Effects of the pattern of spectral spacing on the perceptual fusion of harmonics

Brian Roberts[a] and Albert S. Bregman
Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec H3A 1B1, Canada

(Received 12 May 1991; accepted for publication 8 July 1991)

Harmonic relations are important in the perceptual fusion of frequency components, but other factors can influence how much any given harmonic is integrated into a complex tone. This study considers whether the pattern of spectral spacing can act as such a factor. A complex tone consisting only of odd-numbered harmonics has a regular harmonic spacing of twice the fundamental frequency. The addition of a single even-numbered harmonic will locally disrupt this regular pattern. If harmonic relations alone limit the perceptual integration of the added harmonic, it should be as well integrated into the complex as its odd-numbered neighbors. Subjects were required to "hear out" one of the components of the complex and either to rate its perceived clarity (experiment 1) or to judge its pitch in relation to a pure tone 6% higher or lower in frequency (experiment 2). For a fundamental of 100 or 200 Hz, but not 400 Hz, the even harmonic generally could be heard out more easily than its odd neighbors. It is proposed that the poorer integration of the even harmonic at low fundamentals may result from a cross-channel comparison of amplitude-modulation rates at the output of the auditory filterbank.

PACS numbers: 43.66.Lj, 43.66.Hg, 43.66.Jh [WAY]

INTRODUCTION

Many experimental studies have shown that harmonically related spectral components tend to fuse perceptually into a single tone—a coherent perceptual entity with an associated pitch and timbre. This indicates the operation of a mechanism that can segregate harmonically related components from inharmonic components, thereby reducing the effects of any inharmonic components on the perceptual qualities of the complex tone. Duifhuis *et al.* (1982) proposed a mechanism of this kind, the harmonic sieve, as a means of excluding extraneous sounds from the estimation of low pitch. These authors conceived of the hypothetical harmonic sieve as a mechanism that passes only those components whose frequencies fall (within some tolerance range) at a harmonic multiple of the fundamental ($F0$). Therefore, all the harmonics of a periodic complex tone will be passed when the sieve's fundamental matches that of the complex tone, but inharmonic components will be excluded.

Several experiments have indicated that the principal determinants of the low pitch of a complex tone are the low-numbered harmonics, which are well resolved (Plomp, 1967; Ritsma, 1967; Moore *et al.*, 1985a). For example, Moore *et al.* (1985a) have shown that the dominant harmonics always lie within the first six for complex tones with equal-amplitude harmonics and fundamental frequencies of 100, 200, or 400 Hz. Therefore, a mechanism that operates in the frequency domain, such as the hypothetical harmonic sieve, provides a plausible means of excluding extraneous sounds from the process of pitch estimation. Indeed, the concept of the harmonic sieve has been successfully applied to human pitch judgments of complex tones when a low harmonic has been mistuned from its proper value (Moore *et al.*, 1985a; Moore, 1987). An alternative mechanism, but one related to the concept of the harmonic sieve, has been proposed by Hermes (1988). He presents a time-domain model for determining the low pitch of a complex tone by subharmonic summation. However, the two approaches are equivalent in formal terms and it is currently unclear which approach is more realistic. Therefore, we shall confine our discussion to the harmonic sieve.

A. Harmonic relations and perceptual grouping

Although the concept of the harmonic sieve was originally developed in the context of pitch perception, it has been extended to other aspects of auditory perception. Specifically, the sieve concept has also been applied to judgments of vowel quality (Scheffers, 1983; Darwin and Gardner, 1986; Roberts and Moore, 1990, 1991) and to perceptual grouping (Moore *et al.*, 1985b, 1986; Moore, 1987). Moore and his colleagues asked whether a component that has been rejected by the harmonic sieve is actually heard out from the complex as a separate tone. They reasoned that if this were the case, then the harmonic sieve could also be regarded as a mechanism for the formation of perceptual streams (Bregman, 1978, 1990). They found that only a small degree of mistuning, between 1.3% and 2.1%, was required for subjects to hear a low component of an otherwise harmonic complex tone as a separate tone (Moore *et al.*, 1986). This finding is broadly consistent with a role for the harmonic sieve in perceptual grouping, but these thresholds are still below the limit of 3%–4% beyond which a mistuned harmonic ceases to make its full contribution to the pitch of the complex (Moore *et al.*, 1985a). It seems that a mistuned

[a] Current address: Department of Psychology, University of York, Heslington, York YO1 5DD, England.

harmonic can be heard as a separate tone while still contributing to low pitch—a form of "duplex perception" (Bregman, 1987). However, the results of Moore *et al.* (1985a) show that the exclusion of a mistuned component from the computation of pitch is gradual rather than all or none in character. With this observation in mind, Moore *et al.* (1986) speculated that a component may be weighted less and less in the computation of pitch as it becomes more and more mistuned, and at the same time it becomes more clearly audible as a separate tone.

Merely being a harmonic multiple of $F0$ does not guarantee the total perceptual integration of a given component into a periodic complex tone. For example, if a harmonic in the first-formant ($F1$) region of a voiced vowel begins before or ends after the other harmonics, then its contribution to perceived vowel quality is reduced (Darwin, 1984a,b; Darwin and Sutherland, 1984). Other constraints are known to operate even when a harmonic begins and ends in synchrony with the rest of the components. For instance, a harmonic in a flat-spectrum complex tone becomes audible as a separate tone once its level is sufficiently high with respect to that of its neighbors. As an example, Martens (1981, 1984) found that incrementing the fourth harmonic of a 200-Hz fundamental by about 10 dB made it just audible as a separate tone. A related example is the finding that increasing the level of a harmonic near the $F1$ peak can have less effect on perceived vowel quality than would be predicted on the basis of a complete integration of the extra energy (Darwin, 1984b). In this case, subjects often reported the presence of an extraneous sound accompanying the vowel. Reduced effects on perceived vowel quality were especially evident for manipulations of harmonics on the lower skirt of $F1$, and Darwin (1984b) has hypothesized that this reflects the influence of constraints on plausible spectral shape.

B. Frequency resolution and perceptual grouping

Plomp (1964, 1976) and Plomp and Mimpen (1968) investigated the ability of listeners to hear out a harmonic in a periodic complex tone. Subjects were required to listen to a complex tone with 12 harmonics and to two comparison sinusoidal tones. One tone had a frequency equal to one of the harmonics, and the other had a frequency that was halfway between two harmonics. The subject's task was to choose which sinusoidal tone was actually present in the complex. The results across a wide range of fundamentals showed that subjects could easily hear out the low harmonics, but experienced increasing difficulty as the harmonic number increased. For each fundamental tested, the harmonic number for which listeners made 75% correct responses was defined as the critical number of harmonics that could be heard out. The values of the critical number obtained for the range of fundamentals tested were used to estimate the critical component spacing required to hear out a harmonic in different frequency regions. Across frequency, this measure was found to be very similar to estimates of the critical bandwidth (e.g., Zwicker *et al.*, 1957), suggesting that the data can be regarded as a measure of the ear's resolving power.

A task in which subjects are required to hear out a component from a complex tone can also be used to assess the extent of its perceptual integration into the complex. It is important, however, to be aware of the conceptual difference between resolving power and perceptual segregation. Hartmann *et al.* (1990) emphasize this distinction by contrasting Plomp's experiments with their own investigation of listeners' abilities to hear a mistuned harmonic in an otherwise periodic complex tone. These authors note that the results of Plomp's (1964) study were essentially unchanged when the harmonic complex tone was replaced by a complex comprised of inharmonic components, whereas their own experiments on perceptual segregation depended on inharmonicity as a major cue. Moreover, they argue that this conceptual distinction underlies the contrast between the Plomp and Mimpen (1968) finding that the highest resolvable harmonic of 200 Hz is the fifth and their own finding that listeners can successfully segregate the tenth harmonic of 200 Hz when it has been mistuned by only 1%. The Hartmann *et al.* data demonstrate a greater ability to hear out a given component from a complex than would be predicted on the basis of frequency resolution. A finding of this kind might be regarded as evidence for the operation of a perceptual segregation process.

