Brain and Cognition 129 (2019) 35–39

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Ability to process musical pitch is unrelated to the memory advantage for vocal music[☆]

Michael W. Weiss*, Isabelle Peretz

International Laboratory for Brain, Music, and Sound Research, Canada
Department of Psychology, Université de Montréal, Montréal, Canada

ARTICLE INFO

Keywords:
Amusia
Recognition memory
Timbre
Voice
Pitch

ABSTRACT

Listeners remember vocal melodies better than instrumental melodies, but the origins of the effect are unclear. One explanation for the 'voice advantage' is that general perceptual mechanisms enhance processing of conspecific signals. An alternative possibility is that the voice, by virtue of its expressiveness in pitch, simply provides more musical information to the listener. Individuals with congenital amusia provide a unique opportunity to disentangle the effects of conspecific status and vocal expressiveness because they cannot readily process subtleties in musical pitch. Forty-one participants whose musical pitch discrimination ability ranged from congenitally amusic to typical were tested. Participants heard vocal and instrumental melodies during an exposure phase, and heard the same melodies intermixed with timbre-matched foils in a recognition phase. Memory was better for vocal than instrumental melodies, but the magnitude of the advantage was unrelated to musical pitch discrimination or memory overall. The voice enhances melodic memory regardless of music perception ability, ruling out the role of pitch expressiveness in the voice advantage. More importantly, listeners across a wide range of musical ability can benefit from the privileged status of the voice.

1. Introduction

Listeners remember melodies sung by a voice (without lyrics) better than melodies played on instruments (Weiss, Trehub, & Schellenberg, 2012), but the origins of the 'voice advantage' are unclear. One explanation is that vocal timbre holds a special status as a conspecific signal. Compared to other sounds, human vocalizations elicit enhanced attention from early in life (Vouloumanos & Werker, 2007), enhanced arousal compared to instrumental music (Weiss, Trehub, Schellenberg, & Habashi, 2016), and distinct neural processing (Belin, Zatorre, & Ahad, 2002; Belin, Zatorre, Lafaille, Ahad, & Pike, 2000). A 'conspecific advantage' in attention or arousal could in turn facilitate memory.

An alternative hypothesis is that sung music differs acoustically from instrumental music in ways that benefit memory. Pitch is arguably the primary determinant of melodic identity in Western music, and vocal pitch is expressive relative to other timbres. The vocal apparatus creates coarticulation between adjacent notes in the form of "scoops" in pitch (see Fig. 1). In occasional or everyday singers, such deviations range in subtlety depending on location within a note and are typically smaller than the distance between adjacent notes in the Western scale (i.e., 100 cents or 1 semitone) (Larrouy-Maestri & Pfordresher, 2018, Appendix A). Scoops are perceptually relevant because they often exceed the threshold for detection of a mistuned vocal note in typical listeners (~60 cents; Hutchins, Roquet, & Peretz, 2012). Scoops are musically relevant because they indicate orientation of the intended pitch relative to adjacent notes, which influences perception of intonation (Larrouy-Maestri & Pfordresher, 2018). In theory, compared to a fixed-pitch instrument like a piano, vocal pitch provides additional information about note continuity that could reinforce encoding of auditory sequences.

The separate roles of conspecific status and fine-grained pitch processing in the voice advantage remain unexplored because they are linked acoustically. Removing pitch variation from singing causes the voice to sound artificial or 'autotuned'. An alternative approach is testing individuals with difficulty processing subtle changes in musical pitch.

Roughly 1.5% of the population have lifelong difficulty with musical pitch that cannot be attributed to general auditory deficiencies

---

[☆] Funded by the Canadian Institutes of Health Research and from the program of Canada Research Chairs (IP) and the Québec Research Fund – Nature and Technologies (MWW).

* Corresponding author at: Department of Psychology, Université de Montréal, Marie-Victorin Building, PO Box 6128, Centre-ville STN, Montréal, QC H3C3J7, Canada.

