



Enhancement of speech-relevant auditory acuity in absolute pitch possessors

Nobuo Masataka*

Primate Research Institute, Kyoto University, Inuyama, Japan

Edited by:

Janet F. Werker, The University of British Columbia, Canada

Reviewed by:

Hua Shu, Beijing Normal University, China

Yang Zhang, University of Minnesota, USA

***Correspondence:**

Nobuo Masataka, Primate Research Institute, Kyoto University, Kanrin, Inuyama, Aichi 484-8506, Japan.
e-mail: masataka@pri.kyoto-u.ac.jp

Absolute pitch (AP) is the ability to identify the frequency or musical name of a specific tone, or to identify a tone without comparing it with any objective reference tone. While AP has recently been shown to be associated with morphological changes and neurophysiological adaptations in the planum temporale, a cortical area in the brain involved in speech perception processes, no behavioral evidence of speech-relevant auditory acuity in any AP possessors has hitherto been reported. In order to seek such evidence, in the present study, 15 professional musicians with AP and 14 without AP, all of whom had acquired Japanese as their first language, were asked to identify isolated Japanese syllables as quickly as possible after these syllables were presented auditorily. When the mean latency to the syllable identification was compared, it was significantly shorter in AP possessors than in non-AP possessors whether the presented syllables were those used as Japanese labels representing the 7 tones constituting an octave or not. The latency to hear the stimuli *per se* did not differ according to whether the participants were AP possessors or not. The results indicate the possibility that possessing AP provides one with extraordinarily enhanced acuity to individual syllables *per se* as fundamental units of a segmented word in the speech stream.

Keywords: absolute pitch, syllable perception, music, planum temporale

INTRODUCTION

Absolute pitch (AP) is the ability to identify the frequency or musical name of a specific tone, or to identify a tone without comparing it with any objective reference tone (Martin and Perry, 1999; Levitin and Rogers, 2005; Vanzella and Schnellenberg, 2010). Although whether the extraordinary ability of AP is genetically determined or develops dependently on environmental variables still remains a controversial issue (Levitin and Zatorre, 2003; Drayna, 2007), there is a general consensus that musical training in childhood is important for its acquisition (Zatorre, 2003). Recently, neurophysiological effects of such musical experiences have been investigated using neuroimaging methods. Those investigations have consistently found that while the right hemisphere is in general important for musical processing, increasing musical sophistication causes a shift of musical processing from the right to the left hemisphere (Elbert et al., 1995; Schlaug et al., 1995; Pantev et al., 1998; Tervaniemi et al., 2000). The notion has been confirmed regarding AP by various structural observations of morphological changes in the cortical regions of the planum temporale (PT) in musicians with AP (Keenan et al., 2001; Luders et al., 2004; Schneider et al., 2005; Wilson et al., 2008). Moreover, more recently, the stronger left PT activation has been recorded in AP possessors than in non-AP possessors (including professional musicians; Hirata et al., 1999; Ohnishi et al., 2001; Gaab et al., 2006; Wu et al., 2008; Oechslin et al., 2010).

The left PT is known as Wernicke's area, and is related to language comprehension. The human PT is a roughly triangular region of the superior temporal plane located posterior to the primary auditory field (Steinmetz et al., 1991). It is, on the average, larger in the left hemisphere, suggesting that it may play a specialized

role in language and language lateralization. Why is the left PT involved in music perception in AP possessors, particularly in those of trained musicians? Do they employ a common strategy in music perception and language comprehension? So far, no behavioral evidence of speech-relevant auditory acuity in any AP possessors has been reported. In order to pursue this issue, the present study was conducted.

Here, I have compared the performance of identifying isolated syllables between musicians with AP and musicians without AP. I hypothesized that since AP possessors have special expertise in identifying the tonal label of a specific tone, their sense of identification would be keener than that of non-AP possessors whether the isolated syllables they were asked to identify were the names of musical notes or not.

MATERIALS AND METHODS

PARTICIPANTS

Fifteen professional musicians with AP (referred to below as "AP-possessors": 3 males and 12 females; mean age = 23 years, SD = 3.3; mean practice years = 15.4, SD = 4.3; mean age practice began = 4.2, SD = 3.0), and 14 professional musicians without AP (referred to below as "non-AP possessors": 4 males and 10 females; mean age = 23 years, SD = 3.8; mean practice age = 13.9, SD = 6.2, mean age practice began = 9.1, SD = 5.2) participated in the present study. All of them were born in Japan, had been educated in the Japanese formal education system from the beginning of the entrance into kindergarden at least until the completion of graduation from senior high school, and had acquired Japanese as their first language. None of them reported any hearing impairment.

