

Octave effect in auditory attention

Tobias Borra^{a,b}, Huib Versnel^{c,1}, Chantal Kemner^{d,e}, A. John van Opstal^f, and Raymond van Ee^{a,f,g,h,1}

^aPhysics of Man, Helmholtz Institute, Utrecht University, 3584 CS, Utrecht, The Netherlands; ^bExperimental Psychology, Helmholtz Institute, Utrecht University, 3584 CS, Utrecht, The Netherlands; ^cOtorhinolaryngology and Head and Neck Surgery, University Medical Center Utrecht, 3584 CG, Utrecht, The Netherlands; ^dChild and Adolescent Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, 3584 CG, Utrecht, The Netherlands; ^eDevelopmental Psychology, Social Sciences, Utrecht University, 3584 CS, Utrecht, The Netherlands; ^fBiophysics, Donders Institute for Brain, Cognition, and Behaviour, Radboud University, 6525 AJ, Nijmegen, The Netherlands; ^gBrain, Body, and Behavior, Philips Research, 5656 AE, Eindhoven, The Netherlands; and ^hLaboratory of Experimental Psychology, University of Leuven, B-3000 Leuven, Belgium

Edited by Dale Purves, Duke-National University of Singapore Graduate Medical School, Singapore, and approved August 6, 2013 (received for review August 8, 2012)

After hearing a tone, the human auditory system becomes more sensitive to similar tones than to other tones. Current auditory models explain this phenomenon by a simple bandpass attention filter. Here, we demonstrate that auditory attention involves multiple pass-bands around octave-related frequencies above and below the cued tone. Intriguingly, this “octave effect” not only occurs for physically presented tones, but even persists for the missing fundamental in complex tones, and for imagined tones. Our results suggest neural interactions combining octave-related frequencies, likely located in nonprimary cortical regions. We speculate that this connectivity scheme evolved from exposure to natural vibrations containing octave-related spectral peaks, e.g., as produced by vocal cords.

psychoacoustics | attention band | priming | scene analysis

When listeners are exposed to an audible cue tone, sensitivity is highest for tones around the same frequency and diminishes as a function of distance in frequency to the cue tone (1–5). This process is thought to be mediated by attention (6) and, to date, is characterized by a bandpass filter. The shape of this attention filter resembles the shape of the peripheral auditory filter (1–5, 7, 8).

Attention to a tone can be triggered by various cues. For example, several studies applied a pure tone to assess frequency sensitivity (1–5, 7–9). Subjects can also be cued with harmonic tone complexes, producing a pitch determined by a so-called “missing fundamental” (MF) (10–12). In this case, a frequency is perceived that is not physically present. Cueing with an MF complex heightens sensitivity for targets with the same perceived pitch. The target may be a pure tone (11) or another MF complex (12). In the abovementioned studies, attention was captured involuntarily. Hafer et al. (13) showed that voluntarily directing attention to a frequency relative to a cue frequency (e.g., 1.5 times the cue frequency) also heightens sensitivity for the attended frequency. In summary, auditory sensitivity increases not for what is physically presented, but for what is attended, voluntary or involuntary, and detection of a signal becomes easier when this signal is more similar to the attended sound, in particular when the pitches are similar.

Interestingly, a number of studies have pointed out that two tones differing by an octave are perceived as more similar than when they are separated by any other musical interval (14, 15), not only in humans, but in other species as well (16, 17). The ability to discriminate between simple frequency ratios (such as the octave) is apparent in infants, well before their first birthday (18), suggesting an innate rather than a cultural predisposition for octave-related sensitivity. This perceived similarity of octave-separated tones may therefore be a plausible explanation for the universal occurrence of the octave in western and nonwestern music. This notion is further supported by statistical properties of human speech, which, interestingly, show distinct spectral peaks at octave-related distances (19). Because the shape of the attention filter has generally been studied only over a narrow range of frequencies, typically less than one octave (1–5), it is not

clear whether this octave equivalence would also apply to auditory attention. If so, we hypothesize that cueing on a particular frequency will heighten sensitivity not only for the attended frequency, but for octave-related frequencies as well.

The current study examined how attention in the auditory domain operates over a broad range of frequencies (four octaves) and, by extension, whether sensitivity is also heightened for octave-related frequencies to the attended frequency. We performed three experiments with different cue conditions: physically present cue frequencies (experiment 1), cue frequencies that are perceived, but not present (missing fundamentals; experiment 2), and imagined cue frequencies (experiment 3).

