Hearing Research 240 (2008) 73-79

Contents lists available at ScienceDirect

**Hearing Research** 

journal homepage: www.elsevier.com/locate/heares

# Research paper Dissociation of procedural and semantic memory in absolute-pitch processing

### I-Hui Hsieh, Kourosh Saberi\*

Department of Cognitive Sciences and The Center for Cognitive Neuroscience, University of California, Irvine, CA 92697-5100, United States

#### ARTICLE INFO

Article history: Received 29 July 2007 Received in revised form 12 December 2007 Accepted 4 January 2008 Available online 15 March 2008

*Keywords:* Music Absolute-pitch Memory

#### ABSTRACT

We describe two memory-retrieval systems in absolute-pitch (AP) processing and propose existence of a universal internal pitch template to which subpopulations of musicians selectively gain access through the two systems. In Experiment I, AP and control musicians adjusted the frequency of a pure tone to match the pitch of a visually displayed randomly selected musical note. In Experiment II the same subjects vocally produced within 2 s the pitch associated with a randomly selected musical note label. AP musicians, but not controls, were highly accurate in frequency matching. Surprisingly, both AP and non-AP groups were extremely accurate in voicing the target pitch as determined from an FFT of the recorded voiced notes (i.e.,  $\sigma$  = 0.97, 0.90 semitones, respectively). Spectrogram analysis showed that notes voiced by non-AP musicians are accurate from onset of voicing suggesting that pitch accuracy does not result from an auditory-motor feedback loop. Findings support existence of two memory-retrieval systems for musical pitch: a semantic associative form of memory used by AP musicians, and a more widespread form of procedural memory which allows precise access to internal pitch representations through the vocal-motor system.

© 2008 Elsevier B.V. All rights reserved.

#### 1. Introduction

This study examines the ability of musicians to rapidly produce the pitch of isolated musical notes from long-term memory without feedback or reference to an external acoustic standard. Specifically, we investigate in two experiments the ability of absolute-pitch (AP) and control musicians to retrieve from memory and produce the pitch associated with randomly selected visually displayed musical notes by either vocal production or pure-tone frequency adjustment. The goal is to determine if accuracy of pitch-production through the vocal-motor system is distinct from that of systems that do not engage vocal mechanisms.

Previous studies of AP production have reported conflicting findings depending on task requirements and the mechanism by which pitch is produced. Some have reported large disparities in performance between AP and non-AP musicians, while others have reported that non-AP individuals are more accurate than expected from pitch-identification studies (Petran, 1932; van Krevelen, 1951; Rakowski, 1978; Ross et al., 2004; Zatorre and Beckett, 1989; Siegel, 1974; Wynn, 1972, 1973; Halpern, 1989; Levitin, 1994). No prior study has concurrently examined, within the same subject population, pitch-production accuracy using different production mechanisms. This latter approach is important since it may provide valuable insight into whether different pitch-production mechanisms access internal pitch representations using different retrieval strategies that may explain the apparent contradictions.

We consider here the idea that a universal internal pitch template exists that may be accessed by one of two primary mechanisms: a procedural form of memory retrieval through the vocal-motor system used by most individuals, and a semantic form of retrieval used by AP musicians which draws on associations between pitch categories and symbolic representations (e.g., linguistic, emotional, or spatial). Specifically we examine, in the same subject population of AP and non-AP musicians, two types of pitch-production, one which we suggest invokes a procedural form of pitch memory and one that engages a semantic form of associative memory. In Experiment I, AP and control musicians are allowed either 5 or 30 s to adjust the frequency of a pure tone to match the pitch of a visually displayed target note using a graphical user interface (GUI) slider whose frequency range is randomly shifted on each trial. In Experiment II, musicians vocally produce a target musical note within 2 s. The accuracy with which they voice the target note is determined from the Fourier spectrum of the recorded waveforms. As will be described, a number of surprising findings emerged that point to the existence of a universal finely tuned internal pitch template and a fundamental dissociation between procedural and semantic memory systems in accessing pitch representations.





