A. M. Green, C. E. Chapman, J. F. Kalaska and F. Lepore (Eds.)
Progress in Brain Research, Vol. 191
ISSN: 0079-6123

CHAPTER 7

Perception and action in singing

Sean Hutchins* and Isabelle Peretz

BRAMS Laboratory and Department of Psychology, Université de Montréal, Montréal, Québec, Canada

Abstract: Singing is an important cultural activity, yet many people are hesitant to sing, because they feel they do not sing well. This chapter reviews the work that has been done concerning singing among nonmusicians, focusing on pitch accuracy, which is one of the most important aspects of singing. First, we look at the prevalence of poor pitch singing and examine what it means to be a poor singer. Next, we look at the possible causes of poor singing and examine the possible roles of perceptual deficits, sensorimotor translation deficits, motor control deficits, and feedback deficits. Whereas many previous studies have tried to explain poor singing by positing perceptual problems, we argue that the current evidence supports sensorimotor translation deficits and motor control deficits as the more likely causes. Finally, we examine the neural bases of singing and the possibility of a dual-pathway basis for pitch perception and production. Based on these studies, we suggest changes to improve the singing abilities of poor singers.

Keywords: singing; voice perception; vocal pitch matching.

Introduction

Sing the first few lines of your national anthem right now. Go ahead, sing it. Did you? I bet you didn't. I bet you thought about your national anthem, maybe sung it in your head, maybe whistled it, or even hummed it under your breath, but you probably did not sing it. Why? Most people feel self-conscious about singing, and especially singing alone in front of others. While most people would not hesitate to join into a chorus of Happy Birthday, even at a crowded restaurant, they would not opt to do so in public without the reinforcing effect of others, even though these same people might not hesitate to sing alone in the car or the shower. Singing occupies an odd place in our culture. Whereas in other cultures, people sing regularly as part of their work or in gatherings, in Western society, singing has largely been relegated to a professional activity, where well-trained singers produce music with the benefit of microphones, recording studios, and even auto-tuning, in which the notes produced by the singer are automatically corrected by a computer. Thus, despite the fact that most people listen to music frequently and with pleasure, they

*Corresponding author.
Tel.: +1 514 343 3432; Fax: +1 514 343 2175.
E-mail: sean.michael.hutchins@umontreal.ca

DOI: 10.1016/B978-0-444-53752-2.00010-2

rarely hear live, unadulterated singing. Most instances of singing are heard with a type of perfection imposed upon them, often with a vocal quality that is impossible for a real human voice to reproduce without technological aid. Even more surprising, to an outsider, would be the fact that most Westerners rarely engage with music in a participatory setting, and if you ask people why they do not sing more often, the most common response you will hear is "I don't sing very well."

To the best of our knowledge, no serious cross-cultural study has been performed comparing the singing abilities of people from societies with and without large-scale exposure to professionally recorded music. However, participatory singing occupies a much larger part of people's activities in non-Westernized cultures, and there is certainly not this widespread belief among people that they are unable to sing well. For example, an informal class survey showed that nearly three out of five students indicated that they could not imitate a melody with their voice (Pfordresher and Brown, 2007). It is possible that people from Western societies sing worse, on average, than those from non-Western societies, in which case this would likely be attributable to comparative lack of practice. It is also possible that there is no difference in singing abilities between the two groups. In this case, we would need to explain why so many people think that they sing so poorly. One of the main explanations for this effect could be the professionalization of singing. By hearing mostly audio-engineered examples of trained singers, it is likely that our definition of what constitutes good, or even adequate singing has been shifted.

Regardless of the nature of Western people's singing inadequacies, be they real or only perceived, one simple way of making people feel more comfortable in participating in singing activities is improving their singing abilities. But what does good singing ability actually entail? In a survey of over 1000 music educators, Watts et al. (2003) found that the pitch intonation was the most important factor of determining someone's singing ability. This factor was rated above other factors such as musicality, vocal quality (timbre), range, diction, and others. The importance of intonation was later confirmed by Dalla Bella et al. (2007). It is telling that the most salient examples of vocal audio engineering involve pitch correction, especially auto-tuning, which is now even a part of many live shows. It is also telling that, in shows such as American Idol, the most jarring examples of poor singing tend to come from participants who do not sing in tune. Thus, one of the easiest ways to enhance people's musical performance, and to get people interested in performing, as well as consuming music, is to help them to improve their pitch intonation.

How well do people sing?

How well do people sing? Or, more specifically, how correct or incorrect are people when singing an intended pitch? A few recent studies have attempted to quantify the pitch errors of musically untrained individuals. One such study was conducted by Dalla Bella et al. (2007). They recruited musically untrained singers and asked them to sing "Gens du Pays," a very well known Québecois song typically sung for birthday celebrations. Some of these participants were recruited to come in to sing in a laboratory setting, but the majority of them were approached and sang for the experimenter in a more ecologically valid situation. In these cases, the experimenter approached the participants at a public park and asked them to sing "Gens du Pays" to him for his birthday. These were recorded with a portable microphone and recorder and later brought back to the laboratory for analysis. Among both those who sang in the lab and outdoors, an average of five to six intervals (an interval is the pitch distance between two successive notes) out of 31 were classified as errors. However, these data were not normally distributed: about 30% of singers made 0- or 1-pitch interval errors, whereas a small number of singers made significant errors on over half of all intervals, which would make the song practically unrecognizable without words. Overall, about 80% of all of the song's intervals were sung within one semitone of the intended interval, which the authors interpreted as a reasonably high level of accuracy. However, interestingly, they found that the singers recruited in the park performed at a significantly lower level than those brought into the lab, despite the fact that the laboratory participants were specifically recruited to be nonmusicians, whereas the outdoors singers were not screened for musical ability. This may speak to a selection bias for participants coming to the laboratory, or less motivation for accurate singing among the outdoors singers, but it is worth keeping in mind as an important caveat for other studies of vocal pitch matching that take place in a laboratory.

Another study by Pfordresher and Brown (2007) looked at singing ability through the lens of simpler tasks. They asked 79 participants to reproduce four-note sequences. These sequences were of three types: all the same pitch, two different pitches that formed one interval, or four unique pitches. In this sample, 87% of the participants could sing back the pitches within plus or minus one semitone. They also found that errors in producing the correct pitch interval tended to co-occur with errors in producing the correct absolute pitch. Singers who sang the wrong pitches tended to reduce pitch interval size, and Pfordresher and Brown (2007) concluded that poor singing was primarily related to vocal pitch-matching—rather than interval-matching—ability.

