Perceived continuity and pitch shifts for complex tones with unresolved harmonics

Christopher J. Plack^{a)}

Human Communication and Deafness Division, University of Manchester, Manchester M13 9PL, United Kingdom

Rebecca K. Watkinson

Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom

(Received 5 November 2009; revised 27 July 2010; accepted 27 July 2010)

Brief complex tone bursts with fundamental frequencies (F0s) of 100, 125, 166.7, and 250 Hz were bandpass filtered between the 22nd and 30th harmonics, to produce waveforms with five regularly occurring envelope peaks (“pitch pulses”) that evoked pitches associated with their repetition period. Two such tone bursts were presented sequentially and separated by a silent interval of two periods (2/F0). When the relative phases of the two bursts were varied, such that the interpulse interval (IPI) between the last pulse of the first burst and the first pulse of the second burst was varied, the pitch of the whole sequence was little affected. This is consistent with previous results suggesting that the pitch integration window may be “reset” by a discontinuity. However, when the interval between the two bursts was filled with a noise with the same spectral envelope as the complex, variations in IPI had substantial effects on the pitch of the sequence. It is suggested that the presence of the noise causes the two tones bursts to appear continuous, hence, resetting does not occur, and the pitch mechanism is sensitive to the phase discontinuity across the silent interval.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3479757]

PACS number(s): 43.66.Hg, 43.66.Ba, 43.66.Mk [BCM]

Pages: 1922–1929

I. INTRODUCTION

This paper is concerned with the mechanisms used by the auditory system to determine the pitch of complex tones consisting of harmonics that are unresolved by the peripheral auditory system (harmonic numbers above about eight). Although unresolved harmonics are less dominant in pitch perception than resolved harmonics, understanding these mechanisms is important for a general understanding of purely temporal pitch processing: Unresolved harmonics, by definition, do not convey spectral information to the auditory system and therefore provide experimental control. Furthermore, listeners with sensorineural hearing loss and cochlear implant users are dependent on the information from unresolved harmonics due to reduced frequency selectivity (Moore and Carlyon, 2005). Studying pitch perception for unresolved harmonics can shed light on the deficits experienced by these individuals, and aid the development of signal processing strategies that correct for these deficits.

Fundamental frequency (F0) discrimination for complex tones with unresolved harmonics shows a large effect of duration. For example, a doubling in the number of waveform periods from five to ten can result in a threefold increase in the detectability index, d' (White and Plack, 1998). This is much greater than the optimal increase of $\sqrt{2}$ (1.414) predicted by signal detection theory on the basis of a doubling in the number of independent samples (Viemeister and Wakefield, 1991). However, if two five-period tone bursts are

separated by a brief gap of 5 ms or more, F0 discrimination matches the independent samples prediction, such that d' for discrimination between two pairs of such bursts is $\sqrt{2}$ greater than that for a single burst (White and Plack, 1998).

To account for these and similar findings (Plack and Carlyon, 1995; Plack and White, 2000b; 2000a; Watkinson *et al.*, 2005), it has been suggested that the auditory system uses a long integration time for complex tones consisting of unresolved harmonics, and that this integration time can be “reset” by discontinuities, such as periods of silence, that might indicate the onset of a new auditory object that requires separate analysis (White and Plack, 1998; Plack and White, 2000b; 2000a; Watkinson *et al.*, 2005). When the integration time is reset, the auditory system takes two independent pitch samples and performance reverts to the predictions of signal detection theory. Previously, Bregman *et al.* (1994a; 1994b) had suggested that the auditory system uses neural onset and offset responses to reset pitch processing and to aid auditory scene analysis.

The hypothetical resetting mechanism has also been investigated by observing the pitch changes produced by instantaneous changes in phase. Nabelek (1996) showed that a phase change between two contiguous short pure tones produced a change in pitch, but when the silent interval between the two tone bursts reached a “critical pause duration” of between 8 and 16 ms, this effect was not observed and the two tone bursts appeared to be processed separately. Plack and White (2000b) and Watkinson *et al.* (2005) employed a similar technique to Nabelek (1996), but used complex tone bursts with unresolved harmonics and applied the phase changes to the envelope of the waveform. By manipulating

^{a)}Author to whom correspondence should be addressed. Electronic mail: chris.plack@manchester.ac.uk

the starting phases of the two tone bursts they introduced variations in the interpulse interval (IPI) between the last envelope peak (or “pitch pulse”) of the first burst and the first envelope peak of the second burst. The IPI shifts varied between -0.75 periods and 0.75 periods, where negative numbers indicate a phase advance. When the tone bursts were contiguous, the shifts in IPI produced substantial shifts in pitch. However, when there was a gap between the bursts of one period or more, the shifts in IPI had little effect on the pitch.

