Auditory continuity and loudness computationa)

Stephen McAdams, Marie-Claire Botte, and Carolyn Drake

Citation: The Journal of the Acoustical Society of America **103**, 1580 (1998);
View online: https://doi.org/10.1121/1.421293
View Table of Contents: http://asa.scitation.org/toc/jas/103/3
Published by the Acoustical Society of America

Auditory continuity and loudness computation[a)]

Stephen McAdams[b)]
Laboratoire de Psychologie Expérimentale (CNRS), Université René Descartes, EPHE, 28 rue Serpente, F-75006 Paris, France and Institut de Recherche et de Coordination Acoustique/Musique (IRCAM), 1 place Igor-Stravinsky, F-75004 Paris, France

Marie-Claire Botte and Carolyn Drake
Laboratoire de Psychologie Expérimentale (CNRS), Université René Descartes, EPHE, 28 rue Serpente, F-75006 Paris, France

(Received 21 August 1996; revised 6 October 1997; accepted 4 November 1997)

Sequences composed of alternating bursts of different levels with no silences separating them can give rise to a perception of a continuous sound upon which is superimposed an intermittent stream. These experiments sought to determine how the perceived loudness of the intermittent stream depends on the level difference between higher-level and lower-level bursts in the sequence in cases in which continuity is either heard or not heard. In the main experiment, listeners were asked to adjust the level of continuous or intermittent comparison sequences to match the loudness of components that appeared to be either continuous or intermittent in an alternating-level reference sequence, thus urging them to focus on the two-stream percept. Loudness matches of the continuous comparison stimulus were close to physical levels of the lower-level bursts, whereas matches of the intermittent comparison stimulus were well below the physical levels of higher-level bursts. These results are discussed in terms of Bregman's [*Auditory Scene Analysis* (MIT, Cambridge, MA, 1990)] ''old-plus-new'' hypothesis: The loudness of the intermittent stream should result from the subtraction of the lower level from the higher level under the assumption that the higher-level burst represents a simultaneous mixture of sounds including the continuation of the lower-level burst. Additional experiments verified that, in the absence of the continuity phenomenon, matched levels were very close to the physical levels and that matches to fixed-level continuous and intermittent sequences were precise. The matching results from the main experiment support predictions of neither classical loudness models that do not take auditory organization processes into account nor schema-based models that presume a selection of information from the higher-level burst that does not affect the perceptual content of this burst. The matched levels fell between predictions of models based on subtraction of acoustic pressure and acoustic power, but were very different from subtraction of loudness measured in sones, suggesting that loudness is computed subsequent to auditory organization processes.
[S0001-4966(98)04602-5]

PACS numbers: 43.66.Cb, 43.66.Mk, 43.66.Lj [WJ]

INTRODUCTION

To recover a veridical representation of the acoustic environment, it would be useful for the auditory system to be able to group together acoustic components that originate from the same source into coherent mental descriptions (variously referred to as auditory ''streams,'' ''objects,'' ''images,'' or ''entities''). Once the streams are organized, the auditory system can compute the perceptual attributes (loudness, pitch, timbre, etc.) of the events belonging to each stream. Our experiments aimed to measure the effect of auditory organization on the computation of loudness.

Consider the stimulus sequence in Fig. 1(a) in which a pure tone or a noise signal alternates between a higher level (L_H) and a lower one (L_L). There are several ways that such a signal might be generated, three of which are shown in Fig. 2. Given that the signal might result from several acoustic configurations, it is interesting to understand how the auditory system analyses the situation. One might imagine that a listener would hear an alternating sequence of loud and soft tones or noise bursts (hypothesis 1). If this were the case, one would expect that when listeners are asked to adjust the level of a comparison stimulus to match the loud or soft parts of the reference stimulus, we should obtain matches in the vicinity of L_H and L_L, respectively, with perhaps some deviations due to temporal masking effects and temporal integration of energy within each burst. Such a prediction would be made by classical time-varying loudness models (Zwicker, 1977). This class of model, as currently implemented in several so-called psychoacoustic measurement devices available on the market, does not consider the incident waveform to be

---

[a)]Preliminary work leading to this study was reported at the Troisième Congrès Français d'Acoustique, Toulouse (McAdams *et al.*, 1994a). Portions of the present data were first presented at the ATR Workshop on ''A Biological Framework for Speech Perception and Production,'' Kyoto (McAdams *et al.*, 1994b).

[b)]Address correspondence to S. McAdams, Laboratoire de Psychologie Expérimentale, 28 rue Serpente, F-75006 Paris, France, Electronic mail: smc@ircam.fr

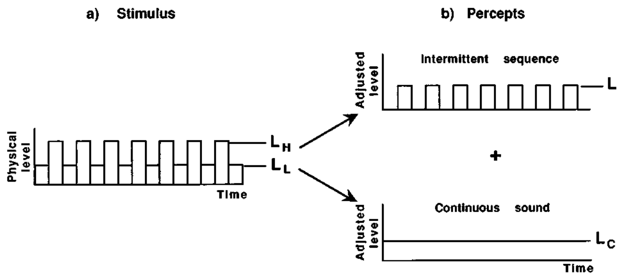

FIG. 1. (a) Stimulus sequence alternating between levels L_L and L_H. (b) Percepts resulting from the alternating sequence: a continuous sound with perceived level L_C and a sequence of intermittent bursts with perceived level L_I.

composed of temporally overlapping sound signals originating from separate sources of overlapping or even identical frequency content.[1]

However, Warren *et al.* (1972) have demonstrated that such a stimulus can be heard in a different way. A level difference between L_H and L_L of about 3–10 dB generally gives rise to a perception not of alternation but of an intermittent sequence superimposed on a continuous sound [Fig. 1(b)]. This perception of continuity is heard under certain conditions that depend on the difference between levels L_L and L_H (Houtgast, 1972; Thurlow, 1957; van Noorden, 1977), on the abruptness of the transition (Bregman and Rousseau, 1991), on the presence or absence of silences between bursts (van Noorden, 1977), as well as on other secondary factors (Houtgast, 1972; Thurlow, 1957; van Noorden, 1975; Verschuure *et al.*, 1976). This phenomenon has been called ''auditory continuity'' (Thurlow and Elfner, 1959) since the low-level signal is heard to continue through the intermittent signal, or ''auditory induction'' (Warren *et al.*, 1972) since the high-level part induces a perception of continuity in the low-level part. Similar types of phenomena have been demonstrated for speech interrupted by noise, in which a continuous speech signal appears to be perceptually restored during the noise burst (see Warren, 1984, for a review). In these cases, the auditory system would appear to have interpreted the signal as being composed of a continuous sound upon which another signal is superimposed. This interpretation depends on the presence of contextual evidence that the restored sound may be present, i.e., there must be no evidence that the low-level sound stopped and, in addition, the peripheral units stimulated by the interrupting sound must include those that would be stimulated by the anticipated fainter sound (Bregman, 1990; Warren *et al.*, 1972). Of particular interest to our present concerns is what this phenomenon might tell us about how the auditory system disentangles the respective perceptual attributes of superimposed signals.

