

# Across-frequency pitch discrimination interference between complex tones containing resolved harmonics

Christophe Micheyl<sup>a)</sup> and Andrew J. Oxenham<sup>b)</sup>

Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, Minnesota 55455

(Received 7 September 2005; revised 29 November 2006; accepted 7 December 2006)

Pitch discrimination interference (PDI) refers to an impairment in the ability to discriminate changes in the fundamental frequency ( $F_0$ ) of a target harmonic complex, caused by another harmonic complex (the interferer) presented simultaneously in a remote spectral region. So far, PDI has been demonstrated for target complexes filtered into a higher spectral region than the interferer and containing no peripherally resolved harmonics in their passband. Here, it is shown that PDI also occurs when the target harmonic complex contains resolved harmonics in its passband (experiment 1). PDI was also observed when the target was filtered into a lower spectral region than that of the interferer (experiment 2), revealing that differences in relative harmonic dominance and pitch salience between the simultaneous target and the interferer, as confirmed using pitch matches (experiment 3), do not entirely explain PDI. When the target was in the higher spectral region, and the  $F_0$  separation between the target and the interferer was around 7% or 10%, dramatic PDI effects were observed despite the relatively large  $F_0$  separation between the two sequential targets (14%–20%). Overall, the results suggest that PDI is more general than previously thought, and is not limited to targets consisting only of unresolved harmonics. © 2007 Acoustical Society of America. [DOI: 10.1121/1.2431334]

PACS number(s): 43.66.Hg, 43.66.Fe, 43.66.Dc, 43.66.Ba [JHG]

Pages: 1621–1631

## I. INTRODUCTION

Harmonic complex tones usually evoke a pitch that corresponds to their fundamental frequency ( $F_0$ ). Most current theories and models of pitch perception assume that the auditory system determines the pitch of a complex tone by integrating information across a relatively wide frequency range. In some cases the integration is based on estimates of the frequencies of individual resolved harmonics (e.g., Goldstein, 1973); in other cases it is based on a “summary autocorrelation function” that sums the periodicities present across all frequency channels (e.g., Licklider, 1951; Meddis and O’Mard, 1997).

In cases where multiple harmonic sound sources are simultaneously present, the auditory system seems able to selectively integrate information from components that fit the harmonic series of the target source for the purpose of estimating its pitch, at least when peripheral interactions do not lead to masking of the target components (e.g., Beerends and Houtsma, 1989). Various models have been designed to account for this ability to extract concurrent pitches when the complexes occupy overlapping spectral regions (e.g., Beerends and Houtsma, 1989; Meddis and Hewitt, 1992; de Cheveigné and Kawahara, 1999). If the sources occupy distinct spectral regions, so that peripheral interactions between the two groups of harmonics are insignificant, it has generally been assumed that the pitch of one source should not interfere with the pitch of the other.

Contrary to this prevailing view, Gockel *et al.* (2004, 2005a) recently showed that listeners’ ability to discriminate small changes in the  $F_0$  of a “target” group of harmonics filtered into a given spectral region could be severely disrupted by the presence of an “interferer” group of harmonics, which was filtered into another spectral region and had a constant  $F_0$ . Note that this effect, which was termed pitch discrimination interference (PDI), occurred despite the fact that the target and the interferer were filtered into remote, nonoverlapping spectral regions, and listeners were instructed to focus on the target sound. This across-frequency PDI phenomenon is superficially reminiscent of another form of interference in pitch discrimination, where a listener’s ability to discriminate differences in the  $F_0$ ’s of two successive target complexes is impaired by the presentation, before and after each target, of an interferer complex filtered into the same spectral region (Carlyon, 1996; Micheyl and Carlyon, 1998; Gockel *et al.*, 1999). Although in the latter case  $F_0$  discrimination is impaired only when the interferer occupies the same spectral region as the target complex, in both cases the interference occurs between harmonic complexes that do not significantly interact within the auditory periphery: in one case the interference is between two complexes filtered *simultaneously* into *different* spectral regions, whereas in the other, it is between two complexes filtered *sequentially* into the *same* spectral region.

The discovery of PDI effects has potentially important implications for theories and models of pitch perception, some of which are discussed by Gockel *et al.* (2004, 2005a). However, it is not yet clear how general or relevant the effect of PDI is for everyday perception. Gockel *et al.* studied the case where the target contained only harmonics numbered above the 16th, which are believed to be

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: cmicheyl@umn.edu

<sup>b)</sup> Electronic mail: oxenham@umn.edu

unresolved in the auditory periphery and produce a relatively weak pitch percept (Plomp, 1964; Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994; Bernstein and Oxenham, 2003). In contrast, most sources in the environment also contain lower-order resolved harmonics, which produce a much more salient pitch than unresolved harmonics, and which appear to be important for the segregation of concurrent harmonic sources differing in  $F_0$  (Carlyon, 1996; Carlyon *et al.*, 2002; Qin and Oxenham, 2005). Thus, an important question is whether PDI also occurs when the target contains peripherally resolved harmonics. If the relative pitch salience of target and interferer governs the amount of PDI observed, as suggested by Gockel *et al.* (2004), one might expect target complexes that contain resolved harmonics in their passband to be far less susceptible to PDI than target complexes that contain only unresolved harmonics. On the other hand, even resolved harmonics affect the overall pitch of a complex to differing degrees (Plomp, 1967; Ritsma, 1967; Moore *et al.*, 1985; Dai, 2000; Gockel *et al.*, 2005b). If the relative dominance of individual resolved harmonics plays a role in PDI, substantial PDI might be obtained between two groups of resolved harmonics if the harmonics in the interferer group are more dominant in determining pitch than those of the target group. The first aim of the present study was therefore to test whether PDI generalizes to conditions in which the target complex contains resolved harmonics.

## II. EXPERIMENT 1: PITCH DISCRIMINATION INTERFERENCE BETWEEN GROUPS OF RESOLVED HARMONICS

### A. Methods

#### 1. Stimuli

The basic stimuli in this study were harmonic complex tones with a nominal  $F_0$  of 250 Hz. Each complex was passed through a digital bandpass filter (8th-order Butterworth) with fixed 3-dB cutoff frequencies of 125–625 Hz (lower spectral region) or 1375–1875 Hz (higher spectral region). These stimulus parameters were chosen, following Shackleton and Carlyon (1994) and Gockel *et al.* (2004), such that the complexes always contained peripherally resolved harmonics in their passbands. As a rule of thumb, it is usually considered that harmonics below about the 10th are resolved; for the lowest average target  $F_0$  tested in this experiment (3 semitones below 250 Hz), the higher spectral region contained the 7th, 8th, and 9th harmonics between the 10-dB cutoff points of the filter.

The interferer  $F_0$ , relative to which the two target  $F_0$ 's on a trial were set, was drawn randomly from a 6-semitone-wide uniform distribution centered on (i.e.,  $\pm 3$  semitones around) 250 Hz. The large rove range, which was roughly twice that used by Gockel *et al.* (2004; 2005a), helped to ensure that listeners could not profitably base their judgments on the absolute  $F_0$  of the target in any one interval, and instead had to compare the  $F_0$ 's of the targets in the two observation intervals presented within each trial. To test whether the use of a larger roving range had a significant impact on performance, and to facilitate comparisons with

the results of Gockel *et al.* (2004, 2005a), we retested the listeners in one condition of the first experiment using a smaller roving range (3 semitones, which is approximately 19% of the lowest  $F_0$  in the range).