E-mail address: michael.weiss@umontreal.ca (M.W. Weiss).

https://doi.org/10.1016/j.bandc.2018.11.011
Received 3 August 2018; Received in revised form 26 November 2018; Accepted 27 November 2018

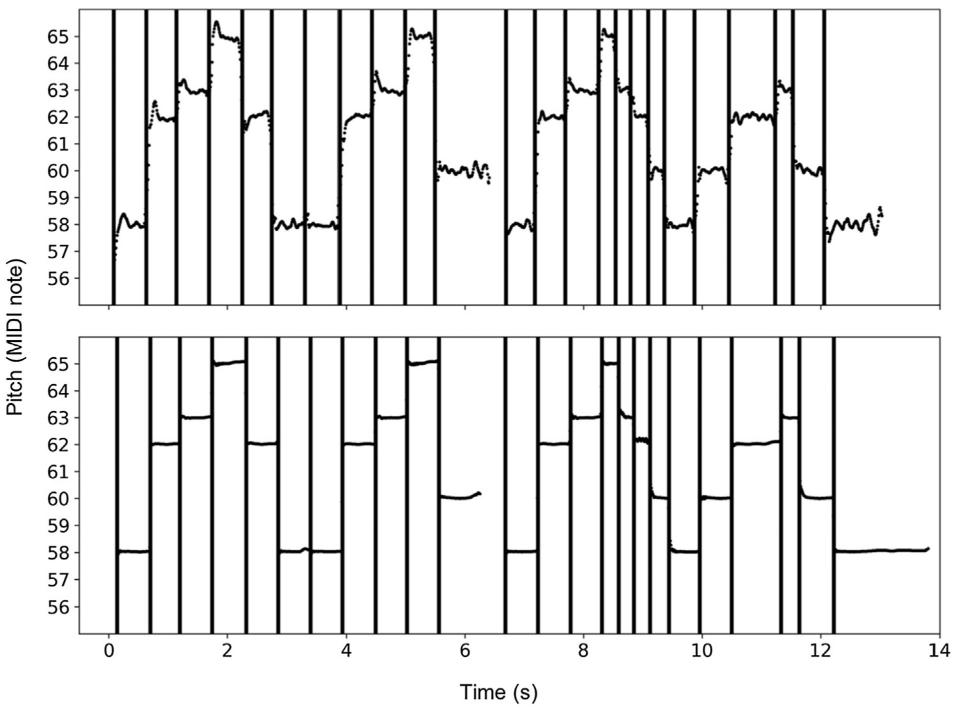

Fig. 1. Pitch expressiveness across timbres. Continuous pitch plotted for a representative vocal performance (upper panel; amateur female singer) and piano performance (lower panel) of the same stimulus melody. Vertical lines are note onsets. Pitch values were extracted using Praat (Boersma & Weenink, 2018) every 10 ms and converted to cents. The vertical axis represents MIDI note values (i.e., 1 semitone or 100 cents apart). Sung notes are dynamic and often contain "scoops" in pitch at onset or offset that connect adjacent notes. Piano notes have fixed pitch.

(Peretz & Vuvan, 2017). This condition is called congenital amusia and refers to a difficulty detecting pitch alterations in a melody of roughly three to seven semitones (i.e., 300–700 cents) or less, even when the alternate pitch violates musical conventions (e.g., being 'out-of-tune'). The amusic brain processes incoming pitch information normally, as evidenced by typical mismatch negativity to out-of-tune notes during electroencephalography (EEG), but exhibits abnormal conscious detection of pitch changes, as indicated by the absence of later components (i.e., P300/P600) and behavior (e.g., Moreau, Jolicœur, & Peretz, 2013).

The structure and functional connectivity of the amusic brain differs from that of typical listeners along a right frontotemporal network (including the right inferior frontal gyrus, the auditory cortex, and the arcuate fasciculus), which, taken together with findings from EEG, suggests abnormal top-down or recurrent feedback (Peretz, 2016). As a result, amusics perform poorly on tasks of long-term memory for music (Vuvan et al., 2018). Amusics also perform worse than controls in short-term memory for timbre and sorting timbres into categories, but critically, they perform normally in timbre discrimination (Marin, Gingras, & Stewart, 2012).

Amusia provides a unique opportunity to contrast the roles of conspecific signal processing and pitch processing in the voice advantage. Both typical and amusic listeners would be expected to benefit from a conspecific advantage because perception of sound source or timbre (i.e., perceiving a voice versus an instrument) is unrelated to the amusia deficit. In contrast, only typical listeners would benefit from inter- and intra-note pitch dynamics in singing because those pitch dynamics occur below amusics' threshold for detecting musical pitch changes.

