Acta Acustica united with Acustica
Vol. 94 (2008) 588 – 593

DOI 10.3813/AAA.918068

Language-Selective Interference with Long-Term Memory for Musical Pitch

I-Hui Hsieh, Kourosh Saberi
Department of Cognitive Sciences, Center for Cognitive Neuroscience, University of California, Irvine, CA 92697-5100, USA

Summary
This study examines linguistic influences on musical pitch processing and provides evidence for a form of language-selective interference with absolute-pitch (AP) memory. We show that voiced solfege syllables whose fundamental frequencies and harmonic structures are digitally shifted to precisely map onto a mismatched musical note can selectively interfere with pitch identification by some but not other AP musicians. Interference diminishes as the stimulus spectrum is increasingly lowpass filtered to remove its broadband speech features. Time reversal of mismatched pitch-syllable "hybrids", which distorts their phase spectra but leaves their amplitude spectra intact, also substantially reduces interference. These findings support recent theories of AP encoding that propose an intrinsic association between linguistic cues and stored pitch representations in extraction and accurate labeling of pitch from long-term memory.

PACS no. 43.75.Cd, 43.66.Lj, 43.66.Hg

1. Introduction

Theories of AP encoding, supported by psychophysical and neuroimaging evidence have proposed an in-built association between stored pitch representations and linguistic processes in facilitating the retrieval and labeling of pitch [1, 2, 3, 4, 5, 6]. Psychophysical evidence for a verbal-encoding model comes from studies that show accurate pitch production in singing familiar songs by non-musicians [1] and from consistency in pitch of spoken words by non-musician natives of countries that use tonal languages [2, 3]. Further evidence for a verbal-coding model comes from neuroimaging studies that have implicated the left auditory association cortex during absolute-pitch processing [5, 7] and from developmental studies which suggest that AP ability requires early musical training during critical periods of *language* acquisition [3, 8, 9, 10].

The present study examines verbal-coding mechanisms of AP by systematically varying the association between the spectral content of a musical note and its voiced linguistic label. Fixed-*Do* solfege system taught during formative stages of language development in many cultures provides a unique case of association between specific linguistic tokens and musical pitches. To our knowledge, two previous studies have examined the effects of mismatching voiced solfege syllables to musical pitch. Itoh *et al.* [11] measured event-related potentials from AP musicians during identification of mismatched stimuli and reported heightened activity in the left auditory association cortex. Miyazaki [12] investigated reaction times (RT) in identification of the pitch of mismatched syllable-pitch stimuli. In the latter study, each presentation of the mismatched stimulus was preceded by an acoustic cue (middle C on the piano) to facilitate pitch identification by both AP and non-AP musicians. He reported that *all* subjects displayed longer RTs in identifying the pitch of a mismatched syllable-pitch stimulus. Miyazaki further reported that AP subjects display longer RTs (relative to non-AP subjects) when attempting to vocally shadow the voiced *label* (syllable) of a mismatched pitch-syllable stimulus. This finding is counterintuitive in that one would expect a voiced label to be shadowed without difficulty [13] but is consistent with the idea that AP subjects involuntarily encode a linguistic representation of a sound's pitch.

One unverified assumption in these studies is that the musician *producing* mismatched stimuli for identification by AP subjects is immune to interference. Stimuli used in the Miyazaki and Itoh *et al.* studies [12, 11] were generated by a trained musician[1] required to voice solfege syllables at a mismatched pitch. There was no objective

[1] These studies do not state whether the trained musician had absolute-pitch. As will be shown in the current paper, even AP musicians can produce errors in vocally generating the pitch of a mismatched pitch-syllable stimulus (σ = 4% or nearly 1 semitone; maximum error = 13% or greater than 2 semitones). These errors are small enough that they do not distort the speech quality of the syllable when spectrally shifted to the exact note frequency, but uncorrected, are over 1 semitone in error for approximately 25% of the recordings and may be greater than 2 semitones in error even if AP musicians are used.