Each harmonic complex was 400 ms in duration, including 20-ms raised-cosine on and off ramps. The harmonics all started in 0-deg (sine) phase, and those not attenuated by the bandpass filter had a level of 45 dB SPL. The harmonics were summed and presented in a background of pink noise with a spectrum level of 12 dB ( $re: 20 \mu\text{Pa}$ ) at 1 kHz [compared to approximately 15 dB SPL in Gockel *et al.*'s (2004) study], and an upper cutoff frequency of 20 kHz. The purpose of this noise was twofold: (1) to prevent listeners from relying on within-channel cues that could have arisen from interactions of components from both spectral regions within auditory filters centered between the two regions; and (2) to mask any possible distortion products. The noise was turned on 300 ms before the onset of the first tone and off 300 ms after the offset of the second tone on each trial, giving it a total duration of 1.9 s. On each trial, a fresh noise sample was obtained by cutting a randomly selected 1.9-s portion of a 10-s burst of pink noise that was generated at the start of each experimental block.

The stimuli were played out with 24-bit resolution at a 50-kHz sampling rate via a LynxStudio LynxOne sound card. After being passed to a headphone buffer (TDT HB6, Tucker Davis Technologies), they were delivered to the listener's left ear through the left earpiece of a Sennheiser HD 580 headset.

#### 2. Procedure

This experiment used a two-interval, two-alternative (2I-2AFC), constant-stimuli procedure. On each trial, two "target" tones were presented successively, separated by a 500-ms interstimulus interval. One of the two complexes had an  $F_0$  of  $F_{0\text{REF}} + \Delta F_0/2$ , and the other an  $F_0$  of  $F_{0\text{REF}} - \Delta F_0/2$ , where  $F_{0\text{REF}}$  was the average  $F_0$  of the two targets and  $\Delta F_0$  was equal to 0.4375%, 0.875%, 1.745%, 3.5%, 7%, 14%, 20%, or 40% of  $F_{0\text{REF}}$ , depending on the condition being tested. The order of presentation of the lower- and higher- $F_0$  tones was randomly chosen on each trial with equal probability. The listener's task was to indicate which interval contained the target with the higher  $F_0$ . Visual feedback was provided after each trial.

The target complexes, which listeners had to discriminate, were either presented alone or gated synchronously with another group of harmonics, which constituted the interferer. The  $F_0$  of the interferer was arithmetically centered between the  $F_0$ 's of the two targets ( $F_{0\text{REF}}$ ) and was always the same in both intervals of a trial, so that it provided no relevant information for the task at hand. In this first experiment, the target complex was always filtered into the higher spectral region (1375–1875 Hz), and the interferer was always in the lower region (125–625 Hz). The  $d'$  values reported in this article are based on 300 trials per listener per condition. The different test conditions resulting from the combination of the  $\Delta F_0$  and target absent/present factors were tested in blocks of 50 trials. Within a given block, the stimulus parameters remained constant. The testing order was pseudorandom, with the constraint that groups of blocks

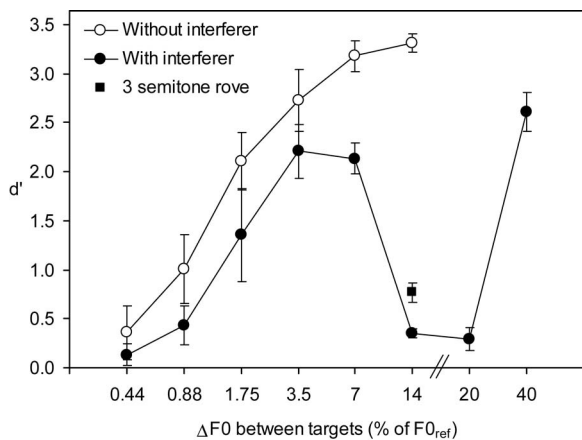


FIG. 1.  $F0$  discrimination performance for a target complex filtered between 1375 and 1875 Hz as a function of the  $F0$  difference between the two target tones,  $\Delta F0$ . The latter is expressed in percent of the reference  $F0$  around which the two target  $F0$ 's were symmetrically placed. Empty symbols show performance for the target alone; solid symbols show performance in the presence of a simultaneous interferer complex, which was filtered into a lower spectral region (125–625 Hz). The solid square corresponds to performance measured using a smaller across-trial  $F0$  roving range of 3 semitones (approximately 19%), instead of 6 semitones (approximately 41%). The error bars indicate standard errors of mean across listeners.

in which the interferer was absent were run before groups of blocks with the interferer, so that listeners were given the opportunity to listen to the target in isolation before they were tested with the interferer.

### 3. Listeners

Five listeners took part in this experiment. All had normal hearing, defined by absolute pure-tone thresholds lower than 15 dB HL at octave frequencies between 250 and 8000 Hz, as measured using a Madsen AC30 audiometer and TDH49 earphones. Listeners were given the opportunity to practice the task with and without interferers, until their performance remained stable across blocks of trials. In addition, four of the five listeners already had over 30 h of practice in experiments involving pitch perception, using harmonic complexes similar to those used here.

### B. Results

The results of this experiment are shown in Fig. 1. In the absence of the interferer (open symbols) performance ( $d'$ ) increased monotonically with  $\Delta F0$ , as expected. At the smallest  $\Delta F0$  tested (0.44%), the mean  $d'$  was not significantly different from zero [one-sample t-test,  $t(4)=1.29$ ,  $p=0.265$ ], indicating chance performance. At the 14%  $\Delta F0$ , performance was essentially at ceiling.<sup>1</sup> Preliminary results showed that performance remained at ceiling at  $\Delta F0$ 's larger than 14% in the no-interferer condition. Therefore, performance at the larger  $\Delta F0$ 's of 20% and 40% was not measured formally.

Performance in the presence of the interferer varied non-monotonically with  $\Delta F0$ , increasing from not significantly different from chance at the 0.44% [ $t(4)=0.136$ ,  $p=0.302$ ] and 0.875% [ $t(4)=0.432$ ,  $p=0.092$ ]  $\Delta F0$ 's, up to well above chance at the 3.5% and 7%  $\Delta F0$ 's, and then showing a marked dip in performance at  $\Delta F0$ 's of 14% and 20%. In

fact, at a  $\Delta F0$  of 20%, performance was not significantly above chance [ $t(4)=0.292$ ,  $p=0.062$ ]. At the largest  $\Delta F0$  tested (40%), performance in the presence of the interferer was well above chance, but still below ceiling.

In order to test whether the interferer had a significant effect overall, a two-way repeated-measures analysis of variance (RMANOVA) was performed, with the interferer (present/absent) and  $\Delta F0$  as within-subject factors, and the  $d'$  values from each subject in each condition as the dependent variable. The results showed significant main effects of interferer [ $F(1,4)=321.92$ ,  $p<0.001$ ] and  $\Delta F0$  [ $F(4,7)=25.54$ ,  $p=0.001$ ], and an interaction between these two factors [ $F(7,28)=28.91$ ,  $p<0.001$ ]<sup>2</sup>. Planned comparisons (multiple paired t-tests) further revealed a significant interferer effect ( $p<0.05$ ) at all  $\Delta F0$ 's except for 0.44% (where performance was at chance with and without an interferer) and, unexpectedly, for 3.5%.

The solid square in Fig. 1 represents the average performance measured in the same listeners using a smaller across-trial  $F0$  roving range of 3 semitones (approximately 19%), instead of 6 semitones (approximately 41%), with the interferer present. As can be seen, the results of this control condition confirm the previous observation of poor performance at the 14%  $\Delta F0$ , indicating that this effect was not simply due to our use of a large  $F0$  roving range. As expected, performance with the smaller rove range in the absence of the interferer remained near ceiling and is not shown here.

### C. Discussion

The results demonstrate significant interference effects in pitch discrimination between two groups of resolved harmonics filtered into nonoverlapping spectral regions. This finding complements the earlier results of Gockel *et al.* (2004, 2005a), which demonstrated PDI with target complexes containing only unresolved harmonics. These results raise two important questions, one general and one specific: First, why does PDI occur between two groups of resolved harmonics? Second, why did performance in the presence of the interferer vary nonmonotonically with  $\Delta F0$ , with a marked dip in performance at  $\Delta F0$ 's of 14%–20%?