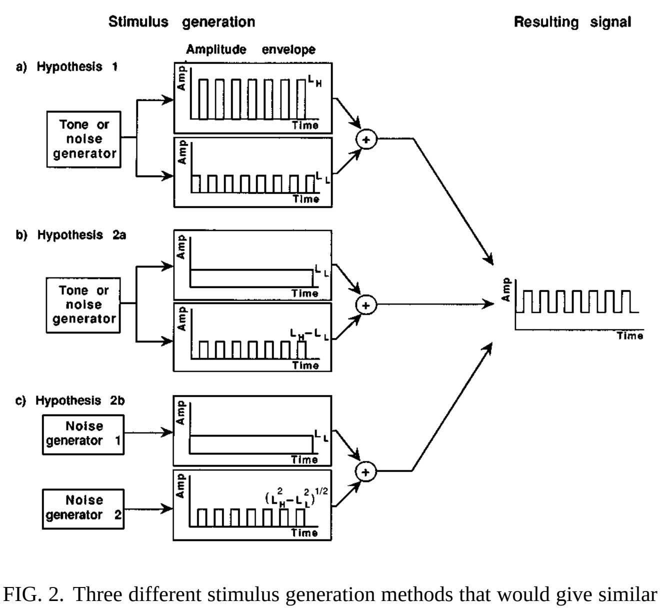

FIG. 2. Three different stimulus generation methods that would give similar resulting signals. Methods a, b, and c correspond to hypotheses 1, 2a, and 2b, respectively (see text).

Bregman (1990) proposes the existence of a general-purpose, bottom-up perceptual heuristic, called the ''old-plus-new'' strategy, which makes some qualitative predictions about this phenomenon: an interpolation is performed between the properties of the lower-level sounds occurring before and after the higher-level interrupting sound (Ciocca and Bregman, 1987). However, this computation is performed only if the auditory information indicates that the low-level sound could have been present during the occurrence of the high-level sound and that the transition between the high-level and low-level sounds is not a continuous one. Subsequently, the signal in the time interval occupied by the high-level sound is interpreted as resulting from a mixture of the low-level (old) sound and an additional (new) sound. The computation of the loudness of the intermittent stream would thus be based on a *subtraction* of the level of the restored part of the continuous sound from the global level of the intense part of the sequence. This kind of subtractive mechanism was first proposed by Warren (1982).[2]

According to this strategy, if we ask listeners to adjust the level of a comparison stimulus to match the loudness of either the continuous or the intermittent parts of this reference stimulus (see Fig. 1), we should obtain an adjusted level in the vicinity of L_L for the continuous part (L_C) and an adjusted level for the intermittent part (L_I) that would depend on the underlying psychological scale used by the subtraction mechanism to derive the loudness of this latter part. If a continuous sound and an intermittent sequence of identical frequency content are added in phase and presented to one or both ears [Fig. 2(b)], one might expect a law computed on acoustic pressure ($L_I = L_H - L_L$, where L is in units of pressure) (hypothesis 2a), whereas for similar, independent stimuli with incoherent phase relations [Fig. 2(c)] (e.g., signals of unknown properties, or even known signals presented in a reverberant environment), one might expect a law computed on acoustic power ($L_I^2 = L_H^2 - L_L^2$, where L is in units of pressure since power is proportional to the square of pressure) (hypothesis 2b). In both cases, and in contradistinction to the predictions from classical loudness models that do not include an organizational stage in their computations, adjusted level L_I would be *less* than the physical level L_H and the more so as $L_H - L_L$ becomes smaller. Since the periodicity of the waveform may affect the operation of these latter two hypothetical mechanisms, we decided to use both pure-tone and narrow-band noise bursts in our experiments.

Note that these models both presume that the signal is organized into continuous and intermittent parts *prior* to loudness computation. Another approach would presume that the loudness is computed on the raw signal levels, and that auditory organization would occur *subsequent* to this computation, in which case some additional loudness computation process would perform the subtractions in sone units (Stevens, 1957) (hypothesis 2c). For pure tones and narrowband noise, identical results would be obtained from Zwicker's loudness model by subtracting specific loudness patterns and summing the specific loudnesses over critical bands in the residual pattern (Zwicker, 1960; Zwicker and Feldtkeller, 1967; Zwicker and Scharf, 1965). Pressure, power, and loudness calculations give different results for a given $L_H - L_L$ level difference.

All three of the aforementioned models that embody hypotheses 2a, 2b, and 2c presume that some kind of subtractive segregation process is employed, the operation of which is based on the available sensory data. Another possibility has been suggested by Repp (1992). He proposed that when an obliterated phoneme is replaced by a noise burst, its perceptual restoration is illusory since some higher-level, schema-driven, phonological completion is performed in a top-down fashion: the necessary information is selected from the noise burst (the top-down process is thus constrained by the sensory information), but this information is not *subtracted* per se from the noise burst, leaving the perceptual properties of the burst unaffected (hypothesis 3).

To summarize, the competing hypotheses to be tested are the following:

H1) classical loudness model without subtraction;
H2) old-plus-new type subtraction models, the computations of which are based on:
 a) acoustic pressure,
 b) acoustic power,
 c) loudness (in sones);
H3) top-down information-selection model.

Note that H1 and H3 make identical predictions for alternating-level sequences.

To test these various hypotheses, we presented listeners with sequences of events that alternated between a high level (L_H) and a low level (L_L), as in Fig. 1(a). The events were identical in spectral content. Listeners were asked to adjust the level of a comparison stimulus so that its loudness matched that of a specific part of the reference stimulus that varied with the experiment or within a block of trials. Stimulus parameters were varied to test the dependence of adjusted levels L_C and L_I on physical levels L_L and L_H under conditions in which listeners either clearly experienced auditory continuity or could not hear it. In the main experiment (experiment 1), conditions were presented in which continuity was heard. The adjusted levels were compared to those predicted by hypotheses H1, H2a–c, and H3. According to the old-plus-new heuristic, when continuity is not perceived, the classical model should be satisfactory and no subtraction of levels should be evident. Experiments 2 and 3 studied loudness matching to H and L bursts in alternating-level sequences in which continuity was impaired either by sending H and L bursts to separate ears or by introducing silences between them, respectively. Finally, a control experiment (experiment 4) was performed to verify listeners' loudness matching precision with fixed-level intermittent and continuous sequences.

To test for the possibility that temporal integration and/or loudness enhancement affect loudness computation in the continuity phenomenon, various duty cycles between H and L bursts were also employed in experiments 1 and 2. If temporal integration plays a role, one would expect differences in loudness matches between H bursts with durations of 200 ms and 100 ms, as well as between 100-ms L bursts and longer duration L bursts (Zwislocki, 1960). If loudness enhancement effects, originally investigated in two-burst stimuli, can be generalized to alternating sequences, one would expect greater loudness enhancement of L bursts (and thus less loudness difference between L and H bursts) in stimuli with shorter L-burst durations (100 and 200 ms; Zwislocki and Sokolich, 1974).

I. GENERAL METHOD

A. Stimuli

Sequences were composed of one of two types of stimulus bursts: a 1-kHz pure tone or a subcritical band, 140-Hz noise band centered on 1 kHz. Individual bursts had 5-ms linear onset and offset ramps. Sequences with alternating levels were composed of eight low-level (L) bursts interleaved with seven higher-level (H) bursts (Fig. 1). In experiments 1, 2, and 4, the sequences were presented with four different duty cycles to study effects of the time course of loudness growth and decay (D_H/D_L in ms: 200/200, 100/100, 100/300, 100/700). The total duration of the reference stimulus varied with duty cycle, being 3.0, 1.5, 3.1, and 6.3 s, respectively. In experiment 3, only the 100/300 duty cycle was used.

The H and L burst onsets and offsets either overlapped by 2.5 ms (no silence, experiments 1 and 2) or were separated by 30- or 100-ms silent intervals (experiment 3). With overlapping bursts, both pure-tone and noise-band signals were added in phase, i.e., they were derived from the same sound generator as in Fig. 2(b). In this latter condition, the continuity percept is quite strong if the sequence is presented diotically (experiment 1), but it is absent if H and L bursts are presented to separate ears (experiment 2). With silent intervals, continuity is generally absent or quite weak with 30-ms silences at the levels we used and is almost never perceived with 100-ms silences (experiment 3).

