The Influence of Different Structural Features on Felt Musical Tension in Two Piano Pieces by Mozart and Mendelssohn

Moritz Lehne, Martin Rohrmeier,
Donald Gollmann, & Stefan Koelsch
Freie Universität, Berlin, Germany

In tonal music, patterns of tension and resolution form one of the core principles evoking emotions. The experience of musical tension and resolution depends on various features of the music (e.g., dynamics, agogics, melody, and harmony); however, the relative contribution of different features to the experience of tension is less clear. To investigate the influence of different features on subjectively experienced musical tension, we compared continuous ratings of felt musical tension for original and modified versions of two piano pieces by Mendelssohn and Mozart. Modifications included versions without dynamics and without agogics as well as versions in which the music was reduced to its melodic, harmonic, or outer voice components. Additionally, we compared tension ratings with a loudness model. Tension ratings for versions without dynamics, versions without agogics and without dynamics, and outer voice reductions correlated highly with ratings for the original versions for both pieces. Tension rating correlations between melodic or harmonic reductions and original versions, as well as loudness and original ratings, differed between pieces and appeared to depend on the relative importance of the feature in the respective piece. In addition, qualitative analyses suggested that felt tension and resolution depend on phrase structure, local harmonic implications, and global syntactic structures of the pieces. Altogether, results indicate that discarding expressive features such as dynamics and agogics largely preserves tension-resolution patterns of the music, whereas the contributions of harmonic and melodic structure depend on the way in which they are employed in the composition.

Received: July 21, 2012, accepted April 10, 2013.

Key words: musical tension, music-evoked emotion, continuous rating, structural features, loudness

A striking feature of music is its ability to evoke strong emotional experiences in the listener (Juslin & Västfjäll, 2008; Koelsch, 2010). However, the principles underlying the evocation of emotions by music are still not fully understood. In particular, it is not clear how structural features of a major-minor tonal piece (melody, meter, rhythm, harmony, loudness, and timbral stucture including instrumentation and texture) relate to the emotional experience of the listener. A concept that can help to elucidate this relationship is *musical tension*.

Musical tension refers to the continuous change of tension and relaxation that is usually experienced when listening to a piece of Western tonal music. Because musical tension is strongly linked to processes such as expectancy build-up, violation or fulfillment of expectancies, to the anticipation of resolution after a breach of expectancy, and to the eventual resolution of such a breach, musical tension plays an important role in the emotional aspects of music listening (see also Huron, 2008; Koelsch, 2012; Margulis, 2005; Meyer, 1956; Narmour, 1992; Rohrmeier & Koelsch, 2012). This is corroborated by empirical research showing that subjective ratings of musical tension correlate with ratings of discrete emotions (sadness, fear, and happiness) and physiological responses during music listening (Krumhansl, 1997). It has furthermore been shown that subjective ratings of musical tension correlate highly within individuals between different exposures to the same music piece (Krumhansl, 1996) as well as between different groups of persons such as musicians and nonmusicians or school children of different ages (Fredrickson, 1997, 1999, 2000), suggesting relatively consistent and stable underlying cognitive and affective processes.

Various structural features have been identified in mediating musical tension ranging from dynamics, timbre, melodic contour, harmony, tonality, repetition (Nielsen, 1983), phrase structure, and note density (Krumhansl, 1996), to pitch height, loudness, onset frequency, and tempo (Farbood, 2012). From a theoretical perspective, the hierarchical structure of music

Music Perception, volume 31, issue 2, pp. 171–185, issn 0730-7829, electronic issn 1533-8312. DOI: 10.1525/mp.2013.31.2.171

(e.g., Lerdahl & Jackendoff, 1983; Rohrmeier, 2007, 2011; Schenker, 1935) can inform models of tension. For instance, Lerdahl's model of tension (Lerdahl, 1996, 2001; Lerdahl & Krumhansl, 2007) combines predictions based on surface structure and tonal distance as well as higher order hierarchical dependency structure. The notion of an influence of local tonal structure on musical tension has been supported by behavioral studies (Bigand & Parncutt, 1999; Bigand, Parncutt, & Lerdahl, 1996). The most elaborate model of musical tension has been devised by Farbood (2012), and takes into account dynamics, pitch height of melody, bass and inner voices, tempo, onset frequency, and harmony. Additionally, the model accounts for the dynamic nature of the music listening process by incorporating attentional and memory processes.

The aforementioned studies (Bigand & Parncutt, 1999; Bigand et al., 1996; Farbood, 2012; Krumhansl, 1996) show that different structural features are related to the subjective experience of musical tension, and that experienced tension can be modelled based on such features. However, research investigating how the experimental manipulation of specific structural features in real music pieces affects experienced tension is scarce. The aim of the present study, therefore, was to investigate how the elimination or isolation of different structural features affected subjectively experienced musical tension using ecologically valid music stimuli. For this, a behavioral experiment was conducted in which continuous ratings of felt musical tension were acquired for original recordings and different modified versions of two piano pieces by Mendelssohn and Mozart. The structural features investigated were dynamics, agogics, harmony, melody, outer voices (as the most salient voices embodying substantial parts of the musical structure), and loudness. For the modified versions, these features were experimentally manipulated yielding versions without dynamics, without dynamics and without agogics, and versions in which harmony, melody, or outer voices were played in isolation (without dynamics and without agogics). Apart from the modified versions, a loudness model was used to investigate to which degree loudness changes accounted for the experienced tension in the original recording. By comparing the tension ratings of the different versions as well as the loudness estimated by the model, the contribution of the different features on subjectively experienced musical tension was evaluated. In addition, we performed a qualitative music-theoretical analysis investigating which musical events corresponded to the peaks and troughs of the tension profiles.