<sup>\*</sup> Corresponding author. Tel.: +1 949 824 6310; fax: +1 949 824 2307. *E-mail address:* saberi@uci.edu (K. Saberi). URL: http://www.socsci.uci.edu/~saberi/ (K. Saberi).

<sup>0378-5955/\$ -</sup> see front matter @ 2008 Elsevier B.V. All rights reserved. doi:10.1016/j.heares.2008.01.017

#### 2. General materials and methods

#### 2.1. Subjects

Ten trained musicians (5 AP and 5 non-AP) participated in the study. Seven of the subjects were undergraduate piano performance or composition/drama majors in the Music Department at the University of California, Irvine. The other 3 were non-music majors but were highly trained pianists with over 10 years of experience. AP and non-AP groups had average ages of 22 (range 19–27) and 19.2 (range 18–21) years, and had begun formal music training at 5 (range 4-6) and 5.8 (range 4-8) years of age, respectively. AP and non-AP subjects had an average of 14 and 13.2 years experience playing their primary instrument. While subjects typically were trained in more than one instrument, piano was the primary instrument of all 10 subjects. Subjects were recruited either through flyers posted around the Music Department or verbally at music performance classes. Subjects gave their written informed-consent to participate. All protocol were approved by the UC Irvine Institutional Review Board.

#### 2.2. Screening for AP

Subjects were screened for AP ability using protocol similar to those described by Baharloo et al. (1998). Stimuli consisted of 50 pure tones and 50 piano notes presented in two blocks of 50 trials each. A predetermined criterion of 90% accuracy for identifying piano notes and 80% for pure tones was used to qualify a subject as AP (Baharloo et al., 1998; Miyazaki, 1990; Ross et al., 2004; Hsieh and Saberi, 2007). Pure tones were 1 s in duration with 100 ms 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 and Kjær Model 4189), a conditioning amplifier (Nexus, Brüel and Kjær), and a 16-bit A-to-D converter (Creative Sound Blaster Audigy 2ZS). 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). The walls and ceiling of the chamber were covered with 10.2 cm Sonex acoustic foam wedges and the floor with heavy carpeting.

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. A 600 ms burst of white Gaussian noise was presented 600 ms after termination of each stimulus, followed by 1200 ms of silence during which subjects responded. The noise was introduced to reduce iconic (sensory) trace memory cues. Subjects were asked to identify each note by selecting 1 of 12 note labels on GUI 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. (1998). Participants received 1 point for correct identification and 0.5 point for identification within a semitone (e.g., C vs. C#). To qualify as AP, we required a minimum score of 45 points (90%) for piano notes and 40 (80%) for pure tones (maximum = 50 points). Averaged scores across 5 AP subjects were 48.8 ( $\sigma$  = 1.26) for piano notes and 43.8 ( $\sigma$  = 2.36) for pure tones. Non-AP subjects had average scores of 17.0 ( $\sigma$  = 5.79) and 13.2 ( $\sigma$  = 2.93) for piano and pure tones, respectively (chance performance = 8.36 points). The slightly above-chance performance by non-AP musicians is consistent with previous studies (Baharloo et al., 1998; Zatorre et al., 1998; Zatorre, 2003). Restricting scoring to exact identification, AP subjects had an average score or 48.0 ( $\sigma$  = 1.87) or 96% for piano notes and 40.0 ( $\sigma$  = 4.62) or 80% for pure tones. Non-AP subjects

scored 13.8 ( $\sigma$  = 6.97) or 27.6% for piano notes and 7.2 ( $\sigma$  = 3.42) or 14.4% for pure tones (chance performance = 4.1 points or 8.3%).

#### 3. Experiment I: Frequency matching to a target musical note

In Experiment I, we examined the ability of AP and non-AP musicians to adjust, in restricted time intervals of 5 or 30 s, the frequency of a pure tone to match that of a target note selected from a set of 60 musical-note frequencies across 5 octaves.

