

# Something in the Way She Sings: Enhanced Memory for Vocal Melodies

Michael W. Weiss, Sandra E. Trehub, and  
E. Glenn Schellenberg

University of Toronto Mississauga

Psychological Science  
23(10) 1074–1078  
© The Author(s) 2012  
Reprints and permission:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0956797612442552  
<http://pss.sagepub.com>



## Abstract

Across species, there is considerable evidence of preferential processing for biologically significant signals such as conspecific vocalizations and the calls of individual conspecifics. Surprisingly, music cognition in human listeners is typically studied with stimuli that are relatively low in biological significance, such as instrumental sounds. The present study explored the possibility that melodies might be remembered better when presented vocally rather than instrumentally. Adults listened to unfamiliar folk melodies, with some presented in familiar timbres (voice and piano) and others in less familiar timbres (banjo and marimba). They were subsequently tested on recognition of previously heard melodies intermixed with novel melodies. Melodies presented vocally were remembered better than those presented instrumentally even though they were liked less. Factors underlying the advantage for vocal melodies remain to be determined. In line with its biological significance, vocal music may evoke increased vigilance or arousal, which in turn may result in greater depth of processing and enhanced memory for musical details.

## Keywords

memory, melody, music, timbre, voice

Received 11/2/11; Revision accepted 2/24/12

Because recognition of conspecifics is required for reproduction, it is not surprising that numerous species respond differentially to conspecific vocalizations, and that members of social species often recognize individual callers (Ghazanfar et al., 2007; Pollard & Blumstein, 2011). Moreover, specific “voice cells” have been identified in the primate temporal lobe (Perrodin, Kayser, Logothetis, & Petkov, 2011). As for humans, newborns recognize their mother’s voice (DeCasper & Fifer, 1980), and they listen preferentially to speech over nonspeech analogues (Vouloumanos & Werker, 2007). In human adults, the voice activates distinctive cortical regions even in the absence of linguistic content (Belin, Zatorre, & Ahad, 2002), and vocal tones produce larger responses than instrumental tones in areas around Heschl’s gyrus (Gunji et al., 2003). In principle, there could be cognitive consequences of conspecific vocalizations in addition to enhanced signal recognition. For example, nonhuman species with reasonable pitch-processing abilities focus on absolute pitch at the expense of pitch relations (D’Amato, 1988; Hulse & Cynx, 1985). Nevertheless, European starlings (*Sturnus vulgaris*) can be trained to recognize transpositions of conspecific songs, even though they fail to recognize transposed piano melodies after comparable training (Bregman, Patel, & Gentner, 2012).

Surprisingly, music cognition in human listeners is typically studied with stimuli that are relatively low in biological significance, such as instrumental sound patterns or, more commonly, digital analogues of those patterns. Although instrumental sounds are unfamiliar and irrelevant to European starlings, they are culturally appropriate, if not biologically significant, for human adults. The assumption is that musical timbre, which refers to the sound quality or tone color that differentiates instruments or voices of the same pitch, amplitude, and duration (Risset & Wessel, 1999), is largely irrelevant to fundamental perceptual or cognitive processes, such as relational processing or memory. It is conceivable, however, that vocal music, by virtue of its status as the earliest musical form (Mithen, 2005) and its use of a biologically significant timbre, would facilitate various aspects of music processing.

The present study focused on human adults’ memory for melodies—in particular, on the possibility that melodies might be remembered better when presented vocally rather than

## Corresponding Author:

Sandra E. Trehub, Department of Psychology, University of Toronto  
Mississauga, Mississauga, Ontario L5L 1C6, Canada  
E-mail: [sandra.trehub@utoronto.ca](mailto:sandra.trehub@utoronto.ca)

instrumentally. The prevailing belief is that listeners' long-term mental representations of music consist largely of relational pitch and timing information (Krumhansl, 2000). It is clear, however, that listeners remember much more than the pitch and temporal relations of recorded music that they hear frequently. For example, adults remember the pitch level and tempo of their favorite pop songs (Levitin, 1994; Levitin & Cook, 1996), children and adults remember the pitch level of familiar TV theme music (Schellenberg & Trehub, 2003; Trehub, Schellenberg, & Nakata, 2008), and infants remember the pitch level of familiar recordings of lullabies (Volkova, Trehub, & Schellenberg, 2006). Moreover, adults remember incredible detail about timbre, which enables them to identify pop songs from excerpts as brief as 100 and 200 ms (Schellenberg, Iverson, & McKinnon, 1999). Joint encoding of timbre and melody is also evident in listeners' reduced memory for melodies when the timbre changes from exposure to test (Halpern & Müllensiefen, 2008; Peretz, Gaudreau, & Bonnel, 1998).

