Published in final edited form as:

Hear Res. 2010 December 1; 270(0): 143–150. doi:10.1016/j.heares.2010.08.002.

Harmonic pitch: Dependence on resolved partials, spectral edges, and combination tones

Huanping Dai

Department of Speech, Language, and Hearing Sciences, University of Arizona, Tucson, AZ 85721

Abstract

Perceptual weights were estimated in a pitch-comparison experiment to assess the relative influences of individual partial tones on listeners' pitch judgments. The stimuli were harmonic sounds ($F_0 = 200$ Hz) with partials up to the 12th. Low-numbered partials were removed step-by-step, so that the remaining higher-numbered partials would have a better chance of showing any effect. The individual frequencies of the partials were perturbed randomly on each stimulus presentation, and weights were estimated as the correlation coefficients between the frequency perturbations and the listeners' responses. When the harmonic sounds contained all twelve partials, the listeners depended mostly on the low-numbered, resolved partials within the well-established dominance region. As the low-numbered partials were taken out of the dominance region, the listeners began to focus on only the lowest and highest partials at the spectral edges. For one listener, such an edge-listening strategy took the form of relying on nonlinear combination tones. Overall, there was no indication of any influence on pitch from unresolved partials, thus no evidence of contribution to pitch from temporal cues carried by this group of partials. The estimated patterns of weights were well described by the predictions of Goldstein's optimal-processor model. The predicted weights were inversely proportional to the amount of error for estimating the individual frequencies of the partials. The agreement between the predicted and measured weights suggests that, for harmonic sounds, partials whose frequencies are perceived with the best precision will likely have the greatest influence on perceived pitch.

Keywords

harmonic pitch; resolved and unresolved partials; place and temporal cues; perceptual weights; correlation analysis

1. Introduction

Pitch perception plays a crucial role in primary auditory functions including sound source identification, auditory scene analysis, and speech communication (Hartmann, 1988, 1996; Lyon and Shamma, 1997; Plack and Oxenham, 2005). For harmonic sounds, the pitch usually corresponds to the fundamental frequency (F_0). However, just how the ear retrieves

Correspondence: Huanping Dai, Department of Speech, Language, and Hearing Sciences, University of Arizona, 1131 E. 2nd Street, Tucson, Arizona, 85721, Tel: (520) 626-5595, Fax: (520) 621-9901, hdai@email.arizona.edu.

²Moore et al (1984) have shown that, under certain assumptions, Terhardt (1974, 1977's model becomes mathematically identical to Goldstein's model.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

the $F0$ information from the sound still remains unclear. In theory, this information is carried by *every* harmonic partial tone, but different partials convey the information to the central nervous system in different ways. For example, low-numbered partials that are readily resolved into separate auditory filters are most useful in forming a unique pattern of neural excitation with regularly spaced and clearly defined peaks along the basilar membrane. With these partials, the $F0$ information can be extracted via pattern matching schemes, as proposed in several models or algorithms (e.g., Goldstein, 1973; Terhardt, 1972, 1974; Wightman, 1973). In contrast, unresolved, higher-numbered partials will interact with each other within a single auditory filter and produce a periodic output. For these higher partials, the $F0$ information is expressed as the periodicity in the temporal envelope and the temporal fine structure (TFS) of the waveform at the output of the auditory filter. The $F0$ information extracted this way serves as the basis for models of residue pitch (e.g., Schouten, 1940; de Boer, 1956, 1976; de Cheveigne, 2005; Oxenham et al., 2009). Since different pitch theories rely on different partials in extracting the $F0$ information, unique insight about pitch theories can be gained from the dependence of the perceived pitch on individual partials. Thus, determining the relative importance of individual partials is valuable for understanding pitch perception.

Quantitatively, the relative importance of individual partials can be expressed as their perceptual weights. The key assumption is that the harmonic pitch (P) is a linear weighted sum of the estimates of $F0$ from the frequencies of all individual partials,

$$P = \sum_{n=1}^N w_n \hat{F0}_n \quad (1)$$

in which $\hat{F0}_n = \hat{f}_n/n$ represents the estimate of $F0$ based on the estimated frequency of the n^{th} partial ($\hat{f}_n$), and w_n the corresponding weight. The relative importance of each partial is then represented by the magnitude of the weight. Note that individual frequencies can be estimated via either place or temporal mechanisms of the auditory system. Likewise, contemporary views and theories of pitch typically depend on both spectral and temporal information (e.g., Srulovicz and Goldstein, 1983; Houtsma and Smurzynski, 1990; Meddis and Hewitt, 1991a,b; Moore, 1993; Shackleton and Carlyon, 1994; Cohen, Grossberg, and Wyse, 1995; Lyon and Shamma, 1995; Chatterjee and Zwislocki, 1997; Meddis and O'Mard, 1997; Shamma and Klein, 2000). The spectral information is mostly relevant to low-numbered partials; the temporal information, however, is relevant to both low and high numbered partials. The primary interest of this paper is concerned with the different roles of low versus high numbered, or resolved versus unresolved, partials.

