

## A Comparison of Dynamics in Major- and Minor-key Works

BEN TURNER

*Department of Psychology, Ohio State University*

DAVID HURON[1]

*School of Music, Ohio State University*

**ABSTRACT:** An analysis of dynamic markings in 24 paired preludes in major and minor keys shows that the minor mode is associated with a generally lower dynamic level than the major mode. The results are consistent with observations in speech prosody, where sad or depressed speakers show reduced acoustic energy compared with normal or happy speakers.

Submitted 2007 October 15; accepted 2008 January 6.

**KEYWORDS:** *music, minor key, sadness, loudness, speech prosody*

RESEARCH on speech prosody has shown that the average loudness of the voice is related to emotional state or affective intent. For instance, Yildirim et al. (2004) found RMS energy to be the best single parameter for distinguishing emotion classes (p. 2195). In fact, of the parameters studied, they found RMS energy to be the only parameter that differed significantly among all the emotion classes examined (p. 2196). In particular, considerable research suggests that loudness is effective in differentiating sadness from the other emotions: research in speech prosody has strongly implicated low intensity as a cue for sadness (Skinner, 1935; Scherer, 1986; Banse & Scherer, 1996; Cowie et al., 2001). For example, although they confound volume with rate of speech, Siegman and Boyle (1993) found that cardiovascular and self-rating responses were amplified when participants recounted sad or depressing events in 'slow and soft' speech, and suggested that the congruency between the vocal pattern and topic of conversation is responsible for this amplification. Likewise, Sobin and Alpert (1999) created 'acoustic profiles' for each of four major emotions (happiness, sadness, fear, anger), and found that sadness could be distinguished from the other emotions primarily along the dimension of loudness, along with pitch variance.

Cross-cultural studies have found significant differences in prosodic interpretation between cultures. However, the identification of 'sad voice' was found to exceed 70% when recordings of German actors were played in nine different countries on three continents (Scherer, Banse & Wallbott, 2001). After anger, sadness is the best-recognized vocal display in both Western and non-Western cultures (Scherer, 2003, p. 236). Thus, sadness seems to be widely recognizable, although these studies did not posit which characteristics of the sad speech examples rendered them 'sad.'

Parallel research on sadness has also been carried out in the psychology of music. Both adults and children readily identify particular melodies or passages as sounding 'sad' (Terwogt & Grinsven, 1991). Further, in examining the acoustic properties that render a particular musical excerpt 'sad' across the music of three cultures (Japanese, Hindustani, and Western) as judged by Japanese listeners, Balkwill, Thompson, and Matsunaga (2004) list decreased intensity as a consistent property. And in Western music, the minor mode has long been described by music scholars and listeners as having sad, dejected, or serious connotations (Heinlein, 1928; Hevner, 1935).

The research on vocal prosody suggests that sad affect will tend to be associated with lower overall loudness. Since the minor key in Western culture is commonly regarded as associated with sadness, one might predict that works in minor keys would tend to exhibit a lower overall loudness. In this study, we propose to test formally this hypothesis, namely that there is an association between the minor key and reduced loudness in Western music.

Before this hypothesis can be tested, a number of methodological issues must be addressed. Ideally, one would test this hypothesis by making direct measures of live performances in matched major

and minor keys using a sound pressure level meter at a standardized distance, but the number of elements that would have to be controlled (e.g., the concert hall, the ensemble or performer, the conductor, and so forth) render this option unfeasible. Existing sound recordings might also be thought to provide suitable material: for example, one could determine whether RMS amplitude (or volume units) is typically higher for recordings in major and minor keys. However, most recordings are produced so as to maximize the signal-to-noise ratio, and so recordings of nominally ‘quiet’ sounds are boosted in amplitude. In effect, modern recording practice tends to reduce the dynamic range found *in situ*.

Another possible approach might examine the notated dynamic markings in printed musical scores. That is, one could look at markings such as *ff*, *f*, *mf*, *mp*, *p*, *pp*, etc. While dynamic markings do not represent specific loudness levels, they nevertheless convey a general indication of overall loudness. It is not unreasonable to assume that the sequence of dynamic levels from pianissimo to fortissimo correlate with actual loudness levels in performance.

