

## Fusion of simultaneous tonal glides: The role of parallelness and simple frequency relations

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

Adult listeners heard a pure-tone glide ("captor") repeatedly alternating with a complex glide consisting of three simultaneous pure-tone glides. One of the pure-tone glide components in the complex glide matched the captor glide in its frequency center, orientation, and direction. This component was referred to as the "target." The two other glides in the complex glide were referred to as "the frame." In some conditions, the captor captured the target glide perceptually into a sequential stream, causing it to be audible as a separate sound in the mixture. The slope of the target relative to the frame was varied. As the slope difference increased, the target became easier to capture. It was also more easily captured when it was parallel to the frame but did not fall into a simple harmonic relation with it.

In a natural acoustic environment, at any given moment, the vibration of our eardrums is the consequence of the joint effects of many different sources of sound that happen to be active at the same time. Patterns caused by the accidental superimpositions of these sources are usually meaningless. Therefore, we must group the incoming acoustic features into subsets that are likely to have come from the same source (Bregman, 1981). There are two aspects of this grouping that we can refer to as sequential grouping and simultaneous grouping. In sequential grouping, we must decide on the sequence of acoustic components that arise over time from the same source. In simultaneous grouping, the set of spectral components arriving at any one instant of time must be decomposed into the contributions of the individual sources.

It appears that there are heuristics for both sequential and simultaneous grouping that collaborate and compete in order to arrive at a correct decomposition of the incoming signal (Bregman, 1978). Using simple recycling signals, we can trigger these heuristics and cause them to group sounds inappropriately. The rapid recycling of a very short signal will often drive the grouping process to supranormal levels, where it remains stable and where its effects are easily observable.

Some partitioning of the momentary spectrum can be done by comparing spectral components in the two ears: components with different spatial origins ought to be dissociated (see Kubovy, Cutting, & McGuire, 1974). Although differences between the signals in the two ears are very important, even in a single ear there is information telling us which spectral components are probably coming from a single source. Many useful clues arise from

the behavior of spectral components over time. It has, for example, been shown that frequency components with synchronous onsets and offsets tend to be heard as the spectral components of a single sound (Bregman & Pinker, 1978; Dannenbring & Bregman, 1978). It also appears that sequential relations among spectral components play a role. If a complex tone is rapidly alternated with a pure tone that has the same frequency as one of the spectral components in the complex tone, the matching component is captured into a sequential stream with the pure tone (Van Noorden, 1975). This capturing gets weaker as the pure-tone captor is moved further away in frequency from the target spectral component of the complex tone. Furthermore, if the complex tone has only two components, and one is captured by a preceding pure tone to form part of a pure-tone sequence, the remaining component of the complex tone is also heard as a relatively pure tone (Bregman & Pinker, 1978).

Generally speaking, the factors that promote the fusion of spectral components can be described loosely by the Gestalt term "common fate." If spectral components are doing the same thing at the same time (for example, going on and off in synchrony), they probably "belong together" as parts of the same signal (Bregman, 1978, 1981). For example, Kubovy (1981) has pointed out that if only one tone in a complex of tones changes in intensity while the remainder remain at the same intensity, the changing one is heard as segregated from the remainder.

In the world of sound, the individual harmonics arising from a single real-world signal behave in a dependent manner. As the fundamental frequency that generates a set of harmonics is caused to glide up and down, as when we modulate the pitch of our voices, all the individual harmonics move up and down in parallel with it on a log frequency scale. From the listener's point of view, this means that any subset of harmonics that are moving around in parallel on log frequency coordinates should be considered

This research was supported by the Natural Sciences and Engineering Research Council of Canada. Requests for reprints should be sent to Albert S. Bregman, Department of Psychology, McGill University, 1205 Docteur Penfield Avenue, Montreal, Quebec, Canada H3A 1B1.

part of the same signal. The listener should fuse such a subset and compute global properties for it (such as pitch and timbre).

