

The perception of concurrent sound objects through the use of harmonic enhancement: a study of auditory attention

Elena Koulaguina · Brandi Lee Drisdelle · Claude Alain ·
Stephan Grimault · Douglas Eck · François Vachon ·
Pierre Jolicœur

Published online: 13 February 2015
© The Psychonomic Society, Inc. 2015

Abstract When the frequency of one harmonic, in a sound composed of many harmonics, is briefly mistuned and then returned to the ‘in-tune’ frequency and phase, observers report hearing this harmonic as a separate tone long after the brief period of mistuning – a phenomenon called harmonic enhancement. Here, we examined the consequence of harmonic enhancement on listeners’ ability to detect a brief amplitude

notch embedded in one of the harmonics after the period of mistuning. When present, the notch was either on the enhanced harmonic or on a different harmonic. Detection was better on the enhanced harmonic than on a non-enhanced harmonic. This finding suggests that attention was drawn to the enhanced harmonic (which constituted a new sound object) thereby easing the processing of sound features (i.e., a notch) within that object. This is the first evidence of a functional consequence of the after-effect of transient mistuning on auditory perception. Moreover, the findings provide support for an attention-based explanation of the enhancement phenomenon.

E. Koulaguina
Université du Québec à Montréal, Montreal, QC, Canada

B. L. Drisdelle · S. Grimault · P. Jolicœur (✉)
Département de Psychologie, Université de Montréal, C.P. 6128,
Succursale Centre-Ville, Montréal, QC, Canada H3C 3J7
e-mail: pierre.jolicœur@umontreal.ca

B. L. Drisdelle · S. Grimault · P. Jolicœur
Centre de Recherche en Neuropsychologie et Cognition (CERNEC),
Université de Montréal, Montreal, QC, Canada

B. L. Drisdelle · P. Jolicœur
Laboratory for Brain, Music and Sound Research (BRAMS),
Université de Montréal, Montreal, QC, Canada

C. Alain
Rotman Research Institute, Baycrest Centre, Toronto, ON, Canada

S. Grimault
Le Centre National de la Recherche Scientifique (CNRS),
Paris, France

D. Eck
Google Inc., Mountain View, CA, USA
e-mail: deck@google.com

F. Vachon
Université Laval, Quebec City, QC, Canada

P. Jolicœur
Centre de Recherche de l’Institut Universitaire de Gériatrie de
Montréal (CRIUGM), Montreal, QC, Canada

Keywords Complex sound segregation · Mistuned harmonic · Harmonic enhancement

Introduction

The ability to separate and identify concurrent sound objects is paramount in dealing with everyday complex auditory environments. Several acoustic cues contribute to the perceptual organization of overlapping acoustic waves into separate meaningful sources. Along with frequency, intensity, spatial location, and previously learned representations, another factor that has an impact on the perception of co-occurring sounds is their harmonic organization. Both vocal chords and musical instruments normally produce complex sounds, each composed of a fundamental frequency and several additional harmonics, that is, tones with frequencies that are an integer multiple of the fundamental frequency. Human perception is adjusted to the properties of these sound-emitting

sources by treating harmonically related waveforms as one coherent auditory object. However, when waveforms contain a sound element that is inharmonically related to the fundamental frequency, this element stands out as a separate, distinct sound (e.g., Alain, Arnott, & Picton, 2001; Moore, Glasberg, & Peters, 1986).

Harmonic segregation has been examined in recent years by raising, lowering, or entirely interrupting components in a complex sound. The degree of pitch alteration was manipulated as well as the ordinal number of the affected harmonic (e.g., Alain et al., 2001; Moore, Peters, & Glasberg, 1985; Moore et al., 1986). Participants were asked either to report which of two complex tones contained a mistuned harmonic, how many sounds were perceived (e.g., Alain et al., 2001; Moore et al., 1986), or to match the pitch of a mistuned harmonic with an adjustable tone (e.g., Hartmann & Doty, 1996; Hartmann, McAdams, & Smith, 1990; Roberts & Brunstrom, 1998, 2001; Roberts & Holmes, 2006). The likelihood of reporting hearing two distinct sounds instead of one during a mistuned interval was taken as evidence for the segregation of a complex sound into two separate concurrent auditory objects (e.g., Alain et al., 2001; Moore et al., 1986).

