Global loudness of rising- and falling-intensity tones: How temporal profile characteristics shape overall judgments

Emmanuel Ponsot,[1,a)] Patrick Susini,[1] and Sabine Meunier[2]
[1]STMS Laboratory (IRCAM, CNRS, UPMC), 1 place Igor Stravinsky, 75004 Paris, France
[2]Laboratoire de Mécanique et d'acoustique-CNRS, Unité Propre de Recherche 7051, Aix-Marseille University, Centrale Marseille, 4 Impasse Nikola Tesla, CS-40006 Marseille Cedex 13, France

(Received 18 January 2017; revised 19 June 2017; accepted 22 June 2017; published online 19 July 2017)

The mechanisms underlying global loudness judgments of rising- or falling-intensity tones were further investigated in two magnitude estimation experiments. By manipulating the temporal characteristics of such stimuli, it was examined whether judgments could be accounted for by an integration of their loudest portion over a certain temporal window associated to a "decay mechanism" downsizing this integration over time for falling ramps. In experiment 1, 1-kHz intensity-ramps were stretched in time between 1 and 16 s keeping their *dynamics* (difference between maximum and minimum levels) unchanged. While global loudness of rising tones increased up to 6 s, evaluations of falling tones increased at a weaker rate and slightly decayed between 6 and 16 s, resulting in significant differences between the two patterns. In experiment 2, ramps were stretched in time between 2 and 12 s keeping their slopes (rate of change in dB/s) unchanged. In this context, the main effect of duration became non-significant and the interaction between the two profiles remained, although the decay of falling tones was not significant. These results qualitatively support the view that the global loudness computation of intensity-ramps involves an integration of their loudest portions; the presence of a decay mechanism could, however, not be attested.
[http://dx.doi.org/10.1121/1.4991901]

[JFL] Pages: 256–267

I. INTRODUCTION

Global loudness, which has been defined as the overall impression of loudness of stimuli varying in loudness across time (e.g., Susini *et al.*, 2007), is an important psychoacoustical variable when dealing with time scales of several seconds. For instance, the industry and the media need to predict and, therefore, to control, as accurately as possible, how loud a sound will be perceived. The sound sequences considered are generally long (at least a few seconds) and strongly vary in level through time, e.g., the passing-by of an airplane, an advertisement broadcasted on the radio. Overall indicators are thus required to evaluate their *global loudness*, i.e., listeners' overall impressions; the primary purpose is often to control their perceived annoyance, which strongly relies on this variable (see Dittrich and Oberfeld, 2009). Psychoacoustical experiments investigating dynamic loudness perception of sound sequences lasting several seconds have shown that global loudness does not correspond to an average of momentary loudness, i.e., to the average loudness experienced during the stimulus, but is rather strongly influenced by the loudest events (e.g., Kuwano and Namba, 1985; Gottschling, 1999; Kuwano *et al.*, 2003; Susini *et al.*, 2002, 2007). Current indicators of global loudness are all based on this outcome. In the media, the overall loudness of a program is simply taken as the integration of its momentary loudness values (predicted by simplified auditory models), which exceed a certain threshold taken as the mean loudness value minus ten units (see ITU-R BS.1770, 2006; EBU-Recommendation, 2011). Other indicators employed in the industry, such as L_{Aeq}, also rely on the assumption that global loudness could be evaluated by averaging the physical energy of the stimulus (this issue is discussed in Oberfeld and Plank, 2011). Even in the context of more basic psychoacoustics, Zwicker and Fastl (1999) proposed to use the maximum value or a certain percentile (e.g., 95th) of the inferred loudness temporal distribution as a predictor. Glasberg and Moore (2002) suggested using the peak of the "short-term loudness" (STL) or the "long-term loudness" (LTL) time series predicted by their model to estimate the overall loudness of time-varying sounds. However, recent studies conducted with very basic sounds have pointed out a limitation to such assumptions. For instance, it has been shown that 1-kHz tones increasing linearly in level during 2-s over a 15-dB dynamics (i.e., range of level variation) are consistently judged about 3–4 dB louder than their time-reversed, falling versions, and it was demonstrated that this asymmetry could not be accounted for by current loudness models (Ponsot *et al.*, 2015a). It was, in all cases, significantly underestimated. Neither the maximum value of STL or LTL outputs (which is the operation most often considered for "long" stimuli varying slowly in loudness, e.g., Ries *et al.*, 2008; see also Moore *et al.*, 2016, for a discussion) could account for its size.[1] Finally, these results show that, even with very basic 1-kHz stimuli ramping up or down in level, global loudness is not simply based on simple operations (average, maximum) computed on the basis of short-term or long-term loudness patterns. The mechanisms

---

[a)]Electronic mail: emmanuel.ponsot@ircam.fr

that underpin global loudness evaluations of such ramps still remain undetermined.