Received 8 September 2007, revised 20 March 2008,
accepted 23 April 2008.

verification of the accuracy of the resultant complex pitch-syllable "hybrids" and, thus, it is possible that the voiced stimuli in these studies were themselves affected by the same interference effect under investigation. In the current study we have developed pitch-syllable "hybrids" whose fundamental frequencies and harmonic structures are accurately mapped onto a target musical note by shifting their Fourier spectra using appropriate sample-rate conversion. In addition, we examine a number of stimulus conditions beyond those previously examined to determine how the amplitude and phase spectra of a mismatched pitch-syllable hybrid affect pitch identification by AP musicians.

2. Methods

2.1. Subjects

Six AP musicians (5 females) served as subjects. They were recruited from the UCI campus community through flyers and announcements in music-performance classes. Their ages ranged from 19–27 (mean = 22) and all had begun formal music training between the ages of 4 to 6 years. They were paid an hourly wage for participation. The protocol for experiments on human subjects was approved by the University of California's Institutional Review Board.

2.2. Screening for AP

AP ability was verified through a screening test using pure tones and piano notes in a single-interval 12-alternative forced-choice task. Stimuli consisted of 50 pure tones and 50 piano notes presented in two blocks of 50 trials each. Pure tones were 1s in duration with 100ms rise-decay ramps. Piano notes were digitally recorded from a 9-foot Steinway grand piano at UCI's Music Department. Notes were recorded at a sampling rate of 44.1 kHz using a 0.5-inch microphone (Brüel & Kjær Model 4189), a conditioning amplifier (Nexus, Brüel & Kjær), and a 16-bit A-to-D converter (Creative Sound Blaster Audigy 2ZS). Spectral analysis of the recorded notes confirmed that the piano was in tune. Stimuli were presented diotically at a sampling rate of 44.1 kHz through Bose headphones (model QCZ, TriPort) in a double-walled steel acoustically isolated chamber (Industrial Acoustics Company). On each trial, a musical note was randomly selected from C2 to B6 (65.4 to 1975.5 Hz; A4 = 440.0 Hz) with the constraint that two successive notes were at least 2 octaves + 1 semitone apart. Subjects were asked to identify each note by selecting 1 of 12 note labels on GUI (graphical user interface) pushbuttons. Subjects were not provided reference stimuli, practice trials, or feedback at any time during screening or experiments. Responses were scored following protocol similar to those used by Baharloo *et al.* [9] and Hsieh and Saberi [14]. Participants received 1 point for correct identification and 0.5 point for identification to within a semitone (e.g., C vs. C#). A predetermined criterion of 90% accuracy for identifying piano notes and 80% for pure tones was used to qualify a subject as AP [9, 15, 16, 14]. Averaged performance across 6 AP subjects were 98% (σ = 2.4) for piano notes and 86.6% (σ = 4.8) for pure tones (chance performance = 10.4%).

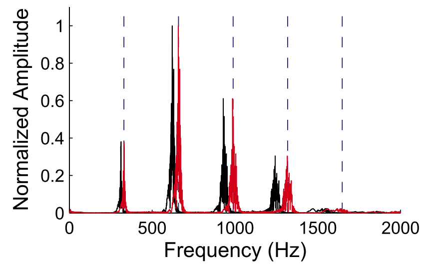

Figure 1. Spectrum of a sung solfege syllable (*Do*; black) error corrected to a mismatched musical note frequency (*Mi*; red). Blue dashed lines show target frequency (329.6 Hz) and its harmonics. The ordinate is normalized amplitude.

2.3. Generation of pitch-syllable hybrid stimuli

Hybrid stimuli consisted of all permutations of 12 musical pitches voiced with 7 solfege syllables based on the fixed-*Do* system. Three AP subjects (2 females, 1 male) recorded the permutations by singing the 12 musical notes in a one-octave range in ascending order using each of the 7 solfeggio syllables (*Do, Re, Mi, Fa, So, La*, and *Si*). The two female subjects were instructed to voice the notes in the 4th octave while the male subject sang in the 3rd octave. Subjects were allowed to rehearse each syllable-pitch combination until they were ready to record. Each of the 84 (12×7) pitch-syllable combinations was recorded for 3s at a sampling rate of 44.1 kHz. The playback digital sampling rate (SR) of the voiced pitch-syllable recordings was individually adjusted for each recorded sample such that its fundamental frequency (F_0), as determined from an FFT of the recorded voice, matched exactly that of a target note frequency and its harmonics (i.e., playback SR = (target F_0/voiced F_0) * 44.1 kHz). Corrections were generally minor ($\overline{X}$ = 1%, σ = 4% across 252 recorded samples) and did not distort the syllable's speech quality as verified by a certified audiologist from an independent laboratory. Figure 1 shows the spectrum of a sung solfeggio syllable in the 4th octave (*Do*; black) error corrected to a mismatched musical note frequency (*Mi*; red). Blue dashed lines show target frequency (329.6 Hz) and its harmonics. Five types of stimuli were used in the current study: 1) All-pass condition consisting of the original unfiltered pitch-syllable hybrids, 2) hybrids that were bandpass filtered to a 4-Hz wide band centered on the hybrid's F_0, 3) hybrids that were lowpass filtered at F_0*1.1, 4) lowpass filtered at F_0*4.1, and 5) unfiltered hybrid stimuli that were time reversed prior to presentation. All stimuli were truncated to 1000 ms in duration and were drawn randomly from mismatched conditions (set of 67 stimuli).

