# Memory for pitch versus memory for loudness

Sylvain Clément, Laurent Demany, and Catherine Semal
*Laboratoire de Neurophysiologie (UMR CNRS 5543), BP 63, Université Bordeaux 2, 146 rue Léo-Saignat, F-33076 Bordeaux Cedex, France*

(Received 14 January 1999; revised 3 May 1999; accepted 28 June 1999)

The decays of pitch traces and loudness traces in short-term auditory memory were compared in forced-choice discrimination experiments. The two stimuli presented on each trial were separated by a variable delay ($D$); they consisted of pure tones, series of resolved harmonics, or series of unresolved harmonics mixed with lowpass noise. A roving procedure was employed in order to minimize the influence of context coding. During an initial phase of each experiment, frequency and intensity discrimination thresholds $[P(C)=0.80]$ were measured with an adaptive staircase method while $D$ was fixed at 0.5 s. The corresponding physical differences (in cents or dB) were then constantly presented at four values of $D$: 0.5, 2, 5, and 10 s. In the case of intensity discrimination, performance ($d'$) markedly decreased when $D$ increased from 0.5 to 2 s, but was not further reduced when $D$ was longer. In the case of frequency discrimination, the decline of performance as a function of $D$ was significantly less abrupt. This divergence suggests that pitch and loudness are processed in separate modules of auditory memory. © *1999 Acoustical Society of America.* [S0001-4966(99)03810-2]

PACS numbers: 43.66.Mk, 43.66.Fe, 43.66.Hg, 43.66.Cb [RVS]

## INTRODUCTION

In order to compare two sounds separated by some delay ($D$), it is of course necessary to memorize the first sound during the delay. Two modes of memory operation were distinguished by Durlach and Braida (1969). In one mode, called the ''trace mode,'' the sensation produced by the second sound is compared to the sensory trace left by the first sound. This comparison may benefit from an overt or covert rehearsal of the trace by the listener (Keller *et al.*, 1995), but its accuracy will strongly depend on $D$. In the other memory mode, called the ''context-coding mode,'' the listener compares instead symbolic (e.g., verbal) representations of the two sounds; these representations result from a categorical evaluation of each sound's relation to a general context of sounds (for instance the set of sounds used in an experiment). This memory mode will be generally less efficient than the trace mode if $D$ is short. However, given that categorical labels can be perfectly remembered for a long time, the context-coding mode can become the most efficient mode if $D$ is long.

The present study was concerned with the organization of auditory memory in the trace mode. Should this memory be viewed as a single ''store'' or is it composed instead of several stores with different properties? Some authors suggested that it includes two stores operating on different time spans: a ''short'' store and a ''long'' store (Wickelgren, 1969; Cowan, 1984). Another hypothesis, on which we focused here, is that there is a set of stores or sub-stores which are specialized in the retention of different perceptual attributes of sound.

This hypothesis has already received some support from psychophysical studies of the interference effects produced by a sound on the sensory trace of a previous sound. In previous experiments from our laboratory (Semal and Demany, 1991, 1993; Semal *et al.*, 1996), listeners were required to make same/different judgments on two periodic test sounds separated by a delay during which other periodic sounds (to be ignored) were presented. The differences to be detected were always differences in period, and thus pitch (but the listeners were not informed of that). It was found that discrimination of the test sounds strongly depended on the pitches of the intervening sounds—as observed before by Deutsch (1972)—but was essentially independent of the intervening sounds' other perceptual attributes. Therefore, these data suggested there is a memory store specialized in the retention of pitch and deaf to any other auditory attribute. Starr and Pitt (1997) obtained analogous results from an experiment in which the differences between the test sounds were differences in timbre (spectral shape). In this case, performance strongly depended on the timbre of the intervening sounds, but not on their pitch, as though timbre (or at least a certain aspect of timbre) was memorized in a specialized store. Let us finally mention data from Botte *et al.* (1992) concerning memory for loudness. In this study, the test sounds and intervening sounds were pure tones at a constant frequency; only intensity was varied. It was found that intensity discrimination between the test tones was not determined by the intervening tones' intensity distance to the test tones, but simply worsened monotonically as the intervening tones' intensity increased. This is in marked contrast with the results of analogous experiments on pitch memory: Here, the crucial factor is the similarity in pitch between the test tones and intervening tones (Deutsch, 1972; Semal and Demany, 1991). The divergence suggests that pitch traces and loudness traces are maintained in separate stores.

Assume that pitch and loudness are indeed processed in separate memory stores. It could then be the case that, in the absence of any intervening stimulus, the trace of a pitch sensation does not fade away with time at the same rate as the trace of a loudness sensation. This would be strong evidence for a separation of stores. The decay of loudness traces (i.e., intensity discrimination) as a function of time was investigated by a number of authors (Kinchla and Smyzer, 1967; Berliner and Durlach, 1973; Berliner *et al.*, 1977; Green *et al.*, 1983; Botte *et al.*, 1992; Lü *et al.*, 1992). There are also some data about the temporal decay of pitch traces (i.e., frequency discrimination: Wolfe, 1886; Harris, 1952; Bachem, 1954; Rakowski, 1972). However, the methods used in these two sets of experiments were widely different. To the best of our knowledge, nobody has compared the decays of pitch and loudness traces using the same subjects and similar procedures. This was the aim of the present study.