#### 3.1. Stimuli

Stimuli consisted of pure tones generated and presented through the apparatus described above. Subjects adjusted an unlabeled GUI slider on the monitor to change the stimulus frequency and pressed a pushbutton on the monitor to hear the 1 s tone. The range of frequencies that could be selected using the slider depended on the target note frequency which itself was randomly chosen on each trial. This slider range was kept constant at <sup>3</sup>/<sub>4</sub> of an octave, but was randomly positioned on each trial with respect to the target note frequency. For example, if the target note was 440 Hz (A), the slider could be adjusted in a <sup>3</sup>/<sub>4</sub> octave range around that frequency, with the 440 Hz point positioned at any location along the slider scale on that trial (left edge, right edge, or any point in between). We chose a <sup>3</sup>/<sub>4</sub> octave range, instead of a full octave, to ensure a single solution and to avoid edge-effects resulting in false alarms. The octave from which a target note was chosen was randomly selected from the 2nd to the 6th octaves on each trial.

#### 3.2. Procedure

At the beginning of each trial, one of 12 notes (i.e., Do(C), Do#(C#),...,Si (B)) was randomly selected without replacement and displayed as text. Subjects adjusted the GUI slider to change the tone frequency and pressed a pushbutton after each adjustment to hear the stimulus. The slider resolution was 1% of total slider distance (0.09 semitone resolution). Subjects were allowed either 5 or 30 s, fixed within a run, to make their adjustments on a trial, after which the stimulus could no longer be played. Notes could be played as many times as the subject wanted during the adjustment interval. Typically, they made 4 to 6 adjustments during the 5 s condition and several more in the 30 s condition. When a final adjustment was made, the subject pressed a pushbutton to record results. The slider was reset to the middle position at the beginning of each trial. A total of 10 adjustment sessions were run for each set of 12 notes and for each adjustment interval (30 s and 5 s).

#### 3.3. Results

Fig. 1 shows results of this experiment. The ordinate represents average deviation of slider-adjusted frequencies from target frequency. Error bars are the standard deviation of mean adjustment error across the five subjects. Chance performance, calculated from a 10,000 trial Monte Carlo simulation is 3.0 semitones. Clearly, AP subjects significantly out-perform non-AP subjects. In the 30 s condition, average deviation from target was 0.51 ( $\sigma$  = 0.08) semitones for AP and 2.30 ( $\sigma$  = 0.312) semitones for non-AP subjects. In the 5 s condition, these values were 0.55 (0.07) and 2.52 (0.40) semitones for AP and non-AP subjects, respectively.

AP subjects typically completed their final adjustment well within 5 s, even in the 30 s condition, and reported that it was an easy task. In contrast, non-AP individuals usually experimented with sounds along the entire range of slider frequencies when



**Fig. 1.** Accuracy of pitch production in a frequency-adjustment task. The stimulus to be adjusted was a pure tone. Mean performance is shown for 5 AP subjects and 5 non-AP subjects. The ordinate shows the average deviation of the adjusted frequency from the target frequency. Error bars are the standard deviation of mean adjustment error across the 5 subjects. Left bars shows data from the 5 s adjustment-interval condition and right bars from the 30 s condition. Chance performance derived from a Monte Carlo simulation is 3.0 semitones.

given sufficient time, suggesting a relative-pitch (RP) cue strategy. We expected the AP group to show little bias and variability in adjustments. It was unclear a priori whether non-AP subjects would show bias (i.e., a small  $\sigma$  would be observed if subjects consistently adjusted the stimulus to a wrong but constant frequency). Analysis of response bias however showed that both AP and non-AP groups made random non-systematic errors with no bias. In the 5 and 30 s conditions respectively, the AP group had an error bias of -0.097 and -0.124 semitones, and the non-AP group also had near-zero biases of 0.266 and 0.0219. Individual-subject bias analysis confirmed these group results. Finally, we saw no significant difference between sharp and "white-key" notes contrary to what had previously been reported for AP subjects (Miyazaki, 1989, 1990; Takeuchi and Hulse, 1991). In the 5 s condition, the AP group had average errors of 0.52 ( $\sigma = 0.14$ ) semitones for white-key notes and 0.57 ( $\sigma$  = 0.14) semitones for sharp notes. This difference was not statistically significant (t(4) = 1.42 ns). For this same condition, non-AP subjects had average errors of 2.47 ( $\sigma$  = 1.19) and 2.59 ( $\sigma$  = 0.84) semitones for white-key and sharp notes, respectively, (t(4) = 0.63 ns). Similar results were obtained for the 30 s condition.

