Sequential F0 comparisons between resolved and unresolved harmonics: No evidence for translation noise between two pitch mechanisms[a]

Christophe Micheyl[b] and Andrew J. Oxenham[c]
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307

(Received 29 May 2003; revised 24 June 2004; accepted 13 August 2004)

Carlyon and Shackleton [J. Acoust. Soc. Am. **95**, 3541–3554 (1994)] suggested that fundamental-frequency (F0) discrimination performance between resolved and unresolved harmonics is limited by an internal ''translation'' noise between the outputs of two distinct F0 encoding mechanisms, in addition to the encoding noise associated with each mechanism. To test this hypothesis further, F0 difference limens (DLF0s) were measured in six normal-hearing listeners using sequentially presented groups of harmonics. The two groups of harmonics presented on each trial were bandpass filtered into the same or different spectral regions, in such a way that both groups contained mainly resolved harmonics, both groups contained only unresolved harmonics, or one group contained mainly resolved and the other only unresolved harmonics. Three spectral regions (low: 600–1150 Hz, mid: 1400–2500 Hz, or high: 3000–5250 Hz) and two nominal F0s (100 and 200 Hz) were used. The DLF0s measured in across-region conditions were well accounted for by a model assuming only two sources of internal noise: the encoding noise estimated on the basis of the within-region results plus a constant noise associated with F0 comparisons across different spectral regions, independent of resolvability. No evidence for an across-pitch-mechanism translation noise was found. A reexamination of previous evidence for the existence of such noise suggests that the present negative outcome is unlikely to be explained by insufficient measurement sensitivity or an unusually large across-region comparison noise in the present study. While the results do not rule out the possibility of two separate pitch mechanisms, they indicate that the F0s of sequentially presented resolved and unresolved harmonics can be compared internally at no or negligible extra cost. [DOI: 10.1121/1.1806825]

PACS numbers: 43.66.Hg, 46.66.Fe [NFV] Pages: 3038–3050

I. INTRODUCTION

Harmonic complex tones generally evoke a sensation of pitch corresponding to their fundamental frequency (F0). This sensation, which is elicited even when the F0 component itself is absent from the physical spectrum, is generally referred to as complex pitch, fundamental pitch, residue pitch, or periodicity pitch. Although the mechanisms mediating periodicity pitch have been studied experimentally for over a century, they remain uncertain. One important open question is whether F0 perception is mediated by a single mechanism in all circumstances (e.g., Meddis and O'Mard, 1997), or whether different mechanisms operate, depending on whether the harmonics are resolved or not in the auditory periphery (e.g., Carlyon and Shackleton, 1994). Resolved harmonics are thought to be represented individually in the auditory periphery, and can be individually ''heard out,'' whereas unresolved harmonics combine within the periphery to produce complex waveforms with a repetition rate corresponding to the F0.

One of the seemingly strongest arguments in favor of the ''two pitch mechanisms'' hypothesis was put forward by Carlyon and Shackleton (1994). They showed that performance in a task where listeners had to compare the F0s of two harmonic complexes presented simultaneously in different spectral regions was significantly worse than predicted in conditions where one group of harmonics was resolved and the other was unresolved, but not when both groups were resolved. They argued that this was consistent with the existence of an extra source of internal noise associated with the ''translation'' between the outputs of the two F0-encoding mechanisms, in addition to the encoding noise associated with each mechanism. Carlyon and Shackleton further noted that existing implementations of a single, autocorrelation-based model of F0 encoding did not account for the finding of better F0 discrimination performance between two resolved complexes than between one resolved and one unresolved complex irrespective of spectral region (see also Carlyon, 1998).

More recently, Carlyon *et al.* (2000) found that thresholds for detecting mistuning between two simultaneous F0-modulated complexes filtered into different spectral regions were larger when the two complexes were of a different resolvability status than when they were of the same resolvability status (at least for modulation rates below 10 Hz). These authors also proposed a model for predicting the

---

[a]Portions of this work were presented at the 145th meeting of the Acoustical Society of America, Nashville, Tennessee, 28 April–2 May 2003 [J. Acoust. Soc. Am. **113**, 2225 (2003)] and at the International Symposium on Hearing, Dourdan, France, 24–28 August 2003.
[b]Electronic mail: cmicheyl@mit.edu
[c]Electronic mail: oxenham@mit.edu

thresholds in simultaneous across-region conditions from those measured in within-region conditions, but they did not test specifically for translation noise between the outputs of two distinct mechanisms for the pitch of resolved and unresolved harmonics in that study. Similarly, an earlier study by Carlyon *et al.* (1992) looked at F0 comparisons between simultaneous F0-modulated complexes, using both resolved and unresolved harmonics, but did not test specifically for translation noise. Thus, the currently available empirical evidence for translation noise remains limited to a single study (Carlyon and Shackleton, 1994).

The present study provides a further test of the existence of translation noise in comparisons between F0 estimates derived from resolved and unresolved harmonics. Thresholds for F0 comparisons between *sequential* harmonic complexes filtered into the same or different spectral regions were measured in six normal-hearing listeners. The nominal F0s and spectral regions were chosen in such a way that the stimuli contained either mostly resolved or only unresolved harmonics. In one set of conditions, the two tones being compared on each trial were filtered into the same spectral region (within-region comparisons). In this situation, the harmonics of both complexes had the same resolvability status. In another set of conditions, the two complexes were filtered into different spectral regions (across-region comparisons). In this situation, depending on which particular combination of spectral regions and nominal F0 was used, the harmonics of the two complexes had either the same or a different resolvability status. By comparing thresholds in these different conditions, we could estimate the size of the encoding noise, the (across-region) comparison noise, and the (across-mechanism) translation noise, and determine the significance of their respective influences.

II. METHODS

A. Listeners

Overall, six subjects (one female, five male, ages between 26 and 46 years) took part in the experiment. All had pure-tone hearing thresholds not exceeding 20 dB HL at octave frequencies between 250 and 8000 Hz. All listeners received substantial training in F0 discrimination using the same stimuli and conditions as used in this study. The listeners were initially trained with stimuli filtered into the same frequency region, so that they could grasp the pitch discrimination task without being confused by timbre differences. Following this initial training phase, the listeners were trained in F0 discrimination with complexes filtered into different spectral regions, until stable thresholds were obtained. Overall, the listeners had received at least 35 h of training before data collection began.

B. Stimuli

The stimulus design was inspired by Shackleton and Carlyon (1994). These authors used a combination of three spectral regions and two F0s in order to tease apart the influence of resolvability from that of F0 or spectral region alone. Our selection of spectral regions was based on a convergence of evidence, suggesting that for F0s between about 100 and 200 Hz, harmonics above about the 10th are unresolved (Hoekstra, 1979; Houtsma and Smurzynski, 1990; Bernstein and Oxenham, 2003). Accordingly, in the ''unresolved'' conditions of the present study, only harmonics above the 10th were contained in the passbands of the spectral regions. We tried to improve on Shackleton and Carlyon's design by choosing spectral regions whose 3-dB passbands always contained at least three harmonics, yet never contained the F0 component itself, and had roughly equal widths on a Cam or ERB_N scale (Glasberg and Moore, 1990; Moore, 2003). In addition, consecutive spectral regions were separated by a roughly equal number of Cams.

TABLE I. Summary of the resolvability conditions produced by combining one of two different nominal F0s with one of three different spectral regions. Entries with the letter R indicate conditions involving essentially resolved harmonics. Entries with the letter U indicate conditions involving only unresolved harmonics.

F0 (Hz)	Low 600–1150 Hz	Mid 1400–2500 Hz	High 3000–5250 Hz
100	R	U	U
200	R	R	U

The stimuli consisted of harmonic complexes filtered into one of three different spectral regions defined by the following corner frequencies: 600–1150 Hz (low), 1400–2500 Hz (mid), or 3000–5250 Hz (high), with spectral slopes on the two sides of 48 dB/oct. Depending on the condition being tested, the two complexes presented successively in each trial were filtered into the same or different spectral regions. Three conditions involved within-region comparisons (low-low, mid-mid, and high-high) and three involved across-region comparisons (low-mid, low-high, mid-high). The nominal F0 of the complexes was either 100 or 200 Hz, leading to 12 combinations of F0 and spectral region. The actual F0s of the complexes presented during the experiment were positioned symmetrically on a logarithmic scale around a reference frequency, which was randomized across trials over a 1-semitone (roughly $\pm 3\%$) range around the nominal F0 (100 or 200 Hz). The spectral regions and F0s were chosen such that the passbands would contain (a) resolved harmonics at both nominal F0s in the low region, (b) only unresolved harmonics at both F0s in the high region, and (c) only unresolved harmonics at the lower F0 but resolved harmonics at the higher F0 in the mid region; this is summarized in Table I.

