Published in final edited form as:

J Acoust Soc Am. 2009 January ; 125(1): 15–18. doi:10.1121/1.3026327.

Reduced contribution of a nonsimultaneous mistuned harmonic to residue pitch: The role of harmonic number

Hedwig E. Gockel^{a)} and Robert P. Carlyon

MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK

Christopher J. Plack

School of Psychological Sciences, University of Manchester, Manchester M13 9PL, UK

Abstract

Listeners adjusted a matching sound so that its pitch equaled that of a subsequent 90-ms complex tone (12 harmonics of a 155-Hz F₀), whose mistuned harmonic ($\pm 3\%$) was presented either simultaneously with or after the other harmonics. The 3rd or the 4th harmonic was mistuned. For both harmonics, pitch shifts were significantly smaller when the mistuned harmonic was presented after rather than simultaneously with the remainder of the target complex and when the component corresponding to the mistuned harmonic was absent (as opposed to present) from the matching sound. Presenting the mistuned component after the remainder of the complex generally reduced the pitch shifts more for the 3rd than for the 4th harmonic; the nonsimultaneous presentation reduced shifts by a factor of 2.8 (3rd) and 1.7 (4th) in the presence, and 6.4 (3rd) and 1.9 (4th) in the absence of the nominally mistuned harmonic from the matching sound. The results are discussed in terms of an analytical-listening and matching strategy that results in larger pitch shifts for a nonsimultaneous mistuned harmonic that has an octave or double-octave relationship to the F₀ than for one that does not, without it necessarily being temporally integrated into the residue pitch of the complex.

I. INTRODUCTION

Several researchers have explored the duration over which information is integrated in the determination of the residue pitch of a complex tone (Houtsma and Goldstein, 1972; Hall and Peters, 1981; Houtsma, 1984; Ciocca and Darwin, 1999; Grose *et al.*, 2002; Gockel *et al.*, 2005b). This question has implications for theories of pitch perception which specify the time constant over which pitch is calculated (Meddis and Hewitt, 1991a; 1991b; Wiegand, 2001).

Some of the studies addressing the question of the integration time for residue pitch exploited the fact that shifting the frequency of a single harmonic in an otherwise harmonic complex tone can lead to a shift in the residue pitch of the complex even when the amount of the mistuning is so large that the mistuned harmonic is heard out as a separate tone (Moore *et al.*, 1985; Moore *et al.*, 1986). The largest shifts in residue pitch have been found when the frequency of the mistuned harmonic is shifted by about 3% from its harmonic value (Moore *et al.*, 1985; Darwin *et al.*, 1992). Ciocca and Darwin (1999) compared the size of the shifts in residue pitch obtained for a mistuned component presented simultaneously with the remainder of the complex with that obtained when the mistuned component was presented immediately after the remainder of the complex. The 90-ms

Corresponding author: Hedwig Gockel, Tel.: +44 1223 355294 ext 703. Fax: +44 1223 359062.

^{a)}Electronic mail: hedwig.gockel@mrc-cbu.cam.ac.uk

harmonic complex had a fundamental frequency (F0) of 155 Hz and its 4th harmonic was mistuned by $\pm 3\%$. When the mistuned component was presented after the remainder of the complex, it could clearly be heard as a separate tone “popping out”, in addition to the richer percept evoked by the complex. Subjects were instructed to ignore this separate higher tone. In experiment one, Ciocca and Darwin found no significant difference between the size of the pitch shifts observed with simultaneous presentation of the mistuned harmonic (mean shift of 0.7 Hz) and with post-remainder presentation (mean shift of 0.5 Hz). In experiment 2, they investigated the effect of silent gaps of various durations between the end of the remainder of the complex and the onset of the mistuned harmonic. The mean pitch shifts for simultaneous and post-target presentation with a 0-ms gap were virtually identical (0.52 and 0.49 Hz). The shifts for the 40- and 80-ms gap conditions were not significantly different from that for the 0-ms gap, but for the 160-ms gap the pitch shift was reduced to zero. Based on their results, Ciocca and Darwin (1999) estimated the integration time for pitch to be around 170-250 ms.