C. Developing a stimulus paradigm

Although there is a body of evidence consistent with the notion of the harmonic sieve as a mechanism underlying the perceptual fusion of low harmonics and the exclusion of inharmonic components, the experimental manipulations employed have generally focused on simple cases. For example, the complex tones employed have comprised all harmonics below a specified upper cutoff frequency, and only one harmonic has been mistuned at any one time. A stimulus paradigm of potential interest is one where more than one harmonic is missing. Stimuli of this kind may provide a means of investigating further the role of harmonic relations in perceptual grouping. Of particular interest is a stimulus for which the incomplete harmonic series itself forms a regular pattern, and that is perceived as a single perceptual entity rather than as several sound sources coincident in time. These two criteria are met by an odd-harmonic series, which has a spectrum consisting only of odd-integer multiples of the fundamental frequency.

A complex tone consisting only of odd-numbered harmonics has a regular pattern of spectral spacing: The frequency separation between each harmonic is twice the frequency of the fundamental. Given this regularity, it is interesting to consider the consequences of adding a single, component at a low, even-integer multiple of $F0$. In physical terms, such a component would be unambiguously harmonic. As an integer multiple of $F0$, it should be passed by the harmonic sieve, like its odd-harmonic neighbors, and its addition should not change the low pitch of the complex tone. Indeed, if harmonic relations alone limit the perceptual integration of the added even harmonic, it should be as well integrated into the complex as its odd neighbors. However, the addition of a single even harmonic violates the equal steps (in linear terms) of spectral spacing, thus locally disrupting the regular spectral pattern of an odd-only harmonic series. By analogy with the visual phenomena described by the Gestalt psychologists (see Koffka, 1935), it might be said that the "good form" of the odd components is not shared by the single even component. This lack of "good form" might be exploited by perceptual grouping mechanisms to segregate the even harmonic from the rest of the complex tone.

I. EXPERIMENT 1

If the pattern of spectral spacing is an important factor in determining the perceptual integration of components in a harmonic complex tone, then a single, low-numbered even harmonic may not integrate fully into a complex tone consisting otherwise only of odd harmonics. In order to evaluate this prediction, a single even harmonic was added to an odd-harmonic complex tone and the degree to which it could be heard out from the complex was compared with the degree to which its neighboring odd harmonics could be heard out. Sections I A–I E describe the method and data analysis employed in this experiment. Experiment 2 employed a related method and analysis. Differences, where they occur, are described in Sec. II.

A. General method

Although the ear's resolving power is sufficient for listeners to hear out the low harmonics of a periodic complex, such a complex is usually heard as a single entity with an associated pitch and timbre. In other words, the complex is perceived "synthetically." For a pure tone component of the complex to be discerned, an analytic mode of listening is required. Martens (1981) notes that this mode of listening is typically induced by some means of directing attention to the appropriate component; for example, by varying its amplitude during the experiment (Duifhuis, 1970) or by allowing the listener to switch freely between the complex and the component in isolation (Plomp, 1964; Plomp and Mimpen, 1968).

There is clearly a close relation between techniques for focusing a listener's attention on a particular component in a complex tone and the phenomena of perceptual grouping. For instance, several studies have established that the frequency proximity of pure tones is a good basis for sequential grouping (e.g., Bregman and Pinker, 1978). Hence preceding the complex tone by a pure tone close in frequency to one of the harmonics might be regarded as providing an alternative perceptual organization for that harmonic. The competition between the two groupings may then be exploited as a means of seeing how strongly that component is fused into the complex. If a given harmonic is well integrated into the complex, it will tend not to group with the preceding tone and will therefore be difficult to hear out. If it is poorly integrated, however, it will tend to be "captured" into a separate stream by the preceding tone and will be easy to hear out. Hence the perceived clarity of the cued component can provide a means of assessing its perceptual integration into the complex. This was the approach employed in experiment 1.

B. Stimuli

The basic stimulus pattern was a sequence of two tones (A and B). Tone A was a pure tone and tone B was a complex tone. Tones A and B were separated by a brief silent interval (50 ms). Tone B was synthesized with one of three fundamental frequencies (100, 200, or 400 Hz). The basic form of tone B consisted of all odd-integer multiples of $F0$, up to and including the 15th harmonic. Four further variants of tone B were derived from the basic form by adding a single even-integer multiple of $F0$, either the second, fourth, sixth, or eighth harmonic. The frequency of tone A was chosen to match that of one of the harmonics 1–9, present in tone B. When tone A matched one of the odd-numbered harmonics, tone B was of the basic form (the "odd-only" complex tone). When tone A matched one of the even harmonics, tone B consisted of the basic form plus the appropriate even harmonic.

All stimuli were generated using a version of the MITSYN software package (see Henke, 1987). With one exception, the complex tones consisted of equal-amplitude components in cosine phase, created by additive sine-wave synthesis. This minimizes the confounding of differences between the even and odd harmonics with those resulting from level differences between the components. The exception was the fundamental component of the complex tone with an $F0$ of 100 Hz. This component was boosted by 6 dB with respect to the other components of the complex, to ensure that its loudness more closely matched that of the other components. The basic form of tone B was presented at a level of 67 dB SPL (C weighting). Tone A was presented at the same level as the corresponding harmonic in tone B. Both tone A and tone B were presented for a duration of 440 ms, including linear onset/offset ramps of 20 ms each.

C. Apparatus

The stimuli were synthesized and played at a sampling rate of 16 kHz by a Compaq DeskPro 386/20 computer controlling a 16-bit D/A converter (Data Translation DT2823). The stimuli were low-pass filtered at 6.5 kHz using both channels of a Rockland Dual Filter Model 852 (48 dB/oct/channel). Stimuli were presented binaurally over Sennheiser HD424 earphones. These earphones are designed to have a "free field" frequency response; i.e., the response at the eardrum is similar to what would be obtained with a flat-response loudspeaker in a free field. The levels of the stimuli were adjusted using a decade voltage divider (General Radio Type 1454-A), and were calibrated with a sound-level meter (General Radio Type 1551-C) connected to the earphones by a flat-plate coupler.

D. Procedure

Twelve subjects, aged from 19–39, were tested individually in an audiometric chamber (Industrial Acoustics Model 1202). All subjects reported having no hearing problems.

Subjects were told that on each trial they would hear a sequence of two tones, A and B, separated by a brief silent interval. Tone A was described as "pure" in quality, and tone B as "rich" in quality. A copy of tone A was said to be embedded in tone B. The subjects' task was to listen carefully to tone A and to rate how clearly they could hear its repetition in tone B, on a scale of 1–5 (perceived clarity). The extremes of the scale (1 and 5) indicated a "very unclear" and "very clear" repetition, respectively. Subjects were asked to ignore the complex tone as much as possible when listening for the sequence of two pure tones. A keyboard was used by the subjects to record their responses. After each response, there was an interval of about one second before the next trial.

