece (Fredrickson, 1999), affect (Nielsen, 1987), and personal preference (Lychner, 1998). There are a variety of factors that are said to determine tension in music, for example dynamics, timbre, melodic contour, harmony, tonality, repetition (Nielsen, 1983), phrase structure, note density (Krumhansl, 1996), pitch height, loudness, onset frequency, and tempo (Farbood, 2012). This paper focuses on cognitive models for *tonal tension*: the anticipation of a harmonic resolution following a sequence of tonal music. The plan is as follows: I give an overview of the proposed models for tonal tension in (I). Then I discuss the main theoretical issues we encounter in research on tonal tension (II). In (III), I propose an experiment that could settle one particular problem, namely the influence of exposure in the perception of tonal tension.

1 Current Models of Tonal Tension

A typical case of tonal tension can be found in leading chords, e.g. in the V7-I relation. In music-theoretical terms, the dominant seventh chord creates tension due to the tritone relation between the leading tone and the seventh of the chord. This tension needs to be resolved to the root and major third of the tonic chord. However, alternative explanations have also been proposed. Bigand et al. (1996) and Bigand & Parncutt (1999) present three accounts of tonal tension: a cognitive approach which assumes some knowledge of the tonal hierarchy; sensory models based on psychoacoustic research; and a model of horizontal movement that considers melodic arrangement.

---

[*] This paper was submitted for the course ‘Cognitive Models of Language and Music,’ taught in spring 2014 at the ILLC. I am grateful to Aline Honingh for very helpful comments.

Cognitive models of tonal tension assume, to various extents, some knowledge of the tonal hierarchy. That is, listeners have internalised chords and keys in such a way that the perceived tonal tension can be related to tonal function, where important events in a tonal hierarchy instil weak or null musical tension, and less important ones create strong musical tension. 'Importance' here can be specified using a *within-key hierarchy* (for example, the tonic, dominant, and sub-dominant chords are more important) and *between-key distances* (keys sharing a greater number of scale notes, such as C major and A minor, are more important). Based on these findings, Lerdahl & Jackendoff (1983) constructed a generative theory of tonal music (*GTTM*) which offers a hierarchical analysis of tonal tension based on musical intuition. Explicit rules are used to generate structures that listeners unconsciously infer from the musical surface of a piece. There are four components to *GTTM*: *grouping structure*, which segments music into motives, phrases, and sections; *metrical structure*, a hierarchy of alternating strong and weak beats; *time-span reduction*, a hierarchy of structural importance of pitches with respect to their position in the grouping and metrical structures; and *prolongational reduction*, a hierarchy that expresses harmonic and melodic tension and relaxation (cf. Schenker (1935/1979)). An example of notation that expresses different degrees of tension can be found in 1.

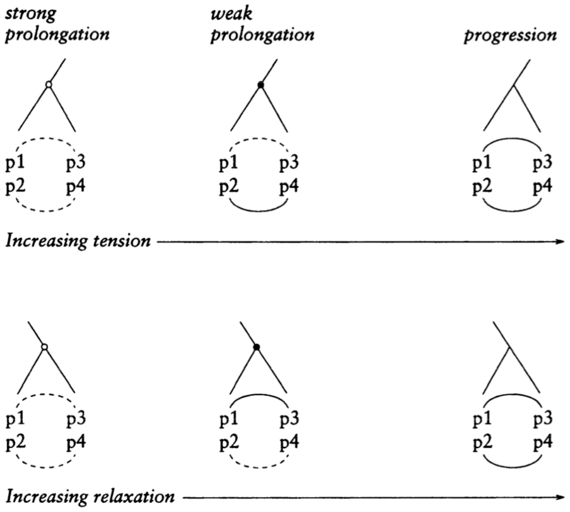

Figure 1: example of *GTTM* prolongational branching from Lerdahl (1996, p. 320)

A different, more formal model is the tonal pitch space model (*TPS*), due to (Lerdahl, 2001; Lerdahl & Krumhansl, 2007). *TPS* maps common pitches, chords, and keys onto a spatial representation to calculate the distances between them. The mapping itself relies on music-theoretic principles (e.g. using the circle of fifths) consistent with empirical findings from varying musical backgrounds and different cultures (Krumhansl, 1990; Krumhansl & Kessler, 1982). Specifically, tonal tension is defined by a formula that computes quantitative predictions of tension and attraction for events in any passage of tonal music. Four components enter into the calculation: (1) a representation of prolongational event structure (in this case the tree notation from *GTTM*), (2) a model of tonal pitch space and all distances within it from *TPS*, (3) a treatment of surface dissonance, and (4) a model of voice-leading attractions (Lerdahl & Krumhansl, 2007, p. 330).

TPS records three values, the first two representing within-key hierarchies, and the third one between-key distances. The diatonic chord distance between chords x and y is defined as

$$\delta(x, y) = i + j + k$$

where i is the number of steps between two regions on the chromatic fifths circle (i.e., distance between two chords with regard to key), j is the number of steps between two chords on the diatonic fifths circle (distance with regard to chord function), and k is the number of distinctive pitch classes in the basic space of y compared to those in the basic space of x. The basic space for a chord consists of its pitch classes at the chromatic, diatonic, triadic, fifths, and root (2).

