ELSEVIER

Neuropsychologia 44 (2006) 479–488

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Vocal amusia in a professional tango singer due to a right superior temporal cortex infarction

Yasuo Terao[a,*], Tomoyuki Mizuno[b], Mitsuko Shindoh[c], Yasuhisa Sakurai[d], Yoshikazu Ugawa[a], Shunsuke Kobayashi[a], Chiyoko Nagai[e], Toshiaki Furubayashi[a], Noritoshi Arai[a], Shingo Okabe[a], Hitoshi Mochizuki[f], Ritsuko Hanajima[a], Shouji Tsuji[a]

[a] *Department of Neurology, Division of Neuroscience, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan*
[b] *Department of Cognitive Neuroscience, Graduate School of Medicine, University of Tokyo, Tokyo, Japan*
[c] *Sophia University, Research Center for Communication Disorders, Tokyo, Japan*
[d] *Department of Neurology, Mitsui Memorial Hospital, Tokyo, Japan*
[e] *Department of Neurology, Tokyo Women's Medical University, Tokyo, Japan*
[f] *Third Department of Internal Medicine, National Defense Medical College, Tokyo, Japan*

Received 4 May 2004; received in revised form 14 May 2005; accepted 17 May 2005

Abstract

We describe the psychophysical features of vocal amusia in a professional tango singer caused by an infarction mainly involving the superior temporal cortex of the right hemisphere. The lesion also extended to the supramarginal gyrus, the posterior aspect of the postcentral gyrus and the posterior insula. She presented with impairment of musical perception that was especially pronounced in discriminating timbre and loudness but also in discriminating pitch, and a severely impaired ability to reproduce the pitch just presented. In contrast, language and motor disturbances were almost entirely absent. By comparing her pre- and post-stroke singing, we were able to show that her singing after the stroke lacked the fine control of the subtle stress and pitch changes that characterized her pre-stroke singing. Such impairment could not be explained by the impairment of pitch perception. The findings suggest that damage to the right temporoparietal cortex is enough to produce both perceptive and expressive deficits in music.

Keywords: Music; Singing; Auditory feedback; Pitch perception; Stroke

---

1. Introduction

Most studies of musical processing have addressed the receptive aspects of music. Clinical and recent neuroimaging studies support the dominant role of the right hemisphere, especially the superior temporal cortex, in musical processing, such as pitch and timbre perception (Samson & Zatorre, 1994; Sidtis & Volpe, 1988; Zatorre & Samson, 1991), musical imagery (Halpern & Zatorre, 1999; Halpern, Zatorre, Bouffard, & Johnson, 2004; Zatorre & Halpern, 1993) or recognition and memory of pitch, timbre and melody (Sidtis & Volpe, 1988; Zatorre, Evans, & Meyer, 1994); the left hemisphere is dominant for processing time and rhythm (Zatorre & Belin, 2001).

In contrast, there have been very few clinical reports on expressive aspects of music including vocal amusia. Only three cases of expressive amusia have been reported in professional musicians in whom the extent of the lesion is clearly defined. Again, all cases presented with right hemispheric lesions that involved the superior temporal cortex in two cases

---

[*] Corresponding author. Tel.: +81 3 3815 5411x33784; fax: +81 3 5800 6548.
E-mail address: yterao-tky@umin.ac.jp (Y. Terao).

0028-3932/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2005.05.013

(Takeda, Bandou, & Nishimura, 1990, extension also into the supramarginal gyrus in one of the cases, McFarland & Fortin, 1982), and the right frontal lobe in the other (Botez & Wertheim, 1959). Including cases involving amateur singers, most reported cases of vocal amusia are related to lesions in the right temporal and the contiguous parietal cortex (Chiba, Obi, Takeda, Sakuta, & Bandou, 1989; Confavreux, Croisile, Garassus, Aimard, & Trillet, 1992; Hofman, Klein, & Arlazoroff, 1993; Ishikawa & Inoue, 1989; Sparr, 2002), or in the right frontal cortex (Mann, 1898; Wertheim, 1969). In none of these cases was a direct comparison of musical performance reported before and after the cortical damage.

Recent neuroimaging studies addressing the neural substrate of singing (Jeffries, Fritz, & Braun, 2003; Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000) have revealed commonalities of cortical activation including the lower motor cortices, anterior cingulate cortex, posterior insula, anterior or posterior parts of the temporal lobe including the Heschl's gyri and the left cerebellum to differing degrees. These studies all concur that the right hemisphere dominates activation when singing is compared with speech production. Consistent with the role of the right superior temporal lobe in music perception, the auditory association cortex was implicated in the use of auditory feedback to modulate vocalization, whereas activation of the lower motor cortex and posterior insula was considered to be involved more with motor performance (Jeffries et al., 2003; Perry et al., 1999; Riecker et al., 2000).

Here, we describe a case of a professional tango singer who displayed both receptive and expressive aspects of amusia after sustaining a lesion mainly in the superior temporal cortex of the right hemisphere. Great variation in musical ability is evident in non-clinical populations. Therefore, documentation of patient's premorbid musical ability and activity as well as a comparison of premorbid to postmorbid performance is crucial for any reliable report of amusia. Through comparison of the patient's pre- and post-stroke musical performance, we discuss the role of this area in music perception and its contribution to the expressive aspects of music.

2. Case report

A 62-year-old right-handed woman, an accomplished professional tango-singer with a three-octave vocal range, consulted our department because her singing was "wrong" and she could no longer perform as a singer. On the morning of July 12, 2002, she woke up and noticed weakness of the left upper and lower limbs, and slight dysarthria. These symptoms cleared after 15 min, and speech was fluent with no signs of aphasia or apraxia on arrival at the hospital 30 min later. Neurologically, her mental status was normal. Neither dysarthria nor any impairment of other cranial nerve functions, including tongue movements, was noted. No weakness or sensory problems of the upper and lower limbs were observed.

She was admitted to a hospital with a diagnosis of cerebral infarction. The doctors administered intravenous ozagrel sodium and edaravone for 2 weeks and stopped a regular dose of conjugated estrogen she had been taking for her postmenopausal syndrome. During her stay in the hospital, she could dance as skillfully as she had during performing days. She could produce trills of 'r' without difficulty, and had no problem in reciting the lyrics in the Spanish songs she used to sing. She could repeat /ta/ at a rate of up to 6–7 Hz, attesting to the normal motor control of tongue. Slight tinnitus was noted during the hospital stay, but disappeared soon after her discharge from the hospital.

