THE PERCEPTUAL DECOMPOSITION OF COMPLEX SOUNDS COMPOSED OF SIMULTANEOUS FREQUENCY GLIDES

by

Ⓒ Howard Steiger

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Arts

Department of Psychology
McGill University
Montreal.

August, 1980

M.A. PSYCHOLOGY

Howard Steiger

THE PERCEPTUAL DECOMPOSITION OF COMPLEX SOUNDS COMPOSED OF SIMULTANEOUS FREQUENCY GLIDES

Abstract

Three experiments studied the contexts in which two simultaneous sine tones, gliding up or down in parallel, would fuse to form a rich-sounding glide, or would be audible as two independent sounds. The pair of simultaneous glides tended to be perceived as separate when they were preceded by another gliding tone (or 'captor') that acted to capture one of them (the 'target' glide) into a sequential stream. The pair of simultaneous glides and the "captor" were rapidly alternated, and the relation between the captor and target glides was varied on three dimensions: (1) the frequency proximity between them, (2) their correspondence in glide orientation and (3) their collinearity (when plotted on a logarithmic scale of frequency). The results suggested that frequency proximity and common orientation, but not common "trajectory" (i.e., collinearity) play a role in causing the alternating captor and target tones to be perceived as a sequential stream, and therefore in "decomposing" the complex glide. A possible neural mechanism was proposed to account for the results, and the functional role of such a mechanism was discussed.

M.A. PSYCHOLOGIE

Howard Steiger

LA DECOMPOSITION PERCEPTUELLE DE SONS COMPLEXES COMPOSÉS DE GLISSANDI SIMULTANÉS

Sommaire

Trois expériences sont présentées dans le but d'étudier l'influence du contexte sur la perception de deux sons sinusoïdaux simultanés, parallélement ascendants ou descendants. Les expériences ont servi à examiner si ces sons étaient perçus comme un glissando composé ou étaient perçus comme étant indépendants. Les deux glissandi simultanés tendaient à être perçus séparément lorsqu'un autre glissando (ou 'capteur') transformait l'un d'entre eux (glissando - cible) en une suite séquentielle. Les deux glissandi simultanés et le glissando "capteur" furent rapidement alternés et la relation entre le "capteur" et la "cible" fut variée sur trois dimensions: (1) la proximité de leur fréquence, (2) la similarité de l'orientation des glissandi, (3) leur collinearité (évaluée sur une échelle logarithmique de fréquence). Les résultats suggèrent que la proximité de fréquences et la similarité d'orientation commune, mais non la ressemblance de la "trajectoire" (i.e., collinearité), influence la perception des sons "capteur" et "cible" alternés, en les transformant en une suite séquentielle et jouent donc un role dans la décomposition d'un glissando complexe. Une interprétation de ces résultats suggérant un éventuel mècanisme neural est proposée, et la fonction de ce mècanisme est discutée.

Acknowledgements

Foremost, I wish to express my gratitude to my thesis supervisor, Dr. Albert S. Bregman, for the insights, guidance and support he has offered both in the preparation of this manuscript and throughout my association with him. I also wish to thank Rhonda Amsel for her statistical advice on this project, and Daniel Salomon for his computing assistance. I am also grateful to Janet Gabel, Justine Sergeant, Tina Dornbusch, Jack Abramson, and Seth Katz, who unselfishly offered their time, advice and support during many phases of this project. I also wish to thank my typist, Ethel Steiger. Finally, I would like to thank Mimi Israel for her help and comments during the data analysis, and the preparation of the figures and manuscript, and above all, for her tolerance during my months of preoccupation with this thesis.

The research for this thesis was supported by a CORE grant awarded to the McGill University Department of Psychology by the Natural Sciences and Engineering Research Council of Canada to support the computer-based laboratory, and by an operating grant awarded to Dr. Albert S. Bregman from the Natural Sciences and Engineering Research Council of Canada.

TABLE OF CONTENTS

INTRODUCTION

Our subjective experience tells us that listening to a musical passage or a speaking voice--or in fact to any sequential auditory pattern which changes in frequency over time--is analogous to visually tracking a moving object through space. The notion that auditory sequences "move" through a "pitch space", much as tangible objects move in three dimensional space, was raised in the early scientific literature by both Helmholtz (1862/1954) and Koffka (1935/1963), and has appeared more recently in a theory of auditory perception developed by Jones (e.g., 1976).

Such an analogy has aroused both curiosity and skepticism among those who study perceptual phenomena--curiosity about why the subjective experience of the analogy is so compelling, and skepticism because the notion has developed from descriptions of similarities between the modalities which are more often metaphorical than physical. Yet recent experimental evidence has suggested that there is more than a metaphorical correspondence between both the pattern processing behaviour and mechanisms of visual and auditory modalities (e.g., Julesz & Hirsh, 1972; Marks, 1978).

Outstanding among such analogies is the resemblance between auditory and visual perceptual organization phenomena. Such phenomena involve the organization of multiple events into units of perception or "gestalts". The principles underlying perceptual organization in vision were thoroughly described in the early psychological literature by the Gestalt psychologists (e.g. Koffka, 1935/1963; Rubin, 1921). More recently, principles resembling those discussed by the Gestaltists have been successfully implemented in computer pattern recognition devices (e.g., Shirai, 1975; Uhr, 1973), suggesting that the principles provide good heuristics for parsing objects out of complex images. Curiously, perceptual organization in audition appears to adhere closely to predictions based on Gestalt laws, which were developed primarily through studies of visual pattern perception.

The law of proximity and similarity

The Gestalt law of proximity and similarity stated that ". . . two parts in a field will attract each other according to their degree of proximity and equality" (Koffka, 1935/1963, p. 166). The law describes the tendency for similar visual events to fuse into a single perceptual unit when they are in close proximity. For example, it accounts for the fact that the dots on the six-dot face of a die appear to form two rows (with 3 dots in each row), rather than three rows (with 2 dots in each row). The elements which are closest, (i.e., each "triplet" of dots), tend to be grouped together during the act of perception.

Perceptual organization by frequency proximity

Similarly, in audition it has been observed that when listening to a melody composed of notes that are widely separated in frequency, listeners often report hearing not one, but a pair of melodies; one consisting of the higher tones and one consisting of the lower tones (Bregman & Campbell, 1971; Miller & Heise, 1950; Norman, 1967; Ortmann, 1926; Van Noorden, 1975). Thus, events widely separated in frequency appear to become perceptually isolated from each other, while those in close frequency proximity seem to group into single perceptual units. Bregman and Campbell (1971) referred to such perceptual isolation of the higher-pitched elements from lower-pitched elements in a tone sequence as "primary auditory stream segregation". They suggested that stream segregation results from preattentive processes which use the frequency relations between tones in an auditory sequence to organize the tones into perceptual units, so that when bimodal distributions of frequencies are present in the sequence, it tends to be split into two perceptual "streams". This conclusion was supported by the observation that listeners erroneously reported the temporal order of sequences containing three interleaved high and low range tones to be a string in which three high tones were followed by three low tones, or in which three low tones were followed by three high tones. They suggested that these erroneous judgments resulted from the fact that the sequences tended to be perceived as forming two separate streams--one containing only the high-pitched tones, and one containing only the low-pitched tones. It was proposed that the segregation of the sequence into two streams prevented listeners from attending to more than one stream at a time, and therefore, from tracking the veridical pattern consisting of alternations between high and low tones. Instead, listeners seemed to report the order of events in a stream containing three tones in one range, and then of the stream containing the three other tones. Thomas and Fitzgibbons (1971) observed that such difficulty in ordering sequences of tones was a function of both the frequency separation and the rate of alternation between the tones. They reported that when tones in a sequence were separated by less than a musical fourth, subjects could correctly order the tones at a rapid 125 msec/ event rate. However, tones at wider separations required progressively slower presentation rates to be correctly ordered. A similar tradeoff between the frequency separation of tones and the rate of presentation of the tones has been reported by Van Noorden (1975).

Perceptual grouping based on proximity in both auditory and visual patterns

Similar observations about the inability of subjects to correctly judge the relations between widely separated events have been made using both auditory and visual stimuli. Kinney (1961) reported that subjects not only had difficulty in temporally localizing a deviant tone (in frequency) with respect to other tones in an auditory sequence, but in horizontally localizing visual events with respect to each other when they were separated by a large vertical distance. Bregman and Achim (1973) demonstrated that the spatially more separate visual events in a rapid temporal sequence tended to segregate into "streams" of perceived movements. The actual spatio-temporal stimulus pattern they used was a repeating sequence of dots presented at eight discrete positions on a vertical axis. The consecutive dots alternated between the higher and lower of the eight positions, and the sequence was presented at three different speeds, by varying the event rate between 57, 114, and 227 msecs/dot. At the slow speed, an illusion of motion of one dot through an irregular pattern of motion (i.e., moving back and forth from high to low positions) was produced. However, at higher speeds, an illusion of motion occurred only between the events in the higher or lower positions, so that an illusion of two moving dots was perceived--one moving between the higher positions, and one moving between the lower positions. Thus, at rapid rates of presentation, the events that were closest to each other (i.e., the sets of high or low dots) appeared to group together, creating two "streams" of apparent motion.

Such apparent motion resulting from the presentation of rapid sequences of spatially separated visual events is referred to as the "phi" phenomenon (e.g., Kolers, 1972). It has been suggested that such movement phenomena result from the activation of mechanisms normally involved in the perception of true movement (Frisby, 1972). The identification of cortical units in cats, which respond only to moving images on the retina, suggests the existence of neural organizations which permit the translation of the successive images projected onto the retina by a moving object into the perception of motion (Hubel & Wiesel, 1959). To account for the "phi" phenomenon, it is assumed that such mechanisms can also be triggered by successive rapid stimulations of discrete regions of the retina. One may also speculate that such mechanisms are responsible for the organization of events into moving visual "streams" on the basis of their proximity to each other. Perhaps, such mechanisms are more sensitive to rapid "jumps" across neighbouring regions of the retina than to rapid jumps across distant regions. Therefore, a rapid pair of visual events exciting neighbouring regions of the retina would trigger a "movement detector", resulting in the perception of movement between the pair. Rapid events more distant from each other might fall outside of the spatio-temporal receptive fields of such detectors, and would thus not induce the apparent motion phenomenon. Possibly, such mechanisms account for the fact that at high presentation rates apparent movement was perceived only between the successive high or successive low elements in Bregman and Achim's (1973) stimulus array. The distance between the successive high and low elements might have been too great to activate motion detectors.

Van Noorden (1975) has adopted the concept of "movement detectors" to explain the tendency for tones in rapid sequences to lose their temporal coherence, and to group into streams on the basis of frequency proximity. He proposes that the auditory system tracks changes in frequency with "pitch-motion" detectors which are more sensitive to "jumps" between neighbouring frequency regions than to rapid jumps across large ranges in frequency. Therefore, only the shifts between successive tones within a critical frequency range, at a given rate of presentation, are conceived to excite "pitch-motion" detectors. "Pitch-motion" would thus be perceived only between the tones in a sequence that were sufficiently close to each other in frequency. No percept of temporal coherence between tones separated by more than the critical range would develop. As a result, instead of perceiving movement in pitch between widely separated sequential tones, a percept of two separate sources (or streams) of "movement" in pitch would occur. Thus, according to Van Noorden's conception, the tendency for tone sequences to split into high- and low-range streams is viewed as a failure in the capacity of the auditory system to track large, rapid frequency changes.

However, Bregman's (e.g., 1978a, 1978b) theory attributes a more ecological significance to "stream segregation". He suggests that the stream segregation phenomenon plays an adaptive role in organizing rapid sequences of events. According to the theory, each stream is the psychological correlate of a physically distinct sound source. This view follows from the fact that in a natural context any single sound source tends not to produce large and rapid shifts in frequency. Therefore, naturally occurring waveforms composed of elements widely separated in frequency are often the product of a summation over many simultaneously active sound sources. Bregman proposes that, perhaps, by organizing the sequential components of such acoustic mixtures that are close in frequency into streams, the auditory system also succeeds in extracting each of the original sources out of the mixture. Thus, the tendency for rapid tone sequences to segregate into perceptual streams on the basis of frequency proximity may reflect a heuristic, which has evolved for the purpose of decomposing naturally occurring acoustic mixtures into their component sources. The role of stream segregation may be to integrate those acoustic components that probably arose from one source into a single auditory stream.

Instead of being viewed as a direct function of a speed/ frequency separation tradeoff between sequential events, "stream segregation" is viewed by Bregman as being a function of context. By considering context, the auditory system is thought to estimate the likelihood that any tone does or does not belong in a stream with other tones. Such a role of contextual information was shown by Bregman (1978d) in experiments in which subjects judged the streaming of the same two alternating tones A and B, when they were placed in different contexts. When A and B were followed in a sequence by two tones (X and Y) at a far lower frequency range, and the ABXY pattern was repeatedly cycled, A and B were heard as grouping into one stream, while the low tones X and Y grouped into a second stream. However, when X was placed close in frequency to A, and Y close in frequency to B, A and B were heard as occurring in separate streams. One stream was formed of alternations of A and X, and the other of alternations of B and X. Thus, the splitting apart of A and B depended on the context of other tones, rather than on the absolute frequency separation between A and B.

If Bregman's view is correct, then by examining the factors which induce sounds to become perceptually isolated from each other, we can hope to develop some understanding of how the auditory system parses individual sound sources out of a complex auditory environment. However, much of the work which has examined stream segregation tendencies has used sequences of steady state tones as stimuli. Since a great many naturally occurring sounds consist largely of glided frequency transitions, the current study sought to examine the factors which induce sequences containing frequency glides to group perceptually into streams. Of interest was whether or not frequency glides could be considered analogous to objects moving along specific trajectories.

The role of frequency glides in auditory perception

Most of the investigations into the role of frequency glides in audition have been conducted in research concerned with speech perception. In natural speech, the glided elements consist of glided formant transitions. Formants result from the filtering by the speech tract of complex waveforms generated at the vocal chords. Glided formant transitions are produced when the centre frequency of the filtering mechanism is continuously adjusted up or down during speech production. For many years, it has been proposed that such glides play an important role in differentiating the units of speech, or phonemes, from each other (e.g., Liberman, Cooper, Shankweiler & Studdert-Kennedy, 1967). For example, the presence or absence of formant transitions has been shown to influence the discrimination of voiced and voiceless consonants (Stevens & Klatt, 1974). Speech sounds containing glides at the onset of voicing are judged as being voiced, while those lacking such glides are perceived as being voiceless.

Several more recent findings suggest that glided transitions also play an important role in maintaining a string of speech sounds as a coherent stream, permitting the sounds to be temporally localized with respect to each other. (If the sounds split into separate streams, such localization should become impossible. This effect would render a speech passage unintelligible). For example, Thomas, Cetti and Chase (1971) observed that it was difficult for subjects to judge the order of a sequence of vowel sounds created by splicing together portions of audiotape, each containing a different vowel sound. In their experiment, the temporal localization of each event relative to the other events became difficult, suggesting that the sequence segregated into more than one auditory stream. Perhaps, the qualitative dissimilarity between the sounds in the sequence, resulting from differences in the frequency spectrum and other acoustic attributes of each sound, caused the sounds to segregate from each other. Similarly, Warren, Obusuk, Farmer and Warren (1969) found that sequences of synthetic speech-like sounds required presentations as slow as 700 msecs/event to be correctly ordered. Curiously, the perception of natural speech demands a correct response to the ordering of phonemes occurring at rates as high as 30 or 40 msecs/event. Apparently then, the sequences employed in both experiments lacked a critical property which must be inherent in natural speech, one which promotes the organization of successive phonemes into a single sequential stream. Since the sequences in both cases lacked glided frequency transitions between successive events, it is likely that glides are a good candidate for "the property" which maintains the coherence of a speech passage. In natural speech, glides connect adjacent phonemes. Cole and Scott (1973) found that removal of the glided transitions in consonant-vowel (CV) syllables caused the perceptual segregation of the consonant from the vowel when such syllables were rapidly repeated over and over again. The finding supports the notion that part of the role of glides in speech is to make the speech train more coherent, and therefore more resistant to the tendency for individual components to segregate on the basis of differences in frequency or other acoustic attributes. Glides might thus indicate to the auditory system which of several successive components to group into a single stream (corresponding to a single speaker's voice), and which to segregate out of the stream.

Bregman and Dannenbring (1973) showed a parallel phenomenon to the Cole and Scott (1973) finding, using stimuli which were in some sense "pure tone analogs" to sequential CV syllables (Note 1). They showed that alternating high and low tones connected by frequency glides resisted the "splitting" tendency, at presentation rates far faster than those required to cause the segregation of an identical sequence of tones which lacked such glided transitions. As well, partial transitions which simply "pointed" from each tone toward the adjacent tone in the sequence were sufficient to improve the coherence of the stream as compared with the condition in which the frequency transition from one tone to the next was completely discontinuous. They suggested that the effect might indicate that auditory mechanisms follow the continuity created by a glided or partially glided transition. Thus, in their experiment, the mechanisms might have become "primed" by the transitions to correctly anticipate the range in which each successive tone would fall, thereby improving the temporal coherence of the stream.

Van Noorden (1975, p. 51) assumes that gradual frequency changes give better stimulation of "pitch-motion detectors"

---

1. Pure tone glides have been used as analogs to formant transitions in speech on several occasions (e.g., Cutting, 1974; Gardner & Wilson, 1979).

than do discontinuous changes in frequency. Thus, in the Bregman and Dannenbring (1973) stimuli, the glided transitions between successive tones may have provided better stimulation of "pitch motion detectors" than when discrete high and low tones were alternated, causing greater temporal coherence of the stream in the glided transition conditions.

The role of continuity in perceptual organization

The preceding evidence suggests that continous changes in frequency over time contribute to perceptual organization, by promoting the perceptual grouping of events which have such continuity between them. The notion that continuity between successive events acts as a "binding force" in perceptual organization was first stated in the Law of Good Continuation (Koffka, 1935). An instance of good continuation occurs when viewing a scene through the slats in a picket fence. The sensory events in the scene become effectively slashed into small vertical strips. However, continuities (or 'good continuation') between the images on either side of the occlusion produced by each slat in the fence cause the scene to be perceived as a whole, uninterrupted image. Principles analogous to "good continuation" have been used with some success in programs designed to enable the computer recognition of partially occluded objects (e.g., Shirai, 1975). Continuity between line segments on either side of an occlusion is taken by such programs to indicate that the line segments are components of a longer uninterrupted line.

A notion analogous to "good continuation" is embodied in a computer program developed by Parsons (1976) to separate a pair of overlapping speech signals. Parsons' system separates the signals essentially by identifying the pair of fundamental frequencies most likely to be associated with the acoustic spectrum at each of small, discrete moments in time, and by tracking the frequency change in those fundamentals over time. The system assumes that pitch will not radically change from one segment to the next, and uses a tracking algorithm to ensure continuity. The tracking algorithm extrapolates from previous trajectories of change in frequency over time, and therefore can deal with situations in which the fundamental tracks from two voices cross each other's paths. In such cases, the program simply extrapolates across the point of crossing, joining the segment prior to the "crossing" with the one after the "crossing" that maintains a continuous trajectory with the first segment. The fact that such use of "trajectory" information provides a fairly successful algorithm for decomposing simultaneous speech signals suggests that it may be important for the auditory system, much like the visual system, to rely on "good continuation" during perceptual organization processes. Therefore, in both modalities certain events may act as "pointers" to subsequent events, so that the "pointed to" events become perceptually integrated with the "pointer". However, to follow a continuity in the frequency change between successive events, (e.g., in order to extrapolate from one gliding segment onto another), the auditory system must first be capable of determining the rate of ascent or descent in frequency over time during each glide. Otherwise, it would be impossible to determine whether or not the two segments were aligned on a continuous path.

The notion that the auditory system tracks continuous changes in frequency over time, (or more specifically, the velocity of pitch change between successive events), is a central component of a recent theory of auditory perception discussed by Jones (1976) and Jones, Maser and Kidd (1978). The theory proposes that the auditory system processes serial patterns by mapping the pattern present in the incoming signal onto a predictive cognitive structure, which uses the velocity of pitch change between events to generate predictions about what events should follow. The function of this mental structure is to "prime" appropriate neural locales in anticipation of the "to be heard" pattern. When the pattern is consistent with the internal predictive structure, it is thought to be easily integrated into a coherent sequence.

Jones proposes that the predictive structure is generated in accordance with the assumption that pitch change in all patterns will be at a constant velocity. Thus, within the context of any pattern, small changes in pitch are expected to occur relatively close in time, while larger changes are expected to occur further apart in time. Therefore, according to the theory, when patterns change at a uniform velocity, they are easier to process serially since they maintain a predictable pattern of change. Being predictable, the positions in time and in a "pitch-space" of the elements of a uniform velocity pattern will correspond to those anticipated by the internal predictive structure. Cognitive mechanisms "primed" to receive events at the appropriate time/"pitch-space" locales easily integrate the incoming serial pattern. Events which do not correspond in time and in the "pitch-space" to the "primed" locales should be more difficult to integrate into a serial pattern. As a result, Jones accounts for "stream segregation" effects as being due to a mismatch between anticipated and actually-occurring events. She proposes that those events in a sequence that do not match the predictive internal sequence (i.e., those which denote a velocity of change inconsistent with the velocity of change between other events in the pattern) are not integrated into the sequence, and therefore "split off" into a separate perceptual stream.