The stimuli were presented as three sets of trials, each corresponding to one of the three fundamental frequencies employed (100, 200, and 400 Hz). The order of these sets was counterbalanced across the 12 subjects (2×6 permutations). A set of trials consisted of ten blocks of trials. Each block consisted of a single presentation, in quasirandom order, of the nine stimuli corresponding to harmonics 1–9 for the appropriate $F0$. The order of presentation within a block was randomized anew for every repetition of that block. Thus the total number of stimuli presented to subjects during the main experimental session was 270 stimuli (3 $F0$s $\times 9$ harmonics $\times 10$ repetitions). Subjects were free to rest at any time.

Before the main session, the subjects were given a two-part practice session. In the first part, subjects were given one repetition of each of the stimuli (27 stimuli $= 9\times 3$). In the second part, they were given two repetitions (54 stimuli). These were given to anchor the judgment scale prior to each run of trials. The order of the three sets of trials was the same as used for that subject in the main session. Data from the practice session were then discarded.

E. Data analysis

Each subject made a total of ten responses per stimulus, on a scale of 1–5. These ten responses were summed to give a perceived clarity score for that stimulus. Thus the minimum score for a given stimulus was 10 and the maximum was 50. The mean score for each stimulus was derived by averaging the scores across the 12 subjects. At no stage in the experiment were any subjects rejected.

The experiment may be described as having a two-way within-subjects design. The two factors, harmonic number and $F0$, had nine and three levels, respectively. An analysis of variance (ANOVA) was used to assess the significance of changes in the mean score across harmonic number and $F0$. This ANOVA gave a highly significant interaction between harmonic number and $F0$ [$F(16,176) = 3.559, P < 0.0001$], so a one-way within-subjects ANOVA was performed separately for each fundamental.

The principal purpose of the data analysis was a planned comparison of the scores for the odd and the even harmonic numbers, which was achieved for each of the three one-way ANOVAs using a contrast within the harmonic number factor. For the harmonic numbers tested, 1–9, five are odd and four are even, so the contrast weights used were $-1/5$ for each odd harmonic and $+1/4$ for each even harmonic.

F. Results and discussion

The results are displayed in Fig. 1. Parts (a), (b), and (c) correspond to fundamentals of 100, 200, and 400 Hz, respectively. Each part shows the mean scores, with standard errors, across harmonic numbers 1–9.

A small but significant effect on the mean score was found for harmonic number for the lower fundamentals [$F0 = 100$ Hz: $F(8,88) = 2.085$, $P = 0.0456$], [$F0 = 200$ Hz: $F(8,88) = 2.243$, $P = 0.0315$]. The planned comparison between the odd- and even-numbered harmonics showed the difference between the mean scores for these two fundamentals to be significant [$F0 = 100$ Hz: $F(1,11) = 18.013$, $P = 0.0014$], [$F0 = 200$ Hz: $F(1,11) = 8.978$, $P = 0.0122$]. For a fundamental of 400 Hz, a highly significant effect on the mean score was found for harmonic number [$F(8,88) = 3.524$, $P = 0.0014$]. In contrast with fundamentals of 100 and 200 Hz, the planned comparison between the odd- and even-numbered harmonics showed that the mean scores for these two groups did not differ significantly [$F(1,11) = 4.128$, $P = 0.0670$].

These observations indicate that when the fundamental of the complex tone was 100 or 200 Hz, the single even harmonic could generally be heard out more clearly than its odd neighbors. The exception was the lowest even harmonic tested (number 2). Apart from the second harmonic, the greater ease with which the single even harmonic could be heard out as a separate perceptual entity (i.e., a pure tone that accompanied the complex tone) implies that its perceptual integration into the complex tone was weaker than that of the odd harmonics. This finding is especially noteworthy, given that the added even harmonic was judged in a situation where the spectral density was greater than when the neighboring odd harmonics were tested (see Sec. I B). In relative terms, this might be expected to reduce the tendency of the even harmonic to segregate from the complex tone, since less well resolved components should be more difficult to hear out (Plomp, 1964, 1976; Plomp and Mimpen, 1968).

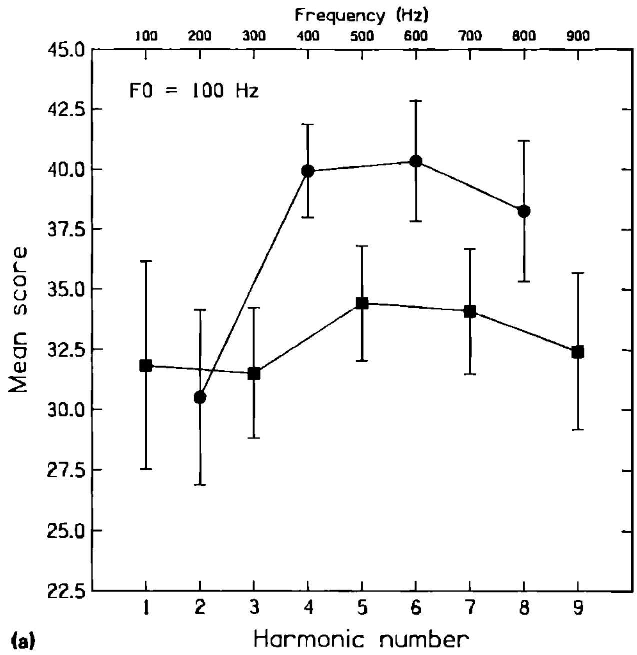
A different picture was observed when the fundamental of the complex tone was 400 Hz. In this case, the added even harmonic did not appear to be easier to hear out than its odd neighbors for any of the harmonic numbers tested. However, before a firm interpretation of these data can be reached, an important issue concerning the nature of the subjects' task must be considered. A judgment of perceived clarity cannot be independently verified, and so it cannot be demonstrated unequivocally that the judgment is governed by how well the subject can hear out the harmonic in tone B that corresponds to tone A. This problem can be addressed by asking subjects to make judgments based on the frequencies of the harmonics in tone B, as such judgments can be independently verified as correct or incorrect.