In order to realize this approach, one might take a random sample of musical scores in both major and minor keys. The range of potential dynamic markings (spanning from *pp* to *ff*) might be regarded as an ordinal scale. For analytic purposes, the markings *pp*, *p*, *mp*, *mf*, *f*, and *ff* might be recoded as numbers from 1 to 6. The distances between successive values may not be interpreted as equivalent in loudness difference, but the ordinal succession still allows the application of appropriate ordinal statistical methods. Although notated dynamics provide only a rough approximation of actual loudness levels, dynamic markings taken at random from individual musical scores in some sample, coded as numbers 1–6, might nevertheless provide an adequate basis on which to test the hypothesis.

A possible objection to a random sample of scores is that minor-key works represent not merely a different ‘mode’ from major-key works, but that minor-key works are biased toward particular genres or classes of music. As a hypothetical example, it might be the case that minor-key works are less likely to employ the voice—and it might also be the case that vocal music tends to be louder than works for solo instrument. Such circumstances might lead a researcher to erroneously conclude that there exists some direct causal connection between the minor key and quieter dynamics.

Ideally, our comparison of works in major and minor keys should involve pairs of works that differ only with respect to major/minor mode. A possible genre that may conform to this requirement is the ‘24-prelude set.’ Inspired by J.S. Bach's *Well-Tempered Clavier*, a number of Western composers have created sets of pieces composed in each of the 12 enharmonic major and minor keys. Typically, these pieces are arranged in pairs: C major/C minor, C# major/C# minor, etc. However, not all sets follow this paired arrangement, and of those that do, the mode that begins each pair may differ. In such sets, order effects cannot be ruled out. Nevertheless, we chose to use such 24-prelude sets as our musical sample.

## METHOD

Works were sampled according to the following procedure. Collected editions were consulted for eight composers: J.S. Bach, Busoni, Chopin, Clementi, Hertz, Hummel, Kalkbrenner, and Rachmaninoff. These composers were chosen, following Beurman (2003), as being writers of at least one 24-prelude set during the baroque, classical, or romantic eras. For each of these composers, volumes of keyboard works containing the 24-prelude set were consulted. In instances where the composer did not originally supply dynamics, we made use of dynamic markings added by a later editor. Such editorial markings are rightly regarded with suspicion by musicians motivated by authentic performance practice. In particular, the dynamic markings added by later editors often involve a more ‘romantic’ interpretation of earlier baroque or classical works. This confound notwithstanding, we elected to use editorial dynamic markings on the reasoning that the purpose of our investigation is to sample musically plausible dynamics, and the markings given by later editors minimally represent a plausible performance interpretation arising from one expert's musical intuitions. In short, our goal is neither to infer the composer's original intention, nor to establish the typical period-appropriate practice, but merely to sample a musical behavior.

The mode (major/minor) for each prelude was determined by examining the first and last chord and comparing these chords to the key signature. In all cases, this led to the same mode (key) as that indicated by the title of the work. Occasionally, a so-called Picardy Third occurs, where the last chord in a nominally minor-mode work is major rather than minor. Of the 24 preludes identified in minor keys, 23 included a first and final minor chord and the key signature was consistent with the relative major for these chords. One of the deemed minor-key works included a Picardy Third, but because this was the only

deviation in the piece, and because the final dynamic was not the one recorded (see below), this piece was still included in the minor-key sample.

Of the 192 possible pieces for sampling (8 composers  $\times$  24 pieces), a subset of just 48 pieces was sampled (6 major/minor pairs from each composer). Individual pieces were selected from the complete set for a single composer by generating a list of random numbers for each work (after predetermining how many pieces were to be used from a single set; care was taken to avoid over- or under-sampling any one set), with each number corresponding to one of the pieces in the set. For the very first piece chosen, the number was taken as randomly generated, and the first notated dynamic marking was recorded. From that point forward, subsequent numbers were chosen such that they conformed to the pattern {Major-Minor-Minor-Major}, in order to ensure a consistent pairing of major and minor works. In this way, any potential biases (such as the order of the pieces within the set or the date of composition) were avoided.

