



Effects of voice on emotional arousal

Psyche Loui^{1,2*}, Justin P. Bachorik¹, H. Charles Li¹ and Gottfried Schlaug¹

¹ Department of Neurology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA

² Department of Psychology, Wesleyan University, Middletown, CT, USA

Edited by:

Anjali Bhatara, Université Paris Descartes, France

Reviewed by:

Mireille Besson, Centre National de la Recherche Scientifique, France
E. Glenn Schellenberg, University of Toronto, Canada

*Correspondence:

Psyche Loui, Department of Psychology, Wesleyan University, Judd Hall 104, 207 High Street, Middletown, 06459 CT, USA
e-mail: ploui@wesleyan.edu

Music is a powerful medium capable of eliciting a broad range of emotions. Although the relationship between language and music is well documented, relatively little is known about the effects of lyrics and the voice on the emotional processing of music and on listeners' preferences. In the present study, we investigated the effects of vocals in music on participants' perceived valence and arousal in songs. Participants ($N = 50$) made valence and arousal ratings for familiar songs that were presented with and without the voice. We observed robust effects of vocal content on perceived arousal. Furthermore, we found that the effect of the voice on enhancing arousal ratings is independent of familiarity of the song and differs across genders and age: females were more influenced by vocals than males; furthermore these gender effects were enhanced among older adults. Results highlight the effects of gender and aging in emotion perception and are discussed in terms of the social roles of music.

Keywords: emotion, music, arousal, perception, gender, aging

INTRODUCTION

The ability to detect emotion in speech and music is an important task in our daily lives. The power of the human voice to communicate emotion is well documented in verbal speech (Fairbanks and Pronovost, 1938; Scherer, 1995) as well as in non-verbal vocal sounds (Skinner, 1935), and the human voice is thought to convey emotional valence, arousal, and intensity (Laukka et al., 2005) via its modification of spectral and temporal signals (Fairbanks and Pronovost, 1938; Bachorowski and Owren, 1995). The use of the human voice to convey emotion is abundant and vital developmentally as in the case of infant-directed speech (Trainor et al., 2000), and can be accurately identified by people of different cultures (Bryant and Barrett, 2008), suggesting that emotion communication may be a universal function of the human voice. Furthermore, the inability to detect emotional signals in voices is associated with psychopathy (Bagley et al., 2009), thus highlighting the importance of emotional identification in the auditory modality in every human functioning.

Music is another form of sound communication that conveys emotional information. To understand the perception of emotions in music, one model that has been validated by psychological and physiological studies is as a two-dimensional space that treats affect as two separable dimensions of valence and arousal (Russell, 1980). This valence-arousal model is well validated with musical stimuli (Balkwill and Thompson, 1999; Bigand et al., 2005; Ilie and Thompson, 2006; Steinbeis et al., 2006; Grewe et al., 2007). Studies investigating why and how music is able to influence its listeners' moods and emotions (Sloboda, 1991; Terwogt and van Grinsven, 1991; Balkwill and Thompson, 1999; Panksepp and Bernatzky, 2002; Gosselin et al., 2007) have identified ratings for musical stimuli that drive changes in each of these two factors independently. Arousal is a measure of perceived energy level, ranging from low (calming) to high (exciting) (Krumhansl, 1997; Gosselin et al., 2007; Sammler et al., 2007). Orthogonally,

valence is the polarity of perceived emotions, and ranges from negative (sad) to positive (happy) (Krumhansl, 1997; Schubert, 1999; Dalla Bella et al., 2001). Multidimensional scaling (MDS) studies have verified that valence and arousal are separable measures, that may be independently manipulated in experimental conditions (Bigand et al., 2005; Vines et al., 2005).

Given that music and the voice may both be strong modulators of emotions, vocal music could be a medium with emotional power. Several studies have investigated the cognition and perception of vocal lyrics in songs. Serafine et al. (1982) studied the effect of lyrics on participants' memory for songs. Results showed that melody recognition was near chance unless the melody's original words (i.e., words that were presented with the music during encoding) were present, suggesting that music and speech were combined into a single coherent object when encoded in the same stream. More recently, Weiss et al. (2012) examined the effect of timbre (including voice) on memory and preference for music. Results showed that melodies with the voice were better recognized than all other instrumental melodies. The authors suggest that the biological significance of the human voice provides a greater depth of processing and enhanced memory.