Plack and White (2000a) investigated whether or not the resetting mechanism is sensitive to illusory continuity. They presented a noise burst between two 250-Hz complex tone bursts separated by a gap such that the percept was of a single continuous burst interrupted by a noise. When the noise was added, performance improved to that observed when the tone really was continuous. The results suggest that the resetting mechanism may be sensitive to the grouping mechanisms underlying perceptual restoration (Warren, 1970; Warren *et al.*, 1972). When the noise is present, there is no evidence that the tone is discontinuous and hence the resetting mechanism is not triggered. An interruption in the regular pattern of pitch pulses is not sufficient to cause resetting.

The second experiment of Plack and White (2000a) affords an intriguing insight into the processing involved in the pitch integration mechanism. Following Plack and White (2000b), they introduced a phase delay of 0.75 periods between two bursts separated by a two-period gap, such that the IPI across the gap was increased by 0.75 periods. When the gap was silent, there was no effect of the phase delay on pitch, consistent with independent processing. However, when a noise with the same spectral envelope as the tone bursts was presented in the gap to produce the illusion of continuity, the phase delay produced a small but significant decrease in the pitch of the pair of bursts. The finding contrasts with that found in a similar situation involving frequency modulation (FM) imposed on a pure tone carrier. Listeners heard the FM continue during the noise interruption but were insensitive to the phase at which the FM resumed after the interruption (Carlyon *et al.*, 2004). This discrepancy may be due in part to the large difference in gap duration between the experiments; 8 ms in Plack and White (2000a) compared to 200 ms in Carlyon *et al.* (2004). It is possible that, over these long intervals, FM is encoded as a “feature” of the sound, and that phase information is discarded.

The finding of Plack and White (2000a) suggests that, in-situations of perceived continuity, the pitch integration mechanism is sensitive to phase relations over intervals of several periods. This has implications for models of pitch perception. For example, the finding is not predicted by models based on the first peak in the autocorrelation function (Meddis and Hewitt, 1991; Meddis and O’Mard, 1997). The phase sensitivity effect suggests further that pitch is extracted from a combination of the temporal information in the two bursts, rather than pitch being derived from each burst independently and then combined. Given the potential importance of the finding for our understanding of pitch integration

mechanisms for complex tones, the aim of the present experiment was to attempt to replicate and extend the results of Plack and White (2000a), using a wider range of F0s and a wider range of phase shifts, to determine the characteristics of the phenomenon and the generality of the finding. A specific aim was to test the hypothesis that the pitch mechanism processes two bursts connected by perceived continuity in the same way as a continuous burst. To this end, the results were compared with the no-gap conditions of Watkinson *et al.* (2005), who measured pitch shifts using the same phase shifts and the same F0s as those tested here.

II. METHOD

A. Stimuli

The stimuli were similar to those used in Experiment 1 of Watkinson *et al.* (2005). Stimuli consisted of complex-tone bursts with F0s of 100, 125, 166.7, and 250 Hz. Each harmonic had a level of 50 dB SPL before filtering. The harmonics were added in sine phase relative to each other (i.e., positive-going zero-crossings were aligned) and band-pass filtered (digital finite impulse response) between the 22nd and the 30th harmonic (3-dB downpoints, 90-dB/octave slope). The overall level of the stimuli was therefore about 60 dB SPL. All filtering occurred after gating. Each tone burst had a duration of 5 waveform cycles (i.e., 50, 40, 30, and 20 ms for F0s of 100, 125, 166.7, and 250 Hz, respectively), and was gated with no ramps.