Some behavioural evidence does appear to support Jones' formulation. Jones (1976) reported that memory for patterns in which pitch change was at a constant velocity was better than that for patterns in which the velocity of change was inconsistent. In addition, Heise and Miller (1951) showed that when a tone deviated too far from the trajectory implied by a pattern of tones, the deviant tone "split away" from the pattern. Such an effect is perfectly consistent with predictions made by Jones' theory. The pitch distance between the deviant tone and its predecessor would have demarcated a jump in pitch at a velocity inconsistent with the change predicted by the anticipatory cognitive mechanism, on the basis of the velocity of pitch change between the previous tones in the sequence. As a result, the mechanism would fail to integrate the deviant tone with the rest of the pattern.

A frequency glide can be viewed as a serial pattern composed of successive excitations of adjacent frequency regions. Since exponential changes in frequency are thought to produce linear changes in subjective pitch (Békésy, 1960), a glide in which the rate of frequency change across equal moments in time is exponential should have a uniform velocity of pitch change. Such a glide should, therefore, represent an easily processed "constant velocity" pattern. Sequences of such glides, aligned so that the rate of pitch change in each glide is continuous with that of the previous glide, should thus also produce easily integrated constant velocity patterns. According to Jones' theory, the auditory system should thus be sensitive to a common "trajectory" of pitch change over time between successive glides, since common trajectory necessarily equates with a constant velocity of change.

Recent electrophysiological data suggest that neural units which might perform some of the functions required to track the trajectory and velocity of frequency change in auditory events do exist. As well, this evidence reveals that the analogy between pattern processing in auditory and visual modalities exists not only at the level of behaviour, but also at the level of neural structures.

Evidence for "movement detectors" in the visual and auditory systems

The identification of neural units in the visual cortex of the cat, which respond selectively to the projection onto the retina of lines in specific positions and at specific orientations (Hubel & Wiesel, 1965), spawned numerous feature detection theories which proposed that even complex percepts were built of various combinations of simple component features registered in cortical units tuned to specific types of stimulation (e.g., Milner, 1974). Detectors capable of registering movement have also been located in the visual system. For example, in rabbits, retinal ganglion cells which respond to both the orientation and speed of moving images on the retina have been reported (Barlow, Hill & Levick, 1964).

Analogous neural units exhibiting specificity to auditory stimuli which "move" in frequency (i.e., frequency-modulated tones) have been found in the cochlear nuclei of cats (Evans & Nelson, 1966; Erulkar, Butler & Gerstein, 1968) and of rats (Møller, 1974), and in the inferior colliculus of cats (Nelson, Erulkar & Bryan, 1966) and rats (Clopton & Winfield, 1974). Orientation-specificity is revealed by cells in the cat brain which respond only to unidirectional frequency modulation. Such responses have been observed in both cortical units (Evans & Whitfield, 1964; Whitfield & Evans, 1965), and at the level of the superior colliculus (Nelson, Erulkar & Bryan, 1966). Whitfield and Evans (1965) identified cortical units in cats, which responded only to ramped frequency modulation. The units showed nearly no response to sinusoidal frequency modulation or to steady state stimulation within the sensitive region. While some units of this type responded throughout the whole cycle of ramp wave stimulation, most of the units responded only to part or all of one half of the ramp modulation cycle. In other words, the response was to the unidirectional frequency change in upward or downward ramps.

Vartanian (1974) suggested that the unidirectional specificity of certain neural units can be explained by the temporal and spatial characteristics of lateral inhibition from side-bands active upon the units. Through stimulation of unidirectionally-responsive units with steady-state tones, he observed asymmetries in lateral inhibition, depending on whether a tone at the centre frequency of the responsive region was preceded by a tone in the frequency region above or below the maximally responsive region of the unit. The strength of the inhibitory effect was also a function of the temporal separation between the tones. Hence, a glide passing through the unit's responsive region at a specific rate in one direction should inhibit the firing of the cell, while a glide at the same rate in the opposite direction should permit the unit to fire continuously as long as the stimulus remained within the cell's receptive field.

Behavioural evidence for the role of such neurons in processing unidirectional frequency sweeps has been shown by Kelly and Whitfield (1971). In their study, cats were trained on a shock avoidance paradigm. Frequency modulated ramps in one direction were used as the "safe" signal, and ramps in the opposite direction were used as the "warning" signal. The discrimination of the two signals was greatly impaired in animals with large bilateral ablations of auditory cortex, and could not be relearned to preoperative baseline levels. However, animals with the same types of lesions have been shown to retain their ability to discriminate high and low tones (Diamond & Neff, 1957). Therefore, cortical lesions may remove neurons required to identify unidirectional frequency modulation, while leaving lower level neurons required for simple pitch discrimination intact. The presence of some residual ability to discriminate the upward and downward sweeps was explained as a probable reflection of the action of intact neurons in the superior colliculus, capable of coding such sensory properties as the orientation of glides.

There is some evidence suggesting that similar neurons showing specificity to unidirectional frequency modulation exist in humans. Gardner and Wilson (1979) report selective adaptation to unidirectional glides, on a two-interval forced-choice task. They used unidirectional 75 msec linear sweeps, modulated between .96 and 1.04 KHz. Repeated presentation of sweeps in either direction yielded direction-specific adaptation. Therefore, adaptation to upward sweeps produced a temporarily elevated detection threshold for upward frequency glides, while adaptation to downward sweeps yielded a similar increase in the detection threshold for downward sweeps. The changes in threshold suggest that adaptation to upward or downward frequency glides resulted in the fatigue of detectors tuned specifically to glides of one orientation or the other. Presumably, the detection threshold for glides of that orientation was elevated until the appropriate neural units had recovered.

A possible function for a neural mechanism responsive to unidirectional glides in frequency is suggested by a finding reported by Delattre, Liberman and Cooper (1955). They showed that the direction of the second formant (F2) transition enables discrimination of different stop consonants. Hence, the selective response of neural units to upward or downward frequency-modulation may play a role in the discrimination of speech sounds.

Vartanian (1974) demonstrated that cortical units in rats, which were sensitive to unidirectional frequency modulation, were not only tuned to specific frequency ranges, but also to specific rates of frequency change. Hence, there is some evidence for a neural substrate capable of performing the functions of "pitch-motion" detection or the tracking of the "trajectory" or velocity of pitch change. Furthermore, the very fact that human subjects can discriminate subtle differences in the rate of ascent or descent in frequency of glides (Nábělek & Hirsh, 1969; Pollack, 1968) suggests that such mechanisms are present in the human auditory system. Although the auditory system could monitor differences in the rates of frequency change in glides of equal duration by comparing the differences between initial and terminal frequencies of the glides, rather than by comparing the actual velocities of frequency change, a finding by Tsumura, Sone and Nimura (1973) suggests that this is not the case. They found that detection tresholds for glided frequency transitions differed from those for frequency transitions which began and ended with steady state segments. This difference was taken to indicate two separate modes for processing frequency transitions. In the former case, they proposed that auditory mechanisms might actually track the rate of frequency change, while in the latter case, a comparison of the initial and terminal frequencies might have been used. Therefore there is some support for the notion that the auditory systems of animals and humans are directly responsive to the rate of frequency change in glided tones.

Possible influences of glides during the perceptual organization of auditory sequences

The preceding evidence suggests several hypotheses concerning the factors which should induce sequences of glided tones to group sequentially, or to segregate perceptually into separate streams. Both behavioural and neurophysiological evidence indicates that the auditory system is sensitive to the orientation of unidirectional frequency glides, even at very brief presentation durations. Therefore, one might hypothesize that orientation information should play an important role in determining whether glides should group together perceptually. Visual patterns composed of identical elements, some at one orientation and the others at another orientation, tend to split perceptually into a figure-ground relation (Beck, 1972). The elements at one orientation become a "figure", and the remaining elements become the "ground". Similarly, one might expect that a group of tones with one glide orientation might segregate perceptually from tones in the same sequence which had a different orientation.

Bregman's (1978a, 1978b, 1978c, 1978d) data leads to the alternate speculation that for sequences composed of glided elements "stream" formation would be a function of the distribution of events in different frequency ranges, much as it appears to be for sequences composed of steady state tones. His suggestion is that if a succession of tones occurs in two or more restricted ranges of frequencies, this provides evidence that a sequence should be segregated into two or more streams. Hence, one would anticipate that a sequence composed of alternating glides in different frequency ranges would segregate into separate streams, whereas one composed of glides in the same frequency range should group into a single sequential stream.

However, Bregman and Dannenbring's (1973) data suggests that continuity in the frequency transition between high and low tones in a sequence reduces the tendency for the sequence to "split" into high and low streams. Therefore, continuity in the frequency change between successive events in a sequence appears to reduce "stream segregation" effects. Thus, the auditory system may follow the "trajectory" of a glided transition, so that the glide can act as a "bridge" across disparate frequency ranges. A similar speculation follows from Jones' (1976) theory, which predicts that a sequence of events sharing a common "trajectory" (or continuity in the direction and velocity of frequency change) between them should be grouped perceptually into a single stream, since the velocity of pitch change would be constant throughout such a sequence. Hence, collinear frequency glides (i.e., glides aligned on the same trajectory when plotted on a logarithmic scale of frequency) might be predicted to group sequentially into a single stream despite the fact that they occur in different frequency ranges. If the auditory system tracks trajectories of frequency change, each glide would "point to" the subsequent glide.

The current study was designed to examine whether continuity in the frequency change in successive glides (i.e., common trajectory), similarity of the frequency range occupied by successive glides, or common orientations of frequency change between successive glides makes a stronger contribution toward maintaining a sequence of glided sounds in a single stream. Therefore, the study would permit the evaluation of the somewhat conflicting expectations discussed above.

The paradigm employed in the current study

In order to examine the roles of trajectory, frequency range, and orientation information in the perceptual grouping of sequential glided tones, the study used a modification of a paradigm first used by Bregman and Pinker (1978). The paradigm exploits the fact that synchrony between the onsets and offsets of sine tones at different frequencies causes the tones to "fuse" together (Dannenbring & Bregman, 1978; Rasch,1978). Several simultaneous sine tones are thus heard, not as a series of simultaneous events, but as a single complex tone with a "rich" spectral texture. Individual sine tones with asynchronous onsets or offsets (of more than 20 msecs) can be "heard out" of such mixtures, despite the fact that they overlap in time with each other. Thus the auditory system appears to use spectral changes over time in order to separate individual components out of a mixture. When an acoustic mixture containing energy at several frequencies is preceded or followed by energy isolated at only one of those frequencies (fx), the auditory system seems to assume that the fx energy does not belong with the other components of the mixture. Instead, it is heard as a separate pure tone against a background mixture formed of the remaining components. Perfectly simultaneous tones provide no such evidence, and are heard as a "fused" complex.

Bregman and Pinker (1978) utilized the tendency for simultaneous pure tones to "fuse" into spectrally "rich" units to show that a priori evidence can be used by the auditory system to decompose complex tones into their sine tone components. In their experiments, a complex tone (composed of two sine components at different frequencies) was rapidly alternated with a pure tone (or 'captor' tone) with an adjustable frequency. They found that when the frequency of the pure 'captor' tone was close to that of one of the components in the complex tone, the complex became decomposed. Rather than sounding like a single "rich" tone, its two sine components were heard as distinct elements. Furthermore, the component close in frequency to the "captor" tone tended to group into a sequential stream with the "captor". The second sinusoidal component of the complex tone was heard as a pure tone, in a separate stream. As the frequency of the adjustable tone was made more distant from the "target" tone in the complex, the streaming effect disappeared. Now, the simultaneous components in the complex remained "fused", so that the whole sequence sounded like the alternation of a pure and a "rich" tone.

Bregman and Pinker's study reveals an important competition between two auditory grouping tendencies. Simultaneous events with synchronous onsets and offsets show a "vertical" tendency, which draws events at different frequencies together into a single "rich-sounding" event. At the same time, sequential events at nearly the same frequency exhibit a mutual sequential attraction which tends to draw the events into a "stream". Their experiment showed that as the adjustable "captor" tone came close enough in frequency to the "target" component of the complex tone, the sequential grouping tendency between the pair of rapidly alternating "captor" and "target" tones became stronger than the "vertical" fusion between the two simultaneous components. As a result, the "target" tone was drawn out of the complex, and into a "horizontal" stream with the adjustable "captor" tone.

In the current study, a reliable measure of the relative strength of sequential (horizontal) grouping between successive frequency glides was required. To develop such a measure, Bregman and Pinker's (1978) paradigm was modified to meet the demands of the current study. Instead of steady state tones, a pure tone glide was rapidly alternated with a complex glide which consisted of two synchronous sine tones (modulated in parallel). Thus, by varying the relation between the single glide (or 'captor' glide) and the "target" glide in the complex, it would be possible to examine how that relation influenced sequential grouping between the captor and target glides. The strength of sequential grouping would be reflected by the resulting reduction in the "vertical fusion between the target glide and the glide in synchrony with it. Hence, a fairly sensitive measurement of the "horizontal" grouping strength between captor and target tones could be achieved.

EXPERIMENT I

Introduction

The first experiment was designed to examine the roles of trajectory, frequency proximity and orientation information in determining the sequential grouping of successive frequency glides. For the purpose of this experiment, the frequency proximity between the glides was defined by the distance in frequency between the average frequencies (on a logarithmic scale) of pairs of glides. As discussed above, the stimuli were analogous to those used by Bregman and Pinker (1978), but differed in that all sine tone components in the stimuli were frequency glides rather than steady state tones, as shown in Figure 1. In Figure 1a, 1b and 1c, a simple frequency-gliding sinusoid is followed by a complex tone formed from two frequency-gliding sinusoids Y and Z. The present experiment attempted to "capture" the glided target component Y imbedded in the two-component glide complex YZ by rapidly alternating the glided "captor" tone X with the glide complex. As in Bregman and Pinker's study, the perceived decomposition of the complex tone into its sine tone components would indicate that sequential streaming between captor (X) and target (Y) tones had occurred.

The relation between the captor tone and the target tone in the complex was manipulated in several ways: (1) To promote a sequential captor-target grouping based on trajectory information, the glided tones were situated in different frequency ranges so that they were collinear when plotted on a logarithmic scale of frequency (see Figure 1a). Thus, each captor glide

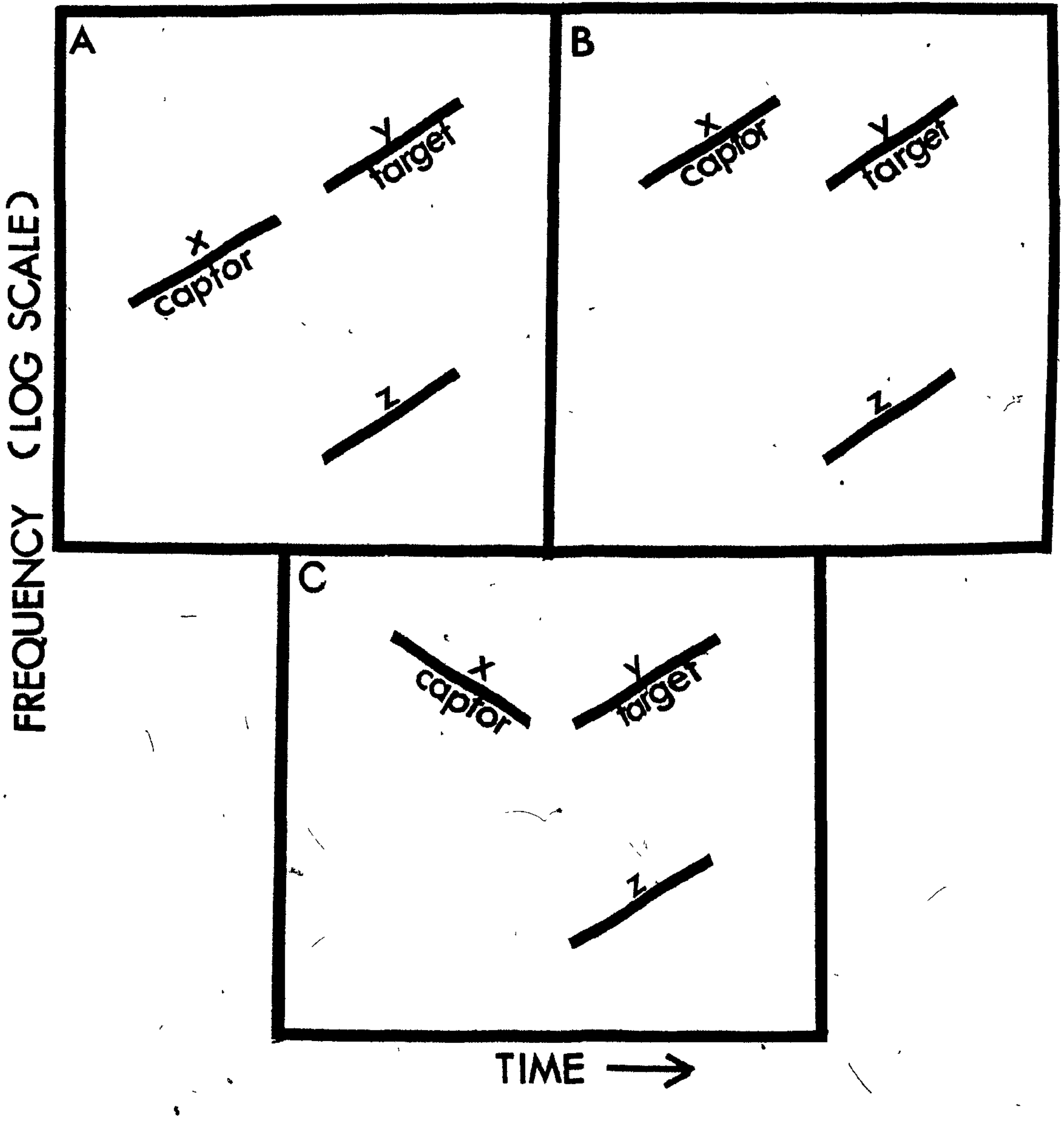


Figure 1. A graphic representation of stimulus patterns used to compare the effects of (A) common trajectories, (B) common frequency ranges and orientations, and (C) different orientations between captor (X) and target (Y) tones. The tone labelled Z is designed to "fuse" with X.

"pointed" upward or downward toward the subsequent target tone, which continued the trajectory created by the captor into a higher or lower frequency range. If the auditory system extrapolates from a trajectory of frequency change, this might tend to cause the captor and target to be organized into a sequential stream. (2) To determine whether or not the sequential grouping of successive glides is more strongly governed by such trajectory information than it is by the similarity between the frequency ranges occupied by the glides, a second condition was designed in which captor and target tones were not collinear, but instead glided across identical frequency ranges (see Figure 1b). Such a pattern might promote sequential grouping of captors (X) and targets (Y) on the basis of the frequency proximity between them. By comparing the decomposition effect in these two conditions, it could be determined whether "frequency proximity" or "trajectory" information was the more crucial determinant of the sequential grouping of captors and targets.

To determine whether the respective orientation of successive glides plays a role in promoting sequential grouping, or whether the sharing of a common frequency range between successive glides is sufficient to induce such grouping, a pair of captor and target tones was also made to glide across the same frequency range, but so that the captor had an opposite orientation to that of the target, as do the X and Y tones in Figure 1c. Therefore, by comparing the sequential (captor-target) grouping tendencies induced by patterns such as those in Figures 1b and 1c, possible influences on sequential grouping effects which were due solely to the respective orientations of captor and target glides could be isolated.

Finally, the effect of different tone durations was examined to determine whether any grouping tendencies (based on frequency proximity, trajectory, or orientation) were dependent on the rate of alternation between the sequential events. By using longer tone durations, the alternation rate of tones in a sequence was made slower. Shorter tones were used to produce a more rapid alternation rate. It was expected that the longer durations would allow more time for trajectory information to accumulate, since more time would be available in which to monitor the velocity of frequency change during each glide. Therefore, sequences consisting of the longer duration tones were expected to promote sequential grouping in the condition containing collinear glides, if the auditory system extrapolates from the trajectory of a gliding tone. Alternately, high alternation rates between events have been shown to encourage stream segregation based on frequency proximity (Bregman, 1978c; Dannenbring & Bregman, 1976; Van Noorden, 1975). The sequences composed of shorter tones might thus accentuate the sequential streaming of X and Y in conditions favouring "frequency proximity" effects, and in turn increase the decomposition of the YZ complex.

Method

Subjects. Twenty-four volunteers, all of whom were university students ranging in age from 20 to 28, were tested. All subjects reported having normal hearing.

Apparatus. Tones were synthesized digitally on a Digital Equipment Corporation (DEC) PDP-11/34 computer. The computer control of stimulus generation was enabled by the use of a software package entitled MITSYN, developed by Henke (1975). The package enables a complete range of auditory stimulus parameters to be specified, and permits the synthesis of polyphonic signals. Digital signals were transformed by a DEC AA11-K 12-bit digital-to-analog converter and output through a DEC Laboratory Peripheral Accelerator (model No. LPA11-K). High frequency components, produced during digital-to-analog conversion of the signals, were filtered out with a Rockland 851 filter, with a low-pass cutoff of 4000 Hz. All waveforms, frequencies, amplitude envelopes, and timing of tones were controlled by the computer.

