EXPLAINING THE EMOTIONAL ENGAGEMENT OF LISTENERS WITH THREE PERFORMANCES OF A SCRIABIN ETUDE

Renée Timmers
Antonio Camurri
Infomus, DIST, University of Genova, Italy

ABSTRACT

Background. Recent years have shown a growing interest into the emotional response to music and into explaining this response on the basis of musical parameters such as loudness, tempo, harmony or melody. Using continuous measurement of self-reported emotional responses, it has become possible, among others, to investigate the intensity of the emotional responses considered and to analyse its relation to the tension and release of music or to continuously varying aspects of the performance.

Aims. This study aims to explain listeners' ratings of emotional engagement for three performances of Skriabin's Etude Op. 8 no. 11 by one pianist. This is first done on the basis of a direct correlation between engagement and parameters of the performance. In a second instance, parameters of the musical structure are added.

Method. A master pianist performed the piece at his best and intensely expressive on a Yamaha Disklavier grand piano. Audio and MIDI information were recorded. 12 participants listened to the performances and continuously rated their emotional engagement with the performances.

Results. Results show that a simple regression model is quite well able to explain the listeners' ratings of emotional engagement on the basis of the tempo and dynamics of the performance, especially the dynamics. They also however show the limitations of this model: 1) the model does well for most, but not all participant, especially not for musicians, 2) it improves when fitted separately for different sections and the effects are therefore not consistent over the entire piece, 3) intercepts are not explained, and 4) there is more variation in the performance than in the ratings. When aspects of the musical structure are added and performance data is smoothened, the regression model increases considerably in power and explains the majority of the emotion data.

Conclusions. Dynamics is an important parameter for the communication of emotional intensity and especially non-musicians follow it when indicating their emotional engagement with the music. Musicians' emotional responses are less predictable and need more and varying parameters to be explained, most importantly polyphony and melodic contour.

1. INTRODUCTION

The emotional response to music seems unpredictable and therefore hard to study and yet there have been an increasing number of studies that do so. The explorative study of Sloboda [4] is exemplary by its investigation of the relation between the emotional responses and aspects of the musical structure: the responses were clearly emotional and responses were remembered to follow certain musical passages that were analyzed for common structural characteristics. More recently several studies have investigated the *perception* of emotion in music more experimentally. The responses to the music did not need to be truly emotional, but the music should be consciously perceived as communicating a certain emotion such as happiness, sadness or anger. Parameters of the musical structure were shown to be responsible for the communication such as its harmony, rhythm, or melodic profile [1], and also features of the musical performance influence the perception of emotions, such as tempo, dynamics and articulation [2].

Besides the type of emotion, the intensity of the response has been investigated using continuous measurement of self-reported emotional responses. The use of continuous measurement also provides the possibility to zoom in and investigate the relation between music and emotional response more locally. For example, Krumhansl [3] related continuous ratings of musical tension to music theoretical predictions of tension and release and to phrase structure. Sloboda and Lehmann [5] investigated the intensity of perceived emotion with performances and found that the points of high emotion often coincided with points that the performers had explicitly mentioned in an interview.

This study is related to the last mentioned studies in that it aims to explain listeners' continuous reports of emotional engagement. We use the term emotional engagement to refer to the task demands of the experiment that instructed the listeners to report the *felt* intensity of emotion. We were especially interested in the role of the musical performance in influencing emotional engagement. Since emotion seems much based on the effect of music structural features, it is for the performer to communicate these. In a similar vein, it is the performer who is responsible for the communication of the ebb and flow, or the tension and release of the music in order for the listener to value the crucial moments. To be more specific, we expected the two main parameters of a musical performance – tempo and dynamics – to communicate the tension and release of the music and to influence the emotional engagement of the listeners.

2. METHOD

2.1. Musical Performances

A professional pianist performed Etude Op. 8 no. 11 by Alexander Skriabin, at a concert that was organised for the experiment's purpose. He first performed without public in a normal manner (to be referred to as p1) and an exaggerated manner (to be referred to as p2) and then performed with public in a normal, concert manner (to be referred to as p3). MIDI and audio recordings were made.