The comparison interval consisted of two tone bursts presented consecutively and separated by an interval of two waveform periods (where the periods were 10, 8, 6, and 4 ms for F0s of 100, 125, 166.7, and 250 Hz, respectively). The standard interval was the same except that the IPI between the bursts was varied by producing a phase change in the envelope of the second tone burst relative to the first. The (nominal) waveform before filtering was given by:

$$a(t) = A \sum_{n=1}^{\infty} \sin[n(2\pi F_0 t + \theta)] \quad (0 \leq t < 5/F_0), \quad (1)$$

$$a(t) = A \sum_{n=1}^{\infty} \sin[n(2\pi F_0 t + \theta + \varphi)] \quad (7/F_0 \leq t < 12/F_0), \quad (2)$$

where t is time in seconds, $a(t)$ is the waveform of the complex, A is the peak amplitude of each harmonic, n is the harmonic number, F_0 is F0 in Hz, θ is the starting phase of the F0 component ($n=1$) of the first tone burst, and φ is the starting phase of the F0 component of the second tone burst relative to the first. As explained in Watkinson *et al.*, (2005) the relative values of θ and φ were varied in the comparison interval to increase or decrease the IPI between the last pitch pulse (envelope peak) of the first burst and the first pitch pulse (envelope peak) of the second burst. IPI shift values were -0.75 , -0.5 , -0.25 , 0 , 0.25 , 0.5 , and 0.75 waveform periods, where 0 represents no shift (i.e., comparison identical to standard), negative numbers represent a phase advance of the second burst relative to the first, and positive numbers

TABLE I. The phase shifts used in the experiment. For each F0 (in Hz), the table shows the starting phase of the first burst (θ), the starting phase of the second burst (φ), the IPI across the gap (in ms), and the IPI shift (in periods).

F0	θ	φ	IPI	Shift
100	$\pi/4$	$3\pi/2$	22.5	-0.75
	$\pi/4$	π	25	-0.5
	π	$\pi/2$	27.5	-0.25
	π	0	30	0
	π	$3\pi/2$	32.5	0.25
	$7\pi/4$	π	35	0.5
	$7\pi/4$	$\pi/2$	37.5	0.75
125	$\pi/4$	$3\pi/2$	18	-0.75
	$\pi/4$	π	20	-0.5
	π	$\pi/2$	22	-0.25
	π	0	24	0
	π	$3\pi/2$	26	0.25
	$7\pi/4$	π	28	0.5
	$7\pi/4$	$\pi/2$	30	0.75
166.7	$\pi/4$	$3\pi/2$	13.5	-0.75
	$\pi/4$	π	15	-0.5
	π	$\pi/2$	16.5	-0.25
	π	0	18	0
	π	$3\pi/2$	19.5	0.25
	$7\pi/4$	π	21	0.5
	$7\pi/4$	$\pi/2$	22.5	0.75
250	$\pi/4$	$3\pi/2$	9	-0.75
	$\pi/4$	π	10	-0.5
	π	$\pi/2$	11	-0.25
	π	0	12	0
	π	$3\pi/2$	13	0.25
	$7\pi/4$	π	14	0.5
	$7\pi/4$	$\pi/2$	15	0.75

represent a phase delay of the second burst relative to the first. Table I lists the conditions in terms of the phase shift in periods and the IPI in ms.

The interval between the two bursts was either silent, or filled with a bandpass “filler” noise. The filler noise had a spectrum level of 30 dB SPL for the tones with an F0 of 100 Hz, 29 dB SPL for the tones with an F0 of 125 Hz, 28 dB SPL for the tones with an F0 of 166.7 Hz, and 26 dB SPL for the tones with an F0 of 250 Hz. The filler noise burst levels were chosen to have the same average spectral density as the tones coming before and after. The noise duration was equal to the duration of the interval between the bursts (i.e., two

waveform periods) and was gated with no ramps prior to filtering. The noise was bandpass filtered into the same frequency region as the complex tones, i.e., a frequency region ranging from the 22nd to the 30th harmonics of the complex tone. An informal listening test by the authors revealed that the noise was sufficient to produce the perception of a continuous tone. Example waveforms of the stimuli with the filler noise are illustrated in Fig. 1.

Finally, the tone bursts were presented in a background noise (generated independently for each interval) with a spectrum level of 15 dB. The noise was low-pass filtered at 1000 Hz for the tones with an F0 of 100 Hz, at 1250 Hz for the tones with F0s of 125 and 166.7 Hz, and at 2500 Hz for the tones with an F0 of 250 Hz. The low-pass noise was gated on (no ramps) 50 ms before the onset of the tone bursts, and gated off 50 ms after the offset of the tone bursts. The noise was presented to mask resolved low-frequency combination tones, although it is likely that some unresolved components between the high-frequency cut-off of the noise and the low-frequency cut-off of the complex tone (such as the cubic difference tone) would not have been masked. Distortion components arising from the filler noise may also have been propagated to this region. However, it is possible that distortion products from the tone bursts could have reduced the strength of the continuity effect (in which case the effect of the filler noise described in the Results section may have been an underestimate of the true effect of continuity), although, as noted above, the percept was of a continuous burst.