Spectral weights for harmonic pitch have been estimated with two approaches. In the first, Moore et al. (1985) estimated the weights by measuring the amount of changes in harmonic pitch caused by mistuning the partials by a certain amount. Let ΔP_n be the change in pitch caused by mistuning the n^{th} partial by Δf_n . Then the weight associated with that partial is derived as $w_n = n(\Delta P_n / \Delta f_n)$. Since the change in pitch in this expression is specific to a given partial, it has to be estimated in separate trials for different partials. As a result, the weights reported in Moore et al. for different partials are obtained separately. Their results show that only the first four or five partials receive large weights. An alternative to the method used by Moore et al. is a correlational approach, in which random, independent perturbations are applied to the frequencies of all partials simultaneously. Perceptual weights are then represented by the correlation between the perturbations and the listeners' responses.^[1] With this approach, weights for *all* partials are estimated under identical conditions.

Using the correlation technique, weights have been estimated in a pitch-comparison task (Dai, 2000). The weighting patterns show that three to four partials closest to 600 Hz, independent of F_0 , dominate the pitch judgment. In particular, the weighting pattern obtained for $F_0 = 200$ Hz is in excellent agreement with that predicted by Goldstein (1973)'s optimal processor model. These weighting patterns are somewhat more sharply tuned, indicating a narrower dominance region, than those obtained by Moore et al. (1985). Furthermore, these results suggest that dominant partials are best described in terms of fixed absolute frequencies (e.g., 600 Hz), rather than fixed harmonic numbers as indicated by the results from Moore et al. (1985). These discrepancies may be resulted from the differences in experimental procedure, or simply from individual differences of the listeners. Qualitatively, the weighting patterns obtained from both studies are consistent with many other results (e.g., de Boer, 1976; Flanagan and Guttman, 1960; Plomp, 1967; Ritsma, 1967, 1970; Patterson and Wightman, 1976; Yost, 1982; Houtsma and Smurzynski, 1990; Gockel, Carlyon, Plack, 2005, 2009; Qin and Oxenham, 2006; Micheyl and Oxenham, 2007), showing a clear dominance of the low-numbered, resolved partials over the high-numbered, unresolved partials in determining the harmonic pitch.

These results lead to the following questions: If the low-numbered partials are systematically removed, how might the weights among the remaining partials be redistributed? In other words, how does harmonic pitch depend on the remaining partials when the dominant partials are absent? Specifically, will the listeners begin to rely on high-numbered, unresolved partials as low-numbered, well-resolved partials are taken away? These questions are the focus of the present study. The extent to which harmonic pitch can and does depend on the high-numbered partials has been illusive. Because these partials are less well resolved by the auditory filters than the lower ones, the F_0 information carried by their individual frequencies is expected to be less reliable. Empirical evidence suggests that, when the F_0 information is mostly temporal (e.g., when the place information is severely degraded or expected to be absent), the perceived pitch is either very weak (e.g., Burns and Viemeister, 1979), or of a limited range (less than 300 Hz, Zeng, 2002), or absent (e.g., Oxenham et al., 2004). Despite the difficulties in demonstrating the effectiveness of the temporal cues carried by the unresolved partials, the possible roles of these partials in pitch perception continue to be of much interest, not only for theoretical but also for practical reasons, such as cochlear-implant applications (e.g., Zeng, 2002; Ciocca et al., 2002; Geurts and Wouters, 2004; Singh et al., 2009). Here we seek to gain some unique insight on the role of the high-numbered, less-well resolved partials by measuring perceptual weighting functions.

In the present study, the correlation approach was used to estimate weighting patterns for harmonic pitch. The results showed that, as the low-numbered partials were systematically removed, the partials at both spectral edges began to assume the dominant roles. However, there was no indication of influence by unresolved partials. This outcome is consistent with the predictions of the optimal-processor model of Goldstein (1973).

2. Assumption underlying the measurements

The key assumption underlying the current measurements, which comes from the optimal-processor model of Goldstein (1973), is that the harmonic pitch is derived entirely from the

¹For further details about the correlation technique, see Richards and Zhu (1994) and Lutfi (1995). Applications of the correlation technique to human perception have their origin in Ahumada and Lovell (1971). Variations of the correlation analysis include classification-image methods in vision research (e.g., Ahumada, 2002), and a conditional-on-a-single-stimulus technique in auditory research (Berg, 1989). Although correlation analysis has been applied mostly in the context of psychophysical tasks with two response alternatives, such as the yes-no, or two-interval, forced-choice task, to uncover a single perceptual template, it can also be applied to tasks with multiple response alternatives to uncover multiple templates (e.g., Dai and Micheyl, 2010).