Method

PARTICIPANTS

Data from 28 participants (aged 20-46 years, $M = 25.4$, 16 female) were included in the analysis (data from three additional participants had been excluded due to missing ratings, negative correlations between a participant's tension rating and average tension ratings, or very fast and up-and-down slider movements of one participant throughout the experiment). Ten participants had received instrument or singing lessons in addition to basic music education at school (instruments and years of training: clarinet: 3 years; violin: 10 and 8 years; piano: 6, 7, 7, and 8 years; guitar: 1 year; trumpet: 6 years; voice: 7 years). All participants gave their written consent and were compensated with course credit for participation.

STIMULI

As stimulus material, Mendelssohn Bartholdy's Venetian Boat Song (Op. 30, No. 6) and the first 24 measures of the second movement of Mozart's Piano Sonata KV 280 were used (the music scores are included in the Appendix). To keep the duration of the experiment reasonable, repetitions indicated in the scores were omitted.

The pieces were performed by a professional pianist on a Clavinova CLP-130 (Yamaha Corporation, Hamamatsu, Japan) from which MIDI data were recorded. This allowed for a selective manipulation of specific parameters of the music. From the original recordings, the following five modified versions were created: 1) a version without dynamics, i.e., all notes were played with the mean MIDI key-stroke velocity value of the piece (Mendelssohn: 42; Mozart: 36); 2) a "deadpan" version without dynamics and without agogics (i.e., all notes were played with the same MIDI key-stroke velocity and without any variations in tempo); 3) a version containing only a harmonic reduction of the piece (i.e., non-chord tones were eliminated and remaining notes of one chord were played synchronously, see Figure 1), presented without dynamics and without agogics (henceforth referred to as *harmony version*); 4) a version containing only the outer voices (i.e., only the highest and lowest voice of the piece, see Figure 1), presented without dynamics and without agogics; and 5) a version that consisted of the top voice only (i.e., a version that contained only the "melody part," see Figure 1), presented without dynamics and without agogics (henceforth referred to as *melody version*). That is, versions 3-5, did not vary in terms of tempo, nor dynamics. The total length of the versions without

FIGURE 1. Score excerpts of measures 6-11 of original and modified versions of the Mendelssohn piece: (a) original; (b) harmonic reduction; (c) outer voice reduction; (d) melodic reduction.

agogics matched the ones of the original versions with agogics (Mendelssohn: 2:27 min; Mozart: 2:01 min).

All resulting MIDI files were used to trigger the VST Plugin "The Grand," an authentic grand piano simulation based on samples of real grand piano recordings, in Steinberg Cubase SL (Steinberg Media Technologies, Hamburg, Germany). From this, audio files were generated (16 bit, 44.1 kHz sampling rate) which were used as final stimulus material.

EXPERIMENTAL DESIGN

Original and modified versions of the two pieces were presented to participants in random order. In addition, the original version of the Mozart piece was presented again at the end of the experiment to evaluate the within-participant consistency of the tension ratings over repeated exposures to a music piece. Thus, in total 13 stimuli were presented to each participant (2 pieces x 6 versions + 1 repetition of the original Mozart piece). Tension ratings were obtained every 10 ms from the position of a slider that was shown vertically on a computer screen and could be moved with the mouse according to the subjectively felt musical tension. A high position of the slider corresponded to a high degree of tension while lower positions indicated lower levels of tension.

EXPERIMENTAL PROCEDURE

For stimulus presentation and data acquisition the software Presentation (Neurobehavioral Systems, Albany, USA) was used. Participants listened to the stimuli via headphones at a comfortable volume level. They were instructed to use the slider to continuously 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.

To familiarize participants with the task, they completed a practice trial during which they could ask questions concerning the task (a three minute excerpt from the second movement of Schubert's Piano Sonata D. 960 was used for the practice trial). Before each stimulus presentation, the slider was reset to the lowermost position. Between stimulus presentations, participants were given the opportunity to take a short rest. After finishing the experiment, participants completed a short questionnaire assessing previous music education, music listening habits, familiarity with the pieces, and additional demographic data (age, sex, and occupation). In total, the duration of one experimental session was approximately 45 min.

DATA ANALYSIS

Each participant's data were converted to z-scores (to discard differences between participants with respect to the slider range used for the tension ratings). To compare ratings of versions that contained agogics (i.e., the original version and the version without dynamics) and versions with constant tempo (i.e., deadpan, harmony, outer voices, and melody versions), tension ratings were temporally aligned. This was done by stretching or compressing ratings within each measure of the versions with agogics to the length of the corresponding measure in the versions without agogics using linear interpolation.

Tension ratings were averaged across participants, separately for each version. Thus, potential order effects in the individual ratings were minimized and variance due to individual rating styles was reduced. Comparisons between different versions were performed on the resulting averaged tension ratings. When comparing continuous rating data, it is common practice to calculate Pearson product-moment correlation coefficients. However, this procedure has been criticized (Schubert, 2002, 2010) due to two problems. First, continuous rating data are usually not normally distributed, rendering parametric statistics inappropriate. Second, the data are serially correlated (i.e., adjacent points of the time-series have more similar values than more distant points), which can lead to inflated correlation results (this is particularly problematic when significance tests are performed, because the large number of data points inflates the degrees of freedom, thus greatly reducing the threshold at which correlation results become significant). To mitigate these problems, this study uses Spearman's rank correlation coefficients as a nonparametric measure of correlation (cf. Schubert, 2010; Vines, Krumhansl, Wanderley, & Levitin, 2006). To reduce serial correlations, the data were downsampled to a sampling rate of 1/3 Hz before calculating correlations between tension ratings (cf. Schubert, 2010).