Stimuli were recorded directly from one output channel of the computer onto audio tape using an AKAI GX-400D-SS tape recorder. The resulting monophonic tapes were played back on a Sony TC-65 tape deck amplified by a Pioneer SA-850011 amplifier. The output signal was presented binaurally over Sennheiser HD-414 stereo headphones. All presentations of experimental stimuli were conducted in an Industrial Acoustics 1202 audiometric chamber. To measure the loudness of the stimuli, a flat-plate coupler was used to connect the headphones to a General Radio Type 1551-C sound-level meter.

Stimuli. All tones used in stimulus sequences were sinusoidal. The frequency glides consisted of exponential changes in frequency. The amplitudes of all tones had roughly exponential attacks and decays of 10 msecs which rose and fell while the tones glided in frequency. All tones in the "slow alternation" condition had 230 msec durations (including the rise and fall times), and tones in the "rapid alternation" condition had 130 msec durations (including the rise and fall times). In both conditions, 20 msec silences separated the tones.

All tones glided across slightly less than one half an octave, and were situated in one of the five half-octave intervals between 256, 362, 512, 724, 1024 and 1448 Hz. The manner in which different combinations of tones gliding across each of the intervals were used to generate different experimental conditions is shown in Figure 2. The schematic representation of glides in the figure can be taken to represent tones with either 130 or 230 msec durations. To generate the 130 msec duration tones, glided tones which traversed the desired half octave interval in 150 msecs were gated off after 130 msecs. To produce 230 msec tones, frequency glides which traversed the desired half octave interval in 250 msecs were gated off after 230 msecs. Thus a sequence formed of successive tones of either duration always had 20 msec silences between the tones. The procedure also ensured that any sequence formed of a pair of ascending or a pair of descending tones from adjacent half octave intervals would align itself perfectly along a single trajectory when the tones were separated by 20 msecs (as shown by the pairs of tones labelled T and Y in the figure). The trajectory relation will be described more fully below.

To facilitate the subsequent description of stimulus patterns, tones will be referred to by the interval in which they fell, even though they were gated off shortly before having

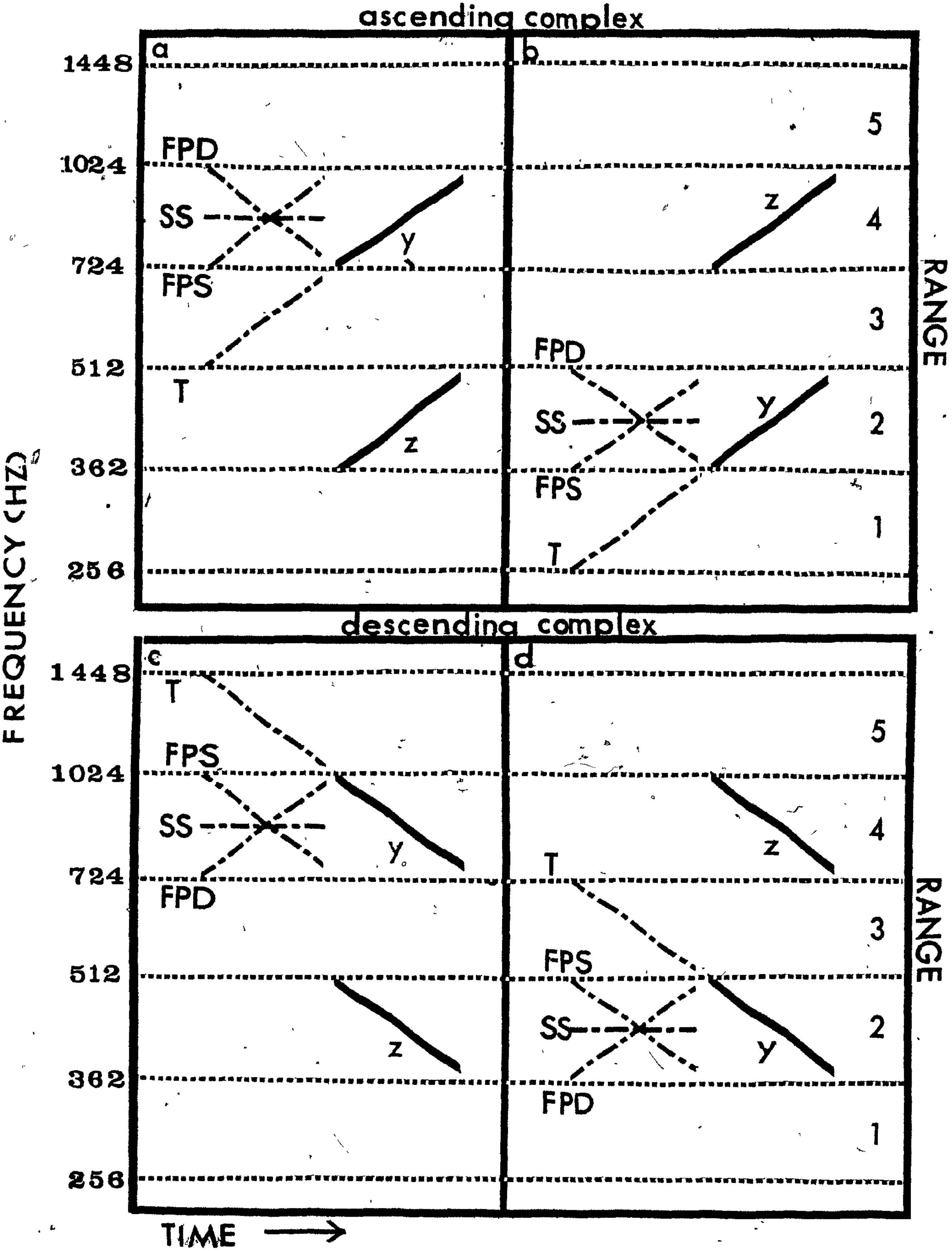


Figure 2. A graphic representation of the stimulus patterns used in Experiment 1. The captors (T, FPS, SS and FPD) are shown in dotted lines. Captors when the upper component of the ascending complex was the target (Y) are shown in "a" above. Captors for the lower component are shown in "b", and those for the descending complex are shown in "c" and "d".

traversed the entire interval. In the figure, the five half octave intervals are referred to as Range 1 to Range 5 from lowest to highest in frequency. Therefore, glides in either direction, falling within the lowest half octave interval, will be referred to as ascending or descending "Range 1" tones, those within the second lowest half octave will be referred to as ascending or descending "Range 2" tones, and so on. The pair of synchronous tones (Y and Z) which formed the complex glides were always situated in Range 2 and Range 4. To produce an upward-gliding complex, a pair of tones were glided upward across Ranges 2 and 4. Similarly, to produce a downward-gliding complex, a pair of tones were glided downward across the same two ranges. Thus, the pair of tones in each complex modulated upward or downward in parallel, at a separation of one octave. The resulting signal was a glided complex tone containing the fundamental frequency and first harmonic. The exact synchrony between onsets and offsets of the tones was found (during pretesting) to be sufficient to cause the tones to be perceived as a single "rich" frequency glide when heard in isolation from other tones.

In different conditions, either the upper or lower of the synchronous tones acted as the "target" tone. The tone that acted as "target" is always labelled Y, and the other tone Z. Several captor tone options are shown with dotted lines in the figure. Only one such option was used in each condition. First, the set of "captors" which had the upper tone of the ascending complex as a target will be described (see Figure 2a).

To produce a condition in which the captor glide was related by a common trajectory to tone Y, the ascending target tone

(Y) in Range 4 was preceded by an ascending pure tone glide crossing Range 3, so that the tones formed a perfectly straight line when plotted on a logarithmic scale of frequency. Such relations between captor and target tones will be referred to as "Trajectory" (T) groupings, and the appropriate captor is labelled T in the figure. By shifting the captor tone to Range 4, while leaving its duration and orientation identical, a captor which fell in the same range as and with the same glide orientation as the target tone was produced. The captor conditions in which the captor had the same orientation and range as the target will be referred to as "Frequency Proximity/Same Orientation" (FPS) grouping (see Figure 2a). To produce a glided captor whose orientation was opposite to that of the Range 4 target glide, but which fell in the same frequency range as the target, a tone gliding downward across Range 4 was used. This condition will be referred to as "Frequency Proximity/Different Orientation" (FPD) grouping (see Figure 2a).

The FPD condition also permitted a test of the hypothesis that the proximity between the terminal frequency of the captor and the onset frequency of the target tone might be a crucial factor in attracting the target tone into a sequential grouping. FPD captors were situated so that the separation in frequency between the end of the captor and the beginning of Y was identical to the separation in frequency (on a logarithmic scale) between the end of the Trajectory (T) conditon captor and the beginning of the subsequent "target" (see Figure 2a). In both conditions, this difference in frequency was relatively small. On the other hand, in the FPS condition, the terminal frequencies of captors and initial frequencies of target tones were very disparate. Therefore, should T and FPD conditions show more evidence than the FPS condition that sequential streaming of captor and target tones was occurring, then it would suggest that the separation between the terminal frequency of the captor and the initial frequency of Y did play a role in causing a sequential attraction between the pair of tones. In addition, any differences between results on T and FPD conditions could not be attributed to inequalities in the frequency separation between the end point of the captor and onset point of the subsequent tone, since the separation in the two conditions was identical.

Those T, FPS and FPD captors which were designed to attract the lower component of the ascending complex are shown in Figure 2b. The captor tones were related to the target in the same manner as were the captors described above. Figures 2c and 2d show how these three types of captors were created so as to capture either the upper or lower glide from a descending complex. The ascending and descending complex conditions were symmetrical to each other. To produce each stimulus sequence, one of the captors was repeatedly alternated with one of the complex glides.

It should be pointed out that there was a slight imbalance between various T conditions. Remember that each condition shown in Figure 2 should be conceived as being a continuous alternation of the captor tone and the complex glide YZ. Note that in the T conditions shown in Figures 2a and 2d, the captor (T) falls between the ranges occupied by tones Y and Z. Therefore, continuous cycles of the pattern not only placed T and Y on the same trajectory, but also placed the Z and T tones of consecutive repetitions of the pattern on collinear paths (see Figure 3a). In the T grouping shown in Figures 2b and 2c, T is above or below the ranges in which Y and Z fall. These patterns therefore allow only tones T and Y to become aligned on a common trajectory (see Figure 3b). Nevertheless, in the T condition both organizations (T with Y and Z with T) potentiate the decomposition of the mixture, if the auditory system extrapolates from tone T to tone Y (or from tone Z to tone T) as the pattern is repeatedly cycled.

In order to keep the trajectory relation perfect on both slow and fast alternation conditions the longer duration glides had to be made to cover a slightly larger frequency range than short duration glides. Were the exactness of the range covered by both long and short duration glides not sacrificed in this manner, then the trajectory between T condition captors and targets would have had to be imperfect at one of the alternation rates, (see Figure 4). The dotted line in the figure shows the imperfect trajectory (extrapolated from the captor glide) which would have resulted in the T condition if the longer duration captor tone was made to cover the same frequency range in 230 msecs as that covered by the shorter duration captor in 130 msecs. A way to alleviate this problem would have been to increase the temporal separation between the captor and target tones on the longer duration condition. However, since such an increase in time between tones might also reduce the tendency for them to group sequentially, by making them perceptually more separate, this alternative was not considered viable.

The size of the excursion in frequency (on a logarithmic

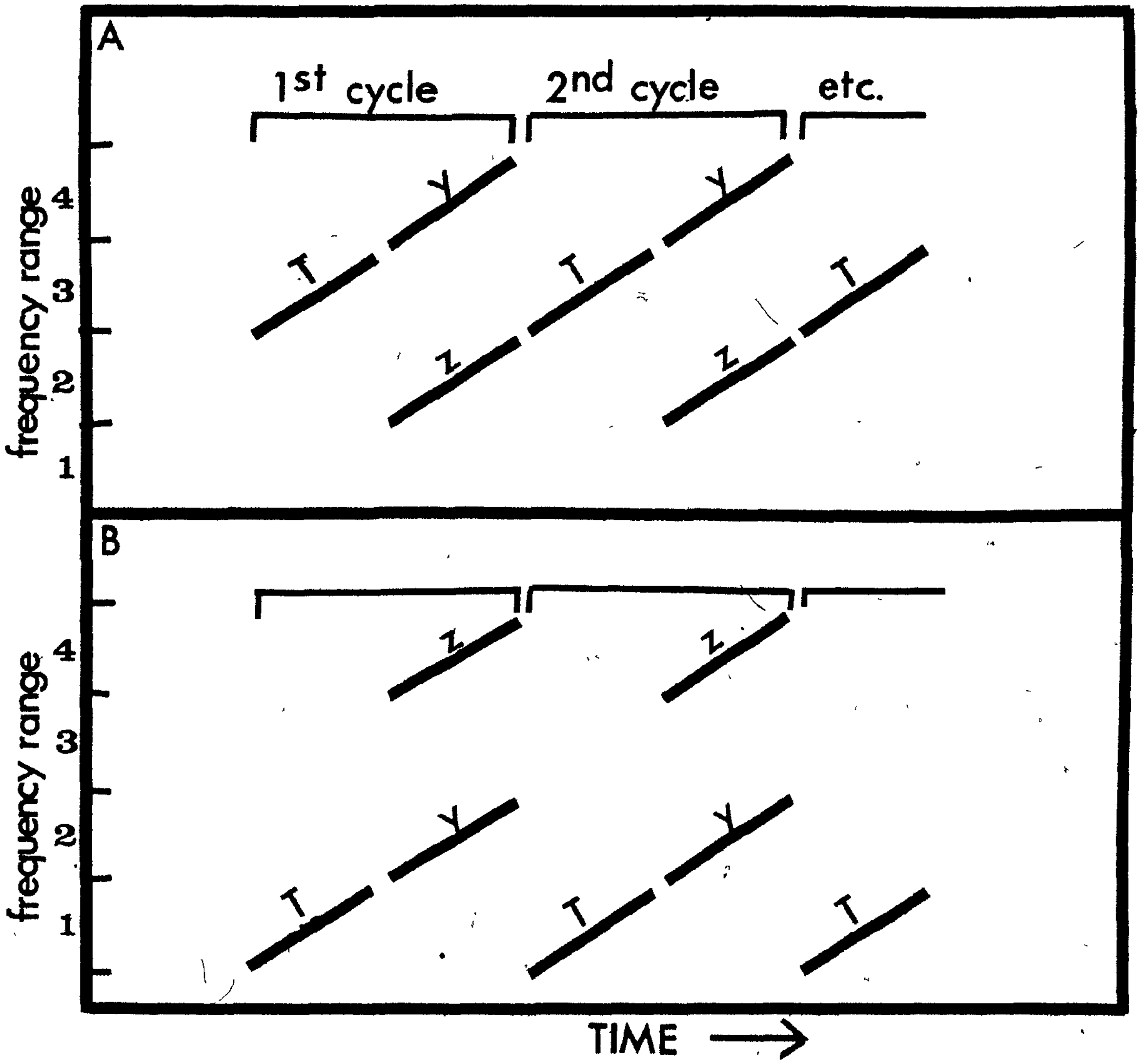


Figure 3. A representation of the difference between various "Trajectory" groupings. In (A) above a trajectory set by T is continued by Y during the first cycle, and that set by Z is continued by T in the second cycle. In (B) only T and Y are aligned on a common trajectory.

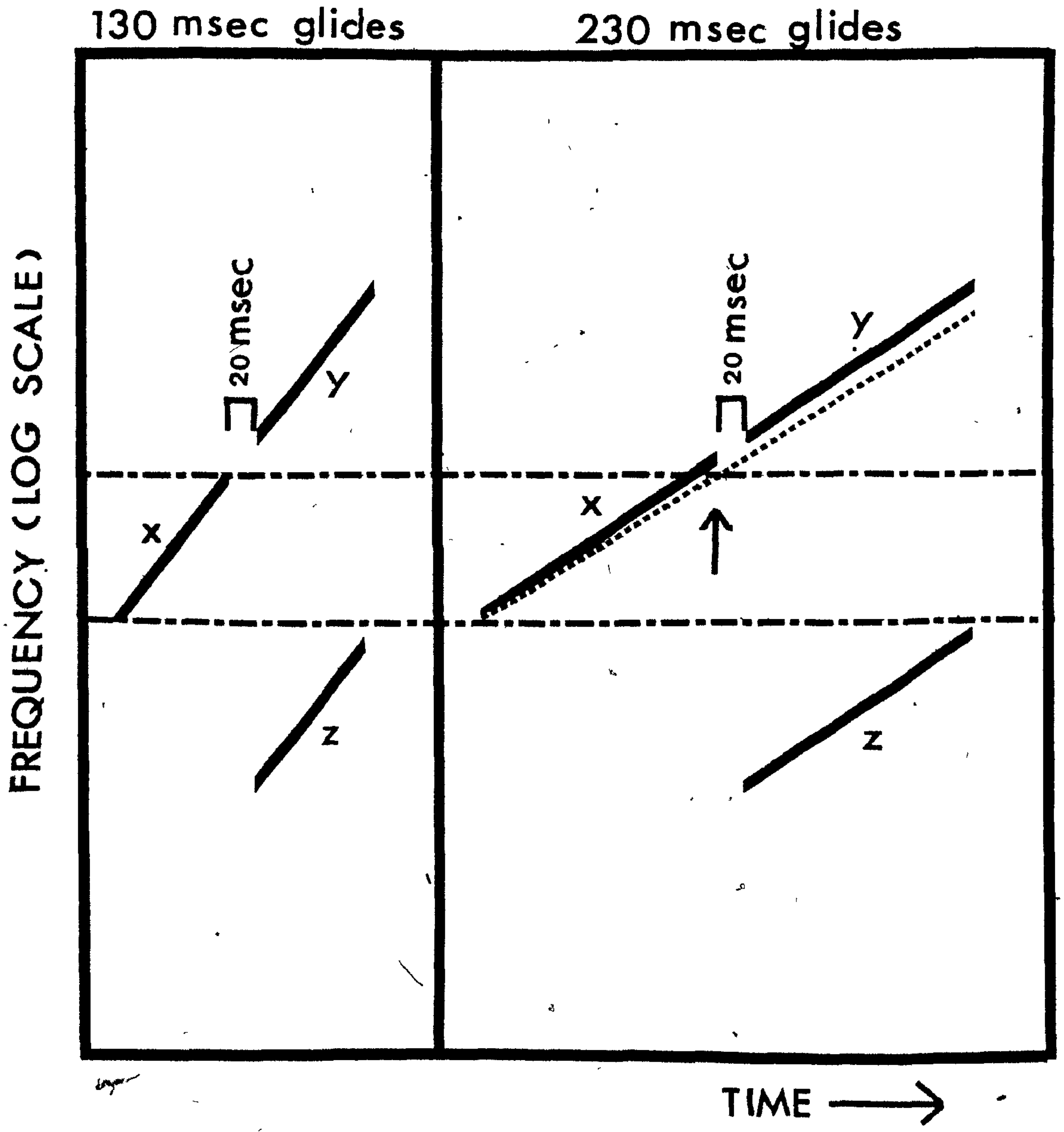


Figure 4. A representation of the faulty trajectory which would have resulted between X and Y had X covered the same frequency range in both "alternation rate" conditions.

scale) of ascending and descending glides of each duration was identical. However, ascending glides began at the lower frequency limit of the appropriate half octave interval, decaying just prior to reaching the upper limit, and descending glides began at the upper limit of the appropriate half octave interval, decaying just prior to reaching the lower limit (see Figure 2). The procedure, which was used because it facilitated the design of collinear glides, caused the actual ranges crossed by ascending and descending glides to be slightly different. This meant that the frequency range covered by FPD captors was slightly different from the range covered by the corresponding target tones which had opposite orientations. While in the FPS condition, captors and targets had the same orientation and therefore covered the same frequency range, FPD captors covered slightly higher frequency ranges in the ascending complex conditions (see Figures 2a and 2b), or slightly lower ranges in the descending complex conditions (see Figures 2c and 2d) than the range covered by the respective target tones. Thus any differences between FPS and FPD conditions would confound orientation effects with a slight inequality in the frequency proximity between captor and target tones across the two conditions. However, such a small discrepancy was considered unlikely to create any observable differences between the conditions. (Nevertheless, Experiment II included a replication of the FPS and FPD conditions and such discrepancies between the tones' ranges were avoided). Although the procedure used to generate the glides caused such small inaccuracies, they were considered negligible, and the procedure also permitted great accuracy on the trajectory conditions. It was therefore considered justified to sacrifice accuracy on the frequency dimension somewhat in favour of accuracy on the trajectory dimension. The precise frequency ranges traversed by all frequency glides used are shown in Table 1.