L_L in the reference stimulus was varied randomly within the A-weighted set $60 \pm \{1,3,5\}$ dB. In all analyses and plots, levels are presented relative to the mean of the roving range (60 dB). L_H was either 2, 6, or 10 dB greater than L_L. The prediction was that when auditory continuity was perceived, the loudness of the intermittent stream would vary systematically with this level difference, always being adjusted to a level below L_H (see Table I). For a mechanism operating on acoustic pressure, ideal listeners perceiving continuity should adjust an intermittent comparison stimulus to levels that are below L_H by 13.7, 6.0, and 3.3 dB, for L_H/L_L level differ-

TABLE I. Subtraction of low level from high level in units of pressure, power, and sones for a difference between H and L bursts of 6 dB ($L_L = 60$ dB SPL, $L_H = 66$ dB SPL, $Pr_{\mathrm{ref}} = 20\ \mu$Pa, $Po_{\mathrm{ref}} = 1$ pW).

	Pressure (μPa) ($Pr = 10$ dB/20$\cdot Pr_{\mathrm{ref}}$)	Power (pW) ($Po = 10$ dB/10$\cdot Po_{\mathrm{ref}}$)	Loudness (sones) ($S = 0.01 \cdot Pr^{0.6}$)
L_L	20 000	1 000 000	3.8073
L_H	40 000	3 981 072	5.7626
$L_H - L_L$	20 000	2 981 072	1.9553
L_I (dB)	60.0	64.7	50.4
$L_I - L_H$ (dB)	−6.0	−1.3	−15.6

ences of 2, 6, and 10 dB, respectively. For a mechanism operating on acoustic power, the adjusted levels should be below L_H by 4.3, 1.3, and 0.5 dB, respectively. And for a mechanism operating on sone units, the adjusted levels should be below L_H by 29.7, 15.6, and 10.1 dB, respectively. These latter values are obtained by converting L_H and L_L from dB to sones (S_H and S_L) by Stevens' law: $S = kp^{0.6}$, where S is the loudness in sones, $k \approx 0.01$, and p is the pressure in μPa (cf. Botte, 1989). Then $S_I = S_H - S_L$, and S_I is reconverted to L_I in dB by the same law in reverse. This relation is only valid for levels above 30 phones (30 dB for pure tones and narrow noise bands in the vicinity of 1 kHz) (cf. Scharf, 1978).

To the contrary, if the classical loudness model or the schema-driven model is appropriate, the level adjusted to match the intermittent part of the sequence should be close to L_H. The same result should obtain when continuity is not perceived (experiments 2 and 3). All models predict that the level adjusted to match the continuous stream when continuity is heard or to the lower-level part of the sequence when continuity is not heard should be near L_L. Departures from the physical levels in reference stimuli not producing auditory continuity may indicate biases induced by the stimulus context and/or by the matching strategy. Such biases would need to be taken into account when interpreting the results for stimuli producing continuity.

B. Procedure

Each experiment was preceded by a familiarization phase in which the stimuli were presented to the subjects who were questioned as to what they heard in order to verify whether or not auditory continuity was perceived for all stimulus conditions. They were also allowed to practice the adjustment procedure. One or two blocks containing all the stimuli for a given condition presented in random order were usually sufficient.

Each trial consisted of the repeated alternation between the reference stimulus and a comparison stimulus. During this alternation, the level of the comparison stimulus could be adjusted with a single-turn potentiometer. Subjects could listen to the alternation as many times as necessary to make a satisfactory loudness match, at which point they signaled the computer to record the level of the comparison stimulus by pressing a button. The listener aligned the turn-pot to a fixed reference point at the beginning of each trial. The starting levels of the reference stimulus were chosen at random from $L_L \pm \{7,8,9,10\}$ dB for that trial. The duration of the silent intervals separating the two stimuli varied with the experiment and will be specified for each one below. For reference stimuli producing continuity, subjects were asked on a given trial either to adjust the level of a continuous comparison stimulus to match the level of what appeared to be continuous in the reference, or to adjust the level of an intermittent sequence to match the level of what appeared to be intermittent in the reference. For sequences not producing continuity, intermittent reference stimuli of similar temporal structure were presented and the subjects were asked to match either the higher or the lower level in the reference stimulus using ear of presentation or duration cues to focus on the target stream (experiments 2 and 3, respectively).

Stimuli were presented in blocks comprising a given burst type (pure-tone or noise-band) and duty cycle. All combinations of burst type and duty cycle tested in a given experiment were completed before any one was repeated. They were block randomized for each subject and five blocks of each type were completed by each subject in each experiment. Different subjects were recruited for each experiment and for the different silent-duration and burst-type conditions in experiment 3. Each experiment was conducted in a series of sessions varying from 60 to 90 min. Subjects were allowed to take breaks between blocks as desired. In all experiments, the dependent variable was the matching ''error'' (in dB) between the adjusted level of the comparison stimulus and the level of the targeted part of the reference stimulus (H or L bursts).

C. Apparatus

Sinusoidal and white noise signals were synthesized at a sampling rate of 20 kHz with 16-bit resolution on an Oros DSP card controlled by a Compaq 386 computer. In the case of alternating-level signals, H and L bursts were processed in different channels. The signals were then filtered with a Kemo VBF/24 bandpass filter with cutoff frequencies of 930 and 1070 Hz and -48 dB/oct slopes. The filtering served both for anti-aliasing and for obtaining the narrow-band noise. The filtered signals were then routed through Charybdis D programmable attenuators with 0.25-dB resolution that were controlled by the computer. The final signals were either sent separately to the two earpieces of a TDH-49 headset for the dichotic conditions or were mixed and sent to both earpieces for the diotic conditions. Experimental sessions took place in an IAC single-walled sound isolation chamber. Subjects adjusted levels for comparison stimuli with a single-turn potentiometer and signaled their satisfaction with the match by pressing a button on the response box. At this point, the computer presented the next trial. Levels at each earpiece were verified using a Bruel & Kjaer 4153 artificial ear and a 2230 sound-pressure meter.

II. EXPERIMENT 1: DIOTIC ALTERNATING-LEVEL SEQUENCES PRODUCING CONTINUITY

The goal of this experiment was to test the main competing hypotheses that the loudness of an alternating sequence organized into continuous and intermittent streams is, on the one hand, partitioned into two quantities that may be computed either on the basis of pressure (H2a), power (H2b), or specific loudness (H2c) subtraction, or, on the other hand, is perceived as corresponding to the physical values presented (H1 and H3).

A. Method

On each trial the reference stimulus (sequence of overlapping H and L bursts) was alternated with an adjustable comparison stimulus that was either continuous or intermittent. The silence separating the end of the reference stimulus and the beginning of the comparison stimulus was 800 ms. A 1500-ms silence separated the end of the comparison stimulus and the next presentation of the reference stimulus. The four duty cycles were presented (200/200, 100/100, 100/300, 100/700). The continuous adjustable comparison stimulus had the same total duration as the reference stimulus and the intermittent adjustable comparison stimulus had the same temporal structure as the H bursts in the reference stimulus. Eight subjects that reported having normal hearing participated in the experiment and were paid for their services. Each subject completed five repetitions of the 48 conditions: 2 burst types $\times 4$ duty cycles $\times 3$ H/L level differences $\times 2$ comparison stimulus types.

B. Results

In the familiarization phase, all subjects reported the continuity percept for each condition, although the effect was weaker for the 2-dB difference in level between L_L and L_H, sometimes heard more as a fluctuating level. Subjects also reported that the bursts composing the intermittent stream in the alternating sequence were degraded in terms of the attack quality and tone color compared with the isolated intermittent sequence (similarly to results reported by Warren *et al.*, 1994).