In our own work, we have endeavored to examine vocal pitch-matching ability in musicians and nonmusicians. We have used simple pitch-matching tasks to measure this ability without any intervening complications, such as time pressure or memory effects. Our results show that almost 50% of nonmusicians failed to match a target pitch to within half a semitone on half of their attempts (Hutchins and Peretz, in preparation-a). This is a considerably higher rate of poor singing than was found in Pfordresher and Brown (2007) and Dalla Bella et al. (2007), who estimated that only about 10–20% of the population should be classified as poor singers. There are a few reasons why this might be the case. First, our task was simpler in nature and, because of this, may have provided fewer cues to aid singers in finding the correct pitch. Whereas the other studies provide tonal contexts in which singers could use their implicit knowledge of key structure and the relationship between notes in a key, the tones in our study were not tonally related, and singers could not use intertrial relationships to help them determine the correct pitch of a trial. In fact, in Pfordresher and Brown's study (2007), where stimulus complexity was manipulated, there was a tendency for better pitch matching in the more complex stimuli among the poorer singers, which indicates that we should see more instances of poor singing in less complex tasks. In addition, poor singing was defined in slightly different ways across these studies. Dalla Bella et al. (2007) studied impromptu song performance and so could not measure pitch matching *per se*, but used only relative interval sizes to determine accuracy. Pfordresher and Brown (2007) measured both absolute and relative accuracy but used absolute accuracy to label singers as good or poor. Both of these, however, used plus or minus one semitone as the demarcation between good and poor singing, whereas we used plus or minus half a semitone. Both of these metrics have some validity to them, but the full semitone demarcation is significantly more forgiving. This can explain some of the difference in measured prevalence of poor singing between the studies, but even if we use the more liberal criterion of a full semitone, our data still show approximately 40% of singers as making more than 50% errors.

This criterion difference does raise an important issue, however, namely, what counts as poor singing? Do listeners think that a note that is half a semitone off from its intended target is out of tune? Or is it closer to a full semitone? One way to test this is to examine the minimum amount of pitch deviation that can be detected by listeners. Just noticeable differences for pitch can vary but are typically estimated at around 5 cents (i.e., 1/20th of a semitone; 100 cents = one semitone; Zwicker and Fastl, 1999). However, these can vary, with some expert listeners reporting just-noticeable differences for pitch even lower and many nonmusicians reporting thresholds closer to 15 or 20 cents. Klatt (1973) estimated the just-noticeable difference of the synthetic vowel / ε / at just under 5 cents. However, naturally sung tones are much more complex: they generally include many upper harmonics, some nonharmonic frequencies, and pitch fluctuations, such as vibrato (sinusoidal pitch variation) and unintended pitch change. These complexities will serve to make pitch judgments more difficult, and it is likely that the just-noticeable differences for sung tones are higher in these cases.

We have examined these perceptual limits for voice recordings. Participants listened to short melodies and single tones, either played on a violin or sung by a trained singer on the syllable/ba/. In the melody trials, the final tone of the melody was pitch-shifted anywhere from −100 to +100 cents, by 10-cent steps. In the single tone trials, we presented two tones, which could be the same or shifted by the same amounts as in the melody condition. In both conditions, the participants decided whether the final note was in tune (i.e., the same note, in single tone trials), the right note, but out of tune, or a wrong note altogether. Note that these types of judgments were shown to be equivalent for single tone and melodic contexts by Warrier and Zatorre (2002). Figure 1 shows the pattern of each type of judgment by nonmusicians for single tones which were sung versus played on the violin. On average, nonmusicians do not notice the tuning difference between two sung tones until they are different by at least 40 cents, whereas they can tell the difference between two violin tones that are only 30 cents apart. We have also replicated these single tone results in tuning judgments of untrained singers and synthesized vocal tones, showing that the higher thresholds for noticing tuning differences are not specific to trained singing voice or the violin.

However, most musical judgments of tuning are not made by asking whether two versions of the same tone are identical. They instead consist of judgments of whether a particular note fits into the musical context, which is what we measured with the melodic context trials. In their tuning judgments, nonmusicians needed the final tone to be mistuned by at least 60 cents before they recognized any mistuning, when the melody was produced by a voice. The smallest consistently noticeable difference was much lower, only 30 cents, when the melody was played on the violin. This difference tells us two things. First, the same tuning deviances are less noticeable when sung, compared with being played on a violin. Listeners seem to be either less likely to notice vocal mistuning than instrumental mistuning or more forgiving to singers than instrumentalists. This may be due to the properties of the voice, such as large amounts of vibrato; however, the violin uses vibrato as well (although not to the same extent). Second, these measurements confirm that one-half semitone (50 cents) is more likely a better criterion than 100 cents (one semitone) for determining whether a sung tone is correct or not. Most tones in the 50–100 cents error range are indeed heard as mistuned, both in an exact pitch-matching context and in a melodic tuning context, and it seems to be the case that if a tone is not produced within one-half semitone of the intended target, it will be heard as mistuned. Using these criteria, it seems the estimates of poor or unreliable singers may be nearly half the population (Hutchins and Peretz, in preparation-a). Thus, it is no surprise that people may be hesitant to sing in public.

What causes poor singing?