(a) octave (root) level:	0												(0)
(b) fifths level:	0							7					(0)
(c) triadic level:	0				4			7					(0)
(d) diatonic level:	0		2		4	5		7		9		11	(0)
(e) chromatic level:	0	1	2	3	4	5	6	7	8	9	10	11	(0)

Figure 2: diatonic basic space, taken from Lerdahl & Krumhansl (2007, p. 331)

TPS assumes that listeners will (i) form the shortest pitch-space path, (ii) use the *principle of good form* (preferring optimal patterns of tension and relaxation), and (iii) look for parallelisms as a gestalt principle. As such, *TPS* can compute tension both sequentially and hierarchically (by employing a tree-analysis of sequences), and thus generate quantitative predictions of tension for any sequence of tones.

Surface dissonance is taken to be largely psychoacoustic, which leads us to the **sensory models** of tonal tension. Surface dissonance can be measured by the calculated *auditory roughness* of individual chords (based on Helmholtz (1877, ch. 10)). Roughness is caused by the perception of rapid amplitude fluctuations in the range of 20–200 Hz. Simultaneous tones close in frequency can cause rough beats—e.g., a minor second played in the medium register of the piano (Pressnitzer et al., 2000, p. 67). It is hypothesised that dissonant (nonharmonic) tones are less stable, and therefore create local tension.

Another psychoacoustic measure is *pitch commonality* between successive chords, determined by (i) the number of tones they have in common, (ii) how close their roots are to each other on the cycle of fifths, and (iii) whether their notes all belong to the same major or minor scale (Bigand & Parncutt, 1999, p. 240). Stability can also be affected by the *voicing*: a triad is more stable if it is in root position, and even more so if the melodic note is on the root. Pitch commonality and voicing can be expressed with a *surface tension rule*, for example in Lerdahl & Krumhansl (2007, p. 334):

$$T_{diss} = f + g + h$$

where f is the chord voicing (1 if the melody is not the chord root, 0 otherwise), g is the inversion (2 if the chord is not in root position, 0 if it is), and h is the sum of all nonharmonic tones (sevenths = 1, diatonic nonharmonic tones = 3, and chromatic nonharmonic tones = 4). (For the sake of simplicity, it is assumed that surface tension is perceived categorically.) One of the caveats of including psychoacoustic considerations is that the perceived pitches of a chord do not necessarily correspond to notated pitches. For instance, an E-flat major triad may weakly imply the pitch C, even though that pitch is not notated, and may not even be physically present (Bigand et al., 1996, p. 130). Implied pitches can therefore confound the analysis.

Horizontal movement has often been neglected in analyses of tonal tension. It suggests that smoother progressions occur when there are small intervals between each voice of successive chords, and that horizontal motion affects the consonance of the chords (Bigand et al., 1996, p. 131). The theoretical assumption here is called *melodic anchoring*: the urge for a less stable pitch to resolve on a subsequent, proximate, and more stable pitch. For example, melodies that contained an unstable tone were given higher ratings of tension when the unstable tone was anchored than when it was not (Bharucha, 1984, 1996). In terms of *TPS*, the smaller the pitch distance between an unstable event and a stable one, the better the resolution of the tension created by the unstable event. The measure for the stability is called *attraction*, and we can also posit a *harmonic attraction rule*, which sums up all the values between the respective voices and divides it by the value obtained from the chord distance. Based on this rule, we can predict that 'the strongest harmonic attraction is from a dominant seventh chord to its tonic, because of the powerful attractions of the leading tone to the tonic and the fourth to the third scale degree and because of the short distance from the dominant to the tonic chord. This is why (aside from statistical frequency) the expectancy for a tonic chord is so high after a dominant-seventh chord (Lerdahl & Krumhansl, 2007, p. 337).

There are only a handful of experiments investigating the relationship between these models. Of them, Bigand et al. (1996) conclude that perceived tension was influenced by the tonal hierarchy (the more important the chord in the hierarchy, the weaker the tension), chordal consonance (minor chords were more tense than major, seventh more tense than triad), and by horizontal motion (the greater the melodic distance traversed by voices, the higher the tension). Lerdahl & Krumhansl (2007) tested a more fine-grained and comprehensive model based on all four components (see the example below). In general, there was a good overlap between the predictions and the data. Their results indicate that listeners hear tension hierarchically more than sequentially. Moreover, attractions need to be incorporated in the model.

2 Theoretical Issues in Modelling Tonal Tension

There are many theoretical issues that need to be disentangled in the discussion. They can be grouped as follows:

1. **Concept of tension**. What amounts to musical tension? What are the theoretical assumptions of the proposed models?

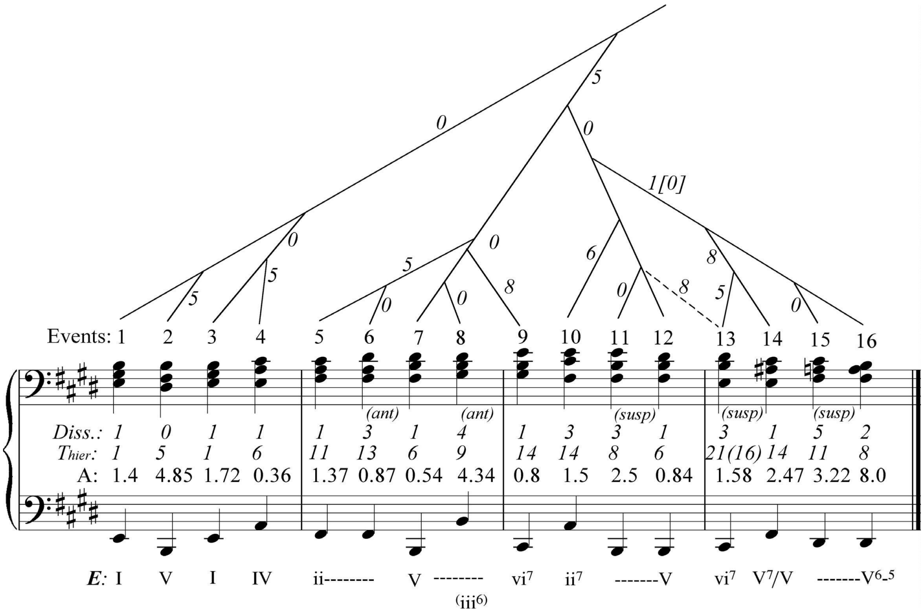

Figure 3: *TPS* analysis of first phrase of Chopin's E major Prelude, taken from Lerdahl & Krumhansl (2007, p. 346)