In addition to the correlation between different versions, the correlation between loudness of the music and tension ratings for the different versions was calculated. Loudness was computed from the unmodified recordings using a Matlab implementation (genesis-acoustics.com/en/loudness_online-32.html) of the loudness model for time-varying sounds by Zwicker and Fastl (1999). Taking a time-varying acoustical signal as input, this model estimates the loudness of the signal as it is subjectively experienced (measured in *sone*). When comparing loudness of the music to the tension ratings, it has to be considered that tension ratings temporally lag behind the musical events they refer to (due to the time the participants need to process the stimulus, and to give a physical response on the slider). To quantify this time lag and correct for it, cross-correlations between loudness and the average tension ratings of the unmodified recordings were calculated. The time point with the highest cross-correlation between the two series was then used as an estimate of the temporal lag of the tension ratings. Before calculating correlation coefficients between loudness and tension ratings, the predictions of the loudness model were temporally shifted to correct for this lag. To make loudness data comparable to the tension ratings, loudness data were temporally aligned to versions without agogics and downsampled to 1/3 Hz (analogous to the procedures described above).

To test the consistency of the tension ratings within participants, the test-retest reliability of the tension ratings was evaluated by calculating Spearman's rank correlation coefficients between the first and second presentation of the original recording of the Mozart piece. (In contrast to the other correlations, this correlation coefficient was not calculated on average ratings but on each participant's individual ratings.)

To gain a better understanding of the musical events mediating the experience of musical tension, we also performed a post-hoc music-theoretically informed qualitative analysis in which we investigated which musical events corresponded to the peaks and troughs of the average tension profiles.

Results

Rating reliability within participants was assessed by calculating the correlation between individual tension ratings of the two presentations of the Mozart piece. Correlation coefficients ranged from –.06 to .88 (M = .52). For five participants rating reliability was relatively low ($\rho < .30$), but we nevertheless included these data sets in the analysis so that results are representative for the general population (notably, excluding these participants yielded results that were highly comparable to the results reported here).

Figure 2 shows individual tension ratings for the two original recordings as well as their average and standard deviation together with the waveform of the audio signal and loudness. The graphs reveal tension profiles with distinct peaks and troughs. Furthermore, visual comparison of the tension profiles with the audio waveform and loudness indicates a relation between tension and loudness (especially for the Mendelssohn piece) that will be investigated in more detail below.

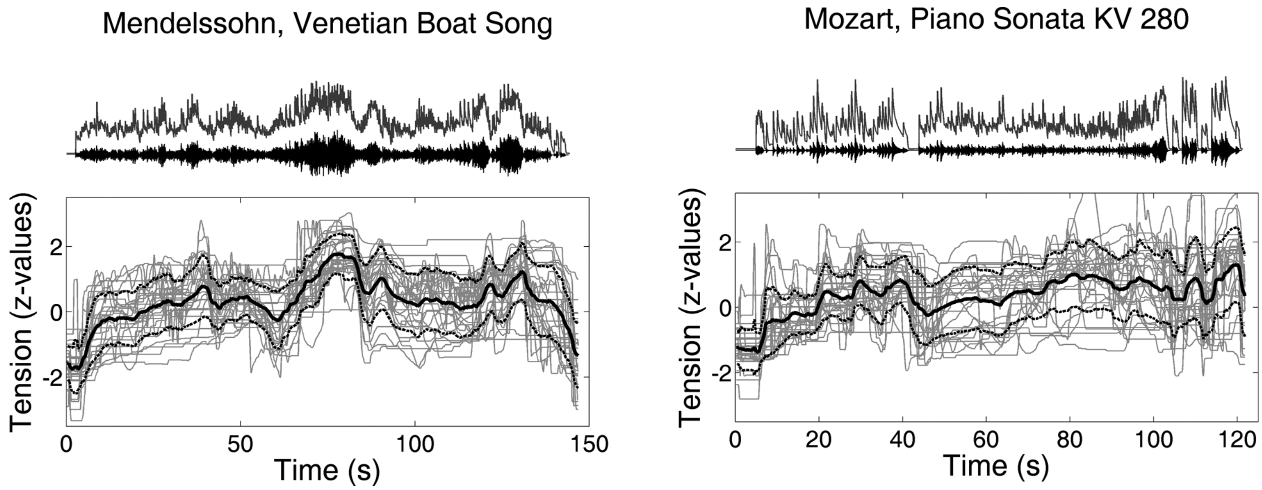

FIGURE 2. Top: Audio waveform (black) with predictions of the loudness model (gray) for the original recordings of the two pieces. Bottom: Individual (gray lines) and average tension ratings (black solid line) $\pm$ standard deviation (black dashed lines).