## 4. Experiment II: Rapid vocal production of an isolated musical note

In Experiment II we examined accuracy of vocal pitch production and compared it to that of frequency adjustment with the aim of determining, first, if the ability to produce an isolated pitch from long-term memory is unique to AP musicians or a more universal attribute, and second, whether pitch-production accuracy is significantly affected by pitch-production mode. We tested this idea by requiring musicians to rapidly retrieve from memory and vocally produce within 2 s the pitch of randomly selected isolated musical notes.

#### 4.1. Subjects and apparatus

The same 10 subjects from Experiment I participated in Experiment II. The microphone assembly and amplifier described in the screening section for recording piano notes were also used to record vocal production of notes. All recordings were conducted in the IAC chamber described earlier. All vocally produced sounds were digitally recorded at a sampling rate of 44.1 kHz on a Dell workstation.

#### 4.2. Procedure

On each trial the name of a note was displayed as text on the screen (So# (G#)) and subjects were instructed to either hum or sing the note in their preferred octave(s). Subjects pressed a GUI pushbutton to begin recording. To minimize the likelihood of note rehearsal, a strict time limit of 2 s was enforced between displaying the target note name and initiation of recording. If the "record" button was not pressed within 2 s it became inactive on that trial (failed trials were fewer than 2%). Subjects were instructed to either sing or hum the note (subject's choice), and were required to initiate and maintain voicing during the 3 s period after pressing the record button (i.e., recording was terminated after 3 s). Some subjects voiced an arbitrary sound such as the syllable "ah" while others sung using Solfeggio syllables (i.e., Do, Re, Mi). During each 12-trial recording session, the 12 notes were displayed on screen randomly without replacement. Each subject completed two recording sessions. Each recorded voiced-note was Fast-Fourier transformed to determine its fundamental frequency. For a 3 s recording, the fundamental could be determined to a resolution of 0.33 Hz (1/duration).

#### 4.3. Results

Fig. 2 shows FFTs of 12 notes of the musical scale vocally produced by a non-AP subject in random order across trials of a single run. Top and bottom panels show white and black piano-key notes, respectively. This subject sang all musical notes in the 4th octave and thus the voiced harmonics (which start in the 5th octave) are outside the range shown. Target frequency is represented by the green-dashed line and the 1-semitone boundary is shown by the cylindrical regions. Voiced fundamental frequencies ( $F_0$ ) for all but one of the 12 notes produced by this non-AP subject fall within the 1-semitone boundary. Average deviation from target note frequency for this non-AP subject was 0.41 semitones with a standard deviation of 0.19. Surprisingly, all non-AP subjects were able to vocally produce the target pitch of all 12 notes accurately within the 2 s time limit and in the absence of an external acoustic reference.

Fig. 3 shows histograms of response errors in semitone units for both the vocal production (left panels) and slider-adjustment (right) tasks. The slider-adjustment distribution is from the 5 s interval condition (results for the 30 s condition were similar). Data are pooled from 5 subjects in each group. As was the case for the slider-adjustment task, we did not observe bias effects in vocal production for either the AP or non-AP groups. The AP group had a near-zero bias of -0.096 semitones. The non-AP group also had a near-zero bias of -0.144. Note that the error variances in vocal pitch production by AP and non-AP groups were nearly identical ( $\sigma = 0.97$  and 0.90 semitones, respectively).

In the slider-adjustment task, however, the AP and non-AP groups produce markedly different performances. AP subjects had a much lower error variance, approximately four times smaller than that observed for non-AP subjects (0.62 vs. 2.64). The staircase distribution shown in the bottom-right panel of Fig. 3 is that which would be expected from chance performance derived from a Monte Carlo simulation of 1000 runs of 600 trials each (12 notes  $\times$  5 subjects  $\times$  10 sessions). The distribution of chance responses is unimodal with an expected value of zero and a standard deviation of 3.69 semitones. While the distribution of responses from the non-AP group was substantially poorer than that of the



**Fig. 2.** Fast-Fourier transform (FFT) of voiced musical notes by a non-AP subject in a single recording session of 12 randomly selected notes (see text). Each trace shows the FFT of a single note. The voiced harmonics are outside the one-octave range and thus not visible. Top panel shows "white-key" notes and bottom panel sharp/flat notes. Green vertical line represents the target frequency and the cylinders show the one semitone boundaries. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article).

