

Spectral integration based on common amplitude modulation

ALBERT S. BREGMAN, JACK ABRAMSON, and PETER DOEHRING
McGill University, Montreal, Quebec, Canada

and

CHRISTOPHER J. DARWIN
University of Sussex, Brighton, England

The perceptual integration of a mixture of two complex tones was studied in experiments on adult subjects. Each tone was formed by amplitude modulation (AM) of a carrier sinusoid of frequency (CF) by a raised sinusoid with modulation frequency (MF). One tone always had CF = 1500 Hz and MF = 100 Hz. The other had different CFs around 500 Hz and different MFs around 100 Hz. Both harmonic and inharmonic partials, produced by AM, were employed. The method involved studying the competition between two perceptual organizations: (1) the fusion of the two complex tones, and (2) the tendency of the higher tone to be stripped out of the mixture by a competing sequential organization. Fusion was best when the higher and lower complex tones had the same MF, even when the resulting partials did not form part of the same harmonic series. When MF was the same for the higher and lower tones and all partials were harmonically related, the tones fused best when the AM applied to the two tones was in phase. Results are discussed in relation to the problem of perceptual separation of two simultaneous voices and are seen as favoring a theory in which basilar membrane outputs that are amplitude-modulated by the same glottal pulse will be allocated to the same voice.

When two or more persons are speaking at the same time, the auditory system of the listener must group those acoustic components that have arisen from a single voice in order to recognize the linguistic message that it carries. Some part of this job could be done by competing speech-sound "recognizers," each looking for and selecting its particular target properties from the incoming mixture. However, such recognizers would need help in cases where two or more of them could accept a subset of the current set of acoustic inputs, and where the choice as to which recognizer best fit the signal depended upon which components had arisen from the same source. It would be very helpful, then, if the spectral components arising from each separate source could somehow be labeled as such by a method that was independent of the speech-sound-recognition process itself.

Previous research has suggested that the auditory system employs a number of heuristic processes to group those components that arise from the same source (Bregman, 1978, 1981). An illustration of their action can be

seen in an experiment by Bregman and Pinker (1978), whose stimulus is diagramed in Figure 1.

B and C are a pair of pure tones presented synchronously. When heard together they sound like a single rich tone. However, if they are rapidly alternated in a repeating cycle with a pure tone, A, whose frequency is the same as that of B, the sequential pair, A and B, form an auditory stream with a pure-tone quality; C is heard in a separate stream, also with a fairly pure quality. The auditory system's choice of whether to group B with C to form a rich sound or to link B with A to form a pure-tone stream depends upon the frequency separation between A and B: the nearer A is to B in frequency, the stronger is its capturing of B out of the BC complex. The relation between B and C also plays a role. The more synchronous B and C are, in onset and offset, the more strongly they merge into a single rich tone.

Speech is subject to similar principles of perceptual grouping. There is evidence that the auditory system tends to group different formants of the human voice together when they share a common fundamental (Broadbent & Ladefoged, 1957), and particularly when two phonetic recognition processes are competing for the same acoustic material (Darwin, 1981).

The present experiments were concerned with inferring the mechanism by which formants are detected as having a common fundamental frequency. We did not actually use vocal sounds but, rather, a simplified laboratory stimulus similar to the one shown in Figure 1, in which

This research was supported by the Natural Sciences and Engineering Research Council of Canada and the Nuffield Foundation of Great Britain. Valuable ideas were contributed by Campbell Searle, Richard Lyon, Mitch Weintraub, and Roy Patterson.

The mailing address of the first three authors is: Psychology Department, McGill University, 1205 Dr. Penfield Avenue, Montreal, P.Q., Canada H3A 1B1. C. J. Darwin's mailing address is: Department of Psychology, University of Sussex, Brighton BN1 9QL, England.

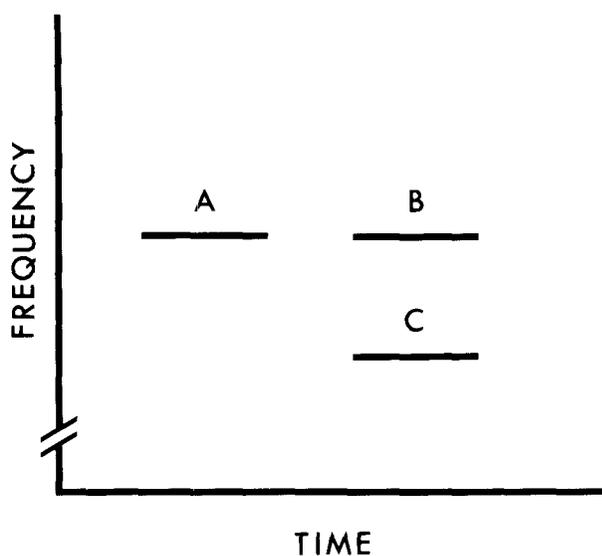


Figure 1. A schematic drawing of a single cycle of a stimulus in which tone A captures tone B out of fusion with tone C. In Bregman and Pinker's experiment (1978), A, B, and C were pure sine tones. In the present experiments, each was a complex tone produced by applying AM to a carrier frequency.

certain acoustic relations could be controlled more exactly. The auditory system recovers the spectral distribution in a speech sound by performing a mechanical frequency analysis in the basilar membrane (BM) of the cochlea. Each point along the BM responds with greatest amplitude to a particular frequency. The neural fibers running off from each point carry specific information about the intensity of the acoustic energy in a particular part of the spectrum. As the incoming energy distribution changes over time, the neural activity will change in a corresponding way to yield what has been called a "neural spectrogram." Our earlier question about how the auditory system can collect the frequency components that have arisen from the same voice now translates into a neural question: How can the auditory system detect which neural fibers coming off the BM are carrying information about the same voice?

One possible approach to this problem is to take advantage of the fact that all the harmonics from a single voice at a certain moment are multiples of the same fundamental frequency. Parsons (1976) has created a computer program to untangle two superimposed voices by exploiting this fact. For each brief period of time, it analyzes the incoming signal into harmonic components and then tries to find two fundamentals that could have generated all the harmonics that are present. The human counterpart of Parsons' program might be the system that detects the pitch of complex tones. This system apparently can detect the fundamental of the tone's harmonics and can give rise to a pitch that is the same as that of the fundamental. Presumably this mechanism could be used for allocating portions of the neural spectrogram to a single voice by determining whether or not these portions could all have been generated by the same fundamental.

As plausible as this method sounds, there is another one that could parse the neural spectrum by looking at quite a different source of information, namely the slower fluctuations in the activity of the neural outputs from various regions of the BM in response to the glottal pulse. Each local region of the basilar membrane can be viewed as a filter that is sensitive to a particular band of frequencies. There is a law in wave mechanics that describes a tradeoff in temporal versus spectral resolution: the less selective a filter is with respect to frequency, the more rapidly it responds (Lyon, 1982; Searle, 1982). Most of the filters implemented by the BM are fairly broadly tuned (Kiang, Watanabe, Thomas, & Clark, 1965), and therefore respond rapidly. As each glottal pulse passes over one of the basilar membrane filters, the neural output of the latter increases, then partially subsides before the next glottal pulse passes over it. As the next glottal pulse reaches it, its rate of firing increases again, then diminishes again, and so on. These bursts of firing, frequency-locked to the glottal pulse, appear in every filter that is broad enough to respond to more than one harmonic of the voice input. Therefore, redundant information about the fundamental resides in the burst activity of a very large number of channels.