frequencies of the individual partials. The amplitudes and phases of the partials are deemed irrelevant. Specifically, the frequency of each partial provides an estimate of F_0 , and the final estimated fundamental frequency, $\hat{F}_0$, can be expressed as a linear weighted sum of the individual estimates of F_0 . This principle is generalized to all partials. That the higher partials are less well unresolved only implies that estimates of their frequencies are subject to greater amount of error. Let $\eta = \langle \eta_1, \eta_2, \dots, \eta_{12} \rangle$ represent the frequencies of the partials normalized by the partial number: $\eta_n = f_n/n$. The pitch of the standard, P_0 , can be expressed as

$$P_0 = \sum_{n=K}^{12} w_n \eta_n + \xi_0 \quad (2-1)$$

in which K represents the lowest partial number, and ξ_0 represents the internal noise. The pitch of the perturbed sound, P , can be expressed as

$$P = \sum_{n=K}^{12} w_n (\eta_n + \Delta \eta_n) + \xi \quad (2-2)$$

in which $\Delta \eta_n$ represents the perturbation, and ξ represents the internal noise from the 2nd observation interval. The difference between the two pitches, $\Delta P = P - P_0$, is then

$$\Delta P = \sum_{n=K}^{12} w_n \Delta \eta_n + \xi - \xi_0 \quad (3)$$

Assume the listeners use the pitch difference ΔP as the decision variable, denoted by Z , in the pitch comparison task, i.e., $Z = \Delta P$. Let R denote the binary response variable. The decision rule is: *Responding 'interval one' ($R = 1$) if and only if $Z < 0$, and 'interval two' ($R = 2$) otherwise*. The weights can be estimated and represented by the correlation coefficient, r , computed between the binary responses, R , and the perturbations imposed on the partials over a total of M trials:

$$w_n = r(\Delta \eta_n, R) = \frac{\sum (\Delta \eta_n R) - M \overline{\Delta \eta_n} \overline{R}}{\sqrt{(\sum \Delta \eta_n^2 - M \overline{\Delta \eta_n}^2)(\sum R^2 - M \overline{R}^2)}} \quad (4)$$

3. Material and methods

The method was in many ways similar to that described in Dai (2000) to which the reader is referred for additional details.

3.1 Listeners

Three normal-hearing college students participated. Each was explained the purpose and potential risk of the study, and signed an IRB approved informed consent.

3.2. Stimuli

All stimuli were harmonic or pseudo-harmonic (with frequencies perturbed slightly from the nominal harmonics) sounds of 200 ms, including 10-ms cosine-squared rise/fall times. Each partial had a level of 60 dB SPL. The fundamental frequency was $F_0 = 200$ Hz. The stimuli

were presented binaurally via headphones to the listener sitting in a sound-attenuating room. In four different conditions, the stimuli included partials 1 to 12 (Condition 1, Fig. 1, top), 3 to 12 (Condition 2), 5 to 12 (Condition 3), and 7 to 12 (Condition 4).

3.3. Procedure

The listeners compared the pitches of a pair of sounds in a two-interval, forced-choice task, indicating which interval contained a higher pitch (Fig. 1, bottom). They were instructed to listen to the pitch of the stimulus as a whole (i.e., *synthetic* mode), rather than the pitches of individual partials (i.e., *analytic* mode). Other than the above instruction given to the listeners explicitly, there was nothing to prevent them from adopting any listening and decision strategies in performing the task.

Each trial began with a standard harmonic sound:

$$S_{Standard}(t, i) = \sum_{n=K}^{12} a \sin[2\pi n F_0 + \theta_{1,n}(i)], \quad (5)$$

in which a represents the amplitude, $\theta_{1,n}(i)$ indicates random phase of each component on each presentation, i indicates the trial number, and $K=1, 3, 5$, or 7 indicates the number of the lowest partial used for a particular condition. The second interval contained the pseudo-harmonic sound. Its form is similar to the standard, except that the frequencies of individual partials were randomly perturbed by an amount of $n\Delta\eta$

$$S_{Comparison}(t, i) = \sum_{n=K}^{12} a \sin\{2\pi n [F_0 + \Delta\eta_n(i)] + \theta_{2,n}(i)\} \quad (6)$$

The random perturbations $\Delta\eta_n(i)$ were drawn from a Gaussian distribution with zero mean and a sigma that was 2% of the fundamental frequency, or 4 Hz. A mistuning exceeding 8% no longer contributes to the overall pitch (Terhardt, 1972, cf. Moore et al., 1985). The probability that a perturbation exceeded 8% of the component frequency is negligible ($p \approx 0.00003$). Note that in the previous study the stimuli were perturbed in both intervals. Here the use of an unperturbed standard, which has the clearest harmonic pitch, was intended to help the listeners focus on the pitch of the complex sound as a whole (*synthetic* mode), rather than the pitch of individual components (i.e., *analytic* mode). Following practice sessions, at least $M=1000$ trials were tested for each listener in each condition.