Few studies have investigated the combination of music and speech in emotion perception. In an investigation of the effects of varying stimulus parameters in music and speech on perceived emotion, Ilie and Thompson (2006) showed that emotional ratings for music and speech concurred in most emotion ratings, except that manipulations of pitch height resulted in different directions of valence change for music and speech. Interaction effects between music and speech were again observed, suggesting that the combination of speech with music may result in complex and non-additive effects on emotion.

As music and speech are both auditory stimuli that vary over time, a fundamental question regarding emotion perception of these auditory sources concerns the time-course of emotional

responses. Approaches that have been used to investigate the time-course of emotion perception in music include online responses made during the presentation of music, and offline responses made after hearing musical excerpts. Using both offline techniques of categorization and MDS (Perrot and Gjerdingen, 1999; Bigand et al., 2005), subjective emotional ratings performed after hearing short musical stimuli showed that a musical segment as short as 250 ms in duration is sufficient to elicit a reliable emotional response. However, these emotional ratings were influenced by the *post-hoc* cognitive appraisal of emotional content within music after their presentation, as well as the emotional experience elicited by music during its presentation. Using continuous emotional ratings in the two-dimensional space of valence and arousal maximizes the influence of emotion perceived online during the presentation of musical stimuli (Schubert, 2004). In previous work using the two-dimensional continuous paradigm (Bachorik et al., 2009), participants took an average of 8.3 s to initiate movement signifying an emotional judgment.

The present study adopts both continuous (online) and discrete (offline) subjective ratings to investigate effects of vocals on perception of arousal in music. In addition to exploring the effects of vocals on arousal in music in a temporally sensitive manner, further questions arise concerning the factors that moderate participants' emotional response to the presence of vocals in songs. As previous studies have shown that age and gender may contribute to personality characteristics, which in turn influence musical preference (Rentfrow and Gosling, 2003), we examined the interaction of arousal ratings with age and gender, while controlling for effects of familiarity on arousal ratings. Subjects were presented with excerpts from two versions of well-known songs, one with vocals and one without (with all other variables in the songs being the same), and made continuous as well as discrete ratings of perceived arousal, as well as familiarity ratings, for each version of each song.

MATERIALS AND METHODS

PARTICIPANTS

Fifty participants (25 females and 25 males) were recruited from the greater Boston metropolitan area via advertisements in daily newspapers. Participants ranged from 19 to 83 years of age (median = 37), and were representative of the Boston metropolitan area in their ethnic distribution. All participants reported having no neurological and/or psychiatric disorders and had normal IQ as assessed by Shipley abstract scale scores (Shipley, 1940). Written informed consent, approved by the Institutional Review Board of the Beth Israel Deaconess Medical Center, was obtained from all participants. Each participant was reimbursed at an hourly rate for participating.

STIMULI

The stimuli consisted of 32 unique musical excerpts, each 60 s long. Vocal and instrumental versions of 16 songs were chosen from commercially available songs (see **Table 1** for a list of all songs used). All excerpts were normalized for loudness and each excerpt was briefly faded in (0.5 s) at the beginning of the stimulus and out (0.5 s) at the end. The stimuli were divided into two blocks of 16 trials each; each block consisted of both versions

Table 1 | Excerpts of song stimuli.