As an example, imagine a set containing 24 pieces, for which the list of generated random numbers was 13-22-9-4-15-7-21-18-12. We might decide *a priori* to take only six works from this set, and so would start by taking the 13<sup>th</sup> piece of the 24. Suppose this was the first set we examined, and also that minor and major key pieces alternated, starting with a major key piece as '1.' In this case, we would take the 13<sup>th</sup>, 22<sup>nd</sup>, 4<sup>th</sup>, 15<sup>th</sup>, 7<sup>th</sup>, and 18<sup>th</sup> pieces from the set, representing three each of major and minor and working through the list in order. As can be seen from this example, there is no bias as to where in the set either major or minor key pieces come from. We would then proceed to the next set, create a new random list, and take the first number from that list corresponding to a minor key work.

In Western music, it is common for works to depart from their nominal key through one or more modulations. In the minor key especially, modulation to the relative or parallel major key occurs frequently. Hence, the character or mood of a piece is prone change over its length. In order to minimize such potential confounds, we limited our sampling of dynamics to the first occurrence in the work (comprising the first instance of *pp* through *ff*). In this simple manner, 48 dynamic markings were recorded, 24 each from major and minor mode works.

## RESULTS

Data were of the following form: each observation was defined as occupying a point in a 2 (major|minor)  $\times$  6 (*pp*|*p*|*mp*|*mf*|*f*|*ff*) matrix. For our purposes, mode was considered a categorical variable, while dynamic could be considered ordinal in the weak case, or interval in the strong case. The data are presented in Table 1. Considering first the strong case, the dynamics *pp*..*ff* were scored as values 1..6, and a *t* test comparing the means of the two groups (major,  $M = 3.62$  and minor,  $M = 2.67$ ) was significant,  $t(46) = 2.29$ ,  $p = .027$ .

Considering next the weak case of an ordinal relation among the dynamic data, an initial chi-square analysis revealed no significant differences between minor and major mode pieces in terms of their distributions of dynamics,  $\chi^2(5, N = 48) = 5.68$ ,  $p > .3$ ; however, this test is unable to capture ordinal effects in the data.[2] Therefore, the dynamics were again scored as above, the modes were scored minor = 0 and major = 1, and a correlation was computed between the two groups,  $r = .319$ ,  $p = .027$ .[3] This correlation is interpretable as a chi-square statistic on one degree of freedom ( $M^2 = (N - 1)r^2$ ), which gives us  $M^2(1, N = 48) = 4.78$ ,  $p < .05$ . Furthermore, we can use this chi-square statistic along with the first chi-square statistic to determine the deviation from linearity of the relationship,  $\chi^2(4, N = 48) = 0.89$ ,  $p > .9$ .

With this result, we have license to compute a weaker chi-square statistic by collapsing the dynamics into two categories, '*p*' and '*f*.' While this is a much coarser view of the range of dynamics, it is plausible insofar as any dynamic that has '*p*' as its fundamental component is conceptually diametrically opposed to the same dynamic with '*f*' substituted. This test yields  $\chi^2(1, N = 48) = 4.27$ ,  $p = .04$ .

In examining the size of effect, Cramer's *V* can be used to measure the strength of relationship between the elements of our original chi-square, and gives us a result of 0.34 (0.30 for the collapsed test just discussed). With one degree of freedom along the smaller dimension, a Cramer's *V* of 0.34 represents a moderate size of effect, accounting for  $0.34^2 = 11.8\%$  of the variance observed (Rea & Parker, 1992).[4]

**Table 1.** Number of observations for each dynamic  $\times$  mode pairing.

| Dynamic/<br>Mode | <i>pp</i> | <i>p</i> | <i>mp</i> | <i>mf</i> | <i>f</i> | <i>ff</i> |
|------------------|-----------|----------|-----------|-----------|----------|-----------|
| Minor            | 2         | 14       | 2         | 2         | 4        | 0         |
| Major            | 1         | 9        | 1         | 2         | 9        | 2         |

## DISCUSSION AND CONCLUSION

Our analysis reveals a significant difference between major and minor works with respect to initial dynamic level, regardless of whether one holds to the strong argument that dynamics are (at least nearly) equidistantly distributed, or assumes only that they have an ordinal relationship, which is the stance taken in the present discussion. The analysis is consistent with an inverse linear relationship, with minor-mode works having a significantly lower average dynamic than major-mode works. Note several caveats, however: first, the effect size is rather small, as discussed above. While an association of 0.34 is moderate, it means that much of the cause of the observed pattern of dynamics is left unexplained. Second, results are surely driven by a linear trend in the minor mode dynamics, as results for major mode are much more ambiguous. Looking at Table 1, there is a clear skew toward lower dynamic levels for works in the minor mode, but for the major mode, the distribution is bimodal. This suggests that while minor-mode works are associated with lower initial dynamics, major-mode works are not necessarily associated with higher initial dynamics.