Because she had refrained from singing in the hospital, it was only after her discharge from the hospital on August 3 that she noted that she could no longer carry a tune. Although she never failed to recognize familiar music or songs and perceived no difficulty in discriminating contours of melodies, she was unable to sing along in tune with a compact disc (CD) that recorded her former singing. Having realized her lost acuity for small pitch changes, she never dared to sing alone in public. She did not perceive any difficulty in daily conversation or in hearing speech intonation, but she was often unable to hear the telephone ringing or people speaking on television, and was compelled to increase the volume of these devices. She had to hold the handset of the telephone to her right ear because she often missed what people said when holding it to her left ear. Six months after onset, she still remained unable to carry a tune, especially singing in head voice. She became able to clearly hear other people speaking with a telephone handset at either ear, but she was deprived of much of her musical enjoyment because the piano and other instruments lost their distinct features of timbre: they sounded dull; as if they were being heard from a great distance.

Magnetic resonance imaging of the brain 1 month after her stroke revealed an infarction that mainly involved the right superior temporal gyrus, partly affecting the Heschl's gyrus, and extending to the lower parietal cortex including the supramarginal gyrus and the posterior aspect of the postcentral gyrus. A small extension into the posterior insula was also noted (Fig. 1A–C).

3. Methods

Experiments were performed according to the guidelines of the Ethics Committee of the University of Tokyo. We recruited 10 normal subjects without formal musical training (8 males and 2 females; age (mean ± S.D.) 37.0 ± 9.6 years, range 24–52) to obtain control data for music performance (see below). The following examinations were carried out after written informed consent was obtained from the patient and the subjects according to the Declaration of Helsinki.

Standard neuropsychological assessments included Japanese versions of the Wechsler Adult Intelligence Scale-Revised (WAIS-R), Wechsler Memory Scale-Revised

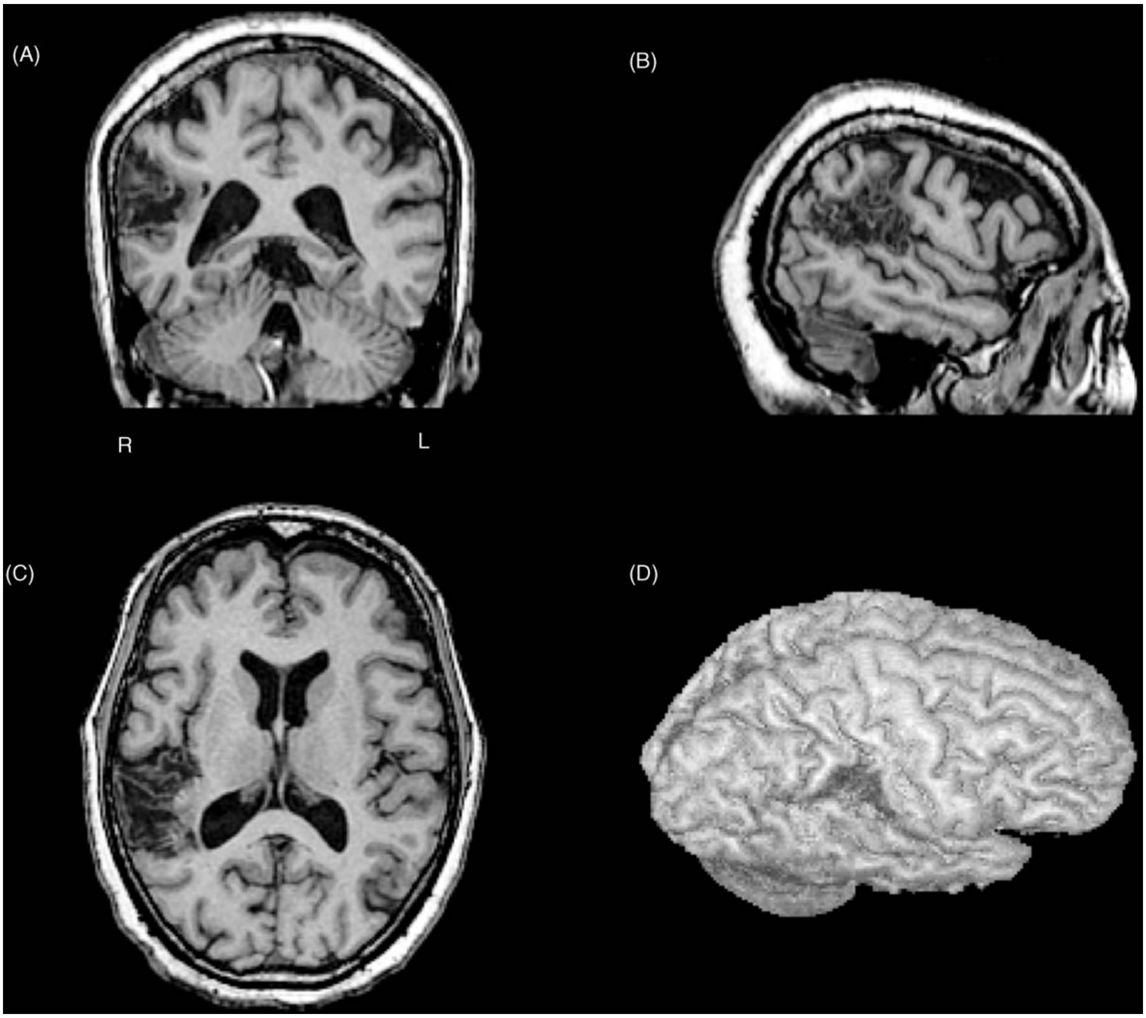


Fig. 1. MRI of the brain. (A) Coronal section, (B) sagittal section, (C) transverse section through the lesion and (D) surface rendering of the MRI showing the lesion as a defect.

(WMS-R) and Western Aphasia Battery (WAB). Seashore measures of musical talents (Seashore, Lewis, & Saetveit, 1960) were also implemented to assess the capacity to perceive pitch, loudness, rhythm, time, timbre and tonal memory. The patient should presumably have displayed above-average capability before her stroke.