ISBN 3-931852-67-9
ISSN 1617-6847

The Skriabin Etude is a slow and lyrical piece (Andante cantabile) in a late Romantic style. According to the pianist, it can be played with a lot of freedom. Theoretically, the piece has a simple A B A with coda structure (A A' B A'' A''' C to be more precise), but the pianist interpreted the line of the music differently: The first main target of the music is a release of tension halfway the B section. Everything preceding this target point is a preparation for this tension release. The A section is anyway preparatory; it leads towards the start of the B section, which is the real beginning of the piece. After this release of tension, the music builds up towards the dramatic return of the theme of the A section. This prepares for the second possible point of tension release halfway the coda at a general pause. The release is however not continued and the piece ends most sad.

2.2. Participants

Twelve people participated in the experiment among them were four musicians.

2.3. Procedure

The participants sat behind a desk with a slider before them. They heard the three performances of the Skriabin Etude over speakers. They indicated to what extent they were emotionally engaged with the music by moving the slider up and down.

3. RESULTS

3.1. Performance Data

From the MIDI recordings, the average key-velocity for each quarter note, which roughly corresponds to the dynamics of the performance, was calculated as well as inter-onset-intervals (IOI's) between successive quarter notes, which is a measure for local duration.

The resulting profiles of quarter note key-velocity and quarter note IOI are plotted in Figure 1, top panels [TIMMERS_ETAL_FIGURE1.JPG]. Separate graphs are plotted for p1, p2 and p3. Vertical dotted lines indicate section boundaries. Bar numbers are given at the bottom. The profiles were highly similar for the three performances: they all started in a slow tempo and with soft dynamics, had considerable crescendi and accelerandi in the A section, a diminuendo and crescendo in the B section accompanied by first a highly variable tempo and thereafter an accelerando, a fast and loud return of the A section with limited variation in tempo and dynamics, a soft and slower repeat of the theme, and a coda that fades away in dynamics and tempo (see Figure 1 top panels).

In addition to this global pattern, the IOI-profile shows the characteristic peaks of phrase-final lengthenings at a fairly high density and large magnitude. Rubato is quite steep throughout the piece, except in the forte return of the A section (A''). The key-velocity profile shows drops in velocity at most phrase-boundaries, though these are balanced by strong crescendi in most sections.

This is the basic pattern of all three performances. Differences between them are that p2 is an exaggerated version of the other performances.

When we consider the relation with musical structure, it becomes clear that local duration especially communicates phrase endings, while dynamics relates to the tension curve of the music. The first follows from the observation that the end of each two-bar phrase coincides with or is immediately followed by a local peak in duration, while this is less the systematic for minima in dynamics. The second follows from the relation between the dynamics and the pianist interpretation of the music as expressed in the interview: Key velocity increases towards the start of the B section, which the performer considers the real beginning of the piece. It decreases towards the release in tension halfway the B section. Thereafter it builds up towards the dramatic return of the theme (most forte passage), and after a contrasting repeat of the theme, preparations are made for the second release in tension halfway the Coda. The soft dynamics at the end, in contrast, are loaded with tension.

3.2. Listeners Data

The indication of emotional engagement was measured at a sampling rate of 10 Hz using a MIDI-slider that had a range from 0 to 127. The average level of the MIDI-slider per quarter note was calculated for each participant separately and averaged over participants. We will refer to this measure as EM (abbreviation for emotion measure).

The agreement between participants was fairly low with an average correlation between participants of 0.34. Still, the average profile within pieces is statistically reliable: the profiles of the mean plus and minus half of the standard deviation per quarter note correlate highly with each other and with the mean and therefore the mean shows a robust trend ($r = 0.96$ on average). In addition, the standard deviation per quarter note is relatively small with respect to the mean per quarter note (on average 1:2.4).

The average EM's for the three performances are highly similar ($r = 0.93$). For all three performances, EM increases towards the B section. It decreases and increases within the B section, reaches a maximum at the return of the theme, and decreases at the repeat of the theme and in the coda.

3.3. Relation between Listeners' and Performer's Data

The main thesis of the study was that the two main parameters of a musical performance – tempo and dynamics – communicate (the performer's interpretation of) the tension and release of the music and influence the emotional engagement of the listeners. In section 3.1, it was suggested that tempo especially communicates phrase-boundaries, while the dynamics is the performer's expression of the intensity of the music. Phrase-boundaries are a potential point of release in emotion, while the musical intensity relates to rises in tension. The following analyses investigate how the listeners' responses relate to this.

