Perception & Psychophysics
1990, 47 (1), 68-73

Fusion of auditory components: Effects of the frequency of amplitude modulation

ALBERT S. BREGMAN, ROBERT LEVITAN, and CHRISTINE LIAO
McGill University, Montreal, Quebec, Canada

The fusion of two amplitude-modulated (AM) tones presented simultaneously was studied. Subjects were presented with an AM tone (A) followed by a copy of itself (B) which was accompanied by another AM tone (C). In different experiments, the subjects were asked either to rate how clearly they could hear Tone B in the BC mixture or whether Tone B was present or not. The stronger the fusion of Tones B and C, the harder it was to "hear out" Tone B. It was found that fusion was strongest when Tones B and C were modulated at the same rate; segregation curves were obtained for the degree of mismatch of modulation frequency.

In the usual complex listening situation, the pressure waves arising from many sources mix together as they enter the ear. An important part of pattern recognition is the recovery of separate descriptions of some of the individual sources of sound that have contributed to the mixture. Therefore, the study of auditory segregation and its opposite, fusion, is the study of the methods used by the auditory system to decompose mixtures of sounds and to group their component frequencies or features together into appropriate descriptions (Bregman, in press).

There are two aspects of this grouping, which we can refer to as *sequential grouping* and *simultaneous grouping*. In sequential grouping, we must decide on the sequence of acoustic components that arise over time from the same source. In simultaneous grouping, the set of spectral components arriving at any one instant must be decomposed into the contributions of the individual sources.

It appears that there are heuristics for both sequential and simultaneous grouping that collaborate and compete in order to arrive at a correct decomposition of the incoming signal (Bregman, 1978, 1981). In simultaneous grouping, or fusion, many useful clues arise from the behavior of spectral components over time. It has, for example, been shown that frequency components with synchronous onsets and offsets tend to be heard as the spectral components of a single sound (Bregman & Pinker, 1978; Dannenbring & Bregman, 1978; Rasch, 1978). Generally speaking, the factors that promote the fusion of spectral components can be described loosely by the Gestalt term, "common fate" (see Koffka, 1935). If spectral components are doing the same thing at the same time (e.g., going on and off in synchrony) they probably "belong together" as parts of the same signal (Bregman, 1978, 1981).

One natural example occurs when a sound (e.g., human speech) is heard against a background of other sounds. The independent sound sources—target and background—will tend to have their own independent patterns of rise and fall in intensity. Evidence from experiments on "comodulation release from masking" has shown that the auditory system can use the correlations in the intensity variations of the different parts of the background spectrum to shield the target from masking (see Grose & Hall, 1989; Haggard, Harvey, & Carlyon, 1985; Hall, 1986; Hall, Haggard, & Fernandes, 1984; Hall, Haggard, & Harvey, 1984; McFadden, 1986).

A second natural case also occurs with speech. This time it is not the the parts of the signal itself that show common amplitude fluctuation, but the parts of the nervous system that have been stimulated by parts of the signal. It can be briefly summarized as follows: As we listen to speech, any individual basilar membrane (BM) "filters" that respond to more than one frequency component will output an amplitude-modulated (AM) signal. A match in the AM frequency in the signals being passed by two such filters can indicate that both filters are responding to frequency components that have arisen from the same signal.

Bregman, Abramson, Doehring, and Darwin (1985) have shown that AM information can serve as an independent clue as to whether acoustic components that occur in different parts of the spectrum have arisen from the same acoustic source. They reported evidence that when two sinusoidal tones, in the regions of the first and second formants of the human voice, respectively, were modulated by the same frequency of AM (about 100 Hz), they tended to be fused more strongly than they would have been if the AM had not been at the same frequency. This was true whether the AM was produced by the beating of harmonically related or harmonically unrelated components. Therefore, the AM itself, independently of the harmonic relationships among the partials, seems to

