

Combination of Spectral and Binaurally Created Harmonics in a Common Central Pitch Processor

HEDWIG E. GOCKEL¹, ROBERT P. CARLYON¹, AND CHRISTOPHER J. PLACK²

¹*MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge, CB2 7EF, UK*

²*Division of Human Communication and Deafness, The University of Manchester, Manchester, M13 9PL, UK*

Received: 7 October 2010; Accepted: 31 October 2010; Online publication: 18 November 2010

ABSTRACT

A fundamental attribute of human hearing is the ability to extract a residue pitch from harmonic complex sounds such as those produced by musical instruments and the human voice. However, the neural mechanisms that underlie this processing are unclear, as are the locations of these mechanisms in the auditory pathway. The ability to extract a residue pitch corresponding to the fundamental frequency from individual harmonics, even when the fundamental component is absent, has been demonstrated separately for conventional pitches and for Huggins pitch (HP), a stimulus without monaural pitch information. HP is created by presenting the same wideband noise to both ears, except for a narrowband frequency region where the noise is decorrelated across the two ears. The present study investigated whether residue pitch can be derived by combining a component derived solely from binaural interaction (HP) with a spectral component for which no binaural processing is required. Fifteen listeners indicated which of two sequentially presented sounds was higher in pitch. Each sound consisted of two “harmonics,” which independently could be either a spectral or a HP component. Component frequencies were chosen such that the relative pitch judgement revealed whether a residue pitch was heard or not. The results showed that listeners were equally likely to perceive a residue pitch when one component was dichotic and the other was spectral as when the components were both spectral or both dichotic. This suggests that there exists a single mechanism for the

derivation of residue pitch from binaurally created components and from spectral components, and that this mechanism operates at or after the level of the dorsal nucleus of the lateral lemniscus (brainstem) or the inferior colliculus (midbrain), which receive inputs from the medial superior olive where temporal information from the two ears is first combined.

Keywords: residue pitch, dichotic pitch, diotic pitch, Huggins pitch

INTRODUCTION

A tonal percept can be produced by a pure tone or a narrowband noise (NBN; see Fig. 1A, top). A pitch can also be produced by presenting a wideband noise diotically, i.e., the same white noise to both ears, except for a narrow frequency region in which the noise is decorrelated between the two ears (see Fig. 1A, bottom). For this “Huggins pitch” (HP) stimulus (Cramer and Huggins 1958), the stimulus presented to each ear alone sounds just like a white noise. However, when the stimulus is presented binaurally, the listener perceives not only a noise, coming from the center of the head, but also a faint tone with a pitch that corresponds to the center frequency of the narrow band that is interaurally decorrelated. The tonal percept is lateralized to one ear or the other in a way that varies idiosyncratically across listeners (Raatgever and Bilsen 1986; Zhang and Hartmann 2008). HP stimuli produce a clear musical pitch, supporting melody recognition (Akeroyd et al. 2001). In contrast to a pure tone or narrowband noise, for which pitch information

Correspondence to: Hedwig E. Gockel · MRC Cognition and Brain Sciences Unit · 15 Chaucer Road, Cambridge, CB2 7EF, UK. Telephone: +44-1223-355294; fax: +44-1223-359062; email: hedwig.gockel@mrc-cbu.cam.ac.uk