As a psychological test for pitch discrimination, we implemented a standard psychophysical forced-choice procedure adapted from previous studies (Johnsrude, Penhune, & Zatorre, 2000; Peretz et al., 2002). Each block of trials comprised 200 pairs of digitalized 700-ms square wave tones, separated by a 500 ms silence. Stimuli were created by digital synthesis sampling at 44.1 kHz using a computer freeware package (Wavegene Ver. 1.20, efu, Japan). They were presented to subjects using Superlab Pro Ver. 2 (Cedrus Corp., San Pedro, CA, USA) through loudspeakers (Multimedia Speaker System MS-71WH; Elecom Co. Ltd., Osaka, Japan) at a level of 70 dB sound pressure level (SPL). One tone was a reference tone having either a fundamental frequency of C4 (262 Hz), C5 (524 Hz) or C6 (1048 Hz). It was used in each block of trials. The other tone was identical in pitch in half of the trials (100 trials), but took a different pitch value in the remaining trials (10 trials for each pair of mismatch tones). The order of the presentation of the standard or the other tone was randomized among the trials. When the reference tone was C5, the different pitch values corresponded to B5 (988 Hz) or C4 sharp (277 Hz) for the 11-semitone distance, E5 (660 Hz) or A4 flat (416 Hz) for the 4-semitone distance, D5 sharp (623 Hz) or A4 (440 Hz) for the 3-semitone distance, D5 (588 Hz) or B4 flat (466 Hz) for the 2-semitone distance and C5 sharp (555 Hz) or B4 (494 Hz) for the 1-semitone distance. The same inter-pitch distances were used for the other two reference tones. Trials of different inter-pitch distances were presented intermixed in a random order. Holding a response pad furnished with three response buttons (RB-620; Cedrus Corp.), the subjects were instructed to press the rightmost key whenever they detected a rising pitch change in the tone pairs, the leftmost key when detecting a falling pitch change or the middle key when detecting no pitch change. The error rate was plotted as a function of the pitch distance in the rising or falling pitch. Reaction time (RT) was measured from the time of cessation of the second tone to the button press.

In the test for pitch reproduction, square-wave tones of 3 s duration and various pitches (sample tones) were presented at a level of 70 dB SPL. The subjects were required, immediately following their cessation, to produce a voice of the same pitch and approximately the same duration. The target notes had fundamental frequencies ranging from A2 to E5 (in Hz) [110 (A2), 123 (B2), 131 (C3), 147 (D3), 165 (E3), 175 (F3), 196 (G3), 220 (A3), 247 (B3), 262 (C4), 294 (D4), 330 (E4), 350 (F4), 392 (G4), 440 (A4), 494 (B4), 524 (C5), 588 (D5) and 660 (E5)]. Each normal subject was tested within their optimal range or the pitch range that each subject was capable of producing, i.e. between the minimum and maximum of fundamental frequencies that each subject used for reproducing the presented pitches. Voices were recorded by a sound recorder (Microsoft [R] sound recorder, Version 5.1; Microsoft Corp., Phoenix, AZ, USA) into a wave file format through a headset connected to a PC, once for each of the target tones in the optimal range. Three registers are usually recognized in the voice of trained singers: chest, normal (middle) and head voices, which each refer to the tonal quality of the lower, middle, higher register and the sensations felt while singing. Thus, the patient was asked to sing in these three voices, and each of these three singings was recorded for each target note. Spectral analysis of the recordings was performed by a commercially available software package (SP4WIN Custom Ver. 2.0, NTT-AT, Tokyo, Japan). Although an automatic frequency extraction algorithm provided by this software was used, we checked the recorded samples, and listening to the recorded files one by one, to make sure that the frequency from the algorithm represented the actual fundamental frequency of the recorded voice, and not half or double of it. Subsequently, the pitch value was averaged across the entire duration of each of the 3-s samples, which was then plotted as a function of the presented pitch to see how accurately the subjects were able to match the pitch of their voice to the presented tone (Figs. 4 and 5). Reproducing the pitch immediately after the sample tone ceased reduced the working memory load for this task. The performance was compared with those of 10 normal musically unselected subjects.

To analyze the test results of pitch discrimination statistically, repeated measures analysis of variance (ANOVA) was conducted with two factors – semitone distance (0–4-, 11-semitones in the ascending and descending directions) and subject (normal subjects versus the patient) – to reveal any deviations in the performance (error rate and reaction time) from normal range at any of the semitone distances. The significance criterion was set at $P < 0.05$. Post hoc analysis using Bonferroni/Dunn's correction for multiple comparisons was carried out to see what differences contributed to the significant differences detected by ANOVA. For the pitch production test, 10 normal subjects took the same test and the 95% confidence range of their performance was calculated. Performance of the patient outside this confidence range of normal subjects was considered abnormal.

We also recorded and compared the pre- and post-stroke singing performances of the patient. The recording made before the stroke was taken from a song that the patient used to sing during her concerts; it was usually sung to a piano accompaniment and recorded pre-stroke on a CD that had been distributed commercially before her stroke. Six months after the stroke, the patient was invited to sing the same song: (a) without piano accompaniment, either with or without auditory feedback of her voice (in the latter case, auditory feedback was blocked with white noise), (b) listening to the CD recording or (c) listening to the piano accompaniment alone (i.e. without the recording of her voice as in the CD). The three tasks were repeated on three occasions separated by 2–3 months, each in a different order. Since we obtained a consistent result, we combined the data for the three recordings (Fig. 7). The post-stroke singing performances were compared with that of her pre-stroke singing. Excerpts of 25 s duration, representing recordings of the corresponding sections of the song before and after stroke, were subjected to spectral analysis by SP4WIN Custom Ver. 2.0 (Fig. 7). The passage consists of three portions: the first had a gradually descending pitch; the second had an overall ascending pitch in which abrupt rises occurred at three time periods; the last had a fall in pitch (black arrows in Fig. 7A). Finally, to show how much the patient deviated from the intended pitch, we compared the pitch values of the correct and sung versions for each note, i.e. we computed the instantaneous pitch differences between the pre-stroke singing recorded in CD and the post-stroke singings under various conditions, and plotted this as a function of time (Fig. 7D).