Artist	Title	One-min selection
Whitney Houston	I will always love you	1:30–2:30
Bette Midler	Wind beneath my wings	0:30–1:30
Donna Summer	Last dance	0:30–1:30
Bryan Adams	Everything I do	0:30–1:30
Sonny and Cher	I got you babe	0:30–1:30
The Carpenters	Close to you	0:30–1:30
Simple minds	Don't you forget about me	0:30–1:30
Madonna	Like a virgin	0:30–1:30
Lionel Richie and Diana Ross	Endless love	0:30–1:30
Barbara Streisand	The way we were	0:30–1:30
Mr. Mister	Broken wings	0:30–1:30
Alanis Morsette	You oughta know	0:30–1:30
The Police	Every breath you take	0:30–1:30
Gloria Gaynor	I will survive	0:30–1:30
R.E.M.	Losing my religion	0:30–1:30
The Beatles	Can't buy me love	0:30–1:30

(vocal/instrumental and instrumental only) of 8 songs. Excerpts ranged in tempo between 49 and 177 beats per minute.

Experiments were conducted using an Apple Powerbook G4 with a 15.4" LCD screen using custom-made stimulus presentation software (Sourcetone, LLC). Audio was presented via Altec Lansing AHP-712 headphones, and participants used a mouse and a Flightstick Pro USB joystick to input their responses to the stimuli.

PROCEDURE

Over the course of two separate testing sessions, each participant completed two trial blocks. Order of trial block presentation was counterbalanced between subjects. Each of the 16 excerpts in each trial block was played in a randomized order, and for each stimulus presentation, the participant's task was the same: to use the joystick to respond, in real time, to the levels of emotional valence (defined as positive or negative emotion induced by the music) and arousal (defined as a stimulating or calming feeling induced by the music) of the music via an onscreen cursor in a two-dimensional grid. The joystick controlled the motion of the cursor in a 640 × 640 resolution grid, and data about the position of the joystick and the position of the cursor was sampled with a frequency of 10 Hz. Centering the joystick caused the cursor to stop moving but did not center the cursor in the grid onscreen.

After the end of each musical excerpt, subjects had additional tasks to rate the degree of valence and arousal perceived in each excerpt (on a scale of 0–4, where 4 is highest, 2 is neutral, and 0 is lowest). Participants also provided subjective ratings of familiarity (on a scale of 0–4, with 0 being "never heard" and 4 being "actively listen to; personally own song") after rating the degree of emotional arousal and valence.

DATA ANALYSIS

Continuous ratings for valence and arousal (X and Y axes on the two-dimensional rating space, respectively) were digitized and

exported for each trial of each subject from the stimulus presentation program and analyzed using in-house software. Pairwise *t*-tests were conducted for each time point comparing subjects' valence and arousal ratings for vocal and instrumental versions of each song. A false-discovery rate *post-hoc* adjustment was used to minimize Type I error.

Discrete valence and arousal ratings were used as the dependent variable in a mixed design ANOVA with between-subject factors of age (two levels: old vs. young, with a median split at the age of 37) and gender (male vs. female) and the within-subject factor of song vocals (instrumental vs. vocals). Paired *t*-tests were run comparing music with and without vocals in familiarity, liking, chills, and intense emotional responses.