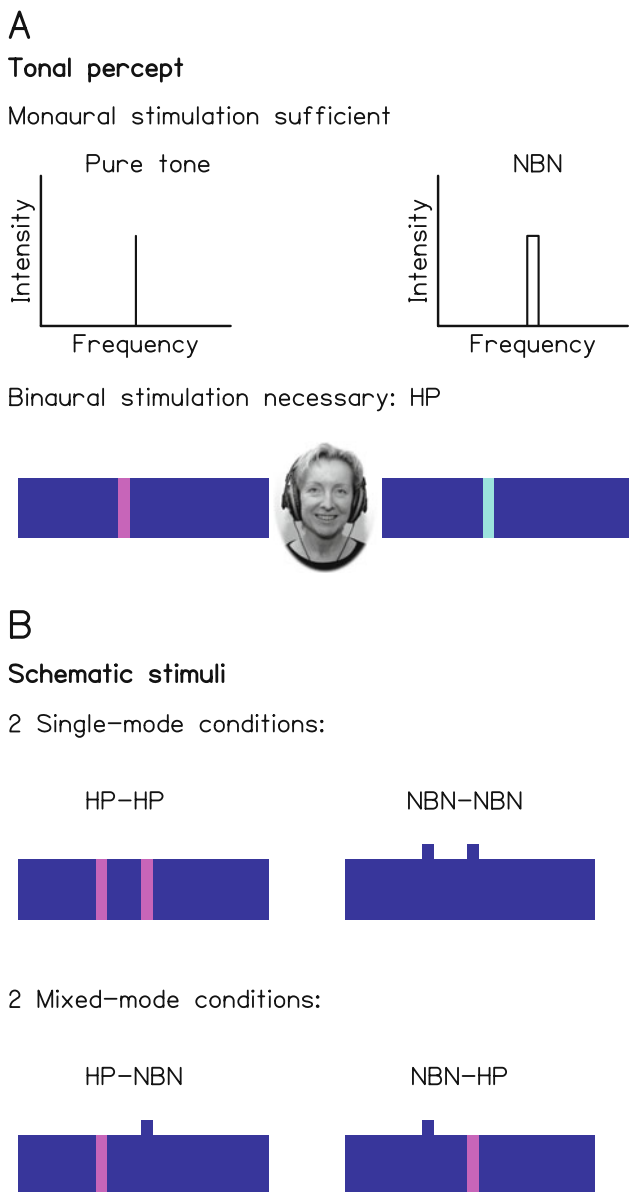


FIG. 1. Schematic diagram of pitch stimuli. **A** Single component stimuli that evoke a tonal percept when presented monaurally (*top*) and for which binaural presentation is necessary (*bottom*). **B** Stimuli with two harmonics used in the present experiment.

is available monaurally at the level of the auditory nerve, the perception of HP depends on fine timing information from the two ears being combined, and thus, depends on central processing (beyond the cochlear nucleus). There is physiological evidence that this processing occurs in the medial superior olive (MSO), and that the results of this brainstem processing can be measured in the inferior colliculus (IC; Palmer and Shackleton 2002).

A complex tone containing several pure tone or NBN harmonics (with spectral components centered on integer multiples of the fundamental frequency, F_0) can lead to the perception of a single pitch, the residue pitch (also called low pitch or virtual pitch),

corresponding to the F_0 , even in the absence of energy at the F_0 itself. The perception of a single residue pitch rather than that of many single components plays a crucial role in analyzing complex auditory scenes in everyday life (Bregman 1990; Darwin and Carlyon 1995). In humans, the ability to extract the residue pitch is present from an early age (Montgomery and Clarkson 1997). The vast majority of studies of residue pitch have used stimuli that do not depend on binaural interactions. However, a residue pitch in the absence of the fundamental can also be perceived when several HP components, centered at harmonic frequencies, are presented (Bilsen 1977; Gockel et al. 2009). For complex HP, binaural presentation is of course essential. This ability to derive a residue pitch, demonstrated for spectral components on the one hand and for binaural components on the other hand, could be due to either a common pitch mechanism or to two different pitch mechanisms, perhaps operating at different stages in the auditory system. The present study investigated whether or not the derivation of a residue pitch from conventional spectral components on the one hand and from binaurally created components on the other is mediated by a common mechanism or by two different mechanisms.

METHODS

Behavioral task

In a two-interval two-alternative forced-choice task, listeners indicated which of the two stimuli had the higher pitch. No feedback was provided in the experiment proper. Each stimulus consisted of two components whose frequencies were explicitly chosen such that the listener's pitch judgment would indicate one interval if a residue pitch was perceived (i.e., if the subject listened synthetically and heard the low or virtual pitch), but would indicate the other interval if no residue pitch was perceived (i.e., if the subject listened analytically to the individual components).

Stimuli

The stimuli were complex tones consisting of two "harmonics" of three different F_0 s: a 400-Hz F_0 complex with components at 400 and 800 Hz (1st and 2nd harmonics), a 267-Hz F_0 complex with components at 533 and 800 Hz (2nd and 3rd harmonics), and a 160-Hz F_0 complex with components at 640 and 800 Hz (4th and 5th harmonics); Table 1. Note that, as the F_0 decreases, the frequency of the lower harmonic increases, while that of the upper harmonic is constant. Thus, pitch judgments based on analytic listening would go in the opposite

TABLE 1

Complex tones that are compared with each other in the main experiment		
<i>F0</i> (Hz)	<i>Frequencies</i>	<i>Harmonic Numbers</i>
400	400+800	1st and 2nd
267	533+800	2nd and 3rd
160	640+800	4th and 5th

As the F_0 decreases (left-most column), the frequency of the lower component increases (center column)

direction to those based on the residue pitch, i.e., based on synthetic listening. There were four conditions (Fig. 1B). In condition HP-HP, both components were HPs. In condition NBN-NBN, both components were NBNs. In condition HP-NBN, the lower component was a HP, while the upper component was a NBN, and in condition NBN-HP, the lower component was a NBN, while the upper component was a HP. The first two and the last two will be referred to as single-mode and mixed-mode conditions, respectively.