AP group, it was nonetheless better than chance (non-parametric Kolmogorov–Smirnov Z = 3.09, p < 0.01).

age latencies suggest that no significant rehearsal strategy was employed by either group in vocal pitch production.

As instructed, subjects vocally produced the target note in their preferred voice octave. This was typically the 4th octave with an occasional note produced in the 3rd octave (usually by a male subject). The target notes in the slider-adjustment task, however, were sampled from the 2nd through the 6th octaves. For non-AP subjects, we reanalyzed the 5 s slider-adjustment data for target notes restricted to the 4th octave. This average deviation was 2.50 semitones, close to that of the entire range of target note frequencies, and not significantly different than that for other octaves; t(4)0.08, p = 0.94. In addition, to determine if non-AP subjects used part of the recording interval to rehearse the target note before initiation of voicing, we compared AP and non-AP voicing latencies defined as the time interval between pressing the record key and onset of voicing (note that subjects had 2s to press the record key after the note was visually displayed). The average voicing latency was 117 ms ( $\sigma$  = 147) for the AP group and 53 ms ( $\sigma$  = 51) for the non-AP group. This difference was not statistically significant; t(8) = 0.924, p = 0.383. Subjects from both groups often initiated voicing quite rapidly at or just prior to pressing the record button, resulting in a large proportion of zero latencies. These short aver-

#### 5. Discussion

The finding that non-AP musicians are highly accurate in vocal production of an isolated musical note was unexpected given their inaccuracy in frequency matching. We propose that in the broader population (of at least musicians) an internal pitch template must exist with narrowly tuned categories to which non-AP subjects gain access through a procedural vocal-motor form of memory retrieval. Such a universal template has been speculated on in recent years from the relative accuracy with which non-musicians sing familiar songs (notwithstanding that songs contain relative spectrotemporal and context cues; Levitin, 1994; Levitin and Rogers, 2005). Additional support for a universal template comes from the work of Deutsch (2002) and Deutsch et al. (2006), who have demonstrated that speakers of tonal languages (e.g., Vietnamese or Mandarin) are remarkably accurate in repeated reproduction of the pitch of tonal words.

The precise mechanism for vocal access to internal pitch representations is unclear. One possible mechanism might be an acous-



**Fig. 3.** Comparison of error distributions in the vocal-production (left panels) and frequency-adjustment (right) tasks. Top panels show data from AP and bottom panels from non-AP subjects. The staircase in the bottom-right panel shows the distribution of chance responses derived from a Monte Carlo simulation. The frequency-adjustment data are from the 5 s response-interval condition.

tic sensorimotor feedback loop that allows real-time recalibration of one's own vocal pitch via auditory feedback in the initial stages of pitch-production. Two observations argue against this explanation. First, non-AP subjects accurately produce the target  $F_0$  immediately from the onset of voicing with no significant frequency drift (i.e., stable to within one semitone). Top panel of Fig. 4 shows a sample spectrogram from a sliding 50 ms temporal window for the note "*Mi*" voiced by a non-AP musician. The red horizontal line represents the target frequency of 329.6 Hz. The bottom panel shows the voiced note's FFT. There is no significant shift in frequency during vocal production. The syllable begins with a brief consonant where the spectral splatter is observed followed by the steady state vowel. Nearly all voiced notes examined had spectrograms similar to that shown in this figure.

