Perceptual and Motor Skills, 2009, 108, 219-228.

HEMISPHERIC SPECIALIZATION FOR LINGUISTIC PROCESSING OF SUNG SPEECH[1,2]

SERENA K. YELLE AND GINA M. GRIMSHAW

California State University San Marcos *Victoria University of Wellington*

Summary.—The two hemispheres of the brain play complementary roles in song perception, with the left hemisphere specialized for processing the linguistic aspects of song and the right hemisphere specialized for the processing of melody. However, very little is known about how language and melody interact. The present study tested the hypothesis that right hemisphere linguistic processing would be facilitated by the presence of melody. In a dichotic listening paradigm, participants (8 men, 43 women) performed a linguistic task while listening to spoken or sung speech. Contrary to the hypothesis, left hemisphere specialization for linguistic processing was identical whether the sentences were spoken or sung.

Listening to songs is a human activity that has received surprisingly little attention from psychologists. Although considerable progress has been made toward understanding the cognitive and neural systems which underlie language comprehension (e.g., Martin, 2003) and knowledge about analogous systems which underlie music perception has been growing (e.g., Peretz & Zatorre, 2005), much less is known about how these systems interact in song perception. The interaction between linguistic and melodic processes is of further interest because such interaction is also interhemispheric, with linguistic and melodic processes lateralized primarily to the left and right hemispheres, respectively (Kimura, 1961, 1964; Bartholomeus, 1974; Schön, Gordon, & Besson, 2005), although such lateralization is relative and not absolute.

In many ways, melody is similar to emotional prosody, the patterns of pitch, intensity, and duration which reflect the emotional state of the speaker. In fact, several researchers have argued for homology between melody and emotional prosody (Patel, Peretz, Tramo, & Labreque, 1998; Juslin & Laukka, 2003a, 2003b; Magne, Schön, & Besson, 2003; Nicholson, Baum, Kilgour, Koh, Munhall, & Cuddy, 2003), with the perception of both associated with activity in the right temporal cortex (Buchanan, Lutz, Mirzazade, Specht, Shah, Zilles, & Jäncke, 2000; Zatorre, Belin, & Penhune, 2002; Peretz & Zatorre, 2005). Like song perception, emotional speech perception requires the coordinated effort of left and right hemispheric processes (Ivry

---

[1]Address correspondence to Gina M. Grimshaw, Ph.D., School of Psychology, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand or e-mail (gina.grimshaw@vuw.ac.nz).
[2]We thank Alexis Hamilton for assistance with data collection and Laury Flora for his vocal talents.

DOI 10.2466/PMS.108.1.219-228

& Robertson, 1998; Poeppel, 2003; Friederici & Alter, 2004; Boemio, Fromm, Braun, & Poeppel, 2005). Such interactions between melodic and linguistic processes in song perception might mirror interactions between prosodic and linguistic processes in emotional speech perception.

It has previously been demonstrated that the presence of emotional prosody modulates hemispheric specialization for linguistic processing (Grimshaw, Kwasny, Covell, & Johnson, 2003; Grimshaw, Séguin, & Godfrey, in press). In a dichotic listening paradigm, participants listened for a target word, a task that typically yields a right ear advantage which reflects left hemisphere specialization for linguistic processing. In one condition, words were spoken in a neutral tone of voice, and the expected right ear advantage was observed. However, when the words were spoken in a sad voice, the right ear advantage was attenuated in the accuracy data and eliminated in the RT data. Furthermore, this attenuation was driven primarily by an improvement in performance for the left ear, suggesting that the presence of emotional prosody (which was irrelevant for the task) facilitated right hemispheric linguistic processing. Thus, while the left hemisphere is certainly superior to the right in linguistic processing, this lateralization is relative, not absolute. Under conditions of emotional prosody, the right hemisphere's linguistic competence is seen to emerge.