All stimuli were generated digitally in MATLAB. The HP stimuli were generated from a 500-ms Gaussian noise that was low-pass filtered at 2 kHz. They were generated in the spectral domain by first applying a fast Fourier transform (FFT) to the noise and then modifying the phases of one of two matched buffers representing the left and right channels. The modification consisted of linearly increasing (the original) phase as a function of frequency over a range of 0 to 2π radians between 3% below and 3% above the center frequency of the chosen component(s). Applying an inverse FFT to the two spectral buffers gave the signal waveforms for the left and right channels. Note that after this phase modification, each channel still contains noise as, within each channel, the phases are still random. Thus, there are no monaural spectral cues for pitch available. The HP stimuli had a spectrum level of 35 dB (re 20 μ Pa). The NBN(s) that were added diotically to the HP stimuli (or to the diotic low-pass noise for the NBN-NBN condition) extended from 3% below to 3% above the center frequency of the chosen component(s) and had a spectrum level of 41 dB (re 20 μ Pa). This relatively low spectrum level of the NBN was chosen following some informal listening as to give a pitch percept that was approximately matched in salience, i.e., that seemed equally loud, to that of the HP components.

During training, the complex tones had the same F_0 s as in the main experiment (400, 267, and 160 Hz) but they contained all harmonics up to 2.4 kHz. The harmonics were sinusoids added in random phase and presented in Gaussian noise that was low-pass filtered at 4 kHz. The root mean square (rms) level of the complex tone was 3 dB below that of the noise.

The stimulus duration was 500 ms, including 40-ms raised-cosine onset and offset ramps. The silent interval between the two intervals within a trial was 500 ms. The stimuli were played out using a 16-bit digital-to-analog converter (CED 1401 plus), with a sampling rate of 40 kHz. Stimuli were passed through an antialiasing filter (Kemo 21C30) with a cutoff frequency of 15 kHz (slope of 96 dB/oct) and presented via the two channels of Sennheiser HD650 headphones at an rms level of about 68 dB SPL.

Experimental procedure

Listeners indicated which of the two stimuli had the higher pitch, without receiving feedback. Each F_0 complex was compared with each of the others. Within a given trial, both stimuli were from the same condition. Before each experimental block of 120 trials (ten trials for each condition and F_0 comparison in a randomized order across condition and F_0 comparison), listeners had a short training block of 30 trials that was intended to help them hear a residue pitch when only two harmonics were present. In the training block, feedback was provided. Overall, in the main experiment, 100 trials were collected in each condition for each F_0 comparison for each listener.

Initially, five subjects were tested without a short training block before each experimental block. At that point, it seemed difficult to find subjects who heard a residue pitch with only two harmonics present in the *single-mode* conditions; only two of the five were able to do so. As synthetic listening with only two harmonics was a prerequisite for the subjects in the present study, it was considered worthwhile to introduce a short training block that might help subjects to listen synthetically. The initial five subjects were retested using training blocks (subject numbers 1, 2, 11, 13, and 14 in Fig. 2), and the effect of the presence of the training blocks on synthetic vs analytic listening was evaluated on the basis of the data from these five subjects.

Subjects

Fifteen subjects (mean age=30 years; range, 20–48 years) with self-reported normal hearing were tested. One of them was the first author. Thirteen of them had some degree of musical training. Informed consent was obtained from all subjects. This study was carried out in accordance with the UK regulations governing biomedical research and was approved by the Cambridge Psychology Research Ethics Committee.

Analyses

The percentage of trials in which the listeners' judgments followed the F_0 rather than the spectral