A final condition, using steady state tones as captors, was introduced to determine whether "frequency proximity" effects required the alternating tones to be qualitatively similar (i.e., glides), or whether the presence of a nonglided captor tone with energy inside the frequency range encompassed by the target glide would be sufficient to induce sequential captor-target grouping. This condition will be referred to as "Steady State Control" (SS). Captor tones in the SS control condition were situated at the middle frequency (on a logarithmic scale) of the interval in which the target tone fell. Therefore, to attempt to "capture" the upper tones in either ascending or descending complexes, the SS captors were set at 861 Hz, and to attempt to capture lower tones in either complex, the captors were set at 431 Hz. By setting the frequency values for the SS captors at the midpoints of the intervals in which upper and lower tones in each complex were situated, the SS captors would always be equally close to the middle frequencies of target glides regardless of whether the targets were ascending or descending, despite the slight discrepancy between the ranges subtended by the ascending and descending glides (see Figure 2). Should the SS condition show effects of sequential grouping (i.e., decomposition of the two-tone complex), then it would be clear that the qualitative similarity between tones in a sequence (i.e., whether they are glided or

Table 1: Frequency ranges subtended by the frequency glides in Experiment I

Tone Duration:	SHORT (130 msecs)		LONG (230 msecs)	
Orientation:	Ascending	Descending	Ascending	Descending
RANGE 1	256 - 339 Hz	-	256 - 347 Hz	-
RANGE 2	362 - 480	512 - 386	362 - 490	512 - 378
RANGE 3	512 - 679	724 - 546	512 - 693	724 - 535
RANGE 4	724 - 960	1024 - 773	724 - 981	1024 - 756
RANGE 5	-	1448 - 1092	-	1448 - 1069

not) is not required before sequential grouping will occur.

By virtue of the fact that "Trajectory" captors fell in a different frequency range from the "Frequency Proximity (FP) captors, i.e., FPS, FPD and SS (see Figure 2), differences between the T condition and the others were confounded with possible effects due simply to the frequency range in which each captor fell. However, the fact that the stimuli were designed so that each set of captors appeared in both high and low frequency ranges addressed this problem. Since within each captor condition (i.e., T, FPS, FPD and SS) two captors appeared, one in the high range and one in the low range, any simple effects of frequency range could be discriminated from effects due to the captor type itself.

It can be seen that sequential captor-target grouping, in any of the patterns shown in Figure 2, would be antagonistic to the fusion, or vertical grouping of Y and Z. Therefore, should any of the patterns represented in the figure be heard as the alternation of a pure and complex frequency glide, it could be inferred that tones Y and Z had remained "fused", and thus, that the sequential captor-target attraction was insufficient to overcome the tendency for the synchronous tones (Y and Z) to fuse together. However, should sequential streaming of captors and targets occur, then the two-component complex would become decomposed into its separate sine tone elements. As a result, the pattern should be perceived as containing two streams: one consisting of the rapid alternations of captor and target tones, and the other consisting of slower repetitions of Z-Z-Z-, etc. Thus, the percept obtained in each case served as a criterion which listeners could use to evaluate the sequential grouping tendency in each condition.

Four stimuli were generated for use in practice trials. These stimuli were analogous to the stimuli labelled T and FPS in Figures 1a and 1b respectively, and SS and T in Figures 1c and 1d respectively. However, all frequencies of tones were shifted somewhat from those used in actual testing to avoid specific carryover effects. Therefore, the practice trials were similar, but not identical to the respective experimental stimuli. The first two practice trials described were presented at the 130 msec event rate, and the latter two were presented at the 230 msec event rate.

Audio tapes were prepared containing four different randomizations of 32 trials each. The 32 trials represented the 16 conditions shown in Figure 1, (consisting of 4 levels of 'captor type', 2 levels of 'orientation of the complex glide', and 2 levels of 'frequency range of captor'), each synthesized at two "alternation rates". Each trial consisted of a warning beep, followed two seconds later by 30 repetitions of the pure-tone/complex-tone sequence. A six second silence separated the trials. The four practice trials were recorded at the beginning of each randomization.

While the frequency response of the auditory system is not flat over the whole frequency range covered by the tones in the stimuli described above (Fletcher & Munson, 1933), no attempt was made to equalize the subjective loudness of the tones over the whole range. It was felt that since the Trajectory and Frequency Proximity captor tones appeared in high and low ranges, any effects resulting from differences in loudness between tones at different frequencies could easily be detected and separated from other effects. Stimuli were presented so that the patterns had a mean intensity of approximately 80 dB SPL when measured on the sound-level meter.

Procedure. Subjects were treated individually, and testing was conducted in the audiometric chamber. All instructions were presented on a typewritten form. Subjects were informed that they were participating in an experiment concerned with the perceptual grouping of auditory patterns containing frequency glides. They were told that on each trial they would hear 30 repetitions of a pattern consisting of alternations of a pure tone frequency glide and a pair of simultaneous glides. A diagram depicting such a pattern accompanied the instructions. The term frequency glide was defined as a sound which consisted of a continuous upward or downward shift in frequency. The alternate percepts which the patterns could produce (discussed earlier) were described to the subjects with the aid of diagrams. The subjects were told to call one such percept "fused" (when only two tones were audible on each cycle of the pattern) and the other "decomposed" (when three tones were audible on each cycle). To indicate their judgments of the strength of fusion or decomposition heard in each sequence, subjects were instructed to rate each sequence on a response sheet containing a 7-point scale for each trial. The scale was labelled FUSED at the low end, and DECOMPOSED at the high end. They were told to give a rating of four when the sequence sounded ambiguous, or seemed to fluctuate between the two types of groupings. The instructions also warned that the sequences would be presented at two speeds, and that care should be taken to avoid confusing the changes in speed for rhythmic differences between the possible percepts which could result from each sequence.

Once the subject had read the instructions, the four practice trials were presented, and the subject's ratings of these trials examined. If the subject failed to recognize any differences between the four patterns, the qualitative differences between "fusion" and "decomposition" were described again and the same four practice trials repeated. All but two subjects reported hearing differences between the sequences by the second presentation of the practice trials. The two subjects who could hear no differences were rejected from the experiment.

All experimental conditions were presented (32 trials) with no interruption. Equal numbers of subjects were tested with each of the four randomizations of trial orders. Subjects were randomly assigned to one of the four orders, with the restriction that only six subjects were run on each randomization.

The data measure and analysis

Subjects' ratings were sorted into experimental conditions, and analyzed in a four-way analysis of variance with repeated measures on all factors. The analysis tested two levels of "alternation rate", two levels of the orientation of the glide complex (i.e., upward or downward), the two "frequency ranges of captors" (i.e., whether the captors were from the high or low range sets), and the four "captor types" (i.e., T, FPS, FPD and SS). Two subjects accidentally omitted one response each during the testing session. The missing data points were replaced with cell means. To compensate for the estimation of missing values, two degrees of freedom were subtracted from all error terms (see Table 2).

Results

Figure 5 shows means for each of the "captor type" factors when presented in both high and low frequency ranges, as well as overall means at each level of "captor type". While the ratings on each pair of FPS, SS and T captor conditions were nearly identical regardless of whether the captor was in the high or low range, the FPD condition appears to have produced higher ratings of decomposition when the captor was high rather than low range. The difference was sufficient to produce a significant interaction of "frequency range of captor" and "captor type" factors, $F(3, 67) = 2.84$, $p < .05$. The values at different levels of the "captor type" factor for each level of the "range of captor" factor shown in Figure 5 suggest that decomposition ratings were generally low in the "Trajectory" and FPD conditions (i.e., relatively fused), while FPS and SS captors yielded higher ratings (i.e., more decomposed). The overall means for captor conditions (shown in dotted lines) suggest a similar pattern. The main effect of "captor type" was very reliable, $F(3, 67) = 9.73$, $p < .0001$.

However, to analyze the main effect of "captor type" in a form uncontaminated by influences from the "frequency range of captor" factor, tests of simple main effects using Satterthwaite's degrees of freedom and pooled error terms were performed on the means resulting from the interaction of the two factors (Winer, 1962). The simple effect of "captor type" when only high range captor means were tested was significant, $F(3, 125) = 3.78$, $p < .05$. Similarly, the simple effect of "captor type" was

Table 2: Summary table of the 4-way analysis of variance (alternation rate x orientation of glide complex x range of captor x captor type) performed on decomposition ratings

Source	SS	df	MS	F	p
Alternation Rate (A)	49.68	1	49.68	5.63(5.14)$^{+}$	0.0263(<.05)$^{+}$
Error	202.82	23(21)$^{+}$	8.82(9.66)$^{+}$		
Orientation of Glide Complex (O)	4.56	1	4.56	1.42(1.30)	0.2450(n.s.)
Error	73.69	23(21)	3.20(3.51)		
A x O	19.01	1	19.01	4.03(3.68)	0.0566(n.s.)
Error	108.44	23(21)	4.71(5.16)		
Range of Captor (R)	23.67	1	23.67	3.57(3.26)	0.0715(n.s.)
Error	152.47	23(21)	6.63(7.26)		
A x R	15.99	1	15.99	4.27(3.90)	0.0503(n.s.)
Error	86.18	23(21)	3.75(4.10)		
O x R	2.37	1	2.37	1.08(0.99)	0.3096(n.s.)
Error	50.48	23(21)	2.19(2.40)		
A x O x R	0.11	1	0.11	0.03(0.03)	0.8609(n.s.)
Error	83.22	23(21)	3.62(3.96)		
Captor Type (CT)	184.34	3	61.45	10.02(9.73)	0.0000(<.0001)
Error	423.23	69(67)	6.13(6.32)		
A x (CT)	5.76	3	0.92	0.65(0.63)	0.5842(n.s.)
Error	203.03	69(67)	2.94(3.03)		
Ox (CT)	7.74	3	2.58	0.75(0.73)	0.5251(n.s.)
Error	236.99	69(67)	3.43(3.54)		
A x O x (CT)	22.08	3	7.36	2.26(2.19)	0.0895(n.s.)
Error	225.10	69(67)	3.26(3.36)		
R x (CT)	29.35	3	9.78	2.92(2.84)	0.0400(<.05)
Error	230.92	69(67)	3.35(3.45)		

Table 2--Continued

Source	SS	df	MS	F	p
A x R x (CT)	5.89	3	0.96	0.49(0.48)$^{+}$	0.6899(n.s.)$^{+}$
Error	275.86	69(67)$^{+}$	4.00(4.12)$^{+}$		
O x R x (CT)	0.56	3	0.19	0.06(0.06)	0.9789(n.s.)
Error	202.54	69(67)	2.94(3.02)		
A x O x R x (CT)	6.16	3	2.05	0.72(0.70)	0.5438(n.s.)
Error	196.88	69(67)	2.85(2.94)		

$^{+}$ Adjusted for missing data points

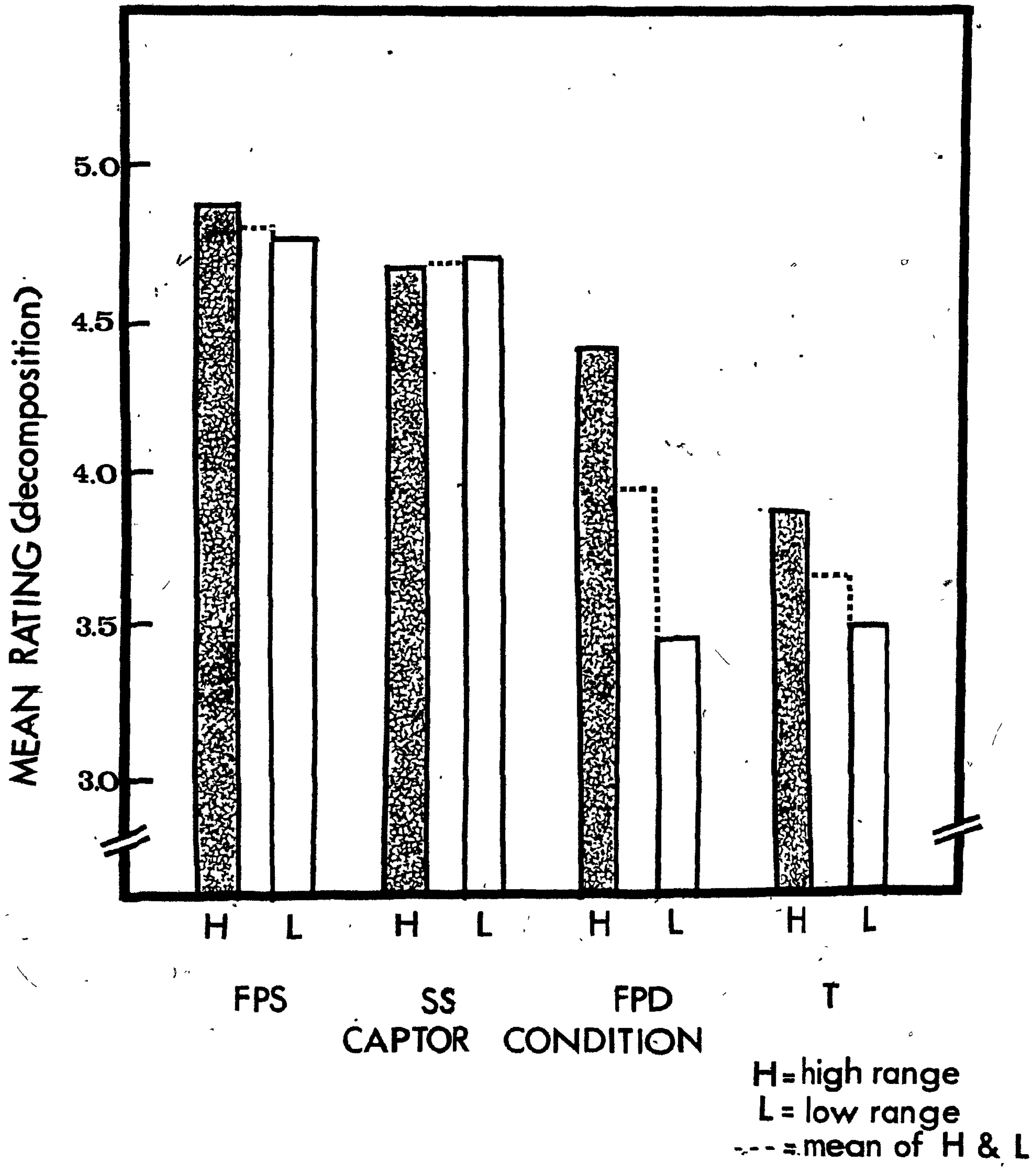


Figure 5. The mean ratings (decomposition) at each level of the "captor type" factor for both high and low range captors. The overall mean at each level of the "captor type" factor is shown in dotted lines.

significant on a test involving only the low range captor means, $F(3, 125) = 11.25$, $p < .01$. None of the differences between the pairs of mean, (i.e., high and low range captors), at each level of "captor type" approached statistical significance on tests of the simple effect of "range of captor", except for the differences between high and low range captors in the FPD condition, $F(1, 80) = 10.82$, $p < .01$. Although results in the FPD condition appeared to be affected by the "range of captor" factor, the fact that the main effect of "captor type" produced such a high F value, relative to that for the interaction of "captor type" and "range of captor" factors, in conjunction with the fact that the simple effect of "captor type" was significant at both levels of the "range" factor, suggests that the main effect of "captor type" reliably reflects true differences between captor conditions. It is, therefore, suggested that the main effect of "captor type" does not misrepresent differences between the conditions by obscuring important effects which result from the interaction of "captor type" and "range of captor" factors.

A set of Tukey tests were performed on pairs of means contributing to the main effect of "captor type", to examine the differences between specific captor conditions more fully. To determine whether or not the conditions in which captors and targets fell in the same frequency range (i.e., FPS, SS and FPD) yielded higher ratings than did the T condition, in which the captors and targets fell in different ranges, each of the former means was tested against the T condition mean. The FPS and T condition means differed significantly, $Q(4, 67) = 6.28$, $p < .01$. Similarly, the difference between SS and T means was significant,

$Q(4, 67) = 5.71$, $p < .01$. However, FPD and T means did not differ reliably, $Q(4, 67) = 1.40$, n. s. To examine whether or not the lower score on the FPD condition could be attributed to the fact that the captor glided in the opposite direction to the target tone on that condition rather than in the same direction as in the FPS condition, the difference between FPS and FPD means was tested, and found to be significant, $Q(4, 67) = 4.88$, $p < .01$. Similarly, the difference between SS and FPD conditions was found to be significant, $Q(4, 67) = 4.31$, $p < .05$. The difference between FPS and SS means was not significant, $Q(4, 67) = 0.57$, n.s.

Table 3 shows means for each "captor type" at both levels of "alternation rate". Note that the mean ratings on all captor conditions are higher (i.e., more decomposed) in the fast alternation rate condition. It appears then that the faster alternation of tones (or shorter tones) produced uniformly more decomposition than did the slower alternation condition (or longer tones) regardless of which captor type was involved. The main effect of "alternation rate" was significant, $F(1, 21) = 5.14$, $p < .05$. The interaction of "alternation rate" and "captor type" factors was not significant, $F(3, 67) = 0.63$, n. s.

Discussion

Since subjects probably adjust their use of a judgment scale to match the range of differences present in any set of stimuli, the actual numerical ratings reported above can be taken to indicate no more than a relative judgment of the fusion or decomposition heard in each sequence. Nevertheless, the differences between conditions can be used to make several inferences about the strength of sequential grouping in various conditions.

Table 3: Mean ratings on each captor condition at each level of the "alternation rate" factor

	CAPTOR TYPE			
	FPS	SS	FPD	T
Rapid (130 msecs/ event)	5.03	5.12	4.16	3.86
Slow (230 msecs/ event)	4.61	4.31	3.71	3.50

The tests of simple effects performed on the means that contributed to the interaction of the "frequency range of captor" with the "captor type" factors suggest that the only significant component of the interaction occurred in the FPD condition. High range FPD captors produced higher ratings of decomposition than did low range captors (see Figure 5), suggesting that the sequential attraction between high range FPD captors and targets (as shown in Figures 2a and 2c) was stronger than that between low range FPD captors and targets (as shown in Figures 2b and 2d). The target tone for high range captors was always the upper harmonic of the two-tone complex, whereas the low range captor's target was the fundamental. One might, therefore, explain the result by supposing that it is easier to extract a harmonic from a complex mixture, than it is to extract the fundamental. Such a phenomenon has been observed previously (Bregman, A. S. Personal communication, June, 1980).

To be consistent with such a supposition, high range captors should have yielded stronger decomposition effects on all captor conditions. However, the ratings on FPS, SS and T conditions do not show such range-related effects. Perhaps the inconsistency is due to the fact that the sequential attraction between FPS and SS captors and targets, which yielded high decomposition ratings, was sufficiently strong to extract either the fundamental or harmonic with equal facility. On the other hand, T captors, which yielded low decomposition ratings, may have produced such a weak sequential attraction to their target tones that neither high or low range captors showed a decomposition effect. According to this reasoning, FPD captors must have exerted only a weak "pull" on the subsequent target tones. Such a weak attraction would be influenced by how strongly the fundamental or harmonic of the two-tone complex resisted being extracted from the complex. The result could, therefore, be interpreted as suggesting that FPD captors exerted a weaker sequential attraction to targets than did FPS or SS captors, but a stronger attraction than did the T captors. Thus, while the difference between FPD and T conditions was not statistically significant, the preceding explanation leads to the deduction that the T condition may have yielded slightly lower decomposition than did the FPD condition.

Similarly, the overall mean rating on the T condition was significantly lower than that on both FPS and SS conditions. Therefore, there is evidence to suggest that the T condition captors grouped more poorly with the targets than did the FPS, SS and possibly the FPD condition captors. In the FPS, SS and FPD conditions, captors fell in the same frequency range as their targets, while the T captors fell in different ranges from their targets. Therefore, the lower judgments of decomposition in the T condition suggest that the frequency proximity between captors and targets, assessed by comparing the similarity between the frequency ranges occupied by the successive tones, generally dominated over trajectory information in determining the sequential grouping of the tones. The proximity between the offset frequency of the captor and the onset frequency of the target was clearly of minor importance in determining sequential grouping effects. Otherwise, the T and FPD conditions in which the distance in frequency between the captor's offset and the target's onset was smaller than the corresponding distances in FPS and SS conditions should have shown the highest rather than lowest ratings.

As a reflection on the auditory system's role as a predictive mechanism, these results imply that a priori events "prime" the auditory system to predict the occurrence of subsequent events in the same or a similar frequency range, rather than events which are continuous with the previous event. As a result, the captors generally attracted targets best when they were in the same frequency range as their targets. Presumably, the captors that fell in the same range as the target tone in the complex were sufficient to "tune" the auditory system to the range in which the target fell, permitting it to isolate the target from the glide complex. The current findings, therefore, support the notion that frequency proximity between events is a major determinant of sequential grouping when stimuli contain frequency glides, much as it is for the grouping of steady state tones (e.g. Bregman & Campbell, 1971; Bregman & Rudnicky, 1975; Van Noorden, 1975). With respect to glided sounds, the proximity between successive tones appears not to be determined by comparing the temporally closest frequency components of the tones (i.e., the terminal frequency of one tone and the initial frequency of the subsequent tone), but by comparing more global descriptions of the frequency characteristics of each tone.