Stimuli were generated digitally with 32-bit resolution and were output via an RME Digi96/8 PAD 24-bit soundcard. The sampling rate was 48 kHz. The stimuli were presented to the right ear via Sennheiser HD 580 headphones. The output from the sound card was fed to the headphones via a patch-panel in the sound booth without filtering or amplification.

B. Procedure

Pitch matches were obtained using the adaptive procedure described by Jesteadt (1980) and by Plack and White (2000b). For each trial, two listening intervals (separated by 500 ms) were presented. Each listening interval contained a tone burst pair. The standard interval contained the shifted stimulus with an F0 of 100, 125, 166.7, or 250 Hz. The

FIG. 1. Example time waveforms for the filler-noise conditions used in the experiment (excluding the low-pass noise). The titles indicate the IPI-shift condition.

comparison interval contained a similar stimulus without the IPI shift (preserved envelope phase), but with the same silent interval as the standard. The F0 of the comparison tone burst pair was varied adaptively using two interleaved adaptive tracks. Initially, the F0 of the comparison tone burst pair was 20% above (higher track) or 20% below (lower track) the F0 of the standard tone burst pair.

Listeners were presented with the standard and a comparison (with either a higher or lower F0), in a random order. The listener's task was to decide which interval contained the higher pitched sound. The higher track used a "two-down one-up" rule. If the listener chose the comparison interval twice in succession then the F0 of the comparison was decreased for the next trial. If the listener chose the standard interval once then the F0 of the comparison was increased for the next trial. The lower track used a "two-up one-down" rule. If the listener chose the comparison interval once then the F0 of the comparison was decreased for the next trial. If the listener chose the standard interval twice in succession then the F0 of the comparison was increased for the next trial. The F0 of the comparison interval was increased or decreased by 4% of the F0 of the standard for the first four reversals, and by 2% thereafter.

Each trial was selected randomly from the higher or lower track. Testing continued until the comparison interval had changed from an increasing to decreasing F0, or vice versa, 16 times for each track. The threshold of the track was calculated from the arithmetic mean of the comparison F0s (in % relative to the F0 of the standard tone burst) for the last 12 transitions for that track. The two tracks bracketed the pitch match. Effectively, the upper track converged on the 70.7% point on the psychometric function for detecting an increase in F0 and the lower track converged on the 29.3% point for detecting an increase in F0 (this is also the 70.7% point for detecting a decrease in F0). The pitch match was taken as the arithmetic mean of the thresholds for the upper and lower tracks. Half the difference between the thresholds from the upper and lower tracks was taken as an estimate of the F0 difference limen (F0DL, the smallest detectable change in the F0).

All the blocks for a single F0 were presented together. The order in which the F0s were presented was random. For each F0, the order of conditions was randomized. Eight pitch matches were collected for each condition. The final estimate was the arithmetic mean of these eight.

Listeners sat in a double-walled sound-attenuated booth and the responses were recorded via a computer keyboard. Listeners could see a computer monitor, through a window in the sound booth, on which "lights" were presented concomitantly with each stimulus interval. No feedback was given. The participants were tested for a maximum of two hours consecutively with breaks, although it was more usual for participants to be tested for one hour at a time.

C. Listeners

Four normally hearing listeners were tested. They were aged between 22 and 60 years. Listeners were given at least two hours training on the task before data collection began.

III. RESULTS

Figure 2 shows the pitch shift results for the individual listeners. As in previous studies (Plack and White, 2000b; Watkinson *et al.*, 2005), the results are plotted relative to the pitch shift for the reference condition with no IPI shift, to correct for any bias in the matching procedure. Such a bias can arise if, for example, the F0DLs for the upper and lower tracks differ. Some of the listeners showed considerable variability in their matches, particularly L1 and L2 at 250 Hz. However, a consistent pattern can be seen in the data, which is summarized by the mean data shown in Fig. 3. For the conditions without the filler noise, there was little consistent pitch shift with shifts in the IPI. However, for the conditions with the filler noise, the pitch shifted downwards for both positive and negative shifts in the IPI relative to the reference. The overall pattern for the noise conditions is broadly consistent across the four F0s.