2. **Global vs. local.** Is musical tension perceived globally or locally? Does harmonic prediction depend on short sequences only, or does tension rise with more complicated hierarchical structures?

3. **Context-sensitivity.** What is the interaction between harmony and other features, especially melody?

4. **Domain-specificity.** Is our ability to perceive tonal tension universal? Does exposure to certain styles of music influence our tension perception?

5. **Relationship to language.** Is there an analogy to be drawn between cognitive models of tonal tension and cognitive models of language?

2.1 The Concept of Tension

A major problem with the existing theoretical models and experimental paradigms is the disagreement on what amounts to tonal tension. The psychoacoustic and music-theoretical accounts seem to pertain to two different things: dissonance which causes a sense of 'uneasiness', and harmonic or melodic expectation which induces 'incompleteness' shortly before the resolution.

The psychoacoustic account relates tension to the basic psychological dimensions of arousal and activity. An increase in intensity within psychoacoustic parameters such as roughness or pitch contour corresponds to an increase in tension. Thus, tonal tension is not a uniquely musical phenomenon (Pressnitzer et al., 2000), and moreover is not necessarily related to expectation. Research has shown that other factors contributing to musical tension may well have an evolutionary benefit, such as 'the role of increasing, "looming" loudness change as a basic biological warning signal, and the use of low pitch register in expressing dominance and aggression' (Granot & Eitan, 2011, p. 221). Furthermore, pitch register may be related to expressing dominance, aggression, and threat. These acoustic sounds can be interpreted as 'primeval signals of impending danger, automatically activating a multimodal brain network aimed at immediately dealing with the incoming sound source' (ibid.).

On the other hand, the music-theoretical account relates tension to expectation. According to Huron (2006, p. 328), musical tension is 'anticipation followed by a positively valenced prediction response' and forms part of musical expectancy in general. In other words, tension rises as expectations rise, and resolves when expectations are realised. Bigand et al. (1996, p. 139) regard the two accounts as rival theories: 'we might possibly consider the cognitive theory of tonal hierarchies as a by-product, or an over-theorization, of more elementary psychoacoustical phenomena such as roughness, pitch commonality, or melodic arrangement between chords'. Their results indicate that cognitive models outperform sensory ones, while Pressnitzer et al. (2000) argue that sensory dissonance has a significant effect on experienced tension: for single chords, subjective roughness ratings correlate with tension ratings. The fundamental question here is whether the perception of tonal tension is a top-down, hierarchical structure or a bottom-up, sensory process. In fact, ERP responses to consonant chords within an irregular tonal context suggest that there is independent processing of sensory consonance and harmonic function (Regnault et al., 2001).

However, even within the framework of musical expectancy, it is unclear what causes tonal tension. Generally, musical events defined as instable, incomplete, or open in relation to a tonal hierarchy are assumed to arouse tension, whereas stable, complete, and closed events are perceived as points of relaxation (cf. Dibben, 1999, p. 270). For example, Melo & Wiggins (2003, p. 1) state that 'the degree of tension can be related to how unfinished the piece of music would sound if it stopped at the point'. But tension may also arise due to 'suspense', i.e. the inability to generate any clear expectations.

These two explanations can be linked to Huron's distinction between two kinds of tension response: uncertainty as to *when* the expected event will happen, or *what* the event will be (Huron, 2006, pp. 9). In the first case, listeners anticipate a particular resolution according to their knowledge of 'good' harmonic-melodic cadences. There is little uncertainty as to which note or chord they expect—it is merely a matter of timing. In the latter case, listeners sense a tonal uncertainty, either because several alternatives are all equally probable, or because the current context is too irregular (Granot & Donchin, 2002). This leads to the expectation of *some* event that resolves or at least disambiguates the given sequence, but there is little uncertainty about the timing (the event will ensue). Surprisingly, this issue has been widely neglected by the current experimental paradigms.

The confusion about what is meant by tonal tension is also reflected in the methodology of the experimental paradigms. Two particular problems occurred: first, the concept of tension was not well-explained and so the measurements were likely confounded; second, the procedures themselves may have constrained the tension ratings.