From the adjusted level of the continuous comparison stimulus (L_C) and that of the intermittent comparison (L_I), the dependent variable (matching ''error'') was computed for each ($L_C - L_L$ and $L_I - L_H$, respectively). These values are plotted as a function of the H/L level difference in Fig. 3. For continuous stimuli, plotted data are averaged over burst type, duty cycle, subjects, and repetitions. For intermittent stimuli, data are averaged over subjects, repetitions, and duty cycle, with the exception of noise stimuli for which the stimuli with 200-ms H bursts are plotted separately from those with 100-ms H bursts (mean over 100/100, 100/300, and 100/700 duty cycles). Separate repeated-measures

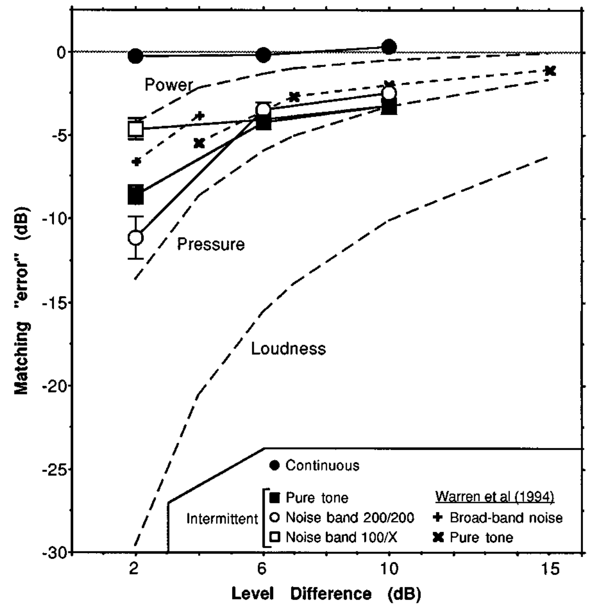

FIG. 3. Summary data for experiment 1. Mean loudness matching ''error'' (see text) as a function of H/L level difference. ''Errors'' for matches to the intermittent stream that are predicted by power, pressure, and loudness subtraction models are shown by dashed lines. All points would lie on the line at zero for models predicting matches to the physical values. All models predict zero error for matches to the continuous stream. For comparison, the data for similar (homophonic) conditions in Warren *et al.* (1994) are shown. (Pure-tone data were derived from Figs. 2 and 3 and broadband noise data from Fig. 7 in that study.) Vertical bars (where visible) show ± 1 standard error.

ANOVAs were performed for continuous and intermittent comparison stimuli with repeated factors burst type (2), H/L duty cycle (4), and H/L level difference (3).

1. Continuous comparison stimuli

There was a nearly perfect match of the comparison stimulus to the level of the continuous stream, matching error being within 1 dB of L_L. There was a significant difference of 0.6 dB between adjusted levels for pure-tone and noise-band stimuli [$F(1,7)=10.67$, $p<0.05$],[3] the stimuli being adjusted 0.3 dB below L_L for pure tones and 0.3 dB above L_L for noise bands. The two means are also both reliably different from the hypothesized value of 0 dB [single sample $t(479)=-2.92$, $p<0.005$; $t(479)=2.98$, $p<0.005$, respectively]. This difference, while significant, is quite small (less than the differential threshold for intensity discrimination, e.g., Luce and Green, 1974). There were no effects of duty cycle or of H/L level difference.

2. Intermittent comparison stimuli

The intermittent stream was adjusted on average to a level less than L_H. The range of the mean matching errors across subjects and repetitions was from -2.4 to -11.2 dB. Classical loudness models would predict mean adjusted intermittent levels near L_H, but these were very rare across subjects (17 out of 192 matching errors averaged across repetitions were above -1.0 dB). The effect of H/L level difference depended on both burst type and duty cycle, as witnessed by the significant triple interaction [$F(6,42)=5.42$,

$p < 0.05$]. For pure-tone stimuli, a separate repeated-measures ANOVA on factors duty cycle and H/L level difference revealed that there was no effect of duty cycle [$F(3,21) = 1.09$, n.s.], while a significant main effect of H/L level difference [$F(2,14) = 6.16$, $p < 0.05$] was found: matching "errors" decreased with increasing level difference. A similar analysis for noise-band stimuli showed that the duty cycle by H/L level difference interaction was significant [$F(6,42) = 5.14$, $p < 0.05$]. This interaction results from the fact that for the 200/200 duty cycle and a 2-dB H/L difference, the comparison stimulus was adjusted to much lower values compared both to those at larger H/L differences for the same duty cycle, as well as to the other duty cycles at the 2-dB difference (all of the cited differences between means were greater than the critical Tukey–Kramer difference of 3.4 dB). There was no effect of H/L level difference for the three duty cycles with 100-ms H bursts [$F(1,42) = 0.65$, n.s., for 2 dB versus 6 dB; $F(1,42) = 1.05$, n.s., for 6 dB versus 10 dB; $F(1,42) = 3.35$, $p < 0.10$, for 2 dB versus 10 dB]. There was no difference between pure-tone and noise-band stimuli at the 6-dB and 10-dB H/L level differences. The source of this triple interaction thus seems to be related to matches for noise-band stimuli at the 2-dB H/L level difference: the adjusted level of the 200/200 duty cycle was low (a value commensurate with those obtained for pure-tone stimuli at all duty cycles) and the adjusted levels of the duty cycles with 100-ms H bursts were high.

C. Discussion

This pattern of data corresponds to predictions of neither the classical loudness models (H1) nor the schema-driven models (H3), i.e., listeners do not systematically adjust levels of the intermittent sequence close to the physically presented levels. We can thus reject both of these classes of models. The data are, however, in qualitative agreement with the prediction of the "old-plus-new" subtraction strategy: the smaller the H/L level difference, the greater the matching error. In fact, L_I was, on average, adjusted even lower than L_I for a difference of 2 dB. This effect is broadly consistent with results obtained by Darwin (1995) for synthetic vowel stimuli and by Warren *et al.* (1994) for pure tones and broadband noise. The matching errors are much smaller in magnitude than those predicted by the loudness subtraction model (see Fig. 3). So whatever the auditory representation of level used for subtracting the continuous portion from the high-level burst, it is clearly not related to loudness as defined in Stevens' power law, suggesting that the signal is organized into streams before loudness is computed rather than afterward. For these homophonic stimuli, the auditory representation of level at the stage of stream organization appears to be closer to physical units like pressure or power. Our data fall between the predictions based on pressure and power subtraction, as do those of Warren *et al.* (1994). Single-sample t-tests, adjusted for multiple tests, were performed on mean matching errors against values predicted by power and pressure subtraction. They revealed that neither subtraction method predicts all of the experimental data. At the 2-dB level difference, matches for noise-band stimuli with the 200/200 duty cycle are not different from the pressure calculation, while those with 100-ms H bursts are not different from the power calculation. At the 10-dB difference, none of the matches is different from the pressure calculation. In all other cases, the mean matching errors are significantly different from both pressure and power calculations. Clearly, neither of these units explains the data and there are some troubling differences due to duty cycle and burst type, particularly at the smallest H/L level difference.

There may be a number of reasons for these discrepancies. Subjects noted that with the 2-dB level difference, the task was more difficult and they were more uncertain in their matches. They also felt the difficulty and uncertainty were increased with the noise stimuli compared to the pure tone stimuli. The former impression is borne out for 2-dB conditions, which have much higher standard deviations (7–8 dB) than do conditions with a larger level difference (2.5–3.0 dB). However, this pattern is very similar for both pure-tone and noise-band stimuli, and indeed for noise stimuli the 200/200 duty cycle has a larger standard deviation than that for the other three duty cycles taken together. This result belies the latter impression and argues against listeners' having difficulty estimating the level in these stimuli due to short-term level fluctuations in narrow-band noise signals. Although our results globally support a subtractive segregation mechanism, they may also suggest that subjects had difficulty in segregating the intermittent stream composed of 100-ms bursts for the 2-dB difference for noise-band stimuli: the stochastic nature of the signal may have hindered the segregation process to some extent making the percept itself somewhat fuzzy. While this interpretation is coherent with the introspective reports of subjects during the familiarization phase described above, seven of the eight subjects adjusted comparison levels *below* the physical level in conformity with predictions of a subtraction model which presumes a segregation of the sequence into two streams.