A three-factor ANOVA (F0 X Noise X IPI Shift) conducted on the individual pitch shift data revealed a highly significant main effect of IPI Shift [$F(6,18)=9.14$, $p=0.00011$]. The main effects of F0 and of Noise were not significant, but there was a highly significant interaction between Noise and IPI Shift [$F(6,18)=9.57$, $p<0.0001$]. No other interactions were significant. To explore this effect further, two two-factor ANOVAs (F0 X IPI Shift) were conducted on the results for the no-noise and noise conditions separately. For the no-noise conditions, none of the effects were significant. For the noise conditions, there was no significant effect of F0, but a highly significant effect of IPI Shift [$F(6,18)=23.86$, $p<0.0001$]. The interaction was not significant. These analyses confirm the visual impression that there was little effect of IPI shift when there was no filler noise, but that there was an effect of IPI shift when filler noise was present.

As described in Section II B, the pitch estimation procedure also produces an estimate of the F0DL for each condition. The individual results are shown in Fig. 4. The F0DLs for the 250 Hz conditions are very high in some cases, which partly explain the variability in the pitch matches. For L1 at 167 Hz, and for L2, L3, and L4 at 250 Hz, the filler noise caused a reduction in the F0DL, as expected from previous results at 250 Hz (Plack and White, 2000a). Individual one-tailed t-tests, Bonferroni corrected for the 16 possible comparisons ($p<0.003$), revealed that the difference between noise and no-noise conditions was significant in all four cases. However, for other cases there was little effect of addition of the noise.

IV. DISCUSSION

A. Perceived continuity and resetting

The present results confirm and extend the finding of Plack and White (2000a), who used a single F0 and a single phase shift. When there is a two-period gap between two complex tone bursts, there is no evidence that the auditory system is sensitive to the envelope phase across the gap, at least with respect to the derivation of pitch. However, when the gap is filled with noise so that the two bursts appear

FIG. 2. Individual pitch shifts for the four F0s as a function of the IPI shift between the bursts and the presence or absence of filler noise. Pitch shifts are plotted relative to the pitch shift for the condition with no IPI shift. Error bars show standard errors.

continuous, the pitch mechanism is sensitive to the phase relations across the gap. The pattern of pitch shifts was consistent across all four F0s tested.

The bold dashed lines in Fig. 3 show the mean results from Watkinson *et al.* (2005) for the no-gap conditions. With the exceptions of the absence of filler noise and the inclusion of conditions without a gap, the stimulus parameters and procedure for that earlier study were identical to those for the present study. It can be seen that the earlier data match the present data quite well, particularly at 250 Hz. For the lower frequencies, there are some discrepancies. In particular, the slight rise in pitch for the -0.25 period IPI shift observed in

the older data is not seen in the present results. Overall, the comparison suggests that the response of the system to two tone bursts separated by filler noise is similar to that observed when there is no gap between the tone bursts, although the absolute effect sizes are sometimes larger for the genuinely continuous bursts.

The discrimination results are largely consistent with the results of Plack and White (2000a), who reported a reduction in the F0DL when a noise was introduced into the gap between two 250-Hz tone bursts. However, the present experiment suggests that this improvement is generally not observed for lower F0s. It is not clear if this is an effect of F0

FIG. 3. Mean pitch shifts for the four F0s as a function of the IPI shift between the bursts and the presence or absence of filler noise. Also shown are the mean pitch shifts for the no-gap conditions of Watkinson *et al.* (2005).

FIG. 4. Individual F0DLs for the four F0s as a function of the IPI shift between the bursts and the presence or absence of filler noise. Error bars show standard errors.

per se, or an effect of the size of the F0DL in the no-noise conditions. Watkinson *et al.* (2005) observed a reduction in F0DLs for continuous 10 pitch-pulse tone bursts compared to two 5 pitch-pulse tone bursts, for F0s of 167 and 250 Hz. However, little benefit of continuity was observed for F0s of 100 and 125 Hz. Similarly, White and Plack (1998) found little improvement in performance for an increase in duration from 5 to 10 pitch pulses for a 62.5-Hz F0. So it is possible that the effects of increased continuous duration, real or perceived, are not beneficial for low F0s over this range of durations.