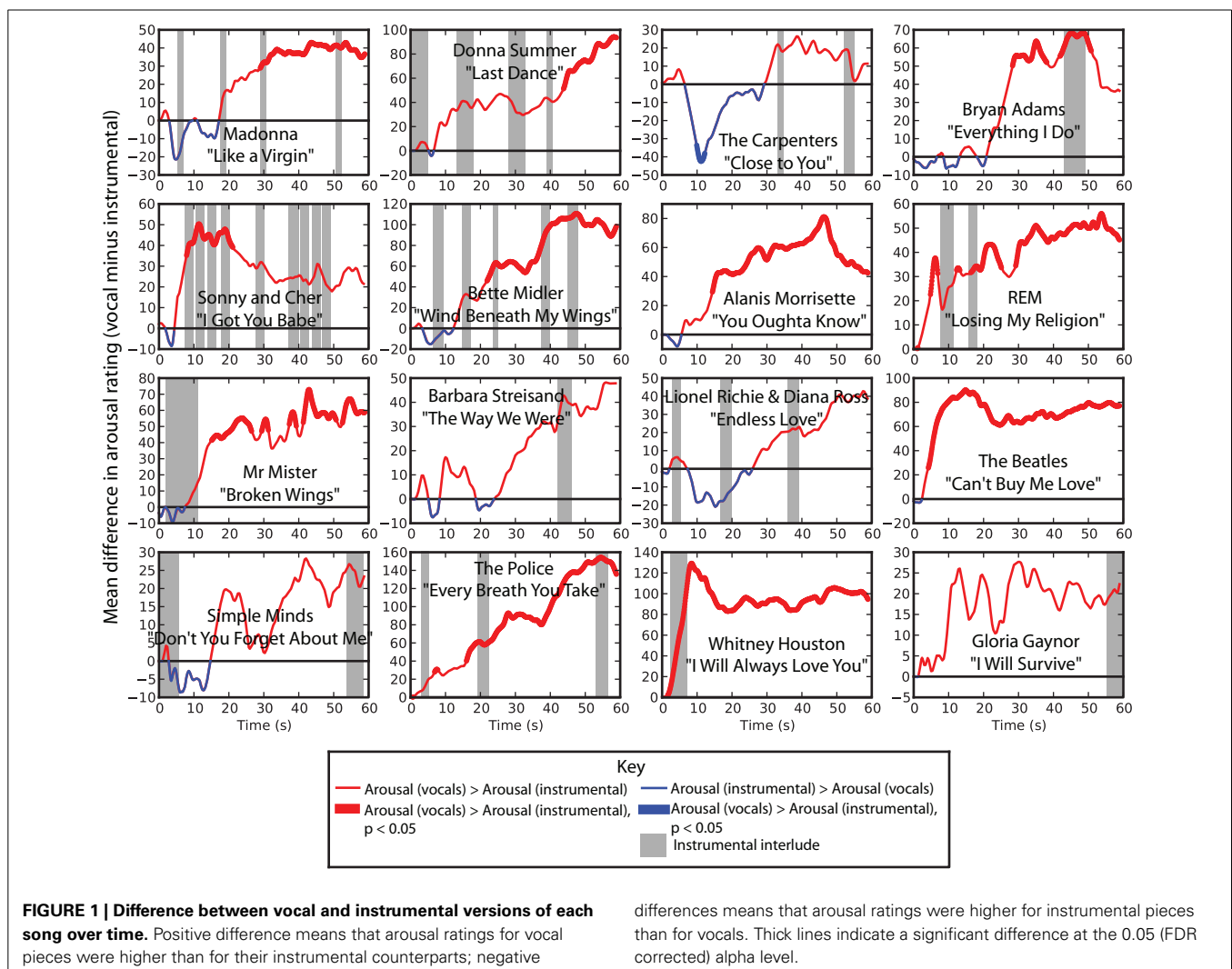
RESULTS

Continuous arousal ratings revealed that the vocal versions were more arousing overall. The average continuous ratings were higher in the vocal version than in the instrumental version in 15 out of 16 songs. This was confirmed using a pairwise *t*-test at every point in the time-series comparing arousal ratings in vocal and instrumental conditions indicating significant difference at

the FDR-corrected alpha level of 0.05 in at least one time point between vocal and instrumental versions in 12 out of 16 songs. Among these 12 songs, 11 showed a significant arousal-enhancing effect of vocals, whereas only one song showed the opposite effect. In contrast to arousal ratings, continuous valence ratings only showed significantly higher valence ratings at the $p < 0.05$ (corrected) level for at least one point in 4 out of 16 songs, and significantly lower valence ratings for at least one point in two songs.

Figure 1 shows the difference between average arousal rating between vocal and instrumental versions as functions of time for each of the 16 songs. Red line segments indicate a higher arousal rating in vocal versions compared to instrumental versions whereas blue line segments indicate the opposite effect. Bold lines indicate significant differences at the $p < 0.05$ (FDR-corrected) level and gray bars behind the graph indicate instrumental interludes within the vocal versions of each song.