4. Results

At the time of examination, the patient no longer experienced tinnitus and she could listen to the television and telephones at a normal volume. Hearing was normal on neurological examinations. Audiogram also revealed normal hearing acuity for pure tones (Fig. 2A; hearing level: right ear 8.8 dB, left ear 11.3 dB [normal range <25 dB; hearing level was defined as [dB (500 Hz) + 2 dB (1000 Hz) + dB (2000 Hz)]/4, where dB (500 Hz), dB (1000 Hz) and dB (2000 Hz) each represent the hearing levels for pure tones of 500, 1000 and 2000 Hz, respectively]) as well as for phonemes and environmental sounds when tested binaurally. Brainstem auditory responses were clearly recorded and the latencies of the *I–V* waves were normal, consistent with normal processing up to the brainstem level (Fig. 2B, left ear stimulation [left lead/right lead]: wave I 1.42 ms, wave II 2.61/2.88 ms, wave III 3.82/3.65 ms, wave IV 5.01/4.95 ms, wave V 5.78/5.93 ms; right ear stimulation [left lead/right lead]: wave I 1.53 ms, wave II 3.02/2.75 ms, wave III 3.69/3.85 ms, wave IV 5.03/5.20 ms, wave V 6.03/5.95 ms).

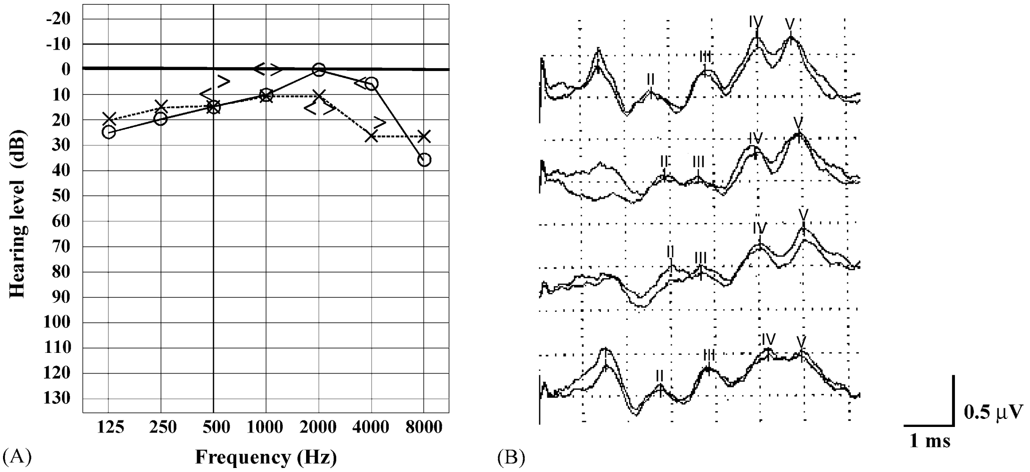


Fig. 2. Audiological examinations of the patient. (A) Audiogram of the patient. The abscissa shows the frequency of presented tone and the ordinate the hearing level in dB. Circles and crosses represent the acuity threshold levels of the left and right ears for air conduction. "<" and ">" stand for those for bone conduction. (B) Brainstem auditory responses of the patient at 90 dB SPL. Each trace represents the average of 2000 responses to a single auditory stimulus. The top two traces are for stimulation of the left ear and the lower two traces for stimulation of the right ear. The first and third traces are recorded from Cz (vertex) linked to the left earlobe, and the second and fourth traces from Cz linked to the right ear lobe.

At 3 months after onset, the patient performed at an average level on a standard intelligence test (WAIS-R full-scale Intelligence Quotient 97, VIQ 103, PIQ 98) and displayed excellent memory for both verbal and non-verbal information (WAB: cortical quotient of 99.1 for both hands) except for a slight impairment in verbal comprehension (9.85/10). When the left ear was tested in isolation at 70 dB SPL, she failed to recognize 19.1% of the phonemes presented such as /ba/, /ga/, /ka/, /to/, /bi/, i.e. stop consonants followed by vowels [a], [i] or [o]. The corresponding figure for the right ear was 0.6%. These results indicated reduced speech discrimination confined to the left ear, which was mainly ascribable to the difficulty in identifying stop consonants. Thus, in the following tests, we always presented the auditory stimuli binaurally, and, before each session, the auditory stimuli were set at a volume optimal for hearing.

Her performance on the Seashore test was far below that expected from a professional singer, with scores corresponding to 34th, 10th, 53rd, 35th, 9th and 24th percentile among the normal population for the pitch, loudness, rhythm, time, timbre and tonal memory subtests, revealing a deficit most pronounced in the loudness and timbre perception, respectively. This performance was unlikely due to the reduced hearing acuity of the patient because before each part of the test, the auditory stimuli were set at an optimal volume to hear, making sure that the subjects heard the sounds clearly. In the pitch discrimination test, she failed to detect pitch changes of one semitone in *circa.* Twenty percent of trials when the pitch was falling for any of the three pitch ranges, and gave "false-positive" responses for pitch changes in ca. 20% of trials when the two tones were actually of identical pitch (Fig. 3A; in this figure, results for the three pitch ranges were combined, either centered on C4, C5 or C6, because the results were consistent across the three pitch ranges). These scores were significantly worse than those for musically unselected nor-

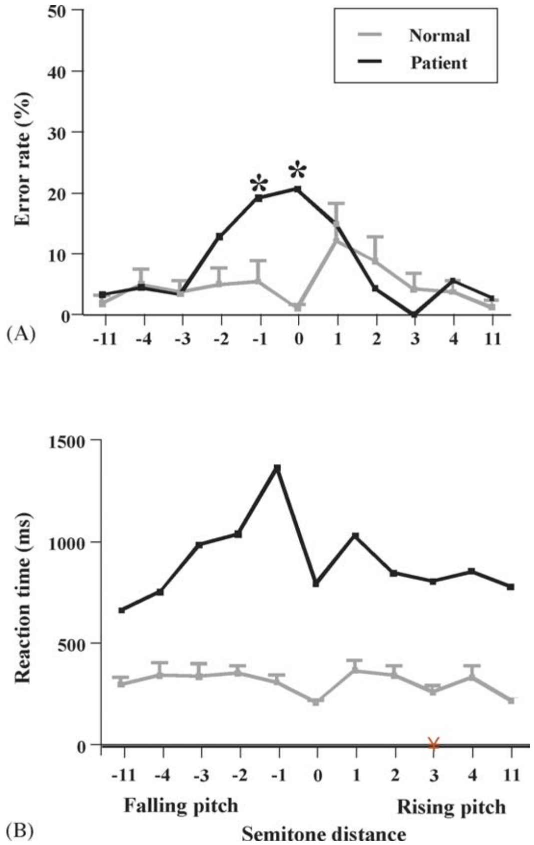


Fig. 3. Performance of pitch perception in the patient. (A) Percentages of error rates (ordinate) in detecting a pitch change in the second tone of a pair for the patient (black curve) and normal subjects (gray curve). The abscissa shows respective pitch distances in semitones in the rising (positive values) and falling pitch (negative values). Asterisks denote significant difference between the patient and normal subjects. (B) Reaction time (RT) measured (abscissa) plotted as a function of pitch distance (ordinate). Error bars in both figures indicate standard errors of the mean across normal subjects.