Any "priming" effect of trajectory which would sensitize the system to events which continue the path set by previous events, seems to be smaller than the frequency proximity effect. Thus in the Trajectory condition, the glide complex was apparently heard as being more fused. The relative inferiority of the T condition in inducing the decomposition of the two-tone complex suggests that the auditory system failed to extrapolate from the trajectory of the captor tones so as to capture the subsequent target tones (which showed good continuity with the captors). Instead, the condition in which captors and targets fell in different ranges yielded a low level of decomposition (i.e., low sequential grouping of sinusoidal components), as one might predict if relying solely on a frequency proximity-based description of the auditory grouping mechanism. It appears then that the auditory system characterized each glide as exciting a specific frequency range, rather than as "pointing to" a specific range.

If the responsive region excited by a tone led only to the prediction that a subsequent event should excite that same region, then one would expect FPS and FPD conditions to have yielded almost identical results, since "captors" and "targets" shared virtually the same frequency range in both conditions. Since the FPS condition yielded a significantly higher mean rating (i.e., greater decomposition) than that on the FPD condition, it appears that another factor was active. Possibly, the fact that FPD captors and targets had opposite orientations biased against sequential grouping, despite the common frequency range of both tones. Therefore, the "similarity" of successive tones, in terms of their glide orientation, may play a role in sequential grouping decisions. One might, therefore, suppose that information extracted by "orientation-specific" glide detectors, analogous to those identified by Evans and Whitfield (1965) is used during sequential grouping operations.

Therefore, a possible argument consistent with all the findings of the current experiment is that the auditory system determines both the frequency range and the orientation of frequency glides. Thus, while the frequency proximity between successive glides influences how well they will group sequentially, the correspondence in orientation of the glides may modulate frequency-based grouping decisions. As a result, glides with corresponding orientations would group better than those with noncorresponding orientations when the frequency proximity between the glides (averaged on a logarithmic scale) was equal. On the other hand, when the glides were far apart (as in the T condition), even those of the same orientation would not readily group into a sequence, since the frequency-based component of the grouping decision would bias against sequential grouping of the tones.

It should be recalled that the apparent difference due to the correspondence in orientation between captors and targets in the FPS and FPD conditions is confounded with the slight discrepancy between the frequency ranges covered by FPD captors and target tones (see Stimuli section). Since such a discrepancy could have biased in favour of frequency-based streaming in the FPS condition, in which the captors and targets shared identical ranges, the difference must be interpreted with caution. However, should the small discrepancy between the ranges covered by captor and target tones in the FPD condition have produced the lower ratings in that condition, it is difficult to comprehend how the overall rating on the T condition, in which the captor and target tones shared no common range at all, is only slightly lower than that on the FPD condition, in which the captors shared a large range in common (see Figure 5). The differences in ratings between FPS, FPD and T conditions are, therefore, not proportional to the differences between the frequency ranges covered by captors and target tones in the three conditions. Therfore, it is suggested that the difference between FPS and FPD conditions cannot be attributed to the difference between the frequency ranges of the captors and targets in the latter condition. A more reasonable explanation is the one offered above, that the relative orientation of the glides accounts for the difference.

An alternate explanation for the differences between captor conditions in the current experiment follows from the Brady, House and Stevens (1961) finding that the subjective pitch of a short duration frequency glide is identified with the terminal frequency of the glide. If the pitch of a glide is determined primarily by its terminal frequency, then captors and targets in the FPS condition, which terminated on the same frequency, would have had subjectively identical pitches. Similarly, the SS captors would be perceived as having nearly the same pitch as the targets, since they fell only slightly above or below the terminal frequency of the targets. The FPD and T captors would be perceived as having progressively more disparate pitches from the pitches of their targets, since the terminal frequencies of captors and targets became progressively further apart in the FPD and T conditions (see Figure 2). If auditory grouping mechanisms induce greatest sequential grouping between tones that have the most similar subjective pitches, then the differences in pitch discussed above might account for the observed results. An experiment to be reported below attempted to isolate whether or not the correspondence between terminal frequencies of successive glides could have been a relevant factor in inducing sequential grouping between the glides in the current study.

While the "alternation rate" factor did not interact with any other factors, it did yield a significant main effect. The higher alternation rate produced uniformly more decomposition across all captor conditions than did the lower alternation rate (see Table 3). Although the tones in each "alternation rate" condition differed slightly in terms of the size of their frequency ranges as well as in terms of their durations (see Figure 4), there is no apparent reason to suppose that such differences in frequency range would cause such an effect. Therefore, the difference can probably be attributed to well-known speed-related grouping tendencies. Generally, it has been found that increasing the speed of alternation between high and low frequency tones also increases the tendency for the tones to segregate into separate sequential streams (e.g., Dannenbring & Bregman, 1976; Van Noorden, 1975). For this reason, the "rapid alternation" rate in the current experiment should have been expected to produce more frequency based sequential streaming, and hence more decomposition than the "slow alternation" rate. However, when the experiment was designed, it was reasonable to assume that the hypothetical trajectory effect would accumulate more evidence when tones were longer and therefore one would expect slow rates to favour such an effect. It is interesting to note that, somewhat counter-intuitively, the slow alternation rate did not improve sequential grouping tendencies in the Trajectory condition relative to those at the rapid alternation rate (see Table 3). The uniformity of the decrease in decomposition at the slow rate, therefore, supports the notion that the trajectory effect did not occur at all. Instead, the result seems to reveal a reduction in only one grouping tendency--based on the similarity in frequency range occupied by successive events and influenced by the rate of alternation of events--which was effective across all conditions. In line with this reasoning, the captors and targets in the T condition, which were further apart in frequency than those in any other condition, should always have shown the weakest sequential grouping, but even more so at the slow alternation rate.

It should be pointed out that it is possible that the stimuli used in the current experiment may have biased slightly against trajectory effects. Remember that each stimulus was a repeating cycle. Note in Figure 2 that it is possible for the Trajectory captors, which always occupied their own separate frequency range, to have attracted each other into a single horizontal stream consisting of T-T-T-, etc., as the sequence was repeated. Since the target (Y) was outside of the range occupied by T, the T-T attraction may have been sufficiently strong to overpower any mutual attraction between T and Y, thereby preventing decomposition of the sequence in the T condition. Such competition between alternate organizations was shown to occur by Bregman (1978d). Alternately, no such possibility was present in the Frequency Proximity conditions, since the target (Y) always fell between each pair of captors and at the same frequency range as each captor (see Figure 2). The tendency for captors to group into a separate stream should thus have been eliminated in Frequency Proximity conditions.

Furthermore, it is possible that the accumulation of information which would promote the sequential attraction of captors and targets into a single stream may have been greatest in the Frequency Proximity conditions. In the Frequency Proximity conditions, evidence favouring sequential captor-target grouping could accumulate between cycles, since a single sequential stream composed of captor-Y-captor-Y, etc., was potentiated. On the other hand, in the Trajectory condition, the evidence favouring a trajectory grouping may have accumulated only within cycles, since the trajectory started afresh on each cycle. This presents a possible bias, since it has been shown that sequential streaming effects are a function of the amount of information favouring sequential grouping which is accumulated during any auditory sequence (Bregman, 1978c).

Experiment III attempted to replicate the finding that frequency proximity information is superior to trajectory information at inducing sequential captor-target streaming effects. Safeguards were taken to eliminate the potential biases against trajectory effects present in the current experiment.

EXPERIMENT II

Introduction

In Experiment I, the results on the FPS and FPD conditions suggested that the correspondence in glide orientation between captor and target glides might affect sequential streaming of the glides. It is, therefore, possible that auditory grouping mechanisms compare the direction of frequency change in successive glides to make decisions concerning the sequential grouping of the glides. However, there was an alternate explanation for the observed effect. As discussed earlier, Brady et al. (1961) have shown that the perceived pitch of a short duration glide is associated with the glide's terminal frequency. Since the FPS captor and target glides had the same terminal frequencies, they may also have been perceived as having identical pitches. Captors and targets in the FPD condition, which had different terminal frequencies, may have been perceived as having different pitches. If sequential grouping mechanisms "prefer" to group more similarly pitched events into sequential streams, then it may explain why sequential grouping was stronger in the FPS than in the FPD condition. While the Brady et al. (1961) finding provides a direct reason to suspect that a disparity between the terminal frequencies of successive glides alone might reduce the sequential grouping of the glides, there is no reason to assume that a disparity between the initial frequencies of successive glides might not also influence their sequential grouping tendencies. Therefore, the stronger sequential streaming of FPS captors and targets over that of FPD captors and targets might be attributed solely to the fact that the former had the same initial frequencies, while the latter had different initial frequencies.

The second experiment attempted to separate possible effects of the correspondence in orientation of successive glides on their tendency to form sequential streams from possible influences due to the correspondence between the terminal or initial frequencies of such glides on sequential stream formation. A similar paradigm to that used in Experiment I was employed.

However, instead of examining only two levels of correspondence between the orientations of the captor and target tones (as in the FPS and FPD conditions of the previous experiment), five levels of correspondence were used. Figure 6 graphically depicts how the five different levels were produced by varying the orientation of the captor in each condition. The figure does not show the third tone which was present in each pattern, a glide parallel to, and simultaneous with the target tone. Since the tones in different conditions had different slopes, when plotted on a logarithmic scale of frequency over time, the different conditions will be referred to as "slope conditions". It was predicted that a monotonic decrement in "decomposition" would occur as the slopes of the captors and targets were made less correspondent. However, the captors and targets in different slope conditions also differed in terms of the degree of correspondence between their initial and terminal frequencies. Figure 6 depicts the fact that as the slopes of the captor and target tones became more dissimilar so did the initial and terminal frequencies of the tones. In order to isolate possible effects of slope from any effects resulting solely from the frequency separation between captors and targets (measured at their initial or terminal frequencies), the current experiment also employed steady state (SS) captor tones. On different conditions, an SS captor was placed at either the initial or terminal frequency of the target. By using SS captors it was possible to evaluate three conflicting hypotheses, as follows:

(1) The "terminal frequency" hypothesis

One might hypothesize that only the terminal frequency of a short glide is used during perceptual grouping operations.

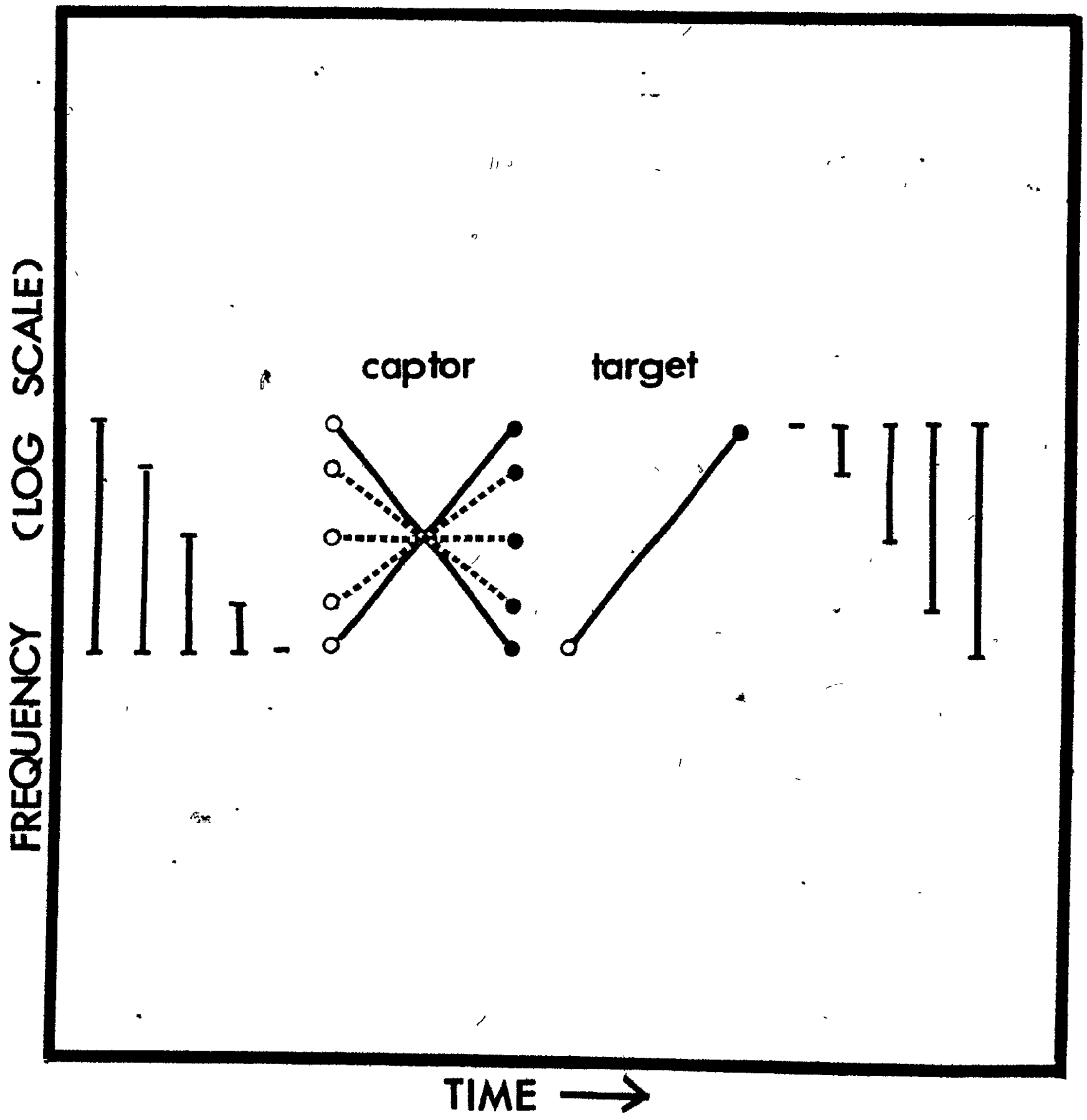


Figure 6. A representation of five captors with slopes progressively more unlike that of a target tone. The vertical lines at the right demonstrate how the difference in frequency between the terminal frequencies of the captor and target tones increases as the slopes of the tones become less correspondent. The vertical lines at the left show corresponding increases in the difference between initial frequencies of the captor and target tones across the "slope conditions".

If so, a monotonic decrement across the slope conditions might be attributed to the increasing separation between the terminal frequencies of the captor and target glides across the conditions. However, if perceptual grouping operations consider only the terminal frequencies of glided tones to arrive at sequential grouping decisions, then the frequency relations between a target glide and a glided captor ending at the same frequency would appear to be the same (to the perceptual grouping mechanism) as the relations between the same target and an SS captor situated at its terminal frequency. The strength of the sequential attraction between both pairs of events should thus be identical. By the same reasoning, an SS captor at the terminal frequency of the target glide should attract the target more strongly than a steady state captor at its initial frequency, since the perceptual grouping mechanism would judge the first captor as being closer in frequency to the target than the latter one.

(2) <u>The "initial frequency" hypothesis</u>

If only the initial frequencies of successive glides are considered during sequential grouping operations, then the predicted monotonic decrement across the slope conditions might be attributed to the fact that the differences between the initial frequencies of the captor and target tones also increased across the conditions. However, to be consistent with such an argument, the difference between the frequency of a steady state tone and the initial frequency of a target glide should have much the same effect on sequential grouping of the tones as the difference between the initial frequencies of the target glide and a glided captor. Thus, an SS captor situated at the initial frequency of a target glide should attract the target at least as strongly as would a glided captor with the same initial frequency as the target. In addition, an SS captor at the initial frequency of the target glide should group better with the target than an SS captor at its terminal frequency.

(3) The "slope" hypothesis

If the predicted monotonic decrement across the various slope conditions is attributable to an additional influence of the correspondence in slope between captors and targets, then SS captors at either the initial or terminal frequency of a target glide should not group as strongly with the target, as a glided captor which had the same slope as the target. The additional slope information provided by the glided captor should increase the sequential grouping tendency on that condition.

Whether the results of the experiment were consistent with one or the other of the above hypotheses would therefore provide a means of isolating whether the correspondence in slope or the correspondence between initial or terminal frequencies of successive glides was more relevant to decisions concerning the sequential grouping of the glides. Since the effect of the two alternation rates used in the previous experiment did not interact with the "captor type" variable, only one alternation rate was employed in the current experiment. It was also arbitrarily decided to attempt to capture only the upper component of the two-tone glided complexes.

Method

Subjects. Twenty-four volunteers, ranging in age from 20 to 37 were tested. All subjects reported having normal hearing, and were drawn from the same population as were subjects in Experiment I.

Apparatus. All equipment used during synthesis and presentation of stimuli was identical to that employed in Experiment I.

Stimuli. All tones used in the stimuli for this experiment were sinusoidal, and frequency glides were exponential. The tones were 230 msecs long, including roughly exponential attacks and decays of 10 msecs (which rose and fell as the tones glided in frequency). Successive tones were separated by 20 msec silences. To generate two-component glide complexes which modulated upward or downward in frequency, the frequency of two sinusoidal tones was modulated up or down in parallel (on an exponential scale) for half an octave. To produce "ascending" glide complexes, the upper tone began at 724 Hz, and the lower tone began an octave below, at 362 Hz. The two tones glided upward to respective terminal frequencies of 1024 Hz and 512 Hz. To produce "descending" glide complexes, a pair of sine tones were modulated downward in frequency; the upper tone gliding from 1024 Hz to 724 Hz, and the lower tone gliding from 512 Hz to 362 Hz. The pair of upper and pair of lower tones in both ascending and descending glide complexes traversed identical frequency ranges. Figure 7 shows the pairs of synchronous tones. The target in each pair is labelled Y, and the other tone Z.

Five captor tones with different slopes were generated. So that all five tones would fall within the same frequency range as the target tone (i.e., the upper component of each complex) and have the same centre frequency (on a logarithmic scale), the

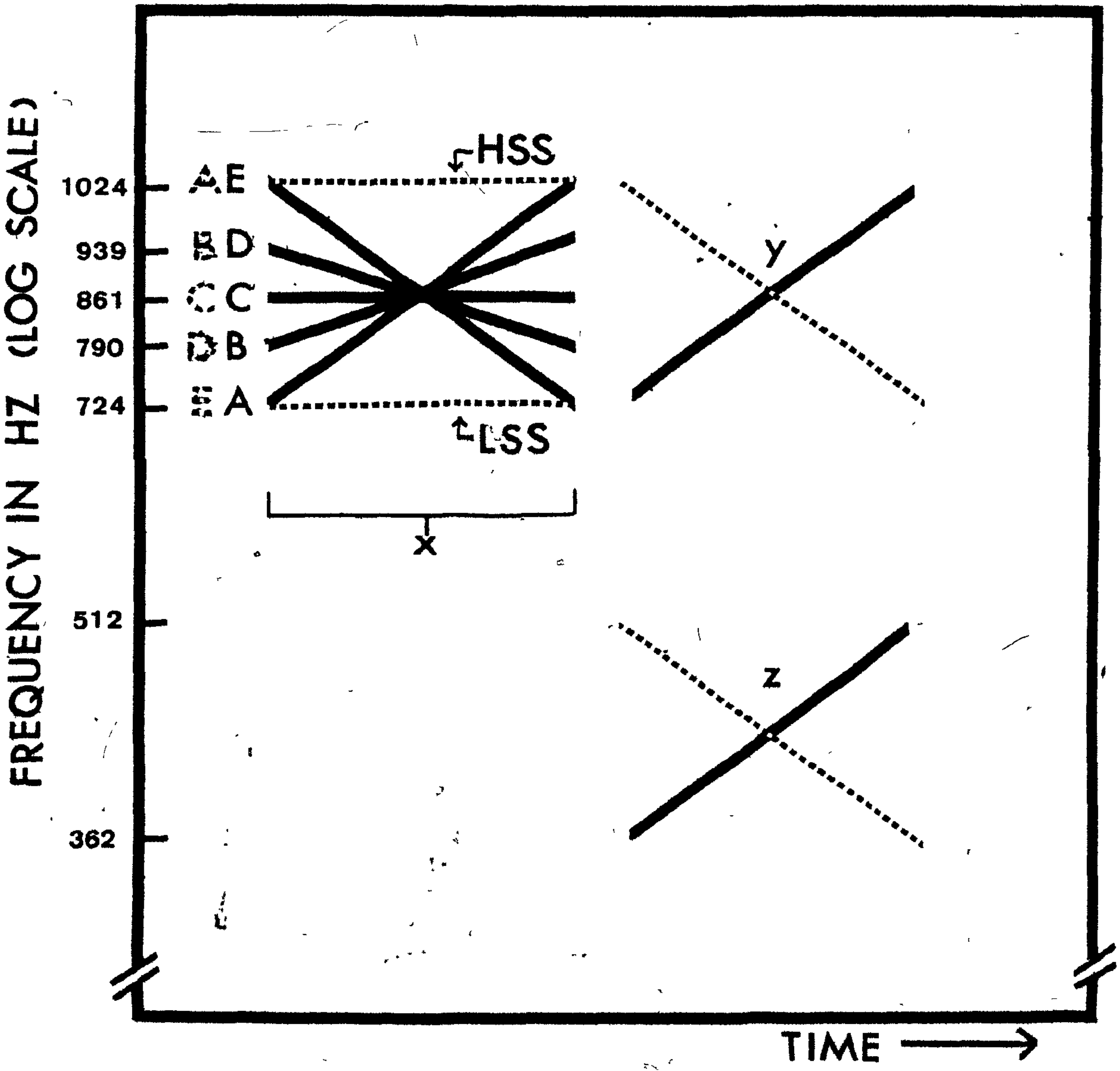


Figure 7. A graphic representation of stimulus patterns used in Experiment 2. The solid lines Y and Z depict the ascending complex, and dotted lines Y and Z depict the descending complex. For the former, the appropriate captors are labelled in solid lettering, and for the latter, captors are labelled in "hatched" lettering. The figure shows one cycle of a continuously repeating pattern.

five frequencies which divided the interval covered by the target tone (i.e., 724 to 1024 Hz) into four equal intervals (on a logarithmic scale) were calculated. In this manner, five frequencies (724, 790, 861, 939, and 1024 Hz) which would serve as the end points for captor tones were generated. These frequencies were used to generate a set of captors representing a continuous range of correspondence in slope to the target tones in both ascending and descending complexes (see Figure 7). The captors are labelled X in the figure. In both ascending and descending complex conditions, the captor labelled A had an identical slope to that of the target (Y). The remaining captors (i.e., B, C, D, and E) had slopes progressively more disparate to that of Y, with E having a completely inverse slope to that of Y. The middle condition C had zero slope. It, therefore, consisted of a steady state captor at the mid frequency (on a logarithmic scale) of Y. Since the different slopes were produced by "rotating" the captor glides in equal increments around a central frequency (on a logarithmic scale), any differences between conditions could not be accounted for by differences in the centre frequency of the various events.