Results

Experiment 1: Physically Present Cue Tone. To first establish a baseline for detection performance without a cue tone, eight subjects performed a two-alternative forced-choice paradigm (2AFC) (Fig. 1A), for frequencies ranging from 250 to 4,000 Hz, at various signal-to-noise ratios (SNRs). To determine for each frequency the SNR at which performance was 75% correct, we fitted a sigmoid function to individual data (Fig. 1A). Next, we tested detection performance using a pure cue tone of 1,000 Hz and a range of target frequencies in background noise, each set to individual 75% performance SNR (Fig. 1B).

Detection performance was best near the cue frequency and, as expected (1–5, 7, 8, 20), gradually decreased as a function of distance to the cue frequency (Fig. 1C). However, note the appearance of distinct peaks at exactly one and two octaves above and below the cue (i.e., at 250, 500, 2,000, and 4,000 Hz, respectively). For example, detection performance for targets of 250 and 500 Hz is 10–15% better with respect to the surrounding

Significance

The human auditory system can filter sounds to select specific information. For example, when attention is “drawn” toward a tone of 1,000 Hz, the ear becomes more sensitive to that tone than to other tones. Here, we report that attention is drawn not only toward one specific frequency, as previously was assumed, but to octave-related frequencies as well. We show that this “octave effect” does not require the physical presence of the attended tone, as it persists for imagined tones: thinking of a tone increases sensitivity for the imagined tone and its octave-related frequencies. We speculate that the octave effect evolved from exposure to natural vibrations containing octave-related spectral peaks, e.g., as produced by the human vocal cords.

Author contributions: T.B., H.V., A.J.v.O., and R.v.E. designed research; T.B., H.V., A.J.v.O., and R.v.E. performed research; T.B., H.V., and A.J.v.O. analyzed data; and T.B., H.V., C.K., A.J.v.O., and R.v.E. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence may be addressed. E-mail: R.vanEe@donders.ru.nl or H.Versnel@umcutrecht.nl.