To improve adults' singing ability, it is important to know the root of their problem. A question like the causes of poor singing can have two types of answers, however. It is tempting to argue that inadequate practice is the cause of poor singing,

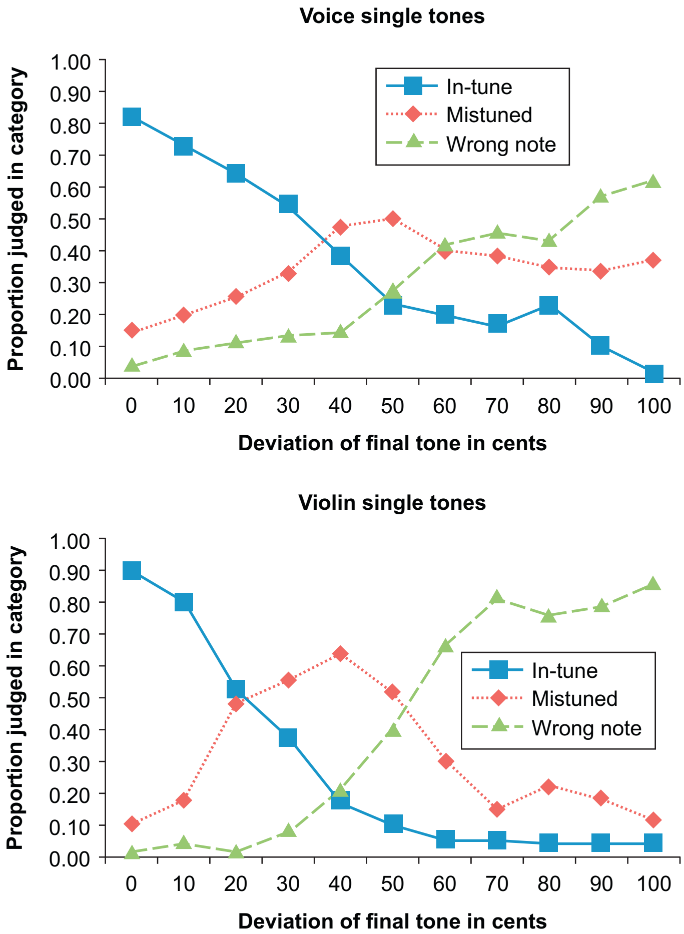

Fig. 1. The pattern of each type of judgment (in tune, the right note but mistuned, or wrong note altogether) by nonmusicians in the single tone context for sung melodies (a) and played on the violin (b).

and it is most likely true that greater amounts of practice, especially among children, will lead to better vocal performance. However, this teleological type of reasoning does not tell us the immediate causes of poor singing and fails to explain many salient cases, such as naturally gifted untrained singers, or people who have been through a comprehensive music education and still fail to match pitch adequately. It also does not give us an idea of what specific benefits practice can have or provide practical advice to music educators who would like to improve their pedagogical techniques and are unsure how to do so. Therefore, we will restrict our discussion to proximal causes of poor singing: the immediate physical and psychological impediments which lead to an erroneous response on a particular trial or set of trials.

Until recently, research on this topic had been quite limited, due to the technological constraints, and there were few studies of normal adult singing ability using acoustical measurements.

Several music education researchers tested children's singing abilities (e.g., Bentley, 1969; Cleall, 1979; Joyner, 1969) as a pedagogical tool, but many of these studies did not use rigorous scientific designs. However, there were some insights into the processes causing poor singing ability. For example, Bentley (1969) noted that monotone singers tended to have lower pitch-discrimination scores. Joyner (1969) followed on this work, evaluating the singing of a sample of "monotone" singers. He posited that three processes were necessary for good singing ability. First, singers must be able to discriminate pitches from each other. Second, singers must be able to recall the organization of pitches in a melody. Finally, singers must possess a vocal instrument capable of producing the intended pitches and able to respond quickly and accurately to their intentions. Joyner also suggested that "monotonism" was caused by a motor dysfunction and that perceptual problems may be a by-product, rather than a cause, of poor singing ability. Since the 1960s, other music educators have continued to research singing abilities in grade school children, using similar methods and elaborating upon these conclusions, often with the explicit aim of improving teaching techniques.

Since the turn of the century, though, there has been a revitalization of singing research. Because of the ease of applying acoustical analyses to digitally recorded data, there have been many studies on the prevalence of good and poor singing, the factors that count as good or poor singing, the root causes of poor singing, and external factors that can affect singing ability. The question of why some people cannot sing in tune is a particularly interesting one because singing is generally not a consciously accessible ability. That is, vocal production is such an automatic process, even among poor singers, that most people do not understand how they are using their bodies when they produce a note or how to adjust their vocal mechanism to change their sound. Those who can sing well typically cannot understand how someone could fail to sing well, and those who cannot sing well have no idea how to improve. Vocal pitch production, in fact, can be even harder to learn about than other automatic tasks, such as walking, because the relevant mechanisms to create vocal pitch are inside the body, and we only have direct access to the final outcome. This is why a lot of singing teachers use metaphor and imagery to try to improve their students' sound, with instructions like "breathe through your navel" and "pretend you are yawning"; these are simply more useful to many singers than "expand your pharyngeal cavity" or "tighten your thyroarytenoid muscles." These types of pedagogical strategies are helpful because of the lack of conscious knowledge of our vocal mechanism, which makes understanding the causes of poor singing all the more difficult to discover.

In their 2007 paper, Pfordresher and Brown listed four possible causes of inaccurate singing. To the three causes listed by Joyner—a perceptual deficit, a production deficit, and a memory deficit—they added a fourth, a sensorimotor mismapping. This refers to an inability to connect an accurately heard pitch to the same vocal response. Taking these into consideration, Fig. 2 shows the steps that must be undertaken in order for someone to sing back a pitch they hear. This figure models immediate reproduction and thus does not account for any possible effects of pitch or song memory. First, a sound is heard and the

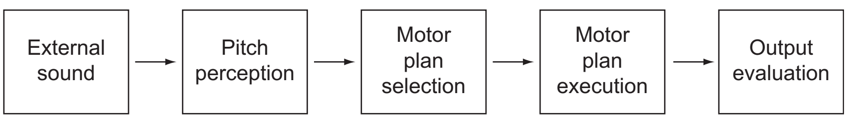

Fig. 2. The steps necessary to accurately sing back the pitch of a presented sound.

brain determines the pitch of the sound (note that this is often, but not always, the same as the fundamental frequency, the primary acoustic correlate of pitch). Second, an appropriate motor plan for producing that pitch must be accessed. Third, the motor plan must be acted upon and must be executed properly. Finally, the output can be evaluated, either from acoustical feedback or feedback from an efference copy of the motor command, and adjustments to the motor plan can be made accordingly. The next section of this review will deal with evidence for deficits at each stage in this process.