There is an additional property of natural acoustic signals that is found when the fundamental moves up or down in frequency. The harmonics retain the same simple frequency-ratio relationships to one another at every instant of time (e.g., the fourth harmonic always bears a 4:1 relation to the fundamental and a 4:3 relation to the third harmonic). One might therefore expect the auditory system to analyze not only for parallel gliding on log frequency coordinates, but for the continuation of simple harmonic relations over time.

Such heuristics would help the listener to deal with the horrendous problem of allocating the large number of spectral components that are audible in complex natural environments to a limited number of acoustic sources. The present experiment looked for evidence that the human auditory system does, in fact, use information of the two types discussed above to determine the fusion of simultaneous glides.

We used as stimuli a set of simultaneous pure-tone glides, gliding linearly on log frequency coordinates. They could be gliding either in parallel or not in parallel. The experiment was addressed toward two questions. First, is it necessary that the glides be exactly parallel? Although it appears that the auditory system, if it is to capitalize upon regularities in the environment, *should* require that the glides be parallel, it is conceivable that it may use a weaker criterion for fusion, possibly one that is easier to compute, such as the criterion that the glides all be ascending (or descending) in frequency. To see how strong the criterion of slope similarity is, it is necessary to vary the slope of co-occurring glides in several steps from the completely parallel case to the opposite-direction case.

Second, the experiment addressed the question of whether the strong fusion of a set of concurrent glides requires not only that the glides be parallel, but also that their frequencies fall into simple harmonic ratios as the partials of a human voice would do. A set of results reported by McAdams (1982a, 1982b, 1984) are relevant to this question. McAdams imposed random frequency jitter (modulation) on subsets of frequency components in a complex tone. When the random frequency modulations of two subsets of partials (say the even and the odd harmonics) were the same, they fused together perceptually to create a single sound, but when the modulations of the two subsets were stochastically independent, the subsets tended to segregate perceptually and be heard as two sounds. This manipulation was one in which the frequency jitter maintained good harmonic relations within each subset of harmonics at every instant of time. However, McAdams went on to apply the same treatment to a non-harmonic series (a series in which the separations of the partials had been "stretched" on log frequency coordinates). Although such a series did not tend, in general, to fuse as well as the harmonic series, the effect of imposing dependent versus independent modulations on the subsets was reported to be quite analogous to the effect

obtained with the strictly harmonic series in some specific stimulus conditions. This result occurred when synthesized voices with "stretched" partials were singing vowels with different pitches and the jitter of the partials was introduced for first one voice and then another, asynchronously. McAdams's results therefore support the idea that parallel frequency change sometimes can make an independent contribution to fusion, even when it is not acting to preserve good harmonic relations over time.

We used a method related to the one first employed by Bregman and Pinker (1978), in which sequential and simultaneous groupings are put into competition, as shown in Figure 1. The stimuli were similar to those used by Steiger and Bregman (1978), except that Steiger and Bregman used *pairs* of simultaneous glides whereas we used the *triplets* of simultaneous glides shown in Figure 1. In the Steiger and Bregman experiments, a pure-tone glide (a "captor glide") was repeatedly alternated with a pair of simultaneous pure-tone glides, one of which (the "target") was in the same frequency range as the captor glide. Under these conditions, the captor sometimes captured the target glide into a stream formed of two sequential glides, thereby breaking the fusion of the target with the glide that accompanied it. When this happened, the listener could hear a cycle composed of an alternation of the captor glide and the target glide; that is, the target did not totally fuse with its accompanying glide. At the same time, he or she could hear, simultaneous with the target, a more or less complex glide, which was the perceptual representation of the glide that accompanied the target.

The captor and target glides of the present experiment are shown in Figure 1. They were made to always have matching slopes and center frequencies, as shown in the figure, since Steiger and Bregman (1978) had shown that this produced the strongest capturing of the target by the