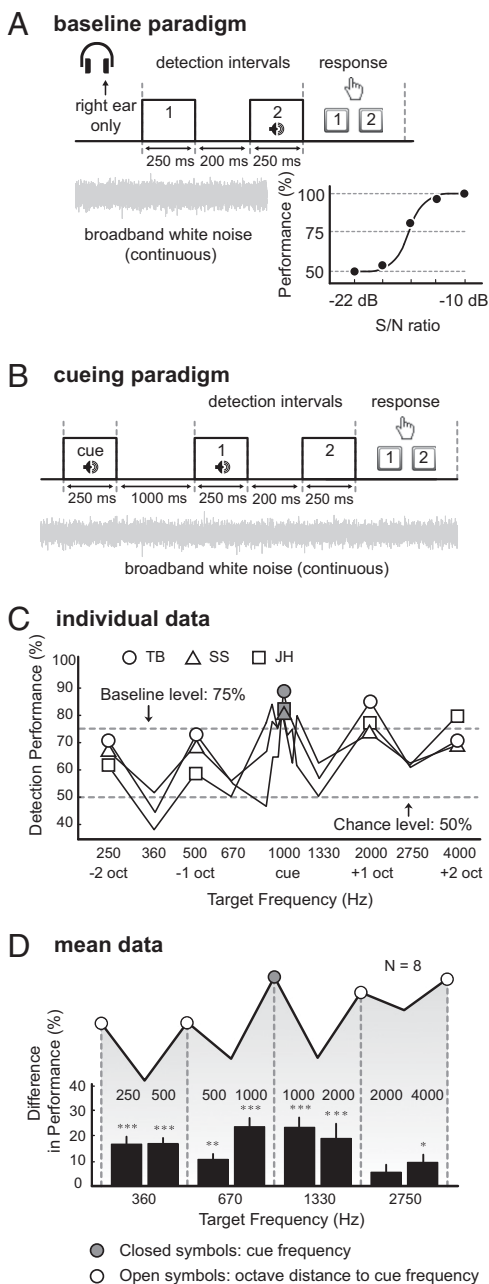


Fig. 1. Paradigm and results for experiment 1. (A) A 75%-performance signal-to-noise ratio (SNR) was determined for 250, 500, 1,000, 2,000, and 4,000 Hz by fitting a sigmoid function to detection performance as a function of SNR. By linear interpolation of a neighboring pair of these baseline SNRs, baseline SNRs for other target frequencies were determined. The slope of the psychometric curve at 75% performance is about 8%/dB. (B) Cueing paradigm. An audible cue tone precedes a two-alternative forced-choice detection task. (C and D) Individual and mean data, showing higher performance for targets identical to the cue tone (1,000 Hz) and octave-related target tones than for other target tones. Bars in D represent peak-trough differences, with the peak frequency indicated above the bar and trough frequency below the bar. Error bar represents SEM. Data were collapsed across subjects providing a single proportion correct per target frequency (41) (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). The scale on the y axis (change-percentage scores) refers to the bars, not to the mean performance scores (connected symbols), which are included for illustrative purposes. "TB" refers to the first author.

frequencies ($P < 0.001$, Z test for proportions) whereas performance for 1,000-Hz targets increased by 23% with respect to 670 and 1,330 Hz ($P < 0.001$; Z test for proportions; Fig. 1D).

The peak-trough differences (bars in Fig. 1D) were smaller at 250, 500, 2,000, and 4,000 Hz than at 1,000 Hz (significant for 500–670 Hz, 2,000–2,750 Hz, and 4,000–2,750 Hz, $P < 0.05$, paired t test). To summarize, we observed a significant relative enhancement of mean detection performance across all subjects (Fig. 1D), not only for targets at the cue frequency but, interestingly, also at octaves above and below the cue frequency.

We performed three control experiments. First, to exclude a possible role for the most peripheral and monaural stages along the auditory pathway (cochlea and the cochlear nucleus) underlying this octave effect, we repeated the experiment, by presenting the cue tone ($f_0 = 500$ Hz) to the left ear, i.e., contralateral to the ear in which the targets and broadband noise were presented. This test produced the same data pattern: higher performance for the cue frequency and at octave distances to the cue (Fig. 2). Second, to exclude the possibility that subjects may have been influenced, or overtrained, by the frequencies they attentively listened to in the baseline experiment, we performed an alternative baseline test in two new subjects at different frequencies (205, 410, 820, 1,640, and 3,280 Hz). After completing this baseline experiment, the subjects immediately performed experiment 1 using a 1,000-Hz cue, again producing performance enhancement sensitivity at octave-related tones (Fig. 3). We finally repeated experiment 1 with subject TB at a higher spectral resolution by applying an increased number of target frequencies (15 frequencies per octave; Fig. 4), which again confirmed the octave effect.

One consequence of our results is that, whereas the attention filter was previously reported to resemble a bandpass filter (1–5, 7, 8, 11–13), a more accurate description would be a filter consisting of multiple pass bands, centered at the cue frequency and at octaves above and below the cue.

Experiment 2: Missing Fundamentals as Cue Tone. To shed more light on the potential neural origin of the octave effect, we performed two further experiments to verify whether the octave effect

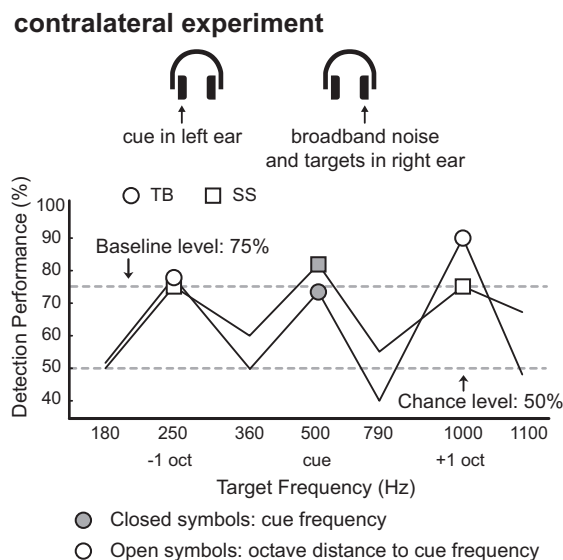
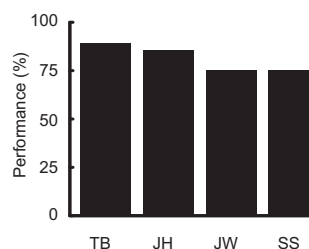
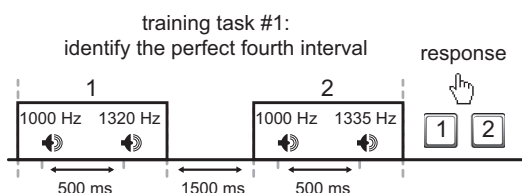