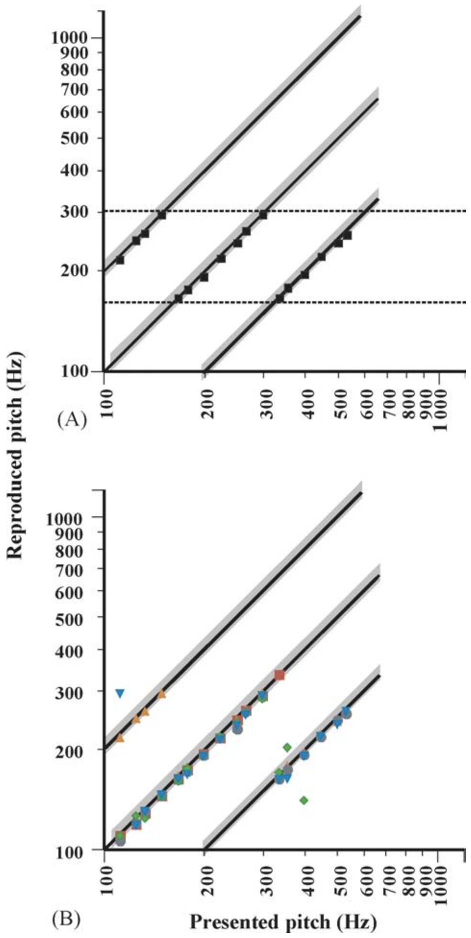


Fig. 4. Performance of pitch reproduction by normal subjects. The logarithm of the fundamental frequency of the reproduced pitch (ordinate) plotted as a function of that of the presented tone (abscissa). (A) Performance of a typical normal subject (female, age 39) without any formal musical training. (B) Performance of normal subjects. Different symbols represent different subjects. Gray shaded areas denote the 95% confidence limit of normal performance. The approximate optimal range of voice is indicated by the horizontal dashed lines in panels A.

mal subjects ($P < 0.02$ corrected for multiple comparisons). For these semitone distances, although fast responses were not encouraged, the patient's reaction time was overall significantly slower than that of the normal subjects (Fig. 3B; $P < 0.0001$ corrected for multiple comparisons). No significant deviation from normal performance was noted for rising pitch changes and pitch changes larger than two semitones (Fig. 3A).

Figs. 4 and 5 show the reproduced pitch plotted against the logarithm of the fundamental frequency of the presented tone for normal subjects and the patient. In this figure, complete coincidence of the presented and reproduced pitches

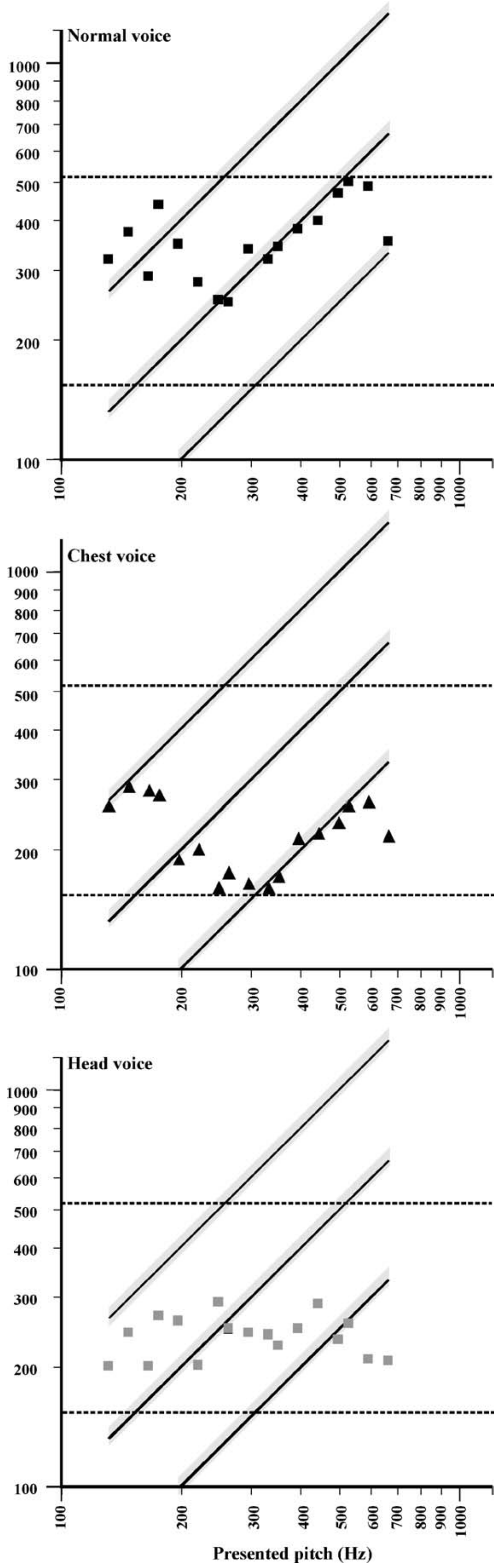


Fig. 5. Patient's performance on the pitch reproduction task. The logarithm of the fundamental frequency of the pitch reproduced by the patient (ordinate) plotted as a function of that of the presented tone (abscissa) for the patient. The top, middle and bottom panels show data for normal, chest and head voices. Conventions as in Fig. 4A and B.