The experiments on musical tension are often not comparable due to the fact that participants were given very different characterisations of (tonal) tension. Lerdahl & Krumhansl (2007, p. 330) write that 'everyone experiences physical tension and relaxation, and it is common to extend the terms to mental and emotional terrains as well. Consequently, it is relatively straightforward to ask experimental participants to respond to degrees of tension and relaxation and thereby elicit consistent interpersonal responses'. Melo & Wiggins (2003), on the other hand, explicitly used 'completeness' to describe tension as 'the uneasy—as opposed to relaxed—sound of the music, and as how "unfinished" the piece would sound if it stopped at that precise moment'. Similarly, in Bigand & Parncutt (1999, p. 242), 'participants listened to musical examples exhibiting different degrees of musical tension. They were told that strong musical tension at the end of a fragment evokes the feeling that there must be a continuation of the sequence. Low musical tension evokes the feeling that the sequence could naturally stop at this point.' In yet another setting, Lehne et al. (2013, p. 173) asked participants to 'indicate the tension of the music as they subjectively experienced it (participants were explicitly instructed not to indicate the amount of tension they thought the music was supposed to express). That is, ratings of *felt* musical tension (in contrast to *perceived* tension, cf. Gabrielsson (2002)) were acquired.' The reference here suggests that the experiment was testing general emotional reactions rather than the specific 'completeness' of a phrase. Lastly, participants 'were allowed to define tension and release according to any criteria that seemed relevant' in Pressnitzer et al. (2000). It seems that many results are confounded by what the experimenters specifically asked the participants to rate.

The measures for tension may also have been influenced by data collection. Lerdahl & Krumhansl (2007, p. 339) proposes two methods. In a *stop-tension task*, the first event was sounded, at which point the participants rated its degree of tension; then the first and second events were sounded and the participants rated the tension of the second event, and so on. In a *continuous-tension task*, the participants move some form of a slider in correspondence with their ongoing experience of increasing and decreasing tension. The advantage of the stop-tension task is that it records the response precisely for the event that is evaluated. However, it is rather artificial and time-consuming for long excerpts. The advantage of the continuous-tension task is that it encourages a spontaneous response to intuitions of tension in real time.

Most experiments used a physical slider called a Continuous Response Digital Interface (CRDI), which consists of a potentiometer mounted in a dial interfaced with a computer. Lerdahl & Krumhansl (2007) and Lehne et al. (2013) used a slider that was shown on a computer screen and could be moved with the mouse. In Melo & Wiggins (2003), a wheel with a spring provided feedback on the position, so the listeners have to apply more force to indicate higher tension, or no force for no tension. Bigand et al. (1996) and Bigand & Parncutt (1999) had participant rate the tension on a discrete 12-point (and 10-point) scale in a stop-tension task. For Pressnitzer et al. (2000), participants had to make forced-choice judgments to the question 'Between the two sounds of this pair, do you hear an evolution from “tension to release” or from “release to tension”?’ Granot & Eitan (2011) asked listeners to give discrete overall tension ratings for short melodic sequences and select text responses that best described how tension was changing during the course of the sequence. The web-based study by Farbood (2012) designed graphical shapes to depict tension changes, and despite the efforts of the experimenters, there was a high correlation between tension shape and pitch contour. To sum up: there is a variety of paradigms for measuring (tonal) tension, and the manner in which the data is collected already reflects some underlying assumptions about tonal tension (e.g. that it is discrete, that it is qualitative rather than quantitative, or that it can be linked to physical force).

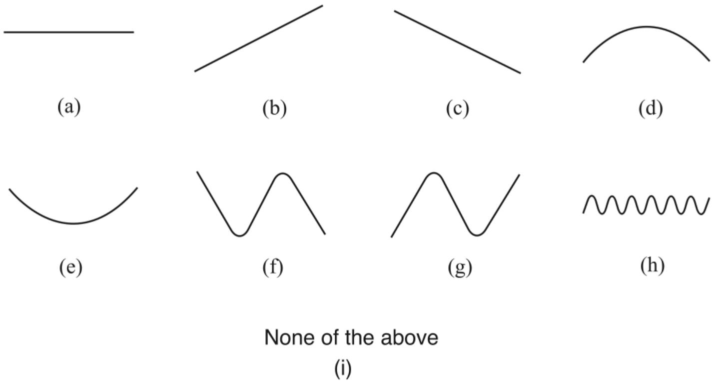

Figure 4: tension shapes used in the web-study by Farbood (2012)

2.2 Gobal vs. Local

Research in tonal tension has mainly focused on the global vs. local debate, i.e. to what extent hierarchical structures contribute to tonal tension. The intuition here is that ‘chords are rarely compared in isolation and the relations between the global context—the key—of a piece and the relations to the local context play a very important role in the perception of tonal harmony’ (De Haas et al., 2010, p. 2).

Bigand & Parncutt (1999) compared hierarchical (global) and sequential (local) listening by constructing two models, respectively, where the hierarchical one uses *inherited values* from *TPS* to calculate the added tension from embedded phrases (see 5). Their results indicate that the hierarchical analysis did not contribute towards tension ratings, and conclude that cadences in a local context play a bigger role than the global hierarchy in which they are embedded. Lerdahl & Krumhansl (2007), on the other hand, found the best correlations when hierarchical structures were incorporated in their post-hoc analysis. According to Lerdahl & Krumhansl (2007, p. 357), the reason for this discrepancy is the absence of surface dissonance and attraction in their model. Of course, the problem with only looking at long sequences is that it tends to test domain-general skills, and so the relation to tonal tension is at times contrived.