In contrast, Gockel *et al.* (2005b) recently observed that pitch shifts *were* significantly reduced when the mistuned harmonic was presented after the remainder of the complex. In addition, pitch shifts were reduced when the matching complex (the complex adjusted in F0 by the subjects) did not contain a component corresponding to the mistuned harmonic. However, Gockel *et al.* mistuned the 3rd rather than the 4th harmonic. They discussed two, not mutually exclusive, explanations for the discrepancy between their results and those of Ciocca and Darwin (1999). Firstly, the 4th harmonic might have a special status due to its double-octave relationship to the F0. For example, Houtsma (1979) reported that an isolated partial with an octave relationship to the fundamental, i.e., with a frequency 2, 4, or 8 times the F0, has the same pitch chroma (see Bachem, 1950) as the harmonic complex. This identity of pitch chromas could lead to the 4th harmonic being more integrated into the pitch of the target complex than the 3rd harmonic, even when slightly mistuned and presented after the remainder of the target. Secondly, rather than basing their judgements on a residue pitch which incorporates the mistuned component, subjects might have used a more analytical listening and matching strategy. This analytical strategy would involve: (a) comparing and trying to minimize the difference between the pitch of the mistuned component that pops out and either the pitch of the corresponding harmonic in the matching sound or the residue pitch of the matching sound, and, also (b) comparing and trying to minimize the difference between the residue pitch of the remainder of the target complex (without the mistuned harmonic) and that of the matching sound. One would have to assume that subjects do both (a) *and* (b) in order to explain why the observed pitch shifts first increase with increasing mistuning and, with further mistuning beyond about 3 %, decrease again (Moore *et al.*, 1985; Darwin, 1992). The finding that pitch shifts were smaller in the absence of the 3rd harmonic from the matching sound than in its presence, lends some support to the idea that subjects might have used such an analytical listening and matching strategy. Note that: (1) such a strategy is more likely to be used in nonsimultaneous than in simultaneous conditions because mistuned components are perceptually segregated to a much higher degree in the former than in the latter case, and (2) the use of such a strategy would lead to larger pitch shifts for the 4th than for the 3rd harmonic because the pitch chroma difference between the mistuned harmonic and the remainder of the target complex is smaller for the 4th than for the 3rd harmonic.

The objective of the present study was to investigate the role of the number of the mistuned harmonic in the contribution of a nonsimultaneous harmonic to residue pitch. Specifically, the question was whether the difference in the conclusions reached by Ciocca and Darwin (1999) and by Gockel *et al.* (2005b) is due to effects specific to the 4th harmonic. For that purpose, in the present study the same subjects were tested with either the 3rd or the 4th harmonic mistuned, and this harmonic could be presented either simultaneously with or after

the remainder of the target complex. Additionally, the importance of the presence of the corresponding harmonic in the adjustable matching complex was assessed.

II. METHOD

The stimuli were similar to those used by Ciocca and Darwin (1999) and identical to those used by Gockel *et al.* (2005b), except that either the 3rd harmonic or the 4th harmonic was mistuned by $\pm 3\%$. The target was a 90-ms complex tone with an F0 of 155 Hz, containing the first 12 harmonics added in sine phase, each with equal amplitude (58 dB SPL). The mistuned harmonic was presented either simultaneously with, or immediately after, the remainder of the target complex ("Post-target condition"). The matching sound was a 90-ms harmonic complex, nominally containing the first 12 harmonics added in sine phase. Depending on the condition, the 3rd or the 4th harmonic was either present in or absent from the matching sound.

The method was the same as that described by Gockel *et al.* (2005b). A two-interval, two-alternative adaptive procedure was used to obtain pitch matches between a matching sound (whose F0 was adjusted, and which was always presented first) and the target complex. Subjects were explicitly instructed to ignore any pure-tone percept in the interval containing the target sound and to match the pitch of the matching complex to that of the target complex.

Eight musically trained subjects participated in all conditions. Their ages ranged from 18 to 34 years, and their quiet thresholds at octave frequencies between 250 and 4000 Hz were within 15 dB of the ANSI (2004) standard. Stimuli were presented monaurally to the left ear for seven subjects and to the right ear for one subject. After about 2-3 h of practice, five subjects were tested first with the 3rd harmonic mistuned and then with the 4th; the order was reversed for the other subjects. The five subjects who were tested first with the 3rd harmonic also participated in part of Gockel *et al.*'s (2005b) study.