3.4. Data analysis

After completion of each condition with each listener, a correlation coefficient was calculated between the 1000 responses, which consisted of 1s and 2s, and the perturbation (η_n) added to each component. For error bounds of estimated weights (e.g., Dai et al., 1996), correlation coefficients (r) may be transformed into Fisher's z values by $z = \tanh^{-1}(r)$, which is distributed approximately Gaussian, with a standard error equal to $1/\sqrt{M-3}$. Here, however, the main concern regarding the magnitude of the obtained correlation coefficients is whether they differ from zero with statistical significance. Under the null hypothesis of zero correlation, the distribution of the correlation coefficient itself is approximately Gaussian with zero mean and a standard error of $\sigma_e = 1/\sqrt{M}$ (Lutfi, 1995). Thus, at a confidence level of about 92%, a weight must differ from zero by more than $2\sigma_e$ for that difference to be considered statistically significant. Accordingly, only partials receiving weights that differ from zero by more than $2\sigma_e$ are considered to have an influence on the pitch judgment.

4. Results

The influence of individual partials on harmonic pitch was assessed and expressed as weights, which were estimated as the correlation coefficients between the frequency perturbations to the partials and the listeners' responses in the pitch-comparison tasks. The estimated weights are presented in Fig. 2.

The three listeners exhibited distinctively different patterns of weights. For S1, when all 12 partials were present, the pitch was mostly influenced by the highest (the 12th) and the two lowest partials (circles). As the lowest partials were removed, the weights maintained a similar pattern, but with the influence growing increasingly stronger for the highest partial while diminishing for the lower ones. For S2, on the other hand, the pitch was at first dominated by the three lowest partials, but soon became dependent on partials at either the lower (squares), higher (diamonds), or both (triangles) spectral edges, displaying the behavior of edge listening. Similar to S1, the influence of the highest partial grew stronger as partials continued to be removed from the low end. Furthermore, once the first six partials were gone (diamonds), both S1 and S2 abandoned the lower partials altogether, clinging instead to the highest partial alone. Finally, S3 exhibited the most unusual behavior. After relying on the first two or three partials in the first two conditions (circles and squares), the listener settled into a unique pattern of a strong positive weight followed by a negative weight with half the magnitude (triangles and diamonds). As discussed later, this weighting pattern presumably reflects the influence of a lower cubic combination tone on the pitch judgment. Unlike S1 or S2, S3 did not rely on the highest partial in any condition.

Despite the apparent individual differences, the results revealed some general trends. When all partials were available (condition 1, circles), the preferred strategy for sampling the F_0 information relied on a group of two to three adjacent partials. When such a strategy took place, it did so only within the first six partials, consistent with previous observations by Moore et al. (1985) and Dai (2000). However, this strategy did not survive the removal of the low-numbered partials; as soon as the first two partials were gone, the listeners began to depend on partials at the spectral edges. Relying on the lower cubic combination tone (for S3) can be regarded as a specific manifestation of edge listening. Finally, and above all, there was no indication that the unresolved partials had any influence on pitch judgment; the estimated weights for these partials stayed consistently around zero.

5. Discussions

A. Combination Tones

Negative weights, such as those observed for S3 (triangles and diamonds in Fig. 2), are anomalous against either previous experiences (e.g., Dai, 2000; Moore et al., 1985) or major pitch theories (e.g., Goldstein, 1973; Wightman, 1973; Terhardt, 1974). The following analysis leads to the premise that the unique weighting pattern observed for S3 implicates cubic-distortion tones.

Let f_K and f_{K+1} be the frequencies of the two lowest partials; the lower cubic distortion tone will then have a frequency of $f_{cubic} = 2f_K - f_{K+1}$. In the 3rd condition, for example, the frequencies are $f_K = 1000\text{Hz}$ and $f_{K+1} = 1200\text{Hz}$, thus $f_{cubic} = 800\text{Hz}$. Effectively, the cubic tone becomes the lower spectral edge, albeit at a very soft level. Suppose S3 depended on the cubic tone in the pitch-comparison task, then the decision variable Z would be proportional to f_{cubic} :

$$Z = cf_{cubic} = 2cf_K - cf_{K+1} \quad (7)$$

where c is a constant of positive value. Equation 7 implies that the weights associated with the two partials, w_K and w_{K+1} , will be $w_K = 2c$ and $w_{K+1} = -c$. The weighting patterns obtained from S3 in the 3rd and 4th conditions are consistent with these predictions both in sign and in relative magnitude, therefore supporting the hypothesis that the cubic distortion tone has influenced S3's pitch judgments.