Perceptual deficits

Musicians and researchers both have long entertained the hypothesis that the reason some people do not sing in tune is that they cannot perceive the pitch accurately. However, in vocal pitch matching, one cannot separate the pitch perception and pitch production stages; it is necessary for a singer to perceive the intended pitch before it can be matched (excluding the relatively rare case of singers with absolute pitch (perfect pitch), who use long-term memory as a basis for their pitch matching). Therefore, one technique used to determine the relationship between the two processes is to measure perception separately, without invoking a vocal matching response and the vocal–motor code associated with it, and correlate this measurement with vocal pitch production abilities. However, the problem with this design is that polling perceptual abilities generally means requiring the use of other cognitive abilities which are not present in vocal pitch matching. Most often, these concern making same/different judgments about a pair of pitches. Several studies have reported correlations between pitch-discrimination abilities and vocal pitch accuracy (e.g., Estis et al., 2009, in press; Moore et al., 2007; Watts et al., 2005). However, there have been many other studies which failed to find such a relationship (e.g., Bradshaw and McHenry, 2005; Dalla Bella et al., 2007; Moore et al., 2008; Pfordresher and Brown, 2007).

The conflicting results make for a complicated story overall, and it is still in doubt whether there actually is a true correlation between pitch-discrimination abilities and vocal pitch accuracy. However, there are some major problems which are endogenous to the paradigm. First, there are some large variations between the methodologies used to measure perception and production abilities between experiments, with different studies using different comparison pitches, different timbres (i.e., the color of a note—this is what makes a piano sound different from a guitar), and different types of subjects. These differences between studies make it difficult to discover what type of consensus may underlie them; the variability in their findings may be more reflective of the methodological differences than the underlying relationship between perception and production.

Another major problem with these studies is their use of percentage of correct responses on the perception task as the variable to correlate with the pitch-matching ability. Percentage of correct responses does not reflect the same kind of measurement as error in vocal pitch matching. Vocal pitch-matching error should be correlated with perceptual error, rather than a discrete correct or incorrect measurement of perception. This average perceptual judgment error is equivalent to the just-noticeable difference for pitch (pitch-discrimination threshold).

The measurement of just-noticeable differences for pitch was done in two different studies. Amir et al. (2003) showed a significant correlation between pitch perception thresholds and vocal pitch-matching ability; this represents a more appropriate type of perceptual measurement than other percent correct measurements. However, this study had some limitations. The pitch-discrimination thresholds were measured at a much higher pitch level than the vocal pitch-matching trials and used unnatural-sounding sine tones. Furthermore, each participant was asked to sing only nine times, which is probably not enough to specify the mean pitch-matching error, especially among poor singers, who tend to be more variable in their responses (Hutchins and Peretz, in preparation-a; Pfordresher et al., in press). Amir et al.'s (2003) measured pitch perception thresholds were also considerably lower than the production errors, making it unlikely that perception error was the limiting factor in vocal pitch-matching ability. Nikjeh et al. (2009) improved upon this design, using complex tones instead of sine tones and measuring perceptual ability at the same pitch height as vocal pitch matching. However, they found a correlation between these two abilities only among trained instrumentalists, but not trained singers or nonmusicians. In addition, their measured pitch perception thresholds were quite high (see, e.g., Hyde and Peretz, 2004; Zwicker and Fastl, 1999 for comparison); they attributed this discrepancy to their stimulus timbre and procedure.

Another problem with these studies of perception is their use of a correlational design to evaluate the perception–production relationship. Even a significant, positive correlation does not give a solid answer as to whether poor perceptual ability causes poor singing performance. It is just as possible that poor singing ability causes poor performance on these perceptual comparison tests (perhaps listeners are using their voice to evaluate whether the two tones are the same) or that both abilities are regulated by some third factor (e.g., attention). Zarate et al. (2010) used a true experimental design to evaluate whether improving perceptual abilities would improve people's singing abilities. Over six sessions, they trained participants to perceive micromelodies, which use very small, nontraditional intervals. These participants showed significant improvements in pitch-discrimination abilities; however, there was no evidence that their vocal pitch-matching abilities had improved. This demonstrates that pitch-discrimination ability was likely not the limiting factor for their singing ability.

In a new study (Hutchins and Peretz, in preparation-a), we measured pitch perception abilities in a different way. Rather than taking a set of same or different judgments, we let participants use a physical slider to adjust a comparison pitch to match a target pitch. This method measures pitch matching with an active process, rather than a passive decision comparing a set of tones. Using the slider approximates the task of vocal pitch matching much better than other types of decision-based tasks and allows us to compare error in a vocal pitch-matching task to error in an instrumental pitch-matching task using the same types of measurements. Both musicians and nonmusicians were very accurate in pitch matching using the slider, to a much greater degree than they were using their voice. Only 6% of our participants (2 out of 31) showed any difficulty using the slider (see Table 1), and we concluded from this data that participants were capable of matching a pitch on a different apparatus, and pitch perception deficits were not responsible for their poor singing.

Although it is a popular theory, the majority of the evidence seems to indicate that poor pitch perception ability does not seem to be a major cause of poor singing. At best, poor singing ability may tend to co-occur with poor performance on pitch-discrimination tasks, but in general, pitch-discrimination abilities tend to be much better than pitch production abilities. This is not what you would expect if the former were the major cause of poor singing and suggests that the real

Table 1. The potential causes of poor singing and the percentage of the population estimated to be affected by deficits in each of these areas (from Hutchins and Peretz, in preparation-a)

Cause of singing problems	Percent of population
Perceptual	6%
Sensorimotor	35%
Motor control	19%
Feedback	Not assessed
No problems	39%

cause is likely in the accessing or execution of the proper motor plan.

However, we would be remiss if we did not mention the notable case where perceptual disability does seem to be at the heart of singing problems. Congenital amusia, commonly referred to as tone deafness, is estimated to affect around 4% of the population (Kalmus and Fry, 1980). Congenital amusics have very poor musical abilities and lack the ability to recognize the tune of songs or perceive wrong notes in familiar melodies, without any history of brain trauma or other neurological disorders. It is thought to be a neurogenetic disorder (Peretz, 2008) which leads to severe pitch perception difficulties (Hyde and Peretz, 2004). Congenital amusics generally have very poor singing abilities (Ayotte et al., 2002; Dalla Bella et al., 2009) and poor pitch-matching abilities (Hutchins et al., 2010), although a small number retains some ability to sing better than one might expect given their perceptual difficulties. In the case of congenital amusia, it does seem that their impaired pitch perception is at least related to their poor singing ability.