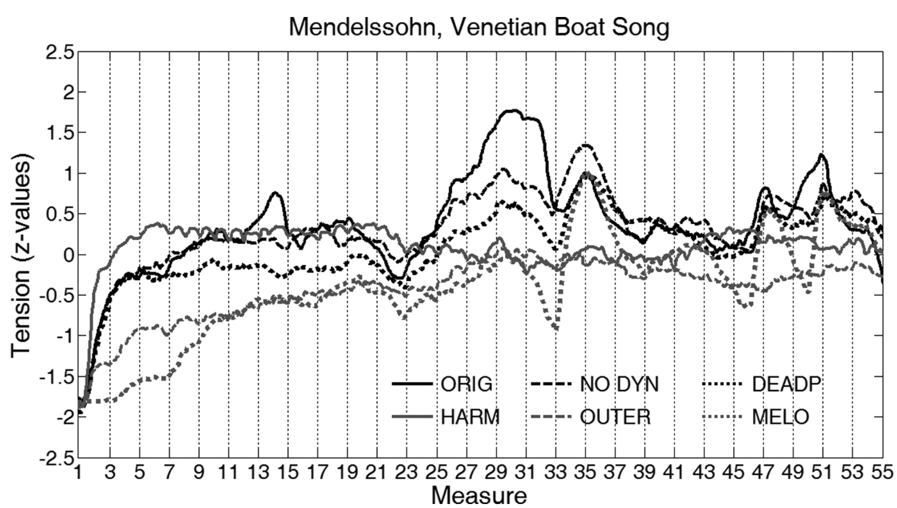

FIGURE 3. Average tension ratings for all versions of the Mendelssohn piece (ORIG: original version; NO DYN: version without dynamics; DEADP: deadpan version without dynamics and expressive timing; HARM: harmony only; OUTER: outer voices only; MELO: melody only).

Average tension profiles for the different versions of the two pieces are shown in Figures 3 and 4. For both pieces, the clearest and highest tension peaks were observed for tension ratings of the original recordings, which on average received higher tension ratings than the versions without expressive features: for the Mozart piece, Wilcoxon signed-rank tests revealed differences between the original and the version without dynamics ($z = -4.06$, $p < .05$), as well as between the original and the deadpan version ($z = -1.98$, $p < .05$); for the Mendelssohn piece only differences between the original and the deadpan version were significant ($z = -4.48$, $p < .05$). Furthermore, changes in experienced tension appeared to be more pronounced for the Mendelssohn piece than for the Mozart piece.

CORRELATION ANALYSIS

Figure 5 shows Spearman's rank correlation coefficients between all possible combinations of average tension ratings of the different versions (the correlation matrices also show correlations to the loudness model which will be treated below).

For the Mendelssohn piece, the rating of the original recording correlated highly with the rating of the version without dynamics ($\rho = .83$), the deadpan version ($\rho = .70$), and the outer voices ($\rho = .71$). Correlation

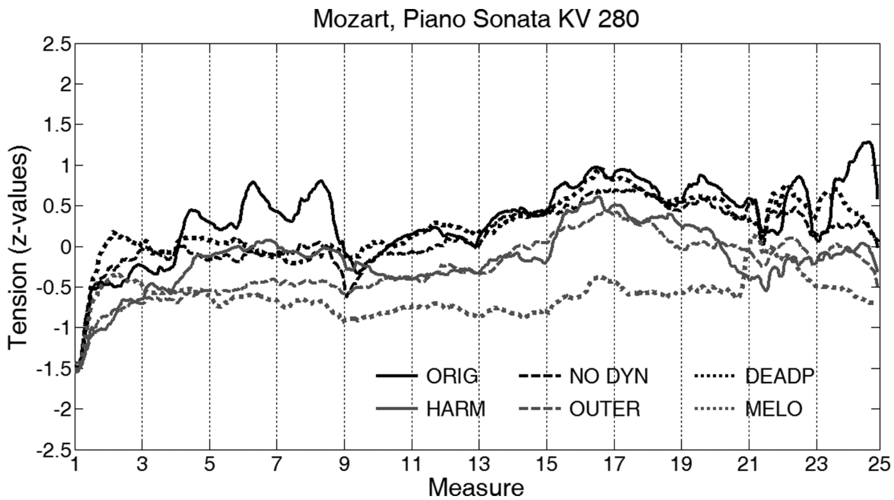

FIGURE 4. Average tension ratings for all versions of the Mozart piece.

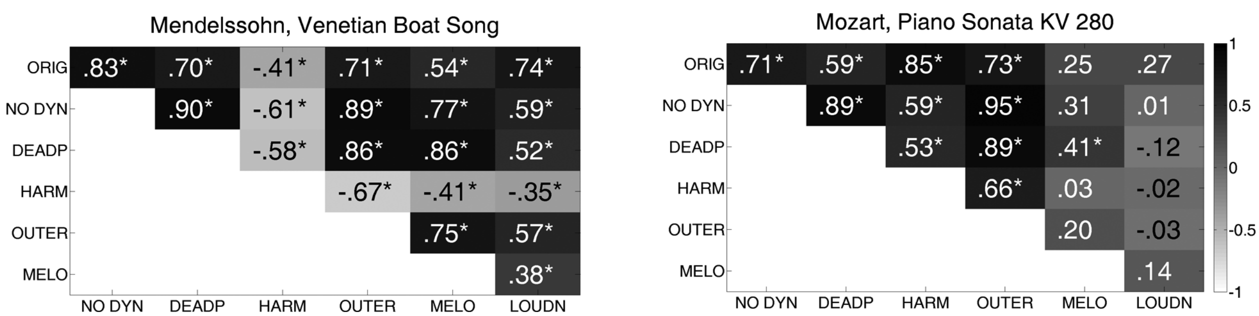

FIGURE 5. Spearman's ρ for average tension ratings of the different versions (ORIG: original version; NO DYN: version without dynamics; DEADP: deadpan version without dynamics and expressive timing; HARM: harmony only; OUTER: outer voices only; MELO: melody only) as well as the loudness model (LOUDN). Significant correlations ($p < .05$) are marked with a star ($df_{\text{Mendelssohn}} = 44$; $df_{\text{Mozart}} = 36$).