Fig. 2. Contralateral experiment. We repeated experiment 1 and changed the cue location to the left ear, i.e., contralateral to the ear in which the targets and broadband noise were presented, and set the cue tone to 500 Hz. Detection performance is shown for two subjects. Seven target frequencies were presented, indicated along the x axis. A total of 50 trials were collected for each target frequency per subject. Local performance peaks occurred at 500 Hz (the cue frequency), at 250 Hz (–1 octave of the cue frequency), and at 1,000 Hz (+1 octave of the cue frequency).

A training task 1

🔊 : 250 ms



B training task 2

🔊 : 250 ms

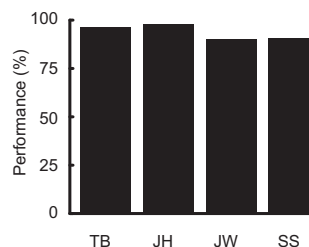
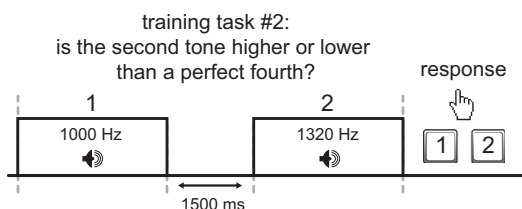


Fig. 6. Training for experiment 3. Training tasks to imagine a perfect fourth above 1,000 Hz for experiment 3. (A) Training task 1: subjects are instructed in a 2AFC task to identify which of two intervals contains a perfect fourth musical interval (1,000 Hz to 1,335 Hz) (in this example, interval 2). Possible frequencies corresponding to a non-perfect fourth interval relative to 1,000 Hz are 1,296, 1,309, 1,322, 1,348, 1,361, and 1,374 Hz. Performance ranged from 75% to 91%. (B) Training task 2: subjects are asked to respond whether the second of two tones is higher or lower than a perfect fourth; possible values of the second tone are equal to those in training task 1, i.e., ranging from 1,296 Hz to 1,374 Hz in 13-Hz increments and excluding the perfect fourth (1,335 Hz). Performance ranged from 89% to 97% for training task 2. A total number of 300 trials were collected per subject for each training task.

at the physical cue of 1,000 Hz and at corresponding octave distances (500 and 2,000 Hz), and, crucially, also at the imagined cue (1,335 Hz) and its octave-related frequencies (667.5 and 2,670 Hz). The peak–trough differences at both cues and low-frequency octave distances were significant (ranging from 9% to 24%).

Two important insights emerge from these data. First, the octave effect also holds for an imagined frequency. Second, both physical and imagined cues lead to enhanced detection. Thus, even though attention was presumably drawn toward the imagined tone, attention toward the physical cue frequency (as in experiment 1) was not “overruled.”

Discussion

We found relative performance enhancement for target tones that differed by multiples of an octave with respect to the cue tone. This improvement holds true for physically presented cue frequencies (experiment 1), cue frequencies at the missing fundamental (experiment 2), and even for imagined cue frequencies (experiment 3). These findings challenge the current views of the auditory attention filter, so far assumed to resemble a single pass band centered at the cue frequency (1–5, 7, 8, 11–13).

Apart from the range of frequencies tested (four octaves versus about one octave), our study differed from previous studies (1–5, 7, 8, 11–13) in the following important aspect. In our experiments, the target frequency was equal to the cue frequency in only 7–14% of the trials (equiprobable) whereas previous studies used a high proportion of the target-cue match (e.g., 75%). This is a relevant difference, as off-cue targets are more suppressed in a high-proportion condition than in a low-proportion condition, as demonstrated by Tan et al. (7), who compared low (20%) and high (75%) proportions of a target-cue match. In other words, the higher the probability that a target frequency is identical to a cue frequency, the more other target frequencies, possibly including those at octave distances, will be suppressed, resulting in decreased detection performance. This probability effect could be the reason that increased sensitivity at

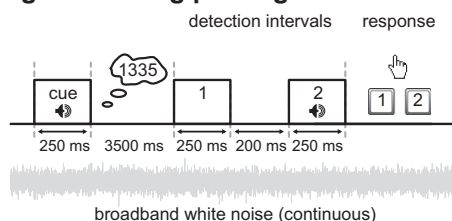
octave distances from the cue has not been reported in previous studies, including those applying targets around an octave distance (8, 9).

