ROYAL SOCIETY OPEN SCIENCE

Research

Author for correspondence:
Yong-Yeol Ahn
e-mail: yyahn@iu.edu

†Present address: Santa Fe Institute, Santa Fe, NM 87501, USA
‡These authors contributed equally to this work.

THE ROYAL SOCIETY PUBLISHING

The Minor fall, the Major lift: inferring emotional valence of musical chords through lyrics

Artemy Kolchinsky[1,†,‡], Nakul Dhande[1,2,‡],
Kengjeun Park[1,3] and Yong-Yeol Ahn[1]

[1]Department of Informatics, Indiana University, Bloomington, IN 47408, USA
[2]Amazon, Seattle, WA 98109, USA
[3]Epic, Verona, WI 53593, USA

Y-YA, 0000-0002-4352-4301

We investigate the association between musical chords and lyrics by analysing a large dataset of user-contributed guitar tablatures. Motivated by the idea that the emotional content of chords is reflected in the words used in corresponding lyrics, we analyse associations between lyrics and chord categories. We also examine the usage patterns of chords and lyrics in different musical genres, historical eras and geographical regions. Our overall results confirm a previously known association between Major chords and positive valence. We also report a wide variation in this association across regions, genres and eras. Our results suggest possible existence of different emotional associations for other types of chords.

1. Introduction

The power of music to evoke strong feelings has long been admired and explored [1–5]. Although music has accompanied humanity since the dawn of culture [6] and its underlying mathematical structure has been studied for many years [7–10], understanding the link between music and emotion remains a challenge [1,11–13].

The study of music perception has been dominated by methods that *directly* measure emotional responses, such as self-reports, physiological and cognitive measurements, and developmental observations [11,13]. Such methods may produce high-quality data, but the data collection process involved is both labour- and resource-intensive. As a result, creating large datasets and discovering statistical regularities has been a challenge.

Meanwhile, the growth of music databases [14–19] as well as the advancement of the field of Music Information Retrieval (MIR) [20–22] opened new avenues for data-driven studies of
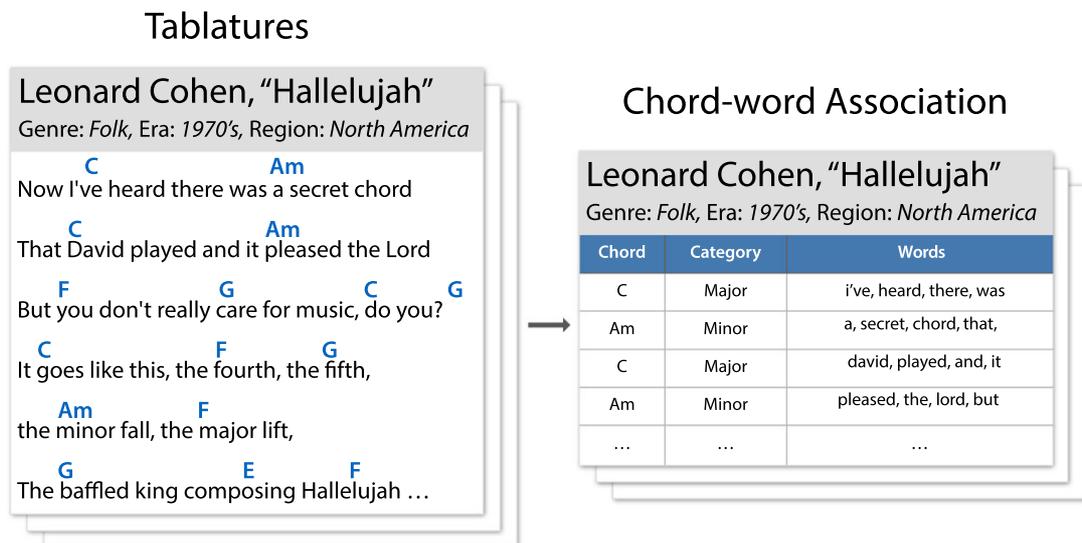
Tablatures

Leonard Cohen, "Hallelujah"
Genre: *Folk*, Era: *1970's*, Region: *North America*

     C   Am
Now I've heard there was a secret chord

     C   Am
That David played and it pleased the Lord

   F   G   C   G
But you don't really care for music, do you?

   C   F   G
It goes like this, the fourth, the fifth,

  Am   F
the minor fall, the major lift,

   G   E   F
The baffled king composing Hallelujah …

Chord-word Association

Leonard Cohen, "Hallelujah"
Genre: *Folk*, Era: *1970's*, Region: *North America*

Chord	Category	Words
C	Major	i've, heard, there, was
Am	Minor	a, secret, chord, that,
C	Major	david, played, and, it
Am	Minor	pleased, the, lord, but
…	…	…

Figure 1. A schematic of our process of collecting guitar tablatures and metadata, parsing chord–word associations and analysing the results using data mining and sentiment analysis. Note that 'genre' is an album-specific label, while 'era' and 'region' are artist-specific labels (rather than song-specific).

music. For instance, sentiment analysis [23–26] has been applied to uncover a long-term trend of declining valence in popular song lyrics [27,28]. It has been shown that the lexical features from lyrics [29–34], metadata [35], social tags [36,37] and audio-based features can be used to predict the mood of a song. There has been also an attempt to examine the associations between lyrics and individual chords using a machine translation approach, which confirmed the notion that Major and Minor chords are associated with happy and sad words respectively [38].

Here, we propose a novel method to study the associations between chord types and emotional valence. In particular, we use sentiment analysis to analyse *chord categories* (e.g. Major and Minor) and their associations with sentiment and words across genres, regions and eras.