Both AP and Non-AP possessors as participants were selected preliminarily with an in-house test which was conducted prior to the present experiment: they heard 108 pure sine wave tones, presented in pseudorandomized order, which ranged from A3 (tuning: A4 = 440 Hz) to A5, with each tone being presented three times. Each tone of the AP test had a duration of 1 s, with a 4-s interstimulus interval. During the intervals, the participants heard brown noise. They had to write down the tonal label immediately after hearing the accordant tone. The whole test unit and its components were created with Adobe Audition 1.5. The accuracy was evaluated by counting correct answers and the semitone errors were taken as incorrect to increase the discriminatory power. The participants were not asked to identify the adjacent octaves of the presented tones because for AP it is a most notable prerequisite to identify the correct chroma. In all, AP possessors were those whose accuracy scores were above 80% (mean = 85.8; SD = 6.6) whereas non-AP possessors were those whose scores were below 10% (mean = 7.2; SD = 3.9).

PROCEDURE

During the experiment as well as during the preliminary screening test for AP, each participant was seated in an attenuation chamber and wore a headphone. While a pure tone was presented using a notebook computer in a screening test, a syllable was chosen for a presented stimulus from a total of the 111 syllables that constitute the Japanese language. A total of 100 isolated syllables, each of which had a duration of 200 ms, were presented to a given participant consecutively with an interstimulus interval being 5 s of silence in a given presentation session, and in all, two such sessions were conducted for each participant. In the first session, the participant was asked to press a key which was located on a table near the participant as quickly as possible when identifying what syllable was the stimulus and to answer it orally (referred to below as “the identification session”). In the second session, the participant was asked to press the key as quickly as possible when hearing the presented sound (referred to below as “the hearing session”). Prior to the identification session, the participant was also instructed not to press the key before recognizing the presented syllable, and that was actually confirmed in each participant by an interview undertaken after the completion of the entire experiment.

As the presented stimuli in each session, the 111 syllables were operationally classified into two categories; seven were those that are used as Japanese tonal labels for seven musical notes constituting an octave, i.e., do (C), re (D), mi (E), fa (F), so (G), ra (A), si (B) (referred to below as “solfege syllables”), and the remaining 104 were those that are not used as note-names (referred to below as “non-solfege syllables”). The 100 isolated syllables presented to each participant in a given session comprised 50 solfege syllables and 50 non-solfege syllables. For each of the two categories, the presented stimuli were randomly chosen. Moreover, all of the stimuli were presented randomly to each participant regarding whether they were solfege syllables or non-solfege syllables.

As a behavioral measure, in both sessions, the interval between the onset of each stimulus presentation and the onset of the subsequent pressing of the button was used. The mean latency to the answer was computed for each participant in each of the two sessions separately with regard to solfege syllables and to non-solfege syllables.

RESULTS

Figure 1 shows the mean latency to press the button of AP possessors and of non-AP possessors in the identification session when the presented stimuli were solfege syllables as well as non-solfege syllables. Throughout the entire experiment, no identification errors were recorded and all the participants answered all the presented syllables correctly. Nonetheless, 2 (AP possessor versus non-AP possessor) \times 2 (solfege syllable versus non-solfege syllable) analysis of variance (ANOVA) revealed a significant main effect [$F(1,27) = 12.176, p = 0.002$], and the mean latency to the presented stimulus was significantly shorter in the AP possessors than in non-AP possessors. The score was not significantly different whether the stimulus was a solfege syllable or a non-solfege syllable [$F(1,27) = 0.012, p = 0.912$]. Interaction between the two main factors was not significant, either [$F(1,27) = 0.042, p = 0.839$].

Figure 2 shows the mean latency to press the button of AP possessors and of non-AP possessors in the hearing session when the presented stimuli were solfege syllables as well as non-solfege syllables.