Sensorimotor translation deficits

Another possible cause of poor pitch matching is that an appropriate motor plan is not selected. In this account, poor singers correctly perceive the intended target pitch but select a motor plan that, even if correctly executed, still results in an errorful output. There are no errors in perception or motor control, but there is an error with the motor plan they choose to enact. There are in fact many different ways to produce the same pitch, so the process of selecting a motor plan is not straightforward, making this a potentially difficult task.

Because they could not find any evidence of problems within the perception or motor systems alone, Pfordresher and Brown (2007) proposed that there was a mismapping between auditory representations and motor representations of pitches. They proposed that among inaccurate singers, there was a regular transformation between the two, such that the inaccuracies are constant. This would suggest that when participants fail to match a pitch, they should always sing the same incorrect pitch in response to the same target tone. However, other studies have shown that inaccurate singers tend to be more variable in their responses than good pitch matchers (Pfordresher et al., in press; Hutchins and Peretz, in preparation-a), which argues against this constant transformation view.

Another factor may be at the root of a possible mismapping problem, though. The timbre of the target tone (i.e., the color or type of the sound) can make a big difference to people's pitch-matching abilities. It is clear that the timbre of a tone can affect its perceived pitch (Krumhansl and Iverson, 1992; Melara and Marks, 1990a,b,c; Pitt, 1994; Warrier and Zatorre, 2002), and prior work has shown that people are better at matching the pitch of target tones when they are more similar to their own voice (e.g., Watts and Hall, 2008), and especially so to actual recordings of their own voice (Moore et al., 2008). In most pitch-matching studies, however, participants are asked to match either an instrumental timbre or a complex-synthesized sound. Thus, it may be that pitch-matching errors occur because the singer cannot determine the proper relationship between the target tone and their vocal response. It is as if different timbres constitute different languages, and poor singers cannot translate between them, unable to find which vocal pitch is equivalent to which target pitch.

In one of the experiments of our recent study (Hutchins and Peretz, in preparation-a), we showed that 35% of singers (Table 1) are able to match synthesized vocal tones (which bear a resemblance to the voice but are not identical) with a slider but are not able to match these synthesized vocal tones with their own voice. These same participants, however, are able to match recordings of their own voice quite accurately. This indicates that neither do they have any problems perceiving the pitch of a tone, nor do they have any motor control issues to prevent them from accurate singing. However, they fail to translate the synthesized vocal tone to the appropriate vocal–motor plan. This is evidence that these participants do have such a sensorimotor translation problem. Because it is due to a lack of translation between timbres, rather than a fixed mismapping, these participants do not make constant errors but produce errors with a great deal of variability. Interestingly, even though they often realize they are matching pitch incorrectly, they do not realize *how* to improve their response. This leads to the interesting conundrum that incorrect singers will sing the same incorrect pitch multiple times when given the opportunity to correct it during a single trial but will sing a different incorrect pitch on a subsequent trial at the same target pitch level. In a follow-up experiment in which we required at least 20 pitch-matching attempts per trial from singers, none showed any significant within-trial pitch changes, but poor singers did show between-trial variability.

It is important to note that this timbral translation explanation is distinct from a pure perceptual problem. Within a timbre, listeners can resolve pitch quite well, easily determining higher–lower and same–different judgments. The problem only arises when the listener needs to make comparisons across timbres, in this case, from their representation of the target tone timbre to their representation of their own voice, and the correct motor plan to produce this. Another type of translation problem that commonly occurs concerns pitch transposition. It has been shown that singers with a high voice can have a hard time imitating singers with a low voice, and vice versa (Clegg, 1966; Goetze et al., 1990; Pfordresher and Brown, 2007). These types of singing errors seem to have to do mainly with a problem in transposing up or down an octave (a common task requirement for singing in groups) and constitute a similar difficulty of finding equivalent pitches, across range rather than timbre. Other translation types of errors may also exist, but these have yet to be identified.

Motor control deficits

A third type of explanation for poor singing abilities is a pure motor problem. That is, poor singers may accurately perceive a pitch and select the correct motor plan but may not have the range or coordination to faithfully execute this plan. The possibility of overt vocal problems is sometimes assessed in studies of vocal pitch matching, but these typically address problems such as vocal range and the ability to adequately sustain a tone. However, motor control problems may in fact lie in other areas, specifically the ability to coordinate one's vocal–motor apparatus to achieve a desired frequency. These types of motor control issues are rarely addressed directly and sometimes may be seen only by ruling out other types of problems. For example, in our recent work on pitch matching (Hutchins and Peretz, in preparation-a), we showed that about 20% of participants (Table 1) could match pitch well on the slider but were impaired in vocal pitch matching both in the case where the target tone was a synthesized vocal tone and where it was a recording of their own voice. Because they neither had direct pitch perception problems nor were good at matching the target in the own voice condition, showing little evidence for a sensorimotor translation deficit, we concluded that these participants had a problem with their motor control of their vocal apparatus. This type of lack of coordination is often observed among amateurs in other physical activities as well. For example, a poor bowler may perceive exactly where he wishes to aim a ball, but may simply not be coordinated enough to do so consistently.