2.4. Procedure

The experiment was run in a randomized block design with each of the 5 stimulus conditions fixed within each of two 50-trial blocks per subject. Each trial was initiated by pressing a GUI 'Start' button. Subjects were required to respond immediately after presentation of each note, and were instructed to select the correct musical pitch ignoring the solfege syllable. On each trial, a randomly selected pitch-syllable hybrid was presented, followed by an 1800 ms ISI during which subjects responded by selecting from 12 musical note labels arranged in 2-rows of GUI pushbuttons on a monitor.

3. Results

An interesting finding was the markedly different results for two groups of AP subjects. Subjects who had received their musical training in the fixed-*Do* solfege system showed substantial interference effect, while those who had received their training in the western tradition of a movable-*Do* system showed no interference in any of the tested conditions[2]. Because of this *post hoc* finding, we plot separately the data from solfege-trained (fixed-*Do*) and western-trained (movable-*Do*) subjects. Of the six subjects, 4 had received their training in the fixed-*Do* system, and two were trained in the movable-*Do* tradition.

Results are shown in Figure 2. The top panel shows pitch-identification accuracy for piano notes (left bars), pure tones (middle), and mismatched syllable-pitch hybrid stimuli (right). Chance performance is shown as the dashed line. Consistent with prior reports all subjects performed near ceiling level for piano notes and above 80% for pure tones. The performances of fixed-*Do* and movable-*Do* AP musicians are nearly identical for both piano and pure-tone conditions. Hybrid stimuli, however, significantly interfered with pitch processing by musicians trained in the fixed-*Do* solfege system but not those trained on the western scale movable-*Do* tradition, $t(4) = 3.327$, $p < 0.05$. Movable-*Do* trained musicians identified the pitch of hybrid stimuli with 90% accuracy, nearly the same as their accuracy in identifying the pitch of pure tones. Fixed-*Do* trained subjects however had substantial difficulty identifying the pitch of hybrid stimuli (∼53% accuracy).

3.1. Effects of lowpass or bandpass filtering hybrid stimuli

Bottom panel of Figure 2 shows the effects of filtering hybrid stimuli to systematically reduce their speech quality. Stimuli were digitally filtered in the frequency domain with a lowpass cutoff equal to either 1.1 or 4.1 times the hybrid's fundamental frequency, or bandpass filtered to within 2-Hz of the fundamental. For comparison, performance for the unfiltered (all-pass) and pure-tone conditions are also shown in this figure. Filtering hybrid stimuli to remove their speech content resulted in a systematic reduction in interference and its complete elimination when a narrowband filter was centered on the syllable's fundamental. An analysis of variance on the data of the bottom panel of Figure 2 showed a significant effect of filter condition ($F_{3,12} = 6.27$, $p < 0.01$), a significant interaction between filter condition and prior form of music training (fixed-*Do* vs. movable-*Do*, $F_{3,12} = 6.59$, $p < 0.01$), but no significant effect of music training, due to convergence of performance at the narrowband-filter condition ($F_{1,4} = 2.55$, n.s.).