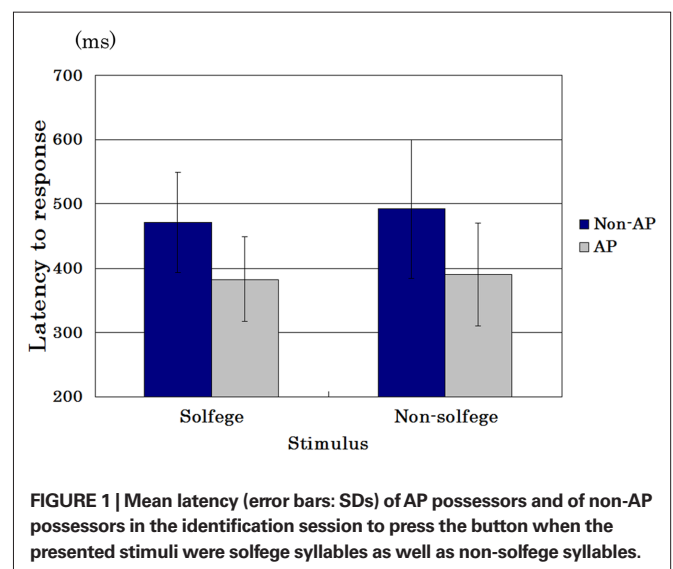


FIGURE 1 | Mean latency (error bars: SDs) of AP possessors and of non-AP possessors in the identification session to press the button when the presented stimuli were solfege syllables as well as non-solfege syllables.

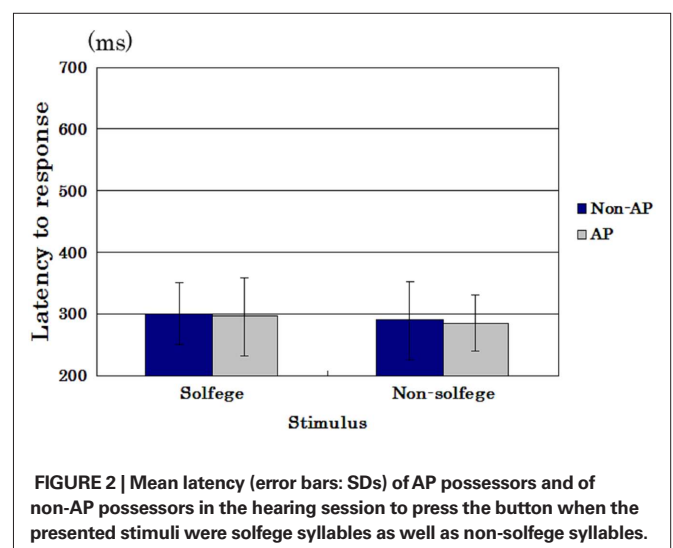


FIGURE 2 | Mean latency (error bars: SDs) of AP possessors and of non-AP possessors in the hearing session to press the button when the presented stimuli were solfege syllables as well as non-solfege syllables.

ANOVA revealed no significant main effects, and the average latency to the presented stimulus was not different between the AP possessors and the non-AP possessors [$F(1,27) = 0.392, p = 0.536$]. The score was not different whether the stimulus was a solfege syllable or a non-solfege syllable [$F(1,27) = 0.963, p = 0.334$]. Interaction between the factors was not significant, either [$F(1,27) = 1.359, p = 0.254$].

DISCUSSION

Previous research demonstrated that the basic auditory capability does not differ between AP possessors and non-AP possessors (Fujisaki and Kashino, 2002). The results of the hearing session in the present study are consistent with that conclusion. Nonetheless, recent neuroimaging studies have provided suggestive evidence for a strong influence of the pitch-processing expertise of AP possessors on their speech perception (Oechslin et al., 2010). The plausibility of this notion was tested in the present experiment, which actually presented suggestive evidence for such a link between musical expertise and speech information processing, and the results of the identification session in the present experiment revealed the fact that AP possessors are significantly superior in basic speech processing to non-AP possessors. Namely, AP possessors were able to identify a given isolated syllable chosen from their first language significantly more rapidly than non-AP possessors could, whether the syllable was one used as musical note or not.