II. EXPERIMENT 2

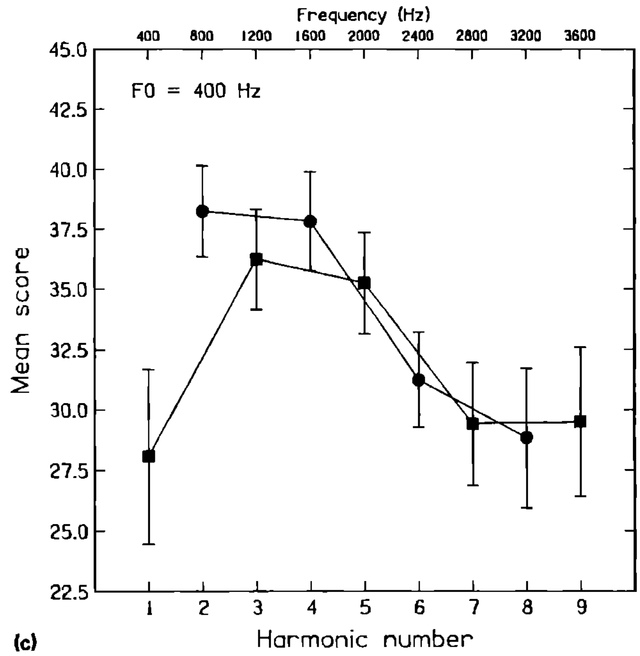
A. Method

1. Stimuli

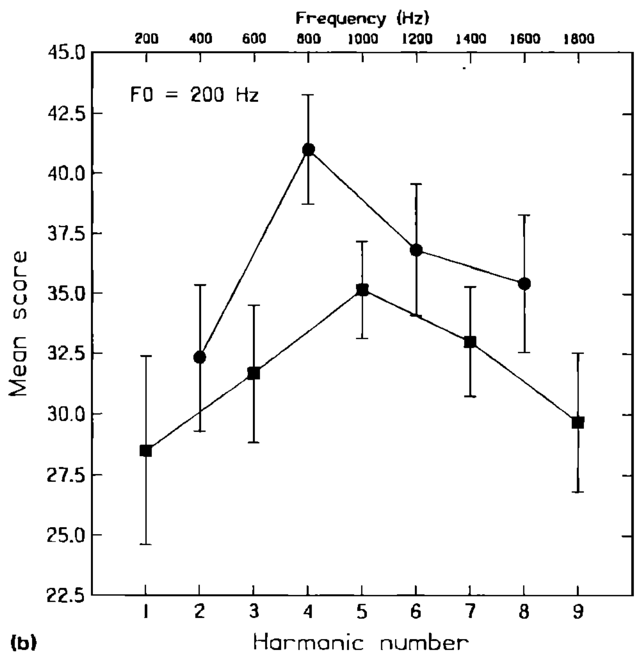
The stimuli employed in experiment 1 were modified for this experiment in the following way. Tone A was mistuned to a frequency either 6% above or below the frequency of the



(a)



(c)



(b)

FIG. 1. Parts (a), (b), and (c) correspond to fundamentals of 100, 200, and 400 Hz, respectively. Each part shows the mean scores, with standard errors, for 12 subjects across harmonic numbers 1–9. The mean scores for odd and even harmonic numbers are represented by filled squares and filled circles, respectively. The scores were derived from ratings of perceived clarity (see the text). The minimum score possible was 10 and the maximum score possible was 50.

corresponding harmonic in tone B. Since the tendency for tone A and the corresponding harmonic to form a sequential stream depends on their frequency proximity (Bregman and Pinker, 1978), the mistuning of tone A will weaken this effect. However, a preliminary investigation found that mistuning tone A by 6% did not greatly reduce the ability of subjects to hear out the corresponding harmonic. Another effect of mistuning tone A is that the frequency proximity of tone A and the next-nearest harmonic in tone B is increased. This effect increases with harmonic number, which might further reduce the tendency for tone A to form a sequential stream with the corresponding harmonic, or perhaps cause it to group with the "wrong" harmonic. However, for the harmonic numbers tested, 1–9, tone A was always nearest in frequency to the appropriate harmonic in tone B. The silent interval between tones A and B was 200 ms in duration, longer than that employed in experiment 1. This value was chosen because some of the subjects in a pilot study reported that the pitch jump between tone A and the corresponding harmonic in tone B sounded more "natural" in quality.

2. Procedure

Though small, a frequency difference of 6% between tone A and the corresponding harmonic in tone B is sufficient to judge whether the two-tone sequence ascends or descends in frequency. Since a correct response probably depends on the ability to discriminate the frequency of the attended harmonic, this judgment can be considered to add an accuracy dimension to the task.

As before, subjects were told that on each trial they would hear a sequence of two tones, A and B, separated by a brief silent interval. Tone A was described as "pure" in quality, and tone B as "rich" in quality. A pure tone close in frequency to tone A was said to be embedded in tone B. The subjects' task was to listen carefully for a sequence of two tones, tone A and the pure tone embedded in tone B. They were required to judge whether the sequence went *down* or *up* in pitch, and also to rate the perceived clarity of the second tone. The total number of stimuli presented to subjects during the main experimental session was 540 ($3F0$s $\times$ 9 harmonics $\times$ 2 mistunings $\times$ 10 repetitions).

All subjects began with a two-stage screening procedure. In the first stage, all the components of tone B were removed, except for the harmonic corresponding to tone A. Thus each stimulus consisted of a sequence of two pure tones, where the second tone was either 6% higher or lower in frequency than the first tone. Subjects were required to listen to the two-tone sequence and to judge whether the pitch of this sequence went down or up. The purpose of this screening procedure was to ensure that all subjects were able to identify the direction of the pitch shift correctly when no competing sounds were present. An ability to do this was a prerequisite for participating in the main session. The total number of stimuli employed was 54 ($3F0$s $\times$ 9 harmonics $\times$ 2 mistunings). Subjects who were unable to reach a criterion of 90% correct on this task were eliminated. Nineteen of the 26 subjects tested were successful.

The second stage of screening employed stimuli similar to those used in the main experimental session, except that all components in tone B other than the harmonic corresponding to tone A were attenuated by 12 dB. Subjects were again required to make a down/up judgment. As before, the total number of stimuli employed was 54. Subjects who were unable to reach a criterion of 85% correct on this task were also discarded. Twelve of the 19 subjects tested were successful.

The remaining twelve subjects were then given a two-part practice session before the main experimental session. Subjects were now required to make both a down/up judgment and a perceived clarity judgment, as in the main session. In the first part, subjects were given one repetition of each of the stimuli (54 stimuli $= 9 \times 3 \times 2$). In the second part, they were given two repetitions (108 stimuli). These were given to familiarize subjects with the response scale and to stabilize their judgments prior to the main session. The order of the three sets of trials was the same as used for that subject in the main session. Data from the practice session were then discarded.

3. Data analysis

The original intention was to derive a composite score for each stimulus by combining the accuracy score for the down/up judgment with the perceived clarity rating. This approach was abandoned for two reasons. First, there are many ways in which such a composite score could be derived, each of which is arbitrary. Second, there are circumstances in which the interpretation of a composite score might be misleading. For example, it is possible for a composite score to be positive even if 50% or more of the down/up judgments are incorrect. In principle, this could occur if the incorrect responses were associated with low clarity ratings and the correct responses were associated with high clarity ratings. Therefore, the analysis presented here is restricted to the accuracy scores for the down/up judgment alone.

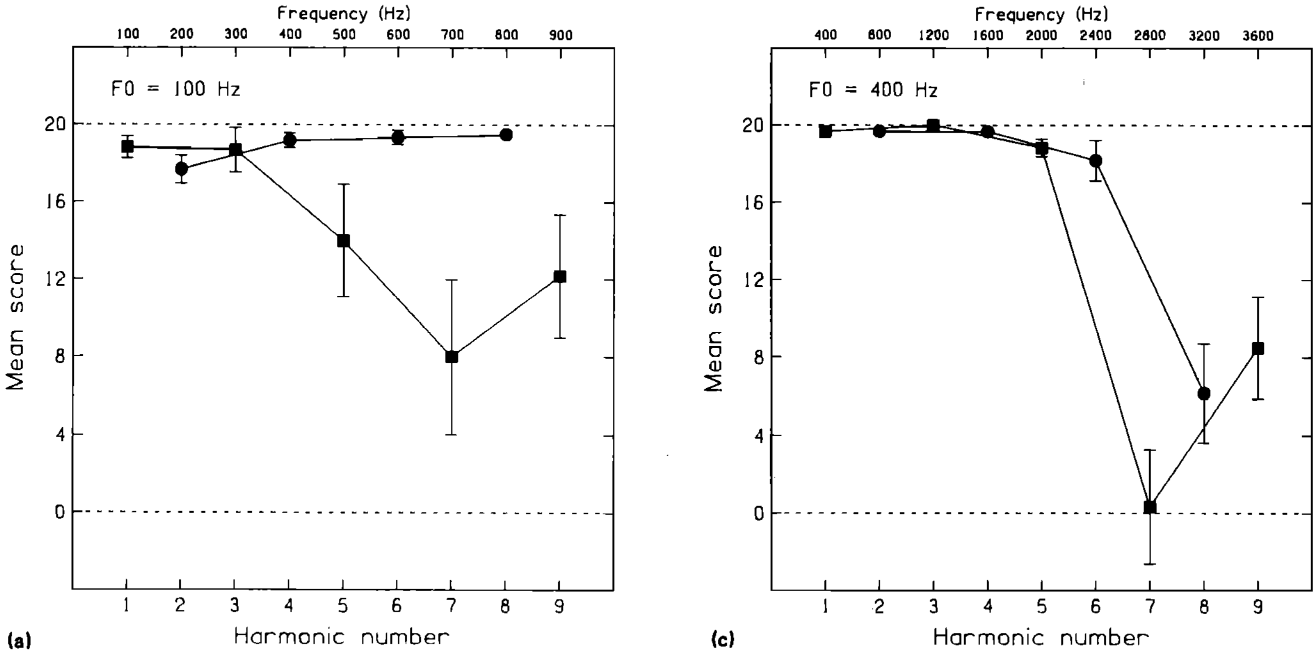
For a given trial, correct and incorrect responses were assigned $+1$ and -1, respectively. Each subject made a total of ten responses per stimulus, and these were summed to give a score for that stimulus. Since the pattern of responses for a given harmonic number was broadly similar for both an upward and a downward mistuning of tone A, these scores were added. Therefore, for a given $F0$ and harmonic number, the minimum score was -20 and the maximum score was $+20$. The mean score for each combination of $F0$ and harmonic number was derived by averaging the scores across the 12 subjects.