The present study used an analogous procedure to assess whether the presence of melody facilitates right hemisphere linguistic processing of lyrics. There are some hints in the literature that the right hemisphere might fare better with song than with speech. In the realm of song production, there are case reports of nonfluent aphasics who have preserved singing in the presence of impaired speech (Yamadori, Osumi, Masuhara, & Okubo, 1977; Morgan & Tilluckdharry, 1982) and reports of patients undergoing the Wada procedure who show greater deficits for speaking than for singing after injection of the left hemisphere (Gordon & Bogen, 1974). Mouth asymmetries in aphasic patients (which presumably reflect activation of the contralateral hemisphere) are rightward for speech but leftward or symmetrical for singing (Graves & Landis, 1985, 1990). The success of melodic intonation therapy in the treatment of aphasia also has been attributed to melodic activation of right hemisphere linguistic processes (Albert, Sparks, & Helm, 1973; Sparks, Helm, & Albert, 1974; Wilson, Parsons, & Reutens, 2006); although this claim is controversial (Hebert, Racette, Gagnon, & Peretz, 2003; Peretz, Gagnon, Hebert, & Macoir, 2004). Finally, neuroimaging research also points to a role for the right hemisphere in song production (Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000; Jeffries, Fritz, & Braun, 2003; Brown, Martinez, & Parsons, 2006).

Although these findings suggest that melody facilitates right hemisphere language production, alternative explanations are viable. For example, it is possible that music produces a general activational arousal that facilitates the function of damaged left hemisphere language substrates in brain-damaged patients. Neuroimaging of song production in intact individuals may be capturing melodic processes that occur during singing, and not linguistic processes. Furthermore, it is also unclear how asymmetries in song production translate to asymmetries in song perception.

The few studies that have examined song perception have used attentional manipulations to compare hemispheric effects when attention is directed toward either the linguistic or the melodic dimension (Bartholomeus, 1974; Goodglass & Calderon, 1977). For example, Bartholomeus (1974) used a dichotic listening task in which participants listened to sung speech. She found a right ear advantage when the task required attention to the words, and a left ear advantage when attention was directed to the melody. However, this research paradigm lacks the critical comparison—that of hemispheric specialization for spoken and sung speech when the linguistic task is held constant. Because this study did not include a spoken condition, it is not possible to know whether the magnitude of the right ear advantage for lyrics was similar to that which would be observed for spoken letter sequences.

On face value, neuroimaging studies seem best suited to answering questions about localization of linguistic and musical perception. However, most neuroimaging research has directly compared perception of spoken and sung speech, and it is unclear if the additional right hemisphere activity typically observed in the musical condition reflects neural activity associated with melodic processing or a change in activity associated with linguistic processing (e.g., Gaab, Gaser, Zaehle, Jäncke, & Schlaug, 2003; Brown, Martinez, Hodges, Fox, & Parsons, 2004; Callan, Kawato, Parsons, & Turner, 2007). A notable exception is a study reported by Schön, *et al.* (2005). In fMRI, they had participants perform a linguistic task (same–different judgment) for pairs of stimuli that were either spoken words, sung words, or vocalises (a single nonsense syllable, sung to a melody), relative to a noise-only baseline condition. Very similar patterns of bilateral activation across temporal regions were observed for all conditions, but critically, a comparison of sung speech with spoken speech showed increased activation in the right middle temporal gyrus. However, this effect was limited to one voxel only and should be considered tentative at best.

The current study used a dichotic listening task to examine hemispheric differences in linguistic processing of spoken and sung sentences. The procedure was identical to that used by Grimshaw, *et al.* (2003), which demonstrated effects of emotional prosody on asymmetry for linguistic processing. In both conditions the task was to process the sentence linguistically, that is, whether the sentence was spoken or sung was irrelevant for the task. It was predicted that a right ear advantage would be observed for spoken speech. However, if the right hemisphere makes a greater linguistic contribution to sung than spoken speech, then the right ear advantage should be reduced, eliminated, or perhaps even reversed to a left ear advantage in the sung condition.

METHOD

Participants

The participants (8 men, 43 women) were undergraduate students whose mean age was 22 yr. All participants wrote with the right hand and had a positive (rightward) score on the Waterloo Handedness Questionnaire (Steenhuis & Bryden, 1989) which assesses hand use for 32 skilled and unskilled activities. Participants had no known hearing deficits and were not musicians[3] (no music lessons after the age of 15 years), with mean instrumental musical training (including school music classes) of 2.3 yr. All were native English speakers or learned English before the age of 5 years, and all provided informed consent at the beginning of the experimental session. The study was approved by the Institutional Review Board at California State University San Marcos.