Overall, the results add further support to the pitch resetting hypothesis (White and Plack, 1998; Plack and White, 2000b; 2000a; Watkinson *et al.*, 2005): when the auditory system determines that there is a discontinuity in a tone, the pitch mechanism treats the portions before and after the discontinuity as distinct, and derives independent pitch estimates. These and previous results suggest that a discontinuity in the regular sequence of pitch pulses is not sufficient to produce resetting if there is a continuity of level caused by the introduction of the filler noise. The phase sensitivity effect suggests further that, in the absence of resetting, pitch is extracted from a combination of the temporal information in the two bursts (a long integration time), rather than pitch being derived from each burst independently and then combined.

B. Implications for models of pitch perception

Watkinson *et al.* (2005) interpreted their results in terms of a modified version of the mean-interval model of pitch perception (Carlyon, 1997; Carlyon *et al.*, 2002). The mean-interval model assumes that the pitch of complex tones containing unresolved harmonics, and the pitch of impulse sequences presented to a cochlear implant, is determined by a weighted mean of the first order IPIs, equivalent to stimulus duration divided by the number of pulses for a regular sequence. The version described by Carlyon *et al.* (2002) was used to account for the pitch of a sequence of pulses with alternating 4- and 6-ms IPIs. The pitch of this sequence was matched to that of an isochronous sequence with a period between 5.6 and 5.7 ms (Carlyon *et al.*, 2002). To account for the finding that the period of the matching stimulus is longer than the mean of the 4- and 6-ms IPIs (i.e., 5 ms), in the model of Carlyon *et al.* longer IPIs are weighted more heavily than shorter IPIs.

Watkinson *et al.* (2005) found that the Carlyon *et al.* (2002) model could not account for their data that showed the same pattern of pitch shifts, relative to the phase shift in periods, across a range of F0s (see Fig. 3). The model failed mainly because the pitch shifts in the data produced by a change in a single IPI were dependent upon the *proportion* of the waveform period by which the single IPI was shifted, not

upon the *absolute* value of the IPI as suggested by the model. In other words, the original model does not weight the IPIs appropriately for low F0s. To account for their results, Watkinson *et al.* (2005) devised a new version of the mean-interval model in which pulses could be “ignored” on a probabilistic basis. Effectively this means that pulses, rather than intervals, are weighted. An important feature of the model is that the weighting, and therefore the reduction in pitch for short IPIs, depends on the IPI *relative* to the most common period of the waveform. The present data provide further support for this feature of the model, since a similar variation in pitch with phase shift in periods was observed here (see Fig. 3).

van Wieringen *et al.* (2003) estimated the pitch of an alternating 4–6 ms IPI sequence in both acoustic and electric (cochlear implant) hearing by comparing with isochronous sequences. The alternating sequence was amplitude modulated in two different phases, so that pulses after either the shorter or longer IPIs were attenuated. For both acoustic and electric stimulation, pitch matches were lower for the condition in which the pulses were attenuated after the shorter interval. The interpretation of van Wieringen *et al.* (2003) was that attenuation of the pulse after the shorter interval made it more likely to be further attenuated due to falling within the auditory-nerve refractory period of the preceding pulse. Consistent with this hypothesis, the authors found little evidence for the effect when they used an 8–12 ms sequence, presumably because the longer IPIs were outside the refractory period. This result is inconsistent with the model described by Watkinson *et al.* (2005).

Using a similar technique, Carlyon *et al.* (2008) estimated the pitches of alternating 3.5–5.5 ms, 4–6 ms, and 4.5–6.5 ms pulse sequences. The ratio of IPI in the matching stimulus to mean IPI in the test stimulus was greater for the sequence with shorter IPIs. This is consistent with the idea that the change in refractoriness between 3.5 and 5.5 ms is greater than that between 4.5 and 6.5 ms. These results are also inconsistent with the model of Watkinson *et al.* (2005) which assumes that weighting of pulses depends on the IPI relative to the common IPI. Carlyon *et al.* also recorded the compound action potential (CAP) in response to pulse sequences from human listeners and from guinea pigs. These revealed that, in a 4–6 ms sequence, the auditory nerve representations of pulses after a 4-ms IPI were indeed attenuated relative to pulses after a 6-ms IPI. However, in additional guinea pig recordings, the relative attenuation was similar for a 3.5–5.5 ms sequence and for a 4.5–6.5 ms sequence. As suggested by the authors, the discrepancy with the behavioral data could be a consequence of refractoriness at a more central stage in the auditory pathway.