B. Results and discussion

The results are displayed in Fig. 2. Parts (a), (b), and (c) correspond to fundamentals of 100, 200, and 400 Hz, respectively. Each part shows the mean scores, with standard errors, across harmonic numbers 1–9. Note that these data indicate a greater-than-chance level of performance for most of the stimuli tested. This suggests that information about the frequency of the attended harmonic is generally preserved for the harmonic numbers and fundamentals tested.

A two-way within-subjects analysis of variance (ANOVA) was used to assess the significance of changes in the mean score across harmonic number and $F0$. This ANOVA gave a highly significant interaction between harmonic number and $F0$ [$F(16,176) = 6.098$, $P < 0.0001$]. The significance of the interaction clearly results from differences in the accuracy of performance across $F0$ for the higher harmonics tested, the fifth and above. For the lower two $F0$s, the accuracy of the down/up judgment is near perfect for the even harmonics, but declines for the odd harmonics, whereas, for an $F0$ of 400 Hz, the accuracy of the judgment declines for both odd and even harmonics. For harmonics up to and including the fourth, the accuracy of the down/up judgment was near perfect for both odd and even harmonics. This marked ceiling effect in the data may have obscured differences between the lower odd and even harmonic numbers in the relative extent of their perceptual integration into the complex tone.

Since the interaction term was significant, a one-way within-subjects ANOVA was performed separately for each of the fundamentals. In each case, a highly significant effect on the mean score was found for harmonic number [$F0 = 100$ Hz: $F(8,88) = 4.605$, $P = 0.0001$], [$F0 = 200$ Hz: $F(8,88) = 12.943$, $P < 0.0001$], [$F0 = 400$ Hz: $F(8,88) = 20.251$, $P < 0.0001$]. An inspection of Fig. 2 indicates that this primarily reflects the fall in mean scores for the odd harmonics with increasing harmonic number. An effect of this kind was anticipated earlier on the grounds that



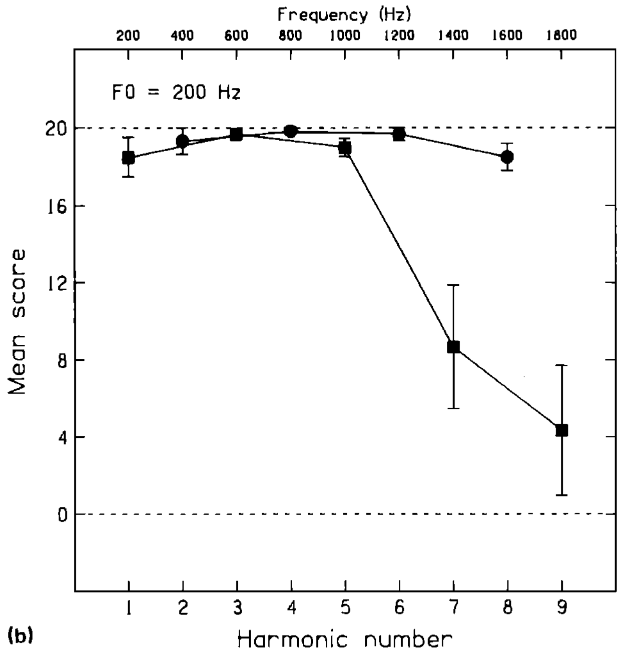


FIG. 2. Parts (a), (b), and (c) correspond to fundamentals of 100, 200, and 400 Hz, respectively. Each part shows the mean scores, with standard errors, for 12 subjects across harmonic numbers 1–9. The mean scores for odd and even harmonic numbers are represented by filled squares and filled circles, respectively. Where an error bar is not visible, it falls within the symbol. The scores were derived from the accuracy of the down/up judgment (see the text). The minimum score possible was -20 and the maximum score possible was $+20$.

the mistuning of tone A increases its frequency proximity to the next-nearest harmonic in the complex tone as harmonic number increases (see Sec. II A 1).

The planned comparison between the odd- and even-numbered harmonics showed the difference between the mean scores for the two to be significant for fundamentals of 100 and 200 Hz [$F(1,11) = 7.470$, $P = 0.0195$; $F(1,11) = 18.704$, $P = 0.0012$], but not for a fundamental of 400 Hz [$F(1,11) = 3.355$, $P = 0.0942$]. The finding that the down/up judgments were more accurate for the even harmonics than for their odd harmonic neighbors for fundamentals of 100 and 200 Hz, but not 400 Hz, is consistent with the pattern of perceived clarity ratings observed in experiment 1. The pattern of results for the lower fundamentals is particularly striking given the experimental design bias against high scores for the even harmonics. As in experiment 1, the even harmonics were tested in a situation where the spectral density was greater than when the neighboring odd ones were tested (see Sec. I B). This might be expected to result in less accurate judgments for the even harmonics. Furthermore, the greater spectral density also results in a greater frequency proximity between tone A and the next-nearest harmonic when the even harmonics are tested, which should further reduce the scores for the even harmonics in relation to their odd neighbors.

C. Conclusions

For fundamentals of 100 and 200 Hz, the results of experiments 1 and 2 indicate that a single even-numbered harmonic can be segregated from a complex tone, consisting otherwise only of odd harmonics, more easily than can the odd-numbered ones. This indicates a role for factors other than harmonic relations in determining the perceptual grouping of the frequency components in the complex—in this instance, the pattern of spectral spacing. These experiments also indicate that the perceptual difference between an added even harmonic and its odd neighbors is attenuated or abolished when the fundamental frequency of the complex tone is 400 Hz.

III. GENERAL DISCUSSION

Experiments 1 and 2 demonstrate that there are conditions for which a single even harmonic is easier to hear out from a complex tone, consisting otherwise only of odd harmonics, than are its neighboring odd harmonics. This result is the opposite of that predicted on the basis of frequency resolution and indicates that, for the lower fundamentals, the single even harmonic is less well integrated into the complex. Current conceptions of the harmonic sieve, or of its formal equivalent in the time domain (subharmonic summation), cannot account for this.