Our listening environment contains lots of rising- and falling-level events lasting a few seconds. Real moving sound sources (e.g., a car passing by) present increasing and decreasing level profiles (induced by the approaching and the receding portions, respectively) and musical sequences are full of *crescendo* and *decrescendo* passages. It is thus particularly valuable to understand how intensity dynamics occurring at this time scale are processed and why asymmetries in global loudness judgments of simple rising vs falling profiles occur. Focusing on the perceptual processing of such contours also provide the possibility to put the present research in the context of the perception of the dynamics of looming/receding sounds (Neuhoff, 1998; for a recent review of the literature, see Olsen, 2014).

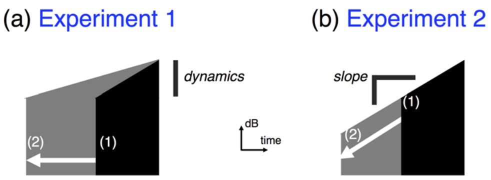
The purpose of the present study was twofold: (1) to better understand the mechanisms underlying the global loudness evaluation of basic 1-kHz rising- and falling-intensity tones of several seconds, and (2) to explore to what extent asymmetries between rising and falling tones are influenced by their temporal profile characteristics. These closely interrelated issues were addressed by means of two psychophysical experiments in which the parameters defining the temporal profiles of these linear rising and falling ramps of sound level, namely, their *duration*, their *slope* (i.e., rate of change in dB/s), and their *dynamics* (i.e., difference between minimum and maximum levels in dB), were manipulated to test specific hypotheses with regard to potential underlying mechanisms.

Most previous studies investigating *global loudness* of simple rising and falling sounds at this time scale employed ramps with the same combination of slope, duration, and dynamics. Most often, the ramps were 2-s long and covered a dynamics of 15-dB, thus, resulting in a slope of 7.5 dB/s (Ponsot *et al*., 2013; Ponsot *et al*., 2015a; Ponsot *et al*., 2015b).

One study examined the effect of the dynamics on the global loudness of 1.8-s rising ramps (Susini *et al*., 2010). For rising ramps having the same maximum level, greater global loudness estimates were found for ramps with 15-dB dynamics as compared to those with 30-dB dynamics. Furthermore, global loudness estimates of these rising ramps were close but slightly lower than those of constant-intensity tones presented at their maximum level. To explain these effects, the authors proposed that global loudness evaluation of rising ramps might involve a certain integration of its level-profile over a temporal window located around the maximum of the stimulus (see Meunier *et al*., 2010). It is important to note at this stage that, while this concept of a temporal integration over the loudest portions of the ramps might belong to or mirror the same class of phenomena as do typical temporal integration of loudness in psychoacoustics, there are significant divergences between the two. In the traditional temporal integration literature, the time constants refer to mechanisms operating at quite short durations ~50–100 ms (e.g., see Buus *et al*., 1997; Hots *et al*., 2014). The temporal integration phenomenon we consider here likely operates at a much coarser time scale and presumably takes place in higher-level cognitive stages (we discussed and illustrated these notions in Ponsot *et al*., 2016; see, in particular, Sec. II and Fig. 1). Such a mechanism, which we will call the *integration mechanism* throughout this paper, is consistent, at least qualitatively, with the observation that when the dynamics of these ramps are decreased while their maximum level remains the same, a greater amount of energy is contained under the temporal window and, hence, global loudness increases. This integration mechanism is also consistent with the fact that rising ramps are perceived softer than their maximum level. According to this hypothesis, if the duration of a ramp is increased but its dynamics is fixed, global loudness should also increase. No experiment directly tested this assumption with simple rising ramps, but there is one study that investigated the influence of the duration on global loudness judgments of time-varying 1-kHz tones, which consisted of sequences of stationary tones plus ramps. For sound sequences made of a 3-s constant plateau followed by a rising ramp, *global loudness* was found to increase gradually when the duration of the ramp increased between 2 and 20 s while its dynamics was kept constant, equal to 20 dB (see Susini *et al*., 2007). This result indirectly supports the idea that a certain integration mechanism might be involved.