Several factors influence the perception of pitch in a complex tone containing harmonics. The ability to hear a mistuned harmonic as a separate tone in a complex sound depends on the ordinal number of the harmonic, the duration and degree of mistuning, and, to a lesser extent, the fundamental frequency (Hartmann, 1988; Hartmann, McAdams, & Smith, 1990; Moore et al., 1985, 1986). The likelihood of perceiving a harmonic as a separate distinct auditory object decreases with increasing harmonic ordinal number and increases with a greater degree of mistuning and longer stimulus duration. Indeed, Moore et al. (1985) demonstrated that the perception of pitch is different for lower and higher harmonics. In their study, participants were instructed to indicate which of two complex tones contained a mistuned harmonic. For lower harmonics (up to the fourth harmonic), the mistuned harmonic was described as “standing out” from the complex sound whereas for higher harmonics, participants were instead sensitive to the change in phase produced by shifting the harmonic from its original frequency. The periodic fluctuation of the waveform produced by the changing phase relationship is heard as beats, which is less audible in sounds containing a lower mistuned harmonic. Moore et al. (1985) demonstrated that stimulus duration has a stronger impact on the perception of mistuning in higher harmonics and that beats appear to provide a cue for a mistuned harmonic when the stimulus has a long duration. However, in lower harmonics, the perception of beats is less audible and therefore the degree of mistuning necessary to perceive a harmonic as a separate auditory object is relatively constant regardless of stimulus duration.

In addition, it was also demonstrated that the mistuning of a harmonic by at least 4 % was needed to hear a mistuned

harmonic as a separate tone in a complex sound (Moore et al., 1986); the effect was stronger for the harmonics with a lower (200 Hz) rather than a higher (400 Hz) fundamental frequency (Alain et al., 2001). Moore et al. (1986) proposed that when a harmonic is mistuned, the underlying mechanism responsible for causing a harmonic to be heard as a separate tone may not function in a gated all-or-none manner, but rather causes the mistuned harmonic to have less and less weight in the overall pitch of the complex sound as it becomes increasingly mistuned. In other words, a harmonic becomes more audible as a separate sound object as it contributes less to the overall complex sound with increasing mistuning. Moreover, the ability to identify a mistuned harmonic as a separate sound decreases with increasing harmonic ordinal number (e.g., Alain et al., 2001; Hartmann et al., 1990), falling below .5 probability above the sixth harmonic (Alain et al., 2001). The likelihood of reporting hearing two sounds instead of one during a mistuned interval was taken as evidence for the segregation of a complex sound into two separate concurrent auditory objects (Alain et al., 2001; Moore et al., 1986). In summary, harmonic segregation can be achieved through several experimental methods by manipulation of complex tone elements such as pitch, duration, and degree of mistuning of a particular harmonic in a harmonic complex.

Previous work in the field of pitch perception (i.e., the masked-excitation model; see Terhardt, 1979; Terhardt, Stoll, & Seewann, 1982a, b) demonstrated that harmonic frequency and masking from neighboring harmonics play important roles in harmonic segregation. Indeed, the discharge rate of auditory nerve fibers of anesthetized cats to a tone at the characteristic frequency of a fiber was reduced when a second tone (a suppressor tone) was presented (Abbas & Sachs, 1976; Houtgast, 1972). Moreover, the inhibitory effect of neighboring tones appears to be stronger in the upward direction (Abbas & Sachs, 1976; Terhardt et al., 1982b). Thus, the prediction would be that a mistuned harmonic will be perceived with a positive upward pitch shift, irrespective of whether the mistuning itself is positive or negative. However, Hartmann and Doty (1996) demonstrated that the pitch perceived by listeners in a mistuned harmonic is more an exaggeration of a mistuning than a positive upward pitch shift: a positive mistuning of a harmonic led to a pitch shift in the positive direction, whereas a negative mistuning led to a pitch shift in the negative direction. Hartmann and Doty (1996) proposed a hybrid model that preserves the notion of mutual inhibition from neighboring harmonics (i.e., the masked-excitation model) and further predicts that the perceived pitch of a mistuned harmonic is instead pulled by the closest neighboring harmonic (the upper harmonic in the case of the positive mistuning, and the lower harmonic for the negative mistuning).

Lin and Hartmann (1998) demonstrated similar exaggeration pitch shift effects, using a pitch-matching task in which participants had to match the pitch of a complex tone

containing a mistuned harmonic with that of a matching sine wave. Importantly, they demonstrated this effect even when neighboring harmonics were omitted. They suggest that a harmonic template is formed when mistuning occurs and allows the detection of a component in a complex sound that does not match it, leading to the “enhancement” of this component due to the difference in actual and anticipated frequencies. Lin and Hartmann (1998) proposed that the mistuned harmonic is thus perceived with a pitch that is an exaggeration of the actual/anticipated frequency difference. One possibility is that this enhanced contrast is related to the deployment of selective attention to the mistuned harmonic when it does not fit the expected frequency template. This latter interpretation is in line with an object-based account of auditory scene analysis (Alain & Arnott, 2000), by which attention is drawn to portions of the sound determined by perceptual grouping principles that shape auditory objects. In this view, such perceptual objects are the basic units of sound to which attention can be deployed (Alain & Arnott, 2000; Shinn-Cunningham, 2008).