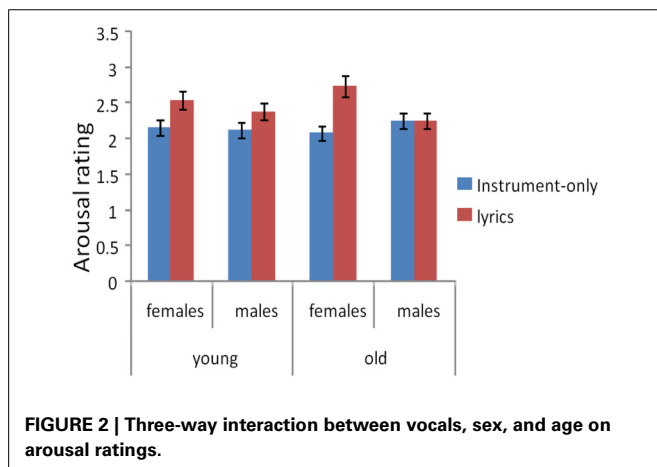
Online ratings indicated that, as shown in **Figure 1**, the arousal-enhancing effect of vocals was more pronounced later within each piece. The trend toward higher arousal ratings in the



vocal versions began at an average of 10 s after the onset of each song, however, this was variable depending on the song (SEM = 2.6 s). The presence of instrumental interludes within each song was uncorrelated with the difference in arousal ratings. Songs that contained non-verbal vocal portions (Whitney Houston, Barbara Streisand, and Mr. Mister songs in the sample) showed a similar effect size as songs containing verbal vocals, suggesting that the presence of the human voice, rather than recognizable words, led to the increase in arousal.

The effect of vocals on arousal was confirmed in discrete as well as continuous arousal ratings. Using the discrete arousal rating as the dependent variable, the mean arousal rating for instrumental versions of the musical excerpts was 2.25 (SEM = 0.07) whereas the mean arousal rating for vocal versions was 2.60 (SEM = 0.06). A highly significant main effect of vocals on arousal was observed, $F_{(1, 96)} = 1389.5, p < 0.001$, indicating that songs with vocals were rated as more highly arousing than their instrument-only counterparts (Figure 2). Participants also reported liking the vocal versions more than the instrumental version, with a mean of 2.75 vs. 2.48, respectively [$t_{(49)} = -3.486, p < 0.001$]. The same effect was not observed in discrete valence ratings [$F_{(1, 96)} = 1.17, n.s.$].

Using discrete arousal ratings as the dependent variable, we next attempted to tease apart the groups of participants who were or were not susceptible to the effects of vocals on arousal by assessing the demographics (gender and age) of each participant and comparing the mean difference between vocals and instrumental versions across demographic groups. A significant main effect of gender was observed for all arousal ratings, with ratings by females being higher [$F_{(1, 96)} = 4.186, p = 0.04$]. Furthermore, a significant interaction between vocals and gender was observed on arousal ratings: $F_{(1, 96)} = 11.9, p = 0.001$, confirming that the positive effect of vocals on arousal ratings was stronger for females than for males (Figure 2). Although no significant main effect of age was present [$F_{(1, 96)} = 0.013, n.s.$], a significant three-way interaction was observed on arousal ratings between gender and age [$F_{(1, 96)} = 4.17, p = 0.04$], with older females being more emotionally influenced by vocals than younger females, but older males being less influenced by vocals than younger males (Figure 2).



Familiarity ratings revealed that participants found songs with vocals to be significantly more familiar than the instrumental version [mean ratings: vocals = 2.63, instrumental = 1.872; $t_{(49)} = -9.319, p < 0.001$]. To investigate the effects of vocals on arousal while controlling for the effect of familiarity, a one-way ANCOVA was conducted on the dependent variable of discrete arousal rating with the factor of vocals (instrumentals vs. vocals), with the covariate of familiarity rating (0 through 4). Results showed a significant effect of vocals [$F_{(1, 97)} = 4.2, p = 0.043$] even with a significant effect of familiarity [$F_{(1, 97)} = 6.3, p = 0.014$], suggesting that the contribution of vocal stimuli to arousal was significant even after controlling for an increase in familiarity for vocal pieces.