Since we wished to investigate the trace mode of auditory memory, it was important to minimize the possible influence of context-coding processes. Several studies of intensity discrimination indicated that the efficiency of context coding decreases as the stimulus range increases (Berliner and Durlach, 1973; Braida and Durlach, 1988). This led us to incorporate a roving procedure in our 2I-2AFC framework: From trial to trial, the period and/or intensity of the first (standard) stimulus were varied randomly within wide ranges. Generally speaking, avoiding the use of fixed standards hindered the formation of precise long-term memories in the course of the experiments (Harris, 1952).

The discriminability of two stimuli separated by $D$ s is determined in part by memory limitations but also depends, more basically, on the precision of their sensory encoding, i.e., on ''sensation noise'' (Durlach and Braida, 1969). For a fair comparison between pitch and loudness trace decays, it is desirable to keep constant the contribution of sensation noise to discrimination performance. We assumed that when $D$ is as short as 0.5 s, the amount of trace decay is negligible and discrimination performance is determined only by sensation noise.[1] Therefore, the two experiments reported here included preliminary measurements intended to select frequency and intensity changes that were equally discriminable for $D = 0.5\,\mathrm{s}$. We then varied $D$ and measured its effect on the discrimination of the corresponding changes.

The temporal decay of a pitch trace may depend on the salience of the initial pitch sensation, or more generally on the spectral properties of the stimulus eliciting the pitch sensation. Salient pitches are evoked by pure tones and by complex tones with harmonics that the auditory system is able to resolve. The pitch of complex tones consisting of unresolvable harmonics is less salient and may be extracted by a specific mechanism (Houtsma and Smurzynski, 1990; Carlyon, 1998). We used these three types of stimuli in experiment 1. Another potentially important factor was stimulus duration: It could be hypothesized that the trace of a long-duration sound decays less rapidly than the trace of a shorter one, all other things being equal. This led us to use very different stimulus durations in experiments 1 and 2.

## I. EXPERIMENT I

### A. Method

#### 1. Subjects

Four listeners without any known hearing deficit served as subjects. Two of them were authors SC and LD; they had a long previous experience with psychoacoustic experiments. The remaining two listeners (MM and MY) were students without such experience. None of the subjects possessed absolute pitch, but each of them had an interest in music and played a musical instrument.

#### 2. Tasks and stimuli

On each trial, two periodic stimuli separated by a silent delay $D$ were presented. Both stimuli (''$S1$'' and ''$S2$'') had a total duration of 500 ms and were gated on and off with 10-ms cosinusoidal amplitude ramps. There were four experimental conditions.

In condition INTENS, $S1$ and $S2$ were 1000-Hz pure tones differing in intensity. The direction of the intensity change was selected at random and the subject's task was to indicate which stimulus was louder. The SPL of $S1$ was randomly selected between 40 and 80 dB.

In the remaining three conditions (FREQ-PURE, FREQ-RES, and FREQ-UNRES), the two stimuli had a constant SPL of 60 dB but different periods. The direction of this change was also a random variable and the subject had to indicate which sound was higher in pitch. The frequency or fundamental frequency ($F0$) of $S1$ was randomly selected between limits specified below, frequency being in each case scaled logarithmically.

In condition FREQ-PURE, the stimuli were pure tones and the frequency of $S1$ varied between 500 Hz and 2000 Hz.

In condition FREQ-RES, the stimuli were complex tones consisting of the first five harmonics of some $F0$. The $F0$ of $S1$ varied between 100 and 400 Hz. Given their low ranks, the harmonics of each tone were resolvable by the auditory system (see, e.g., Plomp, 1976, Chap. 1). They were synthesized at equal amplitudes and added in sine phase.

In condition FREQ-UNRES, the stimuli were bandpass-filtered trains of 50-$\mu$s clicks, and the $F0$ of $S1$ varied between 40 and 200 Hz. The cutoff frequencies of the filters (Stanford Research SR640 and SR645; attenuation rate: about 100 dB/oct) were set at 2860 and 5000 Hz. In order to mask auditory distortion products (Plomp, 1976, Chap. 2), the stimuli were mixed with a white noise that was low-pass filtered at 2860 Hz and presented at 55 dB SPL. The amplitude of the clicks was systematically varied as a function of $F0$, in order to maintain the overall SPL of the stimuli at 60 dB. Given that $S1$ had a maximum $F0$ of 200 Hz and that all stimuli were high-pass filtered at 2860 Hz, the power spectrum of the stimuli consisted of equal-amplitude consecutive harmonics with ranks always exceeding 13. Such harmonics are not resolvable by the auditory system (see, e.g., Houtsma and Smurzynski, 1990).

Subjects were tested individually in a double-walled soundproof booth (Gisol, Bordeaux), using TDH39 earphones through which the stimuli were delivered diotically. Responses were given by pressing one of two buttons on a response box, and feedback was provided immediately: Following each correct response, an LED located just above the corresponding button was turned on for 300 ms; no LED was turned on if the response was wrong.