Another potential problem is that the alternating-level stimulus context may have induced biases in the level matches. It was therefore necessary to verify that listeners adjust comparison stimuli to levels that are close to the physically presented levels in alternating-level stimuli that do not produce continuity. Further, it was also necessary to estimate the precision with which listeners adjust continuous and intermittent sequences to such alternating sequences and to rule out the possibility that alternating-level sequences affect the matching of loudness more generally.

III. EXPERIMENT 2: DICHOTIC ALTERNATING-LEVEL SEQUENCES

This experiment had two goals: verify the prediction that adjusted levels of comparison sequences are close to physically presented levels when continuity is not heard and test the precision of such matches. A reference stimulus was used that had the same temporal configuration as that of the reference stimulus producing continuity in experiment 1, but which did not itself produce the continuity percept. The breakdown of continuity is obtained by sending the H and L bursts to separate ears.

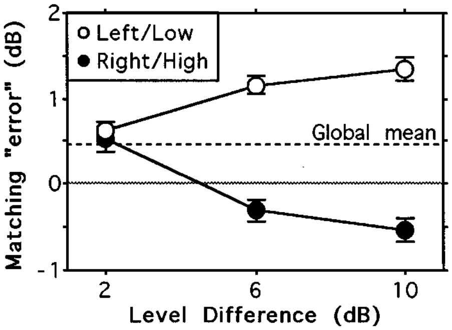

FIG. 4. Summary data for experiment 2. Mean loudness matching ''error'' as a function of H/L level difference for matches to higher-level intermittent sequences in the right ear (Right/High) and lower-level intermittent sequences in the left ear (Left/Low). Vertical bars show ± 1 standard error.

A. Method

H and L bursts were routed to the right and left earphones, respectively. These ear-specific target sequences will be denoted R/H (Right/High) and Le/L (Left/Low). The adjustable comparison stimulus was presented to the target ear and consisted of a series of bursts identical in duration to those of the sequence presented to the same ear in the reference stimulus. At the beginning of the experiment all three level differences were presented to subjects who were asked what they heard. No subject reported a sensation of continuity. Eight subjects reporting no hearing problems were recruited and paid for their participation. Each subject completed five repetitions of the 48 conditions: 2 burst types $\times 4$ duty cycles $\times 3$ H/L level differences $\times 2$ comparison stimulus types (R/H) vs (Le/L).

B. Results

The dependent variable was the ''error'' in adjusted level of the comparison stimulus relative to the physical level presented in the reference stimulus. Mean matching errors are presented in Fig. 4 for each ear as a function of H/L level difference and in Fig. 5 for each ear as a function of duty cycle. The mean errors across burst types, subjects, and repetitions varied from -1.3 to $+2.6$ dB. A repeated-measures ANOVA was performed on factors comparison stimulus type (2), burst type (2), duty cycle (4), level difference (3), and repetitions (5). No difference was found between noise and tone stimuli $[F(1,7)<1]$. The effect of comparison stimulus type was significant $[F(1,7)=14.4,\ p<0.005]$, the global error being -0.1 dB for the R/H sequences and 1 dB for the Le/L sequences. There is a global overestimation of 0.5 dB that was not present in the first experiment for the continuous stream. It may therefore be related to adjusting a simple intermittent sequence to part of an alternating sequence, although Marks (1978) noted a greater sensitivity of the right ear on the order of 1 dB in subjects performing loudness magnitude estimations on binaural stimuli.

A significant interaction between comparison stimulus type and level difference results from a divergence between matching errors for R/H and Le/L conditions $[F(2,14)=31.8,\ p<0.0001]$ (see Fig. 4): listeners increasingly overestimated the level of the Le/L sequence as the level difference increased and they moved from a slight overestimation of R/H sequences at a 2-dB difference to an equivalent underestimation at a 10-dB difference. This divergence is nearly symmetrical about the global average matching error and suggests a bias in matches to the target sequence in the direction of the level of the nontarget sequence.

A significant interaction was also found between comparison stimulus type and duty cycle $[F(3,21)=23.2,\ p<0.0001]$ (see Fig. 5): matching errors were significantly different from the global average in a positive direction for Le/L conditions and in a negative direction for R/H conditions, but only for the 100/100 and 200/200 duty cycles. So the matching bias in the direction of the nontarget sequence appears to disappear when the duration of the L bursts is at least 300 ms. It seems unlikely that this effect is due to temporal integration of loudness over the first 200 ms of the tone (Zwicker and Fastl, 1990), since the effect is symmetric for duty cycles of similar duration in the two ears.

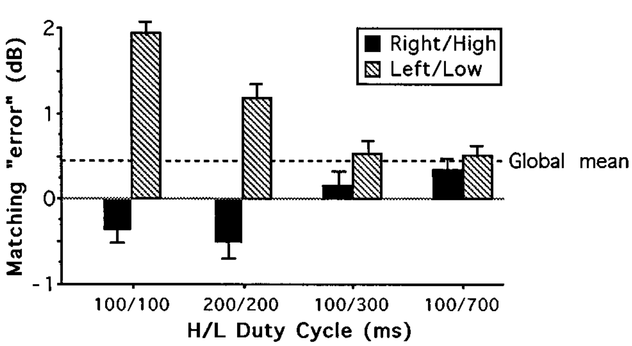

FIG. 5. Summary data for experiment 2. Mean loudness matching ''error'' as a function of duty cycle for matches of Right/High and Left/Low sequences. Vertical bars show $+1$ standard error.

C. Discussion

The first thing to note about these results is that they are strikingly different from those of experiment 1. On average, matching errors are within 2 dB of the physically presented values at all duty cycles for all level differences. Indeed in this experiment departures from perfect matches are greater for greater H/L level difference, whereas in experiment 1 errors were greater for smaller H/L level difference. It would appear that these results, obtained with stimuli not producing the continuity phenomenon, can be roughly predicted by the classical loudness models.

These results may also reflect biases in matches to intermittent sequences embedded in alternating-level sequences. They suggest that an isolated intermittent sequence is heard globally with a level of about 0.5 dB less than the same sequence embedded in an alternating context. They also suggest that increasing the level difference between the embedded target sequence and the nontarget sequence results in an increasing bias in the direction of the nontarget sequence, amounting to about 1 dB for a 10-dB H/L difference. This effect is more pronounced for stimuli with shorter-duration L bursts. The values found in this experiment, however, are neither big enough nor consistently in a given direction to explain the departure of data for matches to intermittent targets in experiment 1 from either pressure or power subtraction predictions. Further, the large difference between noise-band stimuli with 100-ms H bursts and those with 200-ms H bursts that was found in experiment 1 is not found in this experiment. It may be that the dichotic presentation made this set of stimuli too different from those in experiment 1 for the biases revealed to be directly comparable to the former conditions.

IV. EXPERIMENT 3: DIOTIC ALTERNATING-LEVEL SEQUENCES WITH INTER-BURST SILENCES

The aims of this experiment were identical to those of experiment 2. However, in this experiment we used diotically presented stimuli similar in structure to those of the main experiment but in which continuity was not heard. Continuity was broken by introducing brief silences between H and L bursts. The prediction was that matches would be close to physically presented levels.