Viemeister and Bacon (1982) proposed another explanation of the enhancement effect that may involve adaptation of inhibition. When a formerly absent frequency is reinstated, the inhibition produced by this frequency is stronger, causing it to stand out due to the suppressed inhibition of neighboring frequencies. The account predicts that an enhanced component is perceived as more intense within the auditory system than a component that has not been enhanced.

Regardless of perceptual mechanisms underlying the enhancement effect, what can be agreed upon is that this effect exists and successfully segregates a particular harmonic frequency from a complex sound. The main interest of the current paper is what occurs following the mistuning and reinstatement of a harmonic to its original frequency; in particular, how attention is allocated when concurrent sound objects are present (i.e., the harmonic that has been enhanced and the remainder of the complex tone). Several studies on harmonicity focused on the perception of pitch and the contribution of individual components of a complex sound. However, to our knowledge, what happens following the mistuning of a harmonic, in other words the after-effects, has not been examined other than to evaluate the pitch of the enhanced harmonic. Hartmann and Goupell (2006) demonstrated that when a complex sound ended with the pulsed harmonic turned on, the perceived pitch of the harmonic that stands out was close to its original frequency, suggesting this specific harmonic was successfully enhanced. However, they also proposed that when a harmonic is omitted, the masking effect on neighboring harmonics caused by the now omitted harmonic is reduced, consistent with the adaptation of an inhibition model suggested by Viemeister and Bacon (1982). Hartmann and Goupell (2006) suggested the involvement of selective attention in enhancement; although they noted that its role was not yet clear.

It is also possible that harmonic enhancement occurs when a harmonic is removed from an intact complex sound (containing all anticipated harmonic frequencies between the fundamental to the highest harmonic), causing a gap in the harmonic organization. The turning off and on again of a particular harmonic (i.e., creating a gap in this harmonic) may cause strong transient responses such that neuronal activity responding to the harmonic is stronger than for other harmonics, causing it to stand out from the complex sound as a separate tone as does a mistuned harmonic. In the same manner, when a harmonic is mistuned, the original in-tune frequency of that harmonic is also being removed from the intact complex sound, much like a gap, because it is replaced by a frequency with a different pitch. Therefore, this interpretation assumes that in the perception of both mistuned and pulsing harmonics, the same mechanism of sound segregation is involved. According to this interpretation, both mistuned and pulsed harmonics bring about enhancement, which happens either simultaneously with the pitch change, as in the case of the mistuned harmonic, or following a temporal manipulation, as in the case of a pulsing harmonic. This interpretation provides a parsimonious explanation of how two different manipulations – the mistuning of a harmonic and the pulsing of a harmonic – result in the segregation of a pure tone from the complex sound. One could invoke the notion of attention as part of the common mechanism: pulsing or mistuning could draw attention to the frequency band of the harmonic that is different from the others (either in pitch or in temporal pattern).

In order to examine how the enhancement invoked by the brief mistuning of a harmonic could affect the perception of separate sound objects, we designed a signal-detection task in which the signal (here, an amplitude notch) was presented in a complex sound post-harmonic mistuning. An amplitude notch is defined as the momentary decrease in the amplitude of one harmonic. The amplitude depth of the notch determined the difficulty of the task (a smaller decrease in amplitude produced a more difficult detection task). The depth of the notch was varied to obtain data comparable across participants despite the individual variations in auditory capacity. In conditions where a notch was present, the notch would appear after the mistuning on either the previously mistuned harmonic or a different harmonic. In both cases, the notch would occur when all tonal components of the complex sound were back in tune and in phase; that is, in identical physical context, so any effects on notch detection would be the consequence of previous mistuning. This design can show whether the mistuned harmonic is processed distinctly even when it is no longer mistuned. If harmonic enhancement continues post-mistuning, a difference in performance can be expected, depending on whether the notch occurs on the same or on a different harmonic than the previous mistuning.