Apparatus and Stimuli

The dichotic listening task was presented using the SuperLab 2.0 program (Cedrus, USA) on a PC computer through Sony MDR-V150 hi-definition headphones with circumaural cushions. Ten sets of three sentences (of either five or eight syllables) were recorded in both spoken and sung form; sung sentences used one of two melodies (see Fig. 1). Three versions of each sentence were constructed that differed only by a single target word to produce 30 different sentences (e.g., Can you fix the *brain*? Can you fix the *drain*? Can you fix the *train*?). The target words differed only in the initial consonant (/b/, /d/, /p/, or /t/). From each set of three sentences, two dichotic pairs were constructed for which the target words differed in either phonemic voicing or place of articulation but not both. For example, the three sentences cited above produced two pairings, "Can you fix the *brain*/Can you fix the *drain*," for which the target words differed only in place of articulation, and "Can you fix the *drain*/Can you fix the *train*," for which the targets differed only in voicing. This created 20 sentence pairings. Two variations were produced of each sentence-pairing so that each sentence was presented to each ear for a total of 40 unique dichotic sentence pairs.

[3]Here the term nonmusician is used to indicate that these participants do not have formal musical training. This of course does not mean that they are musically naïve or inexperienced, as music is a fundamental human activity.

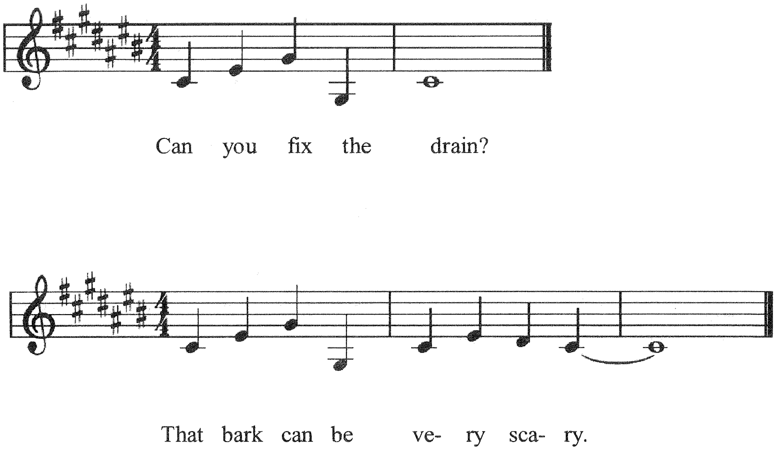

FIG. 1. Melodies presented to participants

The sound files were recorded by a male speaker with musical and vocal training. The sentences were recorded using GoldWave recording software and edited using Audacity. Sentences were dichotically paired by visual inspection of the sound waves so that the onsets of the target words were simultaneous.

Procedure

The procedure is illustrated in Fig. 2. For each trial, a target word was presented centrally on the screen for 800 msec., immediately followed by the dichotic sentences. Participants were required to press the "B" key if the target word was present in either of the sentences and the "N" key if the

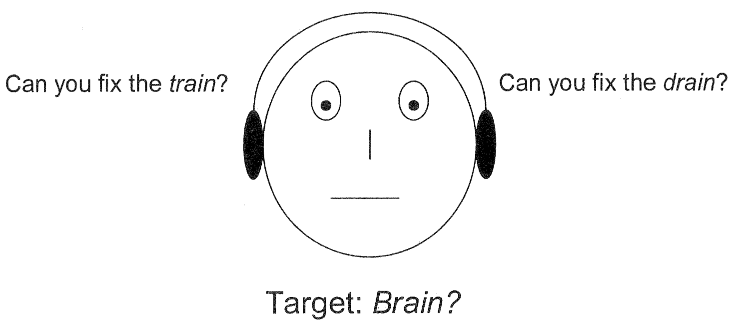

FIG. 2. Schematic representation of an "absent" trial. The participant sees the target word on the screen, followed immediately by the dichotic sentence pairs. On half the trials the target is a nonpresented word, on one-quarter of the trials it is the word in the left ear, and on one-quarter of the trials it is the word in the right ear.