To do so, we adopt a sentiment analysis method that uses the crowd-sourced *LabMT 1.0* valence lexicon [24,39]. Valence is one of the two basic emotional axes [13], with higher valence corresponding to more attractive/positive emotion. The lexicon contains valence scores ranging from 0.0 (saddest) to 9.0 (happiest) for 10 222 common English words obtained by surveying Amazon's Mechanical Turk workers. The overall valence of a piece of text, such as a sentence or document, is measured by averaging the valence score of individual words within the text. This method has been successfully used to obtain insight into a wide variety of corpora [40–43].

Here, we apply this sentiment analysis method to a dataset of guitar tablatures—which contain both lyrics and chords—extracted from ultimate-guitar.com. We collect all words that appear with a specific chord and create a large 'bag of words'—a frequency list of words—for each chord (figure 1). We perform our analysis by aggregating chords based on their 'chord category' (e.g. Major chords, Minor chords, Dominant 7th chords, etc.). In addition, we also acquire metadata from the Gracenote API regarding the genre of albums, as well as era and geographical region of musical artists. We then perform our analysis of associations between lyrics sentiment and chords within the categories of genre, era and region. Details of our methodology are described in the next section.

2. Material and methods

Guitar tabs were obtained from ultimate-guitar.com [18], a large online user-generated database of tabs, while information about album genre, artist era and artist region was obtained from Gracenote API [19], an online musical metadata service.

2.1. Chords–lyrics association

ultimate-guitar.com is one of the largest user-contributed tab archives, hosting more than 800 000 songs. We examined 123 837 songs that passed the following criteria: (i) we only kept guitar tabs and ignored

those for other instruments such as the ukulele; (ii) we ignored non-English songs (those having less than half of their words in an English word list [44] or identified as non-English by the `langdetect` library [45]); (iii) when multiple tabs were available for a song, we kept only the one with the highest user-assigned rating. We then cleaned the raw HTML sources and extracted chords and lyrics transcriptions. As an example, figure 1 shows how the tablature of Leonard Cohen's 'Hallelujah' [46] is processed to produce a chord-lyrics table.

Sometimes, chord symbols appeared in the middle of words; in such cases, we associated the entire words with the chord that appears in its middle, rather than the previous chord. In addition, chords that could not be successfully parsed or that had no associated lyrics were dropped.

2.2. Metadata collection using Gracenote API

We used the Gracenote API (http://gracenote.com) to obtain metadata about artists and song albums. We queried the title and the artist name of the 124 101 songs that were initially obtained from ultimate-guitar.com, successfully retrieving Gracenote records for 89 821 songs. Songs that did not match a Gracenote record were dropped. For each song, we extracted the following metadata fields:

— The geographic *region* from which the artist originated (e.g. *North America*). This was extracted from the highest-level geographic labels provided by Gracenote.
— The musical *genre* (e.g. *50s Rock*). This was extracted from the second-level genre labels assigned to each album by Gracenote.
— The historical *era* at the level of decades (e.g. *1970s*). This was extracted from the first-level era labels assigned to each artist by Gracenote. Approximately 6000 songs were not assigned to an era, in which case they were assigned to the decade of the album release year as specified by Gracenote.

In our analysis, we reported individual statistics only for the most popular regions (*Asia*, *Australia/Oceania*, *North America*, *Scandinavia*, *Western Europe*), genres (top 20 most popular genres), and eras (1950s through 2010s).

2.3. Determining chord categories

We normalized chord names and classified them into chord categories according to chord notation rules from an online resource [47]. All valid chord names begin with one or two characters indicating the root note (e.g. G or Bb) which are followed by characters which indicate the *chord category*. We considered the following chord categories [48]:

— **Major**: A Major chord is a triad with a root, a Major third and a perfect fifth. Major chords are indicated using either only the root note, or the root note followed by M or maj. For instance, F, FM, G, Gmaj were considered Major chords.
— **Minor**: A Minor chord is also a triad, containing a root, Minor third and a perfect fifth. The notation for Minor chords is to have the root note followed by m or min. For example, Emin, F#m and Bbm were considered Minor chords.
— **7th**: A seventh chord has seventh interval in addition to a Major or Minor triad. A **Major 7th** consists of a Major triad and an additional Major seventh, and is indicated by the root note followed by M7 or maj7 (e.g. GM7). A **Minor 7th** consists of a Minor triad with an additional Minor seventh, and is indicated by the root note followed by m7 or min7 (e.g. Fm7). A **Dominant 7th** is a diatonic seventh chord that consists of a Major triad with additional Minor seventh, and is indicated by the root note followed by the numeral 7 or dom7 (e.g. D7, Gdom7).
— **Special chords with '*'**: In tab notation, the asterisk '*' is used to indicate special instructions and can have many different meanings. For instance, G* may indicate that the G should be played with a palm mute, with a single strum, or some other special instruction usually indicated in free text in the tablature. Because in most cases the underlying chord is still played, in this study we map chords with asterisks to their respective non-asterisk versions. For instance, we consider G* to be the same as G and C7* to be the same as C7.
— **Other chords**: There were several other categories of chords that we do not analyse individually in this study. One of these is *Power* chords, which are dyads consisting of a root and a perfect fifth. Because Power chords are highly genre specific, and because they sometimes function musically
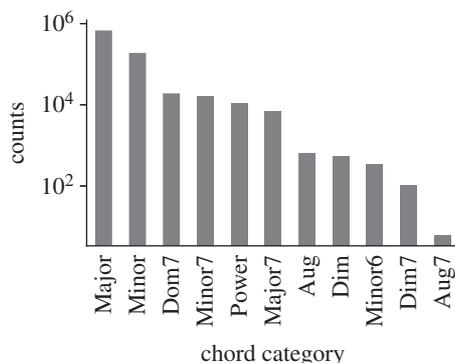
Figure 2. Prevalence of different chord categories within the dataset (note logarithmic scale).

as Minor or Major chords, we eliminated them from our study. For reasons of simplicity and statistical significance, we also eliminated several other categories of chords, such as *Augmented* and *Diminished* chords, which appeared infrequently in our dataset.