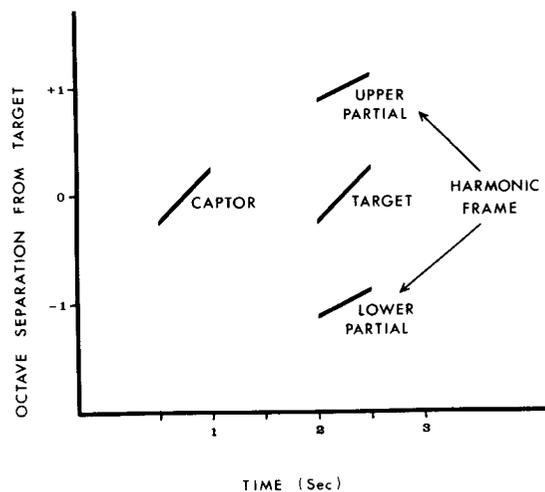


Figure 1. Schematic diagram of one cycle of the repeating stimulus. The slopes of the glides could vary. However, the captor always had the same slope as the target and the upper partial of the harmonic frame always had the same slope as the lower partial. The harmonic frame could also be lowered or raised .25 octaves.

captor. Simultaneous with the target were two other glides, one above the target in frequency and one below it. These two glides always had the same slopes as one another, and their frequencies always maintained a good harmonic ratio. They can be thought of as the "frame" in which the target was embedded. The two glides of the frame defined a gliding harmonic series into which the target glide could be made to fit well or poorly. The main factor that was varied was the relation between the slope of the target and the slope of the frame. We expected that when the target was parallel to its frame (on log frequency by time coordinates), it would fuse most strongly with it, and that the degree of fusion would become gradually less as they became less parallel.

A second factor that we varied was the vertical position (i.e., frequency position) of the frame with respect to the target. In some of the conditions in which the target and frame glides were parallel, the frame was positioned so that the three glides were spaced by octaves throughout their excursions. In other conditions, the frame was displaced up or down as a whole, so as to destroy its octave relations to the target while keeping it parallel to the target. We expected, that, if simple harmonic relations promote fusion, there would be less fusion when the frame was displaced.

The measure of fusion was the resistance of the target to being perceptually grouped with the captor into a sequential stream. This is a reasonable measure of fusion, since sequential grouping and spectral grouping (fusion) compete with one another (Bregman & Pinker, 1978). This resistance was measured in a quantitative way by allowing the subjects to control the amplification of the frame. They were asked to reduce the intensity of the frame glides until the repeating captor-target sequence, that is, a glide repeating twice as often as the frame, could be clearly heard. If the listener had to turn the intensity of the frame down fairly low in order to hear this sequence clearly, it would mean that the target was fusing strongly with its frame.

## METHOD

### Subjects

Eighteen university students (10 female and 8 male) participated as volunteers. They ranged from 20 to 25 years of age, and none reported any known or suspected hearing loss. The results of two subjects were discarded when it was learned, in postexperimental debriefing, that they had misunderstood the original instructions.

### Stimulus Pattern

The stimuli can be described with reference to Figure 1. The individual glides shown as lines in the figure were sinusoidal glides, ascending or descending linearly on log frequency coordinates. Each was 500 msec in duration with additional rise/fall durations of 10 msec. The slope of any glide (target, upper harmonic, or lower harmonic) was set by "rotating" the glide around its frequency midpoint (on a logarithmic scale). In each condition, the captor and target glides were identical, although they varied (together) from condition to condition. For this reason, we can omit the captor in our description of the conditions. The upper and lower glides of the harmonic frame always had the same slope and always were separated by two octaves.

Three slopes were used for the target: (1) rising by .5 octave (in 500 msec), (2) falling by .5 octave, and (3) remaining at a constant frequency ("steady-state glides"). Henceforth, we will express the slopes in octaves/second. Thus, the slopes of the three targets are +1.0, 0.0, and -1.0 octaves/sec. The frequency midpoint for the target was always 724 Hz. The inclusion of both rising and falling targets was intended to control for any orientation bias in the spectral grouping mechanism. The steady-state target condition was included to see whether glides would differ from steady states in their grouping behavior.