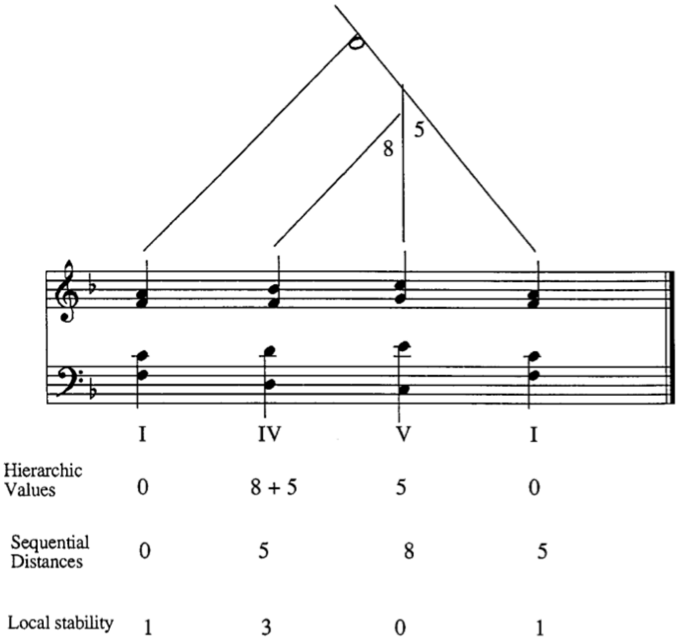

Figure 5: tension shapes used in the web-study by Farbood (2012)

A third approach, suggested by Lehne et al. (2013, p. 179), is that tension perception is influenced both by local harmonic implications and resolutions as well as global overarching syntactic features. For example, local dominants did not affect the tension ratings unless they reflected a deeper structure; on the other hand, the overarching tension increase is often mediated by smaller local tension-resolution patterns between pairs of chords.

2.3 Context-sensitivity

Another issue related to the previous discussion is context-sensitivity. We know that a chord can be very dissonant while having a stable tonal function. For instance, a minor chord with a major seventh that ends some jazz pieces is perceived as stable even though the chord itself is dissonant. Even roughness can be context-sensitive: in blues, ‘not hitting the note’ is often perceived as pleasant. *TPS* cannot account for these features, even when combined with the psychoacoustic models.

Moreover, the influence of other factors are worth examining. For example, both Krumhansl (1996) and Lerdahl (1996, pp) reported a high correspondence between ratings of tonal tension of versions with and without expressive features (i.e., dynamics and agogics). Lerdahl & Krumhansl (2007, p. 349) notes that listeners gravitate toward structures that are more easily processed. ‘When faced with the simple melody and convoluted harmonies of the Chopin, they apparently give the former greater weight than they would under more usual circumstances.’ This contradicts Schmuckler (1989), who thinks that melody and harmony are additive to expectancy. Granot & Eitan (2011) observe that when several parameters intensify concurrently, only the stronger parameter affects the overall tension rating, making activities in the other parameters redundant. Hence, an interactive model for musical tension in general might be required.

One particularly problematic interaction is that between melody and harmony. To begin, a single melody can have an implied harmony, and so confound the harmonic structure (though this usually isn't the case in Western tonal music). More importantly, it may attract attention that is otherwise required for harmony perception. Williams et al. (2011) had music majors listen to Mozart's K 265. When they attended more to harmony, their perception of tonal tension appeared to increase; when they attended more to the melody, their perception of tension appeared to decrease.

The effect of horizontal movement is also intriguing. A study on atonal music has shown that patterns with large descending intervals in the two upper voices are judged to be more stable than larges ascending intervals, whereas for the two lower voices, large ascending intervals are more stable than large descending intervals (Dibben, 1999, p. 287). Schmuckler (1989) already found that small intervals (three semitones or less) were more likely to generate tension than large intervals. Despite many attempts, it is still unclear how we can best quantify attraction (Lerdahl & Krumhansl, 2007, p. 339). It may be worth investigating a hybrid model between tonal tension and melodic expectation, for example that of Margulis (2005).

2.4 Domain-specificity

Very little has been said about the background theory so far: why do we perceive tonal tension, and is it determined by music theory, evolutionary function, or exposure? Already Bigand et al. (1996) note a high degree of agreement between participants, which they take to indicate that the concept of tonal tension is coherent and uniquely identifiable. Moreover, Hackworth & Fredrickson (2012) showed that children perceive musical tension similarly compared to college music majors. Granot & Eitan (2011, p. 220) concludes that the 'consistency with which participants varying in age and music training rate tension within specific musical pieces, regardless of style, … suggest that they possess a clear internal notion of musical tension, which does not rely on explicit explanations or definitions.' The already mentioned evolutionary benefits seem to support this hypothesis.

Of course, more needs to be said if the perception of tonal tension is to be an innate ability. Firstly, it is well-known that children develop an understanding for keys and harmonies between the age of 4 and 7. Secondly, more contemporary theories hold that the perceived consonance of intervals is determined not only by sensory and acoustic factors, but also by musical training, personal preference, and enculturation (Cazden, 1945; Parncutt, 1989). Thirdly, very little work has been done on the relation between the perception of tonal tension and culture. Vassilakis (2008) argues that musical tension is a culture-specific concept, guided by the equally culture-specific musical cues used to organise and recognise them. Tension ratings for Lebanese *mijwiz* playing and Bosnian *ganga* singing show that while roughness is closely related to the non-Western musicians' concept of tension, it is only one of many factors in Western listeners' judgment of musical tension.