From the CAP results of Carlyon *et al.*, it seems likely that refractoriness modifies the representation of pulse trains in the auditory nerve. Hence, it is possible that the reduction in the response to the pulse following a short IPI could have contributed to the reduction in pitch at negative shifts observed in the results of Watkinson *et al.* (2005), since the attenuated pulse may mean that, effectively, a longer IPI (between the pulses before and after the attenuated pulse) is conveyed to the brain. In addition, it is possible that the filler

noise may have affected the representation of the first pulse in the second tone burst in the present experiments. Because of the way the stimuli were generated, for the more negative shifts, the noise ends just before the onset of the first pulse (see Fig. 1). This might be expected to have a refractory effect on this first pulse, reducing its contribution and lowering the pitch.

Refractory effects in the auditory nerve might be expected to be dependent on the *absolute* IPI (and on the *absolute* delay between the offset of the noise and the onset of the pulse). However, the data suggest a pitch shift effect that is related to the delay *relative* to the period of the pulse train. There is little indication in the mean data from the two experiments (Fig. 3) that the same absolute IPI produces the same pitch shift. For example, for the results of Watkinson *et al.*, the -0.5 shift at 250 Hz had a smaller IPI (2 ms) than the -0.5 shifts at the other F0s and the -0.75 shift at 100 Hz (2.5-ms IPI), but the pitch shifts were more negative for the lower F0s. Similarly, in the present data, although there is a slight decrease in pitch for F0s from 100 to 167 Hz for the -0.75 shift, the effect reverses at 250 Hz. Hence there is no compelling evidence from the present data or from those of Watkinson *et al.* that the pitch shifts were dominated by refractory effects, although the possibility remains that refractoriness had a role. For example, relative refractory effects for a fixed IPI may be greater for lower pulse rates, since for higher rates most of the pulses may be partly attenuated by ongoing refractoriness.

Autocorrelation models of pitch perception (Licklider, 1951; Slaney and Lyon, 1990; Meddis and Hewitt, 1991; Meddis and O’Mard, 1997; de Cheveigné, 1998; Balaguer-Ballester *et al.*, 2008) are based on the assumption that the auditory system has the ability to compute the correlation of a delayed version of the stimulus with the original for a range of delays. As described in the Introduction, the present data are problematic for models that depend on an analysis of the early peaks in the autocorrelation function (Meddis and Hewitt, 1991; Meddis and O’Mard, 1997). First, because a shifted stimulus still contains a strong correlation at $1/F_0$, $2/F_0$, and $3/F_0$; only the longer correlation delays are affected (Plack and White, 2000b; Watkinson *et al.*, 2005). Second, because the effects of filler noise suggest that, in certain circumstances at least, the system is sensitive to a phase shift across delays of at least three periods (the gap between the bursts). These data support an approach to pitch derivation based on the Euclidean distance between the autocorrelation functions for the test tone and a reference tone over a wide range of delays, and over a relatively long integration time, such as the recent version of the autocorrelation model proposed by Balaguer-Ballester *et al.* (2008). This model can account for the pitch of the anisochronous 4–6 sequences of Carlyon *et al.* (2002), suggesting that it may be an alternative to mean-interval models for simulating the effects of IPI shifts. Another version of the model published recently can account qualitatively for the effects of resetting and perceived pitch continuity by including descending feedback modulation of pitch integration times (Balaguer-Ballester *et al.*, 2009). When the feedback system detects a change in the stimulus, the integration times are shortened,

effectively resetting the system. The filler noise prevents this, leading to a long integration time that generates the perception of continuity.

V. CONCLUSIONS

- (i) Shifts in a single IPI between two unresolved complex-tone bursts separated by a two-period gap produced no shift in pitch when the gap was silent, but a negative shift in pitch when the gap was filled with filler noise with the same spectral envelope as the tone bursts.
- (ii) The negative pitch shift was observed for both positive and negative changes in the IPI relative to an in-phase reference. The pattern of results was consistent across F0s of 100, 125, 167, and 250 Hz.
- (iii) In conditions of perceived continuity, the pitch mechanism is sensitive to phase changes across a discontinuity, consistent with the hypothesis that the filler noise causes the tone bursts to be processed as if they really were continuous.

ACKNOWLEDGMENTS

Author RW was supported by an EPSRC doctoral training Grant No. (GR/NO7219). We are grateful to the Associate Editor (Brian Moore), and to Bob Carlyon and an anonymous reviewer for constructive and insightful comments on earlier versions of the manuscript.