The grouping of related frequency regions could use this information. If more than one human voice were driving the total system of filters at the same time, the particular channels that were driven primarily by one particular voice would be "labeled" with the same fundamental period. This fact could be exploited by the auditory system if it had the ability to route those channels with the same "periodicity label" into the same subsequent analysis.

We have, then, two candidates for the segregation of concurrent voices. The first uses the set of harmonics that are present to find some small number of fundamentals that will generate all of them, and then collects those harmonics that could have arisen from the same fundamental. The second method gathers together the set of neural frequency channels that are labeled with the same burst periodicity, regardless of the characteristic frequencies of the neural channels themselves.

To obtain information about which method of analysis, if either, is used by the human auditory system, it is necessary to be able to vary the gross periodicity of a signal independently of the membership of its sinusoidal components in a harmonic series. This cannot easily be achieved with voice signals. However, we can crudely simulate the pulse-shaped signal that a single BM channel picks up by amplitude-modulating a pure tone that is located at the characteristic frequency of that channel.

Our first experiment took advantage of the extensive work on the pitch of three-tone complexes (reviewed in Moore, 1982, chap. 4). If a sinusoid is subjected to amplitude modulation (AM) by a raised cosine, it gives rise to a signal that can be described in two distinct, but equivalent, ways (see Figure 2). In the time domain, it appears as a series of amplitude changes in a carrier sine wave, as in Box 1, whereas in the frequency domain, shown in

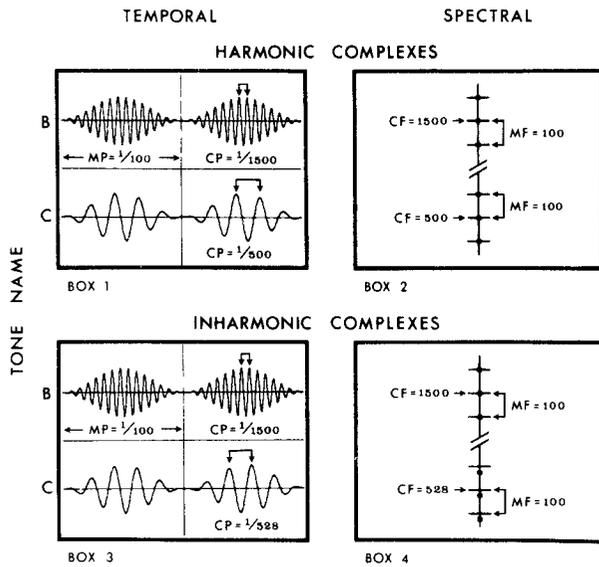


Figure 2. Stimuli of Experiment 1. Each of Boxes 1 to 4 shows a representation of tones B and C. Right-hand boxes, showing frequency components as horizontal lines, are spectral representations of the waveform signals shown in corresponding left-hand boxes. MP = modulation period, CP = carrier period, MF = modulation frequency, CF = carrier frequency.

Box 2, it can be represented as a mixture of three frequencies. The following relation holds between the two representations: the frequency of the carrier tone (CF), seen as the period (CP) of the rapidly varying sinusoid of Box 1, corresponds to the center frequency in the frequency domain representation of Box 2. The frequency of the modulating tone shows up differently in the two representations as well. In the time-domain representation in Box 1, the period (MP) of the modulating tone is seen as the spacing of the large bursts. In the frequency-domain picture, the modulation frequency (MF) becomes the frequency separation of the two flanking tones (sidebands) from the carrier frequency. A final point of considerable importance is that the three frequencies will form part of a harmonic series, whose fundamental is MF, but only when the frequency of the modulator divides evenly into the frequency of the carrier. When this is true, the resulting three-tone complex is "harmonic." An example of a harmonic complex with components 400, 500, and 600 Hz can be obtained by modulating a 500-Hz carrier by a 100-Hz tone (Box 1, lower signal). An inharmonic complex with components 428, 528, and 628 Hz can be obtained by modulating a 528-Hz carrier by a 100-Hz tone (Boxes 3 and 4). In the inharmonic case, the only fundamental common to the three resulting components is 4 Hz, which is below the range heard as a pitch. Although the internal microstructure of the large bursts differs in the cases of inharmonic and harmonic modulation, a crude analysis of the temporal structure would find no difference between a harmonic and an inharmonic complex produced by the same modulation frequency. Both would result in the same periodicity of large bursts (i.e., MP).

In the frequency domain (Box 4), the components of the inharmonic complex are still separated by the 100-Hz MF, but are no longer harmonics of a 100-Hz fundamental. (Locations of harmonics of 100 Hz are indicated by dots on a vertical line in Boxes 2 and 4.) Although a mechanism that looked only at the frequency of large bursts might not discriminate between harmonic and inharmonic complexes, we know that the human pitch-determining system does. The perceived pitch of a harmonic complex is the pitch of the modulating frequency (MF), whereas the pitch of an inharmonic complex is not. One proposed explanation for this fact is that the auditory system tries to find the harmonic series that will best fit the components of the given complex and then outputs the pitch corresponding to the fundamental of this series (Goldstein, 1973). With inharmonic complexes, the fundamental computed in this way will not be the same as the modulation frequency.

The use of harmonic and inharmonic complexes can help us to distinguish between the methods used by the auditory system to group the information arising from different regions of the spectrum. Suppose we were to find that two carriers fused better when amplitude modulated at the same frequency than when modulated at different frequencies, but that the degree of fusion did not depend on whether or not they were harmonic complexes, or even upon whether each of the three-tone complexes was a good fit to the same harmonic series. We could conclude that the grouping mechanism was looking only at the gross periodicity in the neural firing and treating neural outputs that exhibited matching "large-burst" periodicities as arising from the same source.

On the other hand, suppose we created two complexes, one harmonic and one inharmonic, such that although they had different modulation frequencies, they both were heard as having the same pitch. This would mean that, under Goldstein's (1973) assumptions, the same harmonic series was the best fit to both complexes. Suppose, then, that we found that two complexes having the same pitch fused well and those having different pitches fused less well, and that this worked independently of the actual modulation frequency. This would mean that the same mechanism that generates missing fundamentals (or pitch in general) was responsible for the grouping of information from different spectral regions. Following this reasoning, we looked, in Experiment 1, at the fusion of pairs of complexes, matched either on periodicity or on pitch.

The stimulus pattern used in the two experiments resembled that of Bregman and Pinker (1978), described earlier and shown in Figure 1. In the present experiments, each of the three tones, A, B, and C, was a complex tone created through amplitude modulation (AM), in which the amplitude of a sinusoid at a carrier frequency (CF) was modulated by a sinusoid at a modulating frequency (MF). Tones A and B were always identical. The variation in both experiments always concerned the relations between tones B and C. The carrier sinusoids of B and C were always separated in frequency by a ratio of about 3 to 1.

In Experiment 1, two factors were varied orthogonally: (1) whether or not tones B and C had the same MF, and (2) whether or not they had the same pitch. In Experiment 2, the MF applied to B and C was always the same. The experimental variable was whether or not the MF had the same phase in B and C. When the MF was out of phase, however, B's envelope rose whenever C's fell, and vice versa.