To date, there has been no attempt to ascertain whether a biologically significant timbre could enhance memory for melodies. In the present study, we explored adults' memory for unfamiliar melodies taken from British and Irish folk songs. The melodies were presented in four different timbres. Two of the timbres, the voice (i.e., singing the syllable "la") and piano, were highly familiar, and two, the banjo and marimba, were much less familiar. If a familiar timbre makes a melody easy to remember, then vocal and piano melodies should be remembered better than banjo and marimba melodies. By contrast, if the voice has special status, then vocal melodies should be remembered better than all other instrumental melodies, including piano melodies.

In the exposure phase of the study, participants heard several melodies that were assigned to four different timbres. In the recognition phase, they heard the previous melodies as well as novel melodies that were assigned to the same four timbres, and they rated each melody as old or new. Subsequently, they indicated how much they liked each melody, which made it possible to examine potential contributions of timbre appraisals to memory. Finally, they named the timbres and rated their relative familiarity.

## Method

### Participants

Participants were 64 undergraduates (49 women, 15 men; mean age = 20.5 years,  $SD = 2.1$ ), who were recruited without regard to music training. On average, they had 4.1 years of training ( $SD = 4.1$ , range = 0–14, median = 3; positively skewed distribution); 25 participants had taken piano lessons, 5 had taken voice lessons, and 7 others had taken both piano and voice lessons. None had studied the banjo or marimba. Seven additional participants were excluded because of failure to follow instructions ( $n = 1$ ), technical problems ( $n = 1$ ), or

higher "recognition" of new than old melodies, reflecting inattention to the stimuli ( $n = 5$ ).

### Apparatus and stimuli

The stimuli, which were 13 to 19 s in duration, comprised 32 excerpts of unfamiliar folk melodies from the United Kingdom and Ireland (see the Supplemental Material available online for audio examples). All melodies conformed to Western tonality. For the real-instrument condition, each melody was recorded in two common timbres, voice and piano, and two less common timbres, banjo and marimba. For the vocal renditions, an amateur female (alto) singer with a pleasant voice sang all 32 melodies without lyrics (i.e., "la" for each note) in an everyday (nonoperatic) manner to a monophonic backing track (MIDI piano) presented over headphones. Digital editing software was used to pitch-correct and time-correct individual notes. The software centered the average pitch of each note to true tuning, retaining natural qualities like vibrato and amplitude variations but correcting inconsistencies in note timing and pitch. For the instrumental versions, amateur musicians generated live performances of each melody on the piano, banjo, and marimba. They played along with the backing track used for the vocal performances to ensure that the tempo and overall duration of each melody were matched across timbres. In a pretest confirming that digital editing of the vocal melodies did not result in unnatural voice quality, a separate sample of 14 listeners rated how natural each excerpt sounded in each of the timbres. On average, vocal melodies received the highest ratings.

Melodies in one timbre could be more memorable than those in other timbres because of extraneous performance variations. We addressed this issue by including a MIDI condition, in which the instrumental and vocal renditions were more closely matched. Specifically, MIDI data were generated from the vocal performances and used to create versions in digital instrumental timbres (piano, banjo, and marimba). These instrumental versions had notes matched in pitch, duration, and amplitude to the vocal versions (i.e., identical MIDI parameters).

To ensure that intrinsic differences in the memorability of individual melodies were counterbalanced across timbres, we assigned melodies (numbered 1–32) to eight melody-timbre conditions using a modified Latin-square design. In Condition 1, Melodies 1 through 16 were presented during the exposure phase (voice: 1–4; piano: 5–8; banjo: 9–12; marimba: 13–16), and Melodies 17 through 32 served as foils during the recognition phase (voice: 17–20; piano: 21–24; banjo: 25–28; marimba: 29–32). The melodies presented during the exposure phase and those serving as foils in the recognition phase were the same in Conditions 2 through 4 as in Condition 1, but the timbres were rotated. For example, in Condition 2, Melodies 1 through 4 and 17 through 20 were in marimba timbre, Melodies 5 through 8 and 21 through 24 were in voice timbre, Melodies 9 through 12 and 25 through 28 were in piano timbre,