TABLE I. Thresholds measured in the preliminary phase of experiment 1.

| Condition (unit) | INTENS (dB) | FREQ-PURE (cents) | FREQ-RES (cents) | FREQ-UNRES (cents) |
|---|---|---|---|---|
| Subject LD | 1.8 | 8.2 | 5.9 | 108.9 |
| SC | 2.1 | 6.4 | 5.9 | 75.4 |
| MM | 1.7 | 6.9 | 5.3 | 84.6 |
| MY | 1.5 | 8.2 | 7.8 | 108.0 |
| Mean | 1.8 | 7.4 | 6.2 | 94.2 |

### 3. Preliminary measurements

In each of the four conditions defined above, we first determined the amount of stimulus change—in dB or in cents—for which the probability of a correct response was 0.80 when $D$ was 0.5 s. These ''thresholds'' were measured with the adaptive procedure described by Kaernbach (1991). In a given daily session, four blocks of trials were run in each condition. At the outset of a block, the change from $S1$ to $S2$ (in dB or in cents) was large enough to make the task easy. Following each correct response, this change was divided by $1.5^{1/4}$. Following each incorrect response, it was multiplied by 1.5. This continued until 14 reversals had occurred in the variation of the change. The median of the changes used on all trials following the fourth reversal was taken as the threshold.

Subjects were trained until their performances appeared to be stable. This took five 1-h sessions for subjects LD and SC, and nine 1-h sessions for subjects MM and MY. For each condition and subject, the threshold value finally recorded was the median of the last 20 threshold measurements.

### 4. Assessment of memory decay

In this main part of the experiment, $D$ was varied and the previously measured thresholds were used as constant changes from $S1$ to $S2$. (The changes had a constant size, but of course their direction was still a random variable.) In a given daily session, subjects were tested in only one of the four conditions. Each session began with a warm up consisting of 50 trials with $D = 0.5$ s, and then comprised 16 blocks of 20 trials. From block to block, $D$ varied in a sawtooth manner, taking four possible values: 0.5, 2, 5, and 10 s. When $D$ was equal to 0.5 or 2 s, there was a pause of 1 s between each response and the onset of $S1$ in the next trial. When $D$ was equal to 5 or 10 s, the pause had a duration of 5 s; 1 s before its end, a warning visual signal was produced by the LEDs of the response box.

From session to session, the four conditions were used alternately, four times each. Thus, for each subject, condition, and value of $D$, a total of 320 responses were collected. From these 320 responses, we computed four independent $d'$ statistics—one $d'$ per session—as well as the corresponding values of the response bias index $\beta$ (Green and Swets, 1974).

## B. Results

Table I displays the thresholds determined by the preliminary measurements and then used as constant stimulus changes.[2] Note that thresholds were much poorer in condition FREQ-UNRES than in conditions FREQ-PURE and FREQ-RES. This was predictable from the literature on frequency discrimination (e.g., Houtsma and Smurzynski, 1990).

Figure 1 shows the mean of the 16 $d'$ statistics obtained for each condition and value of $D$ in the main part of the experiment. For $D = 0.5$ s, $d'$ had an overall mean of 2.05. This $d'$ value is not very different from 1.68, the value expected from the threshold measurements under the assumption that, in these preliminary measurements, responses were unbiased—i.e., not affected by ''time-order errors''—(Macmillan and Creelman, 1991). Moreover, as we wished, $d'$ did not markedly vary with conditions for $D = 0.5$ s. In condition INTENS, $d'$ strongly decreased when $D$ was increased from 0.5 to 2 s, but $d'$ was not further reduced when $D$ was longer. The decline of $d'$ with $D$ appeared to be more gradual in the FREQ conditions. For each subject, the decline of $d'$ from $D = 0.5$ s to $D = 2$ s was smaller in each of the three FREQ conditions than in condition INTENS. An ANOVA in which sessions were used as the random factor

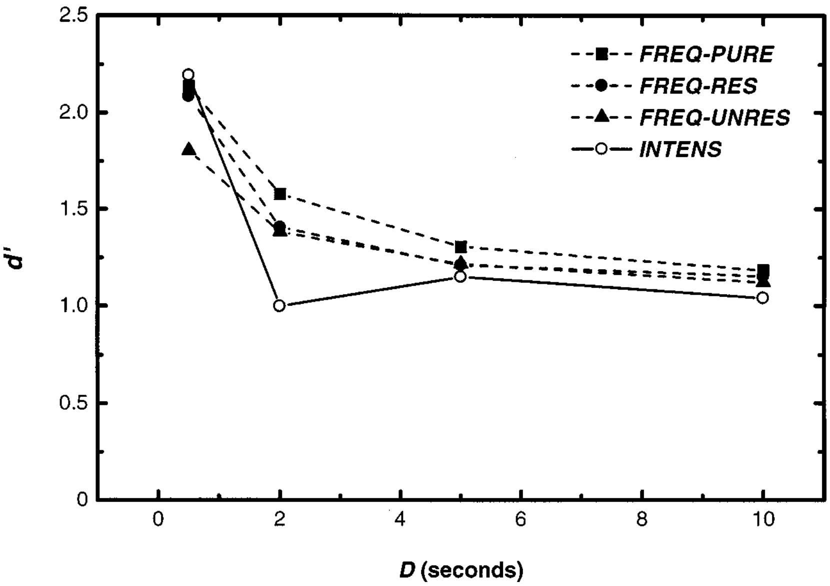

FIG. 1. $d'$ as a function of $D$ in the four conditions of experiment 1. Each data point represents the outcome of 1280 trials (4 sets of 80 trials for 4 subjects).