The subject's task was to detect tone B in the mixture BC and to make some judgment about the sequence AB, the judgment being different in Experiments 1 and 2. The reasoning was that whenever the grouping of B with A into a sequential pattern was easier, B must have been fusing less strongly with C.

EXPERIMENT 1

Method

Selection of subjects. The two experiments were pretested concurrently. What soon became apparent was that some subjects seemed not to be able to respond to differences among the conditions. Therefore, we decided to perform Experiment 1 with selected subjects. We chose these subjects on the basis of a pretest which determined whether they were sensitive to differences in the *phase* of modulation of B and C, since Experiment 1 did not study the effects of phase. Most subjects did not pass the pretest. However, we continued to use it. We reasoned that although the subjects would not necessarily be representative of the general population, their selection on the basis of the pretest could not bias their performance in the main experiment, because the pretest did not involve differences in *frequency* of AM or in *pitch* of AM signals. We will come back to the issue of selected subjects in the General Discussion.

The pretest consisted of a training procedure in which the subjects had to learn to discriminate out-of-phase ABC patterns (in which the modulation applied to B was out of phase with that applied to C) from in-phase ABC patterns (in which the modulation was in-phase). The pretest had much the same form as the main experiment, except for the stimuli themselves, so the general testing method will be given first.

Test procedure. The task of the subject can be called "adjustment with forced choice." Because of the small size of the differences among the percepts generated by the different experimental conditions, it was necessary to use a very sensitive technique. It basically required the subject to compare pairs of experimental conditions to one another and to judge which one sounded most "decomposed." Each condition consisted of four cycles of the tones A, B, and C in the format shown in Figure 1. The conditions differed only in certain properties of C. Let us refer to any two arbitrarily chosen conditions as Conditions X and Y. The subject was asked to determine whether X or Y showed the most decomposition (i.e., B heard most distinctly). He or she heard a long continuous presentation with the following structure: four cycles of Condition X were presented, then four cycles of Y, then four of X, then four of Y, etc., up to a maximum of 10 occurrences in each condition. The subject was required, during the presentation, to write down which condition, X or Y, showed the greatest decomposition and to rate the size of the difference as "little," "medium," or "clear."

There was one further aspect to the method. We knew that the intensity of A and B (relative to C) made a difference in the fusion. Furthermore, we judged that if the experimenter made an arbitrary choice as to the intensity of A and B (relative to C), he might choose an intensity which, for a certain listener, in a certain comparison, might cause both conditions to sound fused. These considerations led us to allow the subject to adjust the gain of tones A and B (in both conditions equally) relative to C to find a point where one condition sounded more fused than the other.

Table 1
The Eight Stimulus Conditions in Experiment 1

Periodicity	Pitch	
	100	105
100	(1) Harmonics of 100 no shift HARMONIC (400, 500, 600)	(2) Harmonics of 100 shifted up (+28 Hz) INHARMONIC (428, 528, 628)
105	(3) Harmonics of 105 shifted down (-28 Hz) INHARMONIC (392, 497, 602)	(4) Harmonics of 105 no shift HARMONIC (420, 525, 630)
Periodicity	Pitch	
	100	95
100	(5) Harmonics of 100 no shift HARMONIC (400, 500, 600)	(6) Harmonics of 100 shifted down (-28 Hz) INHARMONIC (372, 472, 572)
95	(7) Harmonics of 95 shifted up (+24 Hz) INHARMONIC (404, 499, 594)	(8) Harmonics of 95 no shift HARMONIC (380, 475, 570)

Experimental conditions. Tones A and B were always fixed in nature; they were created by modulating a 1500-Hz sinusoidal carrier by a 100-Hz raised cosine. Tone B, then, always had a periodicity of 100 Hz and a pitch related to a missing fundamental of 100 Hz. The experiment manipulated the properties of tone C. Its design can be viewed as $2 \times 2 \times 2$, which breaks down into an upper four-cell table and a lower one (as shown in Table 1). The cells are numbered from 1 to 8 in their upper left-hand corners. Conditions 1 to 4 constitute an orthogonal variation of the pitch and periodicity of tone C.

In Condition 1, C was created from a CF of 500 Hz and a MF of 100 Hz. This generated frequencies at 400, 500, and 600 Hz, as listed in parentheses in the cell. Tone C of Condition 4 matched B neither on pitch nor on periodicity. It was created by CF = 525 and MF = 105, and consisted of the fourth, fifth, and sixth harmonics of 105 Hz. Its pitch was 105 Hz, and so was its periodicity.

Condition 2 had a periodicity of 100 Hz but a pitch that matched that of a 105-Hz tone. It was created by modulating a 528-Hz carrier at 100 Hz. Its component frequencies are given at the right. They constitute an inharmonic complex in that they have no common fundamental in the audio range. They can be considered as being the fourth, fifth, and sixth harmonics of 100 Hz, all shifted up by 28 Hz. This way of describing them is based on the work of Schouten, Ritsma, and Cardozo (1962), who showed that, with a fixed MF, there was a regular rise in pitch as a CF was raised through a series of values that generated inharmonic complexes. (In all inharmonic conditions in our experiment, we matched the pitch empirically to that of the harmonic condition with which it was equated.)

Condition 3 created, as tone C, an inharmonic complex that matched B in pitch but not in periodicity, using CF = 497 and MF = 105. It can be viewed as consisting of the harmonics of 105, all shifted down by 28 Hz.

Conditions 5 to 8 involved C tones that were matches and mismatches to B's periodicity and pitch exactly as in Conditions 1 to 4; but in Conditions 5 to 8, the mismatches were created by moving the pitch and periodicity down to 95 Hz rather than up to 105 Hz. They were a control against the possibility that simply moving the harmonics of C closer to those of B might have an effect over and above the effect of the mismatch.

Stimulus generation and presentation. The stimuli were created by digital synthesis using the MITSYN system (Henke, 1976) run-

ning on a PDP-11/34 computer, output through a 12-bit D/A converter at 12,032 samples/sec and low-pass filtered at 5 kHz by a Rockland Model 852 filter that had a roll-off of 48 dB/octave. It was recorded on a Sony TC-6544 tape deck at 7.5 i.p.s., tones A and B on channel 1 and tone C on channel 2. At the time of testing, channel 1 was passed through the preamplifier stage of an integrated amplifier whose gain could be controlled by the subject before it was mixed with channel 2. The two channels were then mixed in a Marantz 3800 preamplifier and amplified by a Marantz 170DC stereo power amplifier. The mixed signal was then presented binaurally over Sennheiser HD-414 headphones after being low-pass filtered at 8 kHz to reduce tape hiss. The intensity of tone C measured alone was set to 64 dB SPL for the first six subjects and to 74 dB SPL for the remainder.

There was a certain amount of distortion created by the apparatus: tape deck, up to 1.2%; subject-controlled amplifier, up to .05% THD and .05% IMD; final preamplifier, up to .02% THD and .01% IMD; power amplifier, up to .03% THD, up to .03% IMD. For this reason, Experiment 1A was added to replicate some of the conditions of Experiment 1 using equipment with lower distortion. The listeners were tested individually in an IAC 1202 test chamber and were run in two 1-h sessions. There were 64 trials in which all possible permutations of the eight conditions occurred as X-Y pairs, including each condition with itself.