In the present study, amusic and control participants completed a long-term memory task for vocal and instrumental melodies. Two meaningful outcomes are possible. The first possibility is that amusics show a different magnitude of voice advantage than controls, or framed another way, that musical pitch processing ability correlates with the voice advantage. This outcome is not compatible with the conspecific advantage hypothesis, but it is compatible with the vocal pitch expressivity advantage hypothesis. The second possibility is that the voice advantage is present across the sample and unrelated to group or individual musical pitch processing ability. This outcome is not compatible with the vocal pitch expressivity hypothesis, but it is compatible with the conspecific advantage hypothesis.

2. Method

2.1. Participants

Participants were recruited from a database of individuals with or without musical pitch discrimination deficits, as well as from local advertisements. Forty-one participants (29 female, $M = 54.2 \pm 20.3$, $range = 19–79$ years) were included in the final sample. One additional participant was excluded due to deviations from protocol and incomplete data. Years of formal musical training were skewed across the sample ($M = 2.0 \pm 2.5$, $median = 1$, $range = 0–9$ years). Informed consent was obtained from all participants. All participants received token remuneration.

Musical pitch discrimination was measured using the melodic subtests of the Montreal Battery of Evaluation of Amusia or MBEA (Subtests: *scale*, *contour*, *interval*; Peretz, Champod, & Hyde, 2003) and the Online Test for Amusia (Subtests: *scale*, *off-key*; Peretz & Vuvan, 2017). All subtests involve listening to short melodies for melodic changes or violations of Western tonality. In the MBEA melodic subtests, pairs of melodies are presented which match or contain a single pitch change that violates key but maintains contour (*scale* subtest, 30 trials), changes contour but maintains key (*contour* subtest, 30 trials), or changes interval but maintains key and contour (*interval* subtest, 30 trials). Half of the trials of the melodic subtests contained a median pitch change of 400 cents ($mode = 400$, $range = 300–700$). In the *off-key* subtest of the Online Test for Amusia listeners must detect whether an incongruity (out-of-key note) has occurred in a single melody (24 trials). The *scale* subtest in the Online Test for Amusia is identical to the *scale* subtest in the MBEA. Thirty-six participants in the current sample completed both the MBEA and the Online Test for Amusia, two completed the MBEA only, and three completed the Online Test for Amusia only, so musical pitch discrimination was calculated as the average percent correct on all available subtests.

Musical pitch discrimination ability was distributed widely in the current sample from near-chance to near-perfect performance ($M = 74.8 \pm 15.5\%$, $median = 72.5\%$, $range = 49.9–96.5\%$ correct, chance = 50%). The MBEA and Online Test for Amusia were designed to screen individuals for amusia using a performance criterion of 2 SD below the population mean, in addition to other protocol (Vuvan et al., 2018). Across the melodic subtests of the MBEA, the average cutoff score is 72.2% (Peretz et al., 2003). For group analyses in the current study, participants with musical pitch discrimination scores below 72.2% were classified as amusics ($n = 20$) and remaining participants were classified as controls ($n = 21$). For follow-up analyses, musical pitch discrimination was considered as a continuous measure.

Age was not a variable of interest, so participants were recruited to roughly match the age of those with higher and lower musical pitch discrimination ability. The number of older (50 +) controls and amusics was matched ($n = 13$ each) and the number of younger (< 50) controls ($n = 8$) and amusics ($n = 7$) was similar. There was no difference in age between the amusic ($M = 54.7$, $SD = 19.7$, $range = 19.6–76.0$) and control ($M = 53.7$, $SD = 21.3$, $range = 20.9–79.6$) groups, $p > .8$. Measurements of hearing ability were available for most participants older than 50 ($n = 25$ of 26) and did not differ between groups, $p > .1$.