A background noise was used to mask combination tones. As in Bernstein and Oxenham's (2003) study, the background noise was designed to produce roughly equal pure-tone masked thresholds over a wide frequency range. It was characterized by a flat spectral envelope below 600 Hz, and a spectral slope of -2 dB/oct above that frequency. The noise was low-pass filtered at a cutoff of 16 kHz. It started 500 ms before the first interval and ended 500 ms after the second interval, for a total duration of 2500 ms, including 100-ms raised-cosine ramps. The tones had an overall duration of 500 ms, including 20-ms raised-cosine ramps. The spectrum level of the noise in its flat spectral portion (below 600 Hz) was set 30 dB below the level of the individual

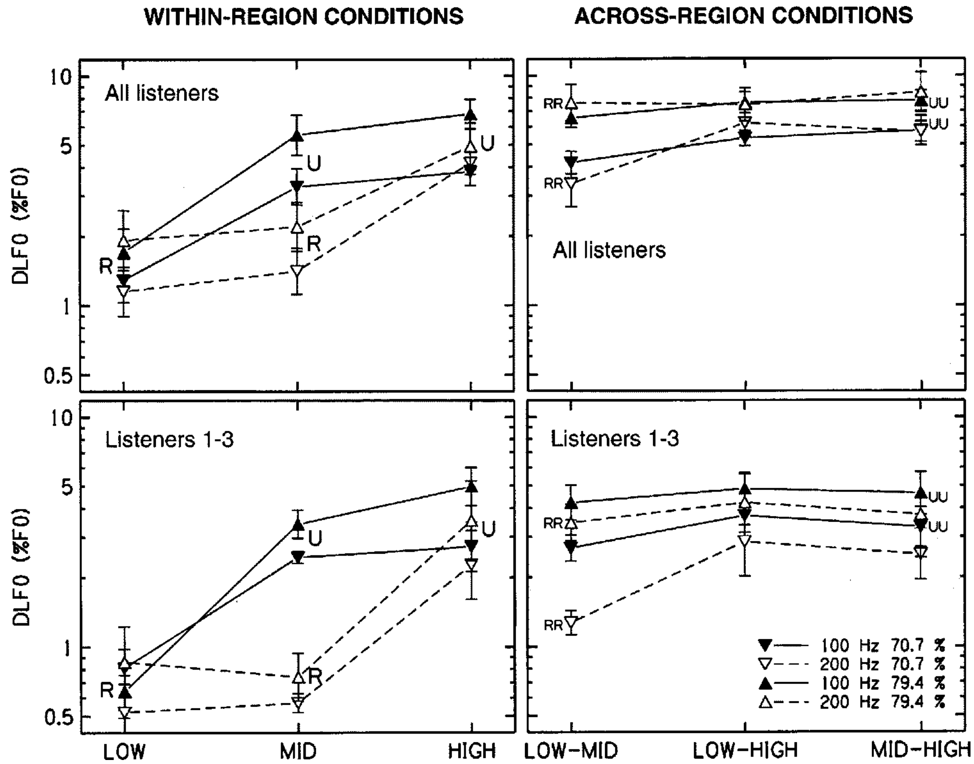

FIG. 1. Mean DLF0s in within-region and across-region conditions. The left-hand panels show within-region data; the right-hand panels show across-region data. The upper panels show averages across all six listeners; the lower panels show averages across listeners 1–3. The different spectral regions used to produce the different conditions are indicated on the abscissa. DLF0s measured using the 100-Hz nominal F0 are indicated by filled symbols connected by solid lines, those measured using the 200-Hz nominal F0 by open symbols connected by dashed lines. For within-region conditions (left-hand panels), DLF0s measured in conditions involving resolved harmonics are indicated by the letter R, and DLF0s measured in conditions involving only unresolved harmonics are indicated by the letter U, next to the corresponding symbols. In across-region conditions (right-hand panels), DLF0s measured in conditions involving two resolved groups are indicated by the letters RR, and DLF0s measured in conditions involving two unresolved groups are indicated by the letters UU, next to the corresponding symbol. (The symbols corresponding to conditions involving one resolved and one unresolved group are not associated to letters in order to avoid cluttering.) Note that while the letters R and U in the within-region panels are associated with more than one symbol, the letters RR and UU in the across-region panels correspond to the single nearest symbol. The 70.7%-correct DLF0s are represented by downward-pointing triangles; the 79.4%-correct DLF0s are represented by upward-pointing triangles. The error bars represent the standard error of the geometric mean across listeners, after removing any overall differences in performance between listeners (i.e., after subtracting their average from the log-transformed DLF0s in each listener).

harmonics within the complexes. The harmonics were added in sine phase and presented at a level of 45 dB SPL per component in the passband.

C. Procedure

DLF0s were measured using an adaptive two-interval, two-alternative forced-choice (2I-2AFC) procedure with either a two-down, one-up or a three-down, one-up rule, tracking the 70.7% and 79.4% correct points, respectively. On each trial, listeners were presented with two successive complex tones differing in F0 by ΔF0. Their task was to judge which of the two intervals contained the complex with the higher F0. The two observation intervals were marked visually, and visual feedback was provided after each trial.

At the beginning of a run, ΔF0 was set to 40%. Depending on the convergence point used (70.7% or 79.4%), ΔF0 was reduced after two or three consecutive correct responses; in all cases, it was increased following an incorrect response. The factor of variation of ΔF0, which was initially 4, was reduced to 2 the first time that ΔF0 went from decreasing to increasing, and to $\sqrt{2}$ the second time (which corresponded to the third overall reversal, a reversal being defined as a change in the direction of variation of ΔF0, either from increasing to decreasing, or from decreasing to increasing). Thresholds were estimated based on the last 12 reversals at the last step size, out of a total of 16 reversals in each track (see below).

Within each run, two independent tracks were randomly interleaved, with each track assigning the higher F0 to a different spectral region. For instance, in the low-mid condition, one track had the higher F0 assigned to the low region and the other track had the higher F0 assigned to the mid region. This allowed us to control for (and calculate) any response biases, or pitch shifts, due to differences in spectral region. Response biases could arise if listeners responded on the basis of spectral region or timbre, rather than periodicity pitch, thus selecting the observation interval that contained the complex filtered in the higher region, not that with the higher F0. Pitch shifts could arise due to a genuine influence of spectral region on pitch (Walliser, 1969; Ohgushi, 1978; Moore and Moore, 2003a,b). Although no such response biases or pitch shifts could occur in within-region conditions (as both tracks were identical), these conditions were run in the same way for consistency. Based on the differences between the raw DLF0s measured in each interleaved pair of tracks, we could estimate the amount of spectral-region bias

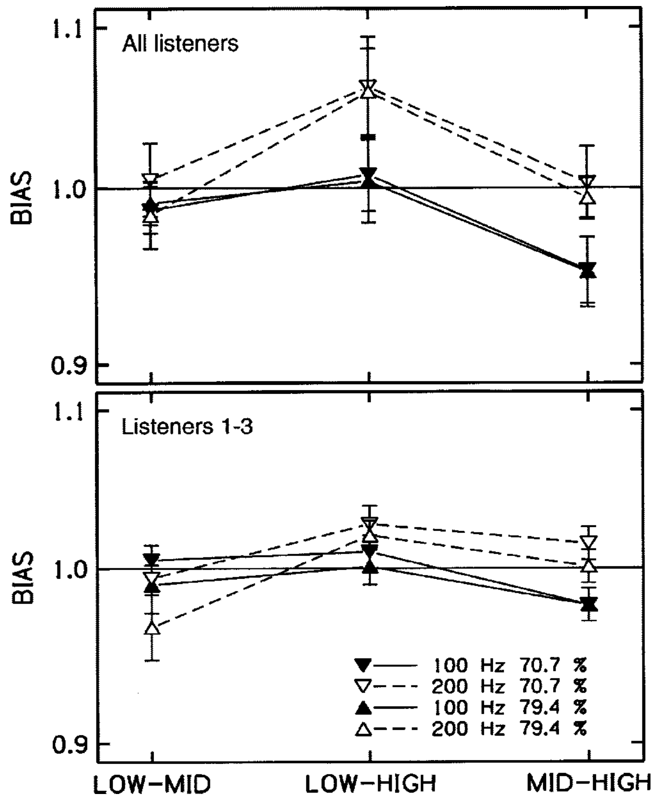

FIG. 2. Spectral-region bias in across-region conditions. The upper panel shows averages across all six listeners; the lower panel shows averages across listeners 1–3. The different spectral regions used to produce the different conditions are indicated on the abscissa. The amount of bias is represented on the ordinate as a multiplicative factor. Values around one indicate no spectral region bias. Values below one indicate a tendency to report the stimulus in the higher spectral region as having the lower F0; values above 1 indicate a tendency to hear the stimulus in the higher spectral region as having the higher F0. Data obtained using a 100-Hz nominal F0 are indicated by filled symbols connected by solid lines; those obtained using a 200-Hz nominal F0 are indicated by open symbols connected by dashed lines. The downward-pointing triangles correspond to 70.7%-correct points, the upward-pointing triangles to 79.4%-correct points. The error bars represent the standard error of the geometric mean across listeners, after removing any overall differences in performance between listeners.

in each condition, and determine the "true" DLF0s (after removing the spectral-region bias) in the corresponding condition, as described in Appendix A. Only the "true" DLF0s, computed as described in Appendix A, were used in the figures, statistical analyses, and model predictions presented in this article. For clarity, these "true" DLF0s are simply referred to as "DLF0s" throughout the paper, the expression "raw DLF0s" being reserved for the original DLF0s measured in each interleaved track of the adaptive procedure. Three "true" DLF0s were obtained in each condition. Pooling across runs and listeners always involved geometric, rather than arithmetic, averages.