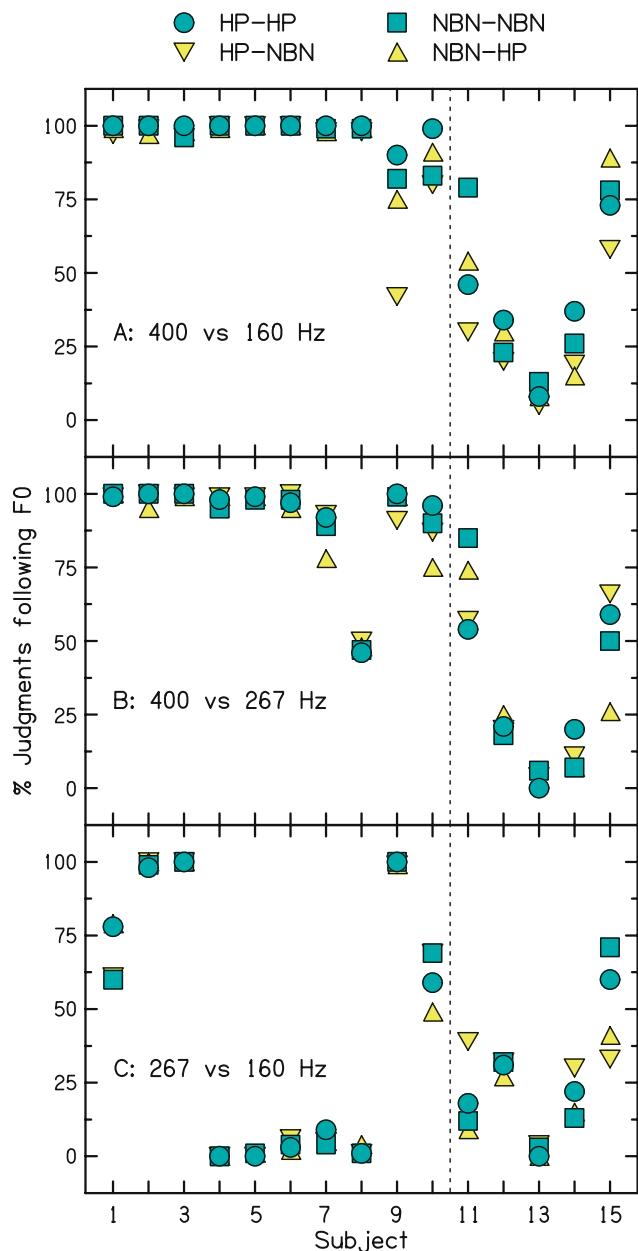


FIG. 2. Percentage of judgments where residue pitch was perceived. **A** Percentage of judgments following the F0 shown for each of the 15 subjects. The F0s of the stimuli to be compared were 400 and 160 Hz. *Turquoise symbols* represent percentages in the two single-mode conditions (*circles*: HP-HP; *squares*: NBN-NBN). *Yellow symbols* represent percentages in the two mixed-mode conditions (*downward pointing triangles*: HP-NBN, the low component is a HP and the high component is a NBN; *upward pointing triangles*: NBN-HP, the low component is a NBN and the high component is a HP). Each data point is based on 100 trials. Sixty independent data points are plotted. **B** as **A**, but for comparison of 400-Hz and 267-Hz F0 stimuli. **C** as **A**, but for comparison of 267-Hz and 160-Hz F0 stimuli. **A–C** The *dashed line* indicates the separation between those listeners who, in the HP-HP or NBN-NBN conditions, reliably heard a residue pitch for at least one of the three F0 comparisons and those who did not.

pitch of the individual components was calculated. These percentages (based on 100 trials) are shown in Figure 2, for each condition and F0 comparison, for all subjects. In the *single-mode* conditions, ten out of the 15 subjects (subject numbers 1–10) consistently heard a residue pitch with only two harmonics present for at least some of the F0 comparisons, i.e., their pitch judgments followed the F0. The remaining subjects either listened mainly analytically for all F0 comparisons in the single-mode conditions (subject 13) or were inconsistent. The data from these five subjects were excluded from further analysis (indicated by the dashed line in Fig. 2), as they would not allow any meaningful conclusions about the ability to hear a residue pitch in the mixed-mode conditions; one cannot expect subjects to perceive a residue pitch in the mixed-mode conditions if they do not perceive a residue pitch in the single-mode conditions. Next, percentages were averaged across the two single-mode conditions and across the two mixed-mode conditions, respectively. All statistical analyses were conducted after applying the rationalized arcsine units (RAU) transformation (Studebaker 1985) on these percentages. Pearson's correlation coefficient was calculated between the mixed-mode RAU-transformed percentages of the time that responses went with F0 and the corresponding single-mode RAU-transformed percentages. *t* tests were conducted, separately on the data for each F0 comparison, to assess the significance of the difference between the RAU-transformed percentages of judgments following the F0 in the single-mode and in the mixed-mode conditions. The best fitting linear regression line was calculated for RAU-transformed percentages of judgments following the F0 in the mixed-mode and in the single-mode conditions, taking into account that both the single-mode and the mixed-mode data include measurement errors. This required minimizing the perpendicular distances of the data points from the fitted straight line, i.e., minimizing the horizontal and vertical distances simultaneously. Note that usually, i.e., in the standard way of doing linear regression, the deviations between data points and fitted straight line are minimized in one dimension only. For example, only the deviations between the measured values of the dependent *y* variable and the predicted *y* values (the *y* values on the regression line for given values of the independent variable *x*) are minimized; the independent variable *x* is assumed to be error-free.

Training

The effect of the training was tested separately for a subgroup of five listeners. The statistical analysis showed that the training did not significantly increase the tendency to perceive the residue pitch rather than the pitches of individual components. Two *t* tests, one

for the mixed-mode conditions and one for the single-mode conditions (with RAU values averaged across the three F0 comparisons), showed that there were no significant differences between the percentages of judgments following the residue pitch before and after training [single-mode, $t(4) = -1.31$, $p=0.26$; mixed-mode, $t(4) = -1.28$, $p=0.27$]. Thus, the tendency to listen analytically to a given set of two-tone harmonic complexes seems to be a stable aspect of perception that is not overcome by listening to multi-harmonic complexes with the same F0s for which the residue pitch is clearly perceived.