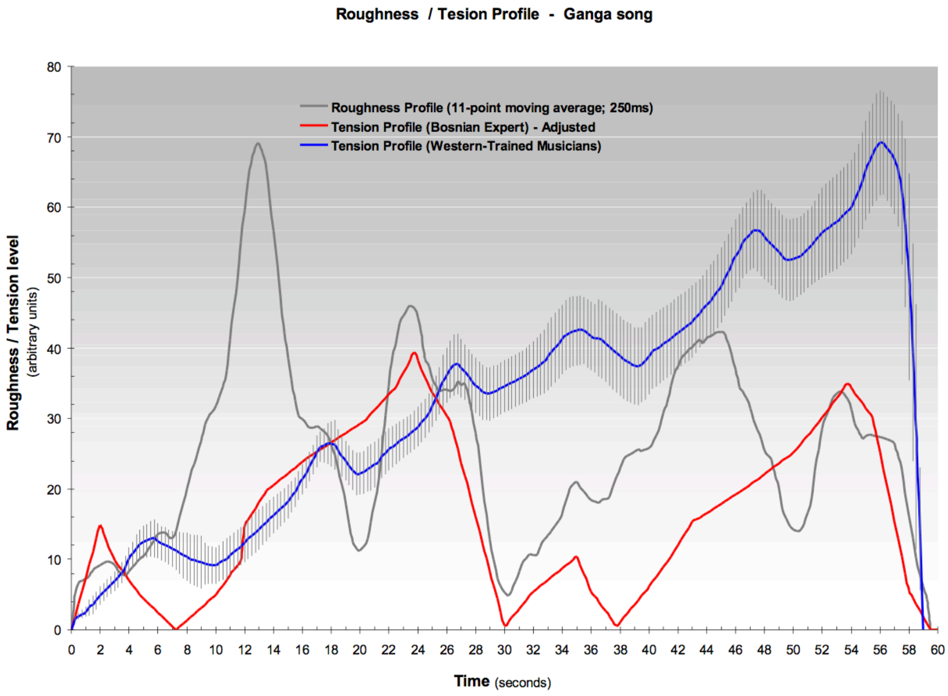

Figure 6: tension curves of ganga performance from Vassilakis (2008, p. 38)

Farbood (2012, p. 390) also notes that most studies have used excerpts from the Western classical repertoire ranging from the mid-Baroque era through the late twentieth century, exceptions being "St. Louis Blues" sung by Nat King Cole and Ella Fitzgerald (Fredrickson & Coggiola, 2003), "We are the Champions" by Queen, and "I Feel Good" by James Brown (Rozin et al., 2004). It is imaginable that different genres of music exhibit different tension patterns (e.g. film music).

Another approach to settle the 'nature vs. nurture' question is to look at the effect of musical training. A number of studies have included musicians and non-musicians. Overall, the data suggests that both groups have similar concepts of musical tension, but musicians tend to outperform non-musicians on tasks specifically related to tonal tension. For Parncutt & Bregman (2000), non-musicians seemed unable to respond to the exact pitch of a tone in a chord. This could be seen as evidence that pitch and chord perception is important for the sensitivity to tonal tension.

Bigand et al. (1996) observes that the predictions of a sensory (roughness) model correlate with tension ratings of musicians more than non-musicians. This seems to suggest that musicians are more trained in hearing pitch differences, and therefore more 'sensitive' to roughness. The results are intriguing, though: if exposure to tonal hierarchies is relevant for tension ratings, we would expect the results of musicians to fit cognitive models instead.

Granot & Eitan (2011, pp. 239) report that differences were not found in the general assessment of tension change. Furthermore, while dynamics influenced all participants, register influenced only the non-musicians, and pitch contour influenced only the musicians. Their hypothesis is that the 'normative' range of pitch for musicians was expanded due to their wider exposure. As to pitch contour, it appears that it only weakly contributes to tension ratings for non-musicians, e.g. rising sequences as more tense than falling sequences when those sequences were also accelerating, but not when sequences were decelerating. Bigand et al. (1996) found that horizontal motion affected non-musicians considerably more than musicians; this could indicate that musicians focus on the overall harmonic landscape rather than a single melody.

Granot & Donchin (2002) show a significant difference between musically trained and untrained people. Generally speaking, non-musicians respond slower and less accurately than musicians but exhibit similar overt behaviour. However, their ERP data implies that non-musicians generate expectancies regardless of the level of theoretical constraints in a tonal sequence, whereas musicians generate specific expectancies in strongly constraining sequences. This confirms the intuitive claims that cognitive models would provide a better fit for musicians. In fact, the study by Farbood (2012) indicates that musicians had greater sensitivity to harmony, tempo, onset frequency, and dynamics, all features they were testing for except for pitch height. It should also be noted that musical training is not always relevant for tension studies. Schmuckler (1989) found that expectancy responses varied depending on the culture of the listener (American, German, or Hungarian), but not on the basis of explicit musical training. Naturally, exposure does not have to come with musical training.

In short, evidence from past studies suggest that musical training and cultural exposure play a significant role in tension perception, despite the fact that overall tension is often rated similarly by people who have had different trainings and exposure. The innateness of tension perception, thus, should be questioned.

2.5 Relationship to Language

Lastly, it is worth noting that there are certain parallels between musical tension and language processing. In language, syntactic and semantic information can provide clues for a sense of completeness or temporal coherence. One could argue, for instance, that a sentence starting with 'if' creates tension that gets resolved by the word 'then'. There have been attempts to employ models from natural language parsing in music. For example, Steedman (1984) offered a generative grammar for chord progressions in twelve-bar blues.