2.2. Stimuli

Stimuli were 24 unfamiliar British and Irish folk melodies (Weiss et al., 2012). Melodies varied in length (13–20 s), time signature (3/4, 4/4, 6/8), tempo (100–130 beats per minute), and number of notes (22–57). Each was recorded in three timbres (voice, piano, and marimba) by amateur musicians. Vocal performances were sung without lyrics (‘*lala*’) by a female vocalist and adjusted to be in-tune using Melodyne (Celemony, Inc.). Importantly, tuning correction in Melodyne (for pitch center and for pitch drift) maintains the dynamics of vocal pitch and sounds natural (see Weiss et al., 2012). An analysis of pitch scoops in the vocal melodies, using pitch values calculated in Praat (Boersma & Weenink, 2018) and the method of Larrouy-Maestri and Pfordresher (2018), showed an average absolute scoop magnitude of 99 cents ($SD = 109$, $max = 831$ cents). Approximately 97% of scoops were lesser in magnitude than the median pitch change on the MBEA (i.e., 400 cents). All stimuli were normalized to a common amplitude (RMS; Audiofile Engineering) and exported in high-quality (44.1 kHz, 16-bit).

For each participant, melodies were assigned to timbre at random ($n = 8$ each for voice, piano, marimba). Half of the melodies in each timbre were assigned at random to be heard during the exposure phase and the test phase (i.e., $n = 4$ ‘old’ melodies per timbre), or the test phase only (i.e., $n = 4$ ‘new’ melodies per timbre).

2.3. Apparatus

Participants were tested individually in a sound-attenuating booth (Industrial Acoustics, Inc.). The study was programmed in a custom script on Mac OS and presented over high-quality headphones (Beyerdynamic DT 990 Pro). Written instructions were presented in both English and French.

2.4. Procedure

Before the task, participants practiced using the program and adjusted the volume to a comfortable level.

In the exposure phase, half of the melodies in the set ($n = 12$, with 4 per timbre) were presented twice in two separately randomized blocks, for 24 trials total. To maintain attention, participants rated each melody from *‘1 – Dislike’* to *‘5 – Like’*. A 5–10 min. break followed the exposure phase, during which participants filled out a background information questionnaire.

After the break, participants completed a surprise recognition test. Participants heard the same 12 melodies as before (i.e., 4 ‘old’ melodies per timbre) intermixed with the remaining 12 melodies in the set (i.e., 4 ‘new’ melodies per timbre) and rated their confidence that each melody was new or old on a scale from *‘1 – Definitely New’* to *‘7 – Definitely Old’*.

3. Results

A preliminary analysis using pairwise comparisons (Bonferroni) confirmed that recognition (i.e., AUC scores; see Recognition) did not differ for melodies played by the piano and marimba, $p > .1$, in line with previous research (Weiss et al., 2012). Instrumental ratings were collapsed in subsequent analyses.

3.1. Liking

Four scores were calculated for each participant by averaging liking ratings by timbre class (voice, instrument) and exposure block (1, 2) separately for amusic and control listeners. A mixed-model ANOVA with within-participant factors for timbre type (voice, instrument) and exposure block (1, 2) and between-participant factor for group (amusic, control), revealed no main effects or interactions, $ps > .1$.

3.2. Recognition

Two scores were calculated for each participant by converting recognition ratings to area under the receiver-operating-characteristic curve (AUC), separately for each timbre class (voice, instrument). AUC provides an unbiased measure of proportion correct by comparing hits and false alarms at all levels of the rating scale, with chance performance at 0.5 and perfect performance at 1.0 (Swets, 1973). A mixed-model ANOVA with a within-participant factor for timbre class (voice, instrument) and a between-participant factor for group (amusic, control) revealed that AUC scores differed by group, $F(1, 39) = 17.33$, $p < .001$, $\eta_p^2 = 0.30$, and by timbre class, $F(1, 39) = 6.57$, $p = .014$, $\eta_p^2 = 0.14$, with no interaction, $F(1, 39) = 1.08$, $p > .3$, $\eta_p^2 = 0.03$. Descriptive statistics (Fig. 2) show that musical memory was better for control than amusic participants across timbres, as in previous research (Peretz et al., 2003), and better for vocal than instrumental melodies across groups, as in previous research (Weiss et al., 2012, 2016). There was no indication that the magnitude of the voice advantage was reduced by amusic status. On the contrary, as visible in Fig. 2, amusics