RESULTS

The main interest of the study was to compare the proportion of trials on which listeners heard a residue pitch in the mixed-mode conditions compared to the single-mode conditions. If there are two separate mechanisms for the derivation of residue pitches from conventional spectral components and from binaurally created components, then listeners should not hear a residue pitch in the mixed-mode conditions, as each component would feed into a different mechanism. If, on the other hand, there is a single mechanism for deriving residue pitch from conventional and dichotic components, then listeners should be equally likely to hear a residue pitch in the mixed-mode and the single-mode conditions because, even in the mixed-mode conditions, the two components would feed into the same process.

Note that the predicted pattern of results for the case of a single mechanism does rely on the absence of salient perceptual differences between the tonal percept that is evoked by the conventional spectral component and the tonal percept that is evoked by the HP. This proviso is needed because salient differences between the two tonal percepts might lead to perceptual segregation (Bregman 1990; Darwin and Carlyon 1995; Moore and Gockel 2002), and thus, promote analytical listening to the individual components rather than synthetic listening to the residue pitch, even if there is only one mechanism for deriving residue pitch. Subjectively, the salience of the tonal percept evoked by the NBN was well matched to that of the HP. Indeed the level of the NBN necessary to achieve this (7 dB higher at the signal frequency than the level of the HP noise), as determined during informal listening, was very similar to the relative levels determined by Plack et al. (2011) in order to produce equal frequency discrimination thresholds for a NBN and a HP component centered at 300 or 600 Hz.

Before we can compare pitch perception in the single-mode and mixed-mode conditions, we must

ensure that judgments for the single-mode stimulation provide clear evidence for the perception of residue pitch. A preliminary analysis of the results showed that, in the single-mode conditions, the tendency to hear a residue pitch varied substantially between subjects and also between F0 comparisons (see Fig. 2). The latter will be discussed below. The former was expected, as it is well-known that when only a small number of harmonics are present, as in our stimuli, a significant proportion of listeners will not perceive a residue pitch for a conventional harmonic complex but rather will listen analytically to individual spectral components (Smooenburg 1970; Schneider et al. 2005a; Schneider et al. 2005b). Two thirds of the subjects tested consistently perceived a residue pitch for at least some of the F0 comparisons in the single-mode conditions, i.e., their pitch judgment consistently followed the F0 (see Fig. 2, subjects 1–10). The data of these ten subjects (towards the left of the dashed line in Fig. 2) were analyzed further and are discussed below.

The important comparison is between the responses of these subjects in the single-mode and the mixed-mode conditions. This is shown in Figure 3, in which the percentage of judgments following the F0, averaged across the two mixed-mode conditions, is plotted as a function of the percentage of judgments following the F0, averaged across the two single-mode conditions. If it does not matter that one harmonic in the complex requires binaural processing, while the other one does not, then all data points should lie around the diagonal. Broadly, this is the data pattern observed.

Listeners' responses in the single-mode and mixed-mode conditions were highly correlated; Pearson's correlation coefficient calculated across all data points was 0.99 ($p<0.001$). They were also highly correlated for each of the three F0 comparisons separately (Pearson's correlation coefficient was 0.95, 0.93, and 1.00 for the 400-Hz vs 160-Hz, 400-Hz vs 267-Hz, and the 267-Hz vs 160-Hz comparison, respectively, with $p<0.001$ for all comparisons). The best fitting straight line calculated for all the data had a slope very close to one (0.97 RAU) and an intercept (value of y when x is zero) very close to zero (-1.38 RAU). These values indicate near identity between the responses in the single-mode and the mixed-mode conditions. In addition, three separate t tests, one for each F0 comparison, showed no significant differences between the averaged percentages of judgments following F0 in the single-mode and in the mixed-mode conditions [400-Hz vs 160-Hz comparison, $t(9) = 2.84$; 400-Hz vs 267-Hz comparison, $t(9) = 1.76$; 267-Hz vs 160-Hz comparison, $t(9) = -0.10$; for all comparisons $p>0.05$, Bonferroni corrected]. Thus, the results clearly show that (1) listeners are able to hear a