Both of these weaknesses likely come about for the following reason: we were investigating the relationship between the minor mode and sadness as though the two were inseparable, that is to say, we assumed that all pieces in a minor key were equally good representatives of sadness as portrayed musically. However, while it is true that the minor key is often associated with sadness in Western music, it is not always so, and any instance where the minor mode is associated with another emotion[5] can reasonably be expected to differ in some regards from that same piece had it been associated with sadness. This can explain why the correlation between mode and dynamic is not higher; any piece in the minor mode that is portraying an emotion characterized by high volume will weaken the correlation. Likewise, although we necessarily had to contrast the minor mode with the major mode, the latter is no more associated exclusively with 'loud' emotions than the former is with 'quiet' ones.

However, because the aim of our study is ultimately to show the association between the minor mode and soft dynamics, the peculiar pattern of dynamics in the major mode, while interesting, is perhaps not entirely relevant to our argument. Dynamics, after all, have a predefined midpoint, separating the '*p*' dynamics from the '*f*' dynamics, so the preponderance of dynamics conceptualized of as 'soft' associated with the minor mode is in support of our hypothesis. Future avenues for research exist in expanding on the points raised above, applying the idea of acoustic profiles from speech prosody research to the study of emotions in music.

## NOTES

[1] Please address all correspondence to: David Huron, School of Music, 1866 College Rd., Ohio State University, Columbus, OH, 43210.

[2] Instead, it treats each dynamic as a completely independent *nominal* group, each having no relation to or bearing on any of the others. This test merely tells us that there aren't significant differences between the number of observations of each dynamic for the two modes, rather than anything about the pattern of observations.

[3] This correlation and the earlier *t*-test are, in fact, mathematically equivalent, but as they represent different underlying assumptions made about the nature of the data, both are included.

[4] Because it has only one degree of freedom, Cramer's *V* here is equivalent to Cramer's  $\phi$ , which like Pearson's *r* can be squared to estimate proportion of variance accounted for.

[5] By this, we mean roughly that the composer wrote the piece to portray another emotion, and common consensus among Western listeners agrees with that emotion label.

## REFERENCES

- Balkwill, L-L, Thompson, W.F., & Matsunaga, R. (2004) Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners, *Japanese Psychological Research*, 46, 337–349.
- Banse, R., & Scherer, K. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70, 614–636.
- Beuerman, E. (2003). *The evolution of the twenty-four prelude set for piano*. D.M.A. dissertation. University of Arizona.
- Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. (2001). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, 18, 32–80.
- Heinlein, C. (1928). The affective characteristics of the major and minor modes in music. *Journal of Comparative Psychology*, 8, 101–142.
- Hevner, K. (1935). The affective character of the major and minor mode in music. *American Journal of Psychology*, 47, 103–118.
- Rea, L. M., & Parker, R. A. (1992). *Designing and conducting survey research*. San Francisco: Jossey-Boss.
- Scherer, K. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99, 143–165.
- Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40, 227–256.
- Scherer, K., Banse, R., & Wallbott, H.G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, 32, 76–92.
- Siegmán, A., & Boyle, S. (1993). Voices of fear and anxiety and sadness and depression: The effects of speech rate and loudness on fear and anxiety and sadness and depression. *Journal of Abnormal Psychology*, 102, 430–437.
- Skinner, E. (1935). A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness. *Speech Monographs*, 2, 81–137.
- Sobin, C., & Alpert, M. (1999). Emotion in speech: The attributes of fear, anger, sadness, and joy. *Journal of Psycholinguistic Research*, 28, 347–365.
- Terwogt, M.M., & Van Grinsven, F. (1991). Musical expression of moodstates. *Psychology of Music*, 19, 99–109.
- Yildirim, S., Bulut, M., Lee, C., Kazemzadeh, A., Busso, C., Deng, Z., Lee, S., & Narayanan, S. (2004). An acoustic study of emotions expressed in speech. In *Proceedings of the International Conference on Spoken Language Processing*. Jeju Island, Korea, October 2004, pp. 2193–2196.