There is evidence that attentional focus can lead to neuronal changes in the auditory cortex (21–23). This notion is relevant with respect to the neuronal origin of auditory perception in relation to learning, cognition, and disorders such as tinnitus (24). The question arises then which potential neuronal mechanisms account for the classic attention filter, and whether they may be reconciled with the octave effect. A candidate mechanism for the classic attention filter (i.e., a relative increase in performance for targets identical to the cue tone) is rapid plasticity. Recordings from ferrets trained to attend to a tone showed a rapid form of plasticity in the spectrotemporal receptive fields of single neurons in primary auditory cortex (21, 22). These changes were facilitative and consistent with enhanced performance during a tone-detection task and were hypothesized to arise from connections between frontal cortex and primary auditory cortex. However, as this rapid type of plasticity seems to be restricted to the attended frequency, and not to extend to neighboring (octave-related) frequencies, it does not seem to account for the octave effect.

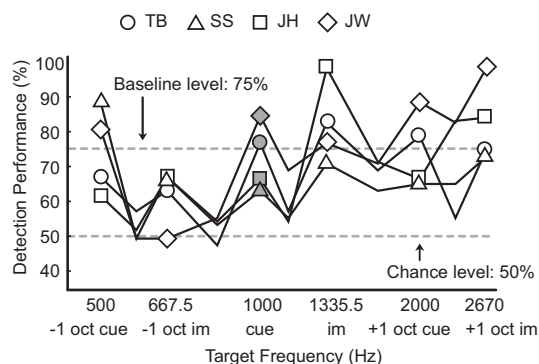
A further candidate mechanism for the octave effect could reside in tuning to multiple, harmonically related frequencies as observed in a subpopulation of mammalian primary auditory cortex neurons (25, 26). However, so far, only a few studies have described this multifrequency tuning, and, in those studies, harmonic relations between the tuning peaks have not been consistently reported (25, 27–29). Interestingly, the response for a neuron with a characteristic frequency of 1,000 Hz is facilitated when a 500-Hz tone is presented either synchronously (25) or within an interval of 400 ms (26). Because the interval to the target in our experiments was considerably longer (1,000 or 1,450 ms), it seems unlikely that this type of time-critical facilitation in primary auditory cortex would underlie the octave effect.

We suggest that involvement of secondary auditory cortex in the octave effect is more likely for the following reasons. First,

A imagined cueing paradigm



B individual data



C mean data

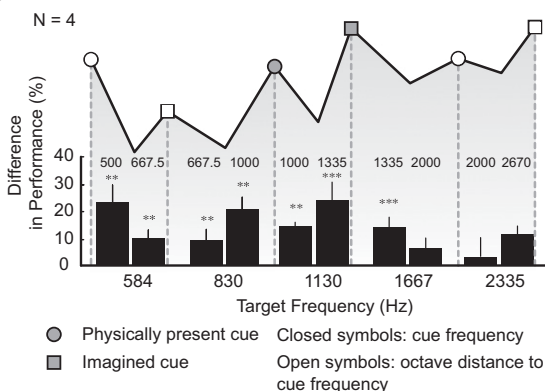


Fig. 7. Paradigm and results for experiment 3. (A) Trained subjects are instructed to imagine a perfect fourth above a presented tone of 1,000 Hz, i.e., 1,335 Hz. This instruction is followed by a two-alternative forced-choice detection task. (B) Individual data. (C) Mean data. Detection performance, for targets identical to the imagined tone (1,335 Hz) and an octave below, is significantly increased. Performance for targets identical to the 1,000-Hz physical-cue tone and an octave below is increased. Same format as Fig. 1 C and D.

the effect also arises for the missing fundamental. It therefore seems plausible that pitch neurons play a role, which are located anterolateral to the primary auditory cortex (30). Second, an imagined tone causes the same octave effect, and functional MRI (fMRI) studies have indicated that imagining a frequency predominantly activates secondary auditory areas (31, 32), rather than primary auditory cortex.