Trial structure. A long burst of noise at a comfortable level announced a new trial. After a 1-sec silence, there were 10 cycles of alternation of Conditions X and Y. Each cycle had the following structure: A short burst of noise announced X. After a 1-sec silence, Condition X was repeated four times with a 25-msec gap between repetitions; after another 1-sec silence, Condition Y was repeated four times, also with a 25-msec silence between repetitions. Then a 1.5-sec silence preceded the short burst of noise that announced X again. After the 10 alternations of X and Y, there was a 4-sec silence preceding the long noise burst that announced the next trial.

Stimulus structure. All tones were 230 msec in duration, including 20-msec rise and 10-msec fall times. The onset envelope was shaped by multiplying it by the square of a linear function that rose from 0 to 1 in 20 msec, and the offset envelope was shaped by the same function falling from 1 to 0 in 10 msec. The time course of one cycle was as follows: 10-msec silence, 230 msec of A, 20-msec silence, 230 msec of B with C, 260-msec silence. For each tone, the carrier always began at 0° phase angle and the modulator began at the phase angle that produced zero amplitude (in the main experiment and pretest) or at the angle that produced maximum amplitude (used only in the pretests for out-of-phase stimuli).

Subjects. Our subjects were 24 adults selected from a population of university students and staff. One of the present authors (A.S.B.) was included. Only about one in three subjects who were given the pretest showed a sufficient ability to discriminate the degree of decomposition obtained in ABC cycles with in-phase modulation of B and C from the degree of decomposition obtained with out-of-phase modulation.

Results

Each condition was compared with each other condition and to itself by subjects. They were asked to choose Condition X or Condition Y and to rate its superiority of decomposition on a 3-point scale. This yielded a score of -3 to +3 for each condition on each trial (the plus sign indicating choice of that condition, the negative sign indicating the choice of its competitor). The sum of these scores for any condition reflects the overall superiority of its "decomposition" as compared with the other conditions.

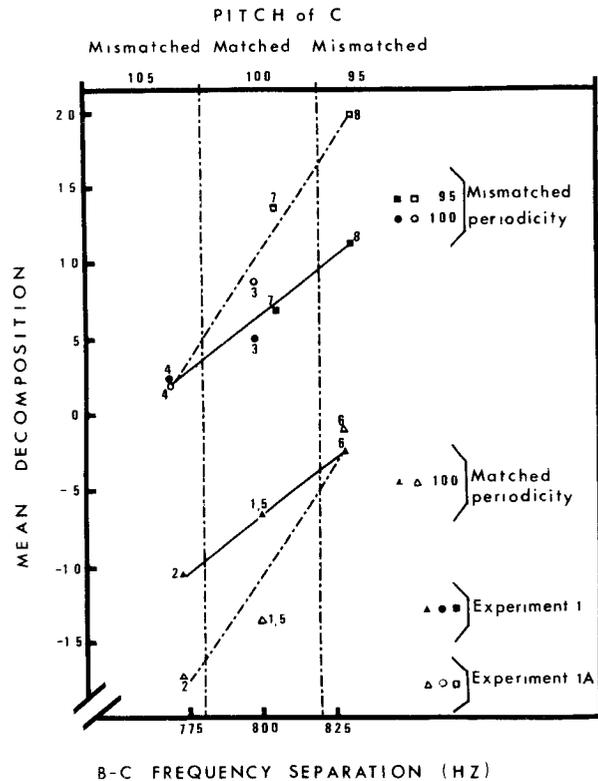


Figure 3. Results of Experiments 1 and 1A. The ordinate shows the mean of decomposition scores for each condition. There are two labelings of the abscissa. The top one shows three categories of the pitch of C and whether or not it is matched to B's pitch. The lower abscissa represents the frequency separation between the lowest component of B and the highest one of C. The parameter is the periodicity of B. The condition number for each point is given above it. The closed symbols show the results for Experiment 1, and the open symbols those for Experiment 1A.

The filled symbols of Figure 3 show the results for the eight conditions of Experiment 1. Mean decomposition scores are shown as a function of pitch (upper abscissa) and frequency separation between B and C (lower abscissa). The parameter is periodicity. The condition number is given beside each data point. The results were analyzed statistically by planned contrasts. Conditions in which C's periodicity matched that of B (i.e., 100 Hz) are shown with triangles. Those in which the periodicities did not match are shown with squares for 95-Hz periodicity of C and circles for 105 Hz. Overall, a periodicity mismatch produced a clearly higher decomposition score [$F(1,22) = 49.3, p < .001$].

The effects of pitch are unclear. If the pitch match of B and C enhanced fusion, mistuning C to either 95 or 105 Hz should have increased the decomposition. Instead, there was a dual effect. Moving pitch down to 95 Hz (seen by contrasting Conditions 6 and 8 with Conditions 5 and 7) significantly increased decomposition [$F(1,22) = 8.51, p < .01$]. However, moving it up to 105 Hz (i.e., contrasting Conditions 2 and 4 with Conditions 1 and 3) ac-

tually *decreased* decomposition (i.e., promoted fusion) [$F(1,22) = 6.05, p < .025$]. This contradiction can be resolved by attributing the "pitch" effects not to the match or mismatch between the pitch of B and C, but simply to how far the harmonics of C were from the harmonics of B in frequency. The lower abscissa shows the frequency separation between the highest harmonic in C and the lowest one in B for the plotted condition. This variable can be seen to fully account for all the effects in the figure except for the quite independent effect of periodicity. The results suggest that as two complex tones are moved closer together spectrally, they tend toward greater fusion.

EXPERIMENT 1A

As a result of spectral analyses performed after Experiment 1 was finished, we judged that the level of intermodulation distortion was not acceptable. Experiment 1A was a replication of Experiment 1 with low-distortion equipment. There were some modifications in the procedure, which will be described.

Method

Subjects and Training. Only five subjects (adults, aged 20 to 23 years, four males and one female) were used in this study. However, because of an improved training procedure, none had to be discarded for failing the pretest.

The training involved explanations of the task with diagrams and gave a careful description of "decomposition" that described it as hearing the "high tone" repeat twice as often as the "low tone." "Fusion" was described as hearing an alternation of the "high" and "low" tones. A careful description was also given of experiences lying between full fusion and full decomposition. Then examples were played and the subject was allowed to increase and decrease the intensity of tones A and B as a pair so as to move B into or out of fusion with C. In addition, the full nature and purpose of the "adjustment with forced choice" procedure was explained to the subjects.

Procedure. The apparatus used in this experiment did not permit the subject to advance to the next trial at will after recording his or her decision. Instead, the pattern of four cycles of Condition X and four cycles of Condition Y repeated exactly eight times on each trial. The pattern of signals to the subject announcing the onsets of new trials and new conditions within a trial were somewhat modified from that of Experiment 1, and the subject was told to adjust the intensity of the A and B tones on a trial only after hearing both of the conditions that were to be compared. The pretest was the same as that of Experiment 1. With the previously noted exceptions, the main testing was the same as in Experiment 1.

Conditions. Because we tested fewer conditions, each subject received two different randomizations of 56 conditions instead of the single randomization of 64 trials used in Experiment 1. The reduction was achieved by eliminating the eight comparisons of conditions with themselves. Tone C was presented at 74 dBA as in the latter part of Experiment 1. Three randomizations of the orders of conditions were created, with each subject receiving a selection of two of them. Each subject required about 6.5 h of testing in total, broken into four sessions.

