

Comparing Word Affect and Tone Affect: Comment on Sun and Cuthbert 2017

YKE SCHOTANUS [1]
Utrecht University

ABSTRACT: In the article “Emotion Painting: Lyric, affect, and musical relationships in a large lead-sheet corpus”, Sun and Cuthbert (2017) explored the correlations between affect-carrying lyrics and musical features such as beat strength, pitch height, consonance, and mode. Several musical features did indeed turn out to be highly correlated with the affect of the lyrics. However, correlations between other features, particularly mode-related musical features and lyric affect, were either insignificant or even contradicted previous research. In the current commentary, it is argued that the difference between the musical features that show significant correlations and those that do not is that the former have a local musical effect whereas the latter tend to affect the mood of a whole phrase or piece, and that the way Sun and Cuthbert estimate lyric affect for sentences or song may not be appropriate. Furthermore, a few remarks are made about the way Sun and Cuthbert treat multi-syllable words and about some basic assumptions concerning the relation between music and lyrics in a song. Nevertheless, the authors are praised for their innovative and interesting work, while several alternative and additional analyses (for example with scale-degree qualia and syncopations) are proposed.

Submitted 2020 May 1; accepted 2020 August 6.

Published 2020 October 22; <https://doi.org/10.18061/emr.v15i1-2.7676>

KEYWORDS: *music, language, emotion, affect, mode, syncopation*

THE connection between music and lyrics is only scarcely investigated in empirical musicology. This is probably due to the fact that ever since Hanslick (1854/2008) criticized the so-called *Affektenlehre*, the ability of music to convey intended emotions or semantic associations has been debated. According to Hanslick, the interpretation of musical emotion and musical meaning is purely individual. Another explanation could be that, in line with music sociologists such as Robinson and Hirsch (1969a, b) and Frith (1987/2007), lyrics are thought to be irrelevant. However, recent research into emotions and musical meaning has shown that the emotions perceived in response to music and the semantic associations evoked by it are not as individual as Hanslick once suggested (e.g., Cespedes-Guevara & Eerola, 2018; Huovinen & Kaila, 2015; Juslin & Laukka, 2003; Schubert, 2013). Moreover, a substantial percentage of listeners take notice of song lyrics (e.g., Condit-Schulz & Huron, 2015; Engels, Slettenhaar, Ter Bogt, & Scholte, 2011; Ter Bogt, Canale, Lenzi, Vieno, & Van den Eijnden, 2019; Schäfer & Eerola, 2020; Schotanus, 2020, p. 135).

To date, most studies that do investigate the relationship between music and lyrics mainly compare overall ratings for the music, or for the song as a whole, with overall ratings for the isolated lyrics (e.g., Ali & Peynircioğlu, 2006; see Schotanus, 2020a, for a review). Yet, there are a few exceptions: Paul and Huron (2010), for example, found a correlation between grief-related words and a breaking voice in country music, and Strykowski (2016) and Lousberg (2018) found correlations between specific pitch patterns and the content of the words aligned with those pitches in madrigals and Gregorian chant, respectively. In addition, Schotanus (2018, in preparation) shows that local musical features (such as an out-of-key note, or a syncopation) can affect the interpretation of a sentence. However, a large study, involving several musical features and more than a thousand songs, had not yet been conducted. Sun and Cuthbert (2017) represents such a study and is therefore a valuable contribution to the research literature.



THE PAPER

In their study, Sun and Cuthbert (2017) investigated the correlations between beat strength, note length, pitch height, consonance, major-minor chord context, tonal certainty, and mode on the one hand, and words with specific affects on the other in a large lead-sheet corpus of 1985 songs. The analysis of word affect was based on the NRC Word-Emotion Association Lexicon (EmoLex, Mohammad & Turney, 2010, 2013). Significant correlations were found between word affect and all musical features except mode. Yet, those between beat strength and note length were more frequently significant and had substantially lower *p*-values than those between word affect and pitch height, consonance, major-minor context, and tonal certainty. Sun and Cuthbert compared their results with the findings of Gabrielson and Lindström (2010) and found that most of the correlations were in line with these findings, except for the fact that happy words turned out not to be combined more often with major chords than sad ones and that there was no effect of mode.