target word was not present. Twenty-five percent of the time the target was presented to the left ear, 25% of the time it was presented to the right ear, and 50% of the time the target was absent (the word from the third, nonpresented sentence in the set). Participants were asked to respond as quickly and accurately as possible after presentation of the sentences, although response times were not recorded. After each response, there was a 1,000-msec. interval until the presentation of the next target. Each dichotic sentence pair was presented four times, for a total of 160 trials in each modality (spoken and sung). The total 320 trials were divided into four blocks consisting of 80 trials each, with block order counterbalanced across participants (sung/spoken/spoken/sung or spoken/sung/sung/spoken). Earphone placement was counterbalanced across participants to control for any mechanical differences in the volume of each speaker.

Data Analysis

A Hit Rate (proportion of times a presented target was identified as present) was calculated for each ear in each modality (spoken and sung). Because an individual's Hit Rate is also influenced by any response bias (for example, a participant who responds "yes" on every trial will have a very high Hit Rate, but actually be very poor at discriminating the target), Hit Rate was adjusted by subtracting the False Alarm Rate for the task (proportion of times a nonpresented target was identified as present) to yield a corrected Hit Rate. A corrected Hit Rate of zero reflects chance performance (Hit Rate = False Alarm Rate).

Results

Hit Rates and corrected Hit Rates were calculated for each person for each condition (spoken left ear, spoken right ear, sung left ear, sung right ear; see Table 1). Two participants were excluded from the analysis because their performance was at or near chance across all conditions. Analyses are therefore based on a sample of 49 participants (41 women, 8 men; see

TABLE 1
Performance as a Function of Ear and Modality

	Hit Rate		False Alarm Rate	
	M	*SD*	*M*	*SD*
Spoken			.12	.09
Left Ear	.65	.14		
Right Ear	.81	.10		
Ear Advantage	.16			
Sung			.12	.07
Left Ear	.62	.14		
Right Ear	.79	.10		
Ear Advantage	.17			

Note.—Positive ear advantages reflect a right ear advantage.

Table 1). Initial omnibus analyses of variance yielded no main effects or interactions involving sex in either dependent measure (all F ratios < 1.0), so reported analyses were collapsed across sex.

Hit Rate Analyses

The primary hypothesis predicted an interaction between ear and modality such that the right ear advantage that should be observed for the spoken condition would be attenuated, eliminated, or reversed for the sung condition. Hit Rates were analyzed in a 2 (modality: spoken, sung) × 2 (ear: left, right) repeated-measures analysis of variance. The analysis gave an overall right ear advantage ($F_{1,47} = 89.57$, $MSE = 1.30$, $p < .001$) and a main effect of modality ($F_{1,47} = 5.24$, $MSE = .03$, $p = .03$) such that participants were more accurate in the spoken condition than in the sung condition. The ear by modality interaction was not significant ($F_{1,48} = .04$, $MSE = .001$, ns), indicating that the magnitude of the right ear advantage was the same for both spoken ($M = .16$) and sung ($M = .17$) speech.

Corrected Hit Rate Analyses

Corrected Hit Rates were similarly analyzed, showing a right ear advantage ($F_{1,47} = 89.57$, $MSE = .02$, $p < .001$) and a main effect of modality which approached significance ($F_{1,47} = 3.53$, $MSE = .008$, $p = .07$). These results parallel those observed in the Hit Rate analysis and indicate that participants did not have differences in response bias which could account for any effects. Notably, the interaction of modality × ear was again not significant ($F_{1,47} = .04$, $MSE = .008$, ns).

Discussion

On the basis of our previous study of the effects of emotional prosody on linguistic processing, it was predicted that the right hemisphere would show a greater contribution to the linguistic processing of sung speech than of spoken speech, leading to a reduction, elimination, or reversal of the typical right ear advantage. Contrary to this prediction, right ear advantages of identical magnitude were observed for the processing of both spoken and sung speech. The left hemisphere appears to process speech equally well whether it is spoken or sung with a melody.