In the present experiment, we mistuned either the third or the fourth harmonic and located the notch on either the third or

the fourth harmonic, yielding four conditions forming a 2×2 within-subject design (see Table 1). Hence, the notch could be located on either the harmonic that was previously mistuned (i.e., Mistuned3-Notch3 or Mistuned4-Notch4) or on a harmonic that remained in tune all along (i.e., Mistuned3-Notch4 or Mistuned4-Notch3). Notch detection performance was evaluated using d' as the main index of sensitivity from the Signal Detection Theory (Macmillan & Creelman, 1991). The four conditions were presented in separate blocks in order to make the calculation of d' more straightforward by reducing the diversity of acoustic events that could apply to the d' formula. We hypothesized that if the segregation can survive the end of the mistuned interval, we will find a differential performance depending on whether the notch is located on the same harmonic as the previously mistuned interval or on a different one.

One interpretation of the enhancement effect is that attention is drawn to the enhanced harmonic enabling the auditory system to maintain the percept of a distinct sound object (despite fitting in with the harmonic structure of the complex sound). In this view, performance in the notch detection task would be expected to improve when the notch is placed on the enhanced harmonic and be hindered when the notch sequence is placed on a different harmonic. Another possibility, however, is that mistuning a harmonic causes a general perturbation of sound processing that is generally detrimental for the detection of weak signals. The present study avoids this possible perturbation by mistuning the harmonic for only a short time and presenting the notch later in the sound, after all harmonics are in tune and in phase. In this way we expect to have a better opportunity to study effects of enhancement and attention in the absence of possible concurrent perturbation from harmonic mismatch.

Method

Participants

Twenty-four young adults, including seven men, from 18 to 35 years of age (mean age = 23.1 years) participated in the experiment. Nineteen participants were right-handed. All participants reported having normal hearing.

Table 1 The four experimental conditions of the present study created by the combination of the notch location (third or fourth harmonic) and mistuned harmonic (third or fourth) within-subject factors

	Notch location	
Mistuning	Third harmonic	Fourth harmonic
Third harmonic	Mistuned3-Notch3	Mistuned3-Notch4
Fourth harmonic	Mistuned4-Notch3	Mistuned4-Notch4

Stimuli and design

To create the auditory stimuli, one complex sound was synthesized at a 44.1-kHz sampling rate using a custom *Matlab* program (Mathworks, Natick, MA, USA). The complex sound was composed of the first eight harmonics of a 200-Hz fundamental frequency (f_1). Hence, eight superimposed pure tones ranged from 200 to 1,600 Hz, with the harmonics representing an integer multiple of f_1 (200, 400, 600, 800, 1,000, 1,200, 1,400, and 1,600 Hz). The duration of the sound was 1,500 ms, including 5-ms rising and falling linear amplitude ramps at the beginning and the end of the sound.

All experimental stimuli (and all harmonics within each stimulus) had the same starting phase and were followed by a brief mistuned interval starting at 700 ms and lasting 100 ms, including 5-ms rising and falling slopes. They differed, however, in the ordinal number of the mistuned harmonic. In half of the stimuli, the third harmonic was mistuned by shifting its original frequency upwards by 16 % (i.e., 696 Hz instead of 600 Hz). In the other half of the stimuli, it was the fourth harmonic that was mistuned upwards by 16 % (i.e., 928 Hz instead of 800 Hz). In a study by Moore et al. (1986), it was found that the mistuning of a harmonic by 4 % is sufficient to hear the harmonic as a separate tone in a complex sound, so our shift of 16 % ensured that the harmonic would be heard as a separate tone.

In addition to the brief mistuning, the experimental sounds also had the possibility of either containing a notch or not. These notches had a length of 30 ms including a 10-ms rising and falling. Thus, there were a total of six sounds presented to each participant; two where the notch was located on either the third or fourth harmonic so that it occurred on the previously mistuned harmonic (in the Mistuned3-Notch3 and Mistuned4-Notch4 conditions), two where the notch was located on a harmonic that was in tune for the whole duration of the stimulus (in the Mistuned3-Notch4 and Mistuned4-Notch3 conditions), and two containing no notch (see Fig. 1). When present, the notch occurred 1,200 ms after the beginning of the sound, that is, 400 ms after all harmonic components were back in tune.