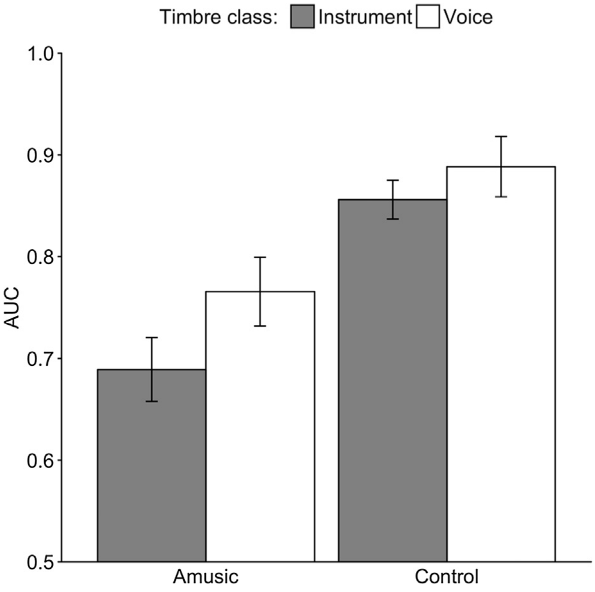

Fig. 2. Mean memory performance. Recognition memory, as measured by area under the receiver operating characteristic curve (AUC; chance = 0.5, perfect = 1.0), was better for vocal than instrumental melodies and better for control than amusic participants. There was no interaction of timbre class and group.

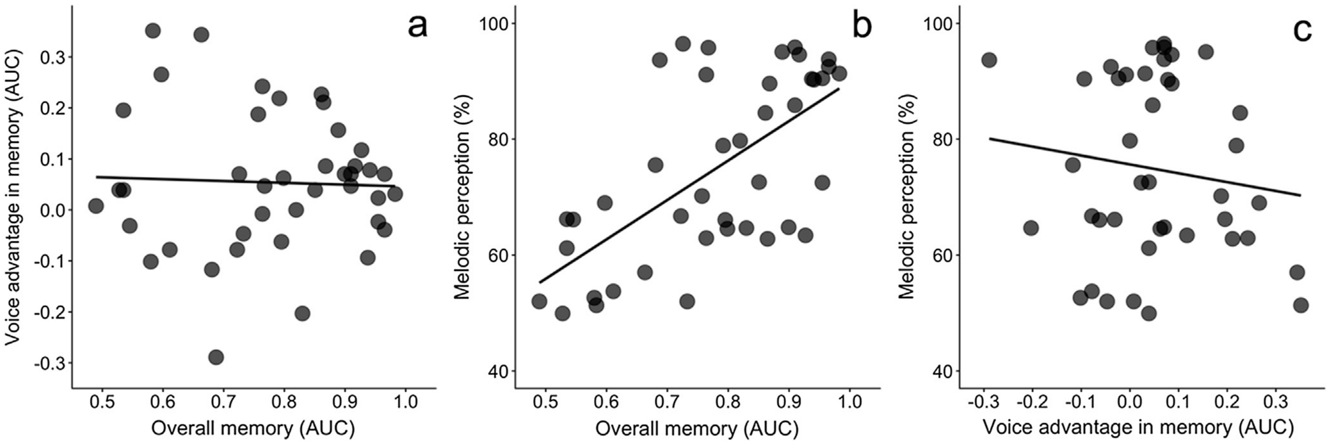

Fig. 3. Memory and musical pitch discrimination. Overall performance on the recognition task was not correlated with the magnitude of the voice advantage in memory (a), but it was positively correlated with melodic perception ability (b). Melodic perception ability did not correlate with the magnitude of the voice advantage in memory (c).

showed a *larger* voice advantage than controls in an absolute sense.

3.3. Individual differences in recognition and musical pitch perception

For each participant, a measure of the voice advantage was calculated as the difference between AUC scores by timbre class (i.e., $\text{AUC}_{\text{Voice}} - \text{AUC}_{\text{Instrument}}$). A measure of overall performance was calculated as AUC for all trials regardless of timbre. There was no correlation between overall memory and the magnitude of the voice advantage, $r(39) = -0.04$, $p > .8$ (Fig. 3a).