D. Apparatus

The stimuli were generated digitally and played out via a soundcard (LynxStudio LynxOne) with 24-bit resolution and a sampling frequency of 32 kHz. The stimuli were then passed to a headphone buffer (TDT HB6) before being presented to the subject via the left earpiece of Sennheiser HD 580 headphones. Subjects were seated in a double-walled sound-attenuating chamber. Intervals were marked by "lights" on a virtual response box, displayed on a computer screen, and subjects responded via a computer keyboard.

III. RESULTS

A. DLF0s

Figure 1 shows the DLF0s measured in the within-region conditions (left) and the across-region conditions (right). The upper two panels show mean results obtained by averaging across all six listeners. The lower two panels show the mean results averaged across only three of the above six listeners. The data of these three listeners were isolated because they showed the lowest thresholds and the clearest dissociation, in the within-region conditions, between conditions that involved resolved harmonics and conditions that did not. Selecting these listeners for illustration will be important when examining the evidence for translation noise in mixed-resolvability conditions. This is because those listeners with the lowest levels of encoding noise (i.e., lower thresholds) are perhaps the most likely to exhibit the effects of any additional noise sources. Accordingly, we ran two sets of analyses in parallel: one on the data of all six listeners and one on the data of listeners 1–3 only.

Consider first the within-region data shown in the left-hand panels. As expected, the 79.4%-correct thresholds were generally higher than the 70.7%-correct thresholds $[F(1,5)=28.95,\ p<0.005]$.[1] Also as expected, conditions involving resolved harmonics (which are indicated by the letter R next to the corresponding symbols in Fig. 1) produced lower thresholds than conditions with unresolved harmonics (which are indicated by the letter U). The average DLF0s in the three selected listeners (lower left panel) are in good agreement with those obtained in earlier studies involving comparable stimuli (Hoekstra, 1979; Shackleton and Carlyon, 1994). Most importantly, the difference between resolved- and unresolved-harmonic conditions observed in these three listeners is at least as large as that observed in these earlier studies.

The right-hand panels in Fig. 1 show the DLF0s measured in across-region conditions. When the whole study group was considered (upper right panel), no significant difference was observed between the DLF0s measured in conditions involving one resolved and one unresolved group, and those measured using both resolved or both unresolved groups (which are marked as "RR" and "UU," respectively, in Fig. 1) $[F(1,5)=0.26;\ p=0.630]$. Furthermore, the DLF0s measured using two resolved groups (i.e., in the low-mid condition with the 200-Hz nominal F0, which is marked as "RR" in Fig. 1) did not differ significantly from those measured using two unresolved groups (i.e., in the mid-high condition with the 100-Hz nominal F0, which is marked as "UU" in Fig. 1) $[F(1,5)=1.77,\ p=0.241]$. In the subset of listeners (listeners 1–3) whose within-region data showed the clearest dissociation between resolved and unresolved conditions, the DLF0s measured using two groups of resolved harmonics were smaller than those measured using two unresolved groups. This difference was marginally significant $[F(1,2)=18.48,\ p=0.050]$.

As in the within-region conditions, a significant difference was found between the 70.7%- and the 79.4%-correct DLF0s [$F(1,5)=47.97$, $p=0.001$ for all six listeners; $F(1,2)=52.84$, $p=0.018$ for the three selected listeners]. The increase in DLF0s between the 70.7%- and 79.4%-correct points, although somewhat variable across conditions, did not differ significantly across conditions and was around 50% on average. Assuming unbiased responding, 70.7% and 79.4% correct in a 2AFC task correspond to d' values of about 0.77 and 1.16, respectively—also an increase of about 50%. This finding of similar proportional increases in DLF0 and d' is consistent with earlier results (Plack and Carlyon, 1995).

B. Spectral-region bias

Figure 2 shows the estimated amount of spectral-region bias in the across-region conditions, calculated as described in Appendix A. The amount of bias is expressed as a multiplicative factor: a value of 1 indicates no bias; values below 1 indicate a tendency to respond that the stimulus in the higher spectral region has the lower pitch; values above 1 indicate a tendency to respond that the stimulus in the higher spectral region has the higher pitch. Although a seemingly large bias was apparent in some conditions (see low-high 200 Hz and mid-high 100-Hz conditions, in which the average bias corresponded to about a semitone), a *post-hoc* statistical analysis, with Bonferroni correction, of the whole-group data failed to reveal any significant biases ($p>0.05$ in all cases). The data from the subset of listeners 1–3, whose within-region DLF0s were the lowest, showed even less departure from no bias.

C. Comparisons with model predictions

As mentioned in the Introduction, Carlyon and Shackleton (1994) proposed that performance in tasks involving F0 comparisons between harmonic complexes filtered into different spectral regions may be limited by three sources of internal noise: encoding noise, across-region comparison noise, and, when the complexes are processed via different mechanisms, across-mechanism translation noise. The size of these different internal noises can be estimated based on the DLF0s measured in different conditions. For instance, the size of the encoding noise for a specific nominal F0 and frequency region can be estimated directly from the DLF0s measured with complexes filtered into the same region, since in that situation only encoding noise should be present. Based on Green and Swets (1966), for a 2AFC task,

$$d'_{2AFC}=\frac{2\Delta}{\sqrt{2\sigma_e^2}}, \tag{1}$$

where Δ is the distance along the relevant internal decision axis between the means of the two distributions of activity (probability densities) evoked by the two stimuli to be compared. The two distributions are assumed to be Gaussian, with common variance σ_e^2. The $_{2AFC}$ subscript after d' is here to avoid confusion with the d' that would be obtained in a yes/no task; given the same Δ and σ_e^2, d'_{2AFC} is $\sqrt{2}$ times d'. Based on the present findings and the earlier results by Plack and Carlyon (1995), showing that in frequency or F0 discrimination tasks d' is proportional to the F0 difference ($\Delta F0$) in Hz between the two stimuli being compared, and assuming for simplicity that the factor relating d'_{2AFC} to $\Delta F0$ expressed in percent of the F0 is constant and equal to unity, the size of the encoding noise[2] can be estimated as

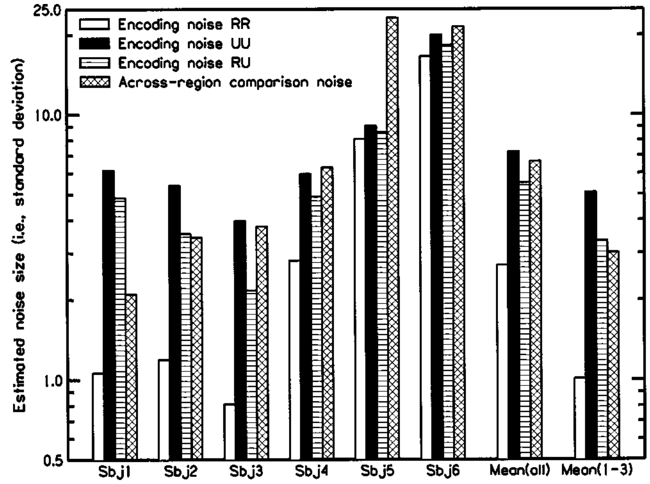

FIG. 3. Estimated sizes (i.e., standard deviations) of the encoding noise associated with comparisons between two resolved groups (empty bars), two unresolved groups (black bars), or one resolved and one unresolved group (horizontal-strip bars), and of the across-region comparison noise (crossed bars). The first six groups of bars correspond to individual data; the remaining two groups of bars show averages across all six listeners and across the three selected listeners. (See text for details.)

$$\sigma_e=\sqrt{2}\frac{\text{DLF0}}{d'_{2AFC}}, \tag{2}$$

where DLF0 is the measured $\Delta F0$ at threshold (in % of F0), and d'_{2AFC} is calculated as two times the z-score (computed using the inverse of the standard cumulative normal function) of the input probability corresponding to the targeted percent correct in the 2AFC task (Green and Swets, 1966).

When the two complexes within a trial are filtered into different spectral regions, the encoding noises in the two observation intervals may differ and across-region comparison noise must be added. In that situation, provided the two complexes being compared both contain either resolved or unresolved harmonics (so that no translation noise is present), one has

$$d'_{2AFC}=\frac{2\text{DLF0}}{\sqrt{\sigma_{e1}^2+\sigma_{e2}^2+\sigma_c^2}}, \tag{3}$$

where DLF0 is the ΔF0 at threshold measured in the considered across-region condition (say, low-mid) at the considered nominal F0 (say, 200 Hz), σ_{e1}^2 and σ_{e2}^2 are the variances of the encoding noises associated with the two considered spectral regions (in this example, low and mid) for the same nominal F0, and σ_c^2 is the variance of the across-region comparison noise in the considered condition (low-mid, 200 Hz F0).