indicated that the effect of $D$ on $d'$ did not significantly differ across the three FREQ conditions $[F(6,18)<1]$, but was reliably different in condition INTENS $[F(3,9)=7.15, P=0.009]$. There was no significant three-way interaction between subjects, conditions, and $D$ $[F<1$ in each case$]$. Similar statistical tests were performed on the absolute values of $\log(\beta)$. Their outcomes were negative in each case. Thus the INTENS and FREQ conditions differed with regard to the effect of $D$ on $d'$ but not with regard to the effect of $D$ on the magnitude of response bias.[3]

### C. Discussion

In the three FREQ conditions, very different stimuli were used. For instance, whereas the tones used in condition FREQ-RES were quasi-vocal sounds, the tones of condition FREQ-PURE had pitches which were generally too high to be sung. More importantly, pitch was much less salient in condition FREQ-UNRES than in conditions FREQ-PURE and FREQ-RES. Yet, the three corresponding memory decays appeared to be similar. This suggests that the decay of a pitch memory trace is independent of the initial pitch sensation. However, this decay appears to differ from the decay of a loudness memory trace when one considers the results obtained in condition INTENS. Apparently, loudness traces decay more rapidly than pitch traces during the first two seconds following the stimulus.

In condition INTENS, $d'$ took similar values, close to 1.0, for $D=2$, 5, and 10 s. A reasonable interpretation of this plateau is that, for $D\geqslant 2$ s, listeners memorized loudness in the ‘‘context-coding’’ mode, which is more resistant to the passage of time than the ‘‘trace’’ mode (Durlach and Braida, 1969). In intensity discrimination tasks, the context-coding mode can be more efficient than the trace mode if the interstimulus interval ($D$) is long and if the overall intensity range is small (Berliner and Durlach, 1973). Here, $S1$ varied within a 40-dB range. This is a wide range in so far as the total dynamic range of the auditory system is barely three times larger. Within 40 dB, however, there are only 22 steps of 1.8 dB (the average threshold for condition INTENS, cf. Table I). By contrast, the 2-oct ranges used in conditions FREQ-PURE and FREQ-RES included more than 300 steps of 7.4 or 6.2 cents (the average thresholds for these two conditions). On this basis, it is reasonable to think that context coding was more profitable in condition INTENS than in conditions FREQ-PURE or FREQ-RES. In condition FREQ-UNRES, on the other hand, the average threshold was only 30 times smaller than the range of $S1$; yet, the effect of $D$ on $d'$ was much more similar to the effect observed in the other two FREQ conditions than to the effect observed in condition INTENS. Hence, it is clear that the form of the decays was not determined *only* by the ‘‘perceptual size’’ of the stimulus ranges.

Nonetheless, the $S1$ stimuli used in condition INTENS had a variable SPL but a fixed frequency whereas the reverse was true for the FREQ conditions. One could imagine that this difference biased in some way the main outcome of experiment 1. In order to demonstrate quite convincingly that pitch traces do not decay in the same manner as loudness traces, it is of course desirable to compare these decays using identical sets of $S1$ stimuli. This is what we did in experiment 2. Another important novelty of experiment 2 was that its stimuli were ten times shorter than those of experiment 1.

TABLE II. Thresholds measured in the preliminary phase of experiment 2.

| Condition (unit) | INTENS (dB) | FREQ (cents) |
|---|---|---|
| Subject LD | 2.3 | 14.8 |
| SC | 3.2 | 10.0 |
| VL | 2.8 | 11.3 |
| EB | 2.1 | 16.3 |
| Mean | 2.6 | 13.1 |

## II. EXPERIMENT 2

### A. Method

Four listeners with normal audiograms participated in this experiment. Two of them were again authors SC and LD. The remaining two listeners (VL and EB) were students with no previous psychoacoustic experience but a strong interest in (popular) music.

The method was basically similar to that used in experiment 1. However, all stimuli were pure tones and had a total duration of 50 ms rather than 500 ms; they were gated on and off with 5-ms cosinusoidal amplitude ramps. Subjects were tested in only two conditions: $S1$ and $S2$ could differ from each other in SPL (condition INTENS) or in frequency (condition FREQ). On each trial, for *both* conditions, the frequency of $S1$ was randomly selected between 1000 and 2500 Hz (using again a logarithmic frequency scale), and its SPL was randomly selected between 42 and 88 dB.

Preliminary threshold measurements were performed as before, the two new conditions being presented alternately. In the main part of the experiment, again, we used the measured thresholds as constant changes from $S1$ to $S2$. Only one condition was presented throughout each experimental session, and the two conditions alternated from session to session. A total of 320 trials were run for each subject, condition, and value of $D$. The corresponding data were analyzed exactly like those of experiment 1.

### B. Results and discussion

The measured thresholds are displayed in Table II. They were larger than those obtained in conditions INTENS and FREQ-PURE of experiment 1—an expected result since the stimuli were ten times shorter.