In this discussion, we consider what factors might be responsible for this finding. Of course, a plausible hypothesis must also be able to account for the two instances in which the single even harmonic does not behave differently from its odd neighbors. First, the extent to which the added even harmonic differs perceptually from its odd neighbors is dependent on the fundamental frequency. Relative to the odd harmonics, significantly higher perceived clarity ratings are seen for the even harmonics when the $F0$ is 100 or 200 Hz, but this effect is much attenuated or absent when the $F0$ is 400 Hz. Second, for the lower $F0$s, the single even harmonic is easier to hear out than its odd neighbors for harmonic numbers 4, 6, and 8, but probably not when the harmonic number is 2. Hence two questions must be addressed: (1) Why is the perceptual integration of the added even harmonic dependent on $F0$? (2) Why, for the lower $F0$s, does harmonic number 2 not behave differently from its neighbors?

The original motivation for performing these experiments was expressed in terms of the Gestalt principle of "good form." It was argued that the addition of a single low-numbered even harmonic to a complex tone consisting otherwise only of odd harmonics would disrupt the pattern of equal spectral spacing. This disruption could be a basis for the perceptual segregation of the added component. However, this proposal cannot of itself explain either the $F0$ dependence or the absence of an effect for harmonic number 2. An increase in $F0$ from 100 to 200 to 400 Hz simply transposes the complex tone through two octaves. Given that auditory filter bandwidth is approximately constant on a logarithmic scale, the resolution of the partials in the complex should be similar. Indeed, auditory filter bandwidths are greater on a log scale at frequencies below 1 kHz (Moore and Glasberg, 1983, 1986), so the resolution of the partials will improve when the fundamental is increased from 100 to 400 Hz. This indicates that the attenuation or absence of a perceptual difference between an added even harmonic and its odd neighbors when $F0$ is 400 Hz cannot be explained in terms of poor frequency resolution.

A. Pitch effects

The relation between the temporal structure of an odd-harmonic complex and its perceived pitch may provide an explanation for the reduced perceptual integration of the single even harmonic observed for the lower $F0$s tested. The waveform of a complex tone consisting of successive in-phase harmonics is a unipolar pulse train (one pulse per period), whereas the waveform of an odd-harmonic complex is a pulse train of alternating polarity (two pulses per period). There is some evidence to suggest that, for low repetition rates, an odd-harmonic complex evokes a pitch that does not correspond to its $F0$ (Flanagan and Guttman, 1960; Warren and Wrightson, 1981). One might speculate that this lack of correspondence between the perceived pitch and the $F0$ for low repetition rates reduces the tendency of the added even harmonic to group with the odd-harmonic complex.

Flanagan and Guttman (1960) investigated the perceived pitch of several patterns of iterated pulses, including a pulse train of alternating polarity (odd harmonics only). They argued that the pitch of the bipolar pulse train would correspond to $F0$ only if the auditory system was sensitive to pulse polarity, otherwise it would correspond to the number of pulses per second (pps). Subjects were required to adjust the repetition rate of the bipolar pulse train until its pitch matched that of a unipolar reference stimulus (successive harmonics). Two modes of pitch perception were found. When the reference $F0$ was above 200 Hz, the pitch of the bipolar pulse train generally corresponded to $F0$, but when the reference $F0$ was below 100 Hz, its pitch generally corresponded to the number of pps (i.e., an octave higher than $F0$). Ambiguous pitch matches were common for reference $F0$s between 100 and 200 Hz, suggesting a transition between the two pitch modes. Warren and Wrightson (1981) reported a similar pattern of results for odd-harmonic tones with complex waveforms, generated by excising a segment from noise and iterating it in alternation with its polarity-inverted form. The region of transition between the two pitch modes in their data was for reference $F0$s between 30 and 200 Hz.

Although these studies demonstrate that an odd-harmonic complex may evoke a pitch that does not correspond to $F0$, there are important reasons for doubting that the $F0$ dependence observed in our data can be accounted for in this way.

(1) It is noteworthy that neither the Flanagan and Guttman (1960) study nor the Warren and Wrightson (1981) study showed any pitch ambiguity for an odd-harmonic complex when $F0$ was 200 Hz. Despite this, experiments 1 and 2 have both demonstrated that an added even

harmonic is significantly easier to hear out than its odd neighbors when $F0$ is 200 Hz.

(2) The Gerson and Goldstein (1978) study casts doubt on whether the odd-harmonic complex we employed would have evoked an ambiguous pitch even when its $F0$ was 100 Hz. They investigated the perceived pitch of stimuli composed of four successive odd harmonics and synthesized with an $F0$ of 100 Hz or in the region of 100 Hz. Two tasks were employed—a pitch-matching task and a musical-interval recognition task. In contrast with the Flanagan and Guttman (1960) findings, they found that the perceived pitch of these stimuli usually corresponded to $F0$.

The difference in the findings of these two studies may result from differences in the stimuli employed, particularly the total number of odd harmonics present, although such an explanation remains speculative. Given the lack of consensus in the literature, the only way to be sure that pitch ambiguity was not an important factor in our study was to test the odd-harmonic complexes used in experiments 1 and 2. Therefore, we asked four of the subjects who had participated in experiments 1 and 2 to take part in an informal pitch-matching task. They were required to match the pitch of the odd-harmonic complexes used in these experiments ($F0 = 100, 200$, and 400 Hz) with a sawtooth wave (successive harmonics, -12 dB/oct slope) of adjustable period. The matches obtained were very close to the actual $F0$ in all cases. Therefore, we consider it unlikely that the $F0$ dependence of the perceptual integration of an added even harmonic reflects a nonfundamental pitch for the lower fundamentals tested.

B. A temporal-pattern hypothesis

The auditory system may exploit temporal information partly to exclude the added even harmonic from the odd-harmonic complex tone. Indeed, any mechanism with limited temporal resolution might account for the $F0$ dependence of our data, since the repetition rate of the waveform doubles with an octave increase in $F0$. A likely source of temporal information is the output of those auditory filters where the frequency resolution is poor enough for more than one of the harmonics to interact.

Interactions between partials unresolved by a given auditory filter give rise to amplitude modulation (AM) at the output of that filter. If the rate of AM were compared across channels, it could provide a means of identifying those components that "belong" together. Since the spectral spacing of harmonics in an odd-harmonic series is twice the fundamental frequency, the envelope of the output of the channels in which components of the complex tone interact will be modulated at a rate of twice $F0$. The addition of a single even harmonic gives a local region in the spectrum where the frequency of the spacing is half this value, the same as the $F0$ frequency. If this component is not fully resolved from its neighbors, this will lead to a local region of the auditory filterbank in which the rate of AM equals the frequency of $F0$. This disruption of the uniform pattern of AM across channels might be detected by a cross-channel comparison of the AM rates at the output of the auditory filterbank. A perceptual grouping mechanism may then use the discrepancy as a means of segregating the even harmonic as the "odd man out."

C. Evidence for cross-channel comparisons

Several recent experiments have demonstrated that the ability to compare the outputs of different auditory filters enhances both signal detection and discrimination. For example, Green and his colleagues have shown that subjects can detect an increment of only 1–2 dB in the level of one component in a complex sound relative to the levels of the other components (see Green, 1988). Green has argued that this is achieved by detecting a change in the shape or "profile" of the spectrum of the sound, an effect he refers to as "profile analysis."