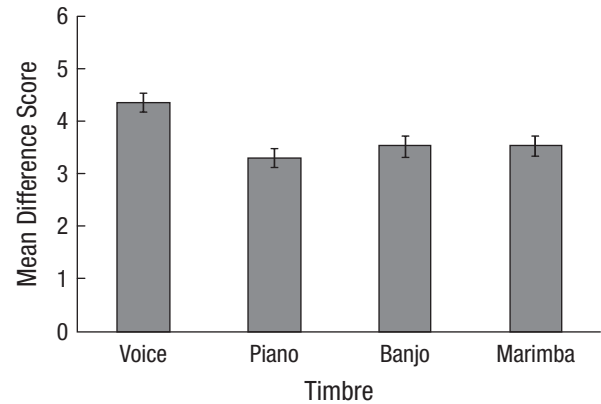
and Melodies 13 through 16 and 29 through 32 were in banjo timbre. Conditions 5 through 8 matched Conditions 1 through 4, respectively, except that the exposure and foil melodies were reversed.

## Procedure

In each recording condition (real instrument, MIDI), 32 participants were assigned to the eight melody-timbre combinations; assignment was random but constrained to ensure a balanced design. Participants were tested individually. They were told that after hearing each melody, they should answer the question on the computer monitor by means of the mouse or keyboard. In the first (exposure) phase of the test session, participants heard 16 melodies, 4 in each timbre. Each melody was presented three times over the course of three blocks of trials; the order of the melodies was randomized separately in each block. To maximize attention to the melodies, we asked participants to indicate whether each melody sounded happy, sad, or neutral. During a 5- to 10-min break following the exposure phase, participants completed a background questionnaire about their hearing health and music training. In the second (recognition) phase, they heard the 16 melodies from the exposure phase (old melodies) and 16 foils (new melodies), 4 in each timbre. They rated their confidence that each melody was old or new on a 7-point scale ranging from 1 (*definitely new*) to 7 (*definitely old*). In the third phase, participants heard all 32 melodies and rated how much they liked each melody on a 5-point scale ranging from 1 (*dislike extremely*) to 5 (*like extremely*). This phase was included to ascertain whether differential appraisals could account for any recognition differences across timbres. In the final phase, participants heard 1 melody from each timbre, then typed the instrument name (voice, piano, banjo, or marimba) and rated their everyday familiarity with that instrument on a 5-point scale from 1 (*very unfamiliar*) to 5 (*very familiar*).

## Results

Timbre identification and familiarity judgments confirmed that the voice and piano timbres were named equally well and were maximally familiar. Both of these timbres were named more readily and judged more familiar than the banjo and marimba timbres. Four difference scores were calculated for each participant by subtracting the average recognition (old/new) rating for new melodies from the average recognition rating for old melodies separately for each timbre. Positive scores reflected recognition of previously heard melodies. A two-way mixed-design analysis of variance (ANOVA) with timbre (voice, piano, banjo, or marimba) as a repeated measure and recording condition (real instrument or MIDI) as a between-subjects variable revealed no main effect of recording condition,  $F < 1$ , and no interaction between recording condition and timbre,  $p > .3$ . Difference scores are shown in

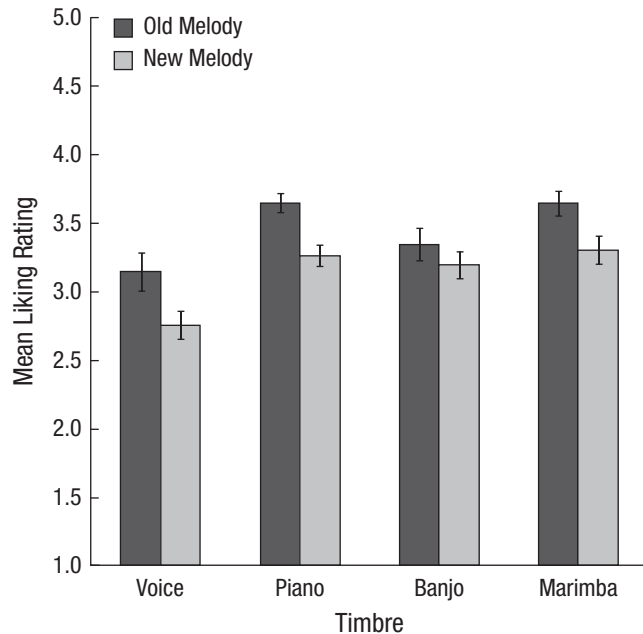


**Fig. 1.** Mean difference score (average recognition rating for old melodies minus average recognition rating for new melodies) as a function of timbre (0 = chance recognition, 6 = perfect recognition). Error bars represent standard errors.