In Meunier *et al*. (2010), it was hypothesized that the same integration mechanism of the loudest portion would act with falling ramps: at equal duration, the global loudness of a falling ramp would decrease when its dynamics is increased, whereas at equal dynamics, its global loudness should increase with duration. However, when dealing with falling-intensity stimuli of a few seconds, there is another phenomenon that needs to be taken into account. Indeed, a number of studies observed that global loudness judgments were greater when the loudness peak was closer to the end of the sequence (Hellbrück, 2000; Susini *et al*., 2002; Kuwano *et al*., 2003). These authors suggested that this might reflect a "memory process": the loudness peak having a smaller impact on the overall evaluation when its encoding in memory is further in time; "recency" was proposed as the candidate mechanism (Susini *et al*., 2002). This effect is strongly related to the "peak-end" rule (Kahneman *et al*., 1993; Schreiber and Kahneman, 2000), which has been specifically introduced and discussed with regard to loudness (Dittrich and Oberfeld, 2009). In what follows, specifically for down-ramps, we will thus simply refer to this phenomenon as the *decay mechanism*, because it is assumed to downsize the influence of a loudness peak as a function of the time lapse between its position and the end of the sound.[2] It is, however, impossible to tell yet beyond which durations this mechanism really starts to be involved and what is the typical rate of decay it consists of. We believe this mechanism could be partly responsible for the asymmetry observed between 2-s rising and falling ramps (Ponsot *et al*., 2015a; Ponsot *et al*., 2015b) and that it might emphasize this asymmetry for longer ramp durations (e.g., 10 s), since the loudness peak of falling tones is then clearly further back away in time (Susini *et al*., 2007). Therefore, it could be hypothesized that, when the duration of a falling ramp is increased while its dynamics is fixed, global loudness judgments result from the product of two mechanisms: (i) the integration mechanism, which increases global loudness, and (ii) the decay mechanism, which decreases global loudness. Whether the sum of these two processes leads, in the end, to an increase or a decrease of global loudness as a function of the duration can, however, not be predicted. In Susini *et al.* (2007), global loudness of sequences containing falling ramps (of fixed dynamics) followed by constant plateau was found to remain fairly constant when ramp duration was increased from 2 to 20 s. However, this result cannot directly be transposed to the present context, e.g., to suggest that the two mechanisms have similar weights in the process, as the presence of a plateau at the end of the sequence might have significantly affected the integration processes specifically related to the ramp itself (more than in the case of rising sequences where the plateau was located at the beginning). The fact that the two mechanisms might potentially be operating simultaneously with falling tones can nevertheless be observed by looking at the evolution of the so-called "asymmetry" (i.e., the difference of global loudness between rising and falling sounds) as a function of their duration when their dynamics are held constant: this asymmetry should increase. The decay mechanism should also be observed independently by increasing falling ramps' duration while keeping their slope constant. Indeed, the integration mechanism, which is only influenced by the shape of the first (loudest) portion of a falling ramp would, when the slope is unchanged, always provide the same global loudness quantity whatever the duration of that ramp is. Such a manipulation should decrease global loudness, directly reflecting the effect of the decay mechanism.

These hypotheses remain, however, somewhat speculative as they are derived from a small number of studies with different experimental procedures, which in some cases, did not use simple rising and falling ramps but more complex sound sequences. The purpose of the present study was, thus, to directly address with the same experimental designs the plausibility that the two proposed mechanisms might be involved in the global loudness processing of rising and falling tones. This was examined, in particular, in the context of direct global loudness judgments using magnitude estimation tasks. Two psychophysical experiments were designed to disentangle the two presumed mechanisms by manipulating the parameters of the ramps (i.e., their slope, duration, and dynamics), which consisted of 1-kHz stimuli either rising or falling linearly in level,[3] like those employed in our previous studies (Ponsot *et al.*, 2015a; Ponsot *et al.*, 2015b). These manipulations are illustrated in Fig. 1. In experiment 1, the ramps were "stretched" in time while keeping their dynamics constant [cf. Fig. 1(a)], resulting in different combinations of slope and duration. This is similar to what was done in Susini *et al.* (2007) with more complex sequences. In this context, as mentioned above, we hypothesized that (1) global loudness of rising tones should increase with duration because of the integration mechanism and (2) global loudness of falling tones should not grow as fast as for rising tones because both the *integration* and the decay mechanism would add. In experiment 2, the ramps were stretched in time in such a way that their slope was kept constant [cf. Fig. 1(b)], resulting in different combinations of dynamics and duration. In that context, we hypothesized that (1) global loudness of rising tones should not vary with duration because, as stated earlier, a constant temporal window located on its loudest portion always integrates the same amount of energy, and (2) global loudness of falling tones would decay over time, providing a direct image of the decay mechanism.