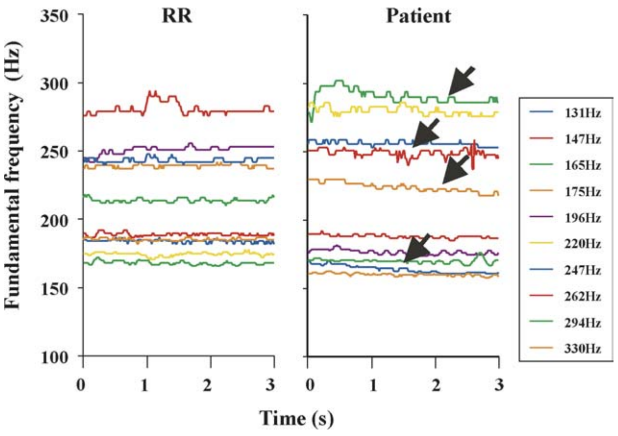


Fig. 6. Fundamental frequency (F_0) contours of reproduced pitches in a normal subject and the patient. The fundamental frequency (F_0) contours averaged for the entire duration of the 3-s samples of reproduced pitches (ordinate) plotted as a function of time (abscissa) in one male normal subject (RR, 40-year-old, left figure) and the patient (right figure). References tones for each of the reproduced pitches are given in the panel on the right side.

would result in a line of unison as indicated by the slanted black line running through the origin. The black slanted lines below and above this line would result if the subject produced a pitch transposed one octave above or below the presented pitch. Normal subjects were generally able to tune their own voice, usually within 10 Hz of the presented pitch, although they shifted one octave above or below this when the presented pitch was below or above their optimal range of voice (Fig. 4A and B). In contrast, the pitch produced by the patient deviated greatly from the presented pitches or shifted one octave below or above this, even when the tones seemed to be well within her optimal range (indicated arbitrarily by the horizontal dashed lines in the figures). The deviation was most remarkable for the head voice (Fig. 5, bottom panel).

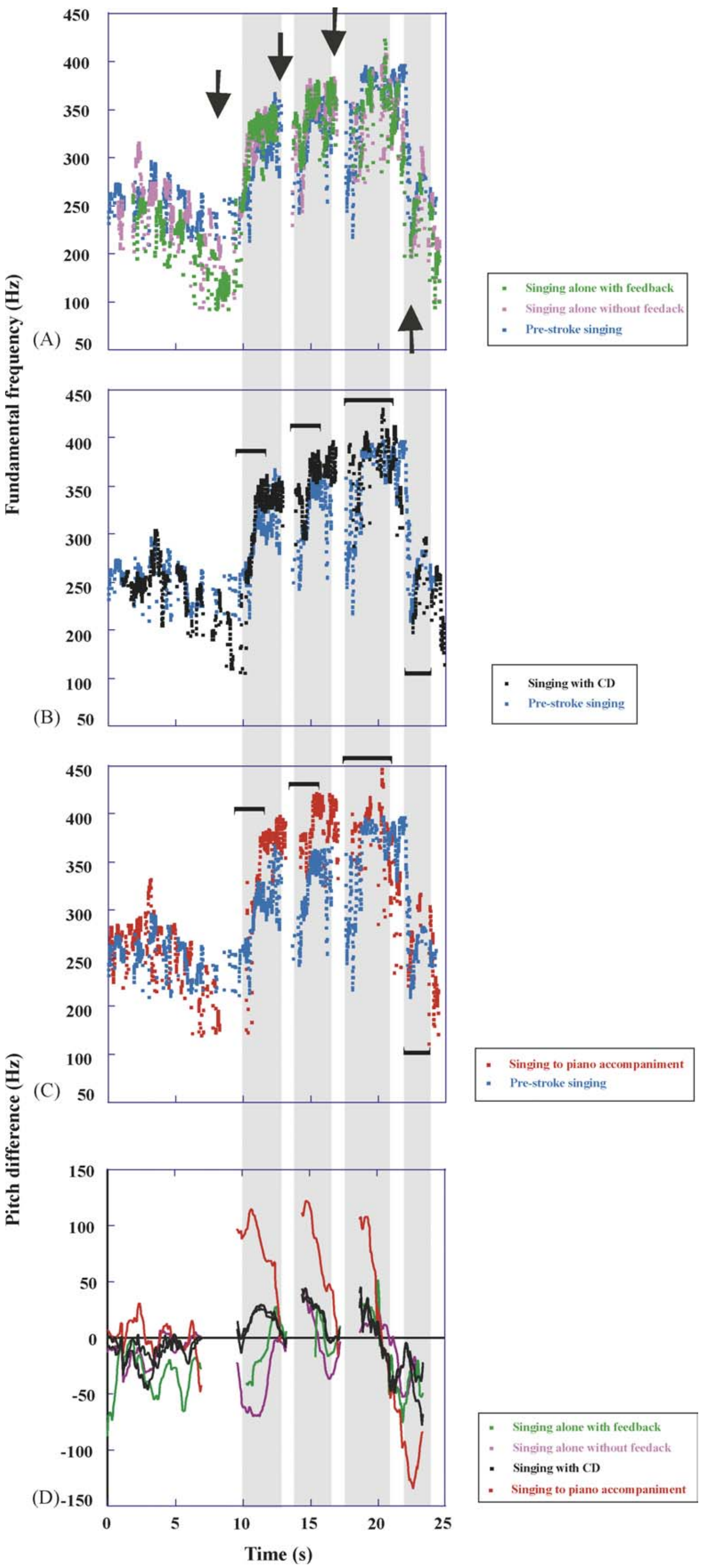
We also computed the variability (standard deviation) of the reproduced pitches for each 3-s sample to show the patient's precision in holding the pitch, even if the pitch itself was not accurate with respect to the target (Fig. 6). Across the 131–330 Hz range, the average of standard deviations for all the produced pitches was 1.77 ± 0.87 Hz (mean ± S.D.) in the patient and 2.21 ± 0.59 Hz on average in normal subjects, i.e. comparable but significantly smaller for the patient ($P < 0.05$). Across the same pitch ranges, however, the overall pitch reproduced by the patient decreased gradually with time (−2.34 ± 1.72 Hz per 1 s of reproduced pitch, Fig. 6, right), whereas the pitch reproduced by the normal subjects, if anything, increased gradually (2.13 ± 1.10 Hz per 1 s of reproduced pitch; significant difference with the patient, $P < 0.05$, Fig. 6, left). Especially when the patient reproduced pitches of 165 and 175 Hz, the rate of descent reached 5.6–5.8 Hz per 1 s, which was sufficiently large to be detected by a naïve listener. Since the patient was able not only to produce a much louder voice than normal subjects but was also able to maintain the volume of voice at a more stable level, this descent was not considered to be due to a problem in expiration.

