ARCHIVES OF ACOUSTICS
33, 4, 441–446 (2008)

AUDITORY INFORMATION PROCESSING FOR SOUND PERCEPTION AND VOCALIZATION IN CASE OF TRITONE PARADOX

Honorata HAFKE

Adam Mickiewicz University
Institute of Acoustics
Umultowska 85, 61-614 Poznań, Poland
e-mail: honorata@spl.ia.amu.edu.pl

(*received June 15, 2008; accepted October 15, 2008*)

The aim of this paper is to answer the question whether "perception-action" dissociation, which is well-documented in vision, may also be found in auditory information processing in case of the tritone paradox. Two experiments were conducted. Sounds with the ambiguous pitch were utilized. They were precisely classified as a note (eg. C, C#, F#), but equivocal as to octave. In the experiment A, participants were asked to aurally estimate pitch shifts of sounds divided by the interval of tritone. In the experiment B, participants were asked to reproduce a pair of tritones, by singing the vowel "a". Results of the two experiments were compared, in order to determine if task specification – perceptive (experiment A) or motor (experiment B) – can cause dissimilar results.

Keywords: tritone paradox, voice fundamental frequency.

1. Introduction

Milner and Goodale [7] postulate two separate visual systems in the human brain. Evidence from both humans and other primates has shown that this distinction between vision for perception and vision for action is reflected in the organization of the visual pathways in the cerebral cortex of primates. Each stream uses visual information in a different way. The function of the ventral stream is the creation of an internal model or perception of the external world. The second one guides the control of object-directed action. In recent years, researchers have attempted to find similar dissociation between action and perception in human audition [9, 10]. They tried to find the exact auditory analogue to the system of visual control of action. A different approach [8] suggests, that in case of audition, the dorsal stream is associated with controlling the voice fundamental frequency. Hafke [8] experiments show that vocal utterances are controlled nonconsciously. Trained singers were asked to produce vowel sounds into a microphone. The sound that each singer produced was fed back to their ears via headphones. Two seconds after the sound production had begun, the auditory feedback was shifted in pitch by a certain degree. After each trial, participants reported whether they were aware of a pitch change or not. It was found that even though the participants were unaware of subtle pitch changes, the fundamental frequency of their vowel production was found to shift slightly in the direction opposite to the pitch shift.

The present experiment is a direct continuation of those studies. During the previous studies, in order to observe the dissociation between perception and motor control, a subliminal experimental situation was created. In a psychoacoustic experiment, values below the perceptual threshold were used. The current experiment employs an acoustic paradox, which tricks the perception stream in a certain way. If both streams are used to extract information concerning sound pitch in different way, it seems possible that the experimental environment could have an influence on the results.

Two experiments were conducted. The tritone paradox described primarily by Deutsch [2–6] had been used. Sounds with the ambiguous pitch were utilized. They were precisely classified as a given note by subjects (eg. C, C#, F#), but equivocal as to octave. In the experiment A, participants were asked to aurally estimate pitch shifts of sounds divided by the interval of tritone. In the experiment B, participants were asked to reproduce a pair of tritones, by singing the vowel “a”. Results of the two experiments were compared, in order to determine if task specification – perceptive (experiment A) or motor (experiment B) – can cause dissimilar results.

2. Method

2.1. Subjects

Twelve adults between 21 and 28 years of age participated in this study. All listeners were qualified as having normal hearing, which was defined as the audiometric threshold of 20 dB HL, or better, for a range from 250 to 8000 Hz (ANSI, 1996). They reported no neurological defects and had no speech or voice disorders. Half of them were musicians.

2.2. Stimuli and procedure

Sounds used in the experiment were identical to those utilized by Deutsch [3–5]. Each sound consisted of 6 sinusoids that were related by octaves, and their amplitudes were determined by a fixed, bell-shaped spectral envelope. All tones were 500 msec in duration, and there were no gaps between tones within a pair. Envelope was placed at four different positions along the spectrum, with peaks at C4 (262 Hz), F# (4370 Hz), C5 (523 Hz), F#5 (740 Hz). Twelve tone pairs were generated under each of the four spectral envelopes, corresponding to the pitch-class pairings C–F#, C#–G, D–G#, D#–A, E–A#, F–B, F#–C, G–C#, G#–D, A–D#, A#–E, and H–F.

The tritone pairs were presented at a level of 84 dB SPL to the subject over Sennheiser HD 600 headphones. Subjects’ voices were recorded with a Shure SM 58 microphone.

The participants were tested in soundproof booths. The subjects were seated in a sound-treated room. Six of them began with the perception part of the experiment (3 musicians, 3 non-musicians), the rest with the motor part.