Our findings may have significant consequences for pure-tone audiometry. As known in audiometry, exposure history has an effect on tone detection. Our data indicate that the detection of, e.g., a 4-kHz tone significantly improves after prior exposure to 2 kHz. At the same time, detection at 3 kHz becomes worse than without prior presentation of the 2 kHz tone. Although we have not examined how long the cueing effect lasts, it may survive successive frequency presentations in typical tone audiometry. If so, the order of tone presentation would influence the audiogram.

One consequence of our results is that attention in the auditory domain is not drawn to an absolute stimulus property such as the frequency of the tone, but rather to a more general perceptual class of the tone, called "tone chroma" (corresponding to the pitch class of the note, ranging in western music from A to G). It is plausible that the octave effect reflects neural connectivity at octave-related frequencies. From an ecological perspective, such connectivity might have evolved from exposure to environmental sounds (e.g., speech, animal vocalizations, and vibrations such as in strings), which contain multiple harmonics such as octaves, which correspond to the second, fourth, eighth, etc. harmonic. Connectivity at octave-related distances has been implicated by harmonic template models (33–37). The octave effect is reflected in the phenomenon of octave equivalence: two tones differing by one octave are perceived as more similar than tones differing by any other musical interval (14, 15, 38, 39), reported in both humans and animals (16, 17). Our results may thus be interpreted as behavioral evidence for a mechanism in the auditory system, possibly mediated by attention, that processes tones based on their chroma, rather than on their absolute frequency. Importantly, the octave effect might explain why both western and nonwestern music is based on a universal interval: the octave.

Materials and Methods

Detection performance was measured using a two-alternative forced choice (2AFC) design. Target tones appeared in one of two 250-ms intervals, separated by 200 ms and indicated on screen with "1" or "2." Gaussian white noise was present during the whole experiment (except for the training tasks for experiment 3), set by the subject to a comfortable level, ranging from 80 to 90 dB sound pressure level (SPL). Audio was generated by a Power Mac G5 computer (sample rate: 50 kHz) and presented through a Texas Instruments TAS3004 sound card and a Philips SBC HP 910 headphone to the right ear. Cue and target-tone stimuli had a duration of 250 ms with 8-ms onset and offset ramps. Sound levels of tones and noise were measured with a Bruel and Kjaer 2260 Investigator and artificial ear 4153; harmonics or subharmonics with levels greater than 37 dB below the level of the fundamental of the tone were not detected. Subjects responded with a keypress of "1" or "2" in which interval they detected the target tone, without feedback. Experiments were self-paced, with the next trial beginning 1,000 ms after each response. Cue tones were presented at a signal/noise ratio of -7 dB: clearly audible. Experiments were written in Matlab 7.2, using the Psychophysics Toolbox extensions (40). Data were acquired in five sessions for each experiment. Experiments were carried out in a sound-attenuated, darkened room, with each experiment run on different days for every subject. All subjects reported normal hearing. Experiments were conducted after obtaining informed consent from the subject. Experimental protocols were approved by the Local Ethics Committee of the University of Utrecht.

Baseline Experiment. For each subject participating in either experiment 1, 2, or 3, a 75%-performance signal-to-noise ratio (SNR) was determined for targets of 250, 500, 1,000, 2,000, and 4,000 Hz by fitting a sigmoid function to detection performance as a function of SNR. Tone conditions (frequency, SNR) were presented in interleaved fashion, randomized per block; each block consisted of 10 trials per condition (250 trials), and the total baseline experiment consisted of 5 blocks (1,250 trials). By linear interpolation of a neighboring pair of these baseline SNRs, baseline SNRs for other target frequencies were determined. The baseline experiment, carried out before experiment 1, was conducted at least 24 h before experiment 1 (with a maximum time in between of 6 d). One subject did not participate in experiments 1 and 2, and performed the baseline experiment 6 d before experiment 3.

Experiment 1. Cue tones (1,000 Hz) were presented 1,250 ms before detection interval 1. Target frequencies were chosen from the following 15 frequencies: 250, 360, 500, 670, 880, 920, 960, 1,000, 1,040, 1,080, 1,120, 1,330, 2,000, 2,750, and 4,000 Hz. Fifty trials were collected per target frequency. Eight subjects (one female, seven male) participated in experiment 1. Four of the subjects reported to have received musical training. Musical training is defined as having practiced a musical instrument (including voice) for more than 1 y.