Gockel *et al.* (2004) suggested that PDI was caused by the more salient pitch of the interferer's resolved harmonics "swamping" the weaker pitch produced by the unresolved harmonics of the target. This explanation cannot account for the present results because both the target and the interferer complexes consisted primarily of resolved harmonics. However, it is still conceivable that the interferer evoked a more salient or dominant pitch than the target, because it was filtered into a lower spectral region, and thus contained lower-numbered harmonics in its passband. Consistent with this hypothesis, Gockel *et al.* (2005b) measured the dominance of individual harmonics for an  $F0$  of 250 Hz (corresponding to that used here), and found that, for 200-ms-long complexes, harmonics 1 to 4 were the most dominant. Dai (2000) also found that for  $F0$ 's of 200 and 300 Hz (comparable to the 250-Hz  $F0$  used here), harmonics 1 through 4 were most dominant, and that the contribution of harmonics above the 6th to the overall pitch was not significant. Here, the inter-

ferer, being filtered between 125 and 625 Hz, typically contained the first three harmonics, while the target, being filtered between 1325 and 1875 Hz, typically contained harmonics between the 5th and the 8th. If the PDI effects observed in this experiment can be accounted for in terms of the relative dominance of harmonics, then exchanging the spectral regions of the target and interferer should lead to a large reduction or elimination of PDI, because then the target would occupy the region with the most dominant harmonics. This prediction was tested in experiment 2.

The answer to the second question, regarding the dip in performance at  $\Delta F0$ 's of about 14%–20%, may be related to the relationships between individual harmonics in the two targets. Generally, harmonics between 5 and 8 were present in the target passband. The frequency differences between harmonics 5 and 6 (20%), 6 and 7 (17%), and 7 and 8 (14%) correspond to the  $F0$  differences at which performance dipped. As the  $F0$  ratio between two identically bandpass-filtered harmonic complexes approaches the frequency ratio between consecutive harmonics in the stimulus spectrum, the frequencies of the harmonics of rank  $n$  and  $n+1$  in the higher- $F0$  complex become close to those of the harmonics of rank  $n+1$  and  $n+2$  in the lower- $F0$  complex. This has two consequences. The first is that the frequency distance between the harmonics of the two complexes reaches a local minimum. This is illustrated in the top panel of Fig. 2, which shows the mean absolute frequency distance between closest-neighbor components across the two targets, as a function of  $\Delta F0$ . As can be seen, this distance first increases with  $\Delta F0$  below 7%, but then decreases to reach a local minimum at the 14%  $\Delta F0$ . The depth and width of this “trough” depend to a large extent on which components are included in the calculation, and how these components are weighted in computing the mean frequency shift. The data shown in Fig. 2 were obtained by considering only components that fell within the 10-dB bandwidth of the higher-region filter, and assuming for simplicity that these components received equal weights. With different assumptions, more or less marked decreases in the mean distance can be observed at the 14% and 20%  $\Delta F0$ 's.

The second consequence is that, as  $\Delta F0$  exceeds half the percentage difference between the two harmonics in the target passband, the pairs of harmonics whose frequencies are most similar between the two targets have different ranks. This produces a decoupling between the direction of the shift in  $F0$  and the shifts in the frequencies of individual harmonics. For instance, when  $F0_{\text{REF}}$  was 250 Hz and  $\Delta F0$  was 14%, the higher- $F0$  target contained harmonics with frequencies of 1605 Hz (6th harmonic) and 1873 Hz (7th harmonic), while the lower- $F0$  target contained harmonics with frequencies of 1628 Hz (7th harmonic) and 1860 Hz (8th harmonic). Thus, a listener basing his/her decisions on the direction of the frequency shifts of individual harmonics in the higher-region passband, or on the average frequency of these harmonics (i.e., the spectral centroid of the target), would perform least accurately for  $\Delta F0$ 's of 14%–20%. This is illustrated in the middle panel of Fig. 2, which shows the proportion of correct responses achieved by an “ideal” (virtual but nonoptimal) observer whose decisions are based on

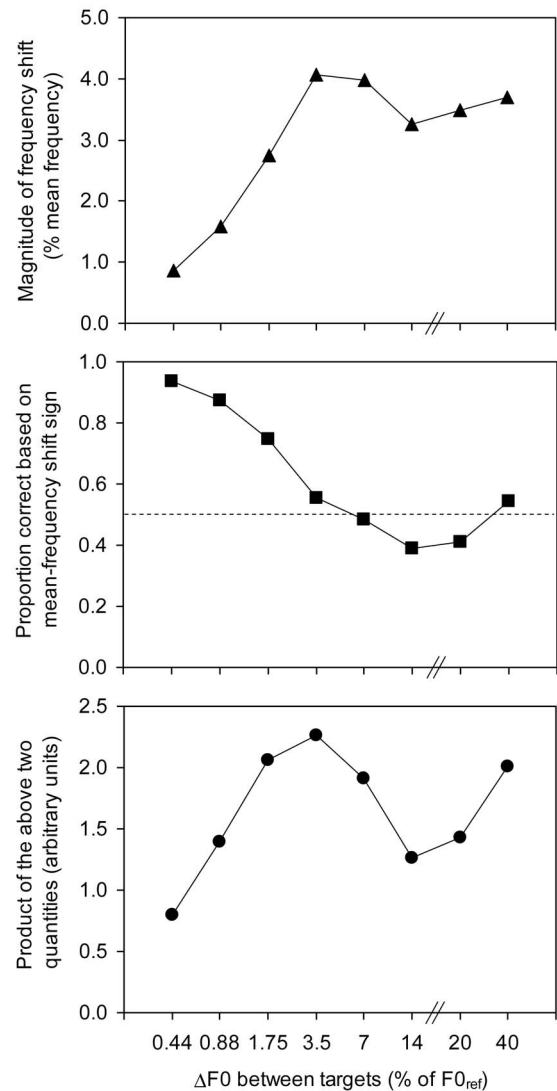


FIG. 2. Variations in the magnitude and sign of the mean frequency distance between closest-neighbor components across the two higher-region targets as a function of the  $\Delta F0$  between them. The upper panel shows the absolute value of the difference between the mean frequencies of two sets of harmonics falling inside the 10-dB bandwidth of the higher-region filter, one set having an  $F0$  of  $F0_{\text{REF}} - \Delta F0/2$ , and the other, an  $F0$  of  $F0_{\text{REF}} + \Delta F0/2$ .  $F0_{\text{REF}}$  was varied over a 6-semitone range, with a uniform distribution, as in the main conditions of experiment 1. The middle panel shows the proportion of cases in which the frequency difference was the same as that of the  $F0$  shift between the first and second targets. This corresponds to the proportion of correct responses for an observer basing his decisions on the sign of the mean frequency difference between the two sets of target components. Chance performance, corresponding to a probability of 0.5, is indicated by the vertical dashed line. The bottom panel simply shows the product of the two above quantities, reflecting a combination of information on both the magnitude and the sign of the shift in the mean frequency of the target components.

the direction of the shift in the mean frequency of the harmonics in the higher region. As can be seen, this proportion reaches a minimum for  $\Delta F0$ 's of 14%–20%.

The predicted proportions of correct responses shown in the middle panel of Fig. 2 are based on simple calculations that do not take into account a number of factors that could affect the performance of real listeners. Clearly, these predictions are not entirely consistent with the experimental data

shown in Fig. 1. For instance, at  $\Delta F0$ 's of 14%–20%, the predicted proportion of correct responses falls below the 0.5 chance line, suggesting that, at these  $\Delta F0$ 's, an observer basing decisions solely on the mean spectral positions of the components in the passband would perform *below* chance. The important point here is that the trend for worst performance at  $\Delta F0$ 's of 14%–20% is qualitatively consistent with the data in Fig. 1. More importantly, the predictions shown in the middle panel of Fig. 2 were derived assuming perfect encoding of the component frequencies. The performance of real listeners is undoubtedly limited by imperfections in the ability to encode stimulus frequencies, and should therefore be expected to increase with the magnitude of the frequency difference between the two observation intervals plotted in the upper panel of Fig. 2. As shown in the bottom panel of Fig. 2, taking both the magnitude and the sign of the frequency shifts into account (which was done here in a crude way, by multiplying the quantity in the top panel by that in the middle panel), one obtains predictions that show the same general qualitative trends as the data in Fig. 1, with an increase in performance with  $\Delta F0$  up to 7% and a dip in performance around 14%–20%.