This research was carried out with the assistance of a grant from the Natural Sciences and Engineering Research Council of Canada, and a Killam Research Fellowship, held by the first author. We acknowledge the assistance of Pierre Abdel-Ahad. Reprints are available from the first author at the Psychology Department, McGill University, 1205 Dr. Penfield Ave., Montreal, Quebec H3A 1B1, Canada.

promote fusion. We know that it is easier for a listener to segregate and identify speech sounds in a mixture when the component voices are at different fundamental frequencies (Brokx & Noteboom, 1982; Halikia, 1985; Scheffers, 1979, 1983). It was proposed by Bregman et al. that an AM frequency-detection mechanism could assist this segregation by favoring the fusion of the formants derived from the same voice.

In the present experiments, following those of Bregman et al. (1985), we used a simpler stimulus than the human voice. We studied the perceptual fusion of a high-frequency and a low-frequency sinusoid as a function of the similarity of the frequency of AM applied to them. Bregman et al. studied the effects of a 5-Hz mistuning of the frequencies of AM applied to two carrier tones and found only a small effect. The present experiments were designed to study fusion over a larger range of mistuning and at different combinations of carrier frequencies.

The degree of fusion was measured using a method reported by Bregman and Pinker (1978). In their experiment, subjects listened to a repeating cycle in which a single, pure tone, A, preceded a mixture of two pure tones, B and C. Tones B and C occurred more or less simultaneously and Tones A and B were at about the same frequency. The perception of the cycle could vary between two extremes, depending upon the conditions of presentation. At one extreme, Tones B and C fused to form a rich tone, BC, which could be heard as alternating with the pure tone A. At the other extreme, Tone A captured Tone B into a sequential cycle so that it did not fuse with Tone C; in this case, the listener heard two pure high-pitched tones on each cycle, A and B, accompanied by a low-pitched tone, C.

In the present experiments, we used a similar cycle of stimuli, except that Tones A, B, and C were amplitude-modulated sinusoids. Tones A and B were always identical, and Tones B and C always started and stopped in synchrony. The main factor that was varied was the frequency difference between the AM applied to Tones B and C.

A secondary factor was the "harmonicity" of Tone C. In some stimuli, Tone C and its "sidebands" fell into a harmonic series, whereas in others they did not. The method for producing harmonic and inharmonic tones can be explained as follows. Suppose we fully modulate a sinusoidal carrier tone of frequency C by a raised cosine whose frequency is MF. Viewed in the frequency domain, this modulation adds two additional frequencies, or sidebands, to accompany the original frequency, C. These are sinusoids (partials) having frequencies C + MF and C − MF and having half the amplitude of tone C. If there exists some fundamental frequency, f_0, such that the carrier frequency C is one of its harmonics, and if the modulation frequency MF is exactly at this fundamental or at an integer multiple of it, then Tone C and its sidebands will be members of the harmonic series of f_0. Otherwise, they will not be. We manipulated harmonic relationships to find out whether the AM effects were independent of the effects of good harmonic relationships between Tones B and C. This was intended as a replication of the findings of Bregman et al. (1985).

A third factor that was varied in the present experiments was the spectral location of Tones B and C.

Experiments 1A and 1B were similar in the tasks used. Listeners were asked to rate the degree to which Tone B was independently audible in the complex tone BC. The carrier frequency of Tone B was positioned at 3000 Hz and that of Tone C at about 2000 Hz. These locations were chosen because they are at prominent formant locations for English vowels. In addition, the AM frequency applied to Tone C was chosen to generate either harmonic or inharmonic sidebands. Experiment 1B was done as a quick check on the general findings of Experiment 1A when it was suspected that the intermodulation distortion introduced by the use of analog tape recorders in Experiment 1A could have influenced the results. Rather than being an exact replication, in Experiment 1B we used stimuli in a different frequency range, so that if the results turned out to be consistent with those of Experiment 1A, they would not only replicate but would extend the scope of the findings. The results of the two experiments were quite similar despite the changes.