Figure 1. The main effect of timbre was significant,  $F(3, 186) = 5.84$ ,  $p = .001$ ,  $\eta_p^2 = .18$ . Follow-up pairwise comparisons (Bonferroni-corrected for multiple tests) revealed significantly higher difference scores (greater recognition) for vocal melodies than for piano, banjo, and marimba melodies, all  $ps < .001$ , but no differences among the three instrumental timbres ( $p$  values indeterminate after Bonferroni correction). In short, listeners exhibited better recognition of melodies presented vocally rather than instrumentally. Moreover, melodies presented in piano timbre were remembered no better than those presented in the less familiar banjo or marimba timbres. Signal detection analyses yielded an identical pattern of findings (see Supplementary Analyses in the Supplemental Material).

Because half of the participants had played the piano for 1 year or more ( $M = 4.95$ ,  $SD = 4.00$ , range = 1–14), we recalculated the ANOVA on difference scores with piano training as an additional independent variable. A main effect of training,  $F(1, 60) = 7.76$ ,  $p = .007$ ,  $\eta_p^2 = .11$ , revealed better memory for piano-trained participants than for their untrained counterparts but no differential enhancement of memory for piano melodies (no two-way interaction),  $F < 1$ .

Liking ratings were examined with a three-way mixed-design ANOVA with timbre (voice, piano, banjo, or marimba) and exposure (old or new) as repeated measures and recording condition (real instrument or MIDI) as a between-subjects variable. As was the case for recognition, there was no main effect or interaction involving recording condition,  $ps > .05$ . Liking ratings, collapsed across recording condition, are shown in Figure 2. There was no interaction between exposure and timbre. A main effect of exposure on liking,  $F(1, 62) = 22.90$ ,  $p < .001$ ,  $\eta_p^2 = .27$ , revealed that participants liked old melodies (heard three times in the exposure phase, once in the recognition phase, and a fifth time in the liking phase) more than foils (heard once in the recognition phase and again in the liking phase). There was also a main effect of timbre,  $F(3,$



**Fig. 2.** Mean liking rating as a function of previous exposure and timbre. Error bars represent standard errors.

186) = 7.94,  $p < .001$ ,  $\eta_p^2 = .11$ : Vocal renditions were liked significantly less than piano, banjo, and marimba renditions,  $ps < .05$ ; liking did not differ among the latter three types of renditions,  $ps > .5$ . Liking ratings were uncorrelated with difference (recognition) scores,  $ps > .05$ .

## Discussion

This study is the first to evaluate the influence of a biologically significant timbre on adults' memory for melodies. The findings provide unequivocal evidence that vocal melodies are remembered better than instrumental melodies. In line with its biological significance, vocal music may evoke increased vigilance or arousal, which in turn may result in greater depth of processing ( Craik & Lockhart, 1972) and enhanced memory for musical details. Unquestionably, the voice is much more familiar than the piano or any other instrumental timbre, but its familiarity is inseparable from its biological significance. It is also possible that listeners more readily encode cues to identity from vocal than from instrumental performances, and that such indexical cues contribute to the recognition of previously heard melodies.

Melodies from the exposure phase, which were heard for the fifth time in the liking phase, were liked more than foils heard for the second time in the liking phase. This finding corroborates previous studies that have demonstrated increased liking for music as a result of increasing exposure, within limits (Halpern & Müllensiefen, 2008; Peretz & Gagnon, 1998; Schellenberg, Peretz, & Viellard, 2008; Szpunar, Schellenberg, & Pliner, 2004). Although vocal melodies were remembered

better than melodies presented in other timbres, increased exposure did not yield differential gains in appraisal for vocal melodies. In fact, old and new vocal melodies were rated less favorably than old and new melodies in other timbres, perhaps because of the interminable repetition of "la la la." In any event, unfavorable evaluations of music are not associated with enhanced recognition (Stalinski & Schellenberg, in press).

Activation of auditory or motor representations is thought to enhance music processing for listeners with training on the instrument of presentation. Instrumental practice generates growth in the auditory cortex and in regions of the somatosensory cortex that correspond to instrument-specific motor movements (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995; Pantev, Roberts, Schulz, Engelien, & Ross, 2001). Such practice results in tightly coupled auditory and motor systems, especially for highly trained musicians (Zatorre, Chen, & Penhune, 2007). In the present study, adults with limited piano training performed better overall than untrained adults, a finding in line with enhanced performance of musically trained listeners on a variety of listening tasks (Strait & Kraus, 2011). Nevertheless, piano-trained participants recognized piano melodies no better than melodies with less familiar timbres. Indeed, there is no evidence to date that melodies are remembered better when presented in familiar instrumental timbres than in unfamiliar timbres (Halpern & Müllensiefen, 2008).