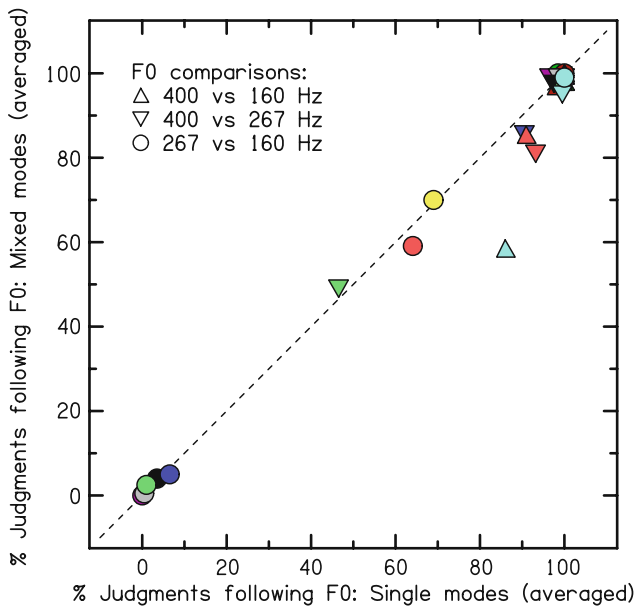


FIG. 3. Perception of residue pitch in mixed-mode as a function of that in single-mode conditions. Percentage of judgments following the F0 when averaged across the two mixed-mode conditions (one HP and one narrowband noise component) as a function of the percentage of judgments following the F0 when averaged across the two single-mode conditions (both components either HPs or narrowband noises) for ten subjects (represented by different colors). Each data point is based on 200 trials. Thirty independent data points are plotted; the cluster of points at the *top right*, indicating cases where the listeners perceived the residue pitch in both the mixed-mode and the single-mode conditions in about 100% of the trials, accounts for 18 of these.

residue pitch when one component is a conventional spectral component and the other is a dichotic component and (2) the ability to perceive a residue pitch is not significantly affected by whether or not the component pitches have a different origin (whether or not they require binaural processing). Therefore, the results clearly indicate the existence of a single mechanism for deriving residue pitch from conventional spectral and binaurally created components.

Curiously, for the 267-Hz vs 160-Hz comparison, five of the subjects seemed consistently to listen analytically (see Fig. 2C, subjects 4–8; Fig. 3, circles in lower left corner), while they listened synthetically for the other F0 comparisons. Verbal reports from two subjects within this group, who, due to their musical training were able to identify which musical intervals they heard in the experiment, indicated that these data were not due to analytic listening but rather to an octave confusion. For the 267-Hz vs 160-Hz comparison, they heard a musical interval corresponding roughly to a minor third. Thus, these subjects perceived the pitch of the complex with F0=160 Hz as that of a complex with F0=320 Hz, and this led them to judge the complex with F0=160 Hz as

being higher than the complex with F0=267 Hz. Overall, it seems that, even for this F0 comparison, subjects listened synthetically. One reason for the octave confusion may be that having heard the other pitches (267 and 400 Hz), the 320-Hz F0 represented a smaller change in chroma (the note of the sound) than did the (true) 160-Hz F0. If so, then our results suggest that, when judging the direction of a pitch change between consecutive tones, listeners often choose the direction that requires the smallest change of chroma between the two sounds, while ignoring the “pitch height,” i.e., which “octave” each sound is in.

DISCUSSION

Most current theories about the processing leading to the perception of a dichotic pitch are based on the assumption that there is an internal “central spectrum” that has a peak at the center frequency of the narrow band that is interaurally decorrelated. It is still a matter of debate how this central spectrum is generated (Raatgever and Bilsen 1986; Culling et al. 1998; Hartmann and Zhang 2003). The dichotic pitch is supposed to be determined by the central spectrum. In the case of a complex pitch (with multiple decorrelated regions), the pitch has been assumed to be determined via a central pattern recognition process (Goldstein 1973; Terhardt 1974) similar to those that can be applied when a sound is presented either to one ear or identically to both ears (Raatgever and Bilsen 1986).

The present results demonstrate a common mechanism for deriving residue pitch from spectral components and dichotic pitch components like the HP, which itself depends on binaural processing. Listeners were just as likely to hear a residue pitch in the mixed-mode conditions as in the single-mode conditions.

These results, taken together with physiological reports, indicate that this common mechanism for deriving the residue pitch is most likely located at or beyond the dorsal nucleus of the lateral lemniscus (DNLL) or the IC in the midbrain. Specifically, there is physiological evidence that binaural processing of temporal information that is necessary for the extraction of a dichotic pitch component occurs in the MSO (Palmer and Shackleton 2002). As the output from MSO cells does not project to other parts of the MSO, but only to higher nuclei, i.e., the DNLL in the brainstem and the IC in the midbrain (Schofield 2005), these are the earliest levels where information on binaural components could be combined, thus allowing the derivation of a residue pitch from dichotic components. Our finding that the pitch mechanism treats dichotic and conventional spectral components alike means that this conclusion is not specific to dichotic pitches such as HP but generalizes

to the wide range of sounds we hear in everyday life. It complements the finding that a representation of residue pitch can be observed in auditory cortex (Bendor and Wang 2005; Hall and Plack 2009) by imposing a constraint on the *earliest* stage of processing at which residue pitch is extracted.