III. RESULTS AND DISCUSSION

The number of pitch matches collected was varied across subjects in order to compensate to a certain degree (depending on the availability of the subject) for differences in the variability of the matches across subjects. At least ten (usually 15) pitch matches were obtained for each subject and condition (and for each direction of mistuning) resulting in a typical standard error of about 0.06-0.07 Hz. The results are presented as the mean *absolute* pitch shift, calculated as half the difference between pitch matches for corresponding positive and negative mistunings.

Figure 1 shows the geometric means and the standard errors of the pitch shifts across subjects. For the 3rd harmonic mistuned (Fig. 1a), the results are very similar to those reported for Experiment 1A of Gockel *et al.* (2005b). This is not unexpected, as data for five of the eight subjects here contributed to the data for eight subjects shown in Fig. 2 of Gockel *et al.* (2005b). In the absence of the 3rd harmonic from the matching sound, the present results are also very similar to those reported for Experiment 2 of Gockel *et al.* (2005b), where either the matching sound was a sine tone (Experiment 2B) or the target complex and the matching sound contained only non-overlapping harmonics (Experiment 2A). As discussed by Gockel *et al.* (2005b), the similarity between results obtained in Experiment 2 and those obtained with a matching sound containing harmonics 1-2 and 4-12 indicates that the results with the latter (and by extension the present results in the absence of the nominally mistuned harmonic from the matching sound) were not due to *perceptual segregation* of the mistuned component from the target complex caused by the absence of

that single specific component in the matching sound. Rather the absence of the nominally mistuned harmonic from the matching sound *per se* seems crucial.

The mean pitch shifts were generally larger for the 4th (Fig. 1b) than for the 3rd harmonic. When the 4th harmonic was present in the matching sound, the size of the pitch shift in the simultaneous condition (0.78 Hz) was bigger than but within the range of the standard error of the mean reported by Ciocca and Darwin (1999) for this condition in their experiment 1. The post-target condition (with the 4th harmonic present in the matching sound) led to a mean pitch shift of about 0.46 Hz, which corresponds well with Ciocca and Darwin's (1999) results.

A repeated-measures three-way ANOVA (with factors harmonic number, timing of the mistuned component and presence of a component corresponding to the mistuned harmonic in the matching complex) was calculated, using the logarithm of the mean pitch shift for each subject and condition as input. Logarithms were used because we were interested in the relative changes across conditions. The ANOVA showed that the three main effects were highly significant [harmonic number: $F(1,7)=33.48$, $p=0.001$; timing: $F(1,7)=138.8$, $p<0.001$; presence: $F(1,7)=32.78$, $p=0.001$] and that all interactions were significant [timing $\times$ harmonic number: $F(1,7)=21.1$, $p<0.01$; timing $\times$ presence: $F(1,7)=13.05$, $p<0.01$; harmonic number $\times$ presence: $F(1,7)=9.36$, $p<0.05$; timing $\times$ harmonic number $\times$ presence: $F(1,7)=13.05$, $p<0.05$]. To assess whether the difference between the simultaneous and the post-target conditions were significant for each harmonic alone, repeated-measures two-way ANOVAs were calculated separately for the 3rd and the 4th harmonics. For the 4th harmonic, the ANOVA showed significant main effects of timing [$F(1,7)=25.77$, $p=0.001$], and of presence [$F(1,7)=15.84$, $p<0.01$], but no significant interaction. For the 3rd harmonic there were again significant main effects of timing [$F(1,7)=112.73$, $p<0.001$], and of presence [$F(1,7)=28.46$, $p=0.001$], and a significant interaction between timing and presence [$F(1,7)=9.31$, $p<0.05$].

In summary, pitch shifts were significantly larger for: (1) the 4th rather than the 3rd harmonic being mistuned; (2) simultaneous rather than post-target presentation of the mistuned harmonic; and (3) presence rather than absence of the nominally mistuned harmonic from the matching sound. Additionally, all interactions were significant. The present results for the 4th harmonic replicate and extend the findings of Gockel *et al.* (2005b) and, in spite of the similarity of the means with those of experiment one (but not experiment two) of Ciocca and Darwin (1999), do not replicate their finding of no significant difference between pitch shifts for simultaneous and post-target presentation of the mistuned component. For the 4th harmonic being mistuned and present in the matching sound, this discrepancy probably was caused by the standard errors within subjects and conditions being larger in Ciocca and Darwin's study (typically around 0.1 Hz, personal communication) than in the current study (about 0.06-0.07 Hz). For their experiment testing the influence of silent gaps of various durations between the mistuned harmonic and the remainder of the target complex, they reported that "The criterion for stopping after five matches was a standard error smaller than 0.4 Hz in all conditions".