In principle, subvocal activity or related motor imagery could have enriched participants' representations of the vocal melodies, but such motor activation is unlikely for unfamiliar musical material. Although the mechanisms underlying the observed effect of vocal timbre are unclear, what is clear is that musical timbres are unequal in terms of their consequences for human listeners. The prevailing use of timbres of convenience in studies of music cognition may limit the insights gained from comparisons of music and language processing and from comparisons of music processing in human and nonhuman listeners. It may also result in underestimation of the impact of music on human listeners.

## Acknowledgments

We thank Stephanie Stalinski, Kyle Heffernan, and Harry Knazan for their musical performances and Aranda Wingsiong for assistance in testing.

## Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

## Funding

The Natural Sciences and Engineering Research Council (NSERC) of Canada (S. E. T. and E. G. S.) and an NSERC-CREATE award (M. W. W.) provided funding for this research.

## Supplemental Material

Additional supporting information may be found at <http://pss.sagepub.com/content/by/supplemental-data>



## References

- Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research*, *13*, 17–26.
- Bregman, M. R., Patel, A. D., & Gentner, T. Q. (2012). Stimulus-dependent flexibility in non-human auditory perception. *Cognition*, *122*, 51–60.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.
- D'Amato, M. R. (1988). A search for tonal pattern perception in cebus monkeys: Why monkeys can't hum a tune. *Music Perception*, *5*, 453–480.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*, 1174–1176.
- Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representations of the fingers of the left hand in string players. *Science*, *270*, 305–307.
- Ghazanfar, A. A., Turesson, H. K., Maler, J. X., van Dinther, R., Patterson, R. D., & Logothetis, N. K. (2007). Vocal-tract resonances as indexical cues in rhesus monkeys. *Current Biology*, *17*, 425–430.
- Gunji, A., Koyama, S., Ishii, R., Levy, D., Okamoto, H., Kakigi, R., & Pantev, C. (2003). Magnetoencephalographic study of the cortical activity elicited by human voice. *Neuroscience Letters*, *348*, 13–16.
- Halpern, A. R., & Müllensiefen, D. (2008). Effects of timbre and tempo change on memory for music. *Quarterly Journal of Experimental Psychology*, *61*, 1371–1384.
- Hulse, S. H., & Cynx, J. (1985). Relative pitch perception is constrained by absolute pitch in songbirds (*Mimus*, *Molothrus*, and *Sturnus*). *Journal of Comparative Psychology*, *99*, 176–196.
- Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, *126*, 159–179.
- Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, *56*, 414–423.
- Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, *58*, 927–935.
- Mithen, S. (2005). *The singing Neanderthals: The origins of music, language, mind, and body*. Cambridge, MA: Harvard University Press.
- Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *NeuroReport*, *12*, 169–174.
- Peretz, I., & Gagnon, L. (1998). Dissociation between recognition and emotional judgments for melodies. *Neurocase*, *5*, 21–30.
- Peretz, I., Gaudreau, D., & Bonnel, A.-M. (1998). Exposure effects on music preference and recognition. *Memory & Cognition*, *26*, 884–902.
- Perrodin, C., Kayser, C., Logothetis, N. K., & Petkov, C. I. (2011). Voice cells in the primate temporal lobe. *Current Biology*, *21*, 1408–1415.
- Pollard, K. A., & Blumstein, D. T. (2011). Social group size predicts the evolution of individuality. *Current Biology*, *21*, 413–417.
- Risset, J.-C., & Wessel, D. L. (1999). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 113–169). San Diego, CA: Academic Press.
- Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review*, *6*, 641–646.
- Schellenberg, E. G., Peretz, I., & Vieillard, S. (2008). Liking for happy- and sad-sounding music: Effects of exposure. *Cognition & Emotion*, *22*, 218–237.
- Schellenberg, E. G., & Trehub, S. E. (2003). Accurate pitch memory is widespread. *Psychological Science*, *14*, 262–266.
- Stalinski, S. M., & Schellenberg, E. G. (in press). Listeners remember music they like. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Strait, D. L., & Kraus, N. (2011). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology*, *2*, 113. doi:10.3389/fpsyg.2011.00113
- Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 370–381.
- Trehub, S. E., Schellenberg, E. G., & Nakata, T. (2008). Cross-cultural perspectives on pitch memory. *Journal of Experimental Child Psychology*, *100*, 40–52.
- Volkova, A., Trehub, S. E., & Schellenberg, E. G. (2006). Infants' memory for musical performances. *Developmental Science*, *9*, 583–589.
- Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, *10*, 159–171.
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*, 547–558.