ISBN 3-931852-67-9
ISSN 1617-6847

A series of multiple regression analyses were done that used quarter note IOI and key velocity to predict EM. This was done directly and with a time-delay of one, two and three quarter notes of the performance data with respect to the listeners' data. The best R^2s obtained are reported. This method was applied both for data of the entire piece and for data of individual sections, and for averaged data (averaged over participants) as well as for data of individual participants. An overview of the results is given in Table 1.

It shows the average R^2 per section or for the entire piece. They are always the means of the R^2s for each performance. The second column shows the average R^2 for analyses of data of individual participants, while the last column shows the averages of R^2s for analyses on averaged data (averaged over participants). It can be seen that on average and for the averaged data, the model is rather strong, but it certainly does not do well for all subjects. Velocity most often contributes significantly to the explanation of variance, but IOI also reaches significance quite often. In 43% of the cases, velocity was the only significant parameter. IOI was the only significant parameter for 19% of the analyses, while velocity and IOI were both significant for 17% of the analyses. The direction of the effects was for key-velocity always positive and for IOI always negative, as predicted. In almost all cases, the intercept is a significant parameter and in some cases it is the only significant one. This indicates that the average height of EM is characteristic for a section or the entire piece. For example, the Skriabin Etude is considered fairly emotional in general. In the first A section, EM rises towards this average height and from this height modulates in relation to performance and musical parameters.

	Av.	Min	Max	Av. Nmus	Av. Mus	Av. EM
Whole	0.28	0.01	0.59	0.30	0.22	0.57
A	0.43	0.04	0.64	0.43	0.42	0.53
A'	0.51	0.1	0.76	0.53	0.46	0.43
B	0.32	0.02	0.70	0.36	0.25	0.43
A''	0.32	0.02	0.72	0.34	0.27	0.43
AC	0.35	0.02	0.73	0.38	0.29	0.58

Table 1: Average R^2s of the multiple regression analyses with key-velocity and quarter note IOI as independent variable and EM as dependent variable. See text for details.

Notably, the performance of the model increases considerably when it is fitted separately for different sections. This adaptation seems quite systematic over participants. Especially for p3, the parameters that significantly contribute to the explanation of variance change consistently with section: IOI is the only significant parameter for sections A and A'', while velocity is the significant parameter for the other sections with some minor exceptions for which none or both of the parameters are significant. The same pattern is found for the other performances, though a little less consistent. One of the reasons for this division could be that in sections A and A'' IOI varies relatively gradually in comparison to key-velocity besides the phrase-final lengthening at the end of the sections. This relates better to the gradual increase of EM in these sections than to the more strongly modulating key-velocity.

Notably as well is the better performance of the model for non-musicians than for musicians. While dynamics is a good predictor of EM for non-musicians, there seems to be more than tempo and dynamics to explain the EM of musicians.

The average EM does strongly correlate with the performance data and especially with key-velocity: the contour of the average EM agrees with the pattern of tension-release outlined by the pianist. For most participants, emotional intensity of engagement does decrease at the end. Some other participants show however an increase or static high emotional engagement from halfway the Coda to the end.

4. IMPROVEMENTS

There were two ways in which we tried to improve the explanation of EM using regression analyses. The first was by smoothing the performance data, since the continuous rating of emotional engagement varied less strongly than the performance variables. This increased all R^2s obtained.

The second improvement was to add some new variables to the regression model. From an extensive analysis of several cues extracted from the audio recordings and the musical score, we selected five cues that appeared most relevant in explaining the variance in EM's: 1) average loudness, extracted from the IPEM ear auditory model [6], which replaced the key velocity feature used in first experiments, 2) tempo, extracted from MIDI (as opposed to IOI's), 3) roughness, which is a measure of sensory dissonance of a sound and is considered to be highly related to texture perception [6], 4) polyphony and 5) melodic contour (both extracted from MIDI). The five features were used to construct regression models of EM's, as described in the previous section. Table 2 lists results of the best regression models obtained by using a combination of two, three and four features, respectively.