Before the experiment, every participant had to take part in a test verifying their perceptive and vocal abilities (one person was rejected after the test). The test was identical to experiments A and B, with the exception of tritones being composed of the simple tonal equivalents of musical notes. Each pair of sounds was repeated three times. Only participants with 100% accuracy in hearing and reproducing the direction of pitch shifts were qualified to the main experiments.

In experiment A an organized pair of tones was presented and the subject had to judge whether the second sound in pair had higher or lower frequency then the first one. In experiment B subjects were asked to reproduce the sounds by singing vowel “a”.

In both experiments the same list sets of tritone pairs was used. Six different lists which consisted of 16 blocks were prepared. Each block contained of 12 randomly ordered tritone pairs generated under one of the spectral envelopes, similarly to experiment by DEUTSCH [3–5]. Trials within a single block were separated by 5-sec pauses, and blocks were separated by 20-sec pauses. During a single day, every person listened to or reproduced two lists, with a 20 min. break. Every pair of tritones was judged and reproduced 24 times.

3. Results and analysis

In experiment A percentage of responses stating that the second tone in a pair had lower frequency than the first one (“down” responses) was calculated and plotted as a function of the pitch class of the first tone in a pair.

In experiment B the voice fundamental frequency from recorded vocalization during each trial was calculated using an algorithm incorporated in the Praat software [1]. Then, for each trial, direction of the fundamental frequency change was determined. Consequently, as in experiment A, a percentage of responses stating that the second tone in a pair had lower frequency than the first one was calculated and plotted as a function of the pitch class of the first tone in a pair.

The averaged results for experiment A are shown in Fig. 1. For three of the envelope sets (C5, F#4, F#5) visible maxima and minima are present, their position, however, differs between envelope sets. Envelope C4 gave a flat average function, which means that in this case, there was the biggest diversity of answers given by subjects. DEUTSCH has suggested that envelope effect is negligible [3]. These averaged data show that, in accordance with the studies of REPP [8] there is a strong envelope effect, which cannot be omitted in further analysis.

In order to verify if the type of conducted experiment is a statistically important factor, an ANOVA variance analysis was conducted for every envelope. The analysis verified, that only in case of envelope: C4 [$F(1, 240) = 5.42$; $p < 0.02$] the type of experiment is statistically important factor.

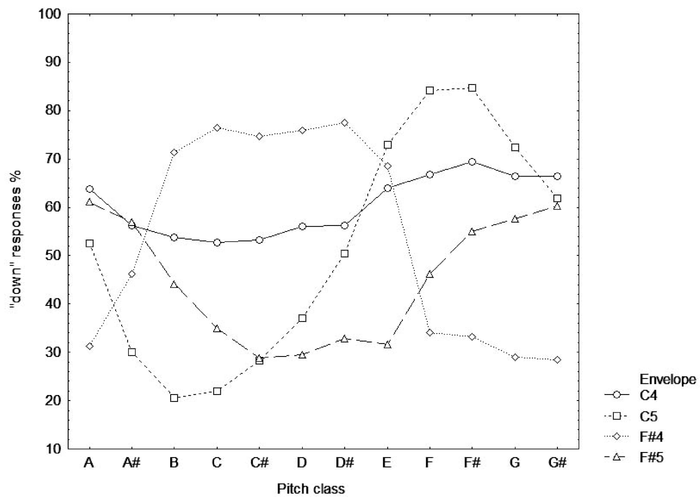

Fig. 1. Average response functions (percentage "down" responses) as a function of initial pitch class of tritone pair for four different envelope sets.

The individual results for experiment A and B for C4 envelope are shown in Fig. 2.