The ability to detect a mistuned harmonic as a separate sound decreases with increasing harmonic ordinal number (e.g., Alain et al., 2001; Hartmann et al., 1990). Because of this, it was anticipated that it would be more difficult to perceive a notch on the fourth harmonic than on the third harmonic. For this reason, task difficulty was determined prior to the experimental trials for each participant by adjusting the amplitude change in the notch to be detected. Once a difficulty level was selected in the training phase for each participant, it remained fixed for that participant for all experimental blocks. The purpose of having multiple difficulty levels was to obtain data comparable across participants despite the individual variations in auditory capacity. The amplitude of the notch was

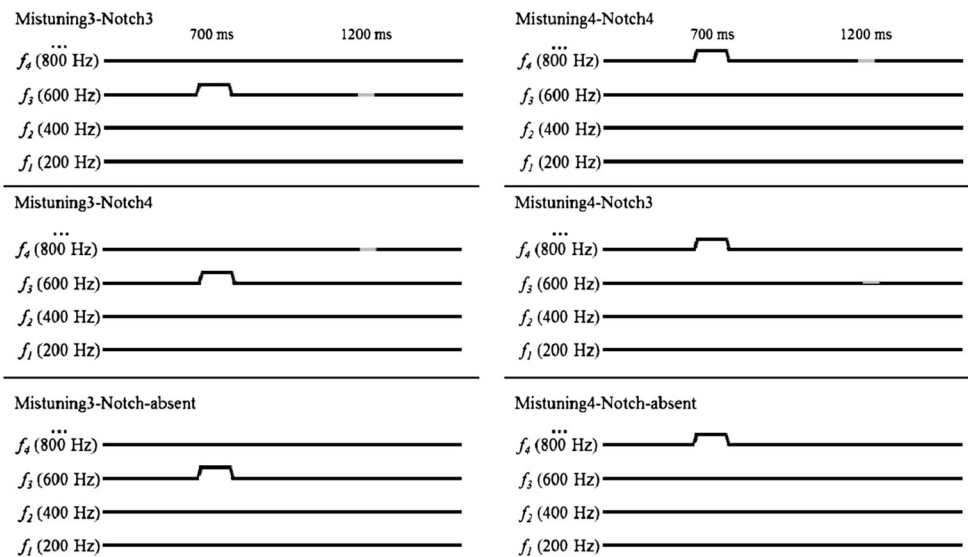


Fig. 1 There were a total of six sounds presented to each participant; two where the notch was located on either the third or fourth harmonic so that it occurred on the previously mistuned harmonic (the first row), two

where the notch was located on a harmonic that was in tune for the whole duration of the stimulus (the second row) and two containing no notch (the third row)

manipulated in order to obtain ten difficulty levels. The level of difficulty was determined by the remaining sound amplitude in the presented notch, which ranged from 0 (100 % of sound amplitude removed; the easiest condition) to 9 (10 % of sound amplitude removed; the hardest condition). For example, a notch with 10 % of the amplitude preserved (i.e., 90 % of the notch amplitude removed) would be easier to detect than a notch with 60 % of the amplitude preserved (i.e., 40 % of the notch amplitude removed).

In total, we obtained 42 sounds. Half of them had the third harmonic mistuned; another half had the fourth harmonic mistuned. Within each group, there was one sound without a notch, ten sounds with a notch placed on the third harmonic, and ten sounds with a notch placed on the fourth harmonic.

Procedure

The experiment took place in a soundproof chamber. Four test blocks were preceded by training blocks that varied in number based on individual performance. Each of the training and testing blocks contained 100 trials (among them 50 trials contained a notch and 50 did not). The task consisted of reporting whether a sound contained a notch or not. Each trial was launched by the participant by pressing the spacebar on a computer keyboard. The participant heard a sound presented binaurally through headphones 500 ms later at a comfortable hearing level, identical for all participants. The offset of the sound was immediately followed by a prompt, after which the participant responded as to whether they heard a notch or not. Half the participants pressed the ‘m’ key with their right hand to indicate the presence of the notch, and the ‘z’ key with their left hand to indicate the absence of the notch. Response key

mappings were reversed for the other half of the participants. Following the response, a feedback screen was presented until the participant chose to continue to the next trial: a green circle indicated a correct response and a red circle signified an incorrect response.

Training was always performed using the conditions in which the notch was located on a non-mistuned harmonic. This allowed equal training for all participants. Pilot data demonstrated that participants had greater difficulty in notch detection when the notch was placed on a non-mistuned harmonic.