Experiments 2A and 2B were done in order to repeat the conditions of Experiments 1A and 1B using an accuracy measure rather than a rating of segregation. Although the same methodology was used, Experiments 2A and 2B are distinguished from one another because they were run at two separate times with different subjects and different carrier frequencies.

EXPERIMENTS 1A AND 1B

Method

Experiment 1A

Procedure. The subjects were told that on each trial they would hear 10 repetitions of a cycle of two sounds, of which the first member was the "target" and the second a "mixture" that contained the target. Their task was to listen carefully to the target and to rate how clearly they could hear it repeating again in the mixture. A rating of 1 was to be used if the target could not be heard at all in the mixture and a rating of 10 meant that it could be very clearly heard.

Stimuli. In all conditions of Experiment 1A, both Tone A and Tone B consisted of a 3000-Hz carrier tone modulated by a raised cosine of 125 Hz. Since 3000 Hz is the 24th harmonic of the modulating frequency (MF), the sidebands are the 23rd and 25th harmonics, and the three partials are part of the same harmonic series.

All of the carrier and modulation frequencies are shown in Table 1. In Experiment 1A, three alternative carrier frequencies were used for Tone C: 2000, 2050, and 1951 Hz. Each was modulated by a range of MFs both higher and lower than the 125-Hz MF applied to Tone B. It was expected that as the MFs of Tones B and C drew closer together, Tones B and C would show greater fusion.

The MFs used for the 2000-Hz carrier tone were always chosen so that the carrier and its sidebands were harmonics of the chosen MF. This was achieved by choosing MFs that divided evenly into

Table 1
AM Frequency (in Hz) Used with Each Carrier Frequency in Experiments 1A, 1B, 2A, and 2B

Carrier Frequency of Tone C	AM Frequency								
	Experiments 1A and 2A								
2000 Hz	200.0	166.7	153.8	142.9	125.0	111.1	100.0	90.9	80.0
2050 Hz	195.2	178.2	151.8	141.4	125.0	110.8	100.0	91.1	80.4
1951 Hz	201.1	169.6	156.1	144.5	125.0	111.5	100.0	90.7	79.6
	Experiments 1B and 2B								
1600 Hz	177.8	145.4	133.3*	123.1*	114.3*	106.7*	100.0*	94.1	84.2
1657 Hz	174.4	149.1	132.6	122.7	114.3	106.9	100.4	94.7	85.0
1545 Hz	181.8	147.1	134.3	123.6	114.3	106.6	99.7	93.7	83.5

Note—In Experiments 1A and 2A, carrier frequency (CF) of Tone B was 3000 Hz; modulation frequency (MF) was 125 Hz. In Experiments 1B and 2B, CF of Tone B was 2400 Hz; MF was 114.3 Hz. *Used only in Experiment 1B.

the carrier frequency. For the 2000-Hz series, therefore, the C tone was always harmonic within itself but its partials fell into the same harmonic series as those of Tone B when the 125-Hz MF was used for Tone C (as well as for Tone B). (One should expect the strongest effects of the tuning and mistuning of the MFs of Tones B and C when both are modulated at the same rate and the harmonics of both tones fall into the same series.) When the 2000-Hz carrier was not modulated at the same rate as Tone B, not only did this C complex fail to fit the harmonic series of Tone B, but it formed a different harmonic series of its own.

When the 2050- and 1951-Hz carriers were used to construct Tone C, the MFs were chosen so as to avoid creating harmonic complexes within Tone C. Whereas the MFs for the 2000-Hz carrier were chosen by dividing the carrier frequency by an integer, those for the other carriers were derived by dividing the carrier frequency by 10.5, 11.5, 12.5, and so forth, thus ensuring a nonharmonic complex. Not all MFs derived in this way were actually used; only those close in value to those used for the 2000-Hz series were employed. The one exception to the generation rule was that the 125-Hz MF was used with all three Tone C carriers to study what would happen when their MFs matched that of Tone B. However, only in the case of the 2000-Hz carrier did this MF create a series of harmonics within Tone C (with the same fundamental frequencies as those of Tone B).