Each of the three targets was paired with 11 different harmonic frames to give a total of 33 conditions. Nine of these frames were used to study the effects of slope differences between the frame and the target. The frequency midpoints of the lower and upper glides of these nine frames were always one octave below and one octave above the midpoint of the target (i.e., at 362 and 1448 Hz). Thus, when the slopes of the target and the harmonic frame coincided, the three remained in a consistent 1:2:4 harmonic relation (i.e., octaves) throughout their duration and were phase-locked, starting together at 0°. The nine frames varied in slope from +2 octaves/sec (+1 octave in 500 msec) to -2.0 octaves/sec, in .5-octave/sec steps, and included the 0.0-octave/sec, or steady-state, condition. For any given target, then, the slope of one of the frames exactly matched it and the others differed from it in varying degrees.

Two additional frame types were paired with each target to investigate the importance of the simplicity of the harmonic relation between the frame and the target upon their fusion. In these two conditions, the slopes of the frame and target were identical, but the target and frame partials were not in the ratio 1:2:4 as before. Instead, the frame was moved up by .25 octaves in one of these conditions and moved down by .25 octaves in the other. Thus, in the latter case, for example, the lower partial of the frame was 1.25 octaves below the target and the upper partial was .75 octaves above the target. Therefore, while the frequency ratios between the target and the partials of the frame remained *constant* throughout the glide, the relations were not as *simple* as before; the glides now maintained the ratios 1:1.68:4 throughout their durations. Notice that the harmonic relations of the two partials of the frame were still simple in relation to one another, in this condition, but that their relation to the target was now more complex.

The stimulus was a repeated alternation of the captor and the target-frame complex. The timing was as follows: captor for 500 msec (plus 10-msec rise/fall times), 1 sec silence, target with frame for 500 msec (plus 10-msec rise/fall times), 2 sec silence, repeat.

### Task

At the beginning of each trial, the intensity of both the captor with the target were at 55 dBA and the intensity of the harmonic frame was set to 95 dBA. At this setting, the captor-target stream could not be heard. Rather, the subject heard the captor alternating with a rich glide composed of the target and its harmonic frame.

Each trial presented 15 cycles of the stimulus, during which the subject was to turn down the intensity of the harmonic frame until the captor-target cycle could be heard clearly. "Hearing the target clearly" involved hearing it as remaining clear throughout its duration (not just at onset or offset) and as having the same quality as the captor. Once lowered, the frame intensity was not to be raised. The subjects were urged to adopt a consistent criterion for "hearing the target clearly" and to maintain it across trials.

Between trials, there was a 7-sec period for the subject to record the amplification setting and a 5-sec period in which to raise the intensity of the frame to its maximum (95 dB) in preparation for the next trial. The beginnings of each of these periods were signaled by brief noise bursts.

### Experimental Design

Each subject received all 33 conditions once. There were four different randomizations of the conditions, each presented to four subjects.

### Apparatus

The stimuli were synthesized on a PDP-11/34 computer (Digital Equipment Corporation) using the MITSYN signal processing software (Henke, 1980). At the time of playback, an ADC Soundshaper II graphic equalizer

(12 bands per channel) was used to produce a flat frequency response of the headphones in the range that was used.

#### Procedure

Oral instructions and diagrams were used by the experimenter to describe to the subjects the patterns of tones and the possible percepts. The subjects were instructed in the use of the apparatus and then received 25 repetitions of a randomly selected stimulus on which to practise the task. They then received 7 more practice trials gradually approximating the conditions of the experiment. Finally, they received the 33 experimental conditions, broken by a rest after 13 trials.

## RESULTS AND DISCUSSION

Figure 2 shows the average frame intensities at which sequential streaming occurred for the rising, falling, and steady-state targets. The plots for the three targets show similar tuning curves, with a minimum when the targets and frames are parallel (0-octave/sec slope difference). The frame intensity had to be lowered by 10 to 12 dB further in the parallel conditions before the target could be perceptually segregated from the frame (as compared with the average nonparallel condition). Because of the similarity and regularity of the three tuning curves, no statistical analysis was used to verify the existence of the effect of target-frame parallelness.

The effects of the simplicity of the target-frame harmonic relations can be seen by comparing the low point in each tuning curve to the two disconnected points that lie directly above it on each graph. The disconnected points represent conditions in which target and frame are parallel but do not have a simple harmonic relation to one another. These points lie consistently in the range of the nonparallel conditions.