Eighteen participants were trained in the Mistuned4-Notch3 condition, and six participants were trained in the Mistuned3-Notch4 condition. To ensure that there was no difference in task performance related to training group, a two-way mixed ANOVA was conducted to test between-subject effects possibly brought about by the difference in training group. Notch location and mistuned location were the dependent factors and whether the notch is located on the same harmonic that was mistuned or not was the dependent variable. No significant between-subjects effect was seen, $F(1, 22) = 0.11, p = .74$. Thus, belonging to a certain practice group did not affect participants’ performance during the test trials. Each participant performed a number of training blocks to adjust the difficulty of the notch (difficulty was increased by having less sound amplitude reduction during the notch and decreased by having more sound amplitude reduction during the notch). A guide for adjusting stimulus parameters during practice was an automatically calculated value of the proportion of false alarms subtracted from the proportion of hits in each block: $p(\text{hits}) - p(\text{false alarms})$. The difficulty adjustment criterion was 0.5–0.6 for the Mistuned4-Notch3 and 0.2–0.3 for the Mistuned3-Notch4 training.

In some cases, the test phase commenced without achieving the desired threshold during the practice block (for example, a notch with 30 % of the sound amplitude removed was chosen if a notch with 20 % of the sound amplitude removed led to performance lower than the sought criterion, and a notch with 40 % of the sound amplitude removed led to a performance higher than the sought criterion). If the situation arose, the intermediate difficulty level was selected for the test without being used in the training. Furthermore, sometimes the difficulty level selected during the training was inaccurate and performance rose or fell drastically during the first experimental block. In this situation, the training phase was re-done so as to achieve a more appropriate difficulty level.

The difference in the required difficulty criterion for the two conditions during training was due to the observation from a pilot study (and from the literature) showing that even in the absence of mistuning, notches were easier to detect on the third than on the fourth harmonic. Hence, lowering the target difficulty criterion for notches located on the fourth harmonic prevented ceiling effects in the easier condition.

Among the eighteen participants who were trained with the Mistuned4-Notch3 condition, fourteen participants were tested with 70 % of the notch amplitude preserved, two participants were tested with 60 %, one participant was tested with 80 %, and one participant was tested with 20 %. The average difficulty level during the test for participants with the Mistuned4-Notch3 training was 66.7 % of amplitude preserved during the notch. Among the six participants who were trained with the Mistuned3-Notch4 condition, three participants were tested with 60 % of the notch amplitude preserved, two participants were tested with 50 %, and one participant was tested with 80 %. The average difficulty level during the test for participants with the Mistuned3-Notch4 training was 60 % of amplitude preserved during the notch. Overall, the average difficulty level in the test was 65 %.

During the testing phase, each participant performed four blocks, one in each of the four experimental conditions: Mistuned3-Notch3, Mistuned4-Notch3, Mistuned3-Notch4, and Mistuned4-Notch4. The order of the blocks was counterbalanced across all participants. During the course of the experiment, the experimenter stayed in another room and only entered the testing room at the end of each block to record the hit rate, false alarm rate, and the difficulty adjustment criterion ($p(\text{hits}) - p(\text{false alarms})$), which were calculated automatically and displayed on the screen. In subsequent analyses, d' was calculated for experimental blocks for all participants.

Results

Figure 2 plots mean d' as a function of notch location and mistuned harmonic. A 2×2 repeated measures ANOVA

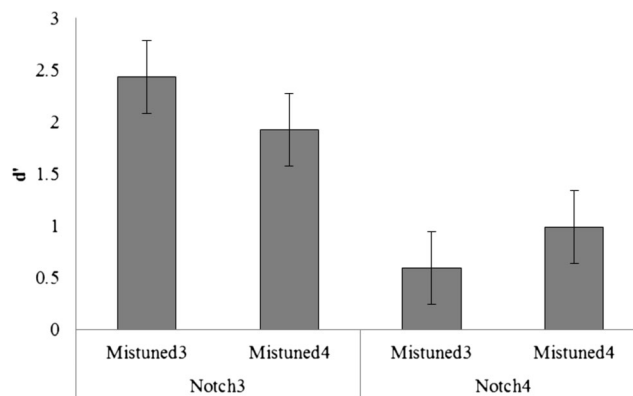


Fig. 2 Mean notch detection performance (expressed using the d') as a function of notch location (third or fourth harmonic) and mistuned harmonic (third or fourth). Error bars represent the within-subject 95 % confidence intervals

performed on these data showed that the main effect of notch location was significant, $F(1, 23) = 83.22, p < .001$, suggesting that participants detected notches on the third harmonic more efficiently than on the fourth harmonic. The main effect of mistuned harmonic was not significant, $F(1, 23) = 0.24, p = .63$, suggesting that notch detection was not affected by whether mistuning occurred on the third or the fourth harmonic. The interaction between mistuned harmonic and notch location was significant, $F(1, 23) = 7.22, p = .013$, showing that performance was better when the notch was located on the harmonic that was previously mistuned than on a harmonic that stayed in tune with the other tonal components.