Each of Tones A, B, and C had the following structure. They were 398 msec long including linear onset/offset amplitude transitions of 20 and 9 msec, respectively. We were concerned that in the conditions in which Tones B and C were modulated at different rates, the first (or last) audible burst in the tone with the higher MF might reach some threshold of audibility before the other tone, leading to segregation by onset/offset asynchrony (Bregman & Pinker, 1978; Dannenbring & Bregman, 1978) rather than by AM frequency. Therefore, we superimposed a white-noise burst over the first and last 10 msec of Tones B and C to render ambiguous the exact starting and ending points of the signal. The noise burst was set at about the same subjective loudness as the tone, as judged by the experimenter.

Each trial consisted of 10 cycles of the A-B-C pattern, followed by 5.7 sec during which the subjects wrote down their answers. Each cycle consisted of Tone A for 398 msec, a 947-msec silence, Tone B accompanied by Tone C for 398 msec, and a 1,288-msec silence.

Eight tapes were created, each containing a different random order of the 27 experimental conditions (nine degrees of mistuning of the AM of Tones B and C, with three carrier frequencies for Tone C). Before being tested on each tape, the subjects were given six practice trials chosen from the 27 conditions to exemplify a representative range of degrees of segregation of the target. These were given to anchor (or reanchor) the judgment scale prior to each run of trials. All eight tapes had the same six practice conditions, in one of three random orders. Each subject judged all eight tapes.

Apparatus. The stimuli were digitally synthesized on a PDP 11/34 computer (Digital Equipment Corporation) using the MITSYN signal-processing system (Henke, 1980). Two channels were used in the synthesis, with channel 1 being used for Tones A and B and channel 2 for Tone C. Synthesized files were output through a 12-bit D/A converter at 13,513 samples per second per channel and low-pass filtered at 4.8 kHz to avoid aliasing, using a Rockland Model 852 filter at the Flat Delay setting. In Experiment 1A, the stimuli were recorded stereophonically on an Akai 4000-SS reel-to-reel tape deck at 7.5 ips. For playback, the signals were output from the tape deck and again low-pass filtered as above to remove noise. They were then passed through a mixer, where the amplitude levels were set. A preamplifier was used to mix the channels to form a binaural signal, which was then passed through a stereo amplifier where the left and right headphone intensities were balanced. The subject listened through Telephonics TDH-49P headphones in an IAC 1202 audiometric test chamber. The levels were 67 dBA for Tones A and B and 71.5 dBA for Tone C as measured by a General Radio 1551-C sound-level meter with a flat-plate coupler. These intensities were chosen on the basis of a pilot experiment in which the level of Tone C relative to Tone B was varied. The intensities mentioned above were chosen to avoid either the complete swamping of the perception of Tone B by Tone C or a too-easy segregation of Tone B from Tone C.

Subjects. Four subjects with experience in listening to experimental stimuli (including one of us—A.S.B.) were used because of the relative difficulty of the discrimination in this experiment. They were all male and varied in age from 21 to 45 years. Each subject listened to the same eight tapes, with each tape consisting of six practice trials followed by 27 experimental trials.

Experiment 1B

A few changes were made in Experiment 1B relative to Experiment 1A. The rating scale ran from 1 to 7 instead of from 1 to 10. Also, a different set of carrier frequencies were used. Those of Tones A and B were at 2400 Hz instead of 3000 Hz and their MF was always set at 114.3 Hz. For Tone C, only one carrier, 1600 Hz, was used. There were only five stimulus patterns instead of the 27 used in Experiment 1A. The MFs that were used for Tone C in Experiment 1B are those marked with an asterisk in Table 1. The MF always generated a harmonic series within Tone C, but only at 114.3 Hz did it generate a harmonic series that matched that of Tone B.