To permit a test of the "initial" and "terminal frequency" hypotheses discussed above, two steady state captors, placed at the initial and terminal frequencies of both ascending and descending targets, were also used. These captors will be called High Range Steady State (HSS) and Low Range Steady State (LSS), respectively (see Figure 7). An interaction of the HSS/LSS factor with the "orientation of glide complex" factor would indicate that the target glides were more strongly attracted to a captor matching either their terminal or initial frequencies.

To generate practice trials, four stimuli analogous to the ascending complex A and E conditions and to the descending complex C and B conditions were generated (see Figure 2). To reduce specific carry-over effects, all frequencies were altered so that stimuli used on practice trials were different from experimental stimuli.

Four different randomizations, each containing 42 experimental trials, were recorded on audio tape. The 42 trials consisted of three replications of each experimental condition (i.e., 7 levels of "captor type" at each of the 2 levels of "orientation of glide complex"). Each trial was preceded by a warning beep, and consisted of 30 repetitions of one of the pure-tone/complex-tone cycles. The trials were separated by six seconds of silence. The four practice trials were recorded at the beginning of each randomization of trial orders. Subjects were randomly assigned one of the four randomized orders, with the restriction that six subjects were run on each randomization.

To ensure that all tones in the stimuli were equal in subjective loudness across their entire range, the loudness of a sine tone (which was ramp modulated at 1 Hz across the entire frequency range employed in the stimuli described above) was measured at the headphones. The output was found to be flat at 80 dB SPL. Human auditory frequency response is also known to be flat at 80 dB SPL across the same frequency range (Fletcher & Munson, 1933). A presentation of experimental stimuli at 80 dB SPL would therefore ensure that all tones in the stimuli were about equal in subjective loudness (at least on average for a group of normal subjects). No individual audiograms were taken.

Procedure. Instructions and treatment of subjects were almost identical to those in the previous experiment. The task remained similar, and an identical seven-point rating scale, labelled FUSED at the lower end and DECOMPOSED at the upper end, was used. Of course, instructions pertaining to the differences in alternation rate of stimuli used in the previous experiment were dropped. Subjects were also informed that the more rapidly alternating pair of pure tones which would be heard in "decomposed" stimuli would always be heard in the higher frequency range, with an isolated tone heard half as often in the lower range. Again, fused stimuli were described as sounding like an even alternation of pure and rich-sounding events. Practice trials were presented in the same manner as in Experiment I. All 42 trials were presented in one session.

The data measure and analysis

The subjects' ratings were sorted into experimental conditions, and the totals of ratings on the three replications in each condition were analyzed in a two-way analysis of variance with repeated measures on both factors. The factors tested were two levels of "orientation of the glide complex" (i.e., ascending or descending) and seven levels of "captor type", five of which represented the correspondence in slope between captors and targets (i.e., A, B, C, D, and E) and two of which represented the SS captor conditions (i.e., HSS and LSS). Subjects occasionally placed their rating marks for two trials on the same copy of the rating scale on their response sheet, and as a result both responses had to be dropped. The missing data points (8 in all) were replaced with cell means. To compensate, 8 degrees of freedom were subtracted from each error term (see Table 4).

Results

The overall mean ratings on captor conditions A through E are shown in Figure 8. The data suggest that as the correspondence between the slopes of captor and target tones was reduced (across conditions A through E), so ratings of perceived decomposition diminished. The main effect of the "captor type" factor was highly significant, $F(6, 130) = 6.35$, $p < .0001$. A test for monotonicity performed on the means for the five "slope" conditions showed the apparent trend described above to be very reliable, $F(1, 130) = 11.43$, $p < .001$. In addition, the interaction of the "slope factor" (when weighted for monotonic trend across conditions A through E), and the "orientation" factor was tested, and found not to approach statistical significance, $F(1, 130) = 0.49$, n. s. In this manner, the consistency of the monotonic trend across conditions A through E in both ascending and descending complex conditions was confirmed. The F-values derived during tests of deviations from monotonicity (across conditions A through E) are shown in Table 4. No significant deviations from monotonicity were observed. Despite the absence of significant deviations from monotonicity, a more liberal test was used to ensure that the apparent upward shift on condition C was not due to more than simple error variance. The observed mean at "C" was tested against that predicted by a straight line fit to the data, using a *t*-test for the difference between a mean and a constant (i.e., by setting the constant to the value predicted at

Table 4: Summary table of the 2-way analysis of variance (orientation of glide complex x captor type) performed on decomposition ratings

Source	SS	df	MS	F	p
Orientation of Glide Complex (0)	97.04	1	97.04	3.36(2.19)$^+$	0.0799(n.s.)$^+$
Error	664.84	23(15)$^+$	28.91(44.32)$^+$		
Slope (S)	631.95	6	105.32	6.75(6.35)	0.0000(<.001)
Error	2154.61	138(130)	15.61(16.57)		
Linear	189.51	1	189.51	12.14(11.43)	<.001(<.001)
Error	2154.61	138(130)	15.61(16.57)		
Quadratic	12.20	1	12.20	0.78(0.74)	n.s.(n.s.)
Error	2154.61	138(130)	15.61(16.57)		
Cubic	8.15	1	8.15	0.52(0.49)	n.s.(n.s.)
Error	2154.61	138(130)	15.61(16.57)		
Quartic	23.79	1	23.79	1.52(1.54)	n.s.(n.s.)
Error	2154.61	138(130)	15.61(16.57)		
0 x S	148.08	6	24.68	2.12(2.00)	0.0550(n.s.)
Error	1607.86	138(130)	11.65(12.37)		
HSS/LSS x 0	115.87	1	115.87	9.95(9.37)	<.01(<.01)
Error	1607.86	138(130)	11.65(12.37)		
Slope (monotonic) x 0	6.03	1	6.03	0.52(0.49)	n.s.(n.s.)
Error	1607.86	138(130)	11.65(12.37)		
HSS ascending versus A descending	64.84	1	64.84	5.57(5.24)	<.05(<.05)
Error	1607.86	138(130)	11.65(12.37)		
LSS descending versus A descending	340.00	1	340.00	29.18(27.49)	<.001(<.001)
Error	1607.86	138(130)	11.65(12.37)		

Table 4--Continued

Source	SS	df	MS	F	p
HSS descending versus					
A descending	1.79	1	1.79	0.15(0.15)[+]	n.s.(n.s.)[+]
Error	1607.86	138(130)[+]	11.65(12.37)[+]		
LSS ascending versus					
A ascending	98.58	1	98.58	8.46(7.97)	<.01(<.01)
Error	1607.86	138(130)	11.65(12.37)		

[+] Adjusted for missing data points

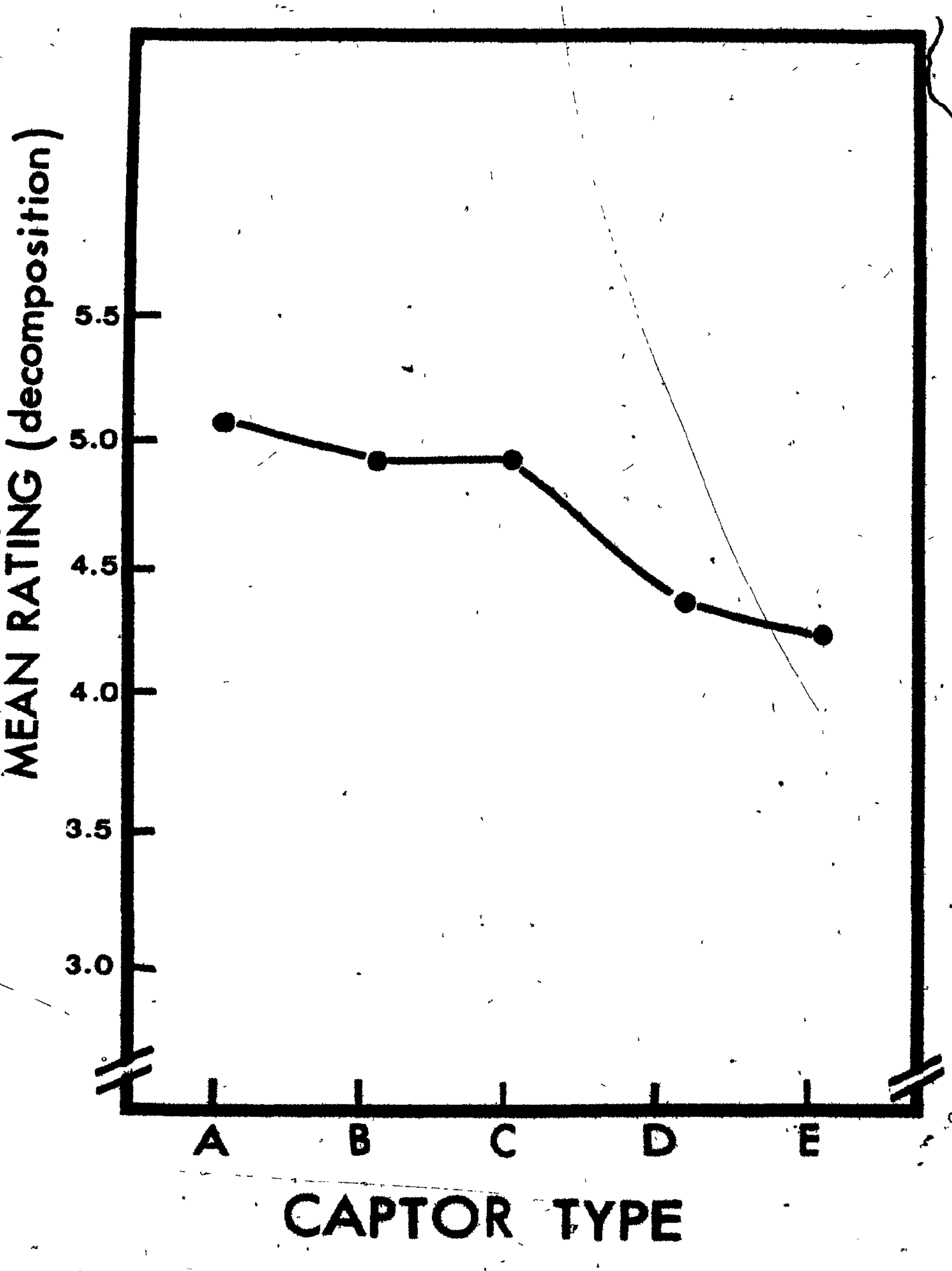


Figure 8. Mean ratings (decomposition) in conditions A through E.

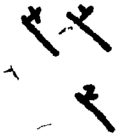
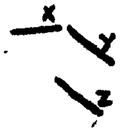
"C" by a linear fit to the end points of the curve). The difference was found to be nonsignificant, $t(46) = 1.82$, n. s. Therefore, the apparent upward shift in condition C was probably not due to any systematic influences on the mean.

Table 5 shows the means on HSS, LSS and A captor conditions, for both ascending and descending targets. The stimulus pattern for the condition associated with each mean is also shown. From Table 5, it appears that each captor produced greater sequential grouping with its target when at the initial rather than terminal frequency of the target. The HSS captor yielded a higher rating on the descending target condition, while the LSS captor yielded a higher rating on the ascending target condition. To determine the reliability of the apparent effect, the interaction of the HSS versus LSS factor with the ascending versus descending target factor was tested on a planned contrast, and found to be significant, $F(1, 130) = 9.37$, $p < .01$. Despite the interaction, LSS means were never higher than HSS means regardless of whether the target tone was ascending or descending. Therefore, it appears that it was not solely the presence of a captor tone at either the initial or terminal frequency of each target which accounts for the results. Instead, the HSS captor appears to have been generally superior at inducing the isolation of both ascending and descending targets from the two-component complex, despite the fact that it was situated at the terminal frequency of the former and at the initial frequency of the latter.

It was also desired to determine whether all the SS captors tended to produce smaller decomposition effects than the condition A captors. Such a difference would suggest that the cor-

Table 5: Mean ratings on HSS, LSS and A captors when alternated with either the ascending or descending glide complexes. (The pattern associated with each condition is shown above each mean. Whether the SS captor was at the terminal or initial frequency of the target glide is also indicated).

<u>CAPTOR</u>

	A	HSS	LSS
Ascending Complex	 4.79	 4.01 terminal	 3.83 initial
Descending Complex	 5.28	 5.15 initial	 3.51 terminal

respondence in slope between the A captors and targets itself provided additional information, over the correspondence in their initial or terminal frequencies, to favour sequential grouping of captor and target tones in that condition. The ratings on the SS captor conditions, when the captors were at their targets' initial frequencies (i.e., HSS in the descending complex condition, and LSS in the ascending complex condition) do not both appear to be as high as the A condition means for the corresponding targets (see Table 5). Planned comparisons revealed that while the "HSS-descending" condition mean was not significantly different from that on the "A-descending" condition, F(1, 130) = 0.15, n. s., the "LSS-ascending" condition mean did differ reliably from the "A-ascending" mean, $\underline{F}(1, 130 = 7.97$, $\underline{p} < .01$. The results suggest, therefore, that the SS captors at the initial frequencies of targets were not uniformly as good at attracting ascending and descending targets as were the corresponding A captors.

The results shown in Table 5 also suggest that both "terminal frequency" captors yielded lower scores than the A condition captors. To determine whether or not they did so reliably, two additional planned comparisons were performed. The difference between the mean on the condition in which the HSS captor was at the terminal frequency of Y (i.e., the HSS-ascending complex condition) was tested against the mean on the A condition for the ascending complex. The HSS mean was significantly lower than the A mean, $\underline{F}(1, 130) = 5.24$, $\underline{p} < .05$. As well, the mean on the condition in which the LSS captor was at the terminal frequency of Y (i.e., the LSS-descending complex condition) was compared to the A-descending complex condition, and found to be reliably lower, $F(1, 130 = 27.49$, $p < .001$.

Since the number of comparisons performed on the data in the current experiment is quite large, the probability that a significant effect might have resulted simply by chance is also quite high. However, the significant F-values on the comparisons fell well above the critical value required to declare each effect significant. Therefore, the effects can be considered to be fairly reliable.

Discussion

The data are consistent with the conception that the strength of sequential grouping between successive pure tone glides is a direct function of the correspondence between the slopes of the glides. As the slopes of successive glides were made less equivalent, perceived decomposition (i.e., sequential X-Y grouping) diminished linearly (see Figure 8). However, to accept the notion that the relative slopes of successive glides affected their sequential grouping tendencies, it must be shown that the correspondence between the terminal or initial frequencies of the glides alone in different conditions was insufficient to have produced the observed effects.

The results on the HSS and LSS captor conditions at each level of the "orientation of glide complex" factor (see Table 5) can be used to infer that the terminal frequencies of the target glides alone were not important in determining their attraction to other glides or steady state events. Each of the SS captors yielded lower decomposition ratings when at the terminal frequency of the target than when at the initial frequency. In addition, both "terminal frequency" captors yielded lower ratings than the corresponding A condition captors. Therefore, the two criteria established to accept the "terminal frequency" hypothesis discussed in the Introduction were not met.

The results shown in Table 5 do, however, suggest that the initial frequency of a glide may play a greater role than its terminal frequency in causing the glide to be attracted to other auditory events. However, if the perceptual grouping of a glide with other events was based solely on the glide's initial frequency, then one would expect that the SS captors would have yielded consistently higher scores when at the initial rather than terminal frequency of their targets. Clearly, this was not so, since the HSS captor at the terminal frequency of the ascending target glide yielded at least as high a score as that of the LSS captor at the initial frequency of the same target (see Table 5). Furthermore, one would predict that if for the purpose of perceptual grouping only the initial frequency of a glide is considered, then each SS condition in which the captor had the same frequency as the initial frequency of the target (i.e., 'HSS-descending complex' and 'LSS-ascending complex' conditions) should have yielded identical results to the A condition with the corresponding target (since the captors and targets had the same initial frequencies in both A conditions). The fact that one of the "initial frequency" captors (i.e., the LSS captor in the ascending complex condition) yielded a significantly lower score than that on the corresponding A condition suggests that the correspondence between the initial frequencies of the captors and targets in the A condition alone cannot account for the higher decomposition rating (or higher sequential grouping) on that condition. There is therefore reason to reject the "initial frequency" hypothesis.

The apparent superiority of each SS captor in attracting Y, when at Y's initial rather than terminal frequency (see Table 5) may be attributable not to the initial frequency factor per se, but rather to the fact that each "initial frequency" captor nearly "connected up" in frequency with Y (see Table 5). It is reasonable to assume that a captor which continuously excited the frequency region corresponding to the region in which Y began, right up until just prior to the onset of Y, would have provided a strong cue to enable the auditory system to isolate Y from the complex. The "terminal frequency" SS captors, which instead "connected up" at the end of Y, may not have provided such a cue. In the "initial frequency" SS captor condition the auditory system would have been primed for the occurrence of Y, while in the "terminal frequency" condition, the cue would have arrived only after the occurrence of Y. If the cue is not as effective retroactively, then this might account for the fact that each captor when in the initial frequency position appeared to attract Y better than when in the terminal frequency position. To account for the overall superiority of the HSS captors over the LSS captors in attracting Y, one might suggest that each captor's attraction to Y may have competed with its attraction to Z. Therefore, the HSS captors, which were always further from Z, may have been less affected by such a competitive attraction than the LSS captors. The HSS captors would thus be somewhat advantaged in attracting Y. The combined effects (i.e., the captors' 'connection' to Y and the competition from Z) may account for the fact that the SS captor that "connected up" with Y and was far from Z (i.e., the 'HSS-descending complex' captor) yielded the highest decomposition rating (see Table 5). However, since the experimental design was not symmetrical with respect to the distance in frequency between HSS and LSS captors and Z, this speculation cannot be verified.

Among captor conditions A through E, the captors that yielded the highest decomposition ratings (e.g., condition A) did not "connect up" with Y in the same manner as the "initial frequency" SS captors. It is, therefore, likely that a different factor from that active on the SS conditions was largely responsible for the captor-target grouping in the other conditions. Possibly the correspondence in slope between the captor and target tones in each condition was such a factor. Perhaps, when the slopes of successive glides corresponded, a bias in favour of sequential grouping occurred and when the slopes did not correspond a bias against such grouping developed.

The main piece of evidence which was used to reject the "initial frequency" hypothesis was the relatively low score of the "initial frequency" LSS condition. Since a competitive attraction from Z may in part account for the low score observed on that condition, we must be cautious in completely discounting the hypothesis. However, Brady, et al.'s (1961) finding, which suggests a dominant role of the terminal frequency of a glide in perception, questions the plausibility of the notion that the initial frequency of a glide dominates during perceptual grouping. Furthermore, the idea that Z may have competed with Y for the attraction of the LSS captor is highly speculative. Even if true, comparable attractions from Z on the captors in conditions A through E might have caused any effects of Z to cancel out across the conditions. Therefore, while further research is required before the "initial frequency" hypothesis can be unequivocally rejected, the notion that the relative slopes of the tones was involved in producing the effects observed in this experiment is preferred at this time.

EXPERIMENT III

Introduction

In Experiment I, the results suggested that frequency proximity information dominates over any effects of trajectory information in determining which successive events will group into a single sequence. However, several possible accounts for the observed inferiority of trajectory effects over frequency proximity effects at inducing sequential grouping of glided tones were raised earlier. To recapitulate, it is possible that the hypothetical trajectory effect is simply weaker than the frequency proximity effect. Therefore, in an experiment in which strong frequency proximity effects were in competition with weaker trajectory effects, the weaker effects would have rated at the low end of the rating scale, even if they were perceptible to subjects, since subjects were probably just comparing one condition with another. The stronger effects would simply have overshadowed the weaker ones. Second, as was pointed out in the Discussion of Experiment I, the possible tendency for captors in the "Trajectory" condition to group with each other sequentially may have competed with any captor-target grouping tendency. Since the "Frequency Proximity" stimuli used in Experiment I would have provided far less opportunity for such captor-captor grouping, they may have been slightly more conducive to sequential captor-target grouping. Finally, it was pointed out earlier that cumulative perceptual evidence supporting sequential grouping may have been greater in the "Frequency Proximity" conditions than in the "Trajectory" conditions employed in Experiment I.