Equipment. Because the audio distortion found in Experiment 1 came mainly from the tape recorder, no tape recorder was used in the replication. Instead, the signal was fed from the computer's filtered D/A output directly into the testing chamber. The same amplification system was used as in Experiment 1. The signal was syn-

thesized at a higher sampling rate (20,000 samples/sec) and low-pass filtered at a lower cut off frequency (2300 Hz) using the same equipment as in Experiment 1.

Stimuli. Stimuli had the same durations as in Experiment 1 and the same rise and fall times, but the onsets and offsets were gated by a different function—a raised quarter sine. The time course of one cycle was slightly altered to increase the time between cycles. 230 msec of the capter, 20 msec silence, 230 msec of tones B and C together and 1.02 sec silence.

Results

The results are shown as the open symbols in Figure 3. They are very similar to those of Experiment 1. The only difference is that the effects are larger. There is a larger effect of the frequency separation between tones B and C. The effect of periodicity is also larger. The increase in the size of these effects is probably the result of using fewer, better trained subjects. Finally, as in Experiment 1, there is no main effect of the mismatch of the pitch of tones B and C that cannot be accounted for by the frequency separation of their partials. We conclude that the results of Experiment 1 were not due to distortion artifacts.

Discussion of Experiments 1 and 1A

The results of Experiments 1 and 1A favor the conclusion that a periodicity match between two spectrally separated signals promotes their fusion, and that it is unimportant whether the frequency components of the two signals fall into the same harmonic series. For example, the best fusing condition of all was Condition 2, in which, although C matches B in periodicity, it belongs to an inharmonic series. The inharmonic C conditions, 2, 3, 6, and 7, yield exactly those levels of decomposition that would be predicted by their periodicity matches and spectral separations from B (as shown by the regular trends in Figure 3).

EXPERIMENT 2

The second experiment used the same pattern of stimuli (as shown in Figure 1) but asked whether tones B and C would fuse better when their modulation envelopes were in phase, that is, C's envelope rising in synchrony with B's.

The method of this experiment differed from that of the previous two. In Experiment 1, the decomposition of B from C was measured by having listeners rate it (in relation to the decomposition of a comparison condition). Rating scales are subject to bias. Although we believe that the subjects were really rating decomposition, it is desirable to be assured, whenever the nature of the question permits it, that they are not actually responding to some other factor generated by the mismatch of the periodicities of B and C. Fortunately, the simpler nature of the question in Experiment 2, and the employment of simpler stimuli, allowed us to design an "accuracy" task. The rationale for this task is as follows: If B truly segregates from C and enters into a sequential stream with

A, then the properties of the sequence A-B should be easier to judge. In particular, it should be easier to detect whether the A-B sequence is going up or going down in frequency. As in Experiment 1, we discovered after the experiment was completed that there was an unacceptable level of intermodulation distortion generated by the analog tape recorder, and so Experiment 2A was performed using a digital tape recorder.

Method

Task. The subjects were told that on each trial they would hear a repeating cycle formed of a single high tone (A) followed by two simultaneous tones (B and C) and that C would be louder than B. On any given trial, tone B could be "higher" or "lower" than tone A. On each trial, the subjects would hear a number of cycles of the stimulus for that condition and then were to decide whether B was "higher" or "lower" and to rate their confidence on a 3-point scale (strong, medium, weak).

Stimuli. There were basically two variables in the experiment: (1) the phase agreement of the AM of tone C with tone B, and (2) the carrier frequency (CF) of tone B, either higher or lower than the CF of tone A. The MF was set at 100 Hz for all tones. Tones A and C were fixed in value throughout the experiments; tone A consisted of CF = 1500 Hz, MF = 100 Hz; tone C had CF = 500 Hz, MF = 100 Hz.

Since the 100-Hz modulation frequency of C divides evenly into its carrier frequency, the resulting frequency components of C (400, 500, 600 Hz) are harmonically related to a fundamental at 100 Hz. We decided to make this harmonic relationship be true also of the components of tone B, whether B was "high" or "low." Otherwise, the high and low versions of B might differ in their harmonic relations to C, enabling the listeners to discriminate "high" from "low" on the basis of this factor. For this reason, the center frequencies of "high B" and "low B" were both divisible by 100 Hz. When B was high, its carrier frequency was 1600 Hz (generating harmonics 14, 15, and 16 of 100 Hz), and when it was low, its carrier was 1400 Hz (generating harmonics 13, 14, and 15 of 100 Hz). Since "high B" and "low B" had the same fundamental, they did not differ in pitch, but rather in "sharpness" or "highness."

In pretesting the stimuli, we noticed that the high version of B was more decomposable from C than the low version was. Presumably this was due to the same factor that caused the effect of B-C frequency separation in Experiment 1. In an attempt to cancel out this difference, the intensity of the low version was set 1 dB greater than the high version. The intensity of tone A and high tone B was 56 dB SPL.

In addition to the high/low factor, we varied the phase match of the AM in tones B and C. In both tones, the carrier began at 0° phase angle. In the in-phase AM condition, the modulating wave started at a value of zero amplitude in both tone B and tone C. In the out-of-phase AM condition, the modulating wave in B started at maximum amplitude (180° shift with respect to the AM of tone C). In order that B not have a sudden onset in the out-of-phase conditions that would produce an onset asynchrony between B and C, a factor known to increase segregation (Bregman & Pinker, 1978), all tones received a further 30-msec shaping of their onsets and offsets using the same envelope function employed in Experiment 1. Three different intensity levels of tone C were employed: 70, 72, or 74 dB SPL.

Each cycle of the ABC stimulus was timed as follows: tone A for 250 msec (including rise/fall times), silence 250 msec, tone B for 250 msec, intercycle silence 500 msec. On each trial, the following sequence of events was presented: a rapid sequence of three tones as a warning signal, a silence of 4 sec, 10 repetitions of the ABC pattern (1.25 sec per cycle). There was a 5-sec pause between trials for the subject's response.

Procedure. After having the task explained to them, the subjects were trained on sample conditions in which tone B was always modulated out of phase with respect to C, and in which C was only 3 dB louder than B. With this stimulus, it was generally easy to isolate B. Subjects were trained until they could successfully identify tone B as being "higher" or "lower" than tone A.

Design. There were three factors in the experiment: (1) the carrier frequency of B (two levels), (2) the AM phase of B (two levels), and (3) the intensity of C (three levels), yielding 12 conditions. There were four replications of each condition, yielding 48 trials. The three different intensity levels were given in separate blocks of trials. These levels were administered in increasing order of difficulty, that is 70, then 72, then 74 dB of tone C. There were six randomizations of the order of conditions within blocks, and these were recorded on separate tapes. There were 5 subjects assigned to each of five tapes and 8 subjects to the sixth tape, for a total of 33 subjects.

Apparatus. The stimuli were digitally synthesized using the same system as in Experiment 1, but at a somewhat higher sampling rate (16,393 samples/sec), low-pass filtered at 7 kHz using the "flat delay" (linear phase) setting of the filter, and then tape recorded. For one-third of the subjects the tape recorder was a Marantz SD-9000 cassette recorder using a 3.75 i.p.s. speed. The testing of the remaining subjects employed the tape deck used for Experiment 1. The signals were played back binaurally to subjects over Sennheiser HD-414 headphones. The subjects were tested individually in an audiometric test chamber.