Stimulus. The stimulus pattern was also slightly different than in Experiment 1A. The tones were lengthened to 430 msec, including 13-msec rise/fall times. The tones were gated on and off by a quarter-sine-wave function. The cycle structure was as follows: Tone A, 0.5 sec silence, Tone B, 0.68 sec silence. There were only 8 repetitions of the stimulus cycle instead of 10. There were 3 sec between trials.

Apparatus. The major difference in the apparatus lay in the replacement of the analog tape recorder by a digital tape recorder consisting of a Sony PCM701-ES digital audio processor and a videotape recorder. Although it would have been better to run the experiment directly from the computer in both experiments, the large number of experimenters using a single laboratory computer made on-line experiments impossible.

Subjects. There were 15 young adult subjects, drawn from a college student population.

Results

First we discuss the results for Experiment 1A. Figure 1 shows the mean segregation ratings for Tone B when it was accompanied by each of the nine MFs of Tone C. This is shown for each of the three carrier frequencies. Each point represents the mean of 32 ratings (eight judgments × 4 subjects). Low scores indicate a high degree of fusion of Tone B with Tone C. All three curves show a sharp increase in fusion when the MF of Tone C was the same as that of Tone B (i.e., 125 Hz). The consistency of the effects for the three carrier frequencies suggests that the main results do not depend upon the existence of a matching harmonic series in the B and C complexes, but are a pure effect of the matching pulsation rates. The curves begin to dip down again at the lower modulation frequencies, probably because they are approaching the condition where the MF of Tone C would be half of the MF of Tone B. Although there was some variation among the curves of the 4 subjects, each subject rated the 125-Hz MF as showing the strongest fusion (least segregation) for each of the three carriers.

The results for Experiment 1B are shown in Figure 2. Basically the same pattern is found as for the analogous set of conditions in Experiment 1A (the 2000-Hz carrier), where Tone B is harmonic within itself. The replication again shows a slightly greater fusion for the 2000-Hz C tone (in which the harmonics of Tones B and C fall into the same harmonic series) at the tuned frequency.

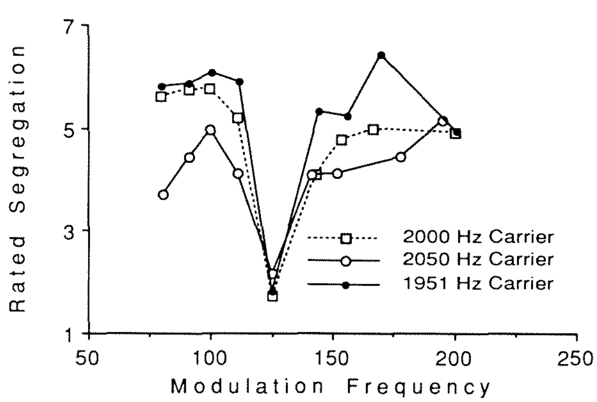

Figure 1. Mean segregation ratings for the target (Tone B) for the three different carrier frequencies (CF) and the nine different modulation frequencies (MF) of Tone C in Experiment 1A. The target always had CF = 3000 Hz and MF = 125 Hz.

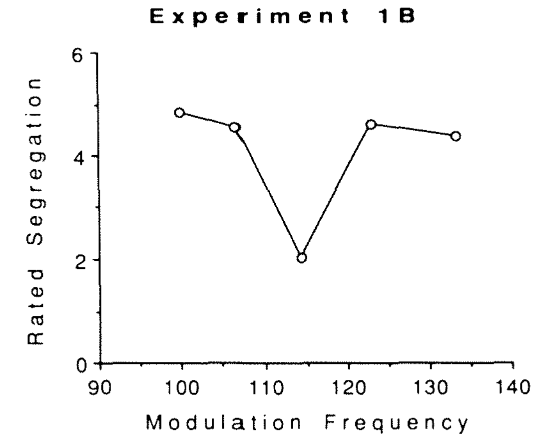

Figure 2. Mean segregation ratings for the target (Tone B) for the three different carrier frequencies (CF) and the five different modulation frequencies (MF) of Tone C in Experiment 1B. The target always had CF = 2400 Hz and MF = 114.3 Hz.