The following experiment attempted to eliminate these potential confounds, while still examining the relative contributions of frequency proximity and trajectory information to the sequential grouping of successive frequency glides. Again, the stimuli consisted of a single glide, followed by a glided complex containing a pair of synchronous glides. To maintain equal grouping tendencies on the basis of slope, the slopes of captor tones and of tones in the complex were always made identical. Since no effects directly due to the direction of the complex glide had been found in previous experiments, all glides were arbitrarily oriented upward. As a measure of sequential grouping of successive tones, it was arbitrarily decided to try to capture the lower tone (or fundamental) of the glide complex, and thereby to try to decompose the complex tone. Instead of presenting continuous cycles of the pattern, repetitions of the pure tone/complex tone sequence were separated by four seconds of silence. Bregman (1978c) has suggested that the streaming mechanism accumulates evidence over successive events, and uses the cumulative strength of this evidence to generate a probability statement concerning how likely it is that a series of events belong in a single stream. The introduction of a four second silent interval was found to clear most bias favouring a particular stream formation, which had developed due to previously accumulated information. Therefore, by imbedding a four second silence between each repetition of the captor/target pair in the current experiment, it was presumed that any tendency for the sequential grouping of tones in successive cycles to occur, would be eliminated. The only remaining tendencies would thus be those for "within cycle" groupings. The introduction of silent pauses between repetitions of the patterns should thus serve two ends: (1) It would eliminate the possibility of captor to captor groupings between cycles of the stimulus pattern, and (2) the accumulation of evidence favouring either trajectory or frequency proximity groupings would be equalized across conditions. Presumably, any mechanism accumulating evidence for either potential grouping would be reset following each four-second silence. Hence repetitions of the pattern were not intended to promote streaming effects, but were simply employed to enable subjects to hear each pure tone/compex tone sequence often enough so that they could make a reliable judgment of whether each presentation of the sequence was "fused" or "decomposed".

To reduce the tendency for stronger effects to "wash out" the hypothetically weaker trajectory effect, captor tones were positioned in seven different frequency regions, ranging from a region equal to that of the target tone (i.e., the fundamental of the glide complex) to one far below it (see Figure 9). The captor at the fourth of these seven positions was situated so that it and the target tone aligned themselves on the same trajectory. It was presumed that if it was only the frequency proximity between the captor and target (assessed by judging the proximity between their average frequencies) that influenced their mutual sequential attraction, then the strength of sequential grouping would diminish as the distance between the captor and target was increased. Therefore, one would predict a decreasing amount of capturing with increasing distance between the captor's and the target's frequency regions. Alternately, if trajectory information also played a role in attracting the target to the captor, then the single "Trajectory" condition should deviate from the smoothly decreasing trend, such that the mean on the "Trajectory" condition would be higher than that predicted by a line fit to the data. Such a result would reflect the fact that the influence of a "trajectory effect" had become added to the effect of frequency proximity on that condition, causing the mean for the "Trajectory" condition to be inflated. Since the two effects should be additive in such a situation, any "trajectory effect" should become visible "on top of" the frequency proximity effect.

Due to the fact that stimuli were not presented in continuous cycles, it became very difficult to achieve clear decomposition effects. Presumably, such effects require the accumulation of evidence over repeated cycles, as suggested by Bregman

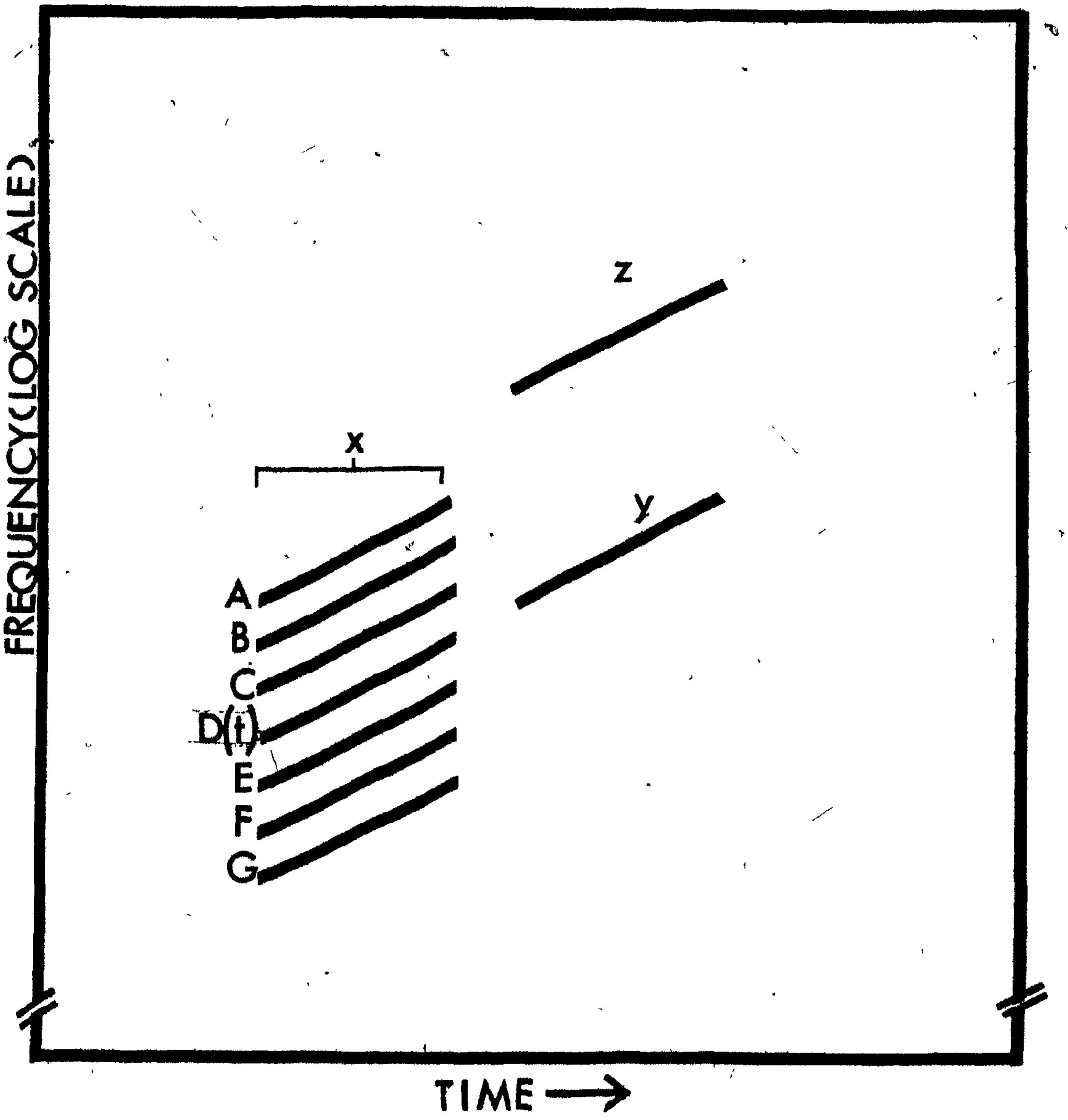


Figure 9. A graphic representation of stimulus patterns in Experiment 3. Captors are labelled X, the target Y, and the third tone (designed to fuse with Y) is labelled Z. The figure shows one cycle of a continuously repeating pattern.

(1978c). However, during pretesting, it was found that decomposition of the mixture was perceptible when a single alternation of captor and complex glides was presented at a 130 msec event rate (with 20 msecs of silence between the tones). Since slower rates would not yield clear decomposition on a single presentation of the pure tone/complex tone pair, only the fast rate was employed in this experiment. At the 130 msec rate, the tones were still clearly audible as frequency glides, and hence it was felt that the hypothetical trajectory effect should still be operative.

Method

Subjects. Twenty volunteers, ranging in age from 19 to 29 were tested. All subjects reported having normal hearing, and were drawn from the same population as those employed in previous experiments.

Apparatus. All equipment used during the synthesis and presentation of stimuli was identical to that employed in the previous experiments.

Stimuli. All tones were sinusoidal, and frequency glides were exponential. Each tone had a 130 msec duration including a 10 msec exponential attack and decay of amplitude (rising and falling as the tone was glided). Each stimulus pattern consisted of a tone followed 20 msecs later by a pair of simultaneous tones.

To generate a rich sounding two-tone complex, the frequency of each of a pair of synchronous sine tones was modulated upward in frequency, in parallel, for just under one half an octave.

In this experiment, the upper tone glided from 1448 to 1919 Hz, and the lower tone glided one octave lower, from 724 to 960 Hz. During pretesting, the complex was judged to sound like a single rich-sounding glide, when heard in isolation from other tones.

To produce captor tones which would precede the complex, and potentiate a sequential grouping between themselves and the lower component of the two-tone mixture, seven different glided sine tones, all with identical durations and slopes to the tones in the complex, were used. To produce strong frequency proximity grouping, a glide identical to the lower component in the complex tone, (i.e., ranging from 724 to 960 Hz), was used. This condition will be referred to as condition A (see Figure 9). Note that in this pattern a sequential grouping of the captor (X) and target (Y) should cause tone Z to be heard as an isolated high pitched event. To create a stimulus pattern which would favour a grouping on the basis of the trajectory between the captor and target tones, a captor tone (X) was glided from 512 to 679 Hz. As a result, the captor (X) and target (Y) were aligned on precisely the same trajectory on a logarithmic scale of frequency. This condition is labelled condition D(t) in Figure 9. The latter condition was produced essentially by gating on and off a single long frequency glide, gliding from the captor onset frequency (512 Hz) to the target offset frequency (960 Hz). Thus it was assured that the captor and target were aligned on precisely the same trajectory.

Captors in condition B and C consisted of tones glided from 645 to 855 Hz and 575 to 762 Hz, respectively. Similarly, three captor conditions were generated in which the captors fell below the D(t) captor in frequency. Condition E, F and G captors glided from 456 to 604, 406 to 539, and 362 to 480 Hz, respectively. In this manner, all glides employed had equal slopes, and traversed the same distance on a logarithmic scale of frequency. Captors across conditions A through G, became progressively further in frequency from the target tone in equal increments (on a logarithmic scale). The glide complex remained the same in all conditions.

Four practice trials, analogous to the patterns shown as conditions A, B, D(t) and G in Figure 8 were generated. However all frequency ranges of tones were altered, so that the practice trials were similar but not identical to experimental trials. Four randomizations of trial orders were recorded on audio tape, with each randomization consisting of 21 trials (i.e., 3 replications of each of the 7 captor tone conditions). The four practice trials were recorded at the beginning of each randomization in the order A, G, B and D(t). Each trial consited of two warning beeps, followed by two seconds of silence. Subsequently, eight repetitions of one stimulus pattern were presented, with four seconds of silence between the repetitions. A single beep was used to signal the beginning of a six second response period between successive trials.

The upper tone in each two-tone complex, which ranged from 1448 to 1919 Hz, fell in a region of heightened frequency response at roughly 80 dB SPL, according to values plotted on the Fletcher-Munson curve (Fletcher & Munson, 1933). Thus the upper tone was probably heard as being louder than any of the other tones. However, since this loudness difference was constant across all conditions, and no attempt to capture the upper tone was made in any experimental condition, it was not necessary to adjust the loudness of the tone. All captor and target tones fell within the frequency region between 362 and 960 Hz. The auditory system shows a flat frequency response at 80 dB SPL within this range. The response of playback equipment was measured at the headphones, and was also found to be flat for an 80 dB sine tone glided up and down across the entire 362 to 960 Hz range. Thus, all captor and target tones had equal subjective loudness (on average for a group of normal subjects).

Procedure. In this experiment, subjects were again asked to judge whether stimuli containing a pure tone glide followed by a pair of simultaneous glides were heard as being "fused" or "decomposed". The subjects were told that "decomposed" stimuli would sound like a pattern containing a pair of rapid glides in the low frequency range, with a third isolated sound audible in the higher range. "Fused" stimuli were described as sounding like a pair of glides, one sounding pure and one sounding rich. Two diagrams depicting "fused" and "decomposed" patterns accompanied the explanation. Subjects were informed that they would hear eight repetitions of each pattern, and that each presentation of the pattern would be followed by a four second silence.

The subjects required considerably more exposure to practice trials than they had required in previous experiments before they could recognize any decomposition effects at all. The procedure employed to train subjects on the task was as follows:

The first practice trial (analogous to condition A in Figure 9) was presented. Although this stimulus represented a "strongly decomposed" pattern, the effect was subtle, and difficult to hear on the first exposure. Only one subject rated the stimulus as being decomposed on the first presentation. For every other subject, the same trial was repeated (approximately 4 times) until the subject reported hearing three clearly audible tones. Then the remaining three practice trials were presented. If the subject felt uncertain about qualitative differences between the stimuli, all four practice trials were presented once more. By this point, all subjects reported that they could hear three tones in some sequences and only two tones in the others, except for one who was rejected from the experiment.

All experimental conditions (21 trials) were presented in a single session. Subjects were randomly assigned one of the four randomized orders, with the restriction that five subjects were run with each order.

The data measure and analysis

Subjects' ratings were sorted into experimental conditions, and the totals of ratings on the three replications in each condition were analyzed in a one-way ANOVA with repeated measures on the seven levels of "captor position" (see Table 6).

Results

Figure 10 shows the mean rating for each of the seven captor conditions. The data suggest that there was a monotonic decrement in the ratings correlated with increasing separation

Table 6: Summary table of the 1-way analysis of variance performed on decomposition ratings in Experiment III

Source	SS	df	MS	F	p
Captor Type (CT)	991.14	6	165.19	11.61	0.0000
Error	1622.00	114	14.23		
Linear	949.00	1	949.00	66.70	0.0000
Error	1622.00	114	14.23		
Quadratic	0.10	1	0.10	0.01	n.s.
Error	1622.00	114	14.23		
Cubic	0.08	1	0.08	0.01	n.s.
Error	1622.00	114	14.23		
Quartic	27.12	1	27.12	1.91	n.s.
Error	1622.00	114	14.23		

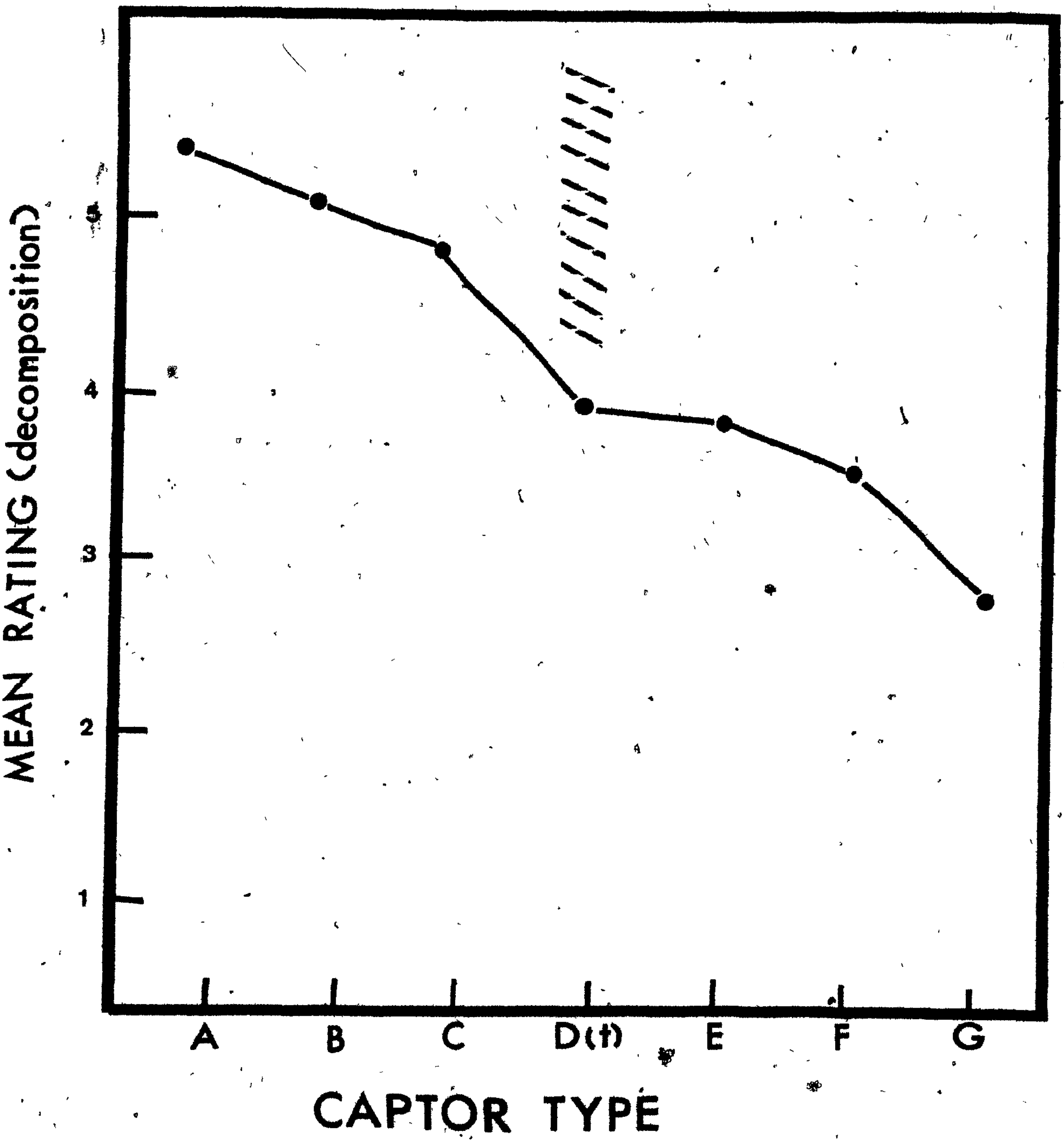


Figure 10. Mean ratings (decomposition) at each level of "captor type". The "hatch marks" indicate the expected region for the mean at D(t), given the existence of a "trajectory effect".

in frequency between captor and target tones (across conditions A through G). There appears to be no upward deviation from the trend in the D(t) condition. This negative result suggests that a "trajectory effect" did not raise the rating on the latter condition, by adding a separate influence favouring decomposition. (The 'hatch marks' in Figure 10 indicate the expected region for the mean at D(t), given the existence of a 'trajectory effect'). The main effect of"captor position" was very reliable, $\underline{F}(6, 114) = 11.61$, $\underline{p} < .0001$. A trend analysis was performed on the data, and resulting F-values are shown in Table 6. Note that while a significant linear trend resulted from the analysis, $\underline{F}(1, 114) = 66.70$, $p < .0001$, all tests for deviations from linearity yielded nonsignificant values.

Discussion

The results of this experiment are very clear cut. The strength of sequential grouping of successive glides appears to have been a direct function of the frequency proximity between the glides. As the captor and target glides were made more distant from each other in mean frequency, the ratings for the patterns became lower (i.e., less 'decomposed'), suggesting that sequential grouping of the captors and targets became weaker across the conditions. This result, therefore, corresponds to the findings of Experiment I, which suggested that glides more proximal to each other in mean frequency are grouped more strongly into sequential streams.

No evidence was obtained to suggest that the auditory system tracks from the trajectory of one gliding sinusoid onto a subsequent sinusoid maintaining the same trajectory of change. It is, therefore, suggested that the auditory system does not use trajectory information during those auditory grouping processes that weigh antagonistic sequential and "vertical" grouping tendencies; at least, such use of trajectory information has not been detectable by the means employed in the current experiment, or in Experiment I.

The finding is somewhat contradictory to a result described by Bregman and Dannenbring (1973). They showed that sequences of alternating high and low tones would remain in a single stream at high presentation rates if short frequency glides at the onsets and offsets of each tone made the tones "point" to each other. High and low tones with no such glides were found to segregate into separate streams when presented in otherwise identical sequences. It was suggested that the result reflected a "trajectory-like" effect, in which the tracking of the rapidly alternating sequence of tones was facilitated by the glides which served to sensitize the auditory system to the frequency region in which each subsequent tone would occur. However, the effect can also be explained in terms of frequency proximity. The short glided segment on each tone, "pointing" to the subsequent tone, also made the tones effectively closer in frequency. Given the result of the current experiment, which shows no evidence for a trajectory effect, and much evidence for frequency proximity effects when stimuli consist of alternating frequency glides, it is probable that frequency proximity effects accounted for Bregman and Dannenbring's finding.

The inability to detect trajectory effects between glided tones in the current study is consistent with a result reported by Halpern (1977). She found that listeners could not fully track either one of a pair of ascending and descending sine tone glides which crossed each other's paths at their mid-points. Listeners tended to track the first part of the ascending glide to the crossing point, and then to follow the second half of the descending glide back down, or to begin by following the downward glide part way and then to follow the second half of the ascending glide back into the higher range. In either case, the trajectory of a single glide appeared not to be extrapolated across the point of crossing. The result, therefore, supports the notion that the role of trajectory information in the perceptual organization of auditory patterns containing glides is minimal in comparison to frequency-based grouping tendencies.