The above hypothesis was formally tested in a replication of the 3rd condition with S3, but in the presence of a continuous masking noise. The noise was low-pass filtered at a cutoff frequency of 800 Hz, which coincided with the frequency of the cubic distortion tone. It had a spectrum level of 25 dB SPL, which was expected to reduce the audibility of the cubic tone significantly without affecting the primary partials of the complex sound. In the presence of the masker, if the listener no longer relies on the cubic tone, the original characteristic weighting pattern (i.e., $w_K = -2w_{K+1}$) is expected to be disrupted. Figure 3 shows the new weighting pattern of S3 in the presence of the masking noise (filled inverted triangles), along with the original result (unfilled inverted triangles from Fig. 2) for comparison. The negative weight at the 6th partial (1200 Hz) disappeared. The new weighting pattern resembles the stereotypical pattern of edge listening. Thus, both the analysis of the weighting pattern and the experimental outcome indicate that S3 had depended on the lower cubic combination tone for the pitch judgment.

It is unclear why S3 was compelled to depend on the cubic tone, given that the tone was much softer than the primary partials. One possible reason is that, even though it was at a much lower level than that of the partials, the cubic tone was better resolved and closer to the dominance region than any of the partials still remaining. However, this reason does not generalize—the other two listeners showed no sign of influence by combination tones. The current finding illustrates the multiplicity in ways that different listeners extract harmonic pitch.

B. Edge listening

Edge listening is reflected by a strong weight occurring at one of the spectral edges. A major finding of the present study is that edge listening plays a prominent role in pitch discrimination of complex sounds. Ten out of twelve estimated weighting functions, with the exception of S2 and S3 for Condition 1 (circles in Fig. 2), exhibited some form of edge listening. Edge listening is not identical to 'edge pitch', even though edge pitch involves edge listening. Likewise, when a listener's spectral weighting function reveals edge listening in a harmonic-pitch discrimination task, it does not necessarily imply that the listener has responded based on an edge pitch (analytic mode) rather than the harmonic pitch (synthetic mode). Rather, edge listening is one of the possible strategies that a listener can adopt for extracting F0 information. However, when the research is concerned with the role of the partials in the middle of the spectrum, a strong inclination of listeners to depend on spectral edges poses a problem. Unless the possibility of edge listening can be ruled out, conclusions about the roles of unresolved partials, or the effectiveness of the temporal cues resulted from the interactions of these partials, will be subject to uncertainty. For this reason, studies of unresolved partials in the perception of harmonic pitch often include steps to limit edge listening.

A popular procedure among pitch researchers for discouraging the listeners from adopting an edge-listening strategy, or relying on edge pitches, is to filter the spectrum with a relatively shallow slope at the edge (e.g., Carlyon and Shackleton, 1994; Micheyl and Oxenham, 2007). The extent to which spectral shaping is effective in suppressing edge listening remains uncertain. To gain some insight on the effect of spectral shaping on edge listening, the measurement of weighting patterns were repeated for S1, but with attenuations applied to the lowest partial (20 dB) and the second lowest partial (10 dB) for each

condition. Figure 4 shows the new weighting pattern (filled symbols), along with the original result (unfilled symbols, replotted from Fig. 2) for comparison.

When all the partials were present (circles in Fig. 4), the spectral shaping cut the weight for the fundamental component by half, but boosted the weights for the 2nd partial significantly, along with a boost of weights for the next two partials (indicated by changes from unfilled to filled circles). When the 3rd partial was the lowest (squares), the spectral shaping cut the weights for the two lowest partials by half, but elevated the weight for the next partial dramatically, from near zero to the highest magnitude. In the next two conditions (triangles and diamonds), no edge listening was observed at the lowest partial in the presence of spectral shaping. Finally, shaping of the lower spectral edge altered the weight for the highest partial in two conditions with the opposite effect, reducing the weight for the first condition (circles) and increasing the weight for the fourth condition (diamonds). Overall, these results revealed two main trends. First, although the spectral shaping did reduce the weight for the lowest partial, even a 20 dB attenuation did not suppress its influence completely. This finding is consistent with the results of S3 showing an influence of a combination tone. Listening to cubic combination tones represents a particular manifestation of edge listening. Second, attenuating the lowest partial(s) tended to enhance the influence of the next partial(s). If this trend continues when a greater amount of attenuation is applied to the current edge to make it inaudible, the next partial that remains audible may be dominant. Since that partial now becomes the new edge, edge listening effectively remains. As a result, edge listening could continue to complicate interpretations of possible roles of unresolved partials.

C. Weighting-pattern predictions by pitch models

Estimated weighting patterns provide a unique way of evaluating pitch models. Because pitch models are usually developed to produce the correct pitch value (i.e., $F0$), in many cases the perceived pitch itself does not provide an effective test of the models. In contrast, different pitch theories predict markedly different weighting patterns (Dai, 2000). For example, the weighting patterns predicted by Goldstein's model fit the estimated patterns more closely than those predicted by the models of Wightman (1973) and Terhardt (1974). The following analysis is focused on the predictions by Goldstein's model. His model is the most straightforward in predicting the weighting patterns estimated from the experimental conditions of the present study. His model is also the most appealing because it makes optimal use of the statistical information available in the input to the pitch processor.