DISCUSSION

Our results indicate that the presence of vocals generally enhances participants' arousal ratings, and were not limited to the effects of familiarity but were moderated by the gender and age of the participant. Vocal sounds and music engage multiple common resources in the brain, resulting in interactions between music and speech as assessed by tasks that tap into perception, cognition and emotion (Serafine et al., 1982; Besson et al., 1998; Ilie and Thompson, 2006). However, little research has investigated the time-course of the impact that vocals may have on arousal perception in music. Using a naturalistic and ecologically valid setting of popular songs with and without vocal content, the present study attempted to address the specific question concerning the relationship between vocals and perceived arousal in music. While the present study uses ecologically valid stimuli and identifies arousal differences attributable to the use of vocals within music, future research may be done to tease apart specific components of the vocals (e.g., words, timbre, sung melody) that most affect perceived arousal.

Based on continuous (online) and discrete (offline) subjective ratings of valence and arousal for identical musical excerpts with and without vocal content, we observed that the presence of vocals generally increases ratings of arousal but not of valence. The emotionally enhancing effect of vocals on arousal is shown in both online (continuous) and offline (discrete) ratings of subjective arousal, and is not limited to verbal lyrics but appears to generalize to non-verbal songs containing the human voice. Online ratings revealed that participants required an average of 10 s (SEM = 2.6 s) of music before differentiating vocal versions from instrumental versions; this was congruent with previous reports using a similar continuous ratings paradigm (Bachorik et al., 2009) showing that participants required an average of 8.3 s to initiate emotional ratings when listening in real time. Furthermore, the enhancing effect of vocals is not limited to familiarity, as shown by an ANCOVA revealing that effects of vocals were significant even after statistically controlling for the contribution of familiarity ratings.

It is interesting to speculate on why valence is less affected by vocals compared to arousal. One possibility is that vocals affected valence both positively and negatively depending on the listener and depending on the song, resulting in increased variability. Another possibility is that valence is already much determined from other structural features of music such

as modality (major vs. minor keys) and melodic contour, leaving little changes that the added vocals could bestow upon the perceived valence of each song. The relative impact of structural features of a piece on its perceived valence vs. arousal may be an avenue for future studies.

As music with vocals has additional components of timbre, melody, and words, the present experiment design could be followed up by assessing the effect of an additional lead instrument on arousal ratings in a non-vocal control condition. However, the selection of the most appropriate additional lead instrument in such a design is non-trivial, as only a highly systematic match in timbre between the voice and the chosen test instrument would provide a true test of the possible confound of voice timbre. Future experiments should seek to identify a timbral match of the voices used in these naturalistic song stimuli in order to define a timbre-matched control condition. Nevertheless, in the current analysis we identify song sections that do not include words as a possible means to de-confound the relationship between voice and lyrics, and as the increase in arousal ratings is observed even for sections of the songs that include non-verbal vocals, the results suggest that the use of vocals, rather than of lyrics within the music, may be driving the increase in arousal.

When offline ratings were compared by the demographic variables of gender and age, results revealed the types of participants who were most sensitive to the arousal-enhancing effect of vocals. Females were more inclined to report perceiving higher arousal in vocal songs compared to males. These effects are exaggerated among older participants. One possible explanation for the gender effect is that the need to detect emotional signals rapidly may be more evolutionarily advantageous for women. Supporting evidence along this possible evolutionary basis of gender-bias in selecting for emotion in vocal content comes from electrophysiological literature showing that the dishabituation of emotional voice content is more robust in females, and is furthermore regulated by estrogen levels (Schirmer et al., 2008). Regarding the three-way interaction between the effect of vocals with gender and with age, one possibility is that the song stimuli—popular songs ranging from the 1960s to the 1990s—chosen for this experiment are more familiar to older individuals than to younger ones. However, the fact that the effect of vocals on arousal was

still significant after controlling for the contribution of familiarity suggests that the influence of vocals on arousal was above and beyond the influence of familiarity. Another possible explanation stems from how individuals of different ages identify with music, with possible sociological effects of changing standards of gender equality throughout the decades that may help explain the observed gender by age interaction. As young adults rely on musical preferences to communicate and understand each other's personality profiles (Rentfrow and Gosling, 2006), it would seem that younger individuals, especially females and individuals who rely on external feedback and social pressures for self-perception, may be more easily aroused by music that is representative of their own culture and the personality profile they wish to convey. Since most popular music is written with vocals, it stands to reason that younger listeners looking to identify themselves with popular taste would find music more arousing when presented with vocals. As the emotional content of songs is highly influenced by our identity as captured by demographic variables such as age and gender, future work should seek to refine our understanding of emotion perception in music and language by placing it in broader sociological and biological contexts.