with the melody version was moderate (ρ = .54). The correlation with the harmonic reduction was negative (ρ = –.41). A similar pattern was observed for the Mozart piece: The rating of the original recording correlated moderately to highly with the rating of the version without dynamics (ρ = .71), the deadpan version (ρ = .59), and the outer voices (ρ = .73). However, in contrast to the Mendelssohn piece, the harmonic reduction correlated highly with the original version (ρ = .85) and the correlation with the melody version was lower and not significant (ρ = .25). This piece-dependant difference for tension ratings of harmony and melody versions was observed consistently in the two correlation matrices: For the Mendelssohn piece, correlations between the rating of the melody version and ratings of the other versions (except harmony) were relatively high, whereas correlation between harmony and the other versions were all negative. This pattern was virtually reversed for the Mozart piece. Here, ratings for the harmony version correlated highly with ratings for the other versions (except melody), while correlations between ratings for the melody version and the other versions were lower and (except for the deadpan version) not statistically significant. Apart from these differences between harmony and melody, the general pattern of the correlations between ratings for different versions were relatively similar for the two pieces. Except for correlations between tension ratings of the melody version and other versions of the Mozart piece (original, no dynamics, harmony, and outer voices), all tension rating correlations were statistically significant ($df_{\text{Mendelssohn}} = 44$; $df_{\text{Mozart}} = 36$; $p < .05$).

Before calculating correlations between loudness and tension, the temporal lag of the tension ratings to corresponding musical events was determined by computing the cross-correlation between loudness and the average tension ratings of the original recordings. For the Mendelssohn piece, the highest correlation was observed at a time lag of 3.2 s. For the Mozart piece, correlation was highest at a lag of 2.0 s. The correlation coefficients reported in the following (also shown in Figure 5) were obtained after correcting for the time lag of the tension ratings. For the Mendelssohn piece, a high positive correlation between loudness and the average tension rating of the original recording was observed ($\rho = .74$). Interestingly, loudness also correlated significantly with tension ratings of versions without dynamics. For the Mendelssohn piece, all correlations with loudness were positive and statistically significant ($df = 44$; $p < .05$). For the Mozart piece, none of the correlations between tension ratings and loudness were statistically significant.

QUALITATIVE ANALYSIS

For the Mendelssohn piece, the most prominent peaks at measures 13, 30, 34, and 50 corresponded to events with dominant function: m. 13 features the first strongly pronounced dominant function of the piece (with $\hat{5}$ in the bass after a long series of $\hat{1}$ pedal bass notes), mm. 30-32 feature the main structural dominant, the diminished chord at m. 34 fulfills an applied dominant function to the *IV*, and the two chords at mm. 46, 60 constitute dominants of the final phrase of the piece (the second and final dominant features the stronger tension rating). However, participants did not simply give high tension ratings to local (or applied) dominants, because other dominants were not associated with peaks, such as the dominant in m. 20 towards the end of the first phrase, m. 23 initiating the motion of the middle section, and m. 42 ending the middle section. In contrast, it appears that participants also attended to the overall organization of the piece when experiencing tension: The lowest rating, except ratings for the beginning, was given at m. 21, the end of the first section and beginning of the middle section, as well as at m. 43 at the end of the middle section and beginning of the final closure.

Many of the events associated with peaks correspond to salient events in the melody line (trills or high notes) as well as dynamics (e.g., sforzato or forte notes at measures 13, 30, 46, and 50). However, when dynamics and agogics were removed, most peaks remained present, yet less pronounced. This suggests that some expressive features employed in the piece enhance the effects of tension that are created by melodic means. In contrast, the tension profile of the harmony version did not show clear peaks and remained relatively flat throughout the piece with a slight overall downward trend. Hence, harmony did not serve as the sole or predominant compositional device to create tension in this piece.

For the Mozart piece, tension profiles also reflect the overall organization of the piece. As in the Mendelssohn piece, the transitions between first, middle, and final parts are reflected in low tension ratings (mm. 9, 21). The first section, mm. 1-8, features five ascending peaks at mm. 2, 3, 4, 6, 8. Each of these peaks corresponds to harmonic events with strong implications: II^6_5, V, VI^6_4, I^4_2. The relaxations of the tension profile correspond with the implied local resolution of the suspensions in mm. 5, 7, 8. The fact that the tension profile constantly rises towards m. 8, even though mm. 4 and 6 are musically identical, seems to reflect the overarching tendency towards the resolution and completion of the phrase towards the final tonic. The tension profile of the middle part reflects the departure from the initial tonic to the *IV* as well as the half cadence with a resolution to the dominant V^6_4–V^5_3 marking the end of the half of this phrase (m. 12). Similarly to the beginning, the harmonically strongly implicative German sixth, diminished, and dominant seventh chords (mm. 13, 15, 16) are associated with ascending sets of peaks along with small local relaxations (reflecting melodic and harmonic relaxation). The musical relaxation around the tonic at mm. 17-18 (after the preceding dominants) is also reflected by a local trough in the tension profile. The deceptive cadence (mm. 19) as well as the (precadential) turns to the relative minor (mm. 21, 23) yield a sudden increase in experienced tension. The subsequent resolutions towards the tonic (as in mm. 19-21, 22, 24) are reflected as a decrease in the tension profile. The cadential V^6_4, as well as the entire final cadential schema at mm. 21-24, provide strong signals of the upcoming end (of the part) and receive the strongest and most pronounced local tension ratings.

Discussion

The aim of the present study was to investigate how the elimination or isolation of different structural features (dynamics, agogics, harmony, outer voices, and melody) influences felt musical tension by comparing tension ratings of ecologically valid original and modified versions of two piano pieces. Modifications featured versions without dynamics, without dynamics and without agogics, and versions in which the music was reduced to its harmonic, melodic, or outer voice component. In addition, we compared tension ratings with the loudness of the music estimated by a standard loudness model.