The four upper panels of Fig. 2 show the $d'$s obtained in the main part of the experiment for each subject. The means across subjects are presented in the bottom panel. Clearly, the overall results are very similar to those obtained in experiment 1. For $D=0.5$ s, $d'$ was close to 2.0 in both conditions. In condition INTENS, $d'$ markedly decreased when $D$ was increased to 2 s, but was approximately constant for $D=2$, 5, and 10 s. In condition FREQ, by contrast, $d'$ declined continuously with $D$ (or, in the case of subject LD, did not decline at all). From $D=0.5$ s to $D=2$ s, $d'$ varied much more in condition INTENS than in condition FREQ for three

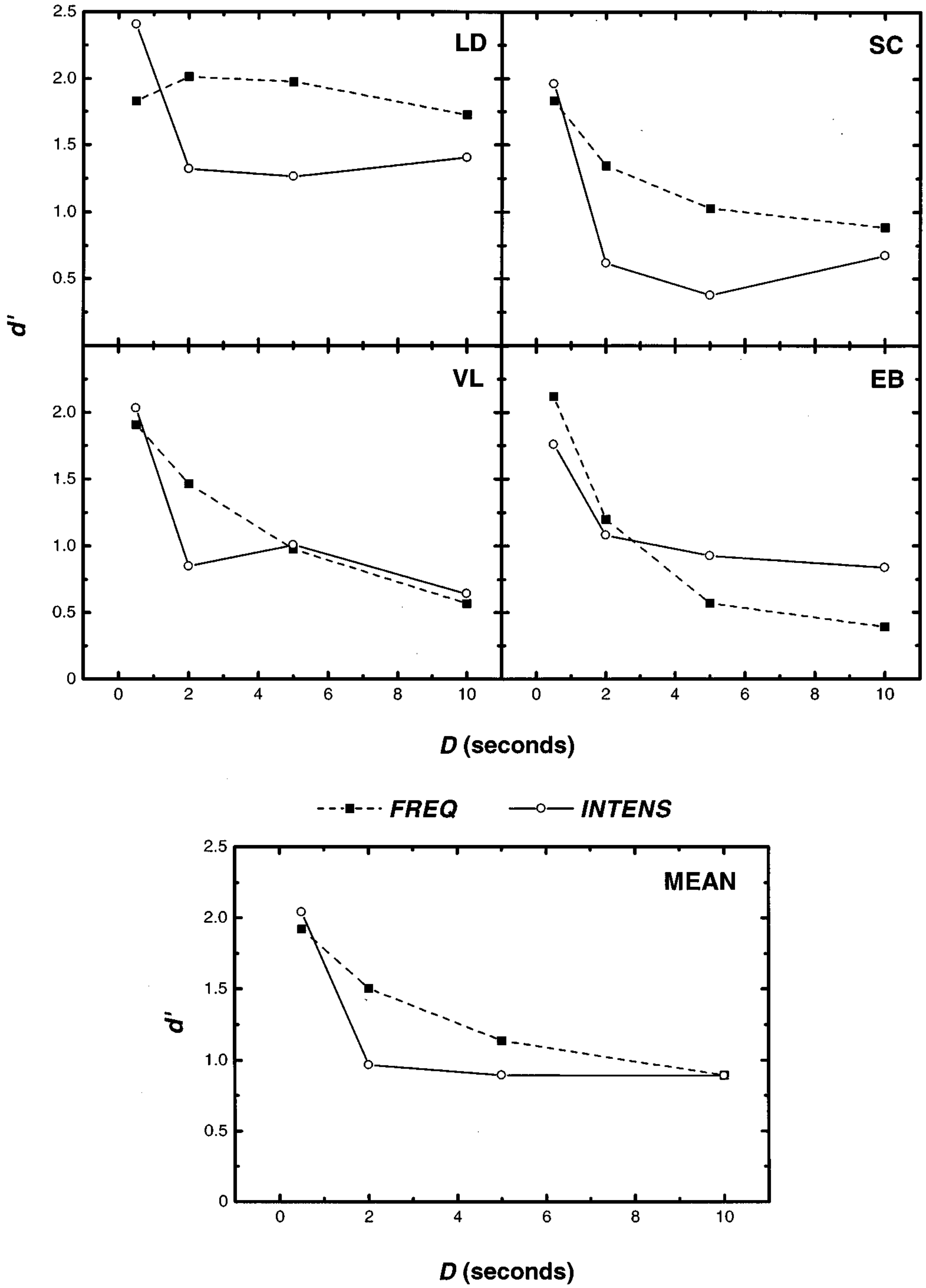

FIG. 2. *d′* as a function of *D* in the two conditions of experiment 2. Four upper panels: results obtained from each of the four subjects. Bottom panel: mean results.

subjects; however, this was not true for the fourth subject (EB). An ANOVA confirmed the existence of a significant interaction between *D* and the condition factor $[F(3,9) = 6.18,\ P = 0.014]$. A similar statistical test performed on $|\log(\beta)|$ rather than $d'$ yielded a negative result $[F(3,9) = 1.10,\ P = 0.399]$.