Although Hansen (2018) may be right that issues with multiple comparisons weaken the results for pitch height, consonance, major-minor context, and tonal certainty, I still think these results are interesting. However, like Hansen, I think future research should attempt to improve the analyses by working with specific hypotheses and by distinguishing more clearly between strictly local effects and overall effects on phrase or song level. Hansen has already elaborated on the possibility that pitch height should be studied on phrase level, and several studies show that there is at least an overall effect of pitch height on song level (Huron, Kinney, & Precoda, 2006; Shanahan & Huron, 2014). In addition, whenever the music is analysed at levels above word level, the language should be analysed at levels above word level as well, and I do not think the way Sun and Cuthbert have tried to do that is appropriate. In order to compare the mode of an entire song with the mode of the lyrics, they have estimated the extent to which the lyrics for a specific song conveys a certain affect by counting the occurrences of this affect within these lyrics. However, song or sentence sentiment is not the sum of word sentiments. For example, ‘Once I was perfectly happy’, predominantly consists of happy words, but it is a sad sentence. Conversely, ‘Hitler is dead!’, has been a joyous message for many people, yet word sentiment is rather negative. Sun and Cuthbert have substantiated their methods by referring to Hu, Chen, and Yang (2009), among others, who have been able to calculate song mood correctly for a large number of songs. However, these authors did not simply add up word sentiments, but rather conducted complex computations in which they reckoned with several aspects of language that could modify the impact of word affect on song affect. A less complex way to establish lyric affect for a whole song or song part is to ask participants to rate the lyrics as a whole. Tiemann and Huron (2010) did so and subsequently found clear correlations between the ratings for lyrics and the ratings for music composed for the same songs.

A promising approach to assessing the correlations between lyric affect and musical affect on the word level would be to compare scale-degree qualia (Arthur, 2018) with word sentiment. Yet, in doing so, one would have to reckon with note position within the measure. Specifically, Sun and Cuthbert already noted that what they call “stopwords” (frequently used neutral words such as “the”) are more often combined with dissonant notes than affect-carrying words, whereas neutral words are less often combined with dissonant notes than affect-carrying words, particularly sad and surprising ones. At first glance, the latter difference makes more sense than the former. However, as Sun and Cuthbert rightly pointed out, stopwords may often be combined with passing notes, which usually occur on weak beats. The dissonance or the instability of those notes (which is supposed to cause the association with an affect) will therefore be less salient, or even negligible. Consequently, notes in relatively weak positions and the words or syllables combined with them should be left aside or weighed less heavily than notes and words or syllables occurring in metrically strong positions, unless such notes occurring at weak positions are indeed syncopated.

Mentioning “syllables” and “syncopations” brings to the fore two other important issues. The word “syllables” refers to an issue that has remained underexposed in Sun and Cuthbert’s study. As many words in the NRC Emotion Lexicon consist of two or more syllables, the question arises as to how Sun and Cuthbert dealt with multi-syllable words. Unfortunately, their statement on this issue is quite ambiguous. They state that “ideally, the note attached to the accented syllable, or an average of the properties of all notes contained within the word would be used” (p. 332) but also note that in this study, given the relatively small number of multi-syllable words, “the assumption that the first note will carry affect-music value is acceptable” (p. 332). If this means that they only took into account the first note combined with a word, this may have confounded their results more than they think. First, it cannot be excluded that affect-carrying

words are more often multi-syllabic than other words; second, several multi-syllabic words start with an unstressed syllable; third, in cases of epenthesis or melisma, the affect-music is partly in the added notes; and fourth, in longer words syllables with secondary stress and even unstressed syllables can be used to express affect. For example, in ‘Amazing grace’, neither the first note nor the note combined with the stressed syllable “-ma-” is a note that conveys the maximum of affect in the word “amazing”. Even though word-stress rules are fully respected, the syllable “-zing” is accentuated in a very expressive way, almost overpowering “-ma-”. Also, the first syllable is accentuated in an expressive way by combining it with a melisma, delaying the rest of the word in a manner that illustrates the sense of awe the word expresses. I think it is therefore more appropriate to assume that all the notes associated with stressed syllables plus all the notes that are in metrically strong positions or that are accentuated otherwise should be considered the affect-carrying notes for a multi-syllable word. Furthermore, both melismas and epentheses should be investigated as a distinguished affect-carrying musical feature.