According to Goldstein (1973)'s model, the harmonic pitch corresponds to the maximum-likelihood estimate of $F0$ based on the frequencies of the individual partials. Let $\hat{F0}_n$ be the estimate of $F0$ based on the n^{th} partial. Then the predicted harmonic pitch (P^T) can be expressed as a weighted average of the individual estimates ($\hat{F0}_n$):

$$P^T = \sum_{n=K}^{12} w_n^T \hat{F0}_n \quad (8)$$

in which w_n^T represents the predicted weight, and K the number of the lowest partial.

Suppose the estimate $\hat{F0}_n$ is subject to an error with a variance of σ_n^2 . The predicted weight is

then $w_n^T = S / \sigma_n^2$ in which $S = \left[\sum_{n=K}^{12} 1 / \sigma_n^2 \right]^{-1}$ is a scaling constant that will scale the sum of the weights to unity (see, e.g., the Appendix of Dai, 2000). Therefore, a partial that produces an estimate of $F0$ with the best precision, or equivalently the least amount of error, will have the strongest influence on the harmonic pitch. To predict the weighting pattern requires

knowing only the variance as a function of frequency, which is referred to hereafter as the variance function: $V(f) = \sigma^2(f)$

A direct estimate of the variance function is still unavailable. One approximation of $\sigma(f)$ comes from the frequency difference limens (DLs) measured within the harmonic sounds, by Moore et al. (1984). The weighting pattern predicted based on their results was an excellent match to that observed by Dai (2000). For the current conditions, predictions of weighting patterns will again be derived based on the frequency DLs measured by Moore et al. Their data are available for stimuli used in only the first and third conditions of the present study. The predicted and observed weighting patterns are compared in Figs 5 and 6. For all cases, weights are scaled to sum to unity.

For the first condition (Fig. 5), the predicted weights (solid line) have a sharp peak at the 3rd partial (600 Hz), which provided an excellent fit to that reported in the previous study (Dai, 2000). In contrast, the mean weights from the present study did not display such a peak, but stayed flat for the first three partials (circles with dashed line). This discrepancy mostly reflects the weighting pattern observed for S1 who depended strongly on the fundamental component. Overall, however, the predicted weighting pattern does capture the general shape of the measured. In particular, both show diminished weights for partials 5 to 11. For the third condition (Fig. 6), the predicted (solid line) and the mean measured (circles with dashed line) weights display a very similar overall pattern, both showing prominent influence from partials at the spectral edges. Taken together, Goldstein's model provides a reasonably good description of the measured weight patterns of the current study. More importantly, it suggests a concrete mechanism for interpreting the weights by linking them to the precision with which the frequencies of the individual partials are perceived.

6. Conclusion

Perceptual weights estimated in pitch-comparison experiments with harmonic sounds revealed relative influences of individual partials on the pitch judgments. When the first twelve partials were present in a harmonic sound, the listeners depended mostly on the low-numbered, resolved partials. When the low-numbered partials were systematically removed, the listeners began to depend on the lowest and the highest of the remaining partials, or the spectral edges. For one listener, such an edge-listening strategy appeared in the form of relying on nonlinear combination tones. There was no indication of influence from unresolved partials, thus no evidence of contribution to pitch from temporal cues carried by these partials. The estimated patterns of weights were well described by the predictions of Goldstein's optimal-processor model, according to which the weights are inversely proportional to the error size (variances) for estimating the individual frequencies of the partials. The good agreement between the predicted and measured weights suggests that, when listening to harmonic sounds, our perceived pitch is likely to be influenced most by the partials whose frequencies are perceived most accurately.

Acknowledgments

I thank William Hartmann, Douglas Keefe, Christophe Micheyl, and Fan-Gang Zeng for helpful discussions, and Tom Christenson, William Hartmann, Christophe Micheyl, Brian Moore, and an anonymous reviewer for their comments and suggestions on previous versions of this paper.