We undertook experiment 2 with the idea that, perhaps, the memory trace of a short tone decays more rapidly than the memory trace of a long tone. The results did not support this idea since they were very similar to those of experiment 1. Concerning the INTENS conditions of both experiments, one can argue that it was *a priori* impossible to observe a faster decay in experiment 2 if, as soon as *D* was equal to 2 s, performance reached a plateau determined by context-coding processes: To be able to demonstrate a difference in decay, we should have used at least one *D* value between 0.5 and 2 s. However, no such objection is possible concerning the FREQ conditions. It is important to note that because *D* was defined as the duration of the silence separating $S2$ from $S1$, differences in stimulus duration were associated with differences in onset-to-onset intervals. From experiment 1 to experiment 2, these intervals were reduced by 450 ms (500−50 ms). Thus they were almost halved when *D* was equal to 0.5 s. The fact that this big relative change did not significantly modify the effects of $D$ on $d'$ suggests that the shortest value of $D$ (0.5 s) was not short enough to truncate the *formation* [or ‘‘acquisition’’ (Wickelgren, 1969)] of an accurate memory trace of $S1$. If such truncatings had occurred for $D = 0.5$ s, they should have been larger for the shorter stimulus duration. Hence, from $D = 0.5$ s to $D = 2$ s, the decrease of $d'$ should have been smaller in experiment 2 than in experiment 1. There was a trend in this direction, for both the INTENS and the FREQ (or FREQ-PURE) conditions; but a comparison between Figs. 1 and 2 shows that these were very small trends.

## III. GENERAL DISCUSSION

In both experiments, we found that the effect of $D$ on $d'$ was not the same in conditions FREQ and INTENS. The difference was largest when $D$ varied from 0.5 to 2 s, and for these small values of $D$ what the difference reveals is almost certainly a divergence in the memory decay of sensory traces: Apparently, the memory decay of a loudness trace is more rapid than the memory decay of a pitch trace. In the INTENS conditions, it is likely that context coding was operative as soon as $D$ was equal to 2 s since $d'$ did not decline when $D$ was longer. But this would only mean that our results underestimated the rate of trace decay for loudness, and thus the divergence between loudness decay and pitch decay.

Three previous papers (Berliner and Durlach, 1973; Berliner *et al.*, 1977; Green *et al.*, 1983) reported experiments in which intensity discrimination (of pure tones) was measured as a function of inter-stimulus interval (i.e., $D$) with a roving procedure. Unfortunately, these three papers do not give a consistent picture of the memory decay of loudness traces. The results obtained by Berliner and his colleagues for wide roving ranges agree rather well with our data. They found that discrimination performance sharply decreases when $D$ increases up to 2.5 s, and that for longer values of $D$ an almost constant performance level is achieved thanks to context-coding processes (discrimination performance becomes similar to *identification* performance). By contrast, according to the results of Green *et al.*, performance does not decrease more between $D = 0.5$ s and $D = 2$ s than between $D = 2$ s and $D = 8$ s. Making sense of this discrepancy (ignored by Green *et al.*) is not easy. It may be significant that whereas Berliner and the present investigators measured $d'$ as a function of $D$ for fixed intensity changes, Green *et al.* measured instead, as a function of $D$, the values of intensity changes yielding a fixed $d'$.

Concerning frequency discrimination, the literature as a whole suggests that discrimination performance declines rather slowly with $D$. Harris (1952) performed on an enormous number of listeners an experiment which was analogous to that of Green *et al.* His stimuli were pure tones and the frequency of the first tone presented on each trial was roved between 950 and 1050 Hz. For $D = 0.1$, 1, 3, and 7 s, the measured discrimination thresholds increased by only 29% (from 4.2 Hz to 5.4 Hz). In the experiment of Green *et al.*, on the other hand, the measured thresholds increased by as much as 250% (from 2.4 dB to 6 dB) when $D$ varied from 0.5 s to 8 s. Therefore, our main finding does not come as a big surprise in the light of previous research. Note that a parallel can be drawn between this finding and the outcome of a recent study on *visual* short-term memory (Magnussen *et al.*, 1996). It was found by Magnussen *et al.* that the spatial frequency of a sinusoidal luminance grating was better memorized than its contrast. In the terminology proposed by Stevens (1966), loudness and perceived contrast are ‘‘prothetic’’ percepts whereas pitch and the perceptual correlate of spatial frequency are ‘‘metathetic’’ percepts. There might be a general law according to which the trace of a metathetic percept decays less rapidly than the trace of a prothetic percept.

From the fact that pitch traces and loudness traces do not decay at the same rate, it seems natural to infer that they are not retained in one and the same sensory store. We mentioned in the Introduction that previous psychophysical experiments already provided evidence for an autonomous processing of pitch (and of a certain aspect of timbre) in auditory memory. Let us point out here that there are also physiological data supporting the hypothesis of multiple and specialized auditory stores. When a listener is presented with a series of identical tones followed by a different tone, the different tone elicits an event-related brain potential called the ‘‘mismatch negativity’’ or MMN (Näätänen *et al.*, 1978). This brain potential is supposed to reflect a preattentive change detection based on a comparison between memory traces (see Schröger, 1997, for a recent review). According to Giard *et al.* (1995), the scalp topographies of the MMNs elicited by pure tones deviating from a repeated standard by either frequency, intensity, or duration vary with the type of stimulus deviance. Thus the corresponding MMNs originate from at least partly distinct neural populations (in the auditory cortex). Another remarkable fact is that the MMN obtained in response to a two-dimensional change in frequency and spatial location, or frequency and duration, or duration and intensity, is equal to the sum of the MMNs elicited by its one-dimensional components, exactly as if each of the combined one-dimensional components elicited its own MMN (see, e.g., Levänen *et al.*, 1993). Disappointingly, however, a similar summation does not seem to occur for combined changes in frequency and intensity (Wolff and Schröger, 1995).