Two correlations assessed the relationship between individual musical pitch discrimination ability and (1) overall memory performance and (2) the magnitude of the voice advantage. There was a positive correlation between musical pitch discrimination and overall memory, $r(39) = 0.64$, $p < .001$ (Fig. 3b), in line with the memory impairment found in amusia. In contrast, there was no correlation between individual musical pitch discrimination ability and the voice advantage, $r(39) = -0.13$, $p > .4$ (Fig. 3c). This result parallels the absence of a group by timbre class interaction in the previous section. The negative direction of the correlation, albeit non-significant, indicates that increasing sample size would not yield a significant *positive* correlation between individual musical pitch discrimination ability and the voice advantage.

A final analysis considered whether age correlated to the measures of interest. There was no correlation between age and the magnitude of the voice advantage, $r(39) = -0.13$, $p > .4$. There was a significant correlation between age and overall memory, $r(39) = -0.36$, $p = .019$, but this result is independent from musical pitch discrimination ability because the ages of amusics and controls were matched during recruitment (see Participants). Indeed, a partial correlation holding constant age revealed a nearly identical relationship as reported above between musical pitch discrimination and overall memory, $r(38) = 0.65$, $p < .001$.

4. Discussion

The current study examined the relationship between pitch processing ability and memory for music in vocal and instrumental timbre in amusic and control listeners with a wide range of musical pitch-processing ability. As in previous research, there was a memory advantage for vocal melodies. The ability to process musical pitch correlated positively with overall musical memory, but importantly, neither musical pitch discrimination ability nor overall musical memory correlated positively with the magnitude of the voice advantage. These findings indicate that the voice advantage is not caused by expressive vocal pitch. Rather, the results are consistent with a voice enhancing memory due to its privileged status as an engaging, conspecific signal.

The current results suggest that the mechanisms involved in perceiving conspecific vocalizations are intact in amusia despite abnormalities in processing musical sounds. Vocal music is rarely used in tests of perception and memory in the amusia literature, but it may confer benefits in music processing. Pitch processing deficits in amusia have been linked to abnormal recurrent processing in a right frontotemporal network (Peretz, 2016). It is possible that the mechanisms involved in the voice advantage might circumvent or attenuate the effects of those anomalous processes. If so, vocal training may be a more consequential rehabilitation strategy than instrumental practice. Preliminary evidence of the effectiveness of treatment is mixed. In one case study, training with vocal music had no long-term effect on music perception ability (Wilbiks, Vuvan, Girard, Peretz, & Russo, 2016), and a small group intervention with vocal training showed only modest signs of improvement in perception (Anderson, Himonides, Wise, Welch, & Stewart, 2012). The role of the voice is impossible to infer from these studies because there was no comparison group with instrumental training.

Finally, the current results raise the possibility of finding a reverse dissociation in participants with a voice-processing deficit but no pitch-processing deficit. For example, individuals with social-communicative disorders (e.g., Autism Spectrum Disorder; ASD) process music normally (Stanutz, Wapnick, & Burack, 2014) but process speech abnormally (Gervais et al., 2004), and may demonstrate no voice advantage. Alternatively, since fronto-temporal connectivity is preserved for singing but not speech in ASD individuals (Sharda, Midha, Malik, Mukerji, & Singh, 2015), a vocal advantage may be observed. Further exploration of the presence or absence of the voice advantage across different populations can help reveal the limits of enhanced cognition from conspecific signals.

5. Conclusion

The processing of vocal music is enhanced compared to instrumental music, even in individuals with deficiencies in musical pitch perception and long-term memory for music (i.e., congenital amusics). Therefore, it is unlikely that the encoding of vocal pitch dynamics plays a significant role in the voice advantage. The current sample represents the widest range of musical ability tested to date on this task, and the results complement previous findings that expert musicians and non-musicians have comparable recognition advantages for vocal music even as they differ in overall recognition (Weiss, Vanzella, Schellenberg, & Trehub, 2015). Collectively the data suggest that the voice advantage is not influenced by musical ability or musical training. If instead the voice advantage results from attention or arousal to conspecific vocalizations (Weiss et al., 2016), then vocalizations have the potential to enhance memory and cognition for a variety of auditory stimuli, musical or otherwise.

Declarations of interest

None.