Experiment 2. Cue tones (a missing fundamental complex consisting of four tones with frequencies of 1,500, 2,500, 3,000, and 3,500 Hz) with a fundamental

frequency (pitch, f_0) of 500 Hz, were presented 1,250 ms before detection interval 1. Target frequencies were chosen from the following seven frequencies: 180, 250, 360, 500, 670, 1,000, and 1,330 Hz. One hundred trials were collected per target frequency. Four subjects (all male) participated in experiment 2, after participating in experiment 1. Subjects were tested at least 2 wk after participating in the baseline experiment. Two of the subjects reported to have received musical training. Note that potential distortion products at the fundamental, which are caused by the cochlea responding to the harmonic complex, are masked by the broadband noise.

Experiment 3. Cue tones (1,000 Hz) were presented 3,500 ms before detection interval 1. Subjects were instructed to imagine a “perfect fourth” interval

above 1,000 Hz (i.e., 1,335 Hz). Target frequencies were chosen from the following 11 frequencies: 500, 584, 667.5, 830, 1,000, 1,130, 1,335, 1,667, 2,000, 2,335, and 2,670 Hz. Fifty trials were collected per target frequency. Four subjects participated in experiment 3. Three subjects were tested at least 1 mo after participating in the baseline experiment, and one subject performed experiment 3 one week after participating in the baseline experiment. Two of the subjects reported to have received musical training.

ACKNOWLEDGMENTS. We thank Dr. S. Shamma for comments and Dr. W. Pestman for statistics. This work was supported by Neuroscience Cognition Utrecht Grant NCU-2010 (to T.B., H.V., C.K., and R.v.E.). R.v.E. was supported by Flemish Grant METH/08/02 assigned to J. Wagemans. A.J.v.O. was supported by Netherlands Grant NWO/ALW 821.02.009.