Human infants are born with the predispositional capability to distinguish all the sounds in all of the world's languages (Kuhl, 2000). By the end of their first year, however, they are on their way to perceiving particularly well the sounds that are important for their native languages (usually around 40 for a given language) whereas their capability to distinguish foreign speech sounds has decreased (Kuhl, 2003). Japanese infants, for example, initially perceive separate sounds for "r" and "l" (as in the words "road" and "load") but lose the ability to hear this "foreign" distinction as they mature and become more adept at recognizing Japanese speech sounds.

Meanwhile, the first recognizable speech infants produce by themselves comprises a single word or what may appear to be a phrase, though at this stage, they are not aware that the words they produce have constituent elements. They do not understand the notion of word or lexical meaning, either. The question that then arises relates to the segmentation problem: how do children discover the structural components of the fluent speech stream without knowing the identity of the target elements? In fact, the infant's task of learning its native language is a daunting one because, unlike written language, spoken language has no obvious markers that indicate the boundaries between words.

The commonest answer to the above question is that language learners use supra-segmental cues to locate boundaries in the speech stream (Werker and Voutoumantos, 2000; Falk, 2009). In particular, it is a well-known fact that long before they can speak, infants are sensitive to the frequencies at which combinations of syllables occur and how they differ within and across word boundaries (Wermke and Mende, 2006). The study used the example of the phrase "pretty baby" and noted that among English words, the likelihood of "ty" following "pre" was higher than the likelihood that "bay" would follow "ty." Thus, with enough repetition, participating 7-month-old infants began to understand that "pretty" was potentially a word, even before they knew what it meant. Prosodic cues embedded

in language also help. In conversational English, for example, the majority of words are stressed on their first syllables, as in the words "monkey" and "jungle." This predominantly strong-weak pattern is reversed in some languages. By the time an English-learning infant is seven and half a months old, he spontaneously perceives words that reflect the strong-weak pattern, but not the weak-strong pattern. Consequently, when infants hear "guitar is" they perceive "taris" as a unit because it begins with a stressed syllable.

Regarding the developmental milestones of vocalizations with segmental features, infants become able to produce them first as canonical babbling around 8 months of age. This period coincides with the period when they become able to segment words from fluent speech (Masataka, 2003). Taken together, these facts indicate that producing a limited, sequentially organized phonemic segment entails extracting words from some portion of the speech stream, which in turn is segmented according to prosodic and more global supra-segmental characteristics. Extracting words enables infants to code them, which in turn enables the infants to recognize them as familiar words (Masataka, 2007). While this should be the usual process by which one's sensitivity to word units in fluent speech typically develops, the results of the present experiment indicate that possessing AP provides a person with extraordinarily enhanced acuity to individual syllables *per se* as fundamental units of a segmented word in the speech stream. Presumably this keen sensitivity allows AP possessors to accomplish speech segmentation with less help of supra-segmental information of the speech compared to non-AP possessors. As it were, AP would enable its possessors to discover the structural components of speech stream, and consequently to rely more on identifying purely isolated linguistic elements involved in the stream than on identifying paralinguistic ones. The development of such a capability is closely related to neurological observations on PT changes that should be a consequence of early special musical experiences.

Admittedly, it is not completely clear how such performance of the isolated syllable identification in terms of reaction time as reported here could be speech segmentation ability because the AP possessors were not examined with regard to their superiority in other non-speech classification processing. Also, the fact should be noticeable that all the participants were native speakers of Japanese, a language having pitch contrast (e.g., Tsujimune, 2007). Thus, information of the pitch of a word has some importance for the identification of the word as it is. In order to link the current data to speech segmentation, the performance of syllable recognition in a context should be compared between AP and non-AP possessors who are, preferably, native speakers of a language that does not have tone distinction or pitch contrast by presenting meaningful phrases or strings of nonsense syllables in which the target syllable can be in the initial, middle or final position. These are the next issues to be investigated in the near future.

ACKNOWLEDGMENTS

The present study was supported by a grant-in-aid from the Ministry of Education, Science, Sports and Culture, Japanese Government (#20243034) as well as by Global COE Research Program (A06 to Kyoto University). I am grateful to Naoko Watanabe for her assistance when conducting the experiment and to Elizabeth Nakajima for her reading the earlier version of the manuscript and correcting its English.