In total, we analysed 924 418 chords (see the next subsection). Figure 2 shows the prevalence of different chord categories among these chords.

2.4. Sentiment analysis

Sentiment analysis was used to measure the valence (happiness versus unhappiness) of chord lyrics. We employed a simple methodology based on a crowd-sourced valence lexicon (`LabMT 1.0`) [24,39]. This method was chosen because (i) it is simple and scalable, (ii) it is *transparent*, allowing us to calculate the contribution from each word to the final valence score and (iii) it has been shown to be useful in many studies [40–43]. The `LabMT 1.0` lexicon contains valence scores ranging from 0.0 (saddest) to 9.0 (happiest) for 10 222 common English words as obtained by surveying Amazon's Mechanical Turk workers. The valence assigned to some sequence of words (e.g. words in the lyrics corresponding to Major chords) was computed by mapping each word to its corresponding valence score and then computing the mean. Words not found in the LabMT lexicon were ignored; in addition, following recommended practices for increasing sentiment signal [24], we ignored emotionally neutral words having a valence strictly between 3.0 and 7.0. Chords that were not associated with any sentiment-mapped words were ignored. The final dataset contained 924 418 chords from 86 627 songs.

2.5. Word shift graphs

In order to show *how* a set of lyrics (e.g. lyrics corresponding to songs in the *Punk* genre) differs from the overall lyrics dataset, we use the *word shift graphs* [27,41]. We designate the whole dataset as the *reference* (baseline) corpus and call the set of lyrics that we want to compare *comparison corpus*. The difference in their overall valence can now be broken down into the contribution from each individual word. Increased valence can result from either having a higher *prevalence* (frequency of occurrence) of high-valence words or a lower prevalence of low-valence words. Conversely, lower valence can result from having a higher prevalence of low-valence words or a lower prevalence of high-valence words. The percentage contribution of an individual word i to the valence difference between a comparison and reference corpus can be expressed as:

$$100 \times \frac{\overbrace{\left(h_i - h^{(\text{ref})}\right)}^{+/-} \overbrace{\left(p_i - p_i^{(\text{ref})}\right)}^{\uparrow/\downarrow}}{\left|h^{(\text{comp})} - h^{(\text{ref})}\right|},$$

where h_i is the valence score of word i in the lexicon, $h^{(\text{ref})}$ and $h^{(\text{comp})}$ are the mean valences of the words in the reference corpus and comparison corpus respectively, p_i is the normalized frequency of word i in the comparison corpus, and $p_i^{(\text{ref})}$ is the normalized frequency of word i in the reference corpus (normalized frequencies are computed as $p_i = n_i / \sum_{i'} n_{i'}$, where n_i is the number of occurrences of word
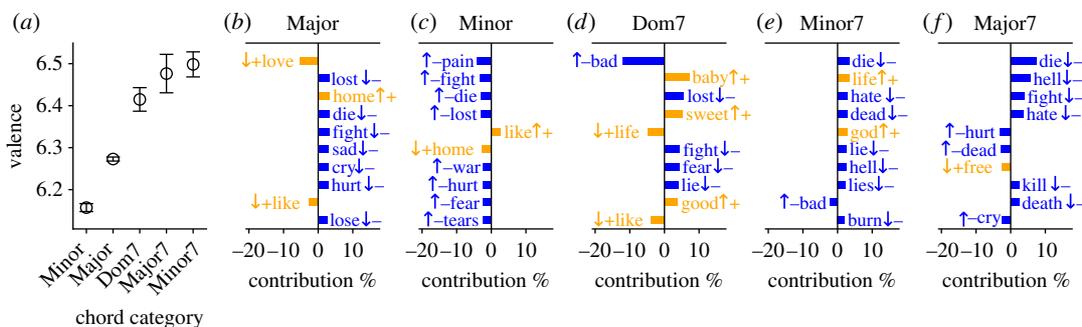
Figure 3. (*a*) Mean valence for chord categories. Error bars indicate 95% CI of the mean (error bars for Minor and Major chords are smaller than the symbols). (*b–f*) Word shift plots for different chord categories. High-valence words are indicated using '+' symbol and orange colour, while low-valence words are indicated by '−' symbol and blue colour. Words that are overexpressed in the lyrics corresponding to a given chord category are indicated by '↑', while underexpressed words are indicated by '↓'.

i). The first term (indicated by '+/−') measures the difference in word valence between word *i* and the mean valence of the reference corpus, while the second term (indicated by ↑ / ↓) looks at the difference in word prevalence between the comparison and reference corpus. In plotting the word shift graphs, for each word we use +/− signs and blue/orange bar colours to indicate the (positive or negative) sign of the valence term and ↑ / ↓ arrows to indicate the sign of the prevalence term.

2.6. Model comparison

In the Results section, we evaluate what explanatory factors (chord category, genre, era and region) best account for differences in valence scores. Using the `statsmodels` toolbox [49], we estimated linear regression models where the mean valence of each chord served as the response variable and the most popular chord categories, genres, eras and regions served as the categorical predictor variables. The variance of the residuals was used to compute the proportion of variance explained when using each factor in turn.

We also compared models that used combinations of factors. As before, we fit linear models that predicted valence. Now, however, explanatory factors were added in a greedy fashion, with each additional factor to minimize the Akaike information criterion (AIC) of the overall model.