References

Anderson, S., Himonides, E., Wise, K., Welch, G., & Stewart, L. (2012). Is there potential for learning in amusia? A study of the effect of singing intervention in congenital amusia. *Annals of the New York Academy of Sciences, 1252*, 345–353. https://doi.org/10.1111/j.1749-6632.2011.06404.x.

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research, 13*, 17–26. https://doi.org/10.1016/S0926-6410(01)00084-2.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403*, 309–312. https://doi.org/10.1038/35002078.

Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [Computer program]. retrieved from *Version, 6*, 43.

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., ... Zilbovicius, M. (2004). Abnormal cortical voice processing in autism. *Nature Neuroscience, 7*, 801–802. https://doi.org/10.1038/nn1291.

Hutchins, S., Roquet, C., & Peretz, I. (2012). The vocal generosity effect: How bad can your singing be? *Music Perception, 30*, 147–159. https://doi.org/10.1525/MP.2012.30.2.147.

Larrouy-Maestri, P., & Pfordresher, P. Q. (2018). Pitch perception in music: Do scoops matter? *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/xhp0000550.

Marin, M. M., Gingras, B., & Stewart, L. (2012). Perception of musical timbre in congenital amusia: Categorization, discrimination and short-term memory. *Neuropsychologia, 50*, 367–378. https://doi.org/10.1016/j.neuropsychologia.2011.12.006.

Moreau, P., Jolicœur, P., & Peretz, I. (2013). Pitch discrimination without awareness in congenital amusia: Evidence from event-related potentials. *Brain and Cognition, 81*, 337–344. https://doi.org/10.1016/j.bandc.2013.01.004.

Peretz, I. (2016). Neurobiology of congenital amusia. *Trends in Cognitive Sciences, 20*, 857–867. https://doi.org/10.1016/j.tics.2016.09.002.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders: The montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences, 999*, 58–75. https://doi.org/10.1196/annals.1284.006.

Peretz, I., & Vuvan, D. T. (2017). Prevalence of congenital amusia. *European Journal of Human Genetics, 25*, 625–630. https://doi.org/10.1038/ejhg.2017.15.

Sharda, M., Midha, R., Malik, S., Mukerji, S., & Singh, N. C. (2015). Fronto-temporal connectivity is preserved during sung but not spoken word listening, across the autism spectrum. *Autism Research, 8*, 174–186. https://doi.org/10.1002/aur.1437.

Stanutz, S., Wapnick, J., & Burack, J. A. (2014). Pitch discrimination and melodic memory in children with autism spectrum disorders. *Autism, 18*, 137–147. https://doi.org/10.1177/1362361312462905.

Swets, J. A. (1973). The relative operating characteristic in psychology. *Science, 182*, 990–1000. https://doi.org/10.1126/science.182.4116.990.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science, 10*, 159–164. https://doi.org/10.1111/j.1467-7687.2007.00549.x.

Vuvan, D. T., Paquette, S., Mignault Goulet, G., Royal, I., Felezeu, M., & Peretz, I. (2018). The montreal protocol for identification of amusia. *Behavior Research Methods, 50*, 662–672. https://doi.org/10.3758/s13428-017-0892-8.

Weiss, M. W., Trehub, S. E., & Schellenberg, E. G. (2012). Something in the way she sings: Enhanced memory for vocal melodies. *Psychological Science, 23*, 1074–1078. https://doi.org/10.1177/0956797612442552.

Weiss, M. W., Trehub, S. E., Schellenberg, E. G., & Habashi, P. (2016). Pupils dilate for vocal or familiar music. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 1061–1065. https://doi.org/10.1037/xhp0000226.

Weiss, M. W., Vanzella, P., Schellenberg, E. G., & Trehub, S. E. (2015). Pianists exhibit enhanced memory for vocal melodies but not piano melodies. *The Quarterly Journal of Experimental Psychology, 68*, 866–877. https://doi.org/10.1080/17470218.2015.1020818.

Wilbiks, J. M. P., Vuvan, D. T., Girard, P. Y., Peretz, I., & Russo, F. A. (2016). Effects of vocal training in a musicophile with congenital amusia. *Neurocase, 22*, 526–537. https://doi.org/10.1080/13554794.2016.1263339.