Rearranging the terms in Eq. (3), one obtains

$$\text{DLF0}=\tfrac{1}{2}d'_{2AFC}\sqrt{\sigma_{e1}^2+\sigma_{e2}^2+\sigma_c^2}. \tag{4}$$

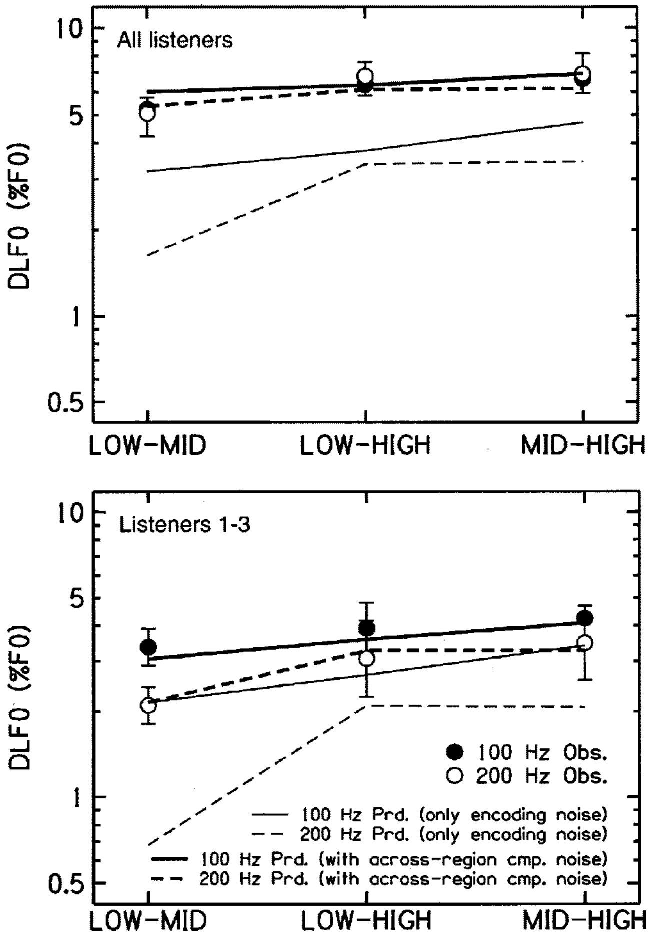

FIG. 4. Comparison between the observed and the predicted DLF0s in the across-region conditions. The upper panel shows averages across all six listeners; the lower panel shows averages across listeners 1–3. The symbols show the geometric mean of the 70.7%- and 79.4%-correct DLF0s measured in each condition. As for Fig. 2, DLF0s obtained using the 100-Hz nominal F0 are indicated by filled symbols and DLF0s measured using the 200-Hz nominal F0 are indicated by open symbols. The error bars represent the standard error of the geometric mean across listeners, after removing any overall differences in performance between listeners. The lines represent predictions. The solid lines correspond to predictions for the 100-Hz nominal F0, the dashed lines to predictions for the 200-Hz nominal F0. The thin lines correspond to predictions using only encoding noise. The thick lines correspond to predictions assuming encoding and across-region comparison noise. (See text for details.)

The variances σ_{e1}^2 and σ_{e2}^2 can be derived directly from the DLF0s measured in within-region conditions (in the above example, low 200 Hz and mid 200 Hz) using Eq. (2). The variance σ_c^2 can then be estimated as the value minimizing the root-mean-square (rms) difference between the mean log-transformed DLF0s predicted using Eq. (4) and the measured (i.e., observed) ones, over all across-region conditions in which the two groups of harmonics are either both resolved or both unresolved.

The best-fitting estimate of the across-region comparison noise size, measured as its standard deviation, σ_c, is shown in Fig. 3 for each of the six listeners. The (geometric) mean noise sizes across all six listeners or just the three selected listeners are shown on the right. The size of the combined encoding noises ($\sqrt{\sigma_{e1}^2+\sigma_{e2}^2}$) associated with F0 comparisons between two groups of resolved harmonics (RR), two groups of unresolved harmonics (UU), and one resolved versus one unresolved group are plotted for comparison. As expected from the DLF0s measured in within-region conditions, the total amount of encoding noise was generally smaller for resolved than for unresolved harmonics. On average, the across-region comparison noise was of roughly the same size as the total amount of encoding noise associated with comparisons between one resolved and one unresolved group.

Before trying to estimate the size of the across-mechanism translation noise, it is worth examining first how well the observed DLF0s in conditions involving one resolved and one unresolved complex can be predicted using only encoding and across-region comparison noise. Figure 4 shows the 75%-correct across-region DLF0s predicted using Eq. (4) with the across-region comparison noise set to zero (thin lines) or to its best-fitting value, held constant across conditions (thick lines). The thin lines show predictions assuming that the DLF0s in across-region conditions were, like those measured in within-region conditions, limited only by encoding noise. "Observed" DLF0s corresponding to roughly 75% correct, which were calculated as the geometric mean of the 70.7%- and 79.4%-correct DLF0s measured in across-region conditions (shown in Fig. 1), are plotted here as symbols. As can be seen, the predictions obtained with the across-region comparison noise set to zero generally resulted in substantial underestimates of the actual DLF0s. On the other hand, the predictions derived with a constant (nonzero) across-region comparison noise, derived through optimal fitting of the DLF0s only in across-region conditions involving two groups of resolved or two groups of unresolved harmonics, provided a good fit to the whole set of across-region DLF0s, including those measured using one resolved and one unresolved group. The observed and predicted (log-transformed) DLF0s were compared using an ANOVA for repeated measures. Only the four combinations of nominal F0 and spectral region that involved across-region comparisons between one resolved and one unresolved group were included in this analysis. No significant difference was found between the predicted and the observed thresholds [$F(1,5)=0.00$, $p=0.973$ for the whole study group; $F(1,2)=0.06$, $p=0.831$ for the subset of three listeners]. This indicates that the present results can be accounted for with a model assuming only two additive sources of internal noise: one source of F0 encoding noise, the size of which varies across conditions, depending primarily on whether the harmonics are resolved or not; and another source of noise, associated with across-region comparisons, the size of which is constant across conditions. No other source of internal noise is needed to explain the present results. Furthermore, adding an extra source of noise in conditions involving comparisons between one resolved and one unresolved group of harmonics would not improve the accuracy of fit because, as can be seen in Fig. 4, the predictions do not systematically underestimate the thresholds observed in these conditions. In other words, the present results provide no evidence for the existence of an across-pitch-mechanism translation noise.

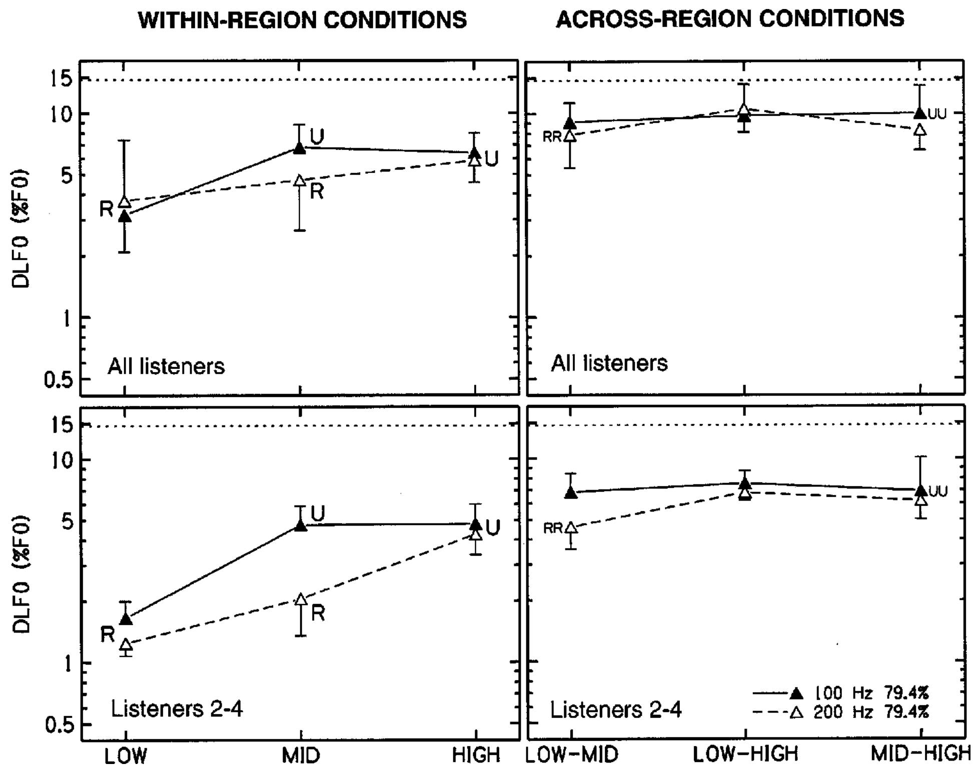

FIG. 5. Mean DLF0s measured in within-region and across-region conditions using a 4-semitone randomization range for the nominal-F0. The general format is the same as for Fig. 1. The only differences are that the upper panels show averages across five listeners (listeners 2–6) and the lower panels show averages across listeners 2–4, and that only 79.4%-correct DLF0s were measured.