- Greenberg GZ, Larkin WD (1968) Frequency-response characteristic of auditory observers detecting signals of a single frequency in noise: The probe-signal method. *J Acoust Soc Am* 44(6):1513–1523.
- Dai HP, Scharf B, Buus S (1991) Effective attenuation of signals in noise under focused attention. *J Acoust Soc Am* 89(6):2837–2842.
- Ison JR, Virag TM, Allen PD, Hammond GR (2002) The attention filter for tones in noise has the same shape and effective bandwidth in the elderly as it has in young listeners. *J Acoust Soc Am* 112(1):238–246.
- Scharf B, Quigley S, Aoki C, Peachey N, Reeves A (1987) Focused auditory attention and frequency selectivity. *Percept Psychophys* 42(3):215–223.
- Yama MF, Robinson DE (1982) Comparison of frequency selectivity for the monaural and binaural hearing systems: Evidence from a probe-frequency procedure. *J Acoust Soc Am* 71:694–700.
- Pashler HE (1998) *Attention* (Psychology Press, p, Hove).
- Tan MN, Robertson D, Hammond GR (2008) Separate contributions of enhanced and suppressed sensitivity to the auditory attentional filter. *Hear Res* 241(1–2):18–25.
- Schlauch RS, Hafter ER (1991) Listening bandwidths and frequency uncertainty in pure-tone signal detection. *J Acoust Soc Am* 90(3):1332–1339.
- Yost WA, Shofner WP (2009) Critical bands and critical ratios in animal psychoacoustics: An example using chinchilla data. *J Acoust Soc Am* 125(1):315–323.
- Moore BCJ (1993) Frequency analysis and pitch perception. *Human Psychophysics*, eds Yost WA, Popper AN, Fay RR (Springer, New York).
- Ebata M, Miyazono H, Kumamaru K, Chisaki Y, Usagawa T (2001) The formation of attentional filters for a missing-fundamental complex tone and frequency-gliding tones. *Acoust Sci Technol* 22:401–406.
- Hafter ER, Saberi K (2001) A level of stimulus representation model for auditory detection and attention. *J Acoust Soc Am* 110(3 Pt 1):1489–1497.
- Hafter ER, Schlauch RS, Tang J (1993) Attending to auditory filters that were not stimulated directly. *J Acoust Soc Am* 94(2 Pt 1):743–747.
- Deutsch D, Boulanger RC (1984) Octave equivalence and the immediate recall of pitch sequences. *Musik Percept* 25:40–51.
- Kallman HJ (1982) Octave equivalence as measured by similarity ratings. *Percept Psychophys* 32(1):37–49.
- Blackwell HR, Schlosberg H (1943) Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *J Exp Psychol* 33:407.
- Wright AA, Rivera JJ, Hulse SH, Shyan M, Neiwirth JJ (2000) Music perception and octave generalization in rhesus monkeys. *J Exp Psychol Gen* 129(3):291–307.
- Schellenberg E (1996) Natural musical intervals: Evidence from infant listeners. *Psych Sci* 7:272–277.
- Schwartz DA, Howe CQ, Purves D (2003) The statistical structure of human speech sounds predicts musical universals. *J Neurosci* 23(18):7160–7168.
- Patterson RD, Moore BCJ (1986) Auditory filters and excitation patterns as representations of frequency resolution. *Frequency Selectivity in Hearing* (Academic, New York), 123–177.
- Fritz J, Shamma SA, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci* 6(11):1216–1223.
- Fritz JB, David SV, Radtke-Schuller S, Yin P, Shamma SA (2010) Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nat Neurosci* 13(8):1011–1019.
- Massoudi R, Van Wanrooij MM, Van Wetter SMCI, Versnel H, Van Opstal AJ (2013) Stable bottom-up processing during dynamic top-down modulations in monkey auditory cortex. *Eur J Neurosci* 37(11):1830–1842.
- Rauschecker JP, Leaver AM, Mühlaus M (2010) Tuning out the noise: Limbic-auditory interactions in tinnitus. *Neuron* 66(6):819–826.
- Kadia SC, Wang X (2003) Spectral integration in A1 of awake primates: Neurons with single- and multipeaked tuning characteristics. *J Neurophysiol* 89(3):1603–1622.
- Brosch M, Schulz A, Scheich H (1999) Processing of sound sequences in macaque auditory cortex: Response enhancement. *J Neurophysiol* 82(3):1542–1559.
- Sutter ML, Schreiner CE (1991) Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *J Neurophysiol* 65(5):1207–1226.
- Noreña AJ, Gourévitch B, Pienkowski M, Shaw G, Eggermont JJ (2008) Increasing spectrotemporal sound density reveals an octave-based organization in cat primary auditory cortex. *J Neurosci* 28(36):8885–8896.
- Pienkowski M, Eggermont JJ (2010) Nonlinear cross-frequency interactions in primary auditory cortex spectrotemporal receptive fields: A Wiener-Volterra analysis. *J Comput Neurosci* 28(2):285–303.
- Bendor D, Wang X (2005) The neuronal representation of pitch in primate auditory cortex. *Nature* 436(7054):1161–1165.
- Halpern AR, Zatorre RJ (1999) When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cereb Cortex* 9(7):697–704.
- Halpern AR, Zatorre RJ, Bouffard M, Johnson JA (2004) Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia* 42(9):1281–1292.
- Cohen MA, Grossberg S, Wyse LL (1995) A spectral network model of pitch perception. *J Acoust Soc Am* 98(2 Pt 1):862–879.
- Shamma SA, Klein D (2000) The case of the missing pitch templates: How harmonic templates emerge in the early auditory system. *J Acoust Soc Am* 107(5 Pt 1):2631–2644.
- Duifhuis H, Willems LF, Sluyter RJ (1982) Measurement of pitch in speech: An implementation of Goldstein's theory of pitch perception. *J Acoust Soc Am* 71(6):1568–1580.
- Goldstein JL (1973) An optimum processor theory for the central formation of the pitch of complex tones. *J Acoust Soc Am* 54(6):1496–1516.
- Terhardt E (1974) Pitch, consonance, and harmony. *J Acoust Soc Am* 55(5):1061–1069.
- Humphreys L (1939) Generalization as a function of method of reinforcement. *J Exp Psychol* 17:361–372.
- Pedersen P (1975) The perception of octave equivalence in twelve-tone rows. *Psychol Music* 3:3–8.
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10(4):433–436.
- MacMillan NA, Creelman CD (1991) *Detection Theory: A User's Guide* (Cambridge Univ Press, Cambridge, UK).