3. Results

3.1. Valence of chord categories

We measure the mean valence of lyrics associated with different chord categories (figure 3*a*). We find that Major chords have higher valence than Minor chords, concurring with numerous studies which argue that human subjects perceive Major chords as more emotionally positive than Minor chords [50–52]. However, our results suggest that Major chords are not the happiest: all three categories of 7th chords considered here (Minor 7th, Major 7th and Dominant 7th) exhibit higher valence than Major chords. This effect holds with high significance ($p \ll 10^{-3}$ for all, one-tailed Mann–Whitney tests).

In figure 3*b–f*, we use the word shift graphs [27] to identify words that contribute most to the difference between the valence of each chord category and baseline (mean valence of all lyrics in the dataset). For instance, 'lost' is a lower-valence word (blue colour and '−' sign) that is underexpressed in Major chords ('↓' sign), increasing the mean valence of Major chords above baseline. Many negative words, such as 'pain', 'fight', 'die' and 'lost' are overexpressed in Minor chord lyrics and under-represented in Major chord lyrics.

Although the three types of 7th chords have similar valence scores (figure 3*a*), word shift graphs reveals that they may have different 'meanings' in terms of associated words. Overexpressed high-valence words for Dominant 7th chords include terms of endearment or affection, such as 'baby', 'sweet' and 'good' while for Minor 7th chords they are 'life' and 'god'. Lyrics associated with Major 7th chords, on the other hand, have a stronger under-representation of negative words (e.g. 'die' and 'hell').
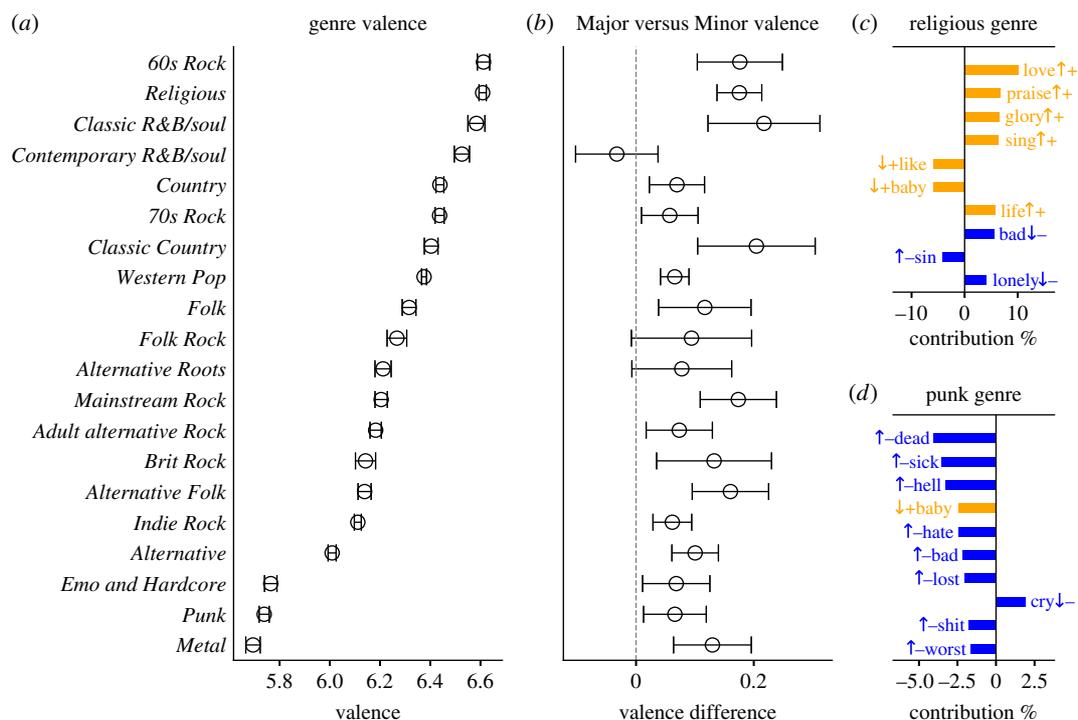
Figure 4. (*a*) Mean valence of lyrics by album genre. (*b*) Major versus Minor valence differences for lyrics by album genre. (*c*) Word shift plot for the *Religious* genre. (*d*) Word shift plot for the *Punk* genre. See caption of figure 3 for explanation of word shift plot symbols.

3.2. Genres

We analyse the emotional content and meaning of lyrics from albums in different musical genres. Figure 4*a* shows the mean valence of different genres, demonstrating that an intuitive ranking emerges when genres are sorted by valence: *Religious* and *60s Rock* lyrics reside at the positive end of the spectrum while *Metal* and *Punk* lyrics appear at the most negative.

As mentioned in the previous section, Minor chords have a lower mean valence than Major chords. We computed the numerical *differences* in valence between Major and Minor chords for albums in different genres (figure 4*b*). All considered genres, with the exception of *Contemporary R&B/Soul*, had a mean valence of Major chords higher than that of Minor chords. Some of the genres with the largest Major versus Minor valence differences include *Classic R&B/Soul*, *Classic Country*, *Religious* and *60s Rock*.

We show word shift graphs for two musical genres: *Religious* (figure 4*c*) and *Punk* (figure 4*d*). The highest contributions to the *Religious* genre come from the overexpression of high-valence words having to do with worship ('love', 'praise', 'glory', 'sing'). Conversely, the highest contributions to the *Punk* genre come from the overexpression of low-valence words (e.g. 'dead', 'sick', 'hell'). Some exceptions exist: for example, *Religious* lyrics underexpress high-valence words such as 'baby' and 'like', while *Punk* lyrics underexpress the low-valence word 'cry'.

3.3. Era

In this section, we explore sentiment trends for artists across different historical eras. Figure 5*a* shows the mean valence of lyrics in different eras. We find that valence has steadily decreased since the 1950s, confirming results of previous sentiment analysis of lyrics [27], which attributed the decline to the recent emergence of 'dark' genres such as metal and punk. However, our results demonstrate that this trend has recently undergone a reversal: lyrics have higher valence in the 2010s era than in the 2000s era.