These considerations suggest a first possible explanation for the dip in performance observed at the 14%–20%  $\Delta F0$ 's in the presence of the interferer. According to this explanation, the interferer may have forced (or at least encouraged) listeners to rely on the spectral positions of the individual harmonics or on the overall spectral location, rather than on the  $F0$  of the target. Earlier studies have shown that in pitch discrimination tasks with small numbers of harmonics, many listeners rely on changes in the spectral position of the harmonics, rather than on changes in the  $F0$  *per se* when no feedback is provided (Smooenburg, 1970; Laguitton *et al.*, 1998; Schneider *et al.*, 2005). Here, this listening strategy may have been promoted by the interferer occupying a lower region, and evoking a more salient pitch, than the target. From this point of view, it appears that, at least in the case of small  $\Delta F0$ 's (below about 3.5%), the measured amount of PDI may represent a lower bound on the effect, because performance could have been even poorer had subjects not been able to use spectral shifts as an alternative cue to the  $\Delta F0$ , in the presence of the interferer.

Another possible explanation for the large PDI effects observed at the 14%–20%  $\Delta F0$ 's, and the fact that at these  $F0$ 's performance in the presence of the interferer dropped to (or close to) chance, is that the internal representation of the target  $F0$  in the central auditory system was almost completely “suppressed” by the more dominant interferer. Specifically, it is conceivable that in a situation where two groups of harmonics compete in determining the perceived pitch, a “harmonic sieve” tuned to the  $F0$  of the dominant group is applied within the spectral region of the other group, and that harmonics that fall in between the teeth of the harmonic comb in that region are suppressed in order to avoid perceptual ambiguity. Under this “harmonic-cancellation” hypothesis, the suppression should be most effective in the situation where the target harmonics in the higher region fall almost exactly in between the comb teeth, as was the case when  $\Delta F0$  was around 14%–20%. Whether the poor perfor-

mance in the presence of the interferer at  $\Delta F0$ 's of 14%–20% is due to listeners relying on misleading spectral cues or to the  $F0$  of the dominant low-region harmonics suppressing the pitch of the harmonics in the higher region, the corresponding PDI effects should be eliminated when the spectral positions of the target and interferer are swapped, so that the target is in the lower spectral region because (a) the harmonics in that region are dominant, and (b) the spectral shifts in that region are coherent with the  $F0$  shifts.

### III. EXPERIMENT 2: PDI WITH THE TARGET IN A LOWER SPECTRAL REGION THAN THE INTERFERER

#### A. Methods

The procedure and listeners were the same as in the previous experiment. The only difference was in the stimuli: the spectral regions of the target and interferer were swapped, so that the target was now filtered into the lower spectral region (125–625 Hz) while the interferer was filtered into the higher region (1375–1875 Hz). Four listeners took part in this experiment, and only the 6-semitone rove range was tested. Because data from these four listeners were also collected under the conditions described in experiment 1, the data from both experiments could be compared using a within-listeners design. Importantly, the order of data collection was counterbalanced, such that two of the listeners completed experiment 2 before embarking on experiment 1, and the other two completed the two experiments in the opposite order. This, combined with the fact that in both experiments listeners' performance was stable across blocks of trials, makes it very unlikely that the differences in results between the two experiments were due to practice effects.

#### B. Results

The results of this experiment are shown in Fig. 3, as squares connected by solid lines. To facilitate a comparison with experiment 1, the data from Fig. 1 are replotted as dotted and dashed lines. Consider first the results obtained in this experiment without the interferer present (solid lines). The results are virtually indistinguishable from those obtained in the comparable conditions of experiment 1 (dashed lines); this was confirmed by the results of a two-way RMANOVA on the  $d'$  data, with spectral region (target in high or low region) and  $\Delta F0$  as factors, which revealed no main effect of spectral region [ $F(1, 4)=0.001$ ,  $p=0.972$ ] and no interaction [ $F(5, 20)=0.451$ ,  $p=0.808$ ]. This outcome is consistent with earlier results showing similar  $F0$  discrimination thresholds for complexes filtered into these spectral regions, as long as they contained resolved harmonics (e.g., Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994). As  $F0$  difference limens are often taken as an indirect measure of pitch salience, it suggests that the salience produced by the resolved harmonics in the upper region was the same as that produced by the resolved harmonics in the lower region, at least when presented in isolation.

Consider next the effect of adding the interferer (compare empty and filled symbols connected by solid lines). Unlike in the previous experiment, where large PDI effects were

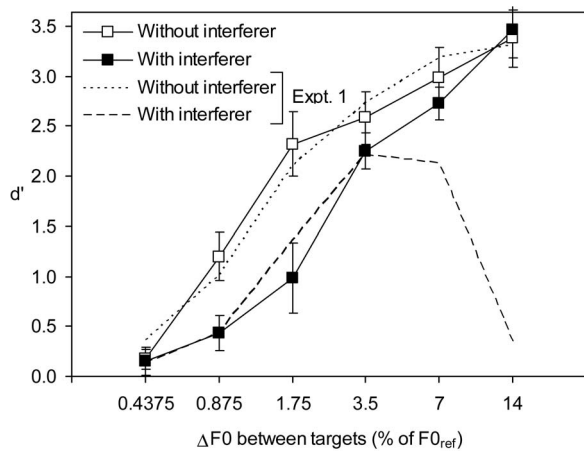


FIG. 3.  $F0$  discrimination performance for a target complex filtered between 125 and 625 Hz as a function of the  $F0$  difference between the two target tones,  $\Delta F0$ , in the absence and presence of a higher-region interferer (1375–1875 Hz).  $\Delta F0$  is expressed in percent of the reference  $F0$  around which the two target  $F0$ 's were symmetrically placed. The data from this experiment are shown as squares connected by solid lines. As in Fig. 1, empty symbols show performance for the target alone; solid symbols, performance in the presence of a simultaneous interferer complex. The error bars indicate standard errors of mean across listeners for the data of experiment 2. To facilitate comparison with experiment 1, the data from experiment 1 (Fig. 1) have been replotted here as dotted and dashed lines; to avoid clutter, the corresponding symbols and error bars are not shown here.

observed at  $\Delta F0$ 's larger than 7%, in this experiment, the interferer effect seems negligible at large  $\Delta F0$ 's. A two-way RMANOVA on these  $d'$  data showed no significant main effect of the interferer [ $F(1, 4)=3.33, p=0.142$ ], but a significant interaction between interferer and  $\Delta F0$  [ $F(1, 4)=5.037, p=0.004$ ]. Planned comparisons (paired t-tests) indicated that the only  $\Delta F0$ 's at which the interferer significantly decreased performance were the 0.875% [ $t=4.30, p < 0.05$ ] and 1.75% [ $t=2.94, p < 0.05$ ]  $\Delta F0$ 's. In fact, at these small  $\Delta F0$ 's, as apparent in Fig. 3, the effect of the interferer was similar to that observed in experiment 1—a three-way RMANOVA on the data collected at these two  $\Delta F0$ 's in experiments 1 and 2 showed no significant main effect of experiment [ $F(1, 4)=0.001, p=0.980$ ] and no significant interaction involving this factor.

### C. Discussion

The results of experiments 1 and 2 indicate that for  $\Delta F0$ 's of 7% or more, the influence of the interferer was heavily dependent on the relative spectral positions of the target and interferer: although the interferer had a large detrimental effect on performance when the target was filtered into the higher spectral region and the interferer into the lower spectral region (experiment 1), it had essentially no effect in the converse situation (experiment 2). This outcome is consistent with the possible explanations described earlier, according to which the PDI effects at  $\Delta F0$ 's larger than 3.5% were due to the pitch evoked by the more dominant harmonics in the lower-region suppressing (partially or completely) that of the harmonics in the higher region, and/or to listeners basing their judgments on spectral shifts in the higher region.