The present results from continuous and discrete ratings, obtained during and after music listening, support the central notion that the combination of vocal and instrumental sounds in music could produce a more pronounced effect on emotional arousal, but not on valence, compared to instrumental music alone. The arousal-enhancing effect of vocals increases over the duration of most songs and is moderated by demographic factors such as age and gender. Results have implications for our understanding of the emotion and meaning of music, and will bear relevance for ongoing efforts to model and predict the emotional content of music (Nagel et al., 2007) for therapeutic as well as commercial applications.

ACKNOWLEDGMENTS

This work was supported by a research grant from Sourcetone, LLC, given to Beth Israel Deaconess Medical Center to support research on music and emotions. Psyche Loui also acknowledges support from the Grammy Foundation and the Templeton Foundation.

REFERENCES

- Bachorik, J., Bangert, M., Loui, P., Larke, K., Berger, J., Rowe, R., et al. (2009). Emotion in motion: investigating the time-course of emotional judgments of musical stimuli. *Music Percept.* 26, 355–364. doi: 10.1525/mp.2009.26.4.355
- Bachorowski, J.-A., and Owren, M. J. (1995). Vocal expression of emotion: acoustic properties of speech are associated with emotional intensity and context. *Psychol. Sci.* 6, 219–224. doi: 10.1111/j.1467-9280.1995.tb00596.x
- Bagley, A. D., Abramowitz, C. S., and Kosson, D. S. (2009). Vocal affect recognition and psychopathy: converging findings across traditional and cluster analytic approaches to assessing the construct. *J. Abnorm. Psychol.* 118, 388–398. doi: 10.1037/a0015372
- Balkwill, L. L., and Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Percept.* 17, 43–64. doi: 10.2307/40285811
- Besson, M., Faita, F., Peretz, I., Bonnel, A. M., and Requin, J. (1998). Singing in the brain: independence of lyrics and tunes. *Psychol. Sci.* 9, 494. doi: 10.1111/1467-9280.00091
- Bigand, E., Filipic, S., and Lalitte, P. (2005). The time course of emotional responses to music. *Ann. N.Y. Acad. Sci.* 1060, 429–437. doi: 10.1196/annals.1360.036
- Bryant, G. A., and Barrett, H. C. (2008). Vocal emotion recognition across disparate cultures. *J. Cogn. Cult.* 8, 135–148. doi: 10.1163/156770908X289242
- Dalla Bella, S., Peretz, I., Rousseau, L., and Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition* 80, B1–B10. doi: 10.1016/S0010-0277(00)00136-0
- Fairbanks, G., and Pronovost, W. (1938). Vocal pitch during simulated emotion. *Science* 88, 382–383. doi: 10.1126/science.88.2286.382
- Gosselin, N., Peretz, I., Johnsen, E., and Adolphs, R. (2007). Amygdala damage impairs emotion recognition from music. *Neuropsychologia* 45, 236–244. doi: 10.1016/j.neuropsychologia.2006.07.012
- Grewe, O., Nagel, F., Kopiez, R., and Altenmüller, E. (2007). Emotions over time: synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion* 7, 774–788. doi: 10.1037/1528-3542.7.4.774
- Ilie, G., and Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Percept.* 23,