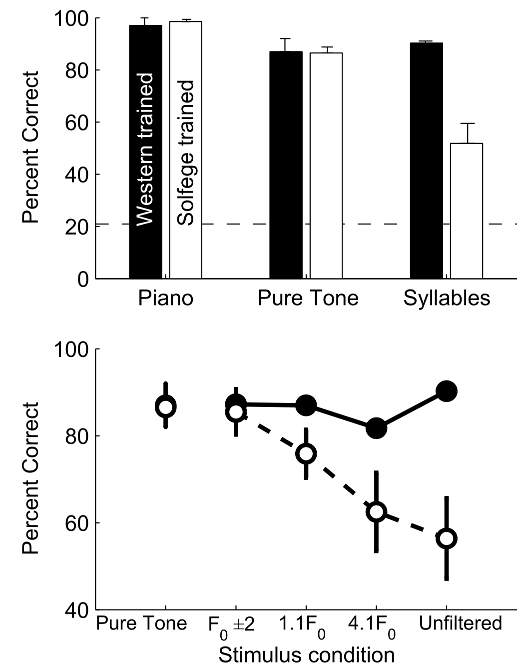

Figure 2. Top: Pitch identification performance from western (movable-*Do*) and solfege (fixed-*Do*) trained AP musicians (left bars = piano notes; middle bars= pure tones; right bars = mismatched hybrid solfege-pitch syllables). Dashed line represents chance performance. Bottom: Filled and open symbols show data from movable- and fixed-*Do* trained AP musicians respectively. Interference declines monotonically as hybrid stimuli are filtered to remove their speech content. Upper performance asymptote of ∼85% accuracy for a 4-Hz bandpass filter centered on the syllable's fundamental frequency is equivalent to pure-tone pitch identification performance. Error bars represent one standard error.

3.2. Temporal-reversal of pitch-syllable hybrids

Temporal reversal of a waveform distorts its phase spectrum but leaves its amplitude spectrum completely intact. Time reversing hybrid solfege syllables, thus, reduces their speech quality but preserves their spectrotemporal statistics and hence part of their speech-like characteristics. Figure 3 shows identification accuracy for time-reversed syl-

[2] While in some western European countries musicians are trained using a fixed-*Do* solfege system, both our western-trained AP subjects were trained in a movable-*Do* system. All four of our fixed-*Do* trained subjects were Asian.

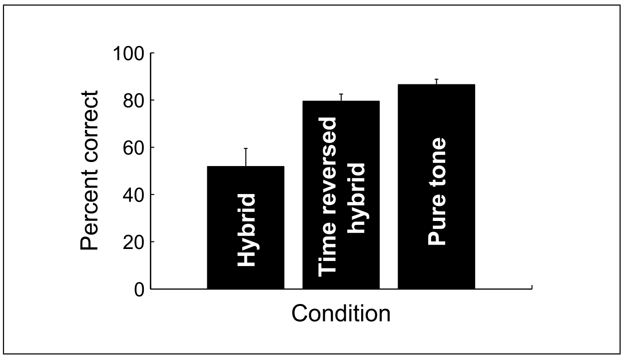

Figure 3. Percent correct identification of time-reversed syllables by fixed-*Do* trained AP musicians. Data for original hybrid and pure tones from the same subjects are also shown for comparison. Error bars represent one standard error.

lables. Results are shown only for fixed-*Do* trained subjects. Scores for original hybrids and pure tones are also shown for comparison. Subjects performed at 79.5% accuracy in the time-reversed condition, a substantial improvement over performance associated with the original hybrid condition ($t(3) = 3.63$, $p < 0.05$), but not as high as that for pure tones, though this latter difference was not statistically significant ($t(3) = 1.60$, n.s.). One possible explanation for this partial but not complete elimination of interference is that a weak vowel-like quality remains in time-reversed syllables which may induce a less effective form of interference.

3.3. Bias and intraspeaker effects

To determine if there were biases in identification of hybrid stimuli toward their voiced labels, we analyzed the distribution of responses for fixed-*Do* subjects conditioned on the stimulus label. Responses were approximately equally distributed across the 12 musical pitches for each of the 7 solfege labels (top panel of Figure 4) suggesting that participants were not biased toward a particular voiced label. For example, the majority of responses to the syllable '*Do*' are not at the response category '*Do*' but are distributed randomly across all response categories. A label-bias effect would result in significantly larger frequency counts along this figure's diagonal. In addition, we considered whether listening to hybrid stimuli generated from one's own voice had a different effect on pitch identification than listening to syllables recorded from other AP subjects. Analysis of the pooled data from only those trials on which syllables from a subject's own voice was used (bottom panel of Figure 4) showed that performance was not significantly different than that for syllables voiced by other speakers ($t(2)=0.48$, n.s.).