Subjects. These were initially 35 adult males and females, between 18 and 37 years of age, recruited from a university population. There was no screening test but the data for 2 subjects of the 35 were dropped; one failed to complete the experiment and the intensity levels were not set correctly for the second.

Results

On each trial, the subject's response was a choice of "high" or "low" accompanied by a confidence rating on a 3-point scale. This can be interpreted as a 6-point scale of rated "highness" running from -3 ("low B," very confident) to +3 ("high B," very confident). These scores were then used to calculate a D score for each subject in each condition of C intensity and B-C phase relations (six D scores per subject). D is an easily calculated nonparametric measure, representing the degree to which subjects could discriminate "high B" from "low B" (Bregman, 1978; Bregman & Campbell, 1971). The values of D can run from -1.0 (total reversal of correct discrimination) to +1.0 (perfect discrimination), with 0 indicating random performance. The D score makes no assumption that subjects' response bias or variability remains constant from condition to condition. A plot of the D scores for the present experiment is given in Figure 4.

The main question was whether the phase relation between the AM of tones B and C would affect the fusion of these tones. Figure 4 shows that when the AM was 180° out of phase, B was more discriminable [$F(1,27) = 6.4, p < .02$]. This happened at all three intensity levels of tone C.

The apparent difference between the discriminability of tone B at different intensities of C is not statistically significant. This was probably due to the small range of intensities used and because the intensity of C increased systematically across the three blocks of trials. The effects of practice in detecting B probably minimized the

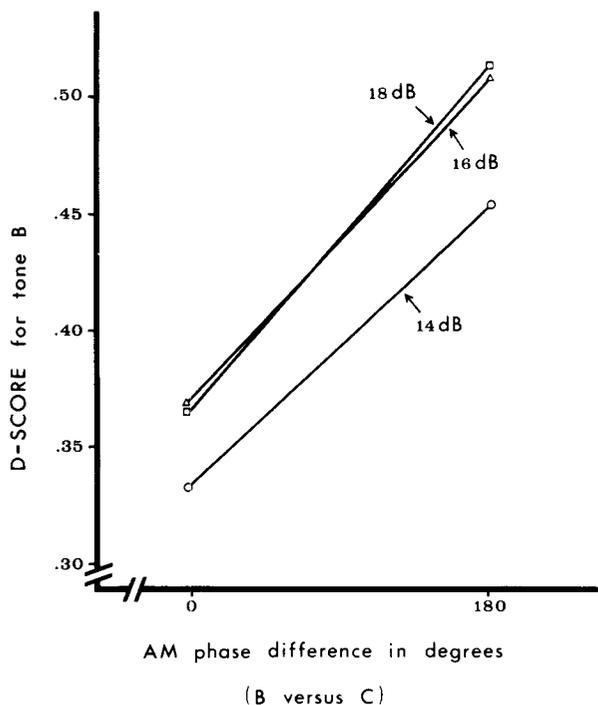


Figure 4. Experiment 2: D score (accuracy of discrimination of "high B") as a function of the difference in the phase of AM in B and C. The parameter is the intensity of tone C relative to tone B.

effects of increasing C's intensity. However, since the only reason for including variations in C's intensity was to try to obtain sufficient variation in B's audibility, the nonsignificance of the intensity effect is unimportant. The only noteworthy point is that the intensity of C did not interact with the phase-of-modulation effect. However, there was a significant interaction between the effects of the intensity of C and the group of subjects (each group of subjects heard a different randomized tape) [$F(10,54) = 2.4, p < .02$]. Since the groups were run consecutively, and since the intensities were always given in the same order, and the tapes contained different orders of conditions, it is not clear what the actual source of this interaction is. In any event, there were no significant interactions of this effect with the phase effect.

There is evidence in the data that we did not adequately equalize the discriminability of high and low Bs when we boosted the low B by 1 dB. If one plots the raw "highness" score assigned to B independently of its correctness, as in Figure 5, several things may be seen. First of all, the high Bs (top line) were rated higher than the low Bs (bottom line) overall. This simply means that the two were discriminable [$F(1,27) = 356.1, p < .0001$]. Second, although high B got a higher rating at the 180° phase conditions than at 0°, low B did not get a lower one. We believe that this arose because the subjects could not really hear the low B well in any condition, and tended to say "low" when they were uncertain. If more uncertain in the 0° condition, they would therefore say "low" more often rather than less often. This would tend to neu-

tralize any effects of phase in the "low-B" condition. Figure 5 confirms the existence of a general bias toward saying "low." "Low-B" responses were coded as negative, and "high-B" responses, as positive. Therefore, if the subjects were unbiased, the line representing the average of responses to all stimuli (middle line) should not be significantly below the zero point; but it is [$F(1,27) = 11.76, p < .01$]. However, the negative bias appears to be somewhat less in the 180° phase condition than in the 0° condition [$F(1,27) = 3.9, p < .06$].

EXPERIMENT 2A

As a result of spectral analyses of the audio tapes performed after Experiment 2 was finished, we judged that the level of intermodulation distortion was not acceptable. Accordingly, the present study replicated it with low-distortion equipment. We also masked the onsets and offsets of tones B and C with white noise to eliminate a possible artifact. There were some additional modifications in the procedure which will be described.

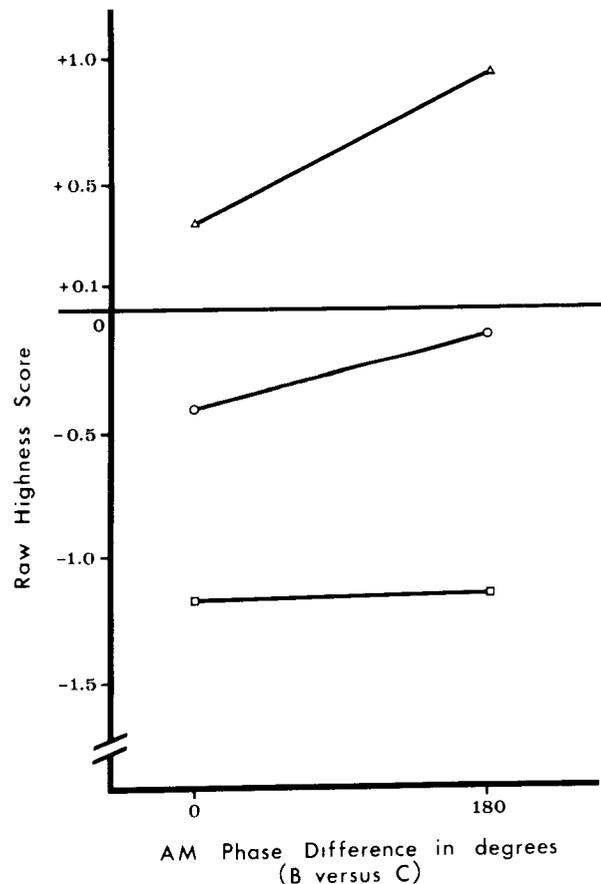


Figure 5. Experiment 2: Mean "highness" scores assigned to tone B as a function of the phase difference of the AM of tone B versus that of tone C. Shown for high B tones (top line) and low B tones (bottom line) and the mean of high and low B tones (middle line).

Method

Task. The task was the same as in Experiment 2.