As in the last section, we analyse differences between Major and Minor chords for lyrics belonging to different eras (figure 5*b*). Although Major chords have a generally higher valence than Minor chords, surprisingly this distinction does not hold in the 1980s era, in which Minor and Major chord valences are similar. The genres in the 1980s that had Minor chords with higher mean valence than Major chords—in other words, which had an 'inverted' Major/Minor valence pattern—include *Alternative Folk*, *Indie Rock* and *Punk* (data not shown).
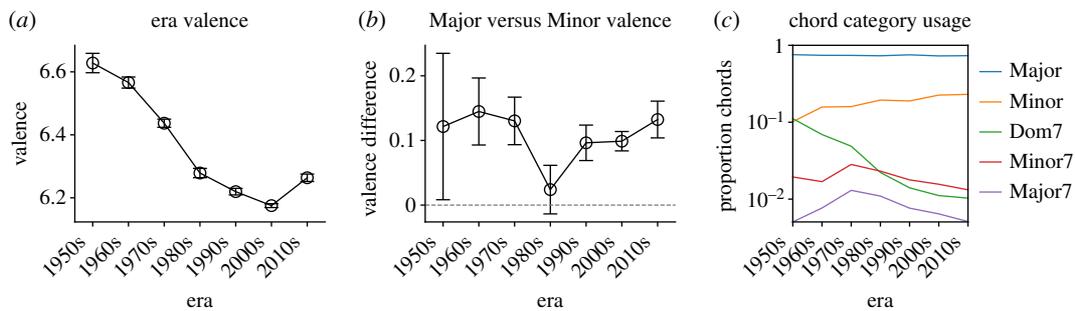
Figure 5. (a) Mean valence of lyrics by artist era. (b) Major versus Minor valence differences by artist era. (c) Proportion of chords in each chord category in different eras (note logarithmic scale).
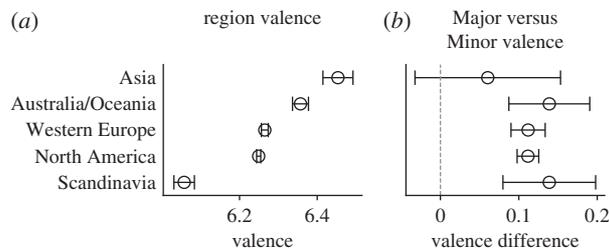

Figure 6. (a) Mean valence of lyrics by artist region. (b) Major versus Minor valence differences by artist region.

Finally, we report changes in chord usage patterns across time. Figure 5c shows the proportion of chords belonging to each chord category in different eras (note the logarithmic scale). Since the 1950s, Major chord usage has been stable while Minor chords usage has been steadily growing. Dominant 7th chords have become less prevalent, while Major 7th and Minor 7th chords had an increase in usage during the 1970s.

The finding that Minor chords have become more prevalent while Dominant 7th chords have become rarer agrees with a recent data-driven study of the evolution of popular music genres [53]. The authors attribute the latter effect to the decline in the popularity of blues and jazz, which frequently use Dominant 7th chords. However, we find that this effect holds widely, with Dominant 7th chords diminishing in prevalence even when we exclude genres associated with Blues and Jazz (data not shown). More qualitatively, musicologists have argued that many popular music styles in the 1970s exhibited a decline in the use of Dominant 7th chords and a growth in the use of Major 7th and Minor 7th chords [54]—clearly seen in the corresponding increases in figure 5c.

3.4. Region

In this section, we evaluate the emotional content of lyrics from artists in different geographical regions. Figure 6a shows that artists from Asia have the highest-valence lyrics, followed by artists from *Australia/Oceania*, *Western Europe*, *North America* and finally *Scandinavia*, the lowest valence geographical region. The latter region's low valence is probably due to the over-representation of 'dark' genres (such as metal) among Scandinavian artists [55].

As in previous sections, we compare differences in valence of Major and Minor chords for different regions (figure 6b). All regions except *Asia* have a higher mean valence for Major chords than Minor chords, while for the *Asian* region there is no significant difference.

There are several important caveats to our geographical analysis. In particular, our dataset consisted of only English-language songs, and is thus unlikely to be representative of overall musical trends in non-English speaking countries. This bias, along with others, is discussed in more depth in the Discussion section.

3.5. Model comparison

We have shown that mean valence varies as a function of chord category, genre, era and region (which are here called *explanatory factors*). We evaluate what explanatory factors best account for differences in
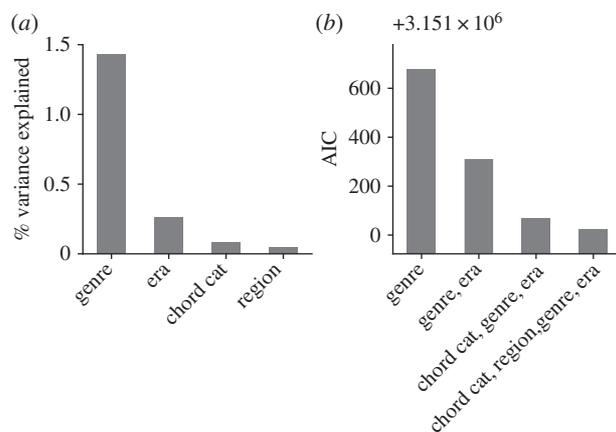
Figure 7. (*a*) Per cent category, genre, era and region as categorical predictor variables. (*b*) AIC model selection scores for models that predict valence using different explanatory factors. The model that includes all explanatory factors is preferred.

valence scores. Figure 7*a* shows the proportion of variance explained when using each factor in turn as a predictor of valence. This shows that genre explains most variation in valence, followed by era, chord category and finally region.