A. Method

Only the 100/300 duty cycle was used. Silences of 30 or 100 ms were introduced to separate the low- and high-level tone bursts. The 30-ms silences were used only with pure-tone stimuli, whereas both pure-tone and noise-band stimuli were tested with 100-ms silences. To focus subjects' attention on the perception of intermittent streams in the reference stimulus, the adjustable comparison stimuli were always intermittent and corresponded identically in temporal structure (300-ms bursts for lower-level and 100-ms bursts for higher-level streams) to the targeted high- or low-level part of the reference stimulus. These comparison stimuli will be denoted Lo/L (long/low) and S/H (short/high). Subjects' verbal reports indicated that 30-ms silences could at times give a weak impression of continuity, but they could also learn not to hear the percept. Van Noorden (1975) had found that with 40-ms pure-tone bursts and silences of 22 ms, the H/L level difference necessary to obtain the continuity effect was over 20 dB, which is well above the maximum level difference employed in this study. For the stimuli with 30-ms silences, some subjects found it difficult to focus on one level at the beginning. To the contrary, the 100-ms silences never gave the continuity percept and presented fewer problems of attentional focus. Three independent groups of subjects were paid for their participation in the experiment. Seven heard pure-tone stimuli with 30-ms silences, eight heard pure-tone stimuli with 100-ms silences, and eight heard noise-band stimuli with 100-ms silences. All reported having normal hearing. Each subject completed five repetitions of the six conditions: 3 H/L level differences $\times 2$ comparison stimulus types (S/H vs Lo/L).

B. Results

The dependent variable was the ''error'' in final adjusted level of the comparison stimulus relative to the physical level presented. Mean matching errors are presented in Fig. 6 for each comparison stimulus type as a function of H/L level difference. The mean matching errors across subjects and repetitions varied from -1.9 to $+1.6$ dB. A mixed ANOVA was performed with independent groups on the combinations of silent-duration and burst-type factors and with repeated measures on factors comparison stimulus type (2), level difference (3), and repetitions (5). There was no difference between the two durations of silence separating H and L bursts [$F(1,20)<1$]. The variability in matches was slightly lower for the 100-ms silences, perhaps due to the better impairment of continuity than was obtained with 30-ms silences. There was a significant interaction of comparison stimulus type, level difference, and burst type [$F(2,40)=16.2$, $p<0.0001$]. For S/H sequences, matching errors were increasingly negative with increased H/L level difference, while the reverse was the case for matches to the Lo/L sequences, mirroring the results from experiment 2. The triple interaction results from the fact that this effect is slightly less marked for pure-tone than for noise-band stimuli (see Fig. 6).

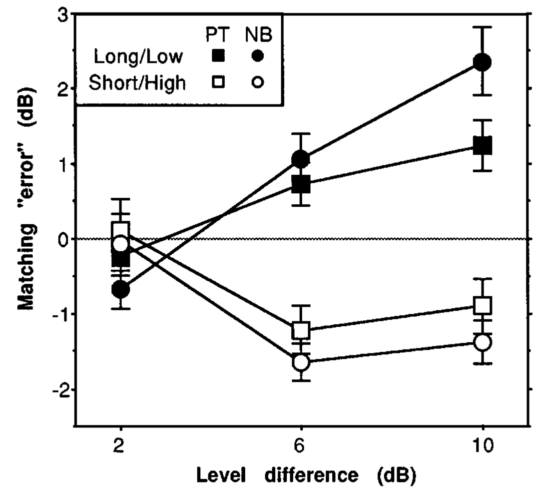

FIG. 6. Summary data for experiment 3. Mean loudness matching ''error'' as a function of H/L level difference for matches to intermittent sequences of higher-level, short-duration bursts (Short/High) or lower-level, long-duration bursts (Long/Low) for both pure-tone (PT: $N=75$) and narrow-band noise stimuli (NB: $N=40$). Vertical bars show ± 1 standard error.

C. Discussion

As for the results of experiment 2, these results are globally consistent with the predictions of classical loudness models. Further, these results also reveal a dependence of adjusted levels on level difference in the alternating sequence. It would seem, therefore, that the context effect of a sequence with alternating levels induces overestimation of lower-level sounds and underestimation of higher-level sounds. This effect is much larger than the small bias found in nonalternating sequences (see experiment 4 below). It is roughly equivalent to that found for dichotic alternating sequences (experiment 2), although the global positive bias in matches present in the latter experiment was not present in the current one, suggesting the previous result may have its origins in the dichotic presentation. Similarly to experiment 2, the matching biases found here are insufficient to explain the departures from pressure or power predictions for intermittent streams in experiment 1, particularly concerning the large difference between pure-tone and noise-band stimuli found in the latter experiment.

V. EXPERIMENT 4: DIOTIC FIXED-LEVEL STIMULI

The goal of this control experiment was to verify whether the intermittence of the reference or comparison stimuli systematically influenced matching errors for either continuous or intermittent sequences.

A. Method

Two types of diotically presented stimuli were used that resemble those in Fig. 1(b): a continuous (CONT) sound and an intermittent (INT) sequence composed of seven sounds having the duration of *H* bursts separated by silences having the duration of *L* bursts. Comparison stimuli were adjusted to match the level of reference stimuli of the same type (INT or CONT).[4] For INT conditions, both pure-tone and noise-band stimuli with four duty cycles were employed. For CONT conditions, four sequence durations were used that corresponded to the total sequence duration for each duty cycle. Eight subjects, all of whom reported having normal hearing, were paid to participate in the study. Each subject completed five repetitions of the 16 conditions: 2 burst types $\times 4$ duty cycles $\times 2$ stimulus types.

B. Results and discussion

The matching precision was good with mean matching errors across burst types, subjects, and repetitions varying from -0.4 to 0.5 dB for INT conditions and from 0.0 to 0.3 dB for CONT conditions. Separate repeated-measures ANOVAs were performed for INT and CONT stimuli on factors burst type (2), duty cycle (4), and repetitions (5). Note that duty cycle corresponds simply to a difference in stimulus duration for CONT stimuli. The only significant effect in both analyses was for the duty cycle factor in INT stimuli [$F(3,21)=4.1$, $p<0.05$]. Matching errors for all duty cycles were positive except for 100/700. However, only one of these mean matching errors was significantly different from zero by single-sample t-tests adjusted for multiple tests. This condition was INT 100/100 for which subjects overestimated the level by about 0.5 dB [$t(79)=3.61$, corrected $p<0.005$]. In general therefore, listeners are quite precise at adjusting both continuous and intermittent sequences, mean matching errors being well within the differential threshold for intensity (Luce and Green, 1974). The differences between adjusted levels and power and pressure predictions in experiment 1 would not appear to be attributable to imprecise level matching between intermittent sequences.

VI. GENERAL DISCUSSION

The data presented above demonstrate that the perceived loudness of an auditory event depends on the way the event sequence is organized perceptually. In the case of an alternating-level sequence perceived as a sequence of intermittent events imposed on a continuous sound, the level of the continuous sound is heard as being equal to the physical level of the lower-level bursts. Level matches of a continuous comparison sound to the continuous stream of the reference stimulus are within 0.4 dB of the physical level and are unaffected by the level difference in the alternating-level context (experiment 1). Indeed, a comparable degree of precision is found for matches to a fixed-level continuous sound (experiment 4). However, matches to the intermittent part of the percept are far below the physical levels presented and are clearly consistent with some kind of subtractive process. The pattern of the data in the present study confirms and extends that of Warren *et al.* (1994) and suggests that the higher-level part of the signal is processed by the auditory system as if it were composed of two parts, each with its own share of the neural input corresponding to the incident energy of the stimulus sequence.[5] Loudness matches fall between values predicted by subtraction based on acoustic pressure and acoustic power. They are very different from computations based on loudness as represented in sone units. The results are thus consistent with the hypothesis that auditory organization takes place prior to loudness computation and that the auditory sensory representation of level at this stage of processing is close to the physical stimulus. For small differences between higher and lower levels, loudness matches also depend both on the relative durations of the *H* and *L* bursts and on the spectral content of the signal. The means for most conditions are nearer the predictions based on pressure with the exception of small H/L level differences for noise-band stimuli with short-duration *H* bursts.