Another phenomenon indicating the importance of cross-channel comparisons in auditory perception was first demonstrated by Hall *et al.* (1984). These researchers showed that the detectability of a pure-tone signal masked by a band of noise centered on the signal is enhanced when the envelope of the masker fluctuates over time, and when the fluctuations are coherent or correlated across different frequency bands. This phenomenon is known as comodulation masking release (CMR). A number of models have been proposed to explain CMR (see Moore, 1990), many of which assume that CMR results from using the temporal structure in one frequency channel to aid detection of a masked signal in another frequency channel. Specifically, these models assume that CMR results from a comparison of envelope modulation patterns at the outputs of auditory filters tuned to different center frequencies (Hall *et al.*, 1984; Buus, 1985; Schooneveldt and Moore, 1987; Moore, 1988). When a signal is added, the modulation pattern at the output of the auditory filter tuned to the signal frequency is altered. Thus the presence of the signal is indicated by a *disparity* in the modulation pattern across different filters.

D. Cross-channel comparisons and auditory grouping

Our temporal-pattern hypothesis assumes that the detection of a disparity in the modulation pattern across different filters reduces the perceptual integration of the added even harmonic into the odd-harmonic complex. Thus we assume not only that the auditory system can perform cross-channel comparisons, but also that such comparisons can be used as a basis for auditory grouping. The latter assumption has recently received support from the work of Hall and his colleagues on CMR. Hall *et al.* (1989) investigated CMR for conditions in which six flanking bands were always comodulated with the on-signal band, but two additional flanking bands (termed "deviant" bands) had envelopes that were independent of the envelope of the on-signal band. They found that CMR was often substantially reduced when the deviant bands were present at spectral locations close to the signal frequency. Hall and Grose (1990) went on to investigate whether the disruptive effects of the deviant bands could be reduced by conditions that favored their integration into an auditory stream independent from that formed by the on-signal band and the comodulated flanking bands.

The basic stimulus employed by Hall and Grose consisted of six bands that were comodulated with the on-signal band, and two deviant bands close in frequency to the on-signal band. The other stimuli employed were derived from this stimulus. One manipulation was to add further deviant bands to the basic stimulus, placed far in frequency from the on-signal band, and sharing a common envelope with the original two deviant bands. These extra bands were termed codeviant bands. The idea was that as more codeviant bands were added, the two deviant bands near the signal frequency would tend to group with them rather than with the on-signal band, thus reducing their disruptive effects. Hall and Grose found that the disruptive effects of the two deviant bands were indeed reduced by increasing the number of codeviant bands. Furthermore, dichotic conditions showed that the disruptive effects could also be reduced by adding codeviant bands in the ear contralateral to the signal. This study provides clear evidence that cross-channel comparisons can be used for auditory grouping.

The Hall and Grose (1990) findings add credence to our suggestion for a temporal-pattern hypothesis, but there is reason to doubt that the phenomenon we have observed can be explained in the same way as CMR. In particular, several studies have shown that CMR decreases as the rate of envelope modulation increases (Hall and Haggard, 1983; Buus, 1985; Hall, 1987), and is small or absent at modulation rates comparable to those evoked by our stimuli. However, Bregman *et al.* (1985) have provided some evidence that auditory grouping can be based on periodicity information at modulation rates more similar to those evoked by our stimuli. They investigated the perceptual integration of a mixture of two complex tones, where each tone was formed by sinusoidal AM of a carrier sinusoid. One tone always had a carrier frequency (CF) of 1500 Hz and a modulation frequency (MF) of 100 Hz. The other had different CFs around 500 Hz and different MFs around 100 Hz. AM was used to produce both harmonic and inharmonic partials. They found that fusion was best when the higher and lower complex tones were modulated at the same rate and in phase, even when the resulting partials did not form part of the same harmonic series. Although the effects they observed were quite small, their study demonstrates the contribution of a periodicity-matching mechanism to perceptual fusion, and one that operates quite independently of the relation of the partials to a common $F0$.

E. The effects of temporal resolution

We have suggested that the auditory system may employ a cross-channel comparison of modulation rates at the output of the auditory filterbank as a means of perceptually segregating an added even harmonic from an odd-harmonic complex. This hypothesis can also provide a plausible explanation for those instances where the single even harmonic does not behave differently from its odd neighbors. For example, when harmonic number 2 is added to the odd-harmonic complex, it is probably too well resolved from its odd neighbors to interact significantly with them at the output of any auditory filters. Hence there will be no detectable local rate of AM to identify harmonic number 2 as the "odd man out" and the reduced perceptual integration evident for harmonic numbers 4, 6, or 8 is not seen.

There are several studies indicating that the temporal resolution of the auditory system may be an important factor underlying performance in our experiments. Temporal resolution is often quantified by measuring the threshold for detecting the amplitude modulation of a white noise as a function of modulation rate. The rate-threshold function so determined is called the temporal modulation transfer function (TMTF). Measures of the TMTF have shown that our sensitivity to modulation decreases with increases in the rate of modulation (Viemeister, 1979; Bacon and Viemeister, 1985). The threshold is independent of modulation rate up to about 16 Hz, but beyond this temporal resolution begins to have an effect and the threshold increases. Above about 1000 Hz, the modulation cannot be detected at all. Furthermore, predictions of the TMTF derived from the temporal-window model of temporal resolution (Moore *et al.*, 1988; Plack and Moore, 1990) suggest an even steeper cutoff for high rates of modulation.

The decrease in sensitivity to AM with an increase in AM frequency may explain the $F0$ dependence of the perceptual integration of the added even harmonic into the odd-harmonic complex. For an $F0$ of 100 Hz, a cross-channel comparison mechanism would need to detect a local region of 100-Hz modulation against a broad region of 200-Hz modulation. For an $F0$ of 400 Hz, however, such a mechanism would need to detect a local region of 400-Hz modulation against a background of 800-Hz modulation. The sensitivity to these high rates of modulation may simply be too poor for a cross-channel comparison to be an effective means of identifying the single even harmonic as disrupting the otherwise regular pattern of spectral spacing. Hence the effect is either greatly attenuated or absent for an $F0$ of 400 Hz.

F. Conclusions and further work

Moore (1990) has suggested that cross-filter comparisons of temporal envelopes are a general feature of auditory pattern analysis, which may play an important role in extracting sounds from noisy backgrounds and in separating competing sound sources. A mechanism of this kind clearly has the potential to explain the phenomena demonstrated by experiments 1 and 2. We suggest, therefore, that the reduced perceptual integration of the added even harmonic at low fundamentals may result from a cross-channel comparison of amplitude-modulation rates at the output of the auditory filterbank. As a preliminary account of our findings, the temporal-pattern hypothesis has the advantage that it provides a basis for predicting the effects of certain stimulus manipulations. Thus it is easily amenable to evaluation by experiment.

If the detection of modulation at the output of the auditory filters plays a key role in the reduced perceptual fusion of the added even harmonic at lower fundamentals, then any means of disrupting this cue should attenuate or abolish the effect at the lower fundamentals. This might be achieved in one of two ways: first, by a phase manipulation of the partials in the complex tone that minimizes the depth of AM at the output of the auditory filters; second, by the addition of a noise of sufficient level and bandwidth to mask the patterns of modulation evoked by the complex tone at the output of the auditory filters. These approaches to evaluating the temporal-pattern hypothesis are the subject of a subsequent paper.

In conclusion, the experiments reported here demonstrate that harmonic relations are not the only factor that can influence the perceptual fusion of simultaneous components in a complex tone, at least for fundamentals of 100 and 200 Hz. Although the underlying mechanism is as yet unclear, it is evident that the pattern of spectral spacing can be an important factor in auditory grouping.