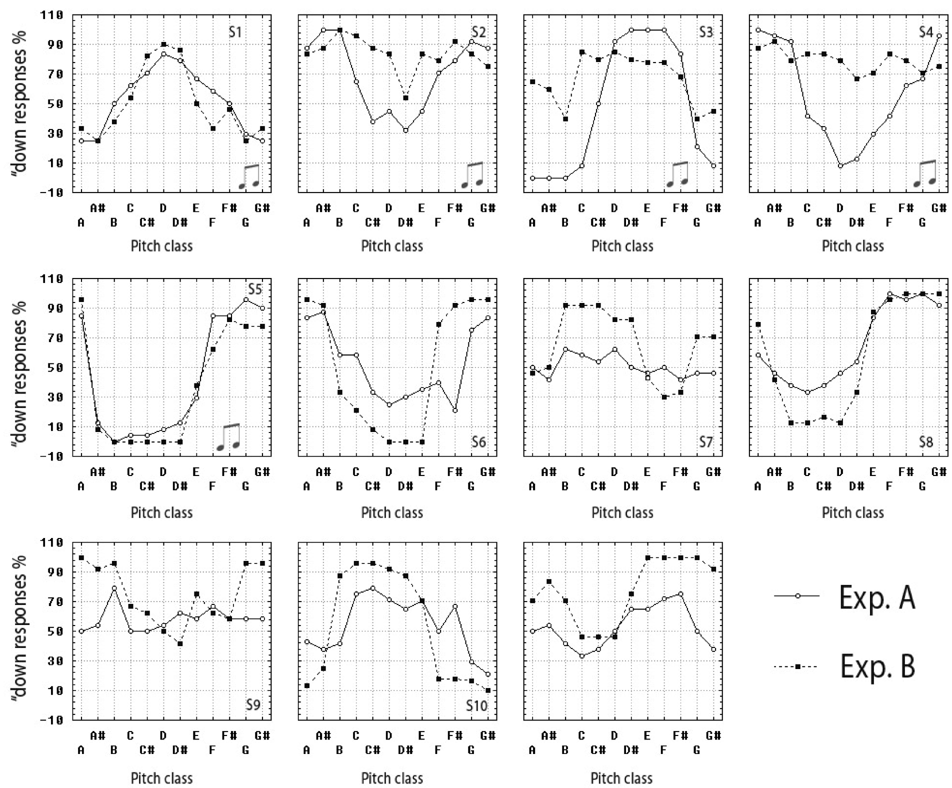

Fig. 2. Individual response functions (percentage "down" responses) as a function of initial pitch class of tritone pair for experiment A and B, for envelope C4. The note on the diagram indicate musicians.

Only listeners 1 and 5 show small differences between the perception (A) and motor (B) parts of the experiments. For all the remaining listeners those differences are significant. It can be observed, that in case of musicians, the maximums and minimums are much more distinct for experiment A. For subjects 2, 3 and 4 (all musicians) motor responses show no signs of order. Almost all answers are reproduced “upward”, regardless of the tritone pair (second sound in a pair was reproduced with lower pitch than the first one). Nevertheless in the perception part of the experiment (A) can be seen, that in accordance with previous research, subjects show clear individual pitch-class template.

The situation is opposite for the non-musicians. In this case the answers are more definite for the experiment B in comparison to the experiment A. In some cases (subject 7 and 9), responses in the perception part are random, without any tendencies, whereas in the motor part subjects show individual pitch-class template.

In both experiments, in spite of listeners being of the same nationality and used the same language, individual differences between pitch-class templates were significant.

4. Discussion

Results of experiments A and B show statistically significant differences only in case of envelope C4. This could be caused by the fact, that for that envelope, maximum energy is focused in the low frequency range. This is the range, in which the voice fundamental frequency is placed. If the motor system is responsible for the control of voice fundamental frequency, then for that envelope the largest differences when using illusory sounds can be expected.

The motor stream is evolutionally older than the perception stream (as is the case with the sense of sight). The perception system is also more complicated. For that reason, the task of singing pairs of sounds gave more determined results for non-musicians, because for them, that type of indirect judgment of pitch was more natural.

In experiment B, there was not possible to effectively cut off the influence of the perception stream. The vocalization occurred after a pair of sound was heard, and therefore the reaction could possibly be also determined by the perception stream. In case of tested musicians, who were educated to properly judge sound pitch, the activity of the motor stream could have strong impact on the perception stream. This could be the reason for deteriorating results in the experiment B.

Acknowledgment

The paper was awarded the Ist prize in the Marek Kwiek Competition.

References

[1] Boersma P., *Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound*, Proceedings of the Institute of Phonetic Sciences, **17**, 97–110 (1993).

[2] Deutsch D., *A musical paradox*, Music Perception, **6**, 275–280 (1986).

[3] Deutsch D., *The tritone paradox: Effects of spectral variables*, Perception & Psychophysics, **41**, 563–575 (1987).

[4] Deutsch D., *The tritone paradox: An influence of language on music perception*, Music Perception, **8**, 335–347 (1990).

[5] Deutsch D., *The tritone paradox: Correlate with the listener's vocal range for speech*, Music Perception, **7**, 371–384 (1991).

[6] Deutsch D., *Mothers and their offspring perceive the tritone paradox in closely similar ways*, Archives of Acoustics, **32**, 3, 517–528 (2007).

[7] Goodale M.A., Milner A.D., *Separate visual pathways for perception and action*, Trends In Neurosciences, **15**, 20–25 (1992).

[8] Hafke H.Z., *Nonconscious control of fundamental voice frequency*, J. Acoust. Soc. Am., **123**, 1, 273–8 (2008).

[9] Repp B.H., *Spectral envelope and context effects in tritone paradox*, Perception, **26**, 645–665 (1997).

[10] Repp B.H., *Compensation for subliminal timing perturbations in perceptualmotor synchronization*, Psychological Research, **63**, 106–128 (2000).

[11] Repp B.H., *Does an auditory perceptual illusion affect on-line auditory action control? The case of (de)accentuation and synchronization*, Exp Brain Res., **168**, 493–504 (2005).