IV. EFFECTS OF INCREASING THE F0 RANDOMIZATION RANGE

A. Rationale

One possible reason why no evidence for an across-pitch-mechanism translation noise was found in the preceding experiment is that listeners avoided comparing the F0s of the two complexes presented in the two observation intervals when one of these complexes was resolved and the other was unresolved. Indeed, listeners could in theory do the task as if it were a one-interval task, ignoring one of the two observation intervals and simply comparing the F0 in the other observation interval to an internal criterion corresponding to the nominal F0. If the F0 in the chosen observation interval were higher than the criterion, the listener would designate that interval as the one containing the higher-F0 complex. Listeners may have been led to use this ‘‘one interval’’ strategy in conditions involving one resolved and one unresolved group because, in those conditions, direct comparisons between the F0s of the two stimuli were perhaps strongly limited by translation noise.

In order to test whether the DLF0s measured in conditions involving one resolved and one unresolved group in experiment I could be accounted for by this type of explanation, we calculated the smallest F0 difference (ΔF0) required to achieve a given percentage of correct responses (70.7% or 79.4%) with this ‘‘one-interval’’ strategy. In order to obtain a conservative test, it was assumed that both the internal criterion and the F0 derived from the selected observation interval were completely noise-free (i.e., perfect coding was assumed). Thus, the ΔF0s indicated below must be regarded as an absolute lower bound on the DLF0s: observed DLF0s smaller than these values cannot be explained in terms of the above-described ‘‘one-interval’’ strategy; DLF0s larger than these values may be explained in terms of this strategy, provided the amount of coding noise is sufficiently small. Naturally, if perfect coding is assumed *and* the nominal F0 of the stimuli is not roved across trials, listeners can achieve 100% correct. If the nominal F0 is roved across trials over a range R (in semitones), then the F0 difference, ΔF0, required to obtain a certain proportion of correct responses, Pc, is simply equal to $2R(Pc-0.5)$ semitones. In experiment I, the nominal F0 of the stimuli was roved over a 1-semitone range (roughly $\pm 3\%$). With this roving range, the formula predicts that 70.7% of correct responses can be achieved with a ΔF0 of only 2.4%, and that 79.4% correct can be achieved with a ΔF0 of 3.5%. Since these lower bounds are smaller than the DLF0s measured in conditions involving one resolved and one unresolved group (see Fig. 1), the possibility that listeners used the ‘‘one interval’’ strategy in these conditions cannot be ruled out. Thus, it is possible in theory that the lack of evidence for translation noise in experiment I was due to the fact that listeners never directly compared the F0s of resolved and unresolved complexes.

One way of encouraging listeners to compare the F0 of the two stimuli presented within each trial is to increase the randomization range of the nominal F0. Accordingly, we retested five out of the six listeners who had taken part in experiment I using the same F0 and region conditions, but this time with an F0 randomization range of 4 semitones instead of just 1 semitone. Using the above formula, it can be shown that with such a randomization range, listeners would need a ΔF0 of at least 14.6% to perform at 79.4% correct. If the DLF0s measured in the second experiment are below this predicted value, one could conclude that listeners were comparing the F0s of the resolved and unresolved stimuli. On the other hand, if the DLF0s are larger than the predicted values above, this would support the view that listeners were relying on an internal reference rather than comparing the two observation intervals. Furthermore, if it turned out that the measured thresholds are significantly larger than the predicted ones specifically in conditions involving one resolved and one unresolved group of harmonics, this would be consistent with the existence of translation noise.

B. Methods

The stimuli and procedure were the same as in the first experiment. The only differences were in the size of the F0 roving range, which was now 4 semitones (i.e., ± 2 semitones around the nominal F0), and in the targeting of only one percent-correct value (79.4%) instead of two. The six listeners who had taken part in experiment I also took part in experiment II. However, one listener (listener 1) became unavailable before the end of that second experiment. Thus, only the data from listeners 2–6 could be retained in the final analyses. As in experiment I, two sets of analyses were run in parallel: one on the data of all five listeners, the other on the data of a subset of three listeners who showed the lowest thresholds and whose within-region data showed the largest difference between conditions that involved resolved harmonics and those that did not. As listener 1 was no longer available, these were now listeners 2–4.

C. Results and discussion

The results of experiment II are shown in Fig. 5. In the across-region conditions, the measured DLF0s were significantly lower on average than the prediction of the internal-reference model (14.6%), even in the whole study group (Fig. 5, left panel) [$t(71)=2.50$, $p<0.01$]. This indicates that listeners were not simply basing their responses on comparisons with a constant internal reference F0.

Furthermore, although a visual comparison between Figs. 1 and 5 suggests that increasing the roving range produced some increase in the DLF0s, statistical comparisons between the 79.4%-correct thresholds measured in the 1- and 4-semitone roving-range conditions in the five listeners who took part in both experiments showed no significant difference. The same analysis, when run on the data of the three selected listeners (listeners 2–4), led to the same outcome. This lack of a significant effect of the roving range is consistent with the notion that listeners were not relying on a fixed internal reference in either experiment I or II.

As with experiment I, the DLF0s measured in within-region conditions were used to estimate the amount of internal noise present in these conditions. These internal noise estimates were then used to predict the DLF0s measured in across-region conditions assuming only across-region (no across-pitch-mechanism translation) noise. Once again, no significant difference was found between the observed and

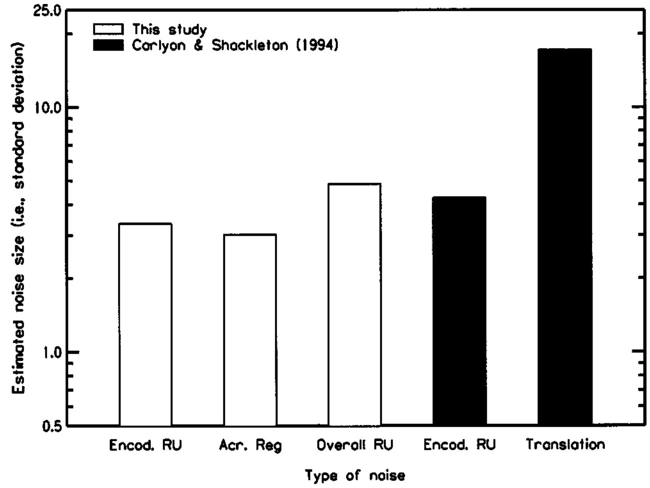

FIG. 6. Estimated size (i.e., standard deviations) of the encoding noise, the across-region comparison noise, the across-pitch-mechanism translation noise, and/or the total noises involved in comparisons between one resolved and one unresolved group in the present study (empty bars) and in the study of Carlyon and Shackleton (1994) (solid bars). (See text for details.)

predicted DLF0s in conditions involving comparisons between one resolved and one unresolved group, independent of whether the analysis was carried out on the data of all five listeners [$F(1,4)=0.22$, $p=0.664$] or of the three selected listeners [$F(1,2)=1.82$, $p=0.310$].

In summary, the apparent lack of translation noise in experiments I and II is not likely to be due to listeners using an internal reference F0 and failing to actually compare F0s across spectral regions. This leaves us with other interpretations to consider in the general discussion: either there is no across-pitch-mechanism translation noise in the auditory system or, at least, the size of this noise is small in comparison to that of the other sources of observer- or measurement-related noise present in this study.

V. GENERAL DISCUSSION

A. Was the across-mechanism translation noise swamped by the across-region comparison noise?

A possible reason why the observed and predicted DLF0s in resolved–unresolved conditions differed very little is that translation noise was swamped by the other sources of internal noise present in the across-region conditions, namely, the encoding and across-region comparison noises. Since both of these sources of internal noise are present whenever listeners compare the F0s of harmonics in different frequency regions, one might argue that if the translation noise is not large enough to have an effect in the presence of these other sources of internal noise, its influence will never be observed, except perhaps under conditions involving comparisons between resolved and unresolved harmonics filtered into the same spectral region (and, consequently, no across-region-comparison noise). However, it is possible that some experimental situations yield less across-region comparison noise than others. For instance, Carlyon and Shackleton (1994) found no significant across-region noise in F0 comparisons between simultaneously presented groups of harmonics (we discuss this finding in more detail below).