Concerning syncopations, according to Brackett (1995), Billy Holiday’s rendition of ‘I’ll be seeing you’ is much sadder than Bing Crosby’s rendition, partly because of Holiday’s syncopations. He is not the only musicologist to suggest that syncopations support the emotional expressiveness of a song. For example, Pattison (2015, March 17) has claimed that off-beat phrasings suggest that there is some subtext to the lyrics. Therefore, it appears to be very interesting to assess whether there are correlations between the occurrence of syncopations and of affect-carrying words. However, this should be done with care. Recent publications show that it is important to distinguish between syncopations that matter and those that will hardly be perceived as such (Condit-Schulz, 2019; Koops, Volk, & De Haas, 2015; Tan, Lustig, & Temperley, 2019; Temperley, 2020). Furthermore, among the syncopations that do matter, the position relative to the beat can influence musical affect. At least, Schotanus (in preparation) finds evidence that stressed syllables occurring either early or late compared to the beat support the impression that the singer is either pressing on or holding back for some reason, but that early ones more often support a sense of urgency, whereas late ones more often support a sense of being upset.

WORDS AND TONES: A FEW THOUGHTS ON TEXTSETTING

A study such as Sun and Cuthbert’s (2017) presupposes not only that specific musical features can convey specific emotions or semantic connotations, but also that, consciously or not, lyricists and composers tend to match affect-carrying words with tones conveying similar affects. Moreover, the title of the study, ‘Emotion painting’, suggests that lyricists and composers do so in order to illustrate the content of their words with music. I am not sure whether this is always the case. First, in many cases, music and lyrics will be matched for the first verse, the chorus and the bridge whereas the second, third, and all subsequent verses are just sung to the melody of the first one. Second, several theories about the relationship between music and lyrics in a song assume that the semantic or emotional content of music and lyrics do not merely coexist in a song but rather interact in relatively complicated ways (see Schotanus, 2020, pp. 40-51). Third, in line with that, composers and lyricists who appear to have consciously matched words and lyrics of every part of the song (such as the ‘poets of Tin Pan Alley’, cf. Furia, 1990, or classical composers such as Mozart, Schubert, and Beethoven), may also juxtapose word affect and musical affect from time to time (Hatten, 2004). They may also try to enrich one affect with another, to denaturalize neutral words, or simply to create prosodic accents signalling the relative importance of specific words or highlighting their ambiguity. A tonic chord, for example, seems to function as a punctuation mark (Schotanus, Koops, & Reed Edworthy, 2018), and out-of-key notes appear to cause an N400 response (a brain potential indicating special interest in the meaning of a word, Schotanus, Eekhof, & Willems, 2018). Accordingly, if they are aligned to ambiguous words, out-of-key notes support ironic, metaphoric, or highly emotional interpretations of the sentences in which these words occur (Schotanus, 2018).

Considering these issues, the absence of significant correlations between affect-carrying words and supposedly affect-carrying musical features does not mean that composers and lyricists do not use these musical features to manipulate word affect, nor that these features do not carry such an affect at all. It is all the more remarkable that Sun and Cuthbert found such correlations in a corpus of popular songs, composed by people who in general will not be aware of the theories and practices concerning word painting preceding Hanslick (1854/2008). This indicates that these practices are either closely related to the way our speech prosody and our body language reflect our emotions (see Huron, 2015), or that they are deeply rooted in our musical culture. Therefore, it seems fruitful to investigate this kind of correlation further. Yet, it is important to find out whether or not such correlations would be clearer in first verses,

choruses, and bridges than in second, third, and subsequent verses, and whether or not the work of lyricists and composers who seem to be quite aware of what they are doing would show deviant results, either because their combinations of music affect and word affect are more salient, or because they are more often purposefully deviant (i.e., contradictory). Furthermore, when interpreting correlations, one should ask whether these correlations show that specific musical features duplicate the affects in the words, or whether they show that they empower them.

ACKNOWLEDGEMENTS

This article was copyedited by Niels Christian Hansen and layout edited by Kelly Jakubowski.

NOTE

[1] Correspondence can be addressed to: Yke Schotanus, Institute for Cultural Inquiry (ICON), Utrecht University, Muntstraat 2A, 3512 EV Utrecht. E-mail: schotschrift.teksten@planet.nl.