Design. In Experiment 2, we used three intensities of tone C in order to maximize our chances of finding at least one level of difficulty of the task in which the performance level was neither too high nor too low to permit the effects of phase to be detectable. Experiment 2A used a more efficient strategy. Two levels of difficulty of the task were created by changing the intensity of the target tone (B). Each subject was first tested on the easier task. If the obtained performance fell below 60%, that subject's data were not used and the subject was dropped from the experiment. If it fell between 60% and 85% correct, the data were used. If it fell above 85%, the data were not used, but the subject came back for a second session of testing on the harder task. If the 2nd day's data fell between 60% and 85% correct, they were used; otherwise they were discarded.

Apparatus. Instead of the analog tape recorder used in Experiment 2, the present experiment employed a digital tape recorder, consisting of a Sony digital audio processor, Model PCM 701-ES, and a Sony Betamax video cassette recorder, Model SL-2401. Recording was done on Sony Dynamicron L-250 HG recording tape, using the 14-bit resolution setting on the digital processor and the Beta II speed on the recorder. At these settings, the rated total harmonic distortion is .007%, and intermodulation distortion is insignificant. At the time of playback, the output gain of the audio processor was set to 0 dB for the first six subjects. This was changed to a value of -6 dB for the remainder, to guard against distortion. The same amplifiers were employed as in Experiment 1, except that the mixing of tones B and C was done at the time of synthesis rather than by analog equipment at the time of playback. An attempt was made to employ a level of gain that minimized distortion at each stage of amplification. The subjects heard the stimuli over Telephonics TDH-49P headphones.

Stimuli. The stimuli had the same general design and were presented to the subjects in a pattern similar to the one used in Experiment 2. However, certain modifications were made. One involved the onsets and offsets of the stimuli. Recall that, under the 180° phase condition, the modulation of tones B and C starts in opposite phase. In effect, this means that the tones reach their first peak amplitude at slightly different times (separated by 5 msec). We were concerned that this onset asynchrony might promote segregation independently of the ongoing phase difference between tones B and C, since previous research had implicated onset and offset asynchrony as a factor in segregation (Bregman & Pinker, 1978; Dannenbring & Bregman, 1978). Accordingly, we masked the onsets and offsets of the tones with white-noise bursts. Each tone was of 250 msec duration, including 20-msec rise/fall times gated by a raised quarter sine function. However, during the 20-msec rise/fall periods, 83-dBA noise bursts were also presented. Tone C was always presented at 74 dBA, the highest level employed in Experiment 2.

Again, as in Experiment 2, tone B had a different intensity, depending on whether it was higher or lower in frequency than tone A, since otherwise "high B" would be more discriminable. The intensities were set on the basis of "equal-clarity" judgments by two observers. In addition, the task was presented at two levels of difficulty by changing the intensity of tone B. In the numbers that follow, the intensity of the tones is expressed relative to the fixed 74-dBA intensity of C. For the "easy" task, high tone B was -27 dBA and low tone B was -23 dBA. For the "hard" task, high tone B was -31 dBA and low tone B was -26 dBA. In general, these levels of tone B were lower than in Experiment 2. Tone A was always set at -27 dBA. The much decreased intensities of the A and B tones relative to C in this experiment were compensated for by using a more extensive training procedure with the subjects.

The cycle of stimuli on each trial was the same as in Experiment 2 except that the interval between the warning signal and the onset of the cycles was decreased to 2 sec and the subject's response time at the end of the trial was decreased to 4 sec.

Training. In the training, subjects received many exposures to each of the stimuli used in the experiment, both singly and in combination, accompanied by verbal descriptions and diagrams. Then the actual pattern used in the experiment was presented, except that in this practice period (but not in the experiment) they could control the intensity of tone C and explore the sound of the B-C mixture at various degrees of clarity of B. Finally, the subjects were given eight trials under conditions identical to those used in the experiment.

Subjects. A total of 31 young adults (13 males and 18 females) between the ages of 18 and 30 were tested. The results of 11 subjects were discarded, 8 who exceeded the upper cutoff performance, 2 who did not reach the lower one, and 1 tested on the wrong loudness levels. Therefore, we were left with the results of 20 subjects.

Results

The D scores reflect the accuracy of deciding whether tone B was high or low. The means and standard deviations for the AM phase conditions were as follows: 0° phase, $M = .54$ ($SD = .07$); 180° phase condition, $M = .72$ ($SD = .08$). Although the D scores were higher than in Experiment 2, the difference between the in-phase and out-of-phase conditions was larger. This difference was statistically significant [$t(19) = 3.2$, $p < .01$].

As in Experiment 2, we plotted the raw "highness" score as a function of AM phase difference, and saw a more symmetrical pattern than the one shown in Figure 5. This time the pattern showed that not only did high B get a higher rating in the 180° condition (1.68) than in the 0° condition (.66), but also low B got a somewhat lower one in the 180° condition (-1.45) than in the 0° condition (-1.3). Although these results are not totally symmetrical, impressionistically one could say that a little under half the asymmetry has been removed. In our discussion of Figure 5, we attributed the asymmetrical pattern to a poor discriminability of low B relative to high B, accompanied by a bias toward saying "low" when unsure. However, in preparing Experiment 2A, we took pains to make high and low B equally discriminable. Although there was again a slight bias toward the "low-B" response in this experiment, it was about half of what it was in Experiment 2.

We conclude from this replication that the original results of Experiment 2 were not due to an artifact arising either from distortion of the signal or from asynchronies in the onset of AM, but were the result of ongoing phase relations between tones B and C.

Discussion of Experiments 2 and 2A

The results showed that the frequency of the B complex tone was more discriminable when its AM was out of phase with C's. However, this effect was somewhat complicated by a general bias toward the "low-B" response. Prior to the experiment, our own personal listening experience had suggested that "low B" was not very audible, and this had led us to boost its intensity relative to "high B," but apparently the 1-dB boost in Experiment 2 was not sufficient and the 4-dB boost in Experiment 2A was just barely adequate. It seems that the proximity of the center frequencies of B and C is an im-

portant factor to consider in the design of experiments on the fusion of modulated tones.

GENERAL DISCUSSION

Weakness of Observed Effects

These experiments have provided some evidence that the information in different frequency regions tends to be fused whenever the modulation rate matches in frequency and in phase. The effects that we observed were, however, quite small, especially in Experiments 1 and 2. Not every subject showed them: In Experiment 1, which concerned AM frequency, about two out of three subjects failed to pass a pretest in which they were required to discriminate stimuli in which the AM of tones B and C was in-phase versus out-of-phase. We must consider whether the rejected subjects never fused signals by periodicity in their interactions with the world or whether they were simply less sensitive to the particular periodicity manipulation used in this experiment. We can imagine a number of reasons that support the second interpretation, all of which suggest that our manipulation produced a less-than-optimal stimulus for the registration of periodicity in the auditory system.

First of all, we know that the effects observed in both Experiments 1 and 2 were susceptible to being strengthened by the more careful training of subjects in Experiments 1A and 2A; so there is no doubt that the effects are real. But training subjects does not make an effect stronger, only more detectable. One reason for the weak effects may be the choice of 500 and 1500 Hz as carrier frequencies. With an amplitude-modulated signal, the auditory system registers the modulation in the time domain only to the extent that it fails to resolve the separate frequency components that it is composed of. In the 500-Hz region (tone C), the ratios between adjacent frequency components are 6:5 and 5:4, values that represent a spacing of about one critical band. Such wide spacing would be expected to lead to a partial resolution of the component frequencies of tone C, with a concomitant drop in the registered AM of the tone. The possibility exists that even the 1500-Hz region might show some resolution of harmonics and therefore less-than-maximum registration of periodicity.