Discussion

The results of Experiments 1A and 1B confirm those of Bregman et al. (1985) in showing the effect of AM frequency matches in promoting fusion. They also confirm the previous results in showing that this effect does not depend on the harmonic series created by the modulation. However, there was a tiny difference in fusion at the 125-Hz MF for the three carrier frequencies: the 2000-Hz frequency with harmonics all related to 125 Hz showed slightly more fusion. This difference was seen again in the subsequent experiments.

The similar pattern in Experiment 1B shows that the segregation curve was probably not an artifact of the apparatus or of the particular choice of frequencies.

EXPERIMENTS 2A AND 2B

These experiments were done in order to repeat the conditions of Experiments 1A and 1B using an accuracy measure rather than a rating of segregation. Although the same methodology was used, Experiments 2A and 2B are distinguished from one another because they were run at two separate times with different subjects and different carrier frequencies.

Method

Stimuli

The stimuli were identical in form to those used in Experiments 1A and 1B, and were synthesized and presented using the same equipment, including the same analog tape recorder as in Experiment 1A. As in Experiments 1A and 1B, Tones A and B were always at the same fixed frequency of both carrier and AM, whereas the frequency parameters of Tone C were varied. However, the task was different. In Experiments 2A and 2B, the B tone was absent on half the trials. On each trial, the listeners decided whether or not they had heard Tone B, answering yes or no, and then rating, on a 3-point scale, the "clarity" with which they had heard the presence or absence of Tone B. Since the subjects were scored on accuracy, these experiments constituted, in effect, a study of masking. Each experimental trial consisted of eight repetitions of the cycle of tones.

In Experiment 2A, Tones A and B were set at 3000 Hz and Tone C in the vicinity of 2000 Hz. As in Experiment 1A, the 27 conditions of MF and carrier frequency shown in the first three rows of Table 1 were used. Each such condition appeared both with and without Tone B, yielding 54 conditions. Each subject judged a different randomization of these 54 conditions in each of four sessions.

Experiment 2B was just like Experiment 2A, except that the set of carrier and AM frequencies shown in the last three rows of Table 1 was used. As in Experiment 2A, the ratio of the carrier frequencies to one another was 3:2, but this time the carriers were lower in frequency; Tones A and B were at 2400 Hz and Tone C was in the neighborhood of 1600 Hz. The MF of 114.29 Hz was chosen for the target because it is a common divisor of 1600 and 2400 Hz. This allowed the target (Tone B) to be a harmonic complex and, as in Experiment 1A, allowed for a set of harmonic conditions (based on the 1600-Hz carrier) and two sets of inharmonic conditions (based on carriers of 1657 and 1545 Hz) for Tone C.

In Experiment 2A, both Tone A and the target tone, B, were presented at 61 dBA, with Tone C (the masker) at 80 dBA. In Experiment 2B, Tone B, but not Tone A, was raised to 67 dBA to achieve approximately the same level of detectability as in Experiment 2A (as shown by pilot testing).

Procedure

The subjects saw a diagram showing two cycles of tones, one with Tone B present and one with Tone B absent. They were told to employ two strategies when trying to detect Tone B. One was to focus on Tone A and to try to hear it repeating in the second sound. The second, recommended for the purpose of reducing false positives, was to listen to the second sound closely and to ask oneself whether there was one sound present or two. If they heard only one, then they were to answer "no," and if they heard two or more (i.e., a mixture), and one of the components sounded like Tone A, even if lower in volume than Tone A, they were to answer "yes." Before the actual test trials were given, the subjects listened to a demonstration tape, which contained an assortment of conditions with Tone B present and absent. The level of the masker was varied and the subjects were informed as to the presence or absence of the target.