REFERENCES

- Ali, S. O. & Peynircioğlu, Z. F. (2006). Songs and emotions: are lyrics and melodies equal partners. *Psychology of Music*, 34(4), 511–534. <https://doi.org/10.1177/0305735606067168>
- Arthur, C. (2018). A perceptual study of scale-degree qualia in context. *Music Perception*, 35(3), 295–314. <https://doi.org/10.1525/mp.2018.35.3.295>
- Brackett, D. (1995). *Interpreting popular music*. Cambridge, UK: Cambridge University Press.
- Cespedes-Guevara, J. & Eerola, T. (2018). Music communicates affects, not basic emotions: a constructionist account of attribution of emotional meanings to music. *Frontiers in Psychology*, 9, 215. <https://doi.org/10.3389/fpsyg.2018.00215>
- Condit-Schultz, N. (2019). Expanding and contracting definitions of syncopation: commentary on Temperley (2019). *Empirical Musicology Review*, 14(1-2), 81086. <https://doi.org/10.18061/emr.v14i1-2.7098>
- Condit-Schultz, N., & Huron, D. (2015). Catching the lyrics: intelligibility in twelve song genres. *Music Perception*, 32(5), 470-483. <https://doi.org/10.1525/mp.2015.32.5.470>
- Engels, R. C. M. E., Slettenhaar, G., ter Bogt, T. F. M. & Scholte, R. H. J. (2011). Effect of alcohol references in music on alcohol consumption in public drinking places. *American Journal on Addictions*, 20(6), 530–534. <https://doi.org/10.1111/j.1521-0391.2011.00182.x>
- Frith, S. (1987/2007). Why do songs have words? In S. Frith (Ed.), *Taking popular music seriously: selected essays* (pp. 209–238). Aldershot, UK: Ashgate Publishing.
- Furia, P. (1990/1992). *The poets of Tin Pan Alley: a history of America's great lyricists* (2nd edition). New York, NY: Oxford University Press.
- Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In P. N. Juslin & J. A. Sloboda (Eds), *Music and emotion: theory and research* (pp. 223–248). Oxford, UK: Oxford University Press.
- Hansen, N. C. (2018). A call for hypothesis-driven, multi-level analysis in research on emotional word painting in music: commentary on Sun & Cuthbert (2018). *Empirical Musicology Review*, 13(3–4). <https://doi.org/10.18061/emr.v13i3-4.6771>

Hanslick, E. (1854/2008). *Vom Musikalisch-Schönen (On the musically beautiful)*. Project Gutenberg-tm. Retrieved on 27th February 2018 from: <http://www.gutenberg.org/ebooks/26949>

Hatten, R. S. (2004). *Interpreting musical gestures, topics, and tropes: Mozart, Beethoven, Schubert*. Bloomington, IL: Indiana University Press.

Huovinen, E. & Kaila, A.-K. (2015). The semantics of musical topoi: an empirical approach. *Music Perception*, 33(2), 217–243. <https://doi.org/10.1525/mp.2015.33.2.217>

Hu, Y., Chen, X., & Yang, D. (2009). Lyric-based song emotion detection with affective lexicon and fuzzy clustering method. In K. Hirata, G. Tzanetakis, & K. Yoshii (Eds), *Proceedings of the 10th International Conference on Music Information Retrieval* (pp. 123–128). International Society for Music Information Retrieval.

Huron, D. (2015). Cues and signals: an ethological approach to music-related emotion. *Signata*, 6, 331–351. <https://doi.org/10.4000/signata.1115>

Huron, D., Kinney, D., & Precoda, K. (2006). Influence of pitch height on the perception of submissiveness and threat in musical passages. *Empirical Musicology Review*, 1(3), 170–177. <https://doi.org/10.18061/1811/24068>

Juslin, P.N. & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological Bulletin*, 129(5), 770–814. <https://doi.org/10.1037/0033-2909.129.5.770>

Koelsch, S., (2011). Toward a neural basis of processing musical semantics. *Physics of Life Reviews*, 8(2), 89–105. <https://doi.org/10.1016/j.plrev.2011.04.004>

Koops, H. V., Volk, A., & De Haas, W. B. (2015) Corpus-based rhythmic pattern analysis of ragtime syncopation. In M. Müller & F. Wiering (Eds), *Proceedings of the 16th International Society for Music Information Retrieval Conference* (pp. 483–489). International Society for Music Information Retrieval.

Lousberg, L. (2018). *Microtones according to Augustine: neumen, semiotics and rhetoric in Romano-Frankish liturgical chant* (Doctoral dissertation, Utrecht University, The Netherlands). Retrieved from <http://dspace.library.uu.nl/handle/1874/369247>.

Pattison, P. (2015, March 17). *Lesson 46: Phrasing. Songwriting: writing the lyrics* [Video]. Coursera. Retrieved From <https://www.coursera.org/learn/songwriting-lyrics/lecture/3BtUo/phrasing>.