Abbreviations

2AFC	two-alternative, forced-choice
DL	difference limen

F0 fundamental frequency
TFS temporal fine structure

References

- Ahumada AJ Jr. Classification image weights and internal noise level estimation. *J Vision*. 2002; 2:121–131.
- Ahumada AJJ, Lovell J. Stimulus features in signal detection. *J Acoust Soc Am*. 1971; 49:1751–1756.
- Berg BG. Analysis of weights in multiple observation tasks. *J Acoust Soc Am*. 1989; 86:1743–1746.
- Burns EM, Viemeister NF. Nonspectral pitch. *J Acoust Soc Am*. 1976; 60:863–869.
- de Boer, E. PhD thesis. Amsterdam: 1956. On the ‘Residue’ in Hearing.
- de Boer, E. On the ‘residue’ and auditory pitch perception. In: Keidel, WD.; Neff, WD., editors. *Handbook of Sensory Physiology*. Vol. V/3. Springer-Verlag; New York: 1976. p. 479-583.
- de Cheveigne, A. Pitch perception models. In: Plack, C.; Oxenham, A.; Fay, R.; Popper, A., editors. *Pitch: Neural Coding and Perception*. New York: Springer Verlag; 2005.
- Chatterjee M, Zwislocki JJ. Cochlear mechanisms of frequency and intensity coding. I. The place code for pitch. *Hear Res*. 1997; 111:65–75. [PubMed: 9307312]
- Ciocca V, Francis AL, Aisha R, Wong L. The perception of Cantonese lexical tones by early-deafened cochlear implantees. *J Acoust Soc Am*. 2002; 111:2250–2256. [PubMed: 12051445]
- Cohen MA, Grossberg S, Wyse LL. A spectral network model of pitch perception. *J Acoust Soc Am*. 1995; 98:862–879. [PubMed: 7642825]
- Dai H. On the relative influence of individual harmonics on pitch judgment. *J Acoust Soc Am*. 2000; 107:953–959. [PubMed: 10687704]
- Dai H, Micheyl C. Psychophysical Reverse Correlation With Multiple Response Alternatives. *J Exp Psych: Human Perception and Performance*. 2010 (in press).
- Flanagan JL, Guttman N. On the pitch of periodic pulses. *J Acoust Soc Am*. 1960; 32:1308–1319.
- Gockel H, Carlyon RP, Plack CJ. Dominance region for pitch: Effects of duration and dichotic presentation. *J Acoust Soc Am*. 2005; 117:1326–1336. [PubMed: 15807021]
- Gockel HE, Carlyon RP, Plack CJ. Reduced contribution of a nonsimultaneous mistuned harmonic to residue pitch: The role of harmonic number. *J Acoust Soc Am*. 2009; 125:15–18. [PubMed: 19173388]
- Goldstein JL. An optimum processor theory for the central formation of the pitch of complex tones. *J Acoust Soc Am*. 1973; 54:1496–1515. [PubMed: 4780803]
- Geurts L, Wouters J. Better place-coding of the fundamental frequency in cochlear implants. *J Acoust Soc Am*. 2004; 115:844–852. [PubMed: 15000196]
- Hartmann, WM. Pitch perception and the segregation and integration of auditory entities. In: Edelman, GM.; Gall, WE.; Cowan, WM., editors. *Auditory Functions*. Wiley; New York: 1988.
- Hartmann WM. Pitch, periodicity, and auditory organization. *J Acoust Soc Am*. 1996; 100:3491–3502. [PubMed: 8969472]
- Houtsma AJM, Smurzynski J. Pitch identification and discrimination for complex tones with many harmonics. *J Acoust Soc Am*. 1990; 87:304–310.
- Lyon, R.; Shamma, S. Computational Strategies for Pitch and Timbre. In: Hawkins, H.; McMullen, T.; Popper, A.; Fay, R., editors. *Auditory Computations*. New York: Springer Verlag; 1997.
- Lutfi RA. Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks. *J Acoust Soc Am*. 1995; 92:1333–1334.
- Meddis R, Hewitt MJ. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification. *J Acoust Soc Am*. 1991a; 89:2866–2882.
- Meddis R, Hewitt MJ. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II. Phase sensitivity. *J Acoust Soc Am*. 1991b; 89:2883–2894.
- Micheyl C, Oxenham AJ. Across-frequency pitch discrimination interference between complex tones containing resolved harmonics. *J Acoust Soc Am*. 2007; 121:1621–1631. [PubMed: 17407899]