These results were examined by performing a two-way ANOVA upon only the conditions that had parallel targets and frames. The two factors were target type (rising, falling, steady-state) and frame position (octave relation, raised, lowered). There was a significant effect of target

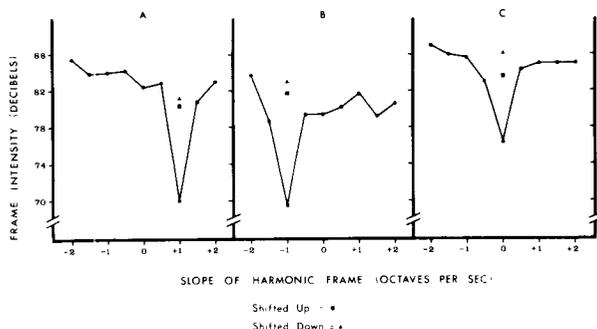


Figure 2. Results: The setting of the intensity of the harmonic frame (in dBA) at which streaming occurred, plotted as a function of the slope of the harmonic frame, for (a) rising target (+1.0 octave/sec), (b) falling target (-1.0 octave/sec), and (c) steady-state target (0.0 octaves/sec). Higher settings indicate that the target could be segregated more easily from the frame. The two isolated points are for the "shifted up" frame (squares) and the "shifted down" frame (triangles).

type [ $F(2,30) = 3.57, p < .05$ ] and of harmonic relation [ $F(2,30) = 4.39, p < .05$ ]. On inspection of Figure 2, it can be observed that the significant effect of target type arises from the fact that the steady-state target always shows less fusion than the rising or falling targets do. This holds true across all conditions, whether the target and frame are spaced by exact octaves or not, and holds across all levels of target-frame slope differences.

Looking at the three tuning curves, another observation is that the dip in the steady-state curve is shallower at the point of optimum tuning than it is for the other two curves. Taken together, these two observations seem to argue that steady states are easier to decompose perceptually than glides are. McAdams (personal communication) has made an observation that is consistent with this possibility. He points out that under favorable conditions it is possible to "hear out" some of the lower harmonics from a complex tone, but that when the tone is subjected to a small amount of frequency modulation, this "hearing out" is prevented. In the present experiment, however, the increased ability to capture a steady-state target may be the result of an enhanced sequential grouping of the target with its captor rather than a diminished grouping of the target with its frame. It may be just another example of the dominant role of frequency proximity in the formation of sequential streams. The steady-state captor is concentrated at one frequency, and so is the steady-state target. Hence, every point within the captor glide is equal in frequency to every point in the target glide, whereas in the ascending or descending glides only corresponding points are equal in frequency.

The significant effect of the harmonic relation bears on one of the original purposes of the experiment. This result suggests that it is not sufficient for glides to move in parallel in order for strong fusion to occur. They must also maintain simple harmonic relations. This finding tends to validate the approach taken by Parsons (1976), whose computer program separates voices by grouping sets of harmonics that have a common fundamental.

We can next ask the question as to whether the tuning curves are gradual or whether there is fusion only when the frame and target are exactly parallel and in octave relations. One way to examine this question is to ask whether the first and second points away from the minimum in each of the turning curves differ from one another, with the point that is closer to the minimum being lower in value. This should hold true for both the two points to the left of the minimum and the two points to its right. A simple way to test this is to count, across every subject and every target type, the number of times this pattern of data is obtained. There are 96 comparisons possible in the data (16 subjects  $\times$  3 captors  $\times$  2 directions of deviation from the minimum of the curve). Of these, 64 show the pattern, 23 show an opposite pattern, and 9 are indecisive (equal data points). The pattern is therefore reliably found ( $p < .001$  by the sign test for the null hypothesis of equal numbers of confirmations and disconfirmations), although the tendency is not a large one. We can conclude that the

agreement of the glide slopes of the target and the frame does not have an all-or-nothing effect and that greater differences yield greater segregation. However, we should note that most of the variation occurs between 0.0 and +0.5 octave/sec of slope difference (i.e., between the minimum and the closest point on either side). We have to remember, though, that the step size by which the slopes were altered was relatively large and that a smaller step size might show better evidence for intermediate degrees of tuning.