In contrast, for  $\Delta F0$ 's less than 3.5%, the interferer was found to have a significant influence regardless of whether it

occupied the lower or the higher spectral region. This outcome indicates that, at these small  $\Delta F0$ 's, either PDI was mediated by mechanisms that depend little on harmonic dominance, or the harmonics in the higher region were not appreciably less dominant in determining pitch than those in the lower region. In an attempt to distinguish between these two possibilities, the final experiment used a pitch-matching procedure to provide a more direct test of what pitch(es) listeners were perceiving when listening to the combination of target and interferer.

## IV. EXPERIMENT 3: PITCH MATCHES

### A. Rationale

The various possible interpretations for the results of the previous two experiments provide some basis for predicting the pitch percepts produced by the experimental stimuli. This experiment used a pitch-matching procedure to gather more direct information regarding the pitches evoked by the simultaneous lower- and higher-region complexes as a function of their  $F0$  separation. Although Shackleton and Carlyon (1994) collected pitch matches using harmonic complexes filtered into the same spectral regions as those used in the present study, their complexes were not presented simultaneously. It is possible that the value and salience of the pitch evoked by one complex is influenced by the simultaneous presence of the other complex, and that PDI is related not to the relative pitch saliences of the two complexes measured in isolation, but to their relative saliences when presented simultaneously. Accordingly, in this experiment, we sought to measure the pitch and relative pitch salience evoked by each group while the other group was present. Listeners adjusted the  $F0$  of a single group of harmonics, which was filtered into the lower or the higher spectral region, to match the pitch of a group of harmonics, which was filtered into the same spectral region but was presented *simultaneously* with a group filtered into the other spectral region. This was done for three different  $F0$  separations between the two simultaneous groups: a “small” separation (1.75%), a large separation (20%), and an intermediate separation (7%).

### B. Method

This experiment used the same basic stimuli as experiments 1 and 2, with 400-ms harmonic complexes filtered into the lower and higher spectral regions, as described above. Each trial of the pitch-matching procedure involved the successive presentation of two sounds. The first, referred to as the “reference” stimulus, consisted of two simultaneous groups of harmonics filtered into the lower and the higher spectral regions, exactly like the target and the interferer in experiment 1. The  $F0$  of the higher-region group was always set relative to that of the group in the lower region, which was roved over a 6-semitone range, as in experiments 1 and 2. Three  $F0$  separations between the two groups were tested: 1.75%, 7%, and 20%; these three  $F0$  separations correspond to those between the target and the interferer for  $\Delta F0$ 's (between the two targets) of 3.5%, 14%, and 40%, respectively, in experiment 1. Both positive and negative  $F0$  separations were tested, meaning that in one set of conditions the  $F0$  of

the group filtered into the higher spectral region was higher than the  $F_0$  of the group filtered into the lower spectral region, while in the other condition, it was lower. Note that this is similar to the situation in experiment 1, where the  $F_0$  of the higher-region target was higher than that of the lower-region interferer in one observation interval, and lower in the other one. The different  $F_0$  separations were tested in random order.

The reference stimulus was followed, after a 500-ms gap, by a “comparison” tone, which was filtered into a single spectral region. In one condition, that region was the lower one; in the other condition, it was the higher one. These two conditions were tested in separate blocks. At the beginning of each matching procedure, the  $F_0$  of the comparison tone was selected randomly from a  $\pm 18$ -semitone interval around the 250-Hz nominal  $F_0$ , in 2-semitone steps.

Listeners were instructed to adjust the pitch of the comparison tone so that it matched as closely as possible the dominant pitch evoked by the stimulus in the first interval. They were informed that the reference stimulus might evoke more than one pitch and that when this happened, they should try to match the pitch that was the most salient to them. By presenting the comparison tone in just one region, we hoped to encourage listeners to attend more selectively to that spectral region in the reference tone. Listeners increased or decreased the  $F_0$  of the comparison tone at will, using six virtual buttons on the computer screen: two large buttons for  $\pm 4$ -semitone (about 26%) steps, two medium-sized buttons for  $\pm 1$ -semitone (about 6%) steps, and two small buttons for  $\pm 1/4$ -semitone (about 1.45%) steps. A “hear again” button was also available. After each button press, the reference and comparison stimuli were regenerated with the appropriate parameters. The stimulus  $F_0$  was roved over a 6-semitone range from one stimulus presentation to the next, keeping the relative  $F_0$  difference between the reference and the comparison tone as last adjusted by the listener. This across-trial roving of  $F_0$  had two objectives: first, to be consistent with experiments 1 and 2, and second, to encourage listeners to compare the two sounds presented on each trial despite their different timbre, rather than rely on some longer-term memory trace.

Each listener performed at least 20 matches in each of the 12 test conditions (six  $F_0$  separations  $\times$  two comparison-tone filtering-region conditions). At the end of each match, the final value assigned to the  $F_0$  of the comparison tone relative to the  $F_0$  of the nominal  $F_0$  (250 Hz) was stored. Pitch-match histograms were produced by counting the number of matches at each  $1/4$ -semitone step over a 2-octave range (i.e.,  $\pm 1$  octave) around the nominal  $F_0$ . In each listener and for each condition, the pitch-match histograms were normalized by dividing the counts in each  $1/4$ -semitone bin by the total number of matches over the whole 2-octave range. In this way, the height of the bars in the resulting histograms directly indicates the proportion of matches made at the corresponding relative frequency, out of the total number of matches for that condition and listener.

In order to estimate the proportion of matches made to the  $F_0$  of the lower- and higher-region groups, the pitch-match histograms obtained in each condition were modeled

as a mixture of two Gaussians, defined each by three free parameters: mean, amplitude, and standard deviation. This was particularly useful for the smallest  $F_0$  separation tested (i.e., 1.75%), as this separation was similar in size to the measurement resolution ( $1/4$ th of a semitone is approximately equal to 1.43%). The best-fitting parameters were determined using a maximum-likelihood procedure, based on the assumption that the proportions of pitch matches in each frequency bin were distributed binomially. The maximum of the likelihood hypersurface was found using the Nelder-Mead simplex algorithm in MATLAB.

Three of the four listeners who had taken part in experiments 1 and 2 were involved in the current experiment. This experiment was performed after experiments 1 and 2.

### C. Results

Pitch-match histograms based on the pooled data of the three listeners who took part in this experiment are shown in Fig. 4. When the comparison tone was filtered into the lower spectral region (left-hand column), the pitch matches were concentrated near the  $F_0$  of the lower-region group in the reference stimulus, which corresponds to the zero mark on the abscissa: 94% of the matches fell within a  $\pm 1/4$ -semitone range around the lower-region reference  $F_0$  at the 1.75% separation (top panel), 80% at the 7%  $F_0$  separation (middle panel), and 92% at the 20% separation (bottom panel). Furthermore, at the 7% and 20% separations, the vast majority (i.e., over 90%) of the matches was closer to the  $F_0$  of the lower-region reference group than to that of the higher-region group. For the 1.75% separation, determining what proportion of the matches was made to each group is less straightforward, due to the similar size of that  $F_0$  separation and the measurement accuracy, as mentioned above. At that separation, the histogram bars on either side of the central bin may reflect inaccurate matches to the  $F_0$  of the lower group and/or matches to the  $F_0$  of the higher group. If it is assumed that the listeners sometimes matched the  $F_0$  of the lower group and sometimes that of the higher group, the proportion of matches corresponding to each group may be estimated by fitting the distribution of matches with a mixture-of-Gaussians with means equal to the  $F_0$ 's of the lower- and higher-region groups, as described in the Methods section. Doing this, we found that the estimated proportion of matches to the  $F_0$  of the lower group was about 66% although, because the two distributions are so close, the proportion of matches assigned to each  $F_0$  remains highly uncertain. Nevertheless, Fig. 4 contains an indication that the higher-region complex had some influence on the listeners' responses at the 1.75%  $F_0$  separation: comparing the heights of the empty and filled histogram bars closest to the two vertical dashed lines in the upper left-hand panel, it can be seen that for negative relative  $F_0$  values (i.e., the dashed line on the left), the empty bar, which corresponds to the case where the  $F_0$  in the higher region was lower than that in the lower region, has a higher amplitude than the filled bar, which corresponds to the case where the  $F_0$  in the higher region was higher than that in the lower region, while for positive relative  $F_0$  values (i.e., the dashed line on the right),