It is possible that variation in valence associated with some explanatory factors is in fact 'mediated' by other factors. For example, we found that mean valence declined from the 1950s era through the 2000s, confirming previous work [27] that explained this decline by the growing popularity of 'dark' genres like *Metal* and *Punk* over time; this is an example in which valence variation over historical eras is argued to actually be attributable to variation in the popularities of different genres. As another example, it is possible that Minor chords are lower valence than Major chords because they are overexpressed in dark genres, rather than due to their inherent emotional content.

We investigate this effect using statistical model selection. For instance, if the valence variation over chord categories can be entirely attributed to genre (i.e. darker genres have more Minor chords), then model selection should prefer a model that contains only the genre explanatory factor to the one that contains both the genre and chord category explanatory factors.

We fit increasingly large models while computing their Akaike information criterion (AIC) scores, a model selection score (lower is better). As figure 7*b* shows, the model that includes all four explanatory factors has the lowest AIC, suggesting that chord category, genre, era and region are all important factors for explaining valence variation.

4. Discussion

In this paper, we propose a novel data-driven method to uncover emotional valence associated with different chords as well as different geographical regions, historical eras and musical genres. We then apply it to a novel dataset of guitar tablatures extracted from ultimate-guitar.com along with musical metadata provided by the Gracenote API. We use word shift graphs to characterize the meaning of chord categories as well as categories of lyrics.

We find that Major chords are associated with higher valence lyrics than Minor chords, consistent with the previous music perception studies that showed that Major chords evoke more positive emotional responses than Minor chords [50–52,56]. For an intuition regarding the magnitude of the difference, the mean valence of Minor chord lyrics is approximately 6.16 (e.g. the valence of the word 'people' in our sentiment dictionary), while the mean valence of Major chord lyrics is approximately 6.28 (e.g. the valence of the word 'community' in our sentiment dictionary). Interestingly, we also uncover that three types of 7th chords—Dominant 7ths, Major 7ths and Minor 7ths—have even higher valence than Major chords. This effect has not been deeply researched, except for one music perception study which reported that, in contrast to our findings, 7th chords evoke emotional responses intermediate in valence between Major and Minor chords [57].

Significant variation exists in the lyrics associated with different geographical regions, musical genres and historical eras. For example, musical genres demonstrate an intuitive ranking when ordered by mean valence, ranging from low-valence *Punk* and *Metal* genres to high-valence *Religious* and *60s Rock* genres.

We also found that sentiment declined over time from the 1950s until the 2000s. Both of these findings are consistent with the results of a previous study conducted using a different dataset and lexicon [27]. At the same time, we report a new finding that the trend in declining valence has reversed itself, with lyrics from the 2010s era having higher valence than those from the 2000s era. Finally, we perform an analysis of the variation of valence among geographical regions. We find that *Asia* has the highest valence while *Scandinavia* has the lowest (probably due to prevalence of 'dark' genres in that region).

We perform a novel data-driven analysis of the Major versus Minor distinction by measuring the *difference* between Major and Minor valence for different regions, genres and eras. All examined genres except *Contemporary R&B/Soul* exhibited higher Major chord valence than Minor chord. Interestingly, the largest differences of Major above Minor may indicate genres (*Classic R&B/Soul*, *Classic Country*, *Religious*, *60s Rock*) that are more musically 'traditional'. In terms of historical periods, we find that, unlike other eras, the 1980s era did not have a significant Major–Minor valence difference. This phenomenon calls for further investigation; one possibility is that it may be related to an important period of musical innovation in 1980s, which was recently reported in a data-driven study of musical evolution [53]. Finally, analysis of geographic variation indicates that songs from the *Asian* region—unlike those from other regions—do not show a significant difference in the valence of Major versus Minor chords. In fact, it is known in the musicological literature that the association of positive emotions with Major chords and negative emotions with Minor chords is culture-dependent, and that some Asian cultures do not display this association [58]. Our results may provide new supporting evidence of cultural variation in the emotional connotations of the Major/Minor distinction.

Finally, we evaluate how much of the variation in valence in our dataset is attributable to chord category, genre, era and region (we call these four types of attributes 'explanatory factors'). We find that genre is the most explanatory, followed by era, chord category and region. We use statistical model selection to evaluate whether certain explanatory factors 'mediate' the influence of others (an example of mediation would be if variation in valence of different eras is actually due to variation in the prevalence of different genres during those eras). We find that all four explanatory factors are important for explaining variation in valence; no explanatory factors totally mediate the effect of others.

Our approach has several limitations. First, the accuracy of tablature chord annotations may be limited because users are not generally professional musicians; for instance, more complex chords (e.g. `dim` or `11th`) may be mis-transcribed as simpler chords. To deal with this, we analyse relatively basic chords—Major, Minor and 7ths—and, when there are multiple versions of a song, use tabs with the highest user-assigned rating. Our manual inspection of a small sample of parsed tabs indicated acceptable quality, although a more systematic evaluation of the error rate can be performed using more extensive manual inspection of tabs by professional musicians.

There are also significant biases in our dataset. We only consider tablatures for songs entered by users of ultimate-guitar.com, which is likely to be heavily biased towards North American and European users and songs. In addition, this dataset is restricted to songs playable by guitar, which selects for guitar-oriented musical genres and may be biased toward emotional meanings tied to guitar-specific acoustic properties, such as the instrument's timbre. Furthermore, our dataset includes only English-language songs and is not likely to be a representative sample of popular music from non-English speaking regions. Thus, for example, the high valence of songs from *Asia* should not be taken as conclusive evidence that Asian popular music is overall happier than popular music in English-speaking countries. For this reason, the absence of a Major versus Minor chord distinction in *Asia* is speculative and requires significant further investigation.