The design of this experiment paralleled that used by Grimshaw, *et al.* (2003) to study emotional prosody. The effects of melody on linguistic processing are clearly different than the effects of emotional prosody, even though both melody and prosody are predominantly processed in the right hemisphere in possibly common right temporal areas (Zatorre, *et al.*, 2002). There are three possible explanations for the discrepancy. First, emotional prosody is clearly a component of language, whereas melody is not. Right hemisphere cortical systems involved in the perception of prosody are largely homologous to speech-processing areas in the left hemisphere, particularly in the middle temporal gyrus (Poeppel, 2003; Boemio, *et al.*, 2005). Prosodic processing systems may be better integrated with right hemisphere linguistic processing than are melodic processing systems.

Alternatively, differences between studies could reflect differences in the spectral and temporal properties of emotional prosody and melody. For example, sad and happy prosodies typically involve pitch changes of over an octave across a continuous range within a sentence (Bänziger & Scherer, 2005), whereas most melodies (and those in the present study) involve pitch changes of less than half an octave, and those pitches are standardized to a tonal system. Temporal characteristics also differ; musical rhythm is highly constrained and predictable, whereas rhythmic changes are important signals of emotional prosody. Similarly, efforts were made in the present study to equate the tempos of the spoken and sung speech, whereas changes in tempo are among the defining characteristics of emotional prosody.

Finally, while most musical events have a strong emotional component (Blood, Zatorre, Bermudez, & Evans, 1999), the simple melodies used here have a minimal affective component. Thus, while the present goal was to produce spoken and sung stimuli that differed only in melodic intonation, it may well be other aspects of music which are analogous to emotional prosody and more likely to engage right hemisphere linguistic systems.

Current conclusions rely on the interpretation of a null effect, which is always problematic. However, the ear advantages for the spoken and sung conditions were almost identical over a substantial number of participants, suggesting that the null effects are not reflective of restricted power. Furthermore, the design mirrored that used in previous studies of emotional prosody, in which significant differences between neutral and emotional speech were observed. The pattern of results observed here is certainly consistent with equivalent left hemisphere specialization for both spoken and sung speech (Gordon & Bogen, 1974; Jeffries, *et al.*, 2003; Morgan & Tillluckdharry, 1982). Although there is some evidence linking the right hemisphere with song production, its role in song perception appears limited to its musical as opposed to its linguistic components.

REFERENCES

ALBERT, M. L., SPARKS, R. W., & HELM, N. A. (1973) Melodic intonation therapy for aphasia. *Archives of Neurology*, 29, 130-131.

BÄNZIGER, T., & SCHERER, K. R. (2005) The role of intonation in emotional expressions. *Speech Communication:* [Special Issue] *Quantitative prosody modelling for natural speech description and generation*, 46, 252-267.

BARTHOLOMEUS, B. (1974) Effects of task requirements on ear superiority for sung speech. *Cortex*, 10, 215-223.

BLOOD, A. J., ZATORRE, R. J., BERMUDEZ, P., & EVANS, A. C. (1999) Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, 2, 382-387.

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience*, 8, 389-395.

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.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006) Language and music side by side in the brain: a PET study of the generation of melodies and sentences. *European Journal of Neuroscience*, 23, 2791-2803.

Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., & Jäncke, L. (2000) Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cognitive Brain Research*, 9, 227-238.

Callan, D. E., Kawato, M., Parsons, M., & Turner, R. (2007) Speech and song: the role of the cerebellum. *The Cerebellum*, 6, 321-327.

Friederici, A. D., & Alter, K. (2004) Lateralization of auditory language functions: a dynamic dual pathway model. *Brain & Language*, 89, 267-276.

Gaab, N., Gaser, C., Zaehle, T., Jäncke, L., & Schlaug, G. (2003) Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. *NeuroImage*, 19, 1417-1426.

Goodglass, H., & Calderon, M. (1977) Parallel processing of verbal and musical stimuli in right and left hemispheres. *Neuropsychologia*, 15, 397-407.

Gordon, H. W., & Bogen, J. E. (1974) Hemispheric lateralization of singing after intracarotid sodium amylobarbitone. *Journal of Neurology, Neurosurgery & Psychiatry*, 37, 727-738.

Graves, R., & Landis, T. (1985) Hemispheric control of speech expression in aphasia: a mouth asymmetry study. *Archives of Neurology*, 42, 249-251.