Further evidence against a feedback-loop explanation comes from measurements of pitch-production accuracy in the absence of acoustic feedback. If an auditory-motor feedback loop provides cues for real-time vocal calibration of musical pitch then eliminating feedback should reduce accuracy. Previous studies of this type have used masking noise during vocal pitch production as a method of eliminating auditory feedback (Ward and Burns, 1978). We, however, have found that even intense masking noise cannot effectively eliminate auditory feedback during voicing since subjects always clearly hear their own voice through bone and tissue conduction. To investigate the effects of eliminating auditory feedback we recruited a deaf cochlear-implanted (CI) musician. The intent, of course, was not a full-scale study of CI musicians, but to simply verify our observations. This musician had become deaf from a genetic disorder in his early 20 s and was deaf for over 30 years. He had received his cochlear implant approximately 1 year ago. With the implant turned "off" he was completely deaf and could not hear his own voice. With the implant "on" he could easily understand speech without lip reading. We tested his ability to produce the pitch of randomly selected musical notes with the cochlear implant turned "off" or "on". Results are shown in Fig. 5. Clearly, this musician can accurately voice the target notes whether the implant is "on" or "off". This finding, together with the accuracy with which non-AP subjects produce pitch from onset of voicing, support the idea that accurate vocal pitch production does not result from real-time auditory calibration of vocal-motor output, and may instead be based on a more intrinsic motor access to pitch representations, the mechanisms of which are not clear at this time. That real-time feedback does not appear to be necessary for accurate vocal pitch production, of course does not mean that long-term absence of feedback or auditory interference from other sounds cannot distort accuracy of vocal pitch (Ward and Burns, 1978; Waldstein, 1990).

In addition to this procedural vocal form of access to pitch memory, our data, as well as other research (Deutsch, 2002; Zatorre, 2003) suggest that AP musicians (as categorized by conventional standards) use a form of semantic associative memory in pitch retrieval and identification. This type of semantic association may take a variety of forms such as associations between pitch and linguistic, emotional, or spatial representations. We interviewed our AP subjects to gain better insight into their strategies for pitch identification. While these descriptions are subjective, they do provide valuable insights into pitch-retrieval mechanisms. All our solfege-trained AP musicians reported that they detect a linguistic quality in the pitch of musical notes. A pure tone at 440 Hz perceptually sounds like the syllable "La". Our western-trained AP musicians reported different and highly individualized forms of associations. One AP musician reported linguistic, emotional, and



**Fig. 4.** Top: Sample spectrogram with a running temporal-integration window of 50 ms from a non-AP subject voicing the note '*Mi*' (329.6 Hz). The target frequency is shown as the red line. Bottom panel shows the note's Fourier spectrum. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article).

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 AP subject described a spatial strategy in which he first rapidly 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 noted 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.

Non-AP subjects on the other hand reported no particular type of strategy in pitch identification; most reported that they felt they were guessing. One non-AP subject reported that she believes she can often accurately identify the note 'A' (La) and therefore tries to use that note as a referent to judging the pitch of other notes. We analyzed this subject's data from the slider-adjustment task and found that her ability to identify 'A' was not significantly different than that of other notes (t(9) = 0.38 ns).

In summary, while only AP subjects were accurate in adjusting a tone frequency to match its pitch to a target note, all subjects were highly accurate in vocally producing the pitch of isolated and randomly selected musical notes. Furthermore, accuracy of vocal production did not appear to depend on real-time auditory calibration of vocal output. Our findings support existence of a common and possibly widespread internal pitch template and two distinct mechanisms for pitch retrieval, a procedural form of vocal-motor access employed by all subjects and a semantic associative form of memory retrieval used by AP musicians. That there are two forms of memory retrieval does not necessarily mean that the same pitch template is accessed by both retrieval mechanisms. The memory systems themselves might be distinct in that pitch memory accessed by vocal mechanisms may be stored in motor areas, separate from that accessed by semantic conditional associations. Finally, we should qualify that the non-AP musicians used in our study are clearly not representative of the broader population of non-musicians. They are highly trained pianists and as such one might consider whether spatial learning of notes on a fixedpitch keyboard or vocal production of visually displayed notes somehow facilitated their performance. Nonetheless, our findings, together with those who have shown better than expected accuracy by non-musicians in vocally reproducing the melody of familiar songs, lend support to the idea that a more common and possibly widespread form of internal pitch representation exists that may be accessed by the vocal system but not aural feedback mechanisms.