REFERENCES

- Drayna, D. T. (2007). Absolute pitch: a special group of ears. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14549–14550.
- Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–307.
- Falk, D. (2009). *Finding our Tongues*. New York: Basic Books.
- Fujisaki, W., and Kashino, M. (2002). The basic hearing abilities of absolute pitch possessors. *Acoust. Sci. Technol.* 23, 77–83.
- Gaab, N., Schulze, K., Ozdemir, E., and Schlaug, G. (2006). Neural correlates of absolute pitch differ between blind and sighted musicians. *Neuroreport* 17, 1853–1857.
- Hirata, Y., Kuriki, S., and Pantev, C. (1999). Musicians with absolute pitch show distinct neural activities in the auditory cortex. *Neuroreport* 10, 999–1002.
- Keenan, J. P., Thangaraj, V., Halpern, A. R., and Schlaug, G. (2001). Absolute pitch and planum temporale. *Neuroimage* 14, 1402–1408.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11850–11857.
- Kuhl, P. K. (2003). Human speech and birdsong: communication and the social brain. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9645–9646.
- Levitin, D. J., and Rogers, S. E. (2005). Absolute pitch perception, coding, and controversies. *Trends Cogn. Sci.* 9, 26–33.
- Levitin, D. J., and Zatorre, R. J. (2003). On the nature of early music training and absolute pitch: a reply to Brown, Sachs, Cammuso, and Falstein. *Music Percept.* 21, 105–110.
- Luders, E., Gaser, C., Jancke, L., and Schlaug, G. (2004). A voxel-based approach to gray matter asymmetries. *Neuroimage* 22, 656–664.
- Martin, O. S. M., and Perry, D. W. (1999). *The Psychology of Music*, 2nd Edn. New York: Academic Press.
- Masataka, N. (2003). *The Onset of Language*. Cambridge: Cambridge University Press.
- Masataka, N. (2007). Music, evolution and language. *Dev. Sci.* 10, 35–39.
- Oechslin, M. S., Meyer, M., and Jancke, L. (2010). Absolute pitch - functional evidence of speech-relevant auditory acuity. *Cereb. Cortex* 20, 447–455.
- Ohnishi, T., Matsuda, H., Asada, T., Atuga, M., Hirakata, M., Nishikawa, M., Katoh, A., and Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cereb. Cortex* 11, 754–760.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., and Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature* 392, 811–814.
- Schlaug, G., Jancke, L., Huang, Y., and Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science* 267, 699–701.
- Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., Dosch, H. G., Bleeck, S., Stippich, C., and Rupp, A. (2005). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat. Neurosci.* 8, 1241–1247.
- Steinmetz, H., Volkman, J., Jancke, L., and Freud, H. J. (1991). Anatomical left-right asymmetry of language related temporal cortex is different in left- and right-handers. *Ann. Neurol.* 29, 315–319.
- Tervaniemi, M., Medvedev, S. V., Alho, K., Pakhomov, S. V., Roudas, M. S., Van Zuijlen, H., and Naatanen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Hum. Brain Mapp.* 10, 74–79.
- Tsujimune, N. (2007). *An Introduction to Japanese Linguistics*. London: Blackwell.
- Vanzella, P., and Schnellenberg, E. G. (2010). Absolute pitch: effects of timbre on note-naming ability. *PLoS ONE* 5, e15449. doi: 10.1371/journal.pone.0015449
- Werker, J. F., and Voutoumantos, A. (2000). Who's got rhythm? *Science* 288, 280–281.
- Wermke, K., and Mende, W. (2006). Melody as a primordial legacy from early roots of language. *Behav. Brain Sci.* 29, 300.
- Wilson, S. J., Lusher, D., Wang, C. Y., Dudgeon, P., and Reutens, D. C. (2008). The neurocognitive components of pitch processing: insights from absolute pitch. *Cereb. Cortex* 19, 624–732.
- Wu, C., Kirk, I. J., Hamm, J. P., and Lim, V. K. (2008). The neural networks involved in pitch labeling of absolute pitch musicians. *Neuroreport* 19, 851–854.
- Zatorre, R. J. (2003). Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. *Nat. Neurosci.* 6, 692–695.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 March 2011; paper pending published: 17 April 2011; accepted: 09 May 2011; published online: 20 May 2011.

Citation: Masataka N (2011) Enhancement of speech-relevant auditory acuity in absolute pitch possessors. *Front. Psychology* 2:101. doi: 10.3389/fpsyg.2011.00101

This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Masataka. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.