Despite these limitations, we believe that our results reveal meaningful patterns of association between music and emotion at least in guitar-based English-language popular music, and show the potential of our novel data-driven methods. At the same time, applying these methods to other datasets—in particular those representative of other geographical regions, historical eras, instruments and musical styles—is of great interest for future work. Another promising direction for future work is to move the analysis of emotional content beyond single chords, since emotional meaning is likely to be more closely associated with melodies rather than individual chords. For this reason, we hope to extend our methodology to study chord progressions.

Authors' contributions. N.D. and Y.Y.A. conceived the idea of analysing the association between lyric sentiment and chord categories, while A.K. contributed the idea of analysing the variation of this association across genres, eras and

References

1. Meyer LB. 1961 *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.

2. Dainow E. 1977 Physical effects and motor responses to music. *J. Res. Music Educ.* **25**, 211–221. (doi:10.2307/3345305)

3. Garrido S, Schubert E. 2011 Individual differences in the enjoyment of negative emotion in music: a literature review and experiment. *Music Percept. Interdiscip. J.* **28**, 279–296. (doi:10.1525/mp.2011.28.3.279)

4. Cross I. 2001 Music, cognition, culture, and evolution. *Ann. N. Y. Acad. Sci.* **930**, 28–42. (doi:10.1111/j.1749-6632.2001.tb05723.x)

5. Zentner M, Grandjean D, Scherer KR. 2008 Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion* **8**, 494–521. (doi:10.1037/1528-3542.8.4.494)

6. Wallin N, Merker B, Brown S. 2001 *The origins of music*. A Bradford book. Cambridge, MA: MIT Press.

7. Fauvel J, Flood R, Wilson R. 2003 *Music and mathematics: from Pythagoras to fractals*. Oxford, UK: Oxford University Press.

8. Tymoczko D. 2006 The geometry of musical chords. *Science* **313**, 72–74. (doi:10.1126/science.1126287)

9. Callender C, Quinn I, Tymoczko D. 2008 Generalized voice-leading spaces. *Science* **320**, 346–348. (doi:10.1126/science.1153021)

10. Helmholtz HL. 2009 *On the sensations of tone as a physiological basis for the theory of music*. Cambridge, UK: Cambridge University Press.

11. Hunter PG, Schellenberg EG. 2010 Music and emotion. *Music Percept.* **36**, 129–164.

12. Juslin PN, Västfjäll D. 2008 Emotional responses to music: the need to consider underlying mechanisms. *Behav. Brain Sci.* **31**, 559–575. (doi:10.1017/S0140525X08005293)

13. Eerola T, Vuoskoski JK. 2013 A review of music and emotion studies: approaches, emotion models, and stimuli. *Music Percept. Interdiscip. J.* **30**, 307–340. (doi:10.1525/mp.2012.30.3.307)

14. International Music Score Library Project. See http://imslp.org.

15. last.fm. See http://last.fm.

16. Bertin-Mahieux T, Ellis DPW, Whitman B, Lamere P. 2011 The million song dataset. In *Proc. of the 12th Int. Society for Music Information Retrieval Conf. (ISMIR2011), Miami, FL, 24–28 October*. International Society for Music Information Retrieval.

17. Huron D. 2002 Music information processing using the Humdrum Toolkit: concepts, examples, and lessons. *Comput. Music J.* **26**, 11–26. (doi:10.1162/014892602760137158)

18. ULTIMATE GUITAR TABS. See http://ultimate-guitar.com.

19. Gracenote. See http://gracenote.com.

20. Downie JS. 2008 The music information retrieval evaluation exchange (2005–2007): a window into music information retrieval research. *Acoust. Sci. Technol.* **29**, 247–255. (doi:10.1250/ast.29.247)

21. Kim YE, Schmidt EM, Migneco R, Morton OG, Richardson P, Scott J, Speck JA, Turnbull D. 2010 Emotion recognition: a state of the art review. In *11th Int. Society for Music Information and Retrieval Conf., Utrecht, The Netherlands, 9–13 August*. International Society for Music Information Retrieval.

22. Yang YH, Chen HH. 2012 Machine recognition of music emotion: a review. *ACM Trans. Intell. Syst. Technol.* **3**, 40. (doi:10.1145/2168752.2168754)

23. Pang B, Lee L. 2008 Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* **2**, 1–135. (doi:10.1561/1500000011)

24. Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM. 2011 Temporal patterns of happiness and information in a global social network: hedonometrics and Twitter. *PLoS ONE* **6**, e26752. (doi:10.1371/journal.pone.0026752)

25. Liu B. 2012 Sentiment analysis and opinion mining. *Synth. Lect. Hum. Lang. Technol.* **5**, 1–167. (doi:10.2200/S00416ED1V01Y201204HLT016)

26. Gonçalves P, Araújo M, Benevenuto F, Cha M. 2013 Comparing and combining sentiment analysis methods. In *Proc. of the 1st ACM Conf. on Online Social Networks, Boston, MA, 7–8 October*, pp. 27–38. ACM.

27. Dodds PS, Danforth CM. 2010 Measuring the happiness of large-scale written expression: songs, blogs, and presidents. *Journal of Happiness Studies* **11**, 441–456. (doi:10.1007/s10902-009-9150-9)

28. DeWall CN, Pond Jr RS, Campbell WK, Twenge JM. 2011 Tuning in to psychological change: linguistic markers of psychological traits and emotions over time in popular US song lyrics. *Psychol. Aesthet. Creat. Arts* **5**, 200–207. (doi:10.1037/a0023195)

29. Yang D, Lee WS. 2004 Disambiguating music emotion using software agents. In *ISMIR, Barcelona, Spain, 10–14 October*, vol. 4, pp. 218–223. International Society for Music Information Retrieval.