The present results have specific implications for models that aim to account for pitch perception using simulations of the response of the peripheral auditory system to auditory stimuli. For example, one popular model simulates the response of the basilar membrane, hair cells, and auditory nerve (AN), and then performs an autocorrelation on the simulated responses of AN fibers (Meddis and Hewitt 1991a, b). Previous findings that listeners can derive a residue pitch from spectral components presented to opposite ears (Houtsma and Goldstein 1972) can be easily accounted for by this class of model by simply assuming that the autocorrelation process receives inputs from AN fibers innervating both ears. This process could be located as early as in the cochlear nucleus (CN; Palmer and Winter 1992), which receives input from both ears (Ingham et al. 2006). However, to the best of our knowledge, there exists no evidence that the CN supports binaural processing of fine temporal information. Substantial modifications of this class of model would be needed to account for the fact that spectral components are readily combined with “components” that do not exist monaurally and can only arise as the result of binaural interactions of temporal information. An outline of such a model has been described by Akeroyd and Summerfield (1999). Another model, which is one of the few to specify the neural processes that derive residue pitch, was proposed by Wiegrebe and colleagues (Wiegrebe and Winter 2001; Wiegrebe and Meddis 2004). They proposed that the initial computation of pitch takes place in “sustained chopper” cells of the CN. However, the lack of interaural temporal processing by the CN means that this model cannot account for the combination of spectral and HP components.

There is evidence that, in speech perception, listeners can combine two formants to determine vowel identity, when one is defined by an increase in intensity and the other is defined by an interaural decorrelation within a narrow frequency band centered on the formant frequency (Akeroyd and Summerfield 2000). While this result might be expected for a speech perception task dependent on identifying high-level units stored in learned cortical representations, the same was not necessarily true for pitch. Gockel et al. (2009) showed that not only can listeners derive a residue pitch from several HP components in the absence of the fundamental, but also that F0 discrimination for a target HP complex can be impaired when another spectral complex tone (the interferer) with F0

close to the mean F0 of the target is presented simultaneously. The amount of interference from a spectral complex interferer was about the same when the target complex was a complex HP as when it was a loudness-matched spectral complex. This interference could arise because all simultaneous components (whether spectral or binaurally created) feed into one mechanism for deriving residue pitch. However, as also discussed by Gockel et al. (2009), there existed the possibility that spectral components and binaurally created components fed into two different pitch mechanisms and that F0 discrimination for the target complex HP was impaired because, at a later stage, either the pitches might not be independently accessed or they might be transformed into a common code. The current study addressed this issue more directly and provided evidence that spectral and HP components are combined into a residue via a single mechanism. It did not, however, address the issue of whether either the pitches of the components or the residue pitch are determined by place or temporal mechanisms or both (Carlyon and Shackleton 1994; Gockel et al. 2004).

To summarize, the mixed-mode stimuli were equally likely to lead to the perception of a residue pitch as the single-mode stimuli, indicating that the mechanism that derives residue pitch does not process components of different origin in a different way (whether or not they require binaural processing). This shows that there exists a single mechanism for the derivation of residue pitch from binaural components and from spectral components. This mechanism is located at or beyond the DNLL or the IC in the midbrain. The current findings may inform future research into the physiology of the perception of residue pitch.

ACKNOWLEDGMENTS

Supported by the Wellcome Trust Grant 088263. We thank Prof. B.C.J. Moore, Dr. M. Davis, three anonymous reviewers, and the editor Barbara Shinn-Cunningham for comments on a previous version of this paper and Dr. B. Glasberg for providing us with a program for best line fits for data with measurement variability in both dimensions.