Discussion

To explore the downstream consequences of harmonic enhancement, we developed a new paradigm to probe perception of a brief amplitude notch either on the enhanced harmonic or on a neighboring harmonic (that was not enhanced). Importantly, when the notch was presented, all harmonics were in tune and in phase. Hence, the perception of enhancement was an after-effect of a previous perturbation of the target harmonic. Overall, participants were better in detecting notches located on the enhanced harmonic than notches placed on a harmonic that was not enhanced.

We hypothesize that the harmonic enhancement after-effect was observed for the pitch corresponding to the initial frequency of the harmonic. This hypothesis is suggested by the fact that the harmonic was returned to the original frequency and the stimulation for all harmonics at the time of presentation of the notch provided strong bottom-up input for the originally perceived pitch both of the harmonic complex and of the enhanced harmonic. Given that we did not explicitly test exactly what pitches were heard by the observers at the time the notches probed detection sensitivity, the hypothesis is still a conjecture that awaits empirical testing.

In previous studies, the harmonic was mistuned for the entirety of the complex sound and the enhancement effect was observed for the frequency of the mistuning itself, that is, the frequency not harmonically related with the fundamental frequency (e.g., Alain et al., 2001; Moore et al., 1986). One advantage of the present experimental design is that the physical differences between a mistuned and an in-tuned harmonic did not confound the results. Even when the harmonic was brought back in tune, an enhancement effect was still observed, and supports any account that proposes a mechanism by which the enhancement effect may become self-sustaining. Importantly, the ongoing enhancement effect provides an input for downstream processing based on object-based accounts of auditory attention (Alain & Arnott, 2000; Shinn-Cunningham, 2008), which supposes that the enhanced harmonic is perceived as a distinct object. According to this interpretation, transient mistuning causes attention to be shifted to the frequency of the disrupted harmonic, leading to enhancement via one of several possible mechanisms (e.g., Hartmann & Doty, 1996; Hartmann & Goupell, 2006; Lin & Hartmann, 1998; Moore et al., 1986; Viemeister & Bacon, 1982). Ongoing attention to the now-enhanced harmonic would perhaps be facilitated by consequences of temporary attentional facilitation, such as a readjustment of relative amplitudes of neural activity in a network in which frequency-tuned cells have mutually inhibitory connections with neighboring frequencies. Ongoing attention to the enhanced tonal component would improve the later detection of signals embedded in that component.

In a recent study, Leung et al. (2011) examined whether attention is drawn to a mistuned harmonic by combining the mistuned harmonic paradigm with a gap detection task. Participants were presented with harmonic complex sounds that may have one tonal component mistuned in the otherwise periodic sound complex. In half the trials, a gap was inserted in one of the harmonics and participants indicated whether the gap was present or not. Leung and colleagues (2011) demonstrated that gap perception was impaired by the presence of a mistuned harmonic and argued that this impairment resulted from the dilution of attention because two auditory objects are perceived. In trials without mistuning, attention could be devoted entirely to a single auditory object (the complex harmonically organized sound). This effect was observed for a wide range of gap durations, and was greater when the mistuned harmonic was perceived as a separate object. At first, these results and ours seem to contradict each other. However, a closer look shows that both situations fit within the framework of the object-based theory of attention. Leung et al. (2011) used a sound that contained a mistuned harmonic for its entire duration. Thus, there are two perceptually distinct sounds that could be attended for the entirety of the auditory stimulation, including during the gap detection. In our task, the transient mistuning of a harmonic during the ongoing sound promoted the pop-out of that harmonic from the sound

compound by virtue of pre-attentive perceptual organization. In turn, we suggest that this pop-out triggered a strong tendency to attend to the enhanced harmonic, which was likely perceived as a new auditory object. Hence, notch detection was facilitated when the notch was located on this attended object which was perceptually enhanced via transient mistuning. In the case of Leung et al. (2011), the presence of an inharmonic component (i.e., the harmonic that is enhanced via mistuning) during gap presentation possibly impaired perception due to reasons other than divided attention (i.e., low-level sensory issues such as beats).

Another possible interpretation of what occurs following harmonic enhancement is that the perturbation sensory input (here, a notch) is mapped on the previously segregated auditory objects rather than on the immediately available auditory scene (here, the now in-tuned complex tone). In this scenario, a notch is mapped on one of the previously segregated auditory objects – either the mistuned harmonic or the rest of the complex sound – rather than on the complex sound as a whole. When the notch is placed on the same harmonic as a mistuning, the auditory object undergoes a larger alteration (i.e., the mistuned harmonic is the only component of that new auditory object and therefore the entirety of the auditory object is altered). However, when the notch is placed on a different harmonic than the mistuned harmonic, only part of the complex sound is altered (in our case, the notch concerns only one-seventh of the second auditory object, that is, only one out of seven non-mistuned harmonics).