When the continuity of such an alternating-level sequence is broken by routing alternate events to separate ears (experiment 2) or by introducing a silence between successive events (experiment 3), mean matches are much closer to the physically presented values and are even at times greater than these values, a situation *never* found when continuity was heard. With the stimuli of experiments 2 and 3, a dependence of the matches to the lower-level bursts on the level difference is found: as the level difference increases, the level of these bursts is progressively overestimated, i.e., in the direction of the higher-level bursts. This overestimation attains about 1–2 dB on average for a level difference of 10 dB. Clearly the matching of loudness in the two cases is influenced by the perceptual context, and these kinds of matching biases are similar to what Poulton (1989) has characterized as ''centering tendencies'' in psychophysical judgment strategies. However, the pattern of results is completely different from that found when the sequence is organized into two streams. In the former case, matches are consistent with classical loudness models (plus judgment biases) whereas in the latter case they are consistent with a subtraction model.

These results are not consistent with effects related to temporal integration, loudness enhancement, loudness adaptation or the ''recalibration'' of sensory input. We can rule out loudness summation as the origin of these effects (Zwicker *et al.*, 1957) since stimuli were confined to a single critical band. We can also rule out loudness adaptation (Botte *et al.*, 1982) as the stimulus sequences were not long enough in duration. Further, it is hard to imagine how loudness enhancement (Zwislocki and Sokolich, 1974) might be playing a role in these stimuli even it if can be considered to generalize to alternating levels: there is clearly no enhancement of the lower-level bursts, the level of the continuous stream being adjusted very near the physical level. However, one might imagine that the loudness of the low-level intermittent stream could be enhanced by a louder continuous stream. This relation is predicted by pressure and power subtraction for the 2-dB H/L level difference, but adjusted levels appear to be very sensitive to both the waveform and the duration of intermittent bursts. It is unclear how loudness enhancement could account for these latter results. Finally, the ''slippery context effect'' or ''recalibration'' of loudness described by Marks and Warner (1991) and Marks (1994), respectively, do not explain these data either. In those studies, the loudness of a soft pure tone could be increased when presented within the context of trials containing relatively higher-level tones at another frequency. First, this effect completely disappears when the tones are of the same frequency (Marks, 1994), and second, the loudness of the higher-level tone is unaffected by the context, while that of the lower-level tone is increased. Their results are thus quite the opposite of ours and most likely reflect a completely different level of the loudness computation mechanism.

A comparison of these data with those obtained in similar sequences that do not produce auditory continuity (experiments 2 and 3) demonstrates that while the alternating context induces matching biases (the bigger the level difference, the greater the compensation of the match in the direction of the nontargeted part of the sequence), this kind of compensation is not at all found for the lower-level continuous stimuli in experiment 1 and the matching ''errors'' for intermittent stimuli are much larger. Further, the pattern of bias is very different in experiments 2 and 3 compared with the deviations from the physical levels in experiment 1. The largest deviations are found for small alternating level differences in experiment 1 and for large differences in experiments 2 and 3. Also, the biases exist for both L and H bursts (in opposite directions) in experiments 2 and 3, while no appreciable deviation is found for the lower-level continuous stimuli in experiment 1. Therefore, we may conclude that the divergence from physical levels in experiment 1 cannot be explained simply by biases induced by the alternating-level context. The results are consistent with a subtractive process based on the perceptual interpretation that H bursts are composed of the continuation of L bursts and an additional superimposed burst.

There remains the problem of determining the nature of the subtraction process which seems to correspond to neither pressure nor power computations. It is possible that the nature of the task may be responsible for these deviations in experiment 1. On a given trial, the listener is asked to focus either on the continuous sound or on the intermittent sequence and to estimate the loudness in order to match it with a comparison stimulus. This focusing may create an imprecise partitioning of the stimulus energy which is mixed together in the same auditory channels since the two streams are spectrally identical. If a pressure subtraction law is used, one could imagine that the focusing process results in too much energy being assigned to the attended stream, which would lead to an overestimation of the levels. The data suggest that if this is the case, a greater difference between the alternating levels results in smaller errors in estimation. They further suggest that for short-duration stimuli with uncertain levels (as is the case with the 100-ms H bursts of narrow-band noise), the overestimation is exaggerated for very small level differences. The question of a possible influence of temporal structure (as well as of experimental instructions) on perceived continuity will be addressed in a subsequent paper (Drake and McAdams, submitted).

If, on the other hand, the auditory system uses power computation, the focusing process would have to result in too little energy being assigned to the target stream. This error would be greatest at small level differences for pure tones and noise bands with longer-duration H bursts, but would be smallest at the same level difference for short-duration H bursts. For larger level differences, the error would be constant for all stimuli tested.

It is important to recall that in the everyday world, pressure calculations would only be appropriate for signals added in coherent phase (an unlikely occurrence) while power calculations would be appropriate for independent signals, the phases of which are unknown. While the stimuli in our experiments were added in coherent phase, Warren *et al.* (1994) demonstrated that randomizing the phase relations between successive bursts had no appreciable effect on the data.

Another possibility is that the auditory system is no more precise than it has to be to get the job done correctly.[6] There may be an inherent ambiguity in the partitioning process, since it is rarely possible on the basis of locally available sensory information to determine the exact properties of the signals that compose a mixture. In more realistic situations than the repetitive alternation between two levels used here, there would normally be separate instants where the ''old'' signal is present alone, the ''new'' signal is present alone, and both signals are mixed. If a significant proportion of the ''old'' signal can be removed at an instant of mixture, the resulting rough estimate of the intensity of the ''new'' one could be corrected by estimates obtained either at earlier or at later instants when it was present by itself. The conjunction of a number of estimates at different instants could lead to a reasonably, although not perfectly, accurate conclusion concerning this property of the target source. This was clearly the case in our study as evidenced by the precise matches to the continuous stream in experiment 1.

VII. CONCLUSIONS

(1) A process akin to the ''old-plus-new'' strategy (Bregman, 1990) seems to operate on alternating-level homophonic sequences. This process subtracts from an interrupting high-level sound the part that is perceptually assigned to another sound perceived to continue through it. The level of the residue thus depends on the level difference between the higher-level sound and the lower-level sound. This result is in contradistinction to predictions by classical loudness models (e.g., Zwicker, 1977) that do not take auditory organization processes into account, or by models that postulate a selection of information on the basis of some kind of mental schema (Repp, 1992) in which the restored

part of the continuous sound is not subtracted from the mixture. However, the law by which the loudness *is* partitioned does not correspond across all conditions to either pressure or power subtraction and is very different from predictions based on loudness subtraction in sone units. The same law would appear to operate on pure tone and narrow-band noise stimuli for larger differences between lower- and higher-level bursts. Observed effects due to the duty cycle of the alternation for noise-band stimuli do not seem to be attributable to the uncertainty in estimation of loudness of short-duration narrow-band noise bursts nor to effects of loudness enhancement.

(2) Loudness matches to fixed-level continuous and intermittent stimuli are very accurate when reference and comparison stimuli have the same temporal structure.

(3) For alternating sequences in which continuity is not heard (dichotically presented sequences or sequences with intervening silences), the level of the softer, intermittent stream is overestimated by an amount that increases with the level difference between higher and lower levels. The reverse was true of the louder stream, suggesting that matching biases are made in the direction of the level of the nontarget stream. Aside from these small matching biases, matches are consistent with predictions of classical loudness models, suggesting that such models are most appropriate for situations in which sound sequences are perceptually organized into a single stream.

(4) The loudness matching biases found in conditions where continuity was not perceived do not explain the deviations measured in the continuity conditions. It is likely that a process akin to the old-plus-new subtraction strategy is employed, though the task used, which requires the listener to focus on a targeted part of the stimulus sequence, may affect the perceived loudness. On the other hand, the auditory system may not need to be more precise in its loudness computations to deal appropriately with real-world situations.