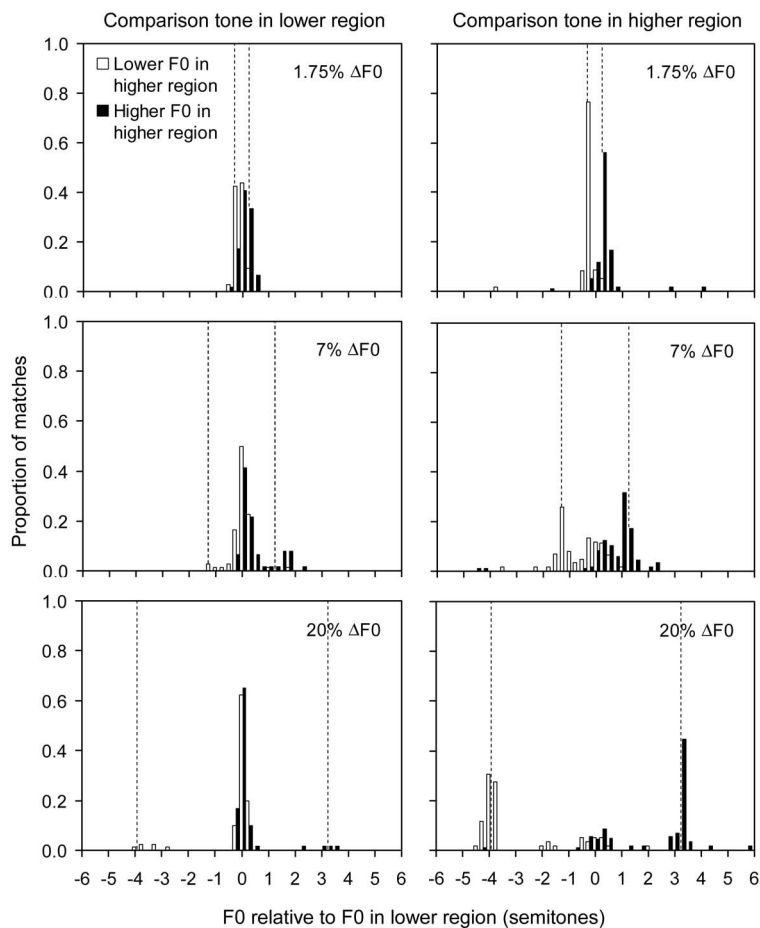


FIG. 4. Pitch-match histograms. The panels in the left-hand column show pitch matches obtained with the comparison tone filtered into the lower spectral region; those in the right-hand column, pitch matches obtained with the comparison tone in the higher spectral region. Each row corresponds to a given  $F_0$  separation between the lower- and higher-region complexes in the reference stimulus: 1.75% (top), 7% (middle), and 20% (bottom); these correspond to the  $F_0$  separations between the target and the interferer in the 3.5%, 14%, and 40%  $\Delta F_0$  conditions of experiment 1. The vertical dashed lines indicate the position of the  $F_0$  of the higher-region complex in the reference stimulus, relative to that of the lower-region complex; the latter always corresponds to zero on the abscissa. The relative  $F_0$  values on the abscissa are expressed in semitones; only matches falling between  $-6$  and  $+6$  semitones (corresponding, roughly, to  $\pm 35\%$ ) around the reference (lower-region)  $F_0$  are shown.

the converse is true. This suggests that, even if listeners always matched the  $F_0$  of the complex in the lower spectral region, the perceived pitch of that complex was shifted towards the  $F_0$  of the higher-region group. Although this shift of the perceived  $F_0$  of the lower-region group in the direction of that of the higher-region group appears to be modest, it may have contributed to the drop in performance produced by the interferer at small  $\Delta F_0$ 's in experiment 2.

When the comparison tone was filtered into the higher spectral region (right-hand column of Fig. 4), the majority of the pitch matches was close to the  $F_0$  of the corresponding higher-region group in the reference stimulus, which is indicated by the vertical dashed lines. However, a substantial proportion of matches was also made to the  $F_0$  of the lower-region group, as indicated by the central peaks (at 0 semitones) in the middle and lower right-hand panels. At the two corresponding  $F_0$  separations of 7% and 20%, the proportions of matches to the  $F_0$  of the lower region were estimated to be 48% and 25%, respectively, indicating that listeners tended to make more matches based on the  $F_0$  of the lower-region group when the  $F_0$  separation between the two groups was 7% than when it was 20%. In both cases, substantially more “across-region” matches were made in conditions where the comparison tone was filtered into the higher spectral region: less than 10% of the matches were across regions when the comparison tone was in the lower spectral region, and the maximum-likelihood mixture-of-Gaussian fits imputed only 3%–4% of matches to that  $F_0$ . At

the 1.75% separation (top-right panel), no central peak can be distinguished. However, this could be due to peak overlap, and should not be taken as evidence that listeners never matched the  $F_0$  of the lower-region group in this condition.

#### D. Discussion

The main finding from this experiment is the asymmetry in the pitch matches at the two larger  $F_0$  separations, depending on which spectral region contained the comparison tone. When the comparison tone was in the lower spectral region, listeners usually matched its  $F_0$  to that of the reference group filtered into the corresponding (lower) spectral region, and across-region matches were rare (less than 10% overall). However, when the comparison tone was in the higher region, listeners made a substantial proportion of matches (between about 20% and 44%) to the  $F_0$  of the lower-region group. This asymmetry was observed at both the 7% and 20%  $F_0$  separations, and is consistent with the hypothesis that when both groups were presented simultaneously, the pitch of the lower-region group was more salient or dominant than that of the higher-region group, even when the higher-region group was emphasized by the presence of the comparison tone.

For the 1.75% separation, things are less clear because, at this small separation, the relative inaccuracy of the pitch matches and the limited measurement resolution make it difficult to determine to which  $F_0$  listeners were matching.

Therefore, the data do not permit a strong conclusion regarding the relative salience of the two groups of harmonics at small  $F_0$  separations.

## V. GENERAL DISCUSSION

### A. Summary of results

The results from experiments 1 and 2 show that listeners' ability to discriminate changes in the  $F_0$  of a group of resolved harmonics can be significantly impaired by a simultaneously presented group of resolved harmonics in a different (nonoverlapping) spectral region. The results extend the earlier findings of Gockel *et al.* (2004, 2005a) by showing that PDI is not restricted to situations in which the target contains only unresolved harmonics. The pitch matches of experiment 3 showed that the harmonics in the lower spectral region seemed to produce a somewhat more salient pitch than those in the higher spectral region. This difference in salience or dominance may explain why, at larger  $\Delta F_0$ 's, PDI was only found when the target was in the higher spectral region (experiment 1). However, the differences in pitch salience are more difficult to reconcile with the fact that the PDI observed at smaller  $\Delta F_0$ 's seemed to be independent of whether the spectral region of the target was higher or lower than that of the interferer.

### B. Correspondence between performance and introspection: Are PDI and perceptual organization related?

When prompted, listeners typically reported that the simultaneous target and interferer were perceived as two separate sounds only when the  $F_0$  separation between them was large, corresponding to about 20% ( $\Delta F_0$  of 40% in our experiments) or more. At  $F_0$  separations between about 3.5% and 10% (i.e.,  $\Delta F_0$ 's between 7% and 20%), most listeners reported hearing only one sound, but did indicate that the stimuli sounded less "fused" than they did at smaller  $F_0$  separations, below 3.5%.