4. Discussion

Results of the current study show that linguistic context substantially interferes with pitch judgments by fixed-*Do*, but not movable-*Do* trained AP musicians. A number of

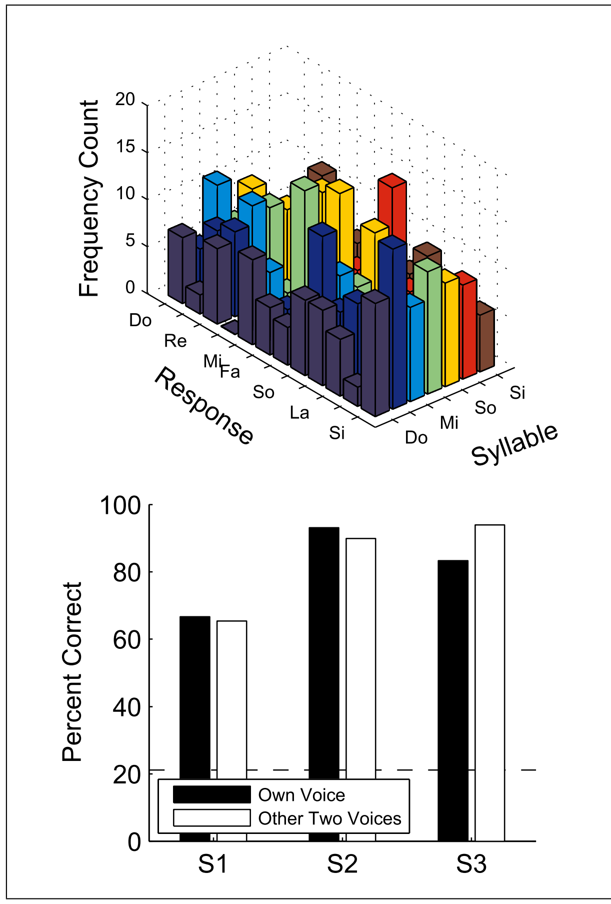

Figure 4. Top panel: Distribution of responses as a function of 7 solfege labels. Data are pooled from 4 solfege-trained (fixed-*Do*) AP subjects. No label-bias effects are observed (see text). Bottom panel: Comparison of pitch-identification accuracy for stimuli generated from one's own voice versus stimuli generated from voices of two other musicians. No significant intraspeaker effects are observed. Two of the speakers were movable-*Do* trained and one was fixed-*Do* trained. Dashed line represents chance performance.

previous studies have reported similar auditory interference effects. Cohen and Martin [17] and McClain [18] for example have shown an increase in response latencies for identifying the pitch of the word "high" spoken at a low pitch (or the word "low" spoken at a high pitch) compared to that for congruent word-pitch conditions. Green and Barber [19, 20] also reported an increase in reaction time for identification of speaker gender when spoken words were semantically opposite to speaker gender, e.g., the word "man" spoken by a woman. Other researchers have reported similar auditory interference effects in processing bilingual cues [21], left- versus right-ear stimulus presentation [22], and across sensory modalities [23, 24, 25].

All these prior studies have reported increases in *reaction times* to incongruent stimuli, as well as low error rates. The latter finding of no decline in *accuracy* of judgments is of course expected from the simple 2-response alternative design used in these experiments, e.g., male vs. female; left vs. right ear. In the current study, we quantified interference as changes in the accuracy with which subjects identified pitch in a single-interval 12-alternative forced-choice task. The substantial decline in accuracy we have observed for mismatched syllable-pitch conditions is likely a result of the more complex 12-alternative design and the 1.8s constraint placed on response interval. Miyazaki [12] who measured reaction times in an experimental design similar to ours (12 alternative) did not report on accuracy of pitch judgments. It would be of interest to determine if subjects in that study also displayed a decline in pitch-identification accuracy in addition to the reported increase in RTs for mismatch pitch-syllable conditions.