Furthermore, as the detrimental influence of differences in spectral region or timbre on DLF0s appears to vary substantially across listeners, it is also conceivable that in some listeners (possibly those with extensive musical training), the size of the across-region comparison noise is so reduced that it becomes negligible. Although some of our listeners had some musical training and all were trained in both the within- and across-region conditions prior to data collection, we cannot rule out the possibility that other listeners might exhibit smaller amounts of across-region comparison noise.

In order to test whether the lack of evidence for translation noise in the present study could be due to the presence of unusually large encoding and/or across-region comparison noise, we compared the size (standard deviation) of the total amount of noise in our across-resolvability conditions to the size of the encoding noise associated with comparisons between one resolved and one unresolved complex in Carlyon and Shackleton's (1994) study. Details on how these estimates were obtained from the data plotted in Carlyon and Shackleton's article are provided in Appendix B. Three important considerations are worth mentioning here: (a) Because Carlyon and Shackleton found no significant across-region comparison noise, the only significant source of internal noise in their simultaneous experiment besides translation noise was encoding noise. (b) Since the performance in simultaneous comparisons between two unresolved complexes in Carlyon and Shackleton's study was enhanced by the detection of pitch pulse asynchronies, the encoding noise for unresolved complexes could only be predicted from the sequential comparisons. (c) While Carlyon and Shackleton's experiment testing for across-region comparisons involved four groups of harmonics on each trial (two in each observation interval), ours involved only two, and in both experiments translation noise was thought to compete with the encoding noise associated with two complexes, not four. These three considerations indicate that the relevant comparison is between the *total* amount of (encoding and across-region) noise associated with across-region comparisons involving one resolved and one unresolved complexes in the present study, and the amount of *encoding* noise associated with corresponding comparisons in Carlyon and Shackleton (1994), as predicted from the sequential-comparison data in that study.

Figure 6 compares the sizes (or standard deviations) of the internal noises computed in the present study (for the three listeners of experiment I) with those estimated from the data plotted in Carlyon and Shackleton's (1994) article (Figs. 6, 7, and 9 in that article). It can be seen that the amount of encoding noise associated with comparisons between one group of resolved and one group of unresolved harmonics in the three selected listeners of the present study is similar to that measured in three listeners by Carlyon and Shackleton (1994). Nevertheless, because of the additional across-region comparison noise found in this study, the total noise potentially competing with translation noise was slightly larger in the present study than in Carlyon and Shackleton (1994) (compare ''overall RU M&O'' with ''encoding RU C&S'' in Fig. 6). In theory, this could have limited the ability of the present study to find evidence for translation noise. However, the reanalysis of Carlyon and Shackleton's data also reveals that the size (standard deviation) of the putative translation noise in their study was about four times larger than the overall size of the other sources of internal noise associated with comparisons between one resolved and one unresolved complex (compare the last two histogram bars in Fig. 6).

If a translation noise of the size found by Carlyon and Shackleton had been present in our study, would it have been detected? More generally speaking, how large should the translation noise have been in order for its influence on the DLF0s to be detectable in the present study? In order to answer this question, we gradually increased the previously estimated size of the internal noise added to the encoding noise in conditions involving comparisons between one resolved and one unresolved group in experiment I, recomputed the DLF0 predictions with the artificially increased internal noise size, and compared the newly predicted thresholds with the observed ones. Repeating this procedure until the ANOVA yielded a significant outcome ($p<0.05$), we found that increasing the size (standard deviation) of the noise added to the total amount of encoding noise in resolved–unresolved comparisons by 75% was sufficient to produce a statistically significant deviation between the predicted and observed thresholds [$F(1,5)=6.74$, $p=0.049$]. A 75% increase in the size (standard deviation) of the across-region comparison noise was found to correspond roughly to a 50% increase in the total amount of (encoding plus across-region comparison) noise measured in the considered conditions. To yield this 50% increase in total noise size, the added (''translation'') noise should have a size slightly (i.e., about 12%) larger than that of the total amount of encoding plus across-region comparison noise already present in the resolved-vs.-unresolved conditions.[3] The size of the translation noise estimated from Carlyon and Shackleton's (1994) data was about four times the total amount of (encoding) noise that was associated with resolved–unresolved comparisons in that study, and it is between 2 and 3.5 times the total amount of (encoding plus across-region comparison) noise measured in the present study (depending on whether one considers the data of all six listeners or only those of the three selected listeners, respectively). All of these ratios are substantially larger than the 12% difference needed to produce a statistical increase in DLF0s here. Thus, had an additional noise with the same relative size as that observed in Carlyon and Shackleton's (1994) study been present here, its influence would have been detected. To check this conclusion, we also tried adding to the total amount of internal noise measured in resolved–unresolved conditions here, another noise with the same absolute (rather than relative) size as the translation noise in Carlyon and Shackleton (1994). As expected, this resulted in a statistically significant difference between the predicted and observed DLF0s in the considered conditions [$F(1,5)=14.51$, $p=0.013$]. These observations indicate that the discrepancy between the present results and those of Carlyon and Shackleton (1994) cannot be explained simply by insufficient statistical power or the presence of unusually large encoding and/or across-region comparison noise in present study.

B. Reevaluating past claims for translation noise

How can the apparent discrepancy between the results of Carlyon and Shackleton (1994), suggesting the existence of a large across-pitch mechanism translation noise, and the present ones, showing no evidence for such noise, be resolved? Carlyon and Shackleton's (1994) evidence for translation noise, just like the present lack thereof, rests entirely on comparisons between observations and predictions. However, in their study, the predictions were derived from performance measures obtained in conditions that differed markedly from the ones in which the predictions were applied. Specifically, performance measures in F0 discrimination between *sequential* groups of harmonics filtered into the *same* spectral region were used to predict performance in F0 comparisons between *simultaneous* groups of harmonics filtered into *different* spectral regions. Because these two types of experimental situations may involve very different internal processes, the interpretation of the differences between predictions and observations in Carlyon and Shackleton's study is uncertain. The evidence for translation noise in that study is further limited by the fact that the authors had to exclude comparisons between observations and predictions in conditions that involved two groups of unresolved harmonics because, as they pointed out, performance in those conditions was probably mediated by the detection of pitch pulse asynchronies between the two complexes. The results obtained in conditions involving comparisons between two unresolved groups being ruled out, the evidence for translation noise in Carlyon and Shackleton (1994) rests entirely on the worse-than-predicted performance in F0 comparisons between one resolved and one unresolved group. However, this finding may be explained by reasons other than translation noise. For instance, when both resolved and unresolved contiguous harmonics are presented simultaneously, resolved harmonics have been shown to dominate the pitch percept (e.g., Plomp, 1967). Possibly, this makes it harder to detect differences in F0 across spectral regions due to some form of masking of the less salient pitch evoked by the unresolved harmonics, or some other form of interference. Recent results by Gockel *et al.* (2004) indicate that such interference effects do exist. These authors showed that F0 discrimination performance between two sequential complexes consisting of unresolved harmonics was impaired by the simultaneous presentation of another complex, filtered in a lower spectral region and containing resolved harmonics. Gockel *et al.* argued that this interference effect might explain the finding of lower-than-predicted performance in Carlyon and Shackleton's conditions involving simultaneous comparisons between a resolved and an unresolved complex.

C. One or two pitch mechanisms?

In view of the above analysis, it appears that the evidence currently available to support the hypothesis that F0 comparisons between resolved and unresolved harmonics are significantly affected by translation noise is, at best, limited. However, it is important to stress that rejecting the hypothesis of the existence of translation noise between the outputs of two different F0-encoding mechanisms for resolved and unresolved harmonics is not tantamount to rejecting the hypothesis that these two F0-encoding mechanisms exist. It is quite possible that the F0 estimates derived from resolved and unresolved harmonics by two distinct F0-encoding mechanisms are already reconciled into a common code before they are compared, or that these F0 estimates, although they are derived through different mechanisms, are expressed in the same form.

There remain some results in the literature that have been interpreted as suggesting the existence of different pitch encoding mechanisms for resolved and unresolved harmonics. Plack and Carlyon (1995) and White and Plack (1998) have shown qualitatively different effects of duration on DLF0s for resolved and unresolved harmonics, which may be construed as an indication that different underlying mechanisms are operating. More recently, Grimault *et al.* (2002) have obtained results that were interpreted as indicating the involvement of different mechanisms in the learning of F0 discrimination with resolved and unresolved harmonics, which also would be hard to reconcile with the assumption that the same mechanism mediates F0-discrimination performance in both cases.