It is not surprising that individuals without much practice at using their voice in this way would lack the coordination to sing consistently on pitch, and the finding that inaccurate singers tend to be inconsistent with their productions (Hutchins and Peretz, in preparation-a; Pfordresher and Brown, 2007, Pfordresher et al., in press) is consistent with this account. However, people do tend to have a great deal of practice using their voice in another manner, that is, speech. Being another vocal method for communication, speech necessarily involves some of the same types of properties of the acoustic signal as singing. Both speech and singing use changes in timbre, amplitude, timing, and pitch to create a meaningful signal. However, unlike singing, the vast majority of people are very skilled at using their voice in precise ways to express what they wish to say. Although pitch is not as important in speech as it is in singing (with the exception of tone languages such as Mandarin, which can use pitch to distinguish between otherwise identical words), pitch is still used to convey meaning in speech in several ways, including marking the focus of a sentence, or indicating whether a sentence is a statement or a question. Most people do in fact manipulate the pitch of their voice in such a way as to create exactly the nuances in meaning that they wish to convey, and it should come as no surprise that vocal communication is for the large part successful. This raises the issue, then, of how people who are good at using pitch cues with their voice in speech could lack the coordination and/or motor control to sufficiently control their vocal pitch in singing. The answer to this apparent puzzle lies in the realization that singing and speech may require two different types of vocal control. Just as a scratch bowler might be a terrible softball pitcher despite the similarities between the two types of movements, the practice at controlling vocal pitch in speech might not carry over to the music domain. Therefore, it is reasonable that many people may be very skilled at using their voice for speech, but not coordinated in using their voice for singing. We have shown that congenital amusics have retained abilities to imitate language just as well as controls, despite reduced pitch sensitivity and reduced singing abilities (Hutchins and Peretz, in preparation-b). In addition, we have found a case of acquired amusia, IR, who after brain damage has lost the ability to produce any sung tone and can only speak the words of songs. IR, however, was unaware that she was not singing during the singing tasks, which implies that she may have a singing-specific motor deficit.

Feedback deficits

The final step of singing production that may be impaired is evaluating feedback and making adjustments based on this feedback. That is, poor singers may perceive a pitch correctly, select an appropriate motor plan, and produce it with reasonable accuracy, but misconstrue their own accuracy and adjust their pitch afterwards in an errorful way. Or, it may be that good and poor singers both tend to make the same initial inaccuracies, but good singers are able to use their feedback to correct their responses, whereas poor singers are unable to do so. We know that feedback from the vocal signal is monitored by singers and speakers from the way that they react to alterations to this vocal feedback. For example, delayed auditory feedback causes severe disruptions to speech (MacKay, 1987), and pitch shifting the auditory feedback of speakers and singers evokes automatic compensatory responses in the opposite direction (Burnett and Larson, 2002; Burnett et al., 1998; Natke et al., 2003), even when these differences are not consciously perceived (Hafke, 2008). Masking the auditory feedback of singers also reduces their pitch-matching accuracy (Anstis and Cavanagh, 1979; Elliott and Niemoeller, 1970; Mürbe et al., 2002; Ternström et al., 1988), which shows that feedback evaluation and adjustments are indeed important to accurate singing. In addition, there is some evidence that enhancing feedback through accompaniment (to provide a concurrent comparison) can aid singers. Wise and Sloboda (2008) found that singers were better at matching pitch when singing along with another singer. However, several studies of elementary school children seem to show that they perform less accurately in a choral context than alone (Clayton, 1986; Goetze, 1986; Joyner, 1971; cited in Goetze et al., 1990), possibly due to the masking effects of the other singers, which may in fact be reducing, rather than enhancing feedback.

However, the majority of evidence we have indicates that deficits in feedback evaluation and adjustment are not a primary cause of poor singing. Pfordresher and Brown (2007) manipulated the feedback they provided to their participants in their pitch-matching tasks and found that masking the feedback made no difference to either good or poor singers, and accompaniment actually led to a small decrement in performance among poor singers (who should have been most aided by this, if their problems were due to feedback). In our study on pitch matching (Hutchins and Peretz, in preparation-a), we showed that neither musicians nor nonmusicians made any large corrections to their pitch across multiple attempts to match the same pitch, but that good and poor singers alike tend to persist on a pitch, whether it is accurate or not, which adds more evidence against the feedback account of poor singing. Although feedback is an important process in the singing system, it is most likely not a major cause of inaccurate singing. In the final section of this chapter, we will briefly examine the neural structures which underlie the steps necessary to singing that we have outlined above.

Neural bases of singing

Several studies have examined brain activity during singing tasks. Perry and collaborators (1999) used positron emission tomography during a simple singing task to identify regions involved in perceiving sounds and in producing sung tones. This study showed that a network involving the right Heschl's gyrus and a posterior region in the right superior temporal plane, which contain primary and higher-order auditory areas, were engaged during singing. In addition, several motor control areas were recruited for the singing task, including the supplementary motor area, the anterior cingulate cortex, the precentral gyrus, the anterior insula, and the cerebellum. These regions are very similar to those recruited during speech production (Özdemir et al., 2006; Paus et al., 1993, 1996). Activity in these same regions has been confirmed in several subsequent studies using more complex singing tasks and involving participants with varying degrees of musical training (Brown et al., 2004; Kleber et al., 2007; Zarate and Zatorre, 2008).

Although the basic elements of the singing network have been identified, it is still difficult to map all of the steps required for singing identified in Fig. 2 to particular brain regions. While perception and motor control areas in the singing network are easily identifiable due to their recruitment in other tasks involving those functions, it is harder to identify areas involved in motor plan selection and feedback evaluation. Zarate and Zatorre (2008) showed that the intraparietal sulcus and the dorsal premotor cortex were both activated in response to changes in auditory feedback during singing and so may be involved in both feedback evaluation and motor plan selection. Foster and Zatorre (2010) also showed intraparietal sulcus activity in response to a pitch task in which transposed melodies are compared, which makes it likely that this region may be involved in pitch translation tasks. The dorsal premotor cortex, however, has been implicated in other types of auditory–motor interactions (Chen et al., 2006; Zatorre et al., 2007) and so may be responsible for adjusting motor output in response to auditory events.

Finally, there might a qualitative difference in the functioning of the ventral and dorsal pathways connecting auditory and motor regions, which underlies the traditional distinction between perception and production areas. In a behavioral study, Hafke (2008) showed that trained singers will make unconscious vocal adjustments in response to very small pitch shifts, even when those shifts are not consciously perceived. A subset of congenital amusics show relatively preserved singing (Dalla Bella et al., 2009) and pitch-matching ability (Hutchins et al., 2010). Some amusics can even reproduce the direction of a pitch interval with their voice but yet be unable to name the direction of the same interval (Loui et al., 2008), and in rare cases, can even sing whole songs without being able to distinguish tone pairs which are easy for normal listeners (Hutchins and Peretz, 2010). This action–perception dissociation bears a resemblance to a type of blindsight in which patients who experience no conscious visual input nonetheless perform much better than chance at goal-directed tasks, such as pointing at an object (Danckert and Rossetti, 2005). Such dissociations lend support to dual-route models (e.g., Hickok and Poeppel, 2007; Milner and Goodale, 1995; Rauschecker and Scott, 2009; Warren et al., 2005), in which the ventral and dorsal pathways support (conscious) perception and (unconscious) action, respectively. Supporting evidence for this was found for music, as well, by Loui et al. (2009), who suggested that the superior and inferior pathways in the arcuate fasciculus, connecting auditory and frontal premotor regions, are separately responsible for pitch perception and production. This may explain some of the dissociations seen in both amusics and normal participants.