Given the impairment in pitch perception, we tested the possibility that the patient's singing deteriorated because the deficit prevented her from monitoring the auditory feedback of her own voice. We compared the patient's post-stroke singing with and without auditory feedback to show the patient's reliance on auditory feedback when singing. The overall time courses of pitch changes in the two post-stroke performances resembled one another remarkably (green and purple dots in Fig. 7A) or even compared with her pre-stroke singing (blue dots). Nevertheless, the deterioration in her post-stroke singing was apparent even to a naïve listener. This deterioration resulted from the loss of fine control of pitch and changes in stress characterizing her former singing; in the passage's descending portion, the time course of pitch change was variable among the plots for the two post-stroke singing samples and for the pre-stroke singing sample. This variability implied that, in parts where the pitch had to stay relatively constant or where it changed only gradually over a period of longer than 1 s, the patient sang with an unstable pitch that deviated gradually either upwards or downwards from the required pitch. Meanwhile, she tended to sing parts of the song at a constant volume with little stress change, where abrupt changes in stress were actually required. Together, these tendencies rendered her singing very monotonous and unmelodious.

She was invited on three different occasions to sing along with a CD that contained a pre-stroke recording of her singing or with piano accompaniment (Fig. 7B). This manipulation should have improved her singing, making it closer to her pre-stroke performance if the patient relied largely on auditory feedback. Indeed, little variability was now noted in the descending portion (initial part of the recording), but some differences emerged in the ascending portion. When the pitch rose or fell by a large interval, she tended to overshoot, so that she eventually ended up in a voice that was higher or lower in pitch than was required (gray zones in Fig. 7B).

The overshoot was more prominent when the patient was asked to sing to a piano accompaniment. Now the pitch deviated more from her pre-stroke singing, again where there were sudden rises or falls in pitch (shown with brackets and gray zones in Fig. 7C). In contrast, in the descending portion where the pitch change was gradual, the deviation from pre-stroke singing was minimal.

We compared pitch values of the correct and sung versions of each note. The instantaneous pitch differences between the pre-stroke singing sample recorded on CD and the post-stroke singing samples under various conditions were plotted as a function of time (Fig. 7D). Again, overshoots and undershoots were prominent where there were sudden rises or falls in pitch, but not where pitch changes were gradual. Interestingly, when the patient sang with piano accompaniment, marked overshoots occurred during rising portions of the song (represented as the three gray zones in Fig. 7);


Fundamental frequency (Hz)
450
400
350
300
250
200
100
(A) 50
▪ Singing alone with feedback
▪ Singing alone without feedack
▪ Pre-stroke singing
450
400
350
300
250
200
100
(B) 50
▪ Singing with CD
▪ Pre-stroke singing
450
400
350
300
250
200
100
(C) 50
▪ Singing to piano accompaniment
▪ Pre-stroke singing
Pitch difference (Hz)
150
100
50
0
-50
-100
(D)-150
0
5
10
15
20
25
Time (s)
▪ Singing alone with feedback
▪ Singing alone without feedack
▪ Singing with CD
▪ Singing to piano accompaniment

undershoots occurred during descending portions. Similar but smaller overshoots and undershoots were noted when the subject sang along with CD. In contrast, slight undershoots for rising portions existed when singing alone with or without feedback.

5. Discussion

Inspired by the idea of modularity of the mind (Fodor, 1983; Gardner, 1983), many studies have attempted to link cognitive subunits of musical abilities to specific lesions, but only fragmentary evidence has been accumulated. Lesions in the right middle to posterior part of the temporal lobe have been linked to receptive types of amusia, as would be expected given that this region is involved in music perception (Samson & Zatorre, 1994; Sidtis & Volpe, 1988; Zatorre & Samson, 1991; Zatorre et al., 1994). On the other hand, expressive amusia, such as instrumental and vocal amusia, has been associated with lesions in the superior temporal lobe (McFarland & Fortin, 1982; Takeda et al., 1990) or the frontal lobe (Botez & Wertheim, 1959) of the right hemisphere, in which cases musical perception was relatively preserved. The right superior temporal lobe lesion in the present case impaired not only the patient's musical perception, but also disrupted expressive aspects of her musical ability, which had qualified her as a professional singer. These cases together raise the question of whether the impaired expressive aspects of music emerged entirely as a result of impairment in music perception, or appeared because the right superior temporal cortex also plays some role in musical motor output per se.

Despite the patient's normal hearing acuity, psychophysical examination revealed a distinct pattern of defect in pitch perception and impaired ability to discriminate changes in loudness and timbre. Indeed, auditory feedback appeared to be essential for singing even overlearned songs to fine-tune the subtle pitch changes in real time (Figs. 6 and 7A–C). Therefore, deterioration of the patient's singing resulted partly from a degraded pitch perception system that prevented her from monitoring the pitch of her voice as precisely as she did prior to stroke.

However, the gravely impaired ability to reproduce the presented pitch is unlikely to be explainable solely as a result of compromised auditory feedback. The deficit in pitch perception was restricted to one or two semitone distances (Fig. 3). It was apparently mild compared to the severe impairment of pitch perception observed in congenital amusia on the 1-semitone distance test (Ayotte, Peretz, & Hyde, 2002; Peretz et al., 2002). In contrast, the reproduced pitch, especially in the head voice, seemed to "stray" largely away from the 95% confidence range of normal performance, lying well outside the 1–2-semitones range, as observed for pitch reproduction especially in the head voice (Fig. 3).

On the other hand, the patient had little problem in producing movements of the tongue, pharyngeal or other vocal muscles after her stroke. Therefore, even if the lesion in the right superior temporal cortex may have affected musical motor output per se, it was not because of defective motor control of the vocal muscles.

The overall similarity between her pre- and post-stroke performance under different conditions of singing was consistent with the notion that because this was an overlearned song, the patient relied more on internal representations such as musical memory or imagery for singing than on auditory feedback (Fig. 7). However, at some parts of the song, the singing performance varied greatly with the singing condition; when the pitch rose or fell by a large interval, the patient's singing tended to overshoot, so that she ended up in a pitch that was higher or lower than was required. Again, these findings are difficult to explain if the patient's singing had deteriorated simply because of the defective auditory feedback, but rather indicated an altered interaction between the vocal output and its auditory feedback.

Both auditory feedback and memory (or imagery) are important in singing because one learns to sing a song first by listening to other people sing and then by trying to match his or her vocal output to a stored memory (Jürgens, 2002). A similar matching of vocal output to stored memory or auditory imagery may be at work when the patient tries to "reproduce" a song she used to sing proficiently before the stroke, whereas disruption of this interaction would result in the deterioration of the exquisite performances of professional singers. Connections between auditory association cortices and other cortical areas may be important for this process because they are involved not only in auditory feedback, but also in sending corollary information to the auditory-decoding regions about intended vocalizations or intended pitches to facilitate the resulting acoustic events (Perry et al., 1999).