We found that the overall shape of the profiles of original versions, versions without dynamics, and deadpan versions (without dynamics and without agogics) resembled each other closely. This was reflected in high correlations between the tension ratings of these versions, indicating that discarding dynamics and agogics preserves a large part of the tension-resolution patterns of the music, and that felt tension is not primarily governed by these expressive features. This is consistent with findings by Krumhansl (1996), who reported a high correspondence between tension ratings of versions with and without expressive features (i.e., dynamics and agogics). Our results support that models of tension based on the tonal structure of a musical piece abstracting from expressive features (e.g., Lerdahl, 1996, 2001) capture a large part of the information relevant for the experience of musical tension. The high correlations between tension ratings of the original recordings and versions reduced to the outer voices furthermore suggest that even when limiting information to these most salient voices, considerable parts of the tension patterns are retained. This confirms that outer voices embody major aspects of the musical structure.

Despite the high correlations between original tension ratings and ratings for versions without dynamics and agogics, discarding these features does have a notable effect on felt musical tension. Average tension ratings of versions without expressive features were significantly lower than for original versions, and the tension profiles were generally flatter, with some of the tension peaks existent in the profiles of the original versions not present or strongly attenuated. This indicates that musical tension can be strongly enhanced by expressive features.

For the Mendelssohn piece, we found a high correlation between loudness and the tension profile of the original version. Interestingly, loudness also correlated significantly with versions without dynamics.[1] This finding indicates a strong redundancy between the dynamics and other structural aspects of a music piece (such as harmony or melody). It underpins that the alignment of expressive features (e.g., dynamics and agogics) and tonal aspects (e.g., melodic or harmonic structure) can enhance the experience of tension, and is a core compositional and performative device that can help to maximize the emotional effect of the music. In the Mendelssohn piece, for example, the highest tension peak (m. 29) reflects the main structural dominant and is prepared by a long crescendo, the rising melody line, the lowest local bass note, and the *fortissimo* and *sforzato* of both repetitions of the chords (as well as expressive details played by the pianist) so that the co-aligned combination of dynamics, melody, and harmony results in a strong experience of increasing tension apparent in the prominent peak in the corresponding tension profile. The redundant use of different features as a compositional device gains further support from a study by Lalitte, Bigand, Kantor-Martynuska, and Delbé (2009), who report high correlations between musical arousal ratings of two original Beethoven sonatas and two atonal counterparts. This stability of musical arousal profiles even in absence of tonal structure indicates that participants responded to features that remained relatively unaltered between the two versions (such as rhythm, note density, or global structure). Assuming that participants' ratings to a large extent depend on tonal structure (as suggested by our results and by previous research, see Krumhansl, 1996; Lerdahl & Krumhansl, 2007), this indicates that tonal aspects tend to covary with nontonal features of the music.

With regard to differences between pieces, our results suggest that melody and harmony contributed differently to experienced tension in the two pieces. For the Mendelssohn piece, correlation with the tension profile of the original version was higher for tension ratings of the melody version than for ratings of the harmony version, whereas the reverse pattern was observed for the Mozart piece. This difference seems to reflect a compositional difference between both pieces: Whereas the rate of harmonic change is slow in the Mendelssohn piece and its melody part plays a major role in shaping the structure (long trills, overarching melodic ascents or descents), dense successions of harmonic implication and resolution patterns govern the Mozart piece to a larger extent. However, the qualitative analysis suggested that central tension peaks in both pieces were driven by harmonic patterns, which is inconsistent with the negative correlations between the tension profile of the harmony versions and the other versions observed for the Mendelssohn piece. These negative correlations may have resulted from the slow rate of harmonic change of the Mendelssohn piece, which rendered the harmonic reduction rather uninteresting to listen to, thus accounting for the relatively flat tension profile of the harmony version and its slight downward trend. The larger contribution of harmony on experienced tension observed for the Mozart piece is also in concordance with results by Williams, Fredrickson, and Atkinson (2011), who showed that focusing more on harmony of a Mozart piece was related to higher tension ratings as compared with attending more to the melody,

[1] Note that even the versions without dynamics still retained some minor loudness variations due to varying note density or short intervals of silence in the music, however, these were negligible compared with the dynamic variations due to the expressive performance.

which indicates a higher importance of harmony in comparison to melody for inducing an experience of tension in Mozart pieces.

Another feature differing between pieces was loudness, which correlated significantly with tension profiles of the Mendelssohn piece but not of the Mozart piece. This also seems to reflect a different importance of this feature in the respective piece that already becomes apparent when comparing the scores of the pieces. Whereas the score of the Mendelssohn piece includes marked *crescendi* and dynamic indications ranging from *pianissimo* to *fortissimo*, the Mozart piece makes more limited use of dynamics with indications ranging from *piano* to *forte*. It seems probable that the larger dynamic variations of the Mendelssohn piece made them stand out more clearly against other musical features, thus accounting for the high correlations between loudness and tension for this piece. This piece-dependant influence of different features on experienced tension indicates that it is problematic to model musical tension based on the assumption that the contribution of different features on experienced tension remains constant over different pieces. Instead, to increase the accuracy of models of musical tension, features should be weighted dynamically depending on the musical context (as, for example, in the model by Farbood, 2012).