Conclusion

We began this review by asking "Why don't people sing more often?" One of the most common answers to this question is that they feel they are not a strong enough singer. Many people would like to be able to sing better, and people often feel very self-conscious about the quality of their voices. However, it also seems to be the case that few poor singers become noticeably better later in life, and one of the reasons is that many voice teachers do not truly understand the roots of poor singing, and especially of poor pitch accuracy—one of the primary determinants of poor singing (Watts et al., 2003). Although most of the research focus lately has been on the role of impoverished pitch perception on poor singing, we believe that the recent evidence points toward poor sensorimotor translation and motor control as being the primary determinants of poor singing.

Music teachers are trained to mainly help young children and can sometimes help them to improve their pitch-matching ability, but teachers have little research to help guide their pedagogical strategy. If we can more fully understand what processes are responsible for poor singing behavior, we can adapt our music pedagogy to be able to better help poor singers of all age levels improve. Our research implies that poor singers need more help in recognizing pitch equivalents between instruments and voices, and in developing better vocal–motor coordination, rather than practice in simple pitch perception. This research may also lead to the creation of other tools to aid training; adapted versions of our slider could prove to be effective teaching tools. Imaging work may also give us further insight as to how these processes of perception, sensorimotor translation, motor control, and use of feedback are carried out in the brain and may lead to further insights on the causes of poor singing and may even help identify people with as yet unexpressed musical talent. In sum, our understanding of poor singing is an important step in helping people to not only enjoy music they get from society but also contribute to the music of society and gain the pleasure of performing.

Acknowledgments

This work was supported by grants from the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada, and a Canada Research Chair in neurocognition of music to I. P. and a fellowship in Auditory Cognitive Neuroscience from the CREATE program of the Natural Sciences and Engineering Research Council of Canada to S. H.

References

Amir, O., Amir, N., & Kishon-Rabin, L. (2003). The effect of superior auditory skills on vocal accuracy. *The Journal of the Acoustical Society of America*, *113*, 1102–1108.

Anstis, S. M., & Cavanagh, P. (1979). Adaptation to frequency-shifted auditory feedback. *Perception & Psychophysics*, *26*, 449–458.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, *125*, 238–251.

Bentley, A. (1969). *Monotones. Music Education Research Papers (1)* London: Novello and Company.

Bradshaw, E., & McHenry, M. A. (2005). Pitch discrimination and pitch matching abilities of adults who sing inaccurately. *Journal of Voice*, *19*, 431–439.

Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., & Parsons, L. M. (2004). The song system of the human brain. *Cognitive Brain Research*, *20*, 363–375.

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice *F*0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, *103*, 3153–3161.

Burnett, T. A., & Larson, C. R. (2002). Early pitch shift response is active in both steady and dynamic voice pitch control. *The Journal of the Acoustical Society of America*, *112*, 1058–1063.

Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *Neuroimage*, *32*, 1771–1781.

Clayton, L. S. (1986). *An investigation of the effect of a simultaneous pitch stimulus on vocal pitch accuracy.* Unpublished master's thesis, Indiana University, Bloomington. Cited in Goetze, M., Cooper, N., & Brown, C.J. (1990). Singing in the general music classroom. *Bulletin of the Council for Research in Music Education*, *104*, 16–37..

Cleall, C. (1979). *Vocal range in young children.* Paper presented at the course in Development of Young Children's Music Skills, University of Reading. Cited in Goetze, M., Cooper, N., & Brown, C.J. (1990). Singing in the general music classroom. *Bulletin of the Council for Research in Music Education*, *104*, 16–37.

Clegg, B. W. (1966). *A comparative study of primary grade children's ability to match tones.* Unpublished master's thesis, Brigham Young University, Provo, UT. Cited in Goetze, M., Cooper, N., & Brown, C.J. (1990). Singing in the general music classroom. *Bulletin of the Council for Research in Music Education*, *104*, 16–37.

Dalla Bella, S., Giguère, J. F., & Peretz, I. (2007). Singing proficiency in the general population. *The Journal of the Acoustical Society of America*, *121*, 1182–1189.

Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2009). Singing in congenital amusia: An acoustical approach. *The Journal of the Acoustical Society of America*, *126*, 414–424.

Danckert, J., & Rossetti, Y. (2005). Blindsight in action: What can the different sub-types of blindsight tell us about the control of visually guided actions? *Neuroscience and Biobehavioral Reviews*, *29*, 1035–1046.

Elliott, L. L., & Niemoeller, A. F. (1970). The role of hearing in controlling voice fundamental frequency. *International Journal of Audiology*, *9*, 47–52.

Estis, J. M., Coblentz, J. K., & Moore, R. E. (2009). Effects of increasing time delays on pitch-matching accuracy in trained singers and untrained individuals. *Journal of Voice*, *23*, 439–445.

Estis, J. M., Dean-Claytor, A., Moore, R. E., & Rowell, T. L. (2010). Pitch-matching accuracy in trained singers and untrained individuals: The impact of musical interference and noise. *Journal of Voice*, *25*, 173–180.

Foster, N. E. V., & Zatorre, R. J. (2010). A role for the intraparietal sulcus in transforming musical pitch information. *Cerebral Cortex*, *20*, 1350–1359.

Goetze, M. (1986). *Factors affecting accuracy in children's singing.* Doctoral dissertation, University of Colorado. Cited in Goetze, M., Cooper, N., & Brown, C. J. (1990). Singing in the general music classroom. *Bulletin of the Council for Research in Music Education*, *104*, 16–37.

Goetze, M., Cooper, N., & Brown, C. J. (1990). Recent research on singing in the general music classroom. *Bulletin of the Council for Research in Music Education*, *104*, 16–37.

Hafke, H. Z. (2008). Nonconscious control of fundamental voice frequency. *The Journal of the Acoustical Society of America*, *123*, 273–278.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, *8*, 393–402.