Recently, Granroth-Wilding & Steedman (2014) proposed that hierarchical structures similar to those associated with prosody and syntax in language can be identified in the rhythmic and harmonic progressions that underlie Western tonal music. The project is not merely to decide whether a musical sequence is grammatical, but also which among a large number of analyses it has. One example for this is harmonic embedding. In Granroth-Wilding and Steedman's analysis, the cadence from *Call Me Irresponsible* has the same form as

$$((\mathrm{D}\sharp\mathrm{o}^7\ \mathrm{Em}^7\ \mathrm{Am}^7)\ \&\ ((\mathrm{E}^7\ \&\ (\mathrm{B}\phi^7\ \mathrm{E}^{7\flat 9}))\ \mathrm{A}^7))$$

Keats ((may or may not) cook) but (certainly eats) beets

Figure 7: Comparison of logical form, Granroth-Wilding & Steedman (2014)

It is an open question whether their approach using a Combinatory Categorical Grammar has a cognitive basis. But it is clearly possible to construct a simple yet powerful statistical parsing model, which can be tested in future studies.

3 Proposal: The Effect of Exposure

The purpose of the following experiment is to investigate the effect of exposure on ratings of tonal tension. We saw that both musical training and enculturation influence tension perception. Furthermore, the majority of the samples used so far are from the classical repertoire. In 2.3, I mention that context can play a role in tension perception: the same chord progressions can be perceived as stable or less stable in different genres. A simple experiment may be able to test the *exposure hypothesis*, which states that harmonic expectation is a result of mere exposure rather than musical training or explicit knowledge of the tonal hierarchy. Previous studies have shown that non-musicians can be as sensitive as musicians to subtle aspects of music harmony (Bigand & Poulin-Charronnat, 2006), and (Honing & Ladinig, 2009) demonstrated that timing judgment were better in the genre listeners were most exposed to. We therefore hypothesise that **tension ratings would be more accurate in the genre the participants were most exposed to**, i.e. the one they prefer. The underlying assumption here is that tonal tension forms part of musical expectancy in general Huron (2006), which is created by exposure.

The hypothesis that tonal tension is a result of exposure has important implications for the models we use. If exposure can be shown to affect tension perception in music, we should extend our current models, or even consider connectionist or memory-based ones. While *TPS* does consider empirical data for modelling the tonal hierarchy, a neural network approach has far more efficient statistical methods to encode and enforce cadential rules, and moreover added neurobiological plausibility. It is also well-known that different rules apply in other genres (e.g. tritone substitutions need to be accommodated). Computational models for tonal organisation have been proposed in the past (Bharucha, 1987; Tillmann et al., 2000) and have the potential to explain the learnability of the tonal hierarchy, even though Lerdahl & Krumhansl (2007, pp. 356) claims that these models ‘do not begin to address the fine distinctions in tension and relaxation that are elicited in the course of listening to music of any length or intricacy.’

3.1 Method

As mentioned above, there are far too many confounding factors in analysing longer pieces of music for tension studies. Both structural expectations and other musical features such as dynamics and pitch height may contribute towards tension perception. Hence we will use short, researcher-composed chord sequences from MIDI piano for the price of less ecologically valid stimuli. Crucially, the 3–5 chords in each sequence will be randomly stacked (‘voicing’) to control for horizontal movement (following Farbood, 2012). The sequences will consist of both cadences (harmonic progressions that establish musical keys) and modulations (harmonic progressions that precipitate key changes), and will be grouped into three genres: classical, rock, and jazz, based on the standard databases (e.g. The McGill Billboard database for rock). For example, a I-V-vi-IV progression will be classified as ‘rock’, whereas a iv7-bVII7-I (backdoor) progression will be classified as ‘jazz’. (In these databases, genres are usually recognised from instrumentation, rhythm, etc. Whether these chord progressions are in fact perceived as belonging to a particular genre should be tested in a prior experiment.)

This experiment will require at least 60 participants (20 for each genre) from online recruitment. They will be asked to indicate their musical preference in terms of age of onset, years of exposure, and frequency of exposure. The experiment can be conducted online with each participant rating up to 90 chord progressions (10 typical sequences for each genre and 3 randomised variations of each sequence). Ratings will be collected for each sequence from a slider that can be continuously moved with the mouse. Given the above set-up, one might argue that the experiment makes the unwarranted assumption that psychoacoustic effects remain similar throughout the groups; that is, dissonant chords (e.g. Imaj9) will contribute equally to tension ratings in all genres. We saw that this is not the case in specific contexts: some dissonant chords can be harmonically stable. Therefore, we will explicitly instruct participants to rate the stability of the chord rather than its dissonance. This can be made clear in an introductory example.

We predict that tension ratings would be higher and more precisely timed for each participant’s preferred genre, where the presumed locus of tension is in the penultimate chord in each progression. With a larger pool of participants and well-annotated databases, we can expand this study to non-Western music as well.

4 Conclusion

I have surveyed the current models for tonal tension and explored the main theoretical issues we encounter. The research proposal could be helpful in motivating a new approach to modelling tonal tension. The models presented here all have different underlying assumptions, methodologies, and empirical validities. The unclarity as to the concept of tonal tension, i.e. whether it arises from expectancy in general or music-theoretic and psychoacoustic factors, has caused some confusion in the literature. We further asked whether it is meaningful to examine tonal tension in isolation, seeing that it interacts with other musical features. This discussion can serve as a starting point for future research.

References

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, *16*(4), 485–518.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, 1–30.

Bharucha, J. J. (1996). Melodic anchoring. *Music Perception*, 383–400.

Bigand, E., & Parncutt, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research*, *62*(4), 237–254.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, *58*(1), 125–141.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? a review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*(1), 100–130.

Cazden, N. (1945). Musical consonance and dissonance: A cultural criterion. *The Journal of Aesthetics and Art Criticism*, *4*(1), 3–11.