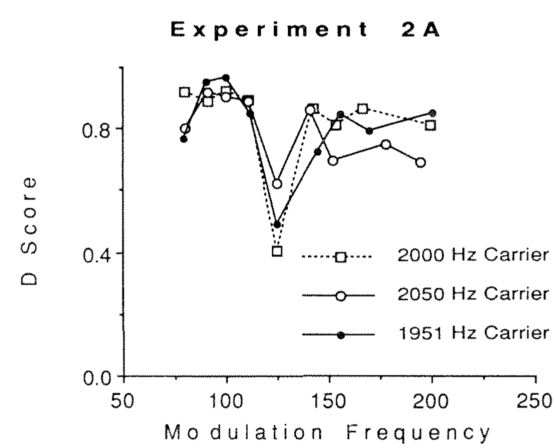

Figure 3. Mean accuracy (D) scores for Experiment 2A as a function of the three carrier frequencies (CF) and the nine modulation frequencies of Tone C. The target (Tone B) always had CF = 3000 Hz and MF = 125 Hz.

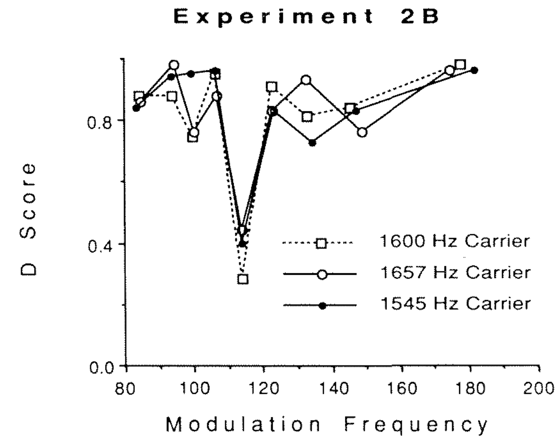

Figure 4. Mean accuracy (D) score for Experiment 2B as a function of the three carrier frequencies (CF) and nine modulation frequencies of tone C. The target (Tone B) always had CF = 2400 Hz and MF = 114.3 Hz.

Subjects

There were 12 subjects in Experiment 2A and 5 in Experiment 2B, selected from a population of young adults working or studying at McGill University.

Results and Discussion

The yes-no judgments and clarity ratings were combined into a 6-point scale running from −3 (clear no) to +3 (clear yes). These were then used to calculate a "D" score for each of the 27 carrier × MF conditions. D is an easily calculated nonparametric measure of the discrimination of the presence or absence of Tone B, which does not assume that the judgments of different subjects, or of the same subject in different conditions, have the same distributions on the rating scale (Bregman & Campbell, 1971). D will have a value of +1 when there is perfect discrimination, 0 when judgments are random and −1 when they are systematically reversed.

Figures 3 and 4 show the segregation curves for detectability of Tone B in Experiments 2A and 2B, respectively, as a function of the AM frequency of Tone C for each of the three carrier frequencies. The same pattern can be seen in these results as in those of Experiments 1A and 1B. Again, Tone B was less audible as a separate sound when its AM matched that of Tone C. In Experiments 2A and 2B, there was a stronger suggestion than in Experiments 1A and 1B that when the AM frequencies of Tones B and C were the same, fusion was strongest when Tone C was a harmonic complex (i.e., when the frequency components of the B and C complexes all fit into the same harmonic series). However, there was no evidence that harmonic complexes are more easily segregated than inharmonic ones when the AM frequencies of Tones B and C do not match. This suggests that under the conditions of the present experiments, the auditory system can combine two sources of information about the presence of a single harmonic series, but cannot benefit from information that two such series are simultaneously present.