Graves, R., & Landis, T. (1990) Asymmetry in mouth opening during different speech tasks. *International Journal of Psychology*, 25, 179-189.

Grimshaw, G. M., Kwasny, K. M., Covell, E., & Johnson, R. A. (2003) The dynamic nature of language lateralization: effects of lexical and prosodic factors. *Neuropsychologia*, 41, 1008-1019.

Grimshaw, G. M., Séguin, J. A., & Godfrey, H. K. (in press) Once more with feeling: the effects of emotional prosody on hemispheric specialisation for linguistic processing. *Journal of Neurolinguistics*.

Hebert, S., Racette, A., Gagnon, L., & Peretz, I. (2003) Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain*, 126, 1838-1850.

Ivry, R. B., & Robertson, L. C. (1998) *The two sides of perception*. Cambridge, MA: MIT Press.

Jeffries, K. J., Fritz, J. B., & Braun, A. R. (2003) Words in melody: an H_2O^{15} PET study of brain activation during singing and speaking. *NeuroReport*, 14, 749-754.

Juslin, P., & Laukka, P. (2003a) Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological Bulletin*, 129, 770-814.

Juslin, P., & Laukka, P. (2003b) Emotional expression in speech and music: evidence of cross-modal similarities. *Annals of the New York Academy of Sciences*, 1000, 270-282.

Kimura, D. (1961) Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 15, 166-171.

Kimura, D. (1964) Left-right differences in the perception of melodies. *The Quarterly Journal of Experimental Psychology*, 16, 355-358.

Magne, C., Schön, D., & Besson, M. (2003) Prosodic and melodic processing in adults and children: behavioral and electrophysiologic approaches. *Annals of the New York Academy of Sciences*, 999, 461-476.

Martin, R. C. (2003) Language processing: functional organization and neuroanatomical basis. *Annual Review of Psychology*, 54, 55-89.

Morgan, O. S., & Tilluckdharry, R. (1982) Preservation of singing function in severe aphasia. *West Indian Medical Journal*, 31, 159-161.

Nicholson, K. G., Baum, S., Kilgour, A., Koh, C. K., Munhall, K. G., & Cuddy, L. L. (2003) Impaired processing of prosodic and musical patterns after right hemisphere damage. *Brain and Cognition*, 52, 382-389.

Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998) Processing prosodic and musical patterns: a neuropsychological investigation. *Brain & Language*, 61, 123-144.

PERETZ, I., GAGNON, L., HEBERT, S., & MACOIR, J. (2004) Singing in the brain: insights from cognitive neuropsychology. *Music Perception*, 21, 373-390.

PERETZ, I., & ZATORRE, R. J. (2005) Brain organization for music processing. *Annual Review of Psychology*, 56, 89-114.

POEPPEL, D. (2003) The analysis of speech in different temporal integration windows: cerebral lateralization as "asymmetric sampling in time." *Speech Communication*, 41, 245-255.

RIECKER, A., ACKERMANN, H., WILDGRUBER, D., DOGIL, G., & GRODD, W. (2000) Opposite hemisphere lateralization effects during speaking and singing at motor cortex, insula, and cerebellum. *NeuroReport*, 11, 1997-2000.

SCHÖN, D., GORDON, R. L., & BESSON, M. (2005) Musical and linguistic processing in song perception. *Annals of the New York Academy of Science*, 1060, 71-81.

SPARKS, R., HELM, N., & ALBERT, M. (1974) Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex*, 10, 303-316.

STEENHUIS, R. E., & BRYDEN, M. P. (1989) Different dimensions of hand preference that relate to skilled and unskilled activities. *Cortex*, 25, 289-304.

WILSON, S. J., PARSONS, K., & REUTENS, D. C. (2006) Preserved singing in aphasia: a case study of the efficacy of melodic intonation therapy. *Music Perception*, 24, 23-36.

YAMADORI, A., OSUMI, Y., MASUHARA, S., & OKUBO, M. (1977) Preservation of singing in Broca's aphasia. *Journal of Neurology, Neurosurgery & Psychiatry*, 40, 221-224.

ZATORRE, R. J., BELIN, P., & PENHUNE, V. B. (2002) Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, 6, 37-46.

Accepted January 5, 2009.