Importantly, the reduction in the size of the pitch shifts in the post-target compared to the simultaneous condition was significantly smaller for the 4th (a factor of 1.8) than for the 3rd harmonic (a factor of 4.2), as shown by the significant interaction between harmonic number and timing of the mistuned harmonic. This supports the hypothesis mentioned in the Introduction that, due to the double-octave relationship, the 4th harmonic has a special status with regard to pitch matches obtained when the mistuned harmonic is presented after the remainder of the complex.

The fact that pitch shifts were larger for the 4th than for the 3rd harmonic mistuned theoretically could indicate that the 4th harmonic generally (whether presented after or simultaneously with the remainder of the complex) has a somewhat special status due to its (near) double-octave relationship to the residue pitch of the target. However, several earlier studies which investigated the relative dominance of individual partials in determining the pitch of complex tones using *simultaneous* presentation of all harmonics, did not report a higher dominance of the 4th harmonic (Moore *et al.*, 1985; Dai, 2000; Gockel *et al.*, 2005a). Thus, it seems most likely that, in the present study, for *simultaneous* presentation, the finding of larger pitch shifts for the 4th than for the 3rd harmonic mistuned was due to a combination of factors such as the relatively low F0 of 155 Hz, the short stimulus duration of 90 ms, and the specific subjects used.

Pitch shifts were significantly smaller in the absence of the nominally mistuned harmonic from the matching sound than in its presence. This was true for both harmonic numbers tested and supports the idea that subjects may partly have used an analytical listening and matching strategy, as described in the Introduction. If the nominally mistuned harmonic is absent from the matching sound, the mistuned component cannot be compared with it, so the matches will be somewhat less influenced by the mistuned harmonic. When the nominally mistuned harmonic is present in the matching sound, the comparison between the mistuned harmonic and the corresponding harmonic in the matching sound may be partly affected by frequency-shift detectors (Demany and Ramos, 2005).

IV. SUMMARY AND CONCLUSIONS

- (1) Pitch shifts were significantly smaller when the matching complex did not contain a component corresponding to the mistuned harmonic, suggesting that an analytical listening and matching strategy contributed to the pitch shifts measured.
- (2) Shifts in matches to the residue pitch were significantly smaller when the mistuned component was presented after the remainder of the target complex than when it was presented simultaneously. This contrasts with Ciocca and Darwin's (1999) *statistical* result for the comparison of simultaneous and nonsimultaneous presentation of a mistuned 4th harmonic in the presence of the 4th harmonic in the matching sound; while the mean results of the present study for the 4th harmonic mistuned were similar to some of the results reported by Ciocca and Darwin (1999), the present standard errors within and across subjects were smaller. The significant reduction in pitch shifts for the post-target conditions replicates and extends the results of Gockel *et al.* (2005b), who mistuned the 3rd harmonic.
- (3) The factor by which the pitch shift decreased when changing from simultaneous to nonsimultaneous presentation of the mistuned component was greater for the 3rd than for the 4th harmonic. This could indicate that nonsimultaneous components which have an octave or double-octave relationship to the F0 are integrated into the estimate of the residue pitch over a *longer* time window than other components. However, an alternative and perhaps more plausible explanation is that listeners use an analytical listening and matching strategy (especially) when the mistuned harmonic is presented nonsimultaneously and therefore pops out. The use of this strategy would result in larger pitch shifts when the mistuned harmonic has an octave or double-octave relationship to the F0 than when it does not (see Introduction). In other words, nonsimultaneous harmonics that have an octave or double-octave relationship to the F0 affect pitch matches more than when they don't have this relationship, without actually being temporally integrated into the residue pitch. Thus, obligatory integration of post-target components for the purpose of determining the residue pitch might be less strong than originally estimated by Ciocca and Darwin (1999), even for the 4th harmonic.