Overall, our results demonstrate that following the temporary mistuning of a harmonic in a complex sound, the harmonic was perceived as a distinct sound object with a pitch at (or close to) the frequency of the in-tune harmonic, an effect known as harmonic enhancement. We hypothesized that attention was preferentially deployed to the enhanced harmonic, and that this facilitated the later detection of a brief and faint amplitude notch when the notch was on the enhanced harmonic. This is the first demonstration of a functional consequence of enhancement other than the ongoing perception of a distinct pitch in the harmonic complex. The results provide evidence for the importance of auditory attention in the perception of pitch and of signals presented in complex auditory scenes.

Acknowledgments This research was funded by a grant from the Natural Sciences and Engineering Research Council of Canada, by infrastructure support from the Canada Foundation for Innovation, and from the Fonds de recherche du Québec, and the Canada Research Chairs program.

References

- Abbas, P. J., & Sachs, M. B. (1976). Two-tone suppression in auditory nerve fibers: Extension of a stimulus-response relationship. *Journal of the Acoustical Society of America*, 59, 112–122.

- Alain, C., & Arnott, S. R. (2000). Selectively attending to auditory objects. *Frontiers in Bioscience*, 5, D202–D212.
- Alain, C., Arnott, S. R., & Picton, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 1072–1089.
- Hartmann, W. M. (1988). Pitch perception and the segregation and integration of auditory entities. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Auditory function* (pp. 623–645). New York: Wiley.
- Hartmann, W. M., & Doty, S. L. (1996). On the pitches of the components of a complex tone. *Journal of the Acoustical Society of America*, 88, 1712–1724.
- Hartmann, W. M., & Goupell, M. J. (2006). Enhancing and unmasking the harmonics of a complex tone. *Journal of the Acoustical Society of America*, 120, 2142–2157.
- Hartmann, W. M., McAdams, S., & Smith, B. K. (1990). Hearing a mistuned harmonic in an otherwise periodic complex tone. *Journal of the Acoustical Society of America*, 88, 1712–1724.
- Houtgast, T. (1972). Psychophysical evidence for lateral inhibition in hearing. *Journal of the Acoustical Society of America*, 51, 1885–1894.
- Leung, A. W. S., Jolicœur, P., Vachon, F., & Alain, C. (2011). The perception of concurrent sound objects in harmonic complexes impairs gap detection. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 727–736.
- Lin, J.-Y., & Hartmann, W. M. (1998). The pitch of a mistuned harmonic: Evidence for a template model. *Journal of the Acoustical Society of America*, 103, 2608–2617.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York, NY: Cambridge University Press.
- Moore, B. C., Glasberg, B. R., & Peters, R. W. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *Journal of the Acoustical Society of America*, 80, 479–483.
- Moore, B. C., Peters, R. W., & Glasberg, B. R. (1985). Thresholds for the detection of inharmonicity in complex tones. *Journal of the Acoustical Society of America*, 77, 1861–1867.
- Roberts, B., & Brunstrom, J. M. (1998). Perceptual segregation and pitch shifts of mistuned components in harmonic complexes and in regular inharmonic complexes. *Journal of the Acoustical Society of America*, 104, 2326–2338.
- Roberts, B., & Brunstrom, J. M. (2001). Perceptual fusion and fragmentation of complex tones made inharmonic by applying different degrees of frequency shift and spectral stretch. *Journal of the Acoustical Society of America*, 110, 2479–2490.
- Roberts, B., & Holmes, S. D. (2006). Grouping and the pitch of a mistuned fundamental component: Effects of applying simultaneous multiple mistunings to the other harmonics. *Hearing Research*, 222, 79–88.
- Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Science*, 12, 182–186.
- Terhardt, E. (1979). Calculating virtual pitch. *Hearing Research*, 1, 155–182.
- Terhardt, E., Stoll, G., & Seewann, M. (1982a). Pitch of complex signals according to virtual-pitch theory: Tests, examples, and predictions. *Journal of the Acoustical Society of America*, 71, 671–678.
- Terhardt, E., Stoll, G., & Seewann, M. (1982b). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, 71, 679–687.
- Viemeister, N. F., & Bacon, S. P. (1982). Forward masking by enhanced components in harmonic complexes. *Journal of the Acoustical Society of America*, 71, 1502–1507.