Our main finding is consistent with the idea that the mnemonic processings of pitch and loudness are *completely* separate, but it is also consistent with a more subtle hypothesis. Assume that the architecture of auditory memory, in what Durlach and Braida (1969) called its ‘‘trace’’ mode, consists of: (1) an all-purpose ‘‘short’’ store retaining *global* ‘‘echoic’’ traces during a limited time; (2) a set of specialized stores permitting each a longer retention of a single auditory attribute. It could then be the case that only the short store is available for the retention of loudness, whereas one of the specialized stores is devoted to pitch. (Of course, some categorical information on loudness could nonetheless be kept for a long time by means of context-coding processes.) Lü *et al.* (1992) assessed psychophysically the decay of a loudness trace and found that it had the same lifetime—about 2 s—as the decay of the neural activation produced by the stimulus in the primary auditory cortex (this neural activation being assessed by magneto-encephalography). One may speculate on this basis that the neural site of the short store is the primary auditory cortex while the specialized stores are located elsewhere.

Finally, let us come back on the fact that for one of the six listeners tested in the present study (subject EB, experiment 2), we found *no* evidence that pitch traces decay less rapidly than loudness traces. It is worthy to note that EB was probably the subject who ranked last in terms of musical practice. This suggests that a correlation might exist between pitch memory and musical experience—a suggestion already made by Pechmann and Mohr (1992). An interesting goal of future research would be to determine if indeed trace decay in auditory memory is correlated with musical experience, and more strongly for pitch traces than for loudness traces.

## ACKNOWLEDGMENTS

This work is a part of the first author's doctoral dissertation. We thank the Conseil Régional d'Aquitaine for its support, as well as Ed Burns and an anonymous reviewer for comments on an earlier version of the manuscript.

[1] This assumption is consistent with the results of Harris (1952) concerning frequency discrimination, as well as those of Kinchla and Smyzer (1967) and Green *et al.* (1983) concerning intensity discrimination. It is not consistent, however, with data reported by Berliner and Durlach (1973) and Berliner *et al.* (1977): According to these authors, for relatively long tone bursts ($\geqslant$500 ms), intensity discrimination worsens significantly as soon as $D$ exceeds 0. Berliner *et al.* (1977) mention that, ''for reasons unknown to [them]'' (p. 1579), their results for fixed standard tones are very different from those obtained by Kinchla and Smyzer (1967). In the present study, we thought that it was not desirable to set $D$ below 0.5 s because two problems may arise if $D$ is very small and the stimuli are rather short: (1) the first stimulus may have a deleterious ''forward-masking'' effect on the second one (Jaroszewski and Rakowski, 1976; Zeng and Turner, 1992); (2) the formation of an accurate memory trace of the first stimulus may not be complete when the second stimulus is presented (Wickelgren, 1969; Massaro, 1975).

[2] For subject MM, during the final experimental session run in each of the three FREQ conditions, the stimulus change was smaller than the threshold indicated in Table I; we respectively used changes of 5.7, 3.0, and 70.0 cents in conditions FREQ-PURE, FREQ-RES, and FREQ-UNRES. These deliberate decreases of the stimulus changes were intended to avoid ceiling effects for $D=0.5$ s: Apparently, MM had not reached her maximum level of performance in the preliminary experimental phase.

[3] The feedback provided on each trial allowed the subjects to reduce or eliminate their ''natural'' responses biases. For this reason, the obtained values of $\beta$ were of little interest by themselves.

Bachem, A. (**1954**). ''Time factors in relative and absolute pitch determination,'' J. Acoust. Soc. Am. **26**, 751–753.

Berliner, J. E., and Durlach, N. I. (**1973**). ''Intensity perception. IV. Resolution in roving-level discrimination,'' J. Acoust. Soc. Am. **53**, 1270–2287.

Berliner, J. E., Braida, L. D., and Durlach, N. I. (**1977**). ''Intensity perception. VII. Further data on roving-level discrimination and the resolution and bias edge effects,'' J. Acoust. Soc. Am. **61**, 1256–1267.

Botte, M. C., Baruch, C., and Mönikheim, S. (**1991**). ''Memory for loudness: the role of loudness contour,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 305–311.

Braida, L. D., and Durlach, N. I. (**1988**). ''Peripheral and central factors in intensity perception,'' in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 559–584.

Carlyon, R. P. (**1998**). ''The effects of resolvability on the encoding of fundamental frequency by the auditory system,'' in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London), pp. 246–254.

Cowan, N. (**1984**). ''On short and long auditory stores,'' Psychol. Bull. **96**, 341–370.