CONCLUSIONS

These experiments were able to show effects of AM mistuning much larger than those shown by Bregman et al. (1985), by using a larger range in the degree of mistuning. They also showed that a measure of masking gave similar results to those of a rating of clarity. The experiments taken together suggest that the pulsing that occurs in the neural outputs at two BM locations promotes the fusion of the spectral information that arrives through these two channels. This occurs even when the sideband frequencies do not themselves promote fusion by virtue of belonging to the same harmonic series. However, the fusion seems to be improved slightly when a harmonic series is involved.

When a human voice is registered in the auditory system, it is probable that the system can detect both a pulsing within each BM filter and a regular harmonic series. There is reason to believe, given the present experiments, that the auditory system makes use of both kinds of information. It is not appropriate, however, to conclude from these experiments that the pulsation-based mechanism is more important in the grouping of spectral regions than is one based on harmonicity. Although this may have been true in our experiments, in which there were very few partials to establish the harmonic series, it may not be true in the human voice, which is rich in harmonics.

REFERENCES

BREGMAN, A. S. (1978). The formation of auditory streams. In J. Requin (Ed.), *Attention and Performance* (Vol. 7, pp. 63-76). Hillsdale, NJ: Erlbaum.

BREGMAN, A. S. (1981). Asking the "what for" question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 99-118). Hillsdale, NJ: Erlbaum.

BREGMAN, A. S. (in press). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

BREGMAN, A. S., ABRAMSON, J., DOEHRING, P., & DARWIN, C. J. (1985). Spectral integration based on common amplitude modulation. *Perception & Psychophysics*, **37**, 483-493.

BREGMAN, A. S., & CAMPBELL, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, **89**, 244-249.

BREGMAN, A. S., & PINKER, S. (1978). Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, **32**, 19-31.

BROKX, J. P. L., & NOTEBOOM, S. G. (1982). Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics*, **10**, 23-36.

DANNENBRING, G. L., & BREGMAN, A. S. (1978). Streaming vs. fusion of sinusoidal components of complex tones. *Perception & Psychophysics*, **24**, 369-376.

GROSE, J. H., & HALL, J. W. (1989). Comodulation masking release using SAM tonal complex maskers: Effects of modulation depth and signal position. *Journal of the Acoustical Society of America*, **85**, 1276-1284.

HAGGARD, M. P., HARVEY, A. D. G., & CARLYON, R. P. (1985). Peripheral and central components of comodulation masking release. *Journal of the Acoustical Society of America*, **78**(Suppl. 1), S63.

HALIKIA, M. H. (1985). *The perceptual segregation of simultaneous sounds.* Unpublished doctoral dissertation, McGill University, Montreal.

HALL, J. W. (1986). The effect of across-frequency differences in masking level on spectro-temporal pattern analysis. *Journal of the Acoustical Society of America*, **79**, 781-787.

HALL, J. W., HAGGARD, M. P., & FERNANDES, M. A. (1984). Detection in noise by spectro-temporal pattern analysis. *Journal of the Acoustical Society of America*, **76**, 50-56.

HALL, J. W., HAGGARD, M. P., & HARVEY, A. D. G. (1984). Release from masking through ipsilateral and contralateral comodulation of a flanking band. *Journal of the Acoustical Society of America*, **76**(Suppl. 1), S76.

HENKE, W. L. (1980). *MITSYN: A coherent family of command-level utilities for time signal processing* (revised and extended by D. Salomon, McGill University). Cambridge, MA: Author.

KOFFKA, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace & World.

MCFADDEN, D. (1986). Comodulation masking release: Effects of varying the level, duration, and time delay of the cue band. *Journal of the Acoustical Society of America*, **80**, 1658-1667.

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

SCHEFFERS, M. T. M. (1979). The role of pitch in perceptual separation of simultaneous vowels. *IPO Annual Progress Report*, **14**, 51-54.

SCHEFFERS, M. T. M. (1983). *Sifting vowels: Auditory pitch analysis and sound segregation.* Unpublished doctoral dissertation, Groningen University, Groningen, The Netherlands.

(Manuscript received July 27, 1988;
revision accepted for publication June 19, 1989.)