To gain better insight into how incongruent pitch-syllable hybrids interfered with pitch judgments in our study, we interviewed our AP subjects about their pitch-identification strategies. While we caution that these descriptions are subjective, they do provide important insights into why auditory interference is observed for some but not other subjects. When asked to describe the percept associated with a *pure tone* at a particular musical-note frequency, all fixed-*Do* trained AP musicians described it as speech-like. A pure-tone frequency of 293.7 Hz is best described as subjectively sounding like "*Re*" while a frequency of 440.0 Hz is heard as "*La*". Hybrid stimuli are described as sounding "wrong" with a pitch whose identification requires substantial attentional effort. AP musicians trained in the movable-*Do* system, however, describe hybrid stimuli as invoking an easily identifiable pitch.

If fixed-*Do* subjects associate a linguistic quality to musical pitch, what do movable-*Do* AP subjects, who did not show interference, encode during identification of musical notes? Unlike fixed-*Do* subjects, movable-*Do* AP musicians report highly individualized and often non-linguistic forms of associations. One such AP musician reported emotional and cross-modal associations. She noted that F# "sticks out like a sore thumb. It sounds really sharp, acid, and bitter. I hear a 'twang' sound when I hear that note." She described B-flat as "a trumpet sound and very comforting" and A-flat as "a beautiful, rich tone…sounds like paradise to me." A second western-trained (movable-*Do*) AP subject described a spatial strategy in which he first identifies, on an imagined piano keyboard, the general spatial location of the note's octave (height) and then its finer position (chroma). He described notes as having no linguistic quality. The note "C" in the fourth octave sounds entirely different than the note "C" in the fifth or other octaves. He stated that other than the fact that, in musical notation, both sounds have been labeled as "C", perceptually they have nothing in common. He further described his strategy as "if you asked me to find Paris on a map of the world…I would first find Europe, then France, then Paris". His strategy was thus based entirely on spatial associations.

Given these individualized coding strategies, one relevant question is whether and under what circumstances could one interfere with pitch identification by western-trained (movable-*Do*) AP musicians. Would hybrid stimuli consisting of western-note labels (C, D, etc.) voiced at a mismatched pitch cause interference? We speculate that this is not the case. Western-trained musicians do not typically use western note labels in pitch voicing, but rather often use movable-*Do* or generic syllables (e.g., *Ah*). Furthermore, the use of non-linguistic coding strategies by at least some of these musicians (e.g., spatial encoding) will likely not be adversely affected by linguistic cues. Clearly, substantial more research is needed to characterize the precise nature of the conditional memory associations used by movable-*Do* trained AP musicians in accessing pitch memory.

While our findings and those of several prior studies [2, 3, 4] support a linguistic component in AP processing, we should note that finding an auditory interference effect *per se* does not necessarily establish a link between pitch *perception* and its putative linguistic code. The Stroop effect in color vision does not specify that color *perception* relies on a linguistic coding strategy, but rather that a language cue can interfere with rapid naming of color. One may thus consider the possibility that a similar kind of language interference underlies identification of the pitch of mismatched pitch-syllable hybrids. Nonetheless, use of a linguistic coding strategy by some AP musicians is consistent with our findings, and is additionally supported by findings from Deutsch *et al.* [3], Levitin [1], and Zatorre [4].

In summary, solfege-trained fixed-*Do* AP individuals showed significant difficulty extracting musical pitch from mismatched linguistic syllables. Lowpass filtering the hybrid stimulus reduced interference, and completely eliminated it when the stimulus was bandpass filtered to within 2 Hz of the note's fundamental frequency. Temporal reversal of hybrid syllables, which preserves their amplitude spectrum but distorts their speech quality, partially reduced interference. Western-trained movable-*Do* AP listeners were immune to pitch-interference from mismatched pitch-syllable hybrids. Our findings support a linguistic-coding strategy for musical pitch retrieval used by fixed-*Do* trained musicians, and a broader form of associative memory used by western-trained AP musicians that may take a variety of forms, including linguistic, emotional, cross-modal, or spatial.

Acknowledgement

We thank Bruce G. Berg, Michael D'Zmura, and Ted Wright for their valuable comments throughout this research. We also thank two anonymous reviewers for their very helpful suggestions. Portions of this work were presented at the annual meetings of the Cognitive Neuroscience Society and the Society for Neuroscience. Work supported by NSF grant BCS0477984.