- 319–330. doi: 10.1525/mp.2006.23.4.319
- Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Can. J. Exp. Psychol.* 51, 336–353. doi: 10.1037/1196-1961.51.4.336
- Laukka, P., Juslin, P. N., and Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cogn. Emot.* 19, 633–653. doi: 10.1080/02699930441000445
- Nagel, F., Kopiez, R., Grewe, O., and Altenmüller, E. (2007). EMuJoy: software for continuous measurement of perceived emotions in music. *Behav. Res. Methods* 39, 283–290. doi: 10.3758/BF03193159
- Panksepp, J., and Bernatzky, G. (2002). Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behav. Processes* 60, 133–155. doi: 10.1016/S0376-6357(02)00080-3
- Perrot, D., and Gjerdingen, R. O. (1999). “Scanning the dial: an exploration of factors in the identification of musical style,” in *Paper Presented at the Proceedings of the 1999 Society for Music Perception and Cognition*.
- Rentfrow, P. J., and Gosling, S. D. (2003). The do re mi’s of everyday life: the structure and personality correlates of music preferences. *J. Pers. Soc. Psychol.* 84, 1236–1256. doi: 10.1037/0022-3514.84.6.1236
- Rentfrow, P. J., and Gosling, S. D. (2006). Message in a ballad: the role of music preferences in interpersonal perception. *Psychol. Sci.* 17, 236–242. doi: 10.1111/j.1467-9280.2006.01691.x
- Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714
- Sammler, D., Grigutsch, M., Fritz, T., and Koelsch, S. (2007). Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology* 44, 293–304. doi: 10.1111/j.1469-8986.2007.00497.x
- Scherer, K. R. (1995). Expression of emotion in voice and music. *J. Voice* 9, 235–248. doi: 10.1016/S0892-1997(05)80231-0
- Schirmer, A., Escoffier, N., Li, Q. Y., Li, H., Strafford-Wilson, J., and Li, W.-I. (2008). What grabs his attention but not hers? Estrogen correlates with neurophysiological measures of vocal change detection. *Psychoneuroendocrinology* 33, 718–727. doi: 10.1016/j.psyneuen.2008.02.010
- Schubert, E. (1999). Measuring emotion continuously: validity and reliability of the two-dimensional emotion-space. *Aust. J. Psychol.* 51, 154–165. doi: 10.1080/00049539908255353
- Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Percept.* 21, 561–585. doi: 10.1525/mp.2004.21.4.561
- Serafine, M. L., Crowder, R. G., and Repp, B. H. (1982). Integration of melody and text in memory for songs. *J. Acoust. Soc. Am.* 72, S45. doi: 10.1121/1.2019896
- Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *J. Psychol.* 9, 371–377. doi: 10.1080/00223980.1940.9917704
- Skinner, E. R. (1935). A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness; and a determination of the pitch and force of the subjective concepts of ordinary, soft and loud tones. *Speech Monogr.* 2, 81–137. doi: 10.1080/03637753509374833
- Sloboda, J. (1991). Music structure and emotional response: some empirical findings. *Psychol. Music* 19, 110–120. doi: 10.1177/0305735691192002
- Steinbeis, N., Koelsch, S., and Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: evidence from subjective, physiological, and neural responses. *J. Cogn. Neurosci.* 18, 1380–1393. doi: 10.1162/jocn.2006.18.8.1380
- Terwogt, M. M., and van Grinsven, F. (1991). Musical expression of moodstates. *Psychol. Music* 19, 99–109. doi: 10.1177/0305735691192001
- Trainor, L. J., Austin, C. M., and Desjardins, R. E. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychol. Sci.* 11, 188–195. doi: 10.1111/1467-9280.00240
- Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., and Levitin, D. J. (2005). Dimensions of emotion in expressive musical performance. *Ann. N.Y. Acad. Sci.* 1060, 462–466. doi: 10.1196/annals.1360.052
- Weiss, M. W., Trehub, S. E., and Schellenberg, E. G. (2012). Something in the way she sings: enhanced memory for vocal melodies. *Psychol. Sci.* 23, 1074–1078. doi: 10.1177/0956797612442552

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 June 2013; accepted: 07 September 2013; published online: 01 October 2013.

Citation: Loui P, Bachorik JP, Li HC and Schlaug G (2013) Effects of voice on emotional arousal. *Front. Psychol.* 4:675. doi: 10.3389/fpsyg.2013.00675

This article was submitted to *Emotion Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2013 Loui, Bachorik, Li and Schlaug. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.