Balaguer-Ballester, E., Clark, N. R., Coath, M., Krumbholz, K., and Denham, S. L. (2009). "Understanding pitch perception as a hierarchical process with top-down modulation," *PLOS Comput. Biol.* **5**, e1000301.

Balaguer-Ballester, E., Denham, S. L., and Meddis, R. (2008). "A cascade autocorrelation model of pitch perception," *J. Acoust. Soc. Am.* **124**, 2186–2195.

Bregman, A. S., Ahad, P., Kim, J., and Melnerich, L. (1994a). "Resetting the pitch-analysis system: 1. Effects of rise times of tones in noise backgrounds or of harmonics in a complex tone," *Percept. Psychophys.* **56**, 155–162.

Bregman, A. S., Ahad, P. A., and Kim, J. (1994b). "Resetting the pitch-analysis system. 2. Role of sudden onsets and offsets in the perception of individual components in a cluster of overlapping tones," *J. Acoust. Soc. Am.* **96**, 2694–2703.

Carlyon, R. P. (1997). "The effects of two temporal cues on pitch judgments," *J. Acoust. Soc. Am.* **102**, 1097–1105.

Carlyon, R. P., Mahendran, S., Deeks, J., Long, C. J., Axon, P., Baguley, D.,

Bleeck, S., and Winter, I. M. (2008). "Behavioral and physiological correlates of temporal pitch perception in electric and acoustic hearing," *J. Acoust. Soc. Am.* **123**, 973–985.

Carlyon, R. P., Micheyl, C., Deeks, J. M., and Moore, B. C. J. (2004). "Auditory processing of real and illusory changes in frequency modulation (FM) phase," *J. Acoust. Soc. Am.* **116**, 3629–3639.

Carlyon, R. P., van Wieringen, A., Long, C. J., Deeks, J. M. and Wouters, J. (2002). "Temporal pitch mechanisms in acoustic and electric hearing," *J. Acoust. Soc. Am.* **112**, 621–633.

de Cheveigné, A. (1998). "Cancellation model of pitch perception," *J. Acoust. Soc. Am.* **103**, 1261–1271.

Jesteadt, W. (1980). "An adaptive procedure for subjective judgments," *Percept. Psychophys.* **28**, 85–88.

Licklider, J. C. R. (1951). "A duplex theory of pitch perception," *Experientia* **7**, 128–134.

Meddis, R., and Hewitt, M. (1991). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification," *J. Acoust. Soc. Am.* **89**, 2866–2882.

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

Moore, B. C. J., and Carlyon, R. P. (2005). "Perception of pitch by people with cochlear hearing loss and by cochlear implant users," in *Pitch: Neural Coding and Perception*, edited by C. J. Plack, A. J. Oxenham, R. R. Fay, and A. N. Popper (Springer, New York), Vol. **24**, pp. 234–277.

Nabelek, I. V. (1996). "Pitch of a sequence of two short tones and the critical pause duration," *Acust. Acta Acust.* **82**, 531–539.

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.

Plack, C. J., and White, L. J. (2000a). "Perceived continuity and pitch perception," *J. Acoust. Soc. Am.* **108**, 1162–1169.

Plack, C. J., and White, L. J. (2000b). "Pitch matches between unresolved complex tones differing by a single interpulse interval," *J. Acoust. Soc. Am.* **108**, 696–705.

Slaney, M., and Lyon, R. F. (1990). "A perceptual pitch detector," *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pp. 357–360.

van Wieringen, A., Carlyon, R. P., Long, C. J., and Wouters, J. (2003). "Pitch of amplitude-modulated irregular-rate stimuli in acoustic and electric hearing," *J. Acoust. Soc. Am.* **114**, 1516–1528.

Viemeister, N. F., and Wakefield, G. H. (1991). "Temporal integration and multiple looks," *J. Acoust. Soc. Am.* **90**, 858–865.

Warren, R. M. (1970). "Perceptual restoration of missing speech sounds," *Sciences (N.Y.)*, **167**, 392–393.

Warren, R. M., Obusek, C. J., and Ackroff, J. M. (1972). "Auditory induction: Perceptual synthesis of absent sounds," *Science* **176**, 1149–1151.

Watkinson, R. K., Plack, C. J., and Fantini, D. A. (2005). "Pitch shifts for complex tones with unresolved harmonics and the implications for models of pitch perception," *J. Acoust. Soc. Am.* **118**, 934–945.

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