Fig. 5. Correlation between voiced fundamental frequency and target frequency during vocal production of musical notes in random order by a deaf cochlear-implanted musician (left panel) and with the implant turned off (right panel). Different symbols represent separate recording sessions. Solid line represents perfect performance.

#### Acknowledgements

We thank Bruce G. Berg, Michael D'Zmura, and Ted Wright for their valuable suggestions throughout this research. We also thank two anonymous reviewers for their helpful comments. 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

- Baharloo, S., Johnston, P.A., Service, S.K., Gitschier, J., 1998. Absolute pitch: an approach for identification of genetic and nongenetic components. American Journal of Human Genetics 62, 224–231.
- Deutsch, D., 2002. The puzzle of absolute pitch. Current Directions in Psychological Science 11, 200–204.
- Deutsch, D., Henthorn, T., Marvin, E., Xu, H-S., 2006. 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, 719–722.
- Halpern, A.R., 1989. Memory for the absolute pitch of familiar songs. Memory and Cognition 17, 572–581.
- Hsieh, I., Saberi, K., 2007. Temporal integration in absolute identification of musical pitch. Hearing Research 233, 108–116.
- Levitin, D.J., 1994. Absolute memory for musical pitch: evidence from the production of learned melodies. Perception and Psychophysics 56, 414–423.
- Levitin, D.J., Rogers, S.E., 2005. Absolute pitch: perception, coding, and controversies. Trends in Cognitive Sciences 9, 26–33.
- Miyazaki, K., 1989. Absolute pitch identification: effects of timbre and pitch region. Music Perception 7, 1–14.
- Miyazaki, K., 1990. The speed of musical pitch identification by absolute-pitch possessors. Music Perception 8, 177–188.

- Petran, L.A., 1932. An experimental study of pitch recognition. Psychological Monographs 42 (6), 1–120.
- Rakowski, A., 1978. Investigations of absolute pitch. In: Asmus, E.P., Jr. (Ed.), Proceedings of the Research Symposium on the Psychology and Acoustics of Music. University of Kansas, Division of Continuing Education, Lawrence, pp. 45–57.
- Ross, D.A., Olson, I.R., Marks, L.E., Gore, J.C., 2004. A nonmusical paradigm for identifying absolute pitch possessors. Journal of the Acoustical Society of America 116, 1793–1799.
- Siegel, J.A., 1974. Sensory and verbal coding strategies in subjects with absolute pitch. Journal of Experimental Psychology 103, 37–44.
- Takeuchi, A.H., Hulse, S.H., 1991. Absolute pitch judgment of black- and white-key pitches. Music Perception 9, 27–46.
- van Krevelen, A., 1951. The ability to make absolute judgments of pitch. Journal of Experimental Psychology 42, 207–215.
- Waldstein, R.S., 1990. Effects of postlingual deafness on speech production: Implications for the role of auditory feedback. Journal of the Acoustical Society of America 88, 2099–2114.
- Ward, W.D., Burns, E.M., 1978. Singing without auditory feedback. Journal of Research in Singing and Applied Vocal Pedagogy 1, 24–44.
- Wynn, V.T., 1972. Measurements of small variations in "absolute" pitch. Journal of Physiology 220, 627–637.
- Wynn, V.T., 1973. Absolute pitch in humans: its variations and possible connections with other known rhythmic phenomena. In: Kerkut, G.A., Phillis, J.W. (Eds.), Progress in Neurobiology, vol. 1, Pt. 2. Pergamon Press, Elmsford, NY, pp. 111– 149.
- Zatorre, R.J., 2003. Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. Nature Neuroscience 6, 692–695.
- Zatorre, R.J., Beckett, C., 1989. Multiple coding strategies in the retention of musical tones by possessors of absolute pitch. Memory and Cognition 17, 582–589.
- Zatorre, R.J., Perry, D.W., Beckett, C.A., Westbury, C.F., Evans, A.C., 1998. Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. Proceedings of the National Academy of Sciences, USA 95, 3172– 3177.