One final observation is that an inspection of Figure 2 shows that for all three targets the curves are asymmetrical, the negative-sloped frames permitting more segregation of the target than the positive-sloped frames do. It is possible that this may be due to an interaction of two factors. First, there is likely to be more masking of the target from the lower partial of the frame than from the upper partial, even at equal log frequency separations. Second, the more negative the slope of the lower partial, the more it is separated from the target at the end of the glide. If the final positions of tones cause more masking than earlier portions, we would expect to find the asymmetry observed in the curves. The assumption of more upward-spreading masking also accounts for the fact that, for all three target slopes, the downward-shifted parallel frame (which took the lower partial further from the target) permitted more segregation of the target than the upward-shifted parallel frame did (see the isolated pairs in Figure 2).

The results of our experiment support the view that the auditory system employs mechanisms that utilize frequency relations among the partials of the gliding complex tone in deciding which partials have arisen from the same source. These may depend upon the phase locking that holds among the harmonics of a fundamental and upon the continued phase locking that occurs despite changes in frequency of the fundamental.

It is tempting to conclude from these results that "common fate" (parallel frequency change) in and of itself does not promote the fusion of simultaneous components, and that it is only the preservation of good harmonic relations throughout a glide in frequency that causes strong fusion to occur. However, we hesitate to totally rule out a contribution from parallel changes. One reason for this hesitation comes from the observation that parallel target-frame positioning produced less fusion with the steady-state target than with the rising or falling targets, suggesting that partials will fuse better when they undergo a common motion.

A second reason for our hesitation is the incompatibility of our results with those of McAdams (1982a, 1982b, 1984) discussed earlier. Recall that his results, unlike ours, showed that a frequency modulation shared by a set of partials of a complex tone did, under certain specific conditions, act to bind that subset together perceptually even when its members were not harmonically related. We therefore prefer to keep open the question, pending further experimental analysis of the differences between

McAdams's results and our own and their dependence upon the particular paradigms used to study the question in our laboratories.

A consideration of the mechanisms by which the auditory system determines the parallelness of change in different frequency regions could be instructive at this point. Two approaches to the solution of this problem suggest themselves. The first method could be referred to as the "harmonicity" approach. The auditory system, using either "place" or "periodicity" information, would perform a separate calculation at each instant of time to determine the one or more fundamentals that could have given rise to the set of frequencies received at that instant. Then it would collect into the same analysis all frequency components related to this fundamental. This seems to be the method favored by engineering approaches to voice separation (e.g., Parsons, 1976). It requires that the frequency components that are to be grouped be strictly harmonically related.

A second approach would be to compute the change that occurred locally in each frequency region at each instant of time (using either local "place" changes or "periodicity" changes), and then to group together those simultaneous frequency components that registered an equal change (on a log frequency scale). This method would be effective with either harmonic or inharmonic partials. It could be referred to as the "direct slope" approach.

The tuning curves that we have obtained in the present experiment seem to us to be more consistent with the direct slope approach. Our argument hinges on the finding that, on the average, a target glide whose slope was mistuned by 1 octave/sec from that of the frame fused better with it than a glide that was mistuned by 2 octaves/sec from it. Yet, in both cases of mistuning, the difference in slope would have been sufficient to rapidly and continuously alter the harmonic relation between target and frame. A computation that looked for harmonic relations at each moment of time (the harmonicity approach) should have found such relations at only very rare moments in both cases. It is not at all clear why such a mechanism should find one of these conditions more "harmonic" than the other. On the other hand, a mechanism that carried out a number of local computations of *slope* and acted to group frequency components on the basis of equality of slope might behave differently. It might not have an all-or-none criterion for equality. Furthermore, a gradual criterion might be particularly visible with the kind of glides that we used in our experiment. With these glides, when there was a mismatch of 1 octave/sec, the mismatch (computed in terms of slope) remained constant at this level throughout the course of the glide, and therefore remained consistently better (from the point of view of a "slope" analyzer) than a mismatch of 2 octaves/sec. Under these conditions, one might expect a "direct slope" mechanism to more strongly fuse glides whose slopes were less discrepant.