D. A note on across-region comparison noise in F0 discrimination

Our finding of significant across-region comparison noise is consistent with a number of earlier studies that have found poorer F0 discrimination when the two stimuli consist of different harmonics than when they contain identical harmonics (e.g., Ritsma, 1963; Faulkner, 1985; Moore and Glasberg, 1990). Moore and Glasberg (1990) showed that DLF0s could be increased when the complexes occupied roughly the same spectral region and even shared many common harmonics. This suggests that even subtle differences in timbre can interfere with the processing of differences in F0. The interference may be due to distraction of the listener's attention away from the relevant sensory dimension (i.e., pitch) by irrelevant variations along another dimension (i.e., timbre). Thus, the expression ''across-region comparison noise,'' which we borrowed from Carlyon and Shackleton (1994), should not be taken too literally, as it may not relate specifically to across-region comparisons, but may instead or also reflect a general distracting effect of timbre differences on F0 comparisons. Whatever mechanism is at the origin of the increase in thresholds from within- to across-region conditions, our quantitative predictions demonstrate that the influence of this mechanism can be modeled successfully as a simple additive source of noise. Indeed, adding a constant amount of noise to the estimated amount of encoding noise associated with each combination of F0 and spectral region condition was sufficient to produce predicted thresholds very close to the observed ones in all the across-region conditions. A model assuming a constant multiplicative across-region-comparison noise, which would predict a proportional upward shift in thresholds from within- to across-region conditions, would clearly not be adequate. It would, for example, predict a larger than observed difference in thresholds between the all-resolved, low-mid 200-Hz condition and the unresolved-resolved, mid-high or low-high 200-Hz conditions.

E. Influence of spectral region on pitch perception

Another result of the present study, which deserves mention in relation to the question of across-region comparison noise, concerns the question of across-region bias. In this study, no significant spectral-region bias was found. This may seem surprising given that the changes in timbre between the two observation intervals were very salient. Previous studies indicate that for harmonic complexes differing in spectral composition, equal F0s do not necessarily produce exactly equal pitches (e.g., van den Brink, 1977; Chuang and Wang, 1978; Singh and Hirsh, 1992). Such influences of spectral region on perceived pitch should have shown up as bias in our study. However, several factors can explain the apparent lack of influence of timbre on pitch in most of the conditions tested in this study. First, our listeners were provided extensive training in the F0-discrimination task; even though they may have been initially tempted to rely more on timbre than on pitch, such an extensive training with trial-by-trial feedback may have eliminated most of this initial inclination. Second, it has been shown that when faced with conflicting pitch and timbre cues (e.g., increasing F0 but decreasing spectral region), most listeners follow periodicity pitch rather than timbre, provided the complexes contain more than two components (Smoorenburg, 1970; Laguitton *et al.*, 1998); our harmonic complexes generally contained more than two components. Finally, various results in the literature indicate that timbre can be varied without affecting pitch (Plomp, 1967; Hafter and Richards, 1988; Demany and Semal, 1993; for consistent results in cochlear-implant listeners, see McKay *et al.*, 2000). These prior results and the present ones, showing little spectral bias, are generally consistent with the idea that, in trained listeners, periodicity pitch is little (or not) influenced by timbre, although it remains possible that our listeners learned to ''recalibrate'' their pitch judgments depending on spectral region.

VI. CONCLUSIONS

The results obtained in this study are consistent with previous studies in showing that (1) DLF0s for harmonic complexes filtered into the same spectral region are significantly worse for unresolved than for resolved harmonics, and that (2) DLF0s for sequential harmonic complexes filtered into different spectral regions are generally worse than DLF0s for sequential complexes filtered into the same spectral region. They reveal that, whatever effect is responsible for worse performance in the across-region conditions, this effect may be modeled simply and adequately as an additive source of internal noise whose size is constant across all conditions tested here. The results provide no evidence for the hypothesis that sequential F0 comparisons between two sequentially presented groups of harmonics, one resolved and the other unresolved, are significantly limited by ''translation'' noise between the internal F0 estimates derived from these two groups. We have shown that the discrepancy between Carlyon and Shackleton's (1994) results and the present ones could not be explained simply by insufficient statistical power or an unusually large across-region comparison noise in the present study: if an additional noise with a size similar to the translation noise in Carlyon and Shackleton's study had been present here, its influence would have been detected. While the present results do not definitely rule out the possible existence of an across-pitch-mechanism translation noise in the auditory system (disproving the existence of something that can be arbitrarily small is a difficult task), at the very least, they show that even if such a noise does exist, its influence on the F0 discrimination thresholds of typical listeners is negligible. Finally, we stress that even if the existence of translation noise could be conclusively ruled out, this would not necessarily rule out the possibility that the F0s of resolved and unresolved harmonics are encoded via different mechanisms in the auditory system. However, if these different mechanisms do exist, it appears that the F0s derived from resolved and unresolved harmonics are already expressed in a similar code at the stage at which F0 comparisons between sequential tone complexes are made.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (NIDCD Grant R01 DC 05216). We thank Hedwig Gockel, Robert Carlyon, Nathaniel Durlach, and an anonymous reviewer for helpful comments on an earlier version of this manuscript. Brian Moore and Christopher Plack are also acknowledged for interesting comments on some of the results at the ISH 2003 Conference, which helped to improve the manuscript.

APPENDIX A: FORMULAS FOR ESTIMATING BIAS AND "TRUE" THRESHOLDS IN ACROSS-REGION CONDITIONS

In across-region conditions, DLF0s were measured using two interleaved tracks. In one track, the stimulus with the higher F0 was filtered into the higher spectral region. In the other track, the stimulus with the higher F0 was filtered in the lower spectral region. If the listeners' decisions are influenced by the spectral region in which the tones are filtered, an effect we refer to here as ''spectral region bias,'' the raw thresholds measured on the two tracks should be different. Spectral region bias might result from a genuine influence of spectral region on perceived F0, so that, for example, tones filtered in a higher spectral region are perceived as having a higher F0 than tones of the same F0 filtered in a lower spectral region. Alternatively, it might result from listeners inadvertently discriminating signals along the ''wrong'' dimension, i.e., timbre instead of periodicity pitch, on some of the trials.

In order to estimate the size and direction of the spectral region bias and determine the ''true'' DLF0s after accounting for this bias, we used an approach based on that described in the Appendix of Oxenham and Buus (2000). The general idea behind this approach, as applied to the present case, is that the influence of spectral region can be modeled as a shift in the value of the perceived F0. A positive shift is taken to indicate that filtering a tone into a higher spectral region increases its perceived F0 relative to some ''true'' internal value. A negative shift indicates that filtering the tone into a higher spectral region decreases its perceived F0 relative to some ''true'' internal value. Oxenham and Buus (2000) considered an additive shift on a logarithmic (dB) scale. Here, we use an equivalent approach, consistent with our use of a multiplicative rule for step-size changes in the DLF0-measurement procedure, by considering a multiplicative shift on a linear (Hz) scale. Accordingly, at threshold, the F0 of the higher-F0 tone in each track can be expressed as a function of the F0 of the lower-F0 tone, as follows:

$$\mathrm{F0}_{hi1}=\mathrm{F0}_{lo1}\cdot\Delta\cdot k, \tag{A1}$$

$$\mathrm{F0}_{hi2}=\frac{\mathrm{F0}_{lo2}}{k}\Delta, \tag{A2}$$

where $F0_{hi1}$ and $F0_{lo1}$ are the F0s of the two stimuli in the track in which the higher-F0 stimulus is presented in the higher spectral region, $F0_{hi2}$ and $F0_{lo2}$ are the F0s of the two stimuli in the track in which the higher-F0 stimulus is presented in the lower spectral region, k is the spectral region bias expressed as a factor, and Δ is the proportional difference between the two F0s corresponding to the ''true'' threshold, also expressed as a factor.

The raw DLF0s for each track are, in percents,

$$\mathrm{DLF0}_1=100\left(\frac{\mathrm{F0}_{hi1}}{\mathrm{F0}_{lo1}}-1\right), \tag{A3}$$

$$\mathrm{DLF0}_2=100\left(\frac{\mathrm{F0}_{hi2}}{\mathrm{F0}_{lo2}}-1\right). \tag{A4}$$

The ''true'' DLF0s, Δ, are then computed as the geometric mean of the raw DLF0s expressed as factors rather than percentages (for simplicity):

$$\Delta=\sqrt{\frac{\mathrm{F0}_{hi1}}{\mathrm{F0}_{lo1}}\cdot\frac{\mathrm{F0}_{hi2}}{\mathrm{F0}_{lo2}}}. \tag{A5}$$

The bias, k, is computed as the square root of the ratio of the raw DLF0s (expressed as factors):

$$k=\sqrt{\frac{\mathrm{F0}_{hi1}}{\mathrm{F0}_{lo1}}\Big/\frac{\mathrm{F0}_{hi2}}{\mathrm{F0}_{lo2}}}. \tag{A6}$$

APPENDIX B: THE SIZE OF THE INTERNAL NOISES IN CARLYON AND SHACKLETON (1994)

The size of the encoding and translation noises associated with across-region comparisons between one resolved and one unresolved group of harmonics in Carlyon and Shackleton (1994) was estimated using the data shown in their Figs. 6 and 9. First, the d' values for sequential comparisons shown in Fig. 6 were used to derive the size of the encoding noise, σ_e, associated with each complex. It is important to note that the d' values in that figure and all the other figures in Carlyon and Shackleton (1994)'s article were computed as $\sqrt{2}$ times the z-score of the measured percent correct in the 2AFC task, so that they correspond to yes/no-task-equivalent d' values—not to be confused with d'_{2AFC}. Accordingly, we have

$$\sigma_e=\frac{\Delta\mathrm{F0}}{d'_{seq}}, \tag{B1}$$

where ΔF0 is the F0 difference between the two stimuli (Carlyon and Shackleton tested 3.5% and 7.1%), and d'_{seq} is the mean across all three listeners of the observed d' values in the considered condition, read from Fig. 6 in Carlyon and Shackleton (1994). Technically, all the variables in this and the following equations should have subscripts indicating to which condition they correspond. These subscripts are not shown here to avoid cluttering.