Hutchins, S., & Peretz, I. (2010). Double dissociation of pitch production and perception. In: *Poster presented at 17th annual meeting of the Cognitive Neuroscience Society, Montreal, April, 2010*.

Hutchins, S., & Peretz, I. (in preparation-a). Why do some people fail to sing in tune? A comparison of pitch perception and production abilities of musicians and nonmusicians.

Hutchins, S., & Peretz, I. (in preparation-b). Perceiving and producing pitch in language.

Hutchins, S., Zarate, J. M., Zatorre, R. J., & Peretz, I. (2010). An acoustical study of vocal pitch matching in congenital amusia. *The Journal of the Acoustical Society of America*, *127*, 504–512.

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science*, *15*, 356–360.

Joyner, D. R. (1969). The monotone problem. *Journal of Research in Music Education*, *31*, 115–124.

Joyner, D. R. (1971). *Pitch discrimination and tonal memory and their association with singing and the larynx.* Master's thesis, University of Reading. Cited in Goetze, M., Cooper, N., & Brown, C. J. (1990). Singing in the general music classroom. *Bulletin of the Council for Research in Music Education*, *104*, 16–37.

Kalmus, H., & Fry, D. B. (1980). On tune deafness (dysmelodia): Frequency, development, genetics and musical background. *Annals of Human Genetics*, *43*, 369–382.

Klatt, D. H. (1973). Discrimination of fundamental frequency contours in synthetic speech: Implications for models of pitch perception. *The Journal of the Acoustical Society of America*, *53*, 8–16.

Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., & Lotze, M. (2007). Overt and imagined singing of an Italian aria. *Neuroimage*, *36*, 889–900.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology. Human Perception and Performance*, *18*, 739–751.

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: A new disconnection syndrome? *The Journal of Neuroscience*, *29*, 10215–10220.

Loui, P., Guenther, F., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology*, *18*(8), R331–R332.

MacKay, D. G. (1987). *The organization of perception and action.* New York: Springer.

Melara, R. D., & Marks, L. E. (1990a). HARD and SOFT interacting dimensions: Differential effects of dual context on classification. *Perception & Psychophysics*, *47*, 307–325.

Melara, R. D., & Marks, L. E. (1990b). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics*, *48*, 169–178.

Melara, R. D., & Marks, L. E. (1990c). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology. Human Perception and Performance*, *16*, 398–414.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* Oxford: Oxford University Press.

Moore, R. E., Estis, J., Gordon-Hickey, S., & Watts, C. (2008). Pitch discrimination and pitch matching abilities with vocal and nonvocal stimuli. *Journal of Voice*, *22*, 399–407.

Moore, R. E., Keaton, C., & Watts, C. (2007). The role of pitch memory in pitch discrimination and pitch matching. *Journal of Voice*, *21*, 560–567.

Mürbe, D., Pabst, F., Hofmann, G., & Sundberg, J. (2002). Significance of auditory and kinesthetic feedback to singers' pitch control. *Journal of Voice*, *16*, 44–51.

Natke, U., Donath, T. M., & Kalveram, K. T. (2003). Control of voice fundamental frequency in speaking versus singing. *The Journal of the Acoustical Society of America*, *113*, 1587–1593.

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2009). The relationship between pitch discrimination and vocal production: Comparison of vocal and instrumental musicians. *The Journal of the Acoustical Society of America*, *125*, 328–338.

Özdemir, E., Norton, A., & Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *Neuroimage*, *33*, 628–635.

Paus, T., Petrides, M., Evans, A. C., & Meyer, E. (1993). Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. *Journal of Neurophysiology*, *70*, 453–469.

Paus, T., Perry, D. W., Zatorre, R. J., Worsley, K., & Evans, A. C. (1996). Modulation of cerebral blood-flow in the human auditory cortex during speech: role of motor-to-sensory discharges. *European Journal of Neuroscience*, *8*, 2236–2246.

Peretz, I. (2008). Musical disorders: From behavior to genes. *Current Directions in Psychological Science*, *17*, 329–333.

Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1999). Localization of cerebral activity during simple singing. *Neuroreport*, *10*, 3979–3984.

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of "tone deafness" *Music Perception*, *25*, 95–115.

Pfordresher, P. Q., Brown, S., Meier, K. M., Belyk, M., & Liotti, M. (2010). Imprecise singing is widespread. *Journal of the Acoustical Society of America, 128,* 2182–2190.

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology. Human Perception and Performance*, *20*, 976–986.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the human auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, *12*, 718–724.

Ternström, S., Sundberg, J., & Colldén, A. (1988). Articulatory F0 perturbations and auditory feedback. *Journal of Speech, Language, and Hearing Research*, *31*, 187–192.

Warren, J. E., Wise, R. J. S., & Warren, J. D. (2005). Sounds do-able: Auditory-motor transformations and the posterior temporal plane. *Trends in Neurosciences*, *28*, 636–643.

Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics*, *64*, 198–207.

Watts, C., Barnes-Burroughs, K., Adrianopoulos, M., & Carr, M. (2003). Potential factors related to untrained singing talent: A survey of singing pedagogues. *Journal of Voice*, *17*, 298–307.

Watts, C. R., & Hall, M. D. (2008). Timbral influences on vocal pitch-matching accuracy. *Logopedics, Phoniatrics, Vocology*, *33*, 74–82.

Watts, C., Moore, R., & McCaghren, K. (2005). The relationship between vocal pitch matching skills and pitch discrimination skills in untrained accurate and inaccurate singers. *Journal of Voice*, *19*, 534–543.

Wise, K. J., & Sloboda, J. A. (2008). Establishing an empirical profile of self-defined “tone deafness”: Perception, singing performance and self-assessment. *Music Scientiae*, *12*, 3–26.

Zarate, J. M., Delhommeau, K., Wood, S., & Zatorre, R. J. (2010). Vocal accuracy and neural plasticity following micromelody-discrimination training. *PLoS ONE*, *5*, e11181.

Zarate, J. M., & Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *Neuroimage*, *40*, 1871–1887.

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. Neurosience*, *8*, 547–558.

Zwicker, E., & Fastl, H. (1999). *Psychoacoustics: Facts and models.* Berlin: Springer-Verlag.