OPEN ACCESS

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

REFERENCES

- AKERROYD MA, SUMMERFIELD AQ (1999) A fully-temporal account of the perception of dichotic pitches. *Br J Audiol* 33:106–107
- AKERROYD MA, SUMMERFIELD AQ (2000) Integration of monaural and binaural evidence of vowel formants. *J Acoust Soc Am* 107:3394–3406
- AKERROYD MA, MOORE BCJ, MOORE GA (2001) Melody recognition using three types of dichotic-pitch stimulus. *J Acoust Soc Am* 110:1498–1504
- BENDOR D, WANG X (2005) The neuronal representation of pitch in primate auditory cortex. *Nature (L)* 436:1161–1165
- BILSEN FA (1977) Pitch of noise signals: evidence for a “central spectrum”. *J Acoust Soc Am* 61:150–161
- BREGMAN AS (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound*. Bradford books. MIT, Cambridge
- CARLYON RP, SHACKLETON TM (1994) Comparing the fundamental frequencies of resolved and unresolved harmonics: evidence for two pitch mechanisms? *J Acoust Soc Am* 95:3541–3554
- CRAMER EM, HUGGINS WH (1958) Creation of pitch through binaural interaction. *J Acoust Soc Am* 30:413–417
- CULLING JF, SUMMERFIELD AQ, MARSHALL DH (1998) Dichotic pitches as illusions of binaural unmasking. I. Huggins’ pitch and the “binaural edge pitch”. *J Acoust Soc Am* 103:3509–3526
- DARWIN CJ, CARLYON RP (1995) Auditory grouping. In: Moore BCJ (ed) *Hearing*. Academic, San Diego
- GOCKEL H, CARLYON RP, PLACK CJ (2004) Across-frequency interference effects in fundamental frequency discrimination: questioning evidence for two pitch mechanisms. *J Acoust Soc Am* 116:1092–1104
- GOCKEL HE, CARLYON RP, PLACK CJ (2009) Pitch discrimination interference between binaural and monaural or diotic pitches. *J Acoust Soc Am* 126:281–290
- GOLDSTEIN JL (1973) An optimum processor theory for the central formation of the pitch of complex tones. *J Acoust Soc Am* 54:1496–1516
- HALL DA, PLACK CJ (2009) Pitch processing sites in the human auditory brain. *Cereb Cortex* 19:576–585
- HARTMANN WM, ZHANG PX (2003) Binaural models and the strength of dichotic pitches. *J Acoust Soc Am* 114:3317–3326
- HOUTSMA AJM, GOLDSTEIN JL (1972) The central origin of the pitch of complex tones: evidence from musical interval recognition. *J Acoust Soc Am* 51:520–529
- INGHAM NJ, BLEECK S, WINTER IM (2006) Contralateral inhibitory and excitatory frequency response maps in the mammalian cochlear nucleus. *Eur J Neurosci* 24:2515–2529
- MEDDIS R, HEWITT M (1991a) Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery. I: Pitch identification. *J Acoust Soc Am* 89:2866–2882
- MEDDIS R, HEWITT M (1991b) Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery. II: phase sensitivity. *J Acoust Soc Am* 89:2883–2894
- MONTGOMERY CR, CLARKSON MG (1997) Infants’ pitch perception: masking by low- and high-frequency noises. *J Acoust Soc Am* 102:3665–3672
- MOORE BCJ, GOCKEL H (2002) Factors influencing sequential stream segregation. *Acta Acust United Ac* 88:320–333
- PALMER AR, WINTER IM (1992) Cochlear nerve and cochlear nucleus responses to the fundamental frequency of voiced speech sounds and harmonic complex tones. In: Cazals Y, Demany L, Horner K (eds) *Auditory physiology and perception*. Pergamon, Oxford
- PALMER AR, SHACKLETON TM (2002) The physiological basis of the binaural masking level difference. *Acta Acust United Ac* 88: 312–319
- PLACK CJ, TURGEON M, LANCASTER S, CARLYON RP, GOCKEL HE (2011) Frequency discrimination duration effects for Huggins pitch and narrowband noise. *J Acoust Soc Am* (in press)
- RAATGEVER J, BILSEN FA (1986) A central spectrum theory of binaural processing: evidence from dichotic pitch. *J Acoust Soc Am* 80:429–441
- SCHNEIDER P, SLUMING V, ROBERTS N, BLEECK S, RUPP A (2005a) Structural, functional, and perceptual differences in Heschl’s gyrus and musical instrument preference. *Ann NY Acad Sci* 1060:387–394
- SCHNEIDER P, SLUMING V, ROBERTS N, SCHERG M, GOEBEL R, SPECHT HJ, DOSCH HG, BLEECK S, STIPPICH C, RUPP A (2005b) Structural and functional asymmetry of lateral Heschl’s gyrus reflects pitch perception preference. *Nat Neurosci* 8:1241–1247
- SCHOFIELD BR (2005) Superior olivary complex and lateral lemniscal connections of the auditory midbrain. In: Winer JA, Schreiner CE (eds) *The inferior colliculus*. Springer, New York
- SMOORENBURG GF (1970) Pitch perception of two-frequency stimuli. *J Acoust Soc Am* 48:924–942
- STUDEBAKER G (1985) A “rationalized” arcsine transform. *J Speech Hear Res* 28:455–462
- TERHARDT E (1974) Pitch, consonance, and harmony. *J Acoust Soc Am* 55:1061–1069
- WIEGREBE L, WINTER IM (2001) Temporal representation of iterated rippled noise as a function of delay and sound level in the ventral cochlear nucleus. *J Neurophysiol* 85:1206–1219
- WIEGREBE L, MEDDIS R (2004) The representation of periodic sounds in simulated sustained chopper units of the ventral cochlear nucleus. *J Acoust Soc Am* 115:1207–1218
- ZHANG PX, HARTMANN WM (2008) Lateralization of Huggins pitch. *J Acoust Soc Am* 124:3873–3887