The total amount of encoding noise, $\sigma_{e\mathrm{tot}}$, associated with F0 comparisons between two complexes filtered into different spectral regions was then calculated as

$$\sigma_{e\mathrm{tot}}=\sqrt{\sigma_{e1}^2+\sigma_{e2}^2}, \tag{B2}$$

where σ_{e1}^2 and σ_{e2}^2 are the variances of the encoding noises associated with the two considered spectral regions. A summary measure of the total amount of encoding noise associated specifically with comparisons between one resolved and one unresolved complex was obtained by averaging the squared values of $\sigma_{e\mathrm{tot}}$ obtained in the different conditions that involved one resolved and one unresolved group.

The size of the translation noise, σ_t, was derived by comparing the above-predicted amount of encoding noise ($\sigma_{e\mathrm{tot}}$) and the total amount of noise estimated using the d' values observed in the simultaneous-comparison conditions involving both resolved and unresolved complexes (Fig. 7 in Carlyon and Shackleton's article). Specifically, the translation noise was calculated as

$$\sigma_t=\sqrt{\left(\frac{\Delta F0}{d'_{sim}}\right)^2-\sigma_{e\mathrm{tot}}^2}. \tag{B3}$$

Equation (B3) can be understood more easily as a rearrangement of the equation below, which gives the predicted performance in the task used by Carlyon and Shackleton (1994) to measure DLF0s between simultaneously presented complexes:

$$d'_{sim}=\frac{\Delta F0}{\sqrt{\sigma_{e\mathrm{tot}}^2+\sigma_t^2}}, \tag{B4}$$

[1]This and all the other statistical analyses in this article were applied to the log-transformed, rather than the untransformed, DLF0s.

[2]The expression ''encoding noise'' is used here for consistency with Carlyon and Shackleton (1994). Admittedly, the amount of internal noise derived using Eq. (2) may incorporate contributions from other sources of internal noise than those associated specifically with the encoding *per se* of the stimulus F0. For instance, it may include memory-related noise associated with the comparison of the F0 of the complex in the second observation interval with the memory trace of the F0 of the complex in the first interval (see Durlach and Braida, 1969). The derived amount of encoding noise subsumes the contributions from all the internal sources of noise involved even when the F0 comparisons are made within the same spectral region.

[3]If σ_{EAT} represents the standard deviation of the combined encoding plus across-region-comparison plus translation noise, σ_{EA} represents the standard deviation of the combined encoding plus across-region-comparison

noise, and σ_T represents the standard deviation of the sole translation noise, we have $\sigma_{EAT} = \sqrt{\sigma_{EA}^2 + \sigma_T^2}$. Knowing that $\sigma_{EAT}/\sigma_{EA} = 1.5$, it follows that $\sigma_T/\sigma_{EA} = \sqrt{1.5^2 - 1}$. The right term in the preceding equation is approximately equal to 1.12, indicating that the translation noise must be roughly 12% larger than the combined encoding plus across-region noise to which it is added in order to cause a 50% increase in the size (standard deviation) of the total amount of noise.

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. (**1998**). ''Comments on ''A unitary model of pitch perception'' [J. Acoust. Soc. Am. **102**, 1811–1820 (1997)],'' J. Acoust. Soc. Am. **104**, 1118–1121.

Carlyon, R. P., and Shackleton, T. M. (**1994**). ''Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms,'' J. Acoust. Soc. Am. **95**, 3541–3554.

Carlyon, R. P., Demany, L., and Semal, C. (**1992**). ''Detection of across-frequency differences in fundamental frequency,'' J. Acoust. Soc. Am. **91**, 279–292.

Carlyon, R. P., Moore, B. C. J., and Micheyl, C. (**2000**). ''The effect of modulation rate on the detection of frequency modulation and mistuning of complex tones,'' J. Acoust. Soc. Am. **108**, 304–315.

Chuang, C. K., and Wang, W. S. (**1978**). ''Psychophysical pitch biases related to vowel quality, intensity difference, and sequential order,'' J. Acoust. Soc. Am. **64**, 1004–1014.

Demany, L., and Semal, C. (**1993**). ''Pitch versus brightness of timbre: Detecting combined shifts in fundamental and formant frequency,'' Music Percept. **11**, 1–14.

Durlach, N. I., and Braida, L. D. (**1969**). ''Intensity perception. I. Preliminary theory of intensity resolution,'' J. Acoust. Soc. Am. **46**, 372–383.

Faulkner, A. (**1985**). ''Pitch discrimination of harmonic complex signals: Residue pitch of multiple component discrimination?'' J. Acoust. Soc. Am. **78**, 1993–2004.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notch noise data,'' Hear. Res. **47**, 103–138.

Gockel, H., Carlyon, R. P., and Plack, C. J. (**2004**). ''Across frequency interference effects in fundamental frequency discrimination: Questioning evidence for two pitch mechanism,'' J. Acoust. Soc. Am. (in press).

Green, D. M., and Swets, D. A. (**1966**). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Grimault, N., Micheyl, C., Carlyon, R. P., and Collet, L. (**2002**). ''Evidence for two pitch encoding mechanisms using a selective auditory training paradigm,'' Percept. Psychophys. **64**, 189–197.

Hafter, E. R., and Richards, V. M. (**1988**). ''Discrimination of the rate of filtered impulses,'' Percept. Psychophys. **43**, 405–414.

Hoekstra, A. (**1979**). ''Frequency discrimination and frequency analysis in hearing,'' Ph.D. dissertation, Institute of Audiology, University Hospital, Groningen, Netherlands.

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.

McKay, C. M., McDermott, H. J., and Carlyon, R. P. (**2000**). ''Place and temporal cues in pitch perception: Are they truly independent?'' ARLO **1**, 25–30.

Meddis, R., and O'Mard, L. (**1997**). ''A unitary model of pitch perception,'' J. Acoust. Soc. Am. **102**, 1811–1820.

Moore, B. C. J. (**2003**). *An Introduction to the Psychology of Hearing* (Elsevier Science, Amsterdam).

Moore, B. C. J., and Glasberg, B. R. (**1990**). ''Frequency discrimination of complex tone with overlapping and non-overlapping harmonics,'' J. Acoust. Soc. Am. **87**, 2163–2177.

Moore, G. A., and Moore, B. C. J. (**2003a**). ''Perception of the low pitch of frequency-shifted complexes by normal-hearing listeners,'' J. Acoust. Soc. Am. **113**, 977–985.

Moore, B. C. J., and Moore, G. A. (**2003b**). ''Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelopes by normally hearing and hearing-impaired subjects,'' Hear. Res. **182**, 153–163.

Ohgushi, K. (**1978**). ''On the role of spatial and temporal cues in the perception of the pitch of complex tones,'' J. Acoust. Soc. Am. **64**, 764–771.

Oxenham, A. J., and Buus, S. (**2000**). ''Level discrimination of sinusoids as a function of duration and level for fixed-level, roving-level, and across-frequency conditions,'' J. Acoust. Soc. Am. **107**, 1605–1614.

Plack, C. J., and Carlyon, R. P. (**1995**). ''Differences in frequency modulation detection and fundamental frequency discrimination between complex tones consisting of resolved and unresolved harmonics,'' J. Acoust. Soc. Am. **98**, 1355–1364.

Plomp, R. (**1967**). ''Pitch of complex tones,'' J. Acoust. Soc. Am. **41**, 1526–1533.

Ritsma, R. J. (**1963**). ''On pitch discrimination of residue tones,'' Int. Audiol. **2**, 34–47.

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.

Singh, P. G., and Hirsh, I. J. (**1992**). ''Influence of spectral locus and F0 changes on the pitch and timbre of complex tones,'' J. Acoust. Soc. Am. **92**, 2650–2661.

Smoorenburg, G. F. (**1970**). ''Pitch perception of two-frequency stimuli,'' J. Acoust. Soc. Am. **48**, 924–941.

Walliser, K. (**1969**). ''Zusammenhänge zwischen dem Schallreiz und der Periodentonhöhe,'' Acustica **21**, 319–328.

White, L. J., and Plack, C. J. (**1998**). ''Temporal processing of the pitch of complex tones,'' J. Acoust. Soc. Am. **103**, 2051–2063.