Paul, B., & Huron, D. (2010). An association between breaking voice and grief-related lyrics in country music. *Empirical Musicology Review*, 5(2), 27–35. <https://doi.org/10.18061/1811/46747>

Robinson, P. & Hirsch, P. M. (1969a). *Teenage response to rock and roll protest songs* [paper presentation]. Annual meeting of the American Sociological Association, San Francisco, CA, September 1–4.

Robinson, P. & Hirsch, P. M. (1969b). It's the sound that does it. *Psychology Today*, 3, 42–45. <https://doi.org/10.1037/e400192009-002>

Schotanus, Y. P. (2018). Out-of-key notes and on-beat silences as prosodic cues in sung sentences. In R. Parncutt and S. Sattmann (Eds), *Proceedings of ICMPC15/ESCOM10* (pp. 395–400). Graz, Austria: Centre for Systematic Musicology, University of Graz.

Schotanus, Y.P. (2020). *Singing as a figure of speech, music as punctuation: A study into music as a means to support the processing of sung language* (Doctoral dissertation, Utrecht University).

Schotanus, Y. P. (2020a). Singing and accompaniment support the processing of song lyrics and change the lyrics' meaning. *Empirical Musicology Review*, 15(1-2), 18-55. <https://doi.org/10.18061/emr.v15i1-2.6863>

Schotanus, Y. P. (in preparation). The effect of timing on the singer's tone of voice.

Schotanus, Y. P., Eckhof, L. S., & Willems, R. M. (2018). Behavioral and neurophysiological effects of singing and accompaniment on the perception and cognition of song. In R. Parncutt and S. Sattmann (Eds), *Proceedings of ICMP15/ESCOM10* (pp. 389–394). Graz, Austria: Centre for Systematic Musicology, University of Graz.

Schotanus, Y. P., Kooops, H. V., & Reed Edworthy, J. (2018). Interaction between musical and poetic form affects song popularity: the case of the Genevan psalter. *Psychomusicology*, 28(3), 127–151. <https://doi.org/10.1037/pmu0000216>

Schubert, E. (2013). Emotion felt by the listener and expressed by the music: literature review and theoretical perspectives. *Frontiers in Psychology*, 4, 837. <https://doi.org/10.3389/fpsyg.2013.00837>

Schäfer, K., & Eerola, T. (2020). How listening to music and engagement with other media provide a sense of belonging: an exploratory study of social surrogacy. *Psychology of Music*, 48(2), 232–251. <https://doi.org/10.1177/0305735618795036>

Shanahan, D., & Huron, D. (2014). Heroes and villains: the relationship between pitch tessitura and sociability of operatic characters. *Empirical Musicology Review*, 9(2), 46–59. <https://doi.org/10.18061/emr.v9i2.4441>

Strykowski, D. K. (2016). Text painting, or coincidence? Treatment of height-related imagery in the madrigals of Luca Marenzio. *Empirical Musicology Review*, 11(2), 109–119. <https://doi.org/10.18061/emr.v11i2.4903>

Sun, S. H. & Cuthbert, M. S. (2017). Emotion, painting: lyric, affect, and musical relationships in a large lead-sheet corpus. *Empirical Musicology Review*, 12(3-4), 327–348. <https://doi.org/10.18061/emr.v12i3-4.5889>

Tagg, P. & Clarida, B. (2003). *Ten little title tunes: towards a musicology of the mass media*. New York, NY: The Mass Media Musicologists' Press.

Tan, I., Lustig, E., and Temperley, D. (2019). Anticipatory Syncopation in rock: a corpus study. *Music Perception*, 36(4), 353–70. <https://doi.org/10.1525/mp.2019.36.4.353>

Temperley, D. (2019). Second-position syncopation in European and American vocal music. *Empirical Musicology Review*, 14(1-2), 66–80. <https://doi.org/10.18061/emr.v14i1-2.6986>

Ter Bogt, T., Canale, N., Lenzi, M., Vieno, A., & van den Eijnden, R. (2019). Sad music depresses sad adolescents: a listener's profile. *Psychology of Music*. Advance online publication. <https://doi.org/10.1177/0305735619849622>

Tiemann, L & Huron, D. (2011). Beyond happiness and sadness: affective associations of lyrics with modality and dynamics. *Empirical Musicology Review*, 6(3), 147–154. <https://doi.org/10.18061/1811/52809>