These informal observations on how the stimuli sound bear some relationship to the pattern of results observed in our formal experiments. At very small  $F_0$  separations, where listeners reported generally hearing a single fused sound, PDI was similar whether the spectral region of the target was above or below that of the interferer (Fig. 3). This may occur because the auditory system integrates the information from both regions and calculates the  $F_0$  based on a sum of the information from all the components (e.g., Goldstein, 1973). However, in the present case, the weights assigned to the different harmonics appear not to be determined by the relative dominance of these harmonics (as estimated in other studies, e.g., Moore *et al.*, 1985; Dai, 2000), because PDI was found not to depend critically on the relative spectral positions of the target and interferer. Alternatively, at these small  $F_0$  separations, PDI could be related to the unitary pitch evoked by the simultaneous target and interferer being less well defined than that evoked by the target alone. This could itself be due to the internal representations of the neighboring  $F_0$ 's merging into a broader peak—an effect that can be observed in the summary autocorrelograms pro-

duced by Meddis and O'Mard's (1997) model in response to stimuli consisting of two simultaneous sets of harmonics with different  $F_0$ 's: as the  $F_0$ 's become closer, the two corresponding peaks in the summary autocorrelogram start to overlap, forming a single, broader (less well-defined) peak.

At very large  $F_0$  separations (e.g., 20%) between the target and the interferer, the  $F_0$  difference between the two simultaneous groups is probably sufficient to overcome grouping based on common onset and offsets, allowing listeners to hear two separate pitches corresponding to the two  $F_0$ 's present. If so, any residual PDI observed at that large  $F_0$  separation (when the target was in the upper spectral region) may have been due to the listeners confusing the pitch of the interferer with that of the target. The large roving range used in the experiments, and the fact that the interferer  $F_0$  was always intermediate to the two target  $F_0$ 's, may have exacerbated this effect. This could explain why larger PDI effects were observed at these large  $F_0$  separations between target and interferer here than in Gockel *et al.*'s (2004) first experiment. In that experiment, the  $F_0$  separation between target and interferer was controlled by increasing the interferer  $F_0$  above that of the target, so that at large separations the targets had a consistently lower pitch than the interferer; this perhaps made it easier for listeners to focus on the target pitch. Here, listeners could not adopt the strategy of listening consistently for the higher or the lower pitch, because the target had a higher pitch than the interferer in one interval, but a lower pitch in the other, and  $F_0$  was roved across trials over a wide range.

At medium  $F_0$  separations ( $\Delta F_0$ 's of 14% and 20%), where the simultaneous target and interferer were generally heard as less fused than at smaller separations, yet not distinctly separated, strong PDI was found when the target was in the (less-dominant) higher spectral region (experiment 1), but no PDI remained when the target was in the (more-dominant) lower spectral region (experiment 2). As discussed above (cf. Sec. II C and Fig. 2), one possible explanation for the marked decrease in performance observed in the presence of the lower-region interferer at these  $\Delta F_0$ 's is that the interferer, evoking a salient pitch, may have encouraged listeners to rely on shifts in the frequencies of individual harmonics (or in the spectral centroid) in the higher region, rather than on pitch changes; as illustrated in Fig. 2, this would have led to poorest performance at  $\Delta F_0$ 's of 14% and 20%, as this is where spectral cues were most ambiguous. Another possible explanation for the large PDI effects at these  $\Delta F_0$ 's, which was also suggested above, is that the lower-region interferer to some extent "suppressed" the pitch that would have otherwise been evoked by the target. The finding in experiment 1, that performance in the presence of the interferer was at (or very close to) chance at these  $\Delta F_0$ 's, suggests that, in these conditions, the suppression was either complete or almost complete. This may appear to be inconsistent with the results of experiment 3, where it was found that even in the condition where the comparison group was in the lower region, listeners sometimes matched its  $F_0$  to that of the higher-region group, indicating that they perceived the pitch of the latter. However, it is important to note that the stimuli in experiments 1 and 3 were not exactly the

same. Whereas in experiment 3 the first stimulus interval contained a single group of harmonics, in experiment 1, when the interferer was present, both intervals contained both groups. The latter design may have made it more difficult for listeners to perceptually segregate the two groups, which in turn may have made it harder to “hear out” the pitch of the less salient higher-region group, and at the same time, promoted suppressive interactions between the two groups. Thus, the fact that listeners did sometimes match to the pitch of the higher-region complex in experiment 3 does not imply that they always perceived the pitch of the higher-region complex in experiment 1.

### C. Implications for models of pitch perception and source segregation

The results are consistent with models that integrate information across distant spectral regions in order to estimate pitch. However, they suggest that the channels on which pitch estimates are based cannot always be selected completely “at will,” based on the spectral region that they occupy. The results also suggest that  $F0$  differences between the target and interferer of more than 10% (i.e.,  $\Delta F0$ 's of more than 20%) are required for complete perceptual segregation based on  $F0$ . Although this difference appears rather large at first sight, particularly when compared with the mistuning of individual harmonics required to hear them as separate objects (Moore *et al.*, 1986), it is in fact compatible with experiments examining the effect of  $F0$  differences on speech segregation (Darwin *et al.*, 2003), where differences in  $F0$  between the target and interfering speech of at least 2 semitones (about 12%) were required for any benefit. This is different from the findings using two concurrent vowels (Scheffers, 1983; Assmann and Summerfield, 1990; Culling and Darwin, 1993); where a difference of 1 semitone (6%) already produces maximal segregation. However, it appears likely that within-channel cues, which were not available in our experiment, are important in mediating this effect (Assmann and Summerfield, 1994; Culling and Darwin, 1994).

### D. Remaining questions

Some difficulties remain in reconciling our results with those of earlier studies. For instance, at small  $F0$  separations, information from both spectral regions appears to be averaged, without regard to the pitch dominance region identified in earlier studies (Plomp, 1967; Ritsma, 1967; Moore *et al.*, 1985; Dai, 2000; Gockel *et al.*, 2005b). Another issue is that the mistuning of one harmonic can lead to its perceptual separation from the rest of the complex at frequency differences of between 1% and 3% (Moore *et al.*, 1986), whereas our results suggest that much larger differences are needed to hear out two  $F0$ 's in different spectral regions. Also, whereas a single mistuned harmonic can be heard as a separate entity while still contributing to the overall pitch of the complex (Moore *et al.*, 1986), here one of the two possible explanations put forward earlier would imply that the  $F0$ 's in two separate spectral regions cease to be integrated at *smaller*  $\Delta F0$ 's than may be required to hear the two regions as separate tones. As suggested by Gockel

*et al.* (2004), part or all of these differences may be due to our use of two spectral regions, separated by noise, as opposed to earlier studies' use of contiguous harmonics within a single spectral region. The spectral separation may make the auditory system more tolerant of (or less sensitive to) mistuning across the two regions, and more willing to accept the two regions as emanating from a single source based on other cues, such as common onset and offset. In this respect, the background noise, which effectively masks the spectral gap between the two regions, may play an important role in fusing the two regions by making spectral continuity plausible in the same way that noise can make temporal continuity plausible (Grose *et al.*, 2002; Darwin and Hukin, 2004; Darwin, 2005). The influence of the background masking noise in pitch perception experiments clearly deserves further investigation. Finally, the present results raise the possibility that listeners modified their strategy in the  $F0$  discrimination task, sometimes relying on  $F0$  information, and sometimes on spectral-pitch or timbre cues, depending on the condition. This could be further investigated by using target complexes that are altered to reduce the utility of spectral cues by, for instance, roving the number of the lowest harmonic present (e.g., Houtsma and Smurzynski, 1990).

### ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (NIDCD Grant R01 DC 05216) and was completed while the authors were in the Research Laboratory of Electronics at the Massachusetts Institute of Technology. The authors are grateful to Hedwig Gockel for the time taken in helping to improve the clarity and presentation of this manuscript. Alain de Cheveigné, John Grose, and an anonymous reviewer also provided helpful comments on earlier versions of the manuscript.

<sup>1</sup>Occasional cases where performance was 100% correct were dealt with by reducing the number of correct responses by one. With 300 trials, the highest possible proportion of correct responses after the application of this correction was approximately 0.997 (i.e., 299/300), which corresponds to a  $d'$  of approximately 3.8. As can be seen in Fig. 1, at the 14%  $\Delta F0$ , the average  $d'$  was around 3.5. This corresponds approximately to a proportion correct of 0.993, or a ratio of 298/300, indicating that at that separation, listeners made about two errors in 300 trials, on average. As this error rate was likely due to occasional attentional lapses, we consider that performance was essentially at ceiling in that condition.

<sup>2</sup>For this analysis, we assumed that the  $d'$  values corresponding to the no interferer condition at the 20% and 40%  $\Delta F0$  were the same as those measured in the same condition at the 14%  $\Delta F0$ . This allowed all levels of the  $\Delta F0$  factor to be included into the analysis. However, we also performed a similar analysis using only  $\Delta F0$ 's up to 14%, for which complete data were measured, both in the absence and in the presence of the interferer. This analysis also showed a significant main effect of interferer [ $F(1, 4)=248.69, p<0.0005$ ]. For this and all other ANOVAs in this study, Greenhouse-Geisser correction was applied whenever appropriate, and the reported  $p$  values are the corrected values. The degrees of freedom indicated between parentheses with each  $F$  statistic are the original (uncorrected) degrees of freedom.

Assmann, P. F., and Summerfield, A. Q. (1990). “Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies,” *J. Acoust. Soc. Am.* **88**, 680–697.

Assmann, P. F., and Summerfield, A. Q. (1994). “The contribution of waveform interactions to the perception of concurrent vowels,” *J. Acoust. Soc. Am.* **95**, 471–484.

- Beerends, J. G., and Houtsma, A. J. M. (1989). "Pitch identification of simultaneous diotic and dichotic two-tone complexes," *J. Acoust. Soc. Am.* **85**, 813–819.
- Bernstein, J. G., and Oxenham, A. J. (2003). "Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number?," *J. Acoust. Soc. Am.* **113**, 3323–3334.
- Carlyon, R. P. (1996). "Encoding the fundamental frequency of a complex tone in the presence of a spectrally overlapping masker," *J. Acoust. Soc. Am.* **99**, 517–524.
- Carlyon, R. P., van Wieringen, A., Long, C. J., Deks, J. M., and Wouters, J. (2002). "Temporal pitch mechanisms in acoustic and electric hearing," *J. Acoust. Soc. Am.* **112**, 621–633.
- Culling, J. F., and Darwin, C. J. (1993). "Perceptual separation of simultaneous vowels: Within and across-formant grouping by  $F_0$ ," *J. Acoust. Soc. Am.* **93**, 3454–3467.
- Culling, J. F., and Darwin, C. J. (1994). "Perceptual and computational separation of simultaneous vowels: Cues arising from low-frequency beating," *J. Acoust. Soc. Am.* **95**, 1559–1569.
- Dai, H. (2000). "On the relative influence of individual harmonics on pitch judgment," *J. Acoust. Soc. Am.* **107**, 953–959.
- Darwin, C. J. (2005). "Simultaneous grouping and auditory continuity," *Percept. Psychophys.* **67**, 1384–1390.
- Darwin, C. J., and Hukin, R. W. (2004). "Limits to the role of a common fundamental frequency in the fusion of two sounds with different spatial cues," *J. Acoust. Soc. Am.* **116**, 502–506.
- Darwin, C. J., Brungart, D. S., and Simpson, B. D. (2003). "Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers," *J. Acoust. Soc. Am.* **114**, 2913–2922.
- de Cheveigné, A., and Kawahara, H. (1999). "Multiple period estimation and pitch perception model," *Speech Commun.* **27**, 175–185.
- Gockel, H., Carlyon, R. P., and Micheyl, C. (1999). "Context dependence of fundamental-frequency discrimination: Lateralized temporal fringes," *J. Acoust. Soc. Am.* **106**, 3553–3563.
- Gockel, H., Carlyon, R. P., and Moore, B. C. J. (2005a). "Pitch discrimination interference: The role of pitch pulse asynchrony," *J. Acoust. Soc. Am.* **117**, 3860–3866.
- Gockel, H., Carlyon, R. P., and Plack, C. J. (2004). "Across-frequency interference effects in fundamental frequency discrimination: Questioning evidence for two pitch mechanisms," *J. Acoust. Soc. Am.* **116**, 1092–1104.
- Gockel, H., Carlyon, R. P., and Plack, C. J. (2005b). "Dominance region for pitch: Effects of duration and dichotic presentation," *J. Acoust. Soc. Am.* **117**, 1326–1336.
- Goldstein, J. L. (1973). "An optimum processor theory for the central formation of the pitch of complex tones," *J. Acoust. Soc. Am.* **54**, 1496–1516.
- Grose, J. H., Hall, J. W., and Buss, E. (2002). "Virtual pitch integration for asynchronous harmonics," *J. Acoust. Soc. Am.* **112**, 2956–2961.
- Houtsma, A. J. M., and Smurzynski, J. (1990). "Pitch identification and discrimination for complex tones with many harmonics," *J. Acoust. Soc. Am.* **87**, 304–310.
- Laguitton, V., Demany, L., Semal, C., and Liegeois-Chauvel, C. (1998). "Pitch perception: A difference between right- and left-handed listeners," *Neuropsychologia* **36**, 201–207.
- Licklider, J. C. R. (1951). "A duplex theory of pitch perception," *Experientia* **7**, 128–133.
- Meddis, R., and Hewitt, M. (1992). "Modeling the identification of concurrent vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **91**, 233–245.
- Meddis, R., and O'Mard, L. (1997). "A unitary model of pitch perception," *J. Acoust. Soc. Am.* **102**, 1811–1820.
- Micheyl, C., and Carlyon, R. P. (1998). "Effects of temporal fringes on fundamental-frequency discrimination," *J. Acoust. Soc. Am.* **104**, 3006–3018.
- Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (1985). "Relative dominance of individual partials in determining the pitch of complex tones," *J. Acoust. Soc. Am.* **77**, 1853–1860.
- Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (1986). "Thresholds for hearing mistuned partials as separate tones in harmonic complexes," *J. Acoust. Soc. Am.* **80**, 479–483.
- Plomp, R. (1964). "The ear as a frequency analyzer," *J. Acoust. Soc. Am.* **36**, 1628–1636.
- Plomp, R. (1967). "Pitch of complex tones," *J. Acoust. Soc. Am.* **41**, 1526–1533.
- Qin, M. K., and Oxenham, A. J. (2005). "Effects of envelope-vocoder processing on  $F_0$  discrimination and concurrent-vowel identification," *Ear Hear.* **26**, 451–460.
- Ritsma, R. J. (1967). "Frequencies dominant in the perception of the pitch of complex sounds," *J. Acoust. Soc. Am.* **42**, 191–198.
- Scheffers, M. T. M. (1983). "Sifting vowels: Auditory pitch analysis and sound segregation," Ph.D. thesis, Groningen University, The Netherlands.
- Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., Dosch, H. G., Bleeck, S., Stippich, C., and Rupp, A. (2005). "Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference," *Nat. Neurosci.* **8**, 1241–1247.
- Shackleton, T. M., and Carlyon, R. P. (1994). "The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination," *J. Acoust. Soc. Am.* **95**, 3529–3540.
- Smoorenburg, G. F. (1970). "Pitch perception of two-frequency stimuli," *J. Acoust. Soc. Am.* **48**, 924–941.