One factor may restrict the generality of results in the current experiment. Since the duration of each glided tone was fairly short, there may have been little time for the auditory system to have accumulated information describing the trajectory of each tone. It is, therefore, possible that a trajectory effect might have been achieved had longer tones been used. However, the durations of tones used in the current study were somewhat longer than the average durations of glided formants which occur in speech. Therefore, if the tracking of trajectory information is to have any practical significance in a normal auditory environment, such information must be processed very rapidly. Thus, even if a trajectory effect would cause a sequential attraction between sinusoidal glides of long durations, it is difficult to imagine what the practical significance of such a phenomenon would be. However, glided transitions in speech are continuous with the speech train, while the glides in the current experiments were discrete. It is possible that the auditory system's sensitivity to continuous trajectory information differs from that for trajectories between discontinuous events. Perhaps the auditory system can use trajectory information, but simply cannot extrapolate from a trajectory across a silence. Therefore, had the captor and target tones in the trajectory conditions been continuous (i.e., one long gliding tone), trajectory information may have succeeded in separating the target from the complex tone. Further research is, therefore, required to determine whether or not the auditory system can utilize trajectory information as a predictive device during the processing of continuous events.

GENERAL DISCUSSION

The current study has shown that a pair of simultaneous gliding sinusoids can be perceived either as a single "rich" event or as separate sinusoidal events. Whether the pair of tones is perceived in one way or the other depends on the tones' relations to other sinusoidal events which precede and follow them. By alternating the simultaneous glides with either a sinusoidal glide, or a steady state sine tone which fell within a similar frequency range as one of the simultaneous components, a bias against fusion of the simultaneous sinusoids and in favour of sequential streaming of successive sinusoids was generally produced. There was also some evidence to suggest that the sequential bonding strength between successive glides, when in the same frequency range, may be modulated by the correspondence in slope of the glides. However, no evidence was developed to suggest that successive glided tones, aligned on a common trajectory tend to attract each other sequentially.

If we assume that when successive tones in a sequence do not group sequentially, it is because the difference between their frequencies falls above some criterion value, then these results, taken together, suggest a specific conclusion. The difference in frequency between successive glided tones appears to be more directly a function of the difference between average measures of each tone's frequency, than the difference in frequency between the end point of one tone and the onset point of the subsequent tone. The explanation would account for the fact that, in the current experiments, when a pair of alternating glides fell in different frequency ranges, they did not group sequentially, even when the frequency change within the second tone was perfectly continuous with that of the first tone. In such a case, the difference in frequency between the tones (based on their average frequencies) would simply have been interpreted as being too large. One can reformulate the speculation above by proposing that "inter-tone" change (in frequency) is determined independently of "intra-tone" change. As a result, mechanisms which assess the extent of frequency change between discrete tones, possibly by comparing an average measure of the frequency of each tone, would be insensitive to a continuity in the pattern of change created by the successive frequency components within each of the tones. Perhaps, the percept of change within a tone (i.e., 'intra-tone' change) is detected at a completely different level, and in a completely different manner from change between tones (i.e., 'inter-tone' change).

Mechanisms sensitive to "intra-tone" frequency change might also be able to detect the slope, or trajectory of change within a gliding tone. The results of Experiment II suggest that the auditory system might possess such sensitivity to differences in the slopes of glided tones. However, Experiments I and III suggest that the system is unable to extrapolate from the trajectory of one glide onto a subsequent glide which maintains the same trajectory. Therefore, trajectories existing between collinear gliding tones appear not to be simple extensions of the trajectory within each of the tones. This finding again suggests that "intra-tone" properties are treated as being distinct from "inter-tone" properties by the auditory system.

To account for the apparent distinction between "intra-" and "inter-tone" properties, we may speculate that auditory mechanisms that describe "inter-tone" relations are autonomous from those that describe "intra-tone" relations. Perhaps at one stage of organization auditory mechanisms only track information about continuous changes (e.g., in frequency) occurring within discrete tones. The discontinuity (in amplitude or in frequency) between successive tones might terminate such tracking processes, and usher in a computation which generates a "summary" of the changes occurring within each tone. Each calculation would therefore result in information pertaining only to the frequency components of a specific tone (e.g., an average frequency), or the components' relations to each other within the tone (e.g., whether they comprise a rising or falling glide), etc. Perhaps, when the auditory system progresses to a description of "inter-tone" relations, the "summaries" of "intra-tone" properties become the raw information for a subsequent level of computations. "Inter-tone" relations might thus be determined by comparing the "summary" for each tone to that for each other tone in a sequence. As a result, relations between discrete components of successive tones, (e.g., whether or not the terminal frequency of one tone was close in frequency to the onset frequency of the subsequent tone) would be "invisible" to the system at this stage, since such relations would have been obscured by previous averaging processes. If auditory grouping mechanisms operate by determining "inter-tone" relations on the basis of such "summary" descriptions, this might explain why the observed "streaming effects" appeared to be based on the "global" similarities in the characteristics of successive tones in a sequence (e.g., their correspondence in frequency range and slope) rather than on continuities between the tones.

A possible physiological substrate for such mechanisms

Hypothetical mechanisms which would underlie such operations can be described, as follows: The locations in which information about each tone in a sequence is accumulated can be conceived to consist of regions in the central nervous system that represent events occurring along the basilar membrane. A frequency analysis performed on any signal would result in the description of frequency components of the signal at each moment in time. Discontinuities in the signal, occurring at the onset or offset of each discrete event would be demarcated as regions in which the energy shifts abruptly to a new level, or is absent for a period of time. If one such event was a glided tone, then it would result in the representation of a continuous spatio-pattern of displacement along a region of the membrane, bounded by an onset and offset discontinuity.

To account for the observations of the current study, it is required that an integration of information over each continuous pattern of displacement (corresponding to a single glide) be performed, but not over discontinuous portions of the pattern (i.e., not across successive tones). Such integration might be performed by "glide detectors" analogous to those described by Evans and Whitfield (1965) and Vartanian (1974). So that the firing of each "detector" would be restricted only to the continuous pattern of displacement produced by a single gliding tone, it is proposed that a discontinuity in amplitude or in frequency at the offset of a glide would cause the "detector" to become temporarily refractory. Therefore, a subsequent glide, even if it followed the same trajectory set by the first glide would not excite the same "detector". Such an organization would account for the apparent absence of frequency tracking between successive collinear glides, and would explain why a pair of collinear glides appear not to be perceived as forming a single continuous trajectory. (If a single glide-sensitive unit could be fired by a pair of collinear gliding tones which fell within the responsive region of the unit, then the common trajectory between the tones would certainly be coded. That such trajectory appeals not to be detected, suggests that some mechanism similar to the one described above prevents successive tones from activating a single glide-sensitive unit). The firing of any single "glide detector" is, therefore, conceived to code only the sensory features of a discrete glided tone. Such features would include its frequency range, and its rate and direction of frequency change. In other words, the "detector" would code the "intra-tone" properties of the glide discussed above.

Auditory grouping mechanisms might utilize information coded in such a manner to determine how a complex signal composed of discrete glides should be organized perceptually. Such processes might simply compare the coded features of successive tones in a sequence and group the events with the most similar "codes" into a stream. Information describing the proximity between the frequency ranges of successive glides and their similarity in terms of slope would be readily available from a comparison of the "codes" embodied in the firing of different "glide detectors". Hence, frequency proximity information and slope information would be expected to play a substantial role in determining the strength of sequential grouping of glided tones. However, a simple comparison of the coded features of successive collinear tones (such as those in the 'T' conditions of the previous experiments) would simply result in the detection of the fact that the tones occupied different frequency ranges. Therefore, the grouping mechanism might organize such tones into a weak sequential stream. Such a stream would not remain integrated when competing with strong "vertical" grouping tendencies.

The functional role of such mechanisms

Through the restricted use of only information concerning the frequency proximity between successive events, and the correspondence between events on dimensions such as slope, the auditory system may meet at least two demands placed on it, when required to separate interleaved components from separate sound sources into perceptual streams. According to Bregman's (1978a) theory, sounds with similar frequencies have a higher probability of resulting from one sound source than do sounds with dissimilar frequencies. By organizing sounds with different frequencies into separate streams, the auditory system is, therefore, likely to produce an appropriate organization of a sequence containing sounds at different frequencies. Another high probability occurrence is that sounds which accidentally fall within the same frequency range, but which arise from separate sound sources have different characteristics on other acoustic dimensions (e.g., their slope of frequency change, loudness, amplitude envelope, etc.). Therefore, by also comparing such features of successive events, qualitatively dissimilar events (which are likely to have arisen from separate sound sources) may be perceptually isolated from each other, even if they share a common frequency range.

However, single sound sources are also capable of generating sequences of sounds which differ in terms of their slopes of frequency change, loudness, etc., (e.g., a human voice). Therefore, the segregation of sounds on the basis of such differences must be performed conservatively. A conservative heuristic for segregating dissimilar sounds which fall in the same frequency range would be to tentatively keep the events in a single stream, unless higher probability groupings with other sounds can be found which would cause the dissimilar events to be isolated from each other in different perceptual streams. For example, in the current experiments, one can conceive that when captor (X) and target (Y) tones fell in the same range, but had very different slopes, there was some negative bias on sequential grouping between the X and Y tones, making the grouping weak and tentative. The presence of an alternate "vertical" grouping between Y and Z would have been taken as the preferred grouping option in such a case, since it separated X and Y. When there was no such bias against sequential X-Y grouping (i.e., when X and Y had the same slopes), there would be less reservation about a sequential X-Y grouping. The X-Y grouping would, therefore, have been given precedence over the Y-Z grouping. Therefore, the auditory system is conceived to consider the relative weighting favouring all available groupings before arriving at final grouping decisions.

There is no compelling reason to assume that the successful organization of components from different sound sources into streams would not also be facilitated by the tracking of trajectories traced by sequential events. Furthermore, the program developed by Parsons (1976) to segregate simultaneous speech signals indicates that trajectory information may be useful in decomposing mixtures composed of such signals. Therefore, the apparent absence of "trajectory effects" in the current study should not be taken to deny the possibility that trajectory information may be used in some contexts by perceptual grouping processes. The current study does, however, suggest that no extrapolation from the trajectory of one sinusoidal glide onto another such glide occurs.

In summary, the current results suggest that there may be independent levels of description in the auditory system. For example, at one level, the relations between the successive frequency components which comprise a glided tone might be described. At such a level neural computations might abstract such properties as the velocity of frequency change within the glide, the frequency range of the glide, etc. Perhaps at a higher level, the information derived from these calculations is used in a new set of computations which describe the relations between successive tones in a sequence. At such a level "inter-tone" relations may be abstracted from the sequence. The results of the current study imply that information extracted at the "intra-tone" level of description may be immiscible with information extracted at the "inter-tone" level. Thus, "intra-tone" trajectories do not appear to extend onto "inter-tone" trajectory, and "intra-tone" frequency relations appear to be irrelevant to the determination of "inter-tone" frequency relations.

This may account for the fact that the current results appear to be somewhat in conflict with predictions based on Jones' (1976) model of perceptual organization. She predicts that a constant velocity of pitch change between successive events should promote the perceptual organization of such events into a single perceptual stream. However, the prediction refers directly to the velocity of change between tones in a sequence (i.e., to 'inter-tone' velocities of pitch change). Since the "constant velocity patterns" in the current study (i.e., in the 'Trajectory' conditions) were created by keeping the velocity of change constant <u>within</u> each of a pair of successive collinear glides (so that the velocity held constant was 'intra-tone' velocity), the velocity of change in the stimulus may have been processed by mechanisms very different from those proposed by Jones to process "inter-tone" velocities. Since there is no reason to assume that patterns in which "intra-tone" velocity is constant should also result in strong perceptual organizations, the absence of a "trajectory effect" in the current study does not necessarily refute Jones' model. It does, at least, however suggest that "intra-tone" properties differ substantially from "inter-tone" properties.

That the auditory system may make such distinctions indicates the need for caution when generating theoretical constructs concerning organizational processes in audition. For clearly, the accuracy of any generalizations about perceptual grouping operations within the auditory system may be limited to a specific level of organization.

References

Barlow, H. B., Hill, R. M., & Levick, W. R. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. Journal of Physiology, 1964, 173, 377-407.

Beck, J. Similarity grouping and peripheral discriminability under uncertainty. American Journal of Psychology, 1972, 85, 1-20.

Békésy, G. von. Experiments in Hearing. New York: McGraw-Hill, 1960.

Brady, P. T., House, A. S., & Stevens, K. N. Perception of sounds characterized by a rapidly changing resonant frequency. Journal of the Acoustical Society of America, 1961, 33, 1357-1362.

Bregman, A. S. Asking the 'what for' question in auditory perception. In M. Kubovy and J. Pomerantz,(Eds.), Perceptual Organization. Hillsdale, N. J.: Laurence Earlbaum Associates, 1978. (a)

Bregman, A. S. The formation of auditory streams, In J. Requin (Ed.), Attention and Performance VII. Hillsdale, N. J.: Laurence Earlbaum Associates, 1978. (b)

Bregman, A. S. Auditory streaming is cumulative. Journal of Experimental Psychology: Human Perception and Performance, 1978, 4, 380-387. (c)

Bregman, A. S. Auditory streaming: Competition among alternative organizations. Perception and Psychophysics, 1978, 23, 391-398. (d)

Bregman, A. S., & Achim, A. Visual stream segregation. Perception and Psychophysics, 1973, 13, 451-454.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. Journal of Experimental Psychology, 1971, 89, 244-249.

Bregman, A. S., & Dannenbring, G. L. The effect of continuity on auditory stream segregation. Perception and Psychophysics, 1973, 13, 308-312.

Bregman, A. S., & Pinker, S. Auditory streaming and the building of timbre. Canadian Journal of Psychology, 1978, 32, 19-31.

Bregman, A. S., & Rudnicky, S. I. Auditory segregation: stream or streams? Journal of Experimental Psychology: Human Perception and Performance, 1971, 1, 263-267.

Clopton, B. M., & Winfield, J. A. Unit responses in the inferior colliculus of rat to temporal auditory patterns of tone sweeps and noise bursts. Experimental Neurology, 1974, 42, 532-540.

Cole, R. A., & Scott, B. Perception of temporal order in speech: the role of vowel transition. Canadian Journal of Psychology, 1973, 27, 441-449.

Cutting, J. E. Two left-hemisphere mechanisms in speech perception. Perception and Psychophysics, 1974, 16, 601-612.

Dannenbring, G. L., & Bregman, A. S. Stream segregation and the illusion of overlap. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 544-555.

Dannenbring, G. L., & Bregman, A. S. Streaming vs. fusion of sinusoidal components of complex tones. Perception and Psychophysics, 1978, 24, 369-376.

Delattre, P.C., Liberman, A.M., & Cooper, F. S. Acoustic loci and transitional cues for consonants. Journal of the Acoustical Society of America, 1955, 27, 769-773.

Diamond, I. T., & Neff, W. D. Ablation of temporal cortex and discrimination of auditory patterns, Journal of Neurophysiology, 1957, 20, 300-315.

Erulkar, S. D., Butler, R. A., & Gerstein, T. L. Excitation and inhibition in cochlear nucleus to frequency modulated tones. Journal of Neurophysiology, 1968, 31, 537-548.

Evans, E. F., & Nelson, P. G. Responses of neurones in cat cochlear nucleus to modulated tonal stimuli. Journal of the Acoustical Society of America, 1966, 40, 1275-1276.

Evans, E. F., & Whitfield, I. C. Responses of cortical neurones to acoustic stimuli varying periodically in frequency. Journal of Physiology, 1964, 172, 52P.

Fletcher, H., & Munson, W. A. Loudness, its definition measurement and calculation. Journal of the Acoustical Society of America, 1933, 5, 82-108.

Frisby, J. P. Real and apparent movement - same or different mechanisms? Vision Research, 1972, 12, 1051-1056.

Gardner, R. B., & Wilson, J. P. Evidence for direction-specific channels in the processing of frequency modulation. *Journal of the Acoustical Society of America*, 1979, *66*, 704-709.

Halpern, L. The effect of harmonic ratio relationships on auditory stream segregation. Unpublished manuscript, McGill University, 1977.

Helmholtz, H. L. F. von. *Sensations of Tone*.(A. Ellis, trans.) New York: Dover Publications, 1954. (Originally published, 1862).

Heise, G. A., & Miller, G. A. An experimental study of auditory patterns. *American Journal of Psychology*, 1951, *64*, 68-77.

Henke, W. L. *MITSYN: An interactive dialogue language for time signal processing*. Cambridge: Michigan Institute of Technology, Research Laboratory of Electronics, 1975.

Hubel, D. H., & Wiesel, T. N. Receptive fields in the cat's striate cortex. *Journal of Physiology*, 1959, *148*, 574-591.

Hubel, D. H., & Wiesel, T. N. Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 1965, *28*, 229-289.

Jones, M. R. Time, our lost dimension: Toward a new theory of perception, attention and memory. *Psychological Review*, 1976, *83*, 323-345.

Jones, M. R., Maser, D. J., & Kidd, G. R. Rate and structure in memory for auditory patterns. *Memory and Cognition*, 1978, *6*, 246-258.

Julesz, B., & Hirsh, I. J. Visual and auditory perception: An essay of comparison. In E. E. David and P. Denes (Eds.), *Human Communication: A Unified View*. New York: McGraw-Hill, 1972.

Kelly, J. B., & Whitfield, I. C. Effects of auditory cortical lesions on discriminations of rising and falling frequency-modulated tones. *Journal of Neurophysiology*, 1971, *34*, 802-816.

Kinney, J. S. Discrimination in auditory and visual patterns. *American Journal of Psychology*, 1961, *74*, 529-541.

Koffka, K. *Principles of Gestalt Psychology*. New York: Harcourt, Brace and World, Inc., 1963 (Originally published, 1935).

Kolers, P. A. Aspects of Motion Perception. Oxford: Pergamon Press, 1972.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. Psychological Review, 1967, 74, 431-461.

Marks, L. E. The Unity of the Senses. New York: Academic Press, 1978.

Miller, G. A., & Heise, G. A. The trill threshold. Journal of the Acoustical Society of America, 1950, 22, 637-638.

Milner, P. M. A model for visual shape recognition. Psychological Review, 1974, 81, 521-535.

Møller, A. R. Coding of sounds with rapidly varying spectrum in the cochlear nucleus. Journal of the Acoustical Society of America, 1974, 55, 631-740.

Nábělek, I., & Hirsh, I. J. On the discrimination of frequency transitions. Journal of the Acoustical Society of America, 1969, 45, 1510-1519.

Nelson, A. G., Erulkar, S. D., & Bryan, J. S. Responses of units of the inferior colliculus to time-varying acoustic stimuli. Journal of Neurophysiology, 1966, 29, 834-860.

Noorden, L. P. A. S. van. Temporal Coherence in the Perception of Tone Sequences. Published Ph. D. thesis, Institute for Perception Research, Eindhoven, The Netherlands, 1975.

Norman, D. A. Temporal confusions and limited capacity processors. Acta Psychologica, 1967, 27, 293-297.

Ortmann, O. On the melodic relativity of tones. Psychological Monographs, 1926, 35, (1, No. 162).

Parsons, T. W. Separation of speech from interfering speech by means of harmonic selection. Journal of the Acoustical Society of America, 1976, 60, 911-918.

Pollack, I. Detection of rate of change of auditory frequency. Journal of Experimental Psychology, 1968, 77, 535-541.

Rasch, R. A. The perception of simultaneous notes such as in polyphonic music. Acustica, 1978, 40, 21-33.

Rubin, E. Visuell wahrgenommene Figuren. Copenhagen: Glydendalske, 1921.

Shirai, Y. Analyzing intensity arrays using knowledge about scenes. In P. H. Winston (Ed.), The Psychology of Computer Vision, New York: McGraw-Hill Book Company, 1975.

Stevens, K. N., & Klatt, D. H. Role of formant transitions in the voiced-voiceless distinction for stops. Journal of the Acoustical Society of America, 1974, 55, 653-659.

Thomas, J. B., Cetti, R. P., & Chase, P. W. Effect of silent intervals on the perception of temporal order for vowels. Journal of the Acoustical Society of America, 1971, 49, 85 (Abstract).

Thomas, J. B., & Fitzgibbons, P. J. Temporal order and perceptual classes. Journal of the Acoustical Society of America, 1971, 50, 86-87 (Abstract).

Tsumura, T., Sone, T., & Nimura, T. Auditory detection of frequency transition. Journal of the Acoustical Society of America, 1973, 53, 17-25.

Uhr, L. Pattern Recognition, Learning and Thought. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973.

Vartanian, I. A. On mechanisms of specialized reactions of central auditory neurons to frequency-modulated sounds. Acustica, 1974, 31, 305-310.

Warren, R. M., Obusek, F. J., Farmer, R. M., & Warren, R. P. Auditory sequence: Confusion of patterns other than speech or music. Science, 1969, 164, 586-587.

Whitfield, I. C., & Evans, E. F. Responses of auditory cortical neurons to stimuli of changing frequency. Journal of Neurophysiology, 1965, 28, 655-672.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, 1962.