De Haas, W. B., Robine, M., Hanna, P., & Veltkamp, R. C. (2010). Comparing harmonic similarity measures. In *Proceedings of the 7th ...*.

Dibben, N. (1999). The perception of structural stability in atonal music: The influence of salience, stability, horizontal motion, pitch commonality, and dissonance. *Music Perception*, *16*(3), 265–294.

Farbood, M. M. (2012). A parametric, temporal model of musical tension. *Music Perception*, *29*(4), 387–428.

Fredrickson, W. E. (1999). Effect of musical performance on perception of tension in gustav hoist's first suite in e-flat. *Journal of Research in Music Education*, *47*(1), 44–52.

Fredrickson, W. E., & Coggiola, J. C. (2003). A comparison of music majors' and nonmajors' perceptions of tension for two selections of jazz music. *Journal of Research in Music Education*, *51*(3), 259–270.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, *5*(1 suppl), 123–147.

Granot, R., & Donchin, E. (2002). Do Re Mi Fa Sol La Ti—Constraints, Congruity, and Musical Training: An Event-Related Brain Potentials Study of Musical Expectancies. *Music Perception*, *19*(4), 487–528.

Granot, R., & Eitan, Z. (2011). Musical Tension and the Interaction of Dynamic Auditory Parameters. *Music Perception: An Interdisciplinary Journal*, *28*(3), 219–246.

Granroth-Wilding, M., & Steedman, M. (2014). A robust parser-interpreter for jazz chord sequences. *Journal of New Music Research*, *43*(4), 355–374.

Hackworth, R. S., & Fredrickson, W. E. (2012). The effect of instruction on children's perceived musical tension in debussy's noël des enfants qui n'ont plus de maisons. *Update: Applications of Research in Music Education*, *31*(1), 72–79.

Helmholtz, H. L. F. v. (1877). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. F. Vieweg und Sohn.

Honing, H., & Ladinig, O. (2009). Exposure influences expressive timing judgments in music. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 281–288.

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. MIT press.

Koelsch, S. (2012). *Brain and music*. John Wiley & Sons.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch* (Vol. 17). Oxford University Press New York.

Krumhansl, C. L. (1996). A perceptual analysis of mozart's piano sonata k. 282: Segmentation, tension, and musical ideas. *Music Perception*, 401–432.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *51*(4), 336.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological review*, *89*(4), 334.

Lehne, M., Rohrmeier, M., Gollmann, D., & Koelsch, S. (2013). The Influence of Different Structural Features on Felt Musical Tension in Two Piano Pieces by Mozart and Mendelssohn. *Music Perception: An Interdisciplinary Journal*, *31*(2), 171–185.

Lerdahl, F. (1996). Calculating Tonal Tension. *Music Perception: An Interdisciplinary Journal*, *13*(3), 319–363.

Lerdahl, F. (2001). *Tonal pitch space*. Oxford University Press.

Lerdahl, F., & Jackendoff, R. S. (1983). *A generative theory of tonal music*. MIT press.

Lerdahl, F., & Krumhansl, C. L. (2007). Modeling Tonal Tension. *Music Perception: An Interdisciplinary Journal*, *24*(4), 329–366.

Lychner, J. A. (1998). An empirical study concerning terminology relating to aesthetic response to music. *Journal of Research in Music Education*, *46*(2), 303–319.

Margulis, E. H. (2005). A Model of Melodic Expectation. *Music Perception: An Interdisciplinary Journal*, *22*(4), 663–714.

Melo, A. F., & Wiggins, G. (2003). A connectionist approach to driving chord progressions using tension. In *Proceedings of the aisb'03 symposium on creativity in arts and science.*

Nielsen, F. (1983). Oplevelse of Musikalsk Spaending [The experience of musical tension]. *Akademisk Forlag, Copenhagen.*

Nielsen, F. (1987). Musical tension and related concepts. *The semiotic web*, *86*, 491–513.

Palmer, C. (1996). Anatomy of a performance: Sources of musical expression. *Music Perception*, 433–453.

Parncutt, R. (1989). *Harmony: A psychoacoustical approach.* Springer-Verlag Publishing.

Parncutt, R., & Bregman, A. S. (2000). Tone Profiles following Short Chord Progressions: Top-down or Bottom-up? *Music Perception: An Interdisciplinary Journal*, *18*(1), 25–57.

Pressnitzer, D., McAdams, S., & Winsberg, S. (2000). Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Perception & ....*

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience*, *13*(2), 241–255.

Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. a critical review. *International Journal of Psychophysiology*, *83*(2), 164–175.

Rozin, A., Rozin, P., & Goldberg, E. (2004). The feeling of music past: How listeners remember musical affect. *Music perception*, *22*(1), 15–39.

Schenker, H. (1935/1979). *Free composition.* New York: Longman.

Schmuckler, M. A. (1989). Expectation in Music: Investigation of Melodic and Harmonic Processes. *Music Perception: An Interdisciplinary Journal*, *7*(2), 109–149.

Steedman, M. J. (1984). A generative grammar for jazz chord sequences. *Music Perception*, *2*(1), 52–77.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychological review*, *107*(4), 885.

Vassilakis, P. (2008). Culture-dependent emotional reactions to music. In *CCRMA*.

Williams, L. R., Fredrickson, W. E., & Atkinson, S. (2011). Focus of attention to melody or harmony and perception of music tension: An exploratory study. *International Journal of Music Education*, *29*(1), 72–81.