ACKNOWLEDGMENTS

This research was supported by a grant from NSERC (Canada). Our thanks go to Peter Bailey, Brian Moore, Pierre Ahad, Chris Darwin, Richard Warren, Bob Carlyon, Bill Hartmann, and Ray Meddis for their valuable advice, criticism, and assistance.

Bacon, S. P., and Viemeister, N. F. (**1985**). "Temporal modulation transfer functions in normal-hearing and hearing-impaired subjects," Audiology **24**, 117–134.

Bregman, A. S. (**1987**). "The meaning of duplex perception: sounds as transparent objects," in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Nijhoff, Dordrecht, The Netherlands), pp. 95–111.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA).

Bregman, A. S., Abramson, J., Doehring, P., and Darwin, C. J. (**1985**). "Spectral integration based on common amplitude modulation," Percept. Psychophys. **37**, 483–493.

Bregman, A. S., and Pinker, S. (**1978**). "Auditory streaming and the building of timbre," Can. J. Psychol. **32**, 19–31.

Buus, S. (**1985**). "Release from masking caused by envelope fluctuations," J. Acoust. Soc. Am. **78**, 1958–1965.

Darwin, C. J. (**1984a**). "Auditory processing and speech perception," in *Attention and Performance X: Control of Language Processes*, edited by H. Bouma and D. G. Bouwhuis (Erlbaum, Hillsdale, New Jersey), pp. 197–210.

Darwin, C. J. (**1984b**). "Perceiving vowels in the presence of another sound: Constraints on formant perception," J. Acoust. Soc. Am. **76**, 1636–1647.

Darwin, C. J., and Gardner, R. B. (**1986**). "Mistuning a harmonic of a vowel: Grouping and phase effects on vowel quality," J. Acoust. Soc. Am. **79**, 838–845.

Darwin, C. J., and Sutherland, N. S. (**1984**). "Grouping frequency components of vowels: When is a harmonic not a harmonic?" Q. J. Exp. Psychol. **36A**, 193–208.

Duifhuis, H. (**1970**). "Audibility of high harmonics in a periodic pulse," J. Acoust. Soc. Am. **48**, 888–893.

Duifhuis, H., Willems, L. F., and Sluyter, R. J. (**1982**). "Measurement of pitch in speech: An implementation of Goldstein's theory of pitch perception," J. Acoust. Soc. Am. **71**, 1568–1580.

Flanagan, J. L., and Guttman, N. (**1960**). "On the pitch of periodic pulses," J. Acoust. Soc. Am. **32**, 1308–1319.

Gerson, A., and Goldstein, J. L. (**1978**). "Evidence for a general template in central optimal processing for pitch of complex tones," J. Acoust. Soc. Am. **63**, 498–510.

Green, D. M. (**1988**). *Profile Analysis* (Oxford U.P., New York).

Hall, J. W. (**1987**). "Experiments on comodulation masking release," in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, New Jersey).

Hall, J. W., and Grose, J. H. (**1990**). "Comodulation masking release and auditory grouping," J. Acoust. Soc. Am. **88**, 119–125.

Hall, J. W., Grose, J. H., and Haggard, M. P. (**1989**). "Effects of flanking band proximity, number, and modulation pattern on comodulation masking release," J. Acoust. Soc. Am. **87**, 269–283.

Hall, J. W., and Haggard, M. P. (**1983**). "Co-modulation—a principle for auditory pattern analysis in speech," Proc. 11th Int. Congr. Acoust. **4**, 69–71.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). "Detection in noise by spectrotemporal pattern analysis," J. Acoust. Soc. Am. **76**, 50–56.

Hartmann, W. M., McAdams, S., and Smith, B. K. (**1990**). "Hearing a mistuned harmonic in an otherwise periodic complex tone," J. Acoust. Soc. Am. **88**, 1712–1724.

Henke, W. L. (**1987**). MITSYN: A coherent family of high-level languages for time signal processing (Software package) (Belmont, MA: Author).

Hermes, D. J. (**1988**). "Measurement of pitch by subharmonic summation," J. Acoust. Soc. Am. **83**, 257–264.

Koffka, K. (**1935**). Principles of Gestalt Psychology (Harcourt and Brace, New York).

Martens, J.-P. (**1981**). "Audibility of harmonics in a periodic complex," J. Acoust. Soc. Am. **70**, 234–237.

Martens, J.-P. (**1984**). "Comment on 'Algorithm for extraction of pitch and pitch salience from complex tonal signals'," J. Acoust. Soc. Am. **75**, 626–628.

Moore, B. C. J. (**1987**). "The perception of inharmonic complex tones," in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, New Jersey), pp. 180–189.

Moore, B. C. J. (**1988**). "Dynamic aspects of auditory masking," in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 585–607.

Moore, B. C. J. (**1990**). "Co-modulation masking release: spectro-temporal pattern analysis in hearing," Br. J. Audiol. **24**, 131–137.

Moore, B. C. J., and Glasberg, B. R. (**1983**). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," J. Acoust. Soc. Am. **74**, 750–753.

Moore, B. C. J., and Glasberg, B. R. (**1986**). "The role of frequency selectivity in the perception of loudness, pitch and time," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 251–308.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1985a**). "Relative dominance of individual partials in determining the pitch of complex tones," J. Acoust. Soc. Am. **77**, 1853–1860.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1986**). "Thresholds for hearing mistuned partials as separate tones in harmonic complexes," J. Acoust. Soc. Am. **80**, 479–483.

Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (**1988**). "The shape of the ear's temporal window," J. Acoust. Soc. Am. **83**, 1102–1116.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1985b**). "Thresholds for the detection of inharmonicity in complex tones," J. Acoust. Soc. Am. **77**, 1861–1867.

Plack, C. J., and Moore, B. C. J. (**1990**). "Temporal window shape as a function of frequency and level," J. Acoust. Soc. Am. **87**, 2178–2187.

Plomp, R. (**1964**). "The ear as a frequency analyzer," J. Acoust. Soc. Am. **36**, 1628–1636.

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

Plomp, R. (**1976**). *Aspects of Tone Sensation* (Academic Press, London, New York).

Plomp, R., and Mimpen, A. M. (**1968**). "The ear as a frequency analyzer II," J. Acoust. Soc. Am. **43**, 764–767.

Ritsma, R. J. (**1967**). "Frequencies dominant in the perception of the pitch of complex sounds," J. Acoust. Soc. Am. **42**, 191–198.

Roberts, B., and Moore, B. C. J. (**1990**). "The influence of extraneous sounds on the perceptual estimation of first-formant frequency in vowels," J. Acoust. Soc. Am. **88**, 2571–2583.

Roberts, B., and Moore, B. C. J. (**1991**). "Modeling the effects of extraneous sounds on the perceptual estimation of first-formant frequency in vowels," J. Acoust. Soc. Am. **89**, 2933–2951.

Scheffers, M. T. M. (**1983**). *Sifting Vowels: Auditory Pitch Analysis and Sound Segregation*, Ph.D. thesis, University of Groningen, The Netherlands.

Schooneveldt, G. P., and Moore, B. C. J. (**1987**). "Comodulation masking release (CMR): Effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band," J. Acoust. Soc. Am. **82**, 1944–1956.

Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based on modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.

Warren, R. M., and Wrightson, J. M. (**1981**). "Stimuli producing conflicting temporal and spectral cues to frequency," J. Acoust. Soc. Am. **70**, 1020–1024.

Zwicker, E., Flottorp, G., and Stevens, S. S. (**1957**). "Critical bandwidth in loudness summation," J. Acoust. Soc. Am. **29**, 548–557.