References

[1] D. J. Levitin: Absolute memory for musical pitch: evidence from the production of learned melodies. Perception and Psychophysics **56** (1994) 414–423.

[2] D. Deutsch, T. Henthorn: Absolute pitch, speech, and tone language: Some experiments and a proposed framework. Music Perception **21** (2004) 339–356.

[3] D. Deutsch, T. Henthorn, E. Marvin, H.-S. Xu: Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. Journal of the Acoustical Society of America **119** (2006) 719–722.

[4] R. J. Zatorre: Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. Nature Neuroscience **6** (2003) 692–695.

[5] R. J. Zatorre, D. W. Perry, C. A. Beckett, C. F. Westbury, A. C. Evans: Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. Proc. Natl. Acad. Sci., USA **95** (1998) 3172–3177.

[6] I. Hsieh, K. Saberi: Dissociation of procedural and semantic memory in absolute-pitch processing. Hearing Research **240** (2008) 73–79.

[7] T. Ohnishi, H. Matsuda, T. Asada, M. Aruga, M. Hirakata, M. Nishikawa, A. Katoh, E. Imabayashi: Functional anatomy of musical perception in musicians. Cerebral Cortex **11** (2001) 754–760.

[8] D. Deutsch: The puzzle of absolute pitch. Current Directions in Psychological Science **11** (2002) 200–204.

[9] S. Baharloo, P. A. Johnston, S. K. Service, J. Gitschier: Absolute pitch: an approach for identification of genetic and nongenetic components. American Journal of Human Genetics **62** (1998) 224–231.

[10] S. Baharloo, S. K. Service, N. Risch, J. Gitschier, N. B. Freimer: Familial aggregation of absolute pitch. American Journal of Human Genetics **67** (2000) 755–758.

[11] K. Itoh, S. Suwazono, H. Arao, K. Miyazaki, T. Nakada: Electrophysiological correlates of absolute pitch and relative pitch. Cerebral Cortex **15** (2005) 760–769.

[12] K. Miyazaki: The auditory Stroop interference and the irrelevant speech/pitch effect: absolute-pitch listeners can't suppress pitch labeling. Paper presented at the 18th International Congress on Acoustics (ICA 2004), Kyoto, Japan, April 4-9, 2004.

[13] P. McLeod, M. I. Posner: Privileged loops from percept to act. – In: Attention and Performance X: Control of Language Processes. H. Bouma, D. G. Bouwhuis (eds.). Erlbaum, Hillsdale, N.J., 1981, 55–66.

[14] I. Hsieh, K. Saberi: Temporal integration in absolute identification of musical pitch. Hearing Research **233** (2007) 108–116.

[15] K. Miyazaki: Absolute pitch identification: effects of timbre and pitch region. Music Perception **7** (1989) 1–14.

[16] K. Miyazaki: The speed of musical pitch identification by absolute-pitch possessors. Music Perception **8** (1990) 177–188.

[17] G. Cohen, M. Martin: Hemisphere differences in an auditory Stroop test. Perception and Psychophysics **17** (1975) 79–83.

[18] L. McClain: Stimulus-response compatibility affects auditory Stroop interference. Perception and Psychophysics **33** (1983) 266–270.

[19] E. J. Green, P. Barber: An auditory Stroop effect with judgments of speaker gender. Perception and Psychophysics **30** (1981) 459–466.

[20] E. J. Green, P. Barber: Interference effects in an auditory stroop task: Congruence and correspondence. Acta Psychologica **53** (1983) 183–194.

[21] J. F. Hamers, W. E. Lambert: Bilingual interdependencies in auditory perception. Journal of Verbal Learning and Verbal Behavior **11** (1972) 303–310.

[22] J. M. Pieters: Ear asymmetry in an auditory spatial Stroop task as a function of handedness. Cortex **17** (1981) 369–379.

[23] P. Walker, S. Smith: Stroop interference based on the synaesthetic qualities of auditory pitch. Perception **13** (1984) 75–81.

[24] P. Walker, S. Smith: Stroop interference based on the multimodal correlates of haptic size and auditory pitch. Perception **14** (1985) 729–736.

[25] P. Walker, S. Smith: The basis of Stroop interference involving the multimodal correlates of auditory pitch. Perception **15** (1986) 491–496.