	# of feat	Av.	Min	Max	Av. NM	Av. M	Av. EM
Whole	2	0.52	0.24	0.75	0.80	0.57	0.80
	3	0.57	0.31	0.79	0.81	0.61	0.82
	4	0.59	0.33	0.80	0.81	0.63	0.83
A	2	0.82	0.61	0.94	0.86	0.91	0.88
	3	0.87	0.66	0.98	0.93	0.96	0.94
	4	0.90	0.71	0.98	0.95	0.97	0.96
A'	2	0.87	0.62	0.97	0.93	0.81	0.91
	3	0.91	0.74	0.98	0.96	0.86	0.95
	4	0.93	0.79	0.98	0.96	0.90	0.96
B	2	0.68	0.43	0.88	0.81	0.60	0.83
	3	0.74	0.53	0.91	0.85	0.68	0.86
	4	0.76	0.60	0.92	0.87	0.72	0.88
A''	2	0.77	0.38	0.96	0.77	0.75	0.79
	3	0.84	0.47	0.97	0.83	0.85	0.86
	4	0.86	0.52	0.98	0.86	0.90	0.91
AC	2	0.68	0.25	0.91	0.93	0.56	0.93
	3	0.74	0.37	0.94	0.95	0.61	0.95
	4	0.77	0.47	0.95	0.95	0.66	0.95

Table 2: Average R^2s of the multiple regression analyses with two, three or four features as independent variable and EM as dependent variable. See text for details.

ISBN 3-931852-67-9
ISSN 1617-6847

The explained variance increases considerably with the addition of new variables, but the effects are not always straightforward.

All features are relevant in EM's of listeners, but the relevance varies between different listeners and different sections of the performance. In regression models of entire performances, loudness and polyphony are the most prominent features. The effect of polyphony is alwayss negative: EM increases with decreasing polyphony. This seems more due to the specific compositional structure of the Skriabin piece than a general rule. The end of phrases and the middle part of the piece have relatively small texture, while the start of the piece has high texture. Tempo and roughness are the other two cues that have relevance in models of EM's for some listeners, while the melodic contour is not relevant.

Individual sections bring to life different relationships, though. In section A, the melodic contour is the most important feature, its effect usually negative, although the reverse is true in the response of one listener. Other cues follow, loudness and roughness being quite important, with the effect of roughness being negative (EM increases with decreasing roughness). Similar results were obtained in the analysis of the repeat of section A, with the melodic contour and loudness being the most important, the effect of melodic contour varying between listeners from positive to negative. In section B, loudness, polyphony and melodic contour are the most prominent features (in this order), while the latter two change places in the reentry of section A (A'') and in the Coda.

The effect of melodic contour is present in all sections, but can be both ways, depending on the listener and section, which again seems partly due to the compositional structure. The A sections have a falling melodic contour, but often increase in intensity. The effect of polyphony is always negative, while roughness varies, depending on its effect in the regression model. On average, tempo does not appear to have a large direct impact on EM's.

5. SUMMARY AND CONCLUSION

The pianist of this study used tempo mainly to communicate phrase boundaries and dynamics to communicate the tension curve of the music. The emotional engagement of listeners underlines this interpretation and was strongly correlated with loudness and the dynamics of the performance. Especially for non-musicians, dynamics accounted for the majority of variance in emotional engagement with higher dynamics corresponding to higher emotional engagement.

In an attempt to improve the explained variance for individual listeners, especially musicians, features of the musical structure and the audio recordings were added. Results were promising in showing high R^2s for combinations of two or three features. On the other hand, the effects of the features were highly variable for participants and depended on local context. The two features that appeared most strong besides loudness were polyphony and melodic contour.

6. REFERENCES

1. Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In Juslin, P., & Sloboda, J. A. (eds.), *Music and Emotion: theory and research* (pp. 223-248). Oxford: Oxford University Press.

2. Juslin, P. (2001). Communicating emotion in music performance: a review and a theoretical framework. In Juslin, P., & Sloboda, J. A. (eds.), *Music and Emotion: theory and research* (pp. 309-337). Oxford: Oxford University Press.

3. Krumhansl, C. L. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception, 13* (2), 401-432.

4. Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music, 19,* 110-120.

5. Sloboda, J. A., & Lehmann, A. C. (2001). Tracking performance correlates of changes of perceived intensity of emotion during different interpretations of a Chopin piano prelude. *Music Perception, 19.*

6. Leman, M., Lesaffre, M. and Tanghe, K. (2001). Introduction to the IPEM Toolbox, *Proceeding of the XIII Meeting of the FWO Research Society on Foundations of Music Research.*

ISBN 3-931852-67-9
ISSN 1617-6847