ACKNOWLEDGMENTS

This research was funded in part by a grant from the French Ministry of the Environment to Marie-Claire Botte shortly before her untimely death. François Banide, Xavier Durot, and Sophie Savel helped run subjects and Anne-Marie Argenti helped with DSP programming. Neil Todd suggested the duty cycle conditions. Albert Bregman, Bertram Scharf, Richard Warren, and one anonymous reviewer made thought provoking comments on an earlier version of the manuscript.

[1]The classical approach to loudness has of course investigated the perception of loudness of individual components of complex signals, such as pure tones in noise (Zwicker and Fastl, 1990), but the models currently published in the literature do not segregate the signal into pure-tone and noise components and then perform the loudness calculation on the separate components. The experimenter has to know where to focus on the model output to perform the appropriate calculations. Further, these results do not generalize to temporally varying signals of constant spectral content, which is the case being considered here.

[2]Other linear effects in loudness have also been investigated. Most notably, the difference in loudness between monaural and binaural presentation can for the most part be explained by a linear summation of the individual loudnesses in sone units of the signals at the two ears (Marks, 1979; Scharf, 1969; Scharf and Fishken, 1970).

[3]Degrees of freedom for the F-statistic are adjusted where necessary with the Geisser–Greenhouse epsilon to compensate for inherent correlation among repeated factors.

[4]All four combinations of intermittent and continuous reference and comparison stimuli were presented to subjects. Only the same-type pairs are reported here as they are the only ones cogent to this study.

[5]It should be noted that subjects' reports of qualitative differences between the intermittent comparison stimulus and the intermittent stream in the reference sequence, in both the present study and in Warren *et al.*'s study, cannot be accounted for by a simple subtraction of stimulus level and may reflect a more complex partitioning of the neural response as suggested by Warren *et al.* (1994).

[6]We are grateful to Albert Bregman for pointing out this ecologically pragmatic possibility to us.

Botte, M.-C., Canévet, G., and Scharf, B. (**1982**). ''Loudness adaptation induced by an intermittent tone,'' J. Acoust. Soc. Am. **72**, 727–739.

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

Bregman, A. S., and Rousseau, L. (**1991**). ''Auditory intensity changes can cue perception of transformation, accompaniment or replacement.,'' Bull. Psychonom. Soc. **29**, 476 (A).

Ciocca, V., and Bregman, A. S. (**1987**). ''Perceived continuity of gliding and steady-state tones through interrupting noise,'' Percept. Psychophys. **42**, 476–484.

Darwin, C. J. (**1995**). ''Perceiving vowels in the presence of another sound: a quantitative test of the ''old-plus-new'' heuristic,'' in *Levels in Speech Communication: Relations and Interactions*, edited by C. Sorin, J. Mariani, H. Meloni, and J. Schoentgen (Elsevier, Amsterdam), pp. 1–12.

Drake, C., and McAdams, S. (**submitted**). ''The continuity illusion: Role of temporal sequence structure.''

Houtgast, T. (**1972**). ''Psychophysical evidence of lateral inhibition in hearing,'' J. Acoust. Soc. Am. **51**, 1885–1894.

Luce, R. D., and Green, D. M. (**1974**). ''Neural coding and psychophysical discrimination data,'' J. Acoust. Soc. Am. **56**, 1554–1564.

Marks, L. E. (**1978**). ''Binaural summation of the loudness of pure tones,'' J. Acoust. Soc. Am. **64**, 107–113.

Marks, L. E. (**1979**). ''A theory of loudness and loudness judgments,'' Psychol. Rev. **86**, 256–285.

Marks, L. E. (**1994**). ''''Recalibrating'' the auditory system: The perception of loudness,'' J. Exp. Psychol. **20**, 382–396.

Marks, L. E., and Warner, E. (**1991**). ''Slippery context effect and critical bands,'' J. Exp. Psychol. **17**, 986–996.

McAdams, S., Botte, M.-C., and Drake, C. (**1994a**), ''Le phénomène de la continuité auditive et la répartition de la sonie entre flux auditifs,'' J. Physique **4**(C5), 383–386.

McAdams, S., Botte, M.-C., Banide, F., Durot, X., and Drake, C. (**1994b**). ''The computation of loudness in the auditory continuity phenomenon,'' ATR Workshop on ''A Biological Framework for Speech Perception and Production,'' Kyoto, ATR Technical Report TR-H-121, pp. 81–86.

Poulton, E. C. (**1989**). *Bias in Quantifying Judgments* (Erlbaum, Hillsdale, NJ).

Repp, B. H. (**1992**). ''Perceptual restoration of a missing speech sound: Auditory induction or illusion?,'' Percept. Psychophys. **51**, 14–32.

Scharf, B. (**1969**). ''Dichotic summation of loudness,'' J. Acoust. Soc. Am. **45**, 1193–1205.

Scharf, B., and Fishken, D. (**1970**). ''Binaural summation of loudness reconsidered,'' J. Exp. Psychol. **86**, 374–379.

Scharf, B. (**1978**). ''Loudness,'' in *Handbook of Perception*, Vol. 4: Hearing, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 187–242.

Stevens, S. S. (**1957**). ''Concerning the form of the loudness function,'' J. Acoust. Soc. Am. **29**, 603–606.

Thurlow, W. R. (**1957**). ''An auditory figure-ground effect,'' Am. J. Psychol. **70**, 653–654.

Thurlow, W. R., and Elfner, L. F. (**1959**). ''Continuity effects with alternately sounding tones,'' J. Acoust. Soc. Am. **31**, 1337–1339.

van Noorden, L. P. A. S. (**1975**). ''Temporal coherence in the perception of tone sequences,'' unpublished doctoral thesis, Eindhoven University of Technology.

van Noorden, L. P. A. S. (**1977**). ''Minimum differences of level and frequency for perceptual fission of tone sequences ABAB,'' J. Acoust. Soc. Am. **61**, 1041–1045.

Verschuure, J., Rodenburg, M., and Maas, A. J. J. (**1976**). ''Presentation conditions of the pulsation threshold method,'' Acustica **35**, 47–54.

Warren, R. M. (**1982**). *Auditory Perception: A New Synthesis* (Pergamon, New York).

Warren, R. M. (**1984**). ''Perceptual restoration of obliterated sounds,'' Psychol. Bull. **96**, 371–383.

Warren, R. M., Bashford, J. A., Healy, E. W., and Brubaker, B. S. (**1994**). ''Auditory induction: Reciprocal changes in alternating sounds,'' Percept. Psychophys. **55**, 313–322.

Warren, R. M., Obusek, C. J., and Ackroff, J. M. (**1972**). ''Auditory induction: Perceptual synthesis of absent sounds,'' Science **176**, 1149–1151.

Zwicker, E. (**1960**). ''Ein Verfahren zu Berechnung der Lautstärke,'' Acustica **10**, 304–308.

Zwicker, E. (**1977**). ''Procedure for calculating loudness of temporally variable sounds,'' J. Acoust. Soc. Am. **62**, 675–682.

Zwicker, E., and Fastl, H. (**1990**). *Psychoacoustics: Facts and Models* (Springer-Verlag, Berlin).

Zwicker, E., and Feldtkeller, R. (**1967**). *Das Ohr als Nachrichtenempfänger* (S. Hirzel Verlag, Stuttgart).

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

Zwicker, E., and Scharf, B. (**1965**). ''A theory of loudness summation,'' Psychol. Rev. **72**, 3–26.

Zwislocki, J. (**1960**). ''Theory of temporal auditory summation,'' J. Acoust. Soc. Am. **32**, 1046–1060.

Zwislocki, J., and Sokolich, W. G. (**1974**). ''On loudness enhancement of a tone burst by a preceding tone burst,'' Percept. Psychophys. **16**, 87–90.