Acknowledgments

This work was supported by EPSRC Grant EP/D501571/1. We thank Brian Moore, Christian Lorenzi, Chris Darwin, Valter Ciocca, Ruth Litovsky and an anonymous reviewer for comments on an earlier version of this paper.

REFERENCES

- ANSI. ANSI S3.6-2004, Specification of Audiometers. American National Standards Institute; New York: 2004.
- Bachem A. Tone height and tone chroma as two different pitch qualities. *Acta Psychologica*. 1950; 7:80–88.
- Ciocca V, Darwin CJ. The integration of nonsimultaneous frequency components into a single virtual pitch. *J. Acoust. Soc. Am.* 1999; 105:2421–2430. [PubMed: 10212423]
- Dai H. On the relative influence of individual harmonics on pitch judgment. *J. Acoust. Soc. Am.* 2000; 107:953–959. [PubMed: 10687704]
- Darwin, CJ. Listening to two things at once. In: Schouten, MEH., editor. *The Auditory Processing of Speech - From Sounds to Words*. Mouton de Gruyter; Berlin: 1992.
- Darwin, CJ.; Buffa, A.; Williams, D.; Ciocca, V. Pitch of dichotic complex tones with a mistuned frequency component. In: Cazals, Y.; Demany, L.; Horner, K., editors. *Auditory Physiology and Perception*. Pergamon; Oxford: 1992.
- Demany L, Ramos C. On the binding of successive sounds: Perceiving shifts in nonperceived pitches. *J. Acoust. Soc. Am.* 2005; 117:833–841. [PubMed: 15759703]
- Gockel H, Carlyon RP, Plack CJ. Dominance region for pitch: Effects of duration and dichotic presentation. *J. Acoust. Soc. Am.* 2005a; 117:1326–1336. [PubMed: 15807021]
- Gockel H, Plack CJ, Carlyon RP. Reduced contribution of a nonsimultaneous mistuned harmonic to residue pitch. *J. Acoust. Soc. Am.* 2005b; 118:3783–3793. [PubMed: 16419823]
- Grose JH, Hall JW, Buss E. Virtual pitch integration for asynchronous harmonics. *J. Acoust. Soc. Am.* 2002; 112:2956–2961. [PubMed: 12509016]
- Hall JW, Peters RW. Pitch for nonsimultaneous successive harmonics in quiet and noise. *J. Acoust. Soc. Am.* 1981; 69:509–513. [PubMed: 7462473]
- Houtsma AJM. Musical pitch of two-tone complexes and predictions by modern pitch theories. *J. Acoust. Soc. Am.* 1979; 66:87–99. [PubMed: 489835]
- Houtsma AJM. Pitch salience of various complex sounds. *Music Percept.* 1984; 1:296–307.
- Houtsma AJM, Goldstein JL. The central origin of the pitch of pure tones: Evidence from musical interval recognition. *J. Acoust. Soc. Am.* 1972; 51:520–529.
- Meddis R, Hewitt M. Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery. I: Pitch identification. *J. Acoust. Soc. Am.* 1991a; 89:2866–2882.
- Meddis R, Hewitt M. Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery. II: Phase sensitivity. *J. Acoust. Soc. Am.* 1991b; 89:2883–2894.
- Moore BCJ, Glasberg BR, Peters RW. Relative dominance of individual partials in determining the pitch of complex tones. *J. Acoust. Soc. Am.* 1985; 77:1853–1860.
- Moore BCJ, Glasberg BR, Peters RW. Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *J. Acoust. Soc. Am.* 1986; 80:479–483. [PubMed: 3745680]
- Wiegand L. Searching for the time constant of neural pitch extraction. *J. Acoust. Soc. Am.* 2001; 109:1082–1092. [PubMed: 11303922]

FIG. 1.

Mean pitch shifts (log scale) and the corresponding standard errors across eight subjects. Figure 1a (top) shows the data for the 3rd harmonic mistuned. The left-hand group of two bars shows pitch shifts obtained for a matching complex that included the 3rd harmonic. The right-hand group of two bars shows the pitch shifts obtained when the 3rd harmonic was absent from the matching complex. The white and the black bars show results for simultaneous and post-target presentation of the mistuned component, respectively. Figure 1b (bottom) as 1a, but for the 4th harmonic mistuned.