Increased sensorimotor coupling may be particularly important for the quality of musical performance (Sergent, 1993). Studying functional MRI of amateurs and professionals, Lotze, Scheler, Tan, Braun, & Birbaumer (2003) showed that the right primary auditory cortex was activated in both amateurs and professionals during execution of performance of Mozart's violin concerto in G major, which was more pro-

Fig. 7. Comparison of the patient's singing before and after stroke. Fundamental frequency of voice (ordinate) plotted as a function of time (abscissa) when the patient sang part of an overlearned song before and after the stroke. (A) Singing alone without piano accompaniment, either with (green dots) or without feedback (purple dots). These were compared with pre-stroke singing (blue dots). (B) Singing along with a CD (black dots) in comparison with the patient's pre-stroke singing (blue dots). (C) Singing to a piano accompaniment alone (red dots) and the patient's pre-stroke singing (blue dots). Brackets as in (B). (D) The instantaneous pitch differences for each note (ordinate) between the pre-stroke performance and the post-stroke performances under various conditions was plotted as a function of time (abscissa). Green curve: singing alone with feedback, purple curve: singing alone without feedback, black curve: singing along with CD, red curve: singing to a piano accompaniment.

nounced for professionals. This activation was considered to reflect the role of the right auditory cortex not only in music perception, but also the increased audio-motor associative connectivity that exists only in the presence of actual performance or real-world stimuli.

The contribution of the posterior insular lesion to a patient's symptoms should be considered in addition to superior temporal lobe involvement. This part of the lesion could have impaired this patient's singing ability because insular activation has been associated with motor performance during singing (e.g. vocalization) rather than auditory feedback in previous neuroimaging studies (Jeffries et al., 2003; Perry et al., 1999; Riecker et al., 2000). Alternatively, in analogy to left insular lesions causing conduction aphasia, the right insular lesion may have compromised pitch processing involving specific connections between the motor and auditory cortical areas, thereby disrupting effective online transformation of auditory imagery or memory into the intended vocal output.

Acknowledgement

We are grateful to Professor John Rothwell for his skillful editing and useful suggestions regarding the manuscript.

References

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

Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C., et al. (1998). Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience*, *10*, 536–540.

Botez, M. I., & Wertheim, N. (1959). Expressive aphasia and amusia following right frontal lesion in a right-handed man. *Brain*, *82*, 186–203.

Brain, W. R. (1941). Visual disorientation with special reference to lesions of the right cerebral hemisphere. *Brain*, *64*, 244–272.

Brunner, H. (1922). Klinische Beiträge zur Frage der Amusie. *Archiv für Ohren, Nasen, Kehle, Kopf Krankheiten*, *109*, 47.

Chiba, A., Obi, T., Takeda, K., Sakuta, M., & Bandou, M. (1989). A case presenting with amusia after hemorrhage in the right putamen. *Rinshou Shinkeigaku*, *29*, 536 [Japanese].

Confavreux, C., Croisile, B., Garassus, P., Aimard, G., & Trillet, M. (1992). Progressive amusia and aprosody. *Archives of Neurology*, *49*, 971–976.

Fodor, G. A. (1983). *The modularity of mind*. Cambridge, USA: MIT-Press.

Gardner, H. (1983). Frames of mind. In *The theory of multiple intelligences*. NY, USA: Basic Books.

Griffiths, T. D., Rees, A., Witton, C., Cross, P. M., Shakir, R. A., & Green, G. G. (1997). Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study. *Brain*, *120*, 785–794.

Halpern, A. R. (2001). Cerebral substrates of musical imagery. *Annals of the New York Academy of Sciences*, *930*, 179–192.

Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, *9*, 697–704.

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, *42*, 1281–1292.

Hofman, S., Klein, C., & Arlazoroff, A. (1993). Common hemisphericity of language and music in a musician. A case report. *Communication Disorders*, *26*, 73–82.

Ishikawa, S., & Inoue, N. (1989). A case with right temporoparietal infarction presenting with motor amusia. *Neuropsychology*, *4*, 156 [Japanese abstract].

Jeffries, K. J., Fritz, J. B., & Braun, A. R. (2003). Words in melody: An H(2)15O PET study of brain activation during singing and speaking. *Neuroreport*, *14*, 749–754.

Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, *123*, 155–163.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavior Reviews*, *26*, 235–258.

Lotze, M., Scheler, G., Tan, H. R., Braun, C., & Birbaumer, N. (2003). The musician's brain: Functional imaging of amateurs and professionals during performance and imagery. *Neuroimage*, *20*, 1817–1829.

Mann, I. (1898). Casuistische Beiträge zur Hirnchirurigie und Hirnlokalisation. *Monatschrift für Psychiatrie und Neurologie*, *4*, 369.

McFarland, H. R., & Fortin, D. (1982). Amusia due to right temporoparietal infarct. *Archives of Neurology*, *39*, 725–727.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, *113*, 1185–1205.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., et al. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron*, *33*, 185–191.

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*, 3453–3458.

Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., & Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport*, *11*, 1997–2000.

Samson, S., & Zatorre, R. J. (1994). Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia*, *32*, 231–240.

Seashore, C. E., Lewis, J. G., & Saetveit, J. G. (1960). *Seashore measures of musical talents*. NY, USA: The psychological corporation.

Sergent, J. (1993). Music, the brain and Ravel. *Trends in Neurosciences*, *16*, 168–172.

Sidtis, J. J., & Volpe, B. T. (1988). Selective loss of complex-pitch or speech discrimination after unilateral lesion. *Brain and Language*, *34*, 235–245.

Sparr, S. A. (2002). Receptive amelodia in a trained musician. *Neurology*, *59*, 1659.

Takeda, K., Bandou, M., & Nishimura, Y. (1990). Motor amusia following a right temporal lobe hemorrhage—a case report. *Rinsho Shinkeigaku*, *30*, 78–83 [Japanese].

Wertheim, N. (1969). The amusias. In P. J. Vinken & G. W. Bruyn (Eds.), *Handbook of clinical neurology: vol. 4* (pp. 195–180). Amsterdam, Netherland: North Holland Publishing Company.

Zatorre, R. J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia*, *23*, 31–41.

Zatorre, R. J. (1989). Intact absolute pitch ability after left temporal lobectomy. *Cortex*, *25*, 567–580.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *11*, 946–953.

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, *14*, 1908–1919.

Zatorre, R. J., & Halpern, A. R. (1993). Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia*, *31*, 221–232.

Zatorre, R. J., & Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain*, *114*, 2403–2417.