At this point, we have to consider the question of whether there is any evidence at all that the auditory sys-

tem does compute the slope of glides. Fortunately, such evidence exists. Steiger and Bregman (1978) reported experiments that used a stimulus similar to the one shown in Figure 1 except that the target was always parallel to its accompanying glide. Instead of varying the slope relation between the target and its accompanying glides, as in the present experiment, Steiger and Bregman varied the slope relation between the captor and target glides. They found that the best capturing of the target glide occurred when the captor and the target had the same slopes. Since these two glides do not co-occur temporally, the only way in which the auditory system could be responsive to a slope similarity would be by comparing a memory of the slope of the first glide with the slope of the incoming second glide; this presupposes a glide-encoding mechanism. Other experiments have shown physiological evidence for the existence of channels sensitive to rate of frequency modulation in the auditory system (Evans & Whitfield, 1964; Vartanian, 1974). Psychophysical data (e.g., Pollack, 1968) and selective adaptation studies (Gardner & Wilson, 1979) have demonstrated that humans are sensitive to the orientation of glided sounds.

The finding in our experiment that argues against the "direct slope" model is the one that shows that in the conditions in which the slopes matched but the harmonicity was reduced, there was less fusion. This finding, however, cannot be viewed as conclusive in the face of the evidence cited above. When the auditory system gives us evidence both for a "direct slope" method of grouping and a "harmonicity" method, perhaps we should believe the evidence and conclude that the system may use both methods, perhaps under different stimulus situations.

#### 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, *Perceptual organization*. Hillsdale, NJ: Erlbaum.
- BREGMAN, A. S., & PINKER, S. (1978). Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, *32*, 19-31.
- DANNENBRING, G. L., & BREGMAN, A. S. (1978). Streaming vs. fusion of sinusoidal components of complex waves. *Perception & Psychophysics*, *24*, 369-376.
- EVANS, E. F., & WHITFIELD, J. C. (1964). Responses of cortical neurones to acoustic stimuli varying periodically in frequency. *Journal of Physiology* (London), *172*, 52P-53P.
- GARDNER, R. B., & WILSON, J. P. (1979). Evidence for direction-specific channels in the processing of frequency modulation. *Journal of the Acoustical Society of America*, *66*, 704-709.
- HENKE, W. L. (1980). *MITSYN: A coherent family of command-level utilities for time signal processing*. (Revised and extended by Daniel Solomon, Dept. of Psychology, McGill University; available from W. L. Henke, 133 Bright St., Belmont, MA 02178.)
- KUBOVY, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, NJ: Erlbaum.
- KUBOVY, M., CUTTING, J. E., & MCGUIRE, R. M. (1974). Hearing with the third ear: Dichotic perception of a melody without monaural familiarity cues. *Science*, *186*, 272-274.
- MCADAMS, S. (1982a, November). *Contributions of sub-audio frequency modulation and spectral envelope constancy to spectral fusion in complex harmonic tones*. Paper presented at the meeting of the Acoustical Society of America, Orlando, Florida.
- MCADAMS, S. (1982b). Spectral fusion and the creation of auditory images. In M. Clynes (Ed.), *Music, mind, and brain*. New York: Plenum Press.
- MCADAMS, S. (1984). The auditory image: A metaphor for musical and psychological research. In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive processes in the perception of art*. Amsterdam: North-Holland.
- 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.
- POLLACK, I. (1968). Detection of rate of change in auditory frequency. *Journal of Experimental Psychology*, *77*, 535-541.
- STEIGER, H., & BREGMAN, A. S. (1978). Capturing frequency components of glided tones: Frequency separation, orientation, and alignment. *Perception & Psychophysics*, *24*, 312-322.
- VAN NOORDEN, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Unpublished doctoral dissertation, Technische Hogeschool Eindhoven, Eindhoven, The Netherlands.
- VARTANIAN, T. A. (1974). On mechanisms of specialized reactions of central auditory neurones to frequency-modulated sounds. *Acoustica*, *31*, 305-310.

(Manuscript received June 4, 1982;  
revision accepted for publication June 22, 1984.)