With respect to the time lag between tension ratings and corresponding musical events, our results are consistent with observations by Schubert (2004), who reported emotion responses 1 to 3 seconds after the respective musical event. Interestingly, the time lag was shorter for the Mozart than for the Mendelssohn piece, which raises the question as to how properties of the music influence the response times of participants. It has been conjectured that faster tempos and loud sudden sounds decrease response times (Schubert & Dunsmuir, 1999, as cited in Schubert, 2010) which is in line with our results.[2]

The qualitative analysis suggests that participants' tension ratings strongly reflected the form of the piece, i.e., different parts and sections, local harmonic implications and resolutions as well as global overarching syntactic features (cf. Lerdahl & Jackendoff, 1983; Rohrmeier, 2011). For instance, the overarching tension increase mediated by several smaller local tension-resolution patterns in the first phrase of the Mozart piece (mm. 1-8) reflects how local and global structure interact in forming patterns of tension and release. The harmonic structure exhibits recursive nesting of harmonic implications from each chord to its successor by cascading syntactic dependencies that are all directed towards the final tonic of the phrase. The tension profile shows that each of these single implications between two chords feature a small tension-release pattern which is itself embedded in the overall rise of tension towards the final tonic of the phrase. This illustrates that experienced tension can be governed by local and global structure at the same time. On the other hand, the finding that not all local dominants—but mostly those that reflect deep structure (in terms of analytic reduction)—affected the tension profile underscores that dominants do not trigger a rise in tension per se, but that their impact depends on the context of the overarching global structure. Both of these examples demonstrate the interplay between local implications and global syntactic implications for establishing musical tension; both local and global implications and dependencies are embraced by theories of tonal syntax (cf. Lerdahl & Jackendoff, 1983, or Rohrmeier, 2011) without the need for two separate models.[3]

Finally, we would like to point out some limitations of the study. First, tension ratings were only tested for two music pieces. To maximize the ecological validity of the study, we used stimuli based on real music pieces instead of artificial stimuli, which comes at the price that stimuli had relatively long durations (thus limiting the amount of different stimuli that can be delivered in one experimental session). To investigate to which degree the results reported here can be generalized to other pieces, research on musical tension has to be extended to stimuli differing on various dimensions such as music genre, tempo, or orchestration. Second, the within-subjects design of the present study, which exposed participants repeatedly to different versions of the same piece, may have resulted in interference effects between different tension ratings because the tension rating of a stimulus may have been influenced in part by a participant's (implicit) memory of prior presentations of the stimulus in a different version (for online learning during experimental tasks compare Rohrmeier, 2009). This may have led to a slight under or overestimation of the correlation coefficients: Correlations between tension ratings of different versions may either have become stronger (the memory of previous exposures may have made tension ratings of different versions more similar) or weaker (repeated exposures may

[2] Although the notated tempo of the Mozart piece is slower than that of the Mendelssohn piece (*Adagio* vs. *Allegretto tranquillo*), the rate of harmonic and melodic change is higher.

[3] Note, however, that this does not entail that the correlation of tension profiles predicted by syntactic theories with human data provides strong evidence that humans do in fact employ such hierarchical representations in music perception.

have made differences between versions more apparent, resulting in more dissimilar ratings). However, because of the randomized stimulus presentation and the averaging across participants, this possible bias would have affected all correlations between different tension ratings in the same way, keeping the relative comparisons between different correlations valid. Last, we only tested the influence of one psychoacoustical feature—loudness—on experienced musical tension. However, other low-level psychoacoustical features may play a role in mediating tension. In particular, sensory dissonance is likely to have an effect on experienced tension, which is indicated by previous research showing that for single chords subjective roughness ratings correlate with tension ratings (Pressnitzer, McAdams, Winsberg, & Fineberg, 2000) and that predictions of a roughness model correlate with tension ratings of musicians (Bigand et al., 1996). To investigate whether these findings generalize to longer music pieces, future research on musical tension should therefore consider also including measures of sensory dissonance into the analysis.

As a critical note, and general limitation of research on musical tension using one-dimensional tension scales, we would also like to emphasize that the fine-structure of emotional activity underlying tension phenomena (including its neural correlates) cannot be grasped adequately by one-dimensional tension values, and subjective ratings of high-level concepts such as tension (and emotion ratings in general) only provide limited insight into the multiple cognitive and affective mechanisms underlying the subjective emotional experience.

As laid out previously (Koelsch, 2012), this is because different structural principles with different affective qualities can give rise to tension or resolution: The *build-up of a musical structure* (which may lead to a rise in tension), a *breach of expectancy* (which also leads to a rise in tension), the *anticipation of resolution* after the breach of expectancy (which usually either maintains, or even increases, tension), and the *resolution of a breach* (leading to release of tension) are qualitatively different phenomena, yet they are not differentiated when measuring tension with a one-dimensional scale. Thus, for example, the tension value of a tonic chord at the beginning of a harmonic sequence with a structural breach is similar, or even identical, to the tension value of a tonic chord at the end of a sequence (both tonic chords have low tension values). However, the underlying affective phenomena are different (*build-up* vs. *resolution*). Such differences in cognitive and affective phenomenology are relevant for investigations in related fields such as neuroscientific investigations of tension phenomena: A working hypothesis suggested recently (Koelsch, 2012) is that a specific brain region (the dorsal striatum) is involved in emotional activity due to anticipation: In a functional neuroimaging study by Koelsch, Fritz, and Schlaug (2008) this region was activated during blocks of chord sequences with irregular chords evoking the anticipation for resolution. A study by Salimpoor, Benovoy, Larcher, Dagher, and Zatorre (2011) showed release of the neurotransmitter dopamine in this region while listeners anticipated a music-evoked frisson (an intensely pleasurable experience often involving goosebumps or shivers down the neck, arms, or spine). Activity changes in another brain region (the amygdaloid complex) appear to be related to the processing of breaches of expectancy (Koelsch et al., 2008), and yet another brain region might be involved in the processing of resolution: In the study by Salimpoor et al. (2011), the anticipated and rewarding frisson itself evoked dopaminergic activity in another brain structure (the ventral striatum, presumably the so-called nucleus accumbens). Thus, the pleasurable and rewarding experience of the resolution of a breach of expectancy might involve activity in different brain structures than those involved in the anticipation of the resolution. Such considerations illustrate that a multidimensional approach to tension might be necessary for fruitful future research.