30. Yang YH, Lin YC, Cheng HT, Liao IB, Ho YC, Chen HH. 2008 Toward multi-modal music emotion classification. In *Advances in Multimedia Information Processing-PCM 2008, Tainan, Taiwan, 9–13 December*, pp. 70–79. Springer.

31. Laurier C, Grivolla J, Herrera P. 2008 Multimodal music mood classification using audio and lyrics. In *7th Int. Conf. on Machine Learning and Applications, 2008. ICMLA'08, San Diego, CA, 11–13 December*, pp. 688–693. IEEE.

32. Hu X, Downie JS, Ehmann AF. 2009 Lyric text mining in music mood classification. *Am. Music* **183**, 2–209.

33. Van Zaanen M, Kanters P. 2010 Automatic mood classification using TF* IDF based on lyrics. In *ISMIR*, pp. 75–80.

34. Mihalcea R, Strapparava C. 2012 Lyrics, music, and emotions. In *Proc. of the 2012 Joint Conf. on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, EMNLP-CoNLL '12*, pp. 590–599. Stroudsburg, PA, USA: Association for Computational Linguistics.

35. Schuller B, Dorfner J, Rigoll G. 2010 Determination of nonprototypical valence and arousal in popular music: features and performances. *EURASIP J. Audio Speech Music Process.* **2010**, 5. (doi:10.1155/2010/735854)

36. Turnbull DR, Barrington L, Lanckriet G, Yazdani M. 2009 Combining audio content and social context for semantic music discovery. In *Proc. of the 32nd Int. ACM SIGIR Conf. on Research and Development in Information Retrieval, Boston, MA, 19–23 July*, pp. 387–394. ACM.

37. Bischoff K, Firan CS, Paiu R, Nejdl W, Laurier C, Sordo M. 2009 Music mood and theme classification: a hybrid approach. In *ISMIR, Kobe, Japan, 26–30 October*, pp. 657–662. International Society for Music Information Retrieval.

38. O'Hara T, Schüler N, Lu Y, Tamir D. 2012 Inferring chord sequence meanings via lyrics: process and evaluation. In *ISMIR, Porto, Portugal, 8–12 October*, pp. 463–468. International Society for Music Information Retrieval.

39. Bradley MM, Lang PJ. 1999 Affective norms for English words (ANEW): instruction manual and affective ratings. Technical Report Citeseer.

40. Bliss CA, Kloumann IM, Harris KD, Danforth CM, Dodds PS. 2012 Twitter reciprocal reply networks exhibit assortativity with respect to happiness. *J. Comput. Sci.* **3**, 388–397. (doi:10.1016/j.jocs.2012.05.001)

41. Mitchell L, Frank MR, Harris KD, Dodds PS, Danforth CM. 2013 The geography of happiness: connecting Twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE* **8**, e64417. (doi:10.1371/journal.pone.0064417)

42. Kloumann IM, Danforth CM, Harris KD, Bliss CA, Dodds PS. 2012 Positivity of the English language. *PLoS ONE* **7**, e29484. (doi:10.1371/journal.pone.0029484)

43. Frank MR, Mitchell L, Dodds PS, Danforth CM. 2013 Happiness and the patterns of life: a study of geolocated tweets. *Sci. Rep.* **3**, 2625. (doi:10.1038/srep02625)

44. SIL-International English Wordlists. See http://www-01.sil.org/linguistics/wordlists/english/wordlist/wordsEn.txt.

45. Danilak M. 2014 langdetect: Language detection library ported from Google's language detection. See https://pypi.python.org/pypi/langdetect/ (accessed 19 January 2015).

46. Ultimate-Guitar "Hallelujah" by Leonard Cohen (ver 6). See http://tabs.ultimate-guitar.com/l/leonard_cohen/hallelujah_ver6_crd.htm.

47. Wright H. Howard's big list of guitar chords. See http://www.hakwright.co.uk/guitar chords/.

48. Benward B, Saker M. 2007 *Music in theory and practice*. Number v. 1 in Music in Theory and Practice. New York, NY: McGraw-Hill.

49. StatsModels: statistics in Python. See http://statsmodels.sourceforge.net/.

50. Crowder RG. 1984 Perception of the major/minor distinction: I. Historical and theoretical foundations. *Psychomusicology: J. Res. Music Cogn.* **4**, 3–12. (doi:10.1037/h0094207)

51. Kastner MP, Crowder RG. 1990 Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Percept. Interdiscip. J.* **8**, 189–201. (doi:10.2307/40285496)

52. Hunter PG, Schellenberg EG, Schimmack U. 2010 Feelings and perceptions of happiness and sadness induced by music: similarities, differences, and mixed emotions. *Psychol. Aesthet. Creat. Arts* **4**, 47–56. (doi:10.1037/a0016873)

53. Mauch M, MacCallum RM, Levy M, Leroi AM. 2015 The evolution of popular music: USA 1960–2010. *R. Soc. Open Sci.* **2**, 150081. (doi:10.1098/rsos.150081)

54. Stephenson K. 2002 *What to listen for in rock: a stylistic analysis*. New Haven, CT: Yale University Press.

55. Scaruffi P. 2009 *A history of rock and dance music*, vol. 2. Silicon Valley, CA: Omniware.

56. Crowder RG, Reznick JS, Rosenkrantz SL. 1991 Perception of the major/minor distinction: V. Preferences among infants. *Bull. Psychon. Soc.* **29**, 187–188. (doi:10.3758/BF03342673)

57. Lahdelma I, Eerola T. 2014 Single chords convey distinct emotional qualities to both naïve and expert listeners. *Psychol. Music* **44**, 37–54. (doi:10.1177/0305735614552006)

58. Malm WP. 2000 *Traditional Japanese music and musical instruments*, vol. 1. Tokyo, Japan: Kodansha international.