Despite the limited periodicity information, we chose to place tones B and C in the 1500- and 500-Hz regions, because these regions are the approximate positions of F2 and F1 in speech signals. We felt that if we could not show an effect of periodicity in these regions, the relation of our findings to speech perception would be rather tenuous.

How is it, then, that if the AM with 500- and 1500-Hz carriers is not strongly registered, people appear, nonetheless, to be able to integrate formants by fundamental frequency? This may be due to the fact that when we hear a normal signal, the same periodicity is being registered in a very large number of BM frequency channels; there is a widespread and redundant encoding of the periodicity. F1 and F2 span a wider frequency range than our B and C complexes do and are actually only peaks in a con-

tinuous spectrum. It may be that the continuity of this spectrum, registering the same periodicity throughout, helps to tie F1 and F2 together. Our signal, on the other hand, imposed the periodicity in only two regions.

A relevant finding has been discussed by Moore (1982, p. 133): When the residue pitch of a signal is raised by increasing the repetition rate (periodicity) of a waveform, pitch changes are obtained only up to a certain repetition rate. However, he reports, higher repetition rates will produce further increases in pitch if a greater number of harmonics is employed. This supports the idea that there is integration of computed periodicity across the spectrum, and that it is more strongly registered when more frequency channels are stimulated.

The most probable reason, however, for the weak effect in Experiment 1 is that the size of the AM frequency mismatch in that experiment was only about 5%. We chose such a small value because Brox and Noteboom (1982) had reported research on the intelligibility of speech masked by speech and had found that, with synthetic voices, differences in the fundamental frequency (f_0) led to improved segregation of the voices and an increase in the intelligibility of the speech. Their data showed that an effect of a different f_0 was visible in the data with differences as small as .5 semitones. In designing Experiment 1, we reasoned that unless we could show an effect at comparably small mistunings of f_0 (in our case, MF), we could not argue that we were investigating the same mechanism as the one involved in the segregation of speech signals. Subsequent unpublished research in our laboratory that did not test the same hypotheses as the studies we are reporting here, but studied the effects of the periodicity mismatch between tones B and C employing wider variations in the mismatch, yielded much more robust results.

Implications of the Results

Despite the reasons for believing that we produced a less-than-optimal simulation of the burst periodicity found in natural vocal signals, we were able to detect a spectral grouping based on periodicity. What does this mean for the existing findings that formants sharing an f_0 tend to group together and to become segregated from formants sharing a different f_0 ? Brox and Noteboom (1982) attributed their finding of voice separation by f_0 to a pitch-detection mechanism of the type proposed by Gerson and Goldstein (1978) and Goldstein (1973). We obtained two results opposed to this hypothesis. The first was that the match of the pitches of tones B and C in Experiments 1 and 1A had no consistent effects upon their perceptual integration. If B and C had the same pitches, then presumably this arose from their having been accepted by the same harmonic template. Yet this did not cause the tones to be fused. However, when B and C had the same periodicity, they tended to fuse (despite mismatches in pitch and in harmonic series).

Experiments 2 and 2A gave additional evidence favoring the existence of a pure periodicity mechanism. The mismatch of phase of AM reduced the fusion of tones B

and C. Since the AM phase does not affect which harmonics are present, a harmonic template analyzer should be blind to phase.

The present experiments do not altogether rule out a contribution from a harmonic- or pitch-matching mechanism. Experiments 1 and 1A might simply have failed to find one due to some limitation in the method. For example, for a 100-Hz fundamental, a 1500-Hz carrier would be on the edge of the "existence region" for residue pitch (Ritsma, 1962). However, if this were the reason for the absence of a pitch effect, it should be expected to apply also to speech sounds where the higher formants would be outside the existence region for their fundamental. Whether or not there is a true effect of pitch that was missed by Experiment 1, the two experiments do show a clear contribution of a periodicity-matching mechanism to fusion, a contribution that operates quite independently of the relation of the harmonics to a common fundamental frequency.

REFERENCES

- BREGMAN, A. S. (1978). The formation of auditory streams. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, NJ: Erlbaum.
- BREGMAN, A. S. (1981). Asking the "what for" question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, NJ: Erlbaum.
- BREGMAN, A. S., & CAMPBELL, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, **89**, 244-249.
- BREGMAN, A. S., & PINKER, S. (1978). Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, **32**, 19-31.
- BROADBENT, D. E., & LADEFOGED, P. (1957). On the fusion of sounds reaching different sense organs. *Journal of the Acoustical Society of America*, **29**, 708-710.
- BROKX, J. P. L., & NOTEBOOM, S. G. (1982). Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics*, **10**, 23-36.
- DANNENBRING, G. L., & BREGMAN, A. S. (1978). Streaming vs. fusion of sinusoidal components of complex waves. *Perception & Psychophysics*, **24**, 369-376.
- DARWIN, C. J. (1981). Perceptual grouping of speech components differing in fundamental frequency and onset-time. *Quarterly Journal of Experimental Psychology*, **33A**, 185-207.
- GERSON, A., & GOLDSTEIN, J. L. (1978). Evidence for a general template in central optimal processing for pitch of complex tones. *Journal of the Acoustical Society of America*, **34**, 1418-1424.
- GOLDSTEIN, J. L. (1973). An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America*, **54**, 1496-1516.
- HENKE, W. L. (1976). *MITSYN: An interactive dialogue language for time signal processing*. Cambridge, MA: M.I.T., Research Laboratory of Electronics.
- KIANG, N. Y.-S., WATANABE, T., THOMAS, E. C., & CLARK, L. F. (1965). *Discharge patterns of single fibers in the cat's auditory nerve*. Cambridge, MA: M.I.T. Press.
- LYON, R. F. (1982). A computational model of filtering, detection, and compression in the cochlea. *International Conference on Acoustics Speech and Signal Processing. May 1982*. Institute of Electrical and Electronics Engineers.
- MOORE, B. C. J. (1982). *Introduction to the psychology of hearing* (2nd ed.). New York: Academic Press.
- PARSONS, T. W. (1976). Separation of speech from interfering speech by means of harmonic selection. *Journal of the Acoustical Society of America*, **60**, 911-918.
- RITSMAN, R. J. (1962). Existence region of the tonal residue. I. *Journal of the Acoustical Society of America*, **34**, 1224-1229.
- SCHOUTEN, J. F., RITSMAN, R. J., & CARDOZO, B. L. (1962). Pitch of the residue. *Journal of the Acoustical Society of America*, **34**, 1418-1424.
- SEARLE, C. L. (1982). Speech perception from an auditory and visual viewpoint. *Canadian Journal of Psychology*, **36**, 402-409.

(Manuscript received August 31, 1983;
revision accepted for publication March 21, 1985.)

26th ANNUAL MEETING OF THE PSYCHONOMIC SOCIETY, INC.

The 26th Annual Meeting of The Psychonomic Society, Inc., will be held in Boston, at the Boston Park Plaza Hotel, on Friday through Sunday, November 22-24, 1985. The program and hotel reservation cards will be mailed to members of the Society in September. A copy of the program will be published in the *Bulletin of the Psychonomic Society*.