Conclusion

The present study investigated the effect of different structural features of music (dynamics, agogics, harmony, outer voices, melody, and loudness) on felt musical tension. Overall, tension ratings for versions without expressive features (dynamics and agogics) correlated highly with ratings of the original recordings, indicating that the general tension-resolution pattern of a music piece is governed essentially by its tonal structure, rather than by expressive features. Adding expressive features, however, can enhance the experience of musical tension. The relative contribution of loudness, as well as melody and harmony, depended more on the special characteristics of individual pieces with more salient features apparently having a stronger impact on felt musical tension. A qualitative analysis suggested that participants are sensitive to core features of harmonic, melodic, and global syntactic musical structure.

Author Note

Martin Rohrmeier is now at the Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States.

This research was supported by the Excellence Initiative of the German Federal Ministry of Education and Research. We would like to thank Yu Fukuda and Vienna Dönnie for helping with the preparation of the music stimuli. We also thank the action editor, Emmanuel Bigand, and two anonymous reviewers for helpful comments and suggestions.

Correspondence concerning this article should be addressed to Moritz Lehne, Cluster of Excellence "Languages of Emotion", Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. E-mail: moritz.lehne@fu-berlin.de

References

Bigand, E., & Parncutt, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research, 62*, 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 and Psychophysics, 58*, 124-141.

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

Fredrickson, W. E. (1997). Elementary, middle, and high school student perceptions of tension in music. *Journal of Research in Music Education, 45*, 626-635.

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

Fredrickson, W. E. (2000). Perception of tension in music musicians versus nonmusicians. *Journal of Music Therapy, 37*, 40-50.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, Special issue 2001-2002*, 123-147.

Huron, D. (2008). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press

Juslin, P. N., & Västfjäll, D. (2008). Emotional repsonses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31,* 559-621.

Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends in Cognitive Sciences, 14*, 131-137.

Koelsch, S. (2012). *Brain and music*. West Sussex, UK: John Wiley & Sons, Ltd.

Koelsch, S., Fritz, T., & Schlaug, G. (2008). Amygdala activity can be modulated by unexpected chord functions during music listening. *NeuroReport, 19*, 1815-1819.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's piano sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception, 13*, 401-432.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology, 51,* 336-353.

Lalitte, P., Bigand, E., Kantor-Martynuska, J., & Delbé, C. (2009). On listening to atonal variants of two piano sonatas by Beethoven. *Music Perception*, *26*, 223-234.

Lerdahl, F. (1996). Calculating tonal tension. *Music Perception, 13*, 319-363.

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

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Lerdahl, F., & Krumhansl, C. L. (2007). Modeling tonal tension. *Music Perception, 24*, 329-366.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception, 22*, 663-714.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago, IL: University of Chicago Press.

Narmour, E. (1992). *The analysis and cognition of basic melodic structures: The implication-realization model.* Chicago, IL: University of Chicago Press.

Nielsen, F. V. (1983). *Oplevelse of musikalsk spaending* [The experience of musical tension]. Akademisk Forlag, Copenhagen.

Pressnitzer, D., McAdams, S., Winsberg, S., & Fineberg, J. (2000). Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Perception and Psychophysics*, *62*, 66-80.

Rohrmeier, M. (2007). A generative grammar approach to diatonic harmonic structure. In C. Spyridis, A. Georgaki, G. Kouroupetroglou, & C. Anagnostopoulou (Eds.), *Proceedings of the 4th Sound and Music Computing Conference* (pp. 97-100). Athens, Greece.

Rohrmeier, M. (2009). Learning on the fly. Computational analyses of an unsupervised online learning effect. In A. Howes, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 9th International Conference on Cognitive Modeling - ICCM 2009* (pp. 372-377). Manchester, UK.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music, 5*, 35-53.

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

Salimpoor, V., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience, 14*, 257-262.

Schenker, H. (1935). *Der Freie Satz. Neue Musikalische Theorien und Phantasien* [Free composition. New musical theories and fantasies.]. Margada, Liège, Belgium.

SCHUBERT, E. (2002). Correlation analysis of continuous emotional response to music: Correcting for the effects of serial correlation. *Musicae Scientiae, Special Issue 2001-2002*, 213-236.

SCHUBERT, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception, 21*, 561-585.

SCHUBERT, E. (2010). Continuous self-report methods, In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion - Theory, research, applications* (pp. 223-253). Oxford, UK: Oxford University Press.

VINES, B. W., KRUMHANSL, C. L., WANDERLEY, M. M., & LEVITIN, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition, 101*, 80-113.

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*, 72-81.

ZWICKER, E., & FASTL, H. (1999). *Psychoacoustics - Facts and models*. Berlin and Heidelberg: Springer-Verlag.

Appendix

Mendelssohn Bartholdy, Venetian Boat Song (Op. 30, No. 6)


31
dim.
pp
sfz
36
dim.
p
41
p
cresc.
f
46
dim.
p
cresc.
f
sfz dim.
51
sfz
p
dim.
pp

Mozart, Piano Sonata KV 280 (Second Movement)