Deutsch, D. (**1972**). ''Mapping of interactions in the pitch memory store,'' Science **175**, 1020–1022.

Durlach, N. I., and Braida, L. D. (**1969**). ''Intensity perception. I. Preliminary theory of intensity resolution,'' J. Acoust. Soc. Am. **46**, 372–383.

Giard, M. H., Lavikainen, J., Reinakinen, K., Perrin, F., Bertrand, O., Pernier, J., and Näätänen, R. (**1995**). ''Separate representation of stimulus frequency, intensity and duration in auditory sensory memory: An event-related potential and dipole-model analysis,'' J. Cognit. Neurosci. **7**, 133–143.

Green, D. M., Kidd, G., and Picardi, M. C. (**1983**). ''Successive versus simultaneous comparison in auditory intensity discrimination,'' J. Acoust. Soc. Am. **73**, 639–643.

Green, D. M., and Swets, J. A. (**1974**). *Signal Detection Theory and Psychophysics* (Krieger, Huntington, NY).

Harris, J. D. (**1952**). ''The decline of pitch discrimination with time,'' J. Exp. Psychol. **43**, 96–99.

Houtsma, A. J. M., and Smurzynski, J. (**1990**). ''Pitch identification and discrimination for complex tones with many harmonics,'' J. Acoust. Soc. Am. **87**, 304–310.

Jaroszewski, A., and Rakowski, A. (**1976**). ''Pitch shifts in post-stimulatory masking,'' Acustica **34**, 220–223.

Kaernbach, C. (**1991**). ''Simple adaptative testing with the weighted up-down method,'' Percept. Psychophys. **49**, 227–229.

Keller, T. A., Cowan, N., and Saults, J. S. (**1995**). ''Can auditory memory for tone pitch be rehearsed?'' J. Exp. Psychol. (Learn., Mem. and Cogn.) **21**, 635–645.

Kinchla, R. A., and Smyzer, F. (**1967**). ''A diffusion model of perceptual memory,'' Percept. Psychophys. **2**, 219–229.

Levänen, S., Hari, R., McEvoy, L., and Sams, M. (**1993**). ''Responses of the human auditory cortex to changes in one versus two stimulus features,'' Exp. Brain Res. **97**, 177–183.

Lü, Z. L., Williamson, S. J., and Kaufman, L. (**1992**). ''Behavioral lifetime of human auditory sensory memory predicted by physiological measures,'' Science **258**, 1668–1670.

Macmillan, N. A., and Creelman, C. D. (**1991**). *Detection Theory: A User's Guide* (Cambridge University Press, Cambridge, U.K.).

Magnussen, S., Greenlee, M. W., and Thomas, J. P. (**1996**). ''Parallel processing in visual short-term memory,'' J. Exp. Psychol. (Hum. Percept. Perform.) **22**, 202–212.

Massaro, D. W. (**1975**). ''Backward recognition masking,'' J. Acoust. Soc. Am. **58**, 1059–1065.

Näätänen, R., Gaillard, A. W. K., and Mäntysalo, S. (**1978**). ''Early selective attention effect on evoked potential reinterpreted,'' Acta Psychol. **42**, 313–329.

Pechmann, T., and Mohr, G. (**1992**). ''Interference in memory for tonal pitch: Implications for a working-memory model,'' Mem. Cognition **20**, 314–320.

Plomp, R. (**1976**). *Aspects of Tone Sensation* (Academic, London).

Rakowski, A. (**1972**). ''Direct comparison of absolute and relative pitch,'' in *Proceedings of the Symposium on Hearing Theory* (Institut voor Perceptie Onderzoek, Eindhoven, Holland).

Schröger, E. (**1997**). ''On the detection of auditory deviations: a pre-attentive activation model,'' Psychophysiology **34**, 245–257.

Semal, C., and Demany, L. (**1991**). ''Dissociation of pitch from timbre in auditory short-term memory,'' J. Acoust. Soc. Am. **89**, 2404–2410.

Semal, C., and Demany, L. (**1993**). ''Further evidence for an autonomous processing of pitch in auditory short-term memory,'' J. Acoust. Soc. Am. **93**, 1315–1322.

Semal, C., Demany, L., Ueda, K., and Hallé, P. A. (**1996**). ''Speech versus nonspeech in pitch memory,'' J. Acoust. Soc. Am. **100**, 1132–1140.

Starr, G. E., and Pitt, M. A. (**1997**). ''Interference effects in short-term memory for timbre,'' J. Acoust. Soc. Am. **102**, 486–494.

Stevens, S. S. (**1966**). ''On the operation known as judgment,'' Am. Sci. **54**, 385–401.

Wickelgren, W. A. (**1969**). ''Associative strength theory of recognition memory for pitch,'' J. Math. Psychol. **6**, 13–61.

Wolfe, H. K. (**1886**). ''Untersuchungen über das Tongedächtniss,'' Philos. Stud. (Wundt) **3**, 534–571.

Wolff, C., and Schröger, E. (**1995**). ''MMN elicited by one-, two-, and three-dimensional deviants,'' J. Psychophysiol. **9**, 374.

Zeng, F. G., and Turner, C. W. (**1992**). ''Intensity discrimination in forward masking,'' J. Acoust. Soc. Am. **92**, 782–787.