- Moore, BCJ. Frequency analysis and pitch perception. In: Yost, WA.; Popper, AN.; Fay, RR., editors. *Human Psychophysics*. Springer-Verlag; New York: 1993. p. 56-115.
- Moore BCJ, Glasberg BR, Peters RW. Relative dominance of individual harmonics in determining the pitch of complex tones. *J Acoust Soc Am*. 1985; 77:1853–1860.
- Moore BCJ, Glasberg BR, Shailer MJ. Frequency and intensity difference limens for harmonics within complex tones. *J Acoust Soc Am*. 1984; 75:550–562. [PubMed: 6699293]
- Oxenham AJ, Bernstein JGW, Penagos H. Correct Tonotopic Representation Is Necessary for Complex Pitch Perception. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(5):1421–1425. [PubMed: 14718671]
- Oxenham AJ, Micheyl C, Keebler MV. Can temporal fine structure represent the fundamental frequency of unresolved harmonics? *J Acoust Soc Am*. 2009; 125:2189–2199. [PubMed: 19354395]
- Plack, C. *The sense of hearing*. Lawrence Erlbaum Associates Inc; 2005.
- Plack, C.; Oxenham, A. The psychophysics of pitch. In: Plack, C.; Oxenham, A.; Fay, R.; Popper, A., editors. *Pitch: Neural Coding and Perception*. New York: Springer Verlag; 2005.
- Patterson R, Wightman F. Residue pitch as a function of component spacing. *J Acoust Soc Am*. 1976; 59:1450–1459. [PubMed: 939878]
- Plomp R. Pitch of complex tones. *J Acoust Soc Am*. 1967; 41:1526–1533. [PubMed: 6075560]
- Qin MK, Oxenham AJ. Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. *Journal of the Acoustical Society of America*. 2006; 119:2417–2426. [PubMed: 16642854]
- Ritsma RJ. Frequencies dominant in the perception of the pitch of complex sounds. *J Acoust Soc Am*. 1967; 42:191–198. [PubMed: 6052077]
- Ritsma, RJ. Periodicity detection. In: Plomp, R.; Smoorenburg, GF., editors. *Frequency Analysis and Periodicity Detection in Hearing*. Sijthoff; Leiden, The Netherlands: 1970. p. 250-263.
- Richards VM, Zhu S. Relative estimates of combination weights, decision criteria, and internal noise based on correlation coefficients. *J Acoust Soc Am*. 1994; 95:423–434. [PubMed: 8120253]
- Schouten, JF. *Five articles on the perception of sounds (1938–1940)*. Institute for Perception; Eindhoven, Netherlands: 1940.
- Shackleton TM, Carlyon RP. The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. *J Acoust Soc Am*. 1994; 95:3529–3540. [PubMed: 8046144]
- Singh S, Kong YY, Zeng FG. Cochlear Implant Melody Recognition as a Function of Melody Frequency Range, Harmonicity, and Number of Electrodes. *Ear Hear*. 2009; 30:160–168. [PubMed: 19194298]
- Shamma S, Klein D. The case of the missing pitch templates: how harmonic templates emerge in the early auditory system. *J Acoust Soc Am*. 2000; 107:2631–2644. [PubMed: 10830385]
- Srulovicz P, Goldstein JL. A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum. *J Acoust Soc Am*. 1983; 73:1266–1276. [PubMed: 6853838]
- Terhardt E. Zur Tonhohenwahrnehmung von Klängen II. Ein Funktionsseheema, *Acustica*. 1972; 26:187–199.
- Terhardt E. Pitch, consonance, and harmony. *J Acoust Soc Am*. 1974; 55:1061–1068. [PubMed: 4833699]
- Wightman F. The pattern-transformation model of pitch. *J Acoust Soc Am*. 1973; 54:407–416. [PubMed: 4759014]
- Yost WA. The dominance region and ripple noise pitch: A test of the peripheral weighting model. *J Acoust Soc Am*. 1982; 72:416–425. [PubMed: 7119284]
- Zeng FG. Temporal pitch in electric hearing. *Hear Res*. 2002; 174:101–106. [PubMed: 12433401]

Figure 1.

Stimuli (top) and procedure (bottom) used in the pitch-comparison experiment. The stimuli were harmonic sounds with $F0 = 200\text{ Hz}$ and varying number of partials in four conditions. The procedure was two-interval, forced-choice, with the listener indicating which interval had higher pitch.

Figure 2.

Weights as a function of harmonic number, from three individual listeners (panels 1 to 3) and their mean (panel 4), and from four conditions (as indicated by different symbols). The two horizontal dotted lines sandwiching the line of zero weight represent plus and minus two standard errors ($2\sigma_e = 2/\sqrt{900}$) from zero correlation. In the panels for individual listeners, data points inside the area bounded by these two lines are deemed no different from zero. The error bars in the mean panel show plus and minus one standard error of the mean. The large sizes of the error bars reflect the considerable individual differences in listening strategy.

Figure 3. Weights estimated from S3 in Condition 3, with (filled triangles) or without (unfilled triangles), same as in Fig. 2) a masking noise low-pass filtered at the frequency of the cubic combination tone.

Figure 4. Weights estimated from S1 with spectral shaping applied to the complex sounds. In each condition, the sound pressure level was 40 dB for the lowest partial, 50 dB for the second lowest partial, and 60 dB for all the rest.

Figure 5. Predicted weights (solid line) and the mean estimated weights (circles with dashed line, same as in Fig. 2) for Condition 1. The predictions were derived from Goldstein (1973)'s pitch model.

Figure 6. Predicted weights (solid line) and the mean estimated weights (triangles with dashed line) for Condition 3. Note that the mean results shown here were different from the mean results presented in Fig. 2 for the same condition; they were computed using S3's weights obtained with the masking noise. Computed this way, the outcome is free from the influence of the combination tone.