(a) Experiment 1

(b) Experiment 2

FIG. 1. (Color online) Schematic representation of the two experiments conducted in the present study where the duration of rising- and falling-intensity ramps was manipulated in different ways [as indicated by the arrows, a rising ramp taken here, as an example, could be stretched from (1) to (2)]. (a) Experiment 1: The ramps were "stretched" in time while keeping their *dynamics* constant. (b) Experiment 2: The ramps were stretched in time while keeping their *slope* constant. A rising ramp is taken as an example here, but the similar operation was realized on falling ramps.

Furthermore, we also wanted to investigate throughout these two experiments the extent to which asymmetries between global loudness of rising and falling tones vary with the manipulated parameters, namely, the slope, the dynamics and the duration of these ramps. Due to the presumed decay mechanism operating with falling tones, we were expecting an increase of the asymmetry with the duration of the time stretching in both experiments. Finally, the experimental design also attempted to determine to which extent the effects of the ramp parameters (slope, duration, dynamics) on both global loudness judgments and their resulting asymmetries depend on the mean intensity of the stimuli. Indeed, we already observed in previous studies that the asymmetry between rising and falling ramps was significantly reduced when the maximum level of the stimuli was higher than 80 dB sound pressure level (SPL), an effect that remained unexplained so far (Ponsot *et al.*, 2015a; Ponsot *et al.*, 2015b). Thus, in both experiments, ramps were presented in different intensity-regions, below and above 80 dB SPL. To complete this investigation, we examined to what extent different global loudness indicators derived from the outputs of the loudness model of Glasberg and Moore (2002) could account for the results collected.

II. EXPERIMENT 1

A. Materials and method

1. Participants

Forty-five participants were recruited for this experiment. They were divided into two groups, performing the experiment under different conditions (see Sec. II A 2 below): group A, 30 participants (15 women, 15 men; age 22–35 years old); group B, 15 participants (8 women, 7 men; age 18–32 years old). All reported normal hearing. They gave their informed written consent according to the Declaration of Helsinki prior to the experiment and were paid for their participation. The participants were naive with respect to the hypotheses under test.

2. Stimuli

The stimuli were 1-kHz pure tones with various durations and intensity profiles. The loudness function of each participant was measured prior to the experiment using 500-ms constant-intensity tones (presented at ten different levels equally spaced between 45 and 90 dB SPL). The data collected for these constant-intensity tones were not used in the paper, except for normalizing the ratings attributed to the ramps (see Sec. II B). In the experiment, tones with rising- and falling-intensity profiles were used; their sound level was linearly varied over 15 dB (i.e., 15-dB dynamics). They were presented in four regions: $R_1 = [60–75]$, $R_2 = [65–80]$, $R_3 = [70–85]$, and $R_4 = [75–90\,\text{dB SPL}]$. Participants of group A were presented with ramps of five different durations (1, 2, 6, 9, and 12 s); participants of group B were presented with another set of three durations (4, 8, and 16 s). The amplitude envelopes of the stimuli were all smoothed with 10-ms linear rise and fall times.

3. Apparatus

The stimuli were generated at a sampling rate of 44.1 kHz with 16-bit resolution using MATLAB. Sounds were converted using a RME Fireface 800 soundcard (Haimhausen, Germany), amplified using a Lake People G-95 Phoneamp amplifier (Konstanz, Germany) and presented diotically through headphones (Sennheiser HD 250 Linear II, Wedemark, Germany). Sound level was calibrated using a Brüel and Kjær artificial ear (type 4153, IEC318, Nærum, Denmark). Participants were tested in a double-walled (IAC) sound-insulated booth at Ircam.

4. Procedure

An absolute magnitude estimation (AME) procedure was used, based on the instructions of Hellman (1982). No standard was given to the participants. Their task was to give a number proportional to the *global loudness* of each sound, i.e., the overall impression loudness over the total sound duration (Ponsot *et al.*, 2013; Ponsot *et al.*, 2015a; Susini *et al.*, 2007; Susini *et al.*, 2010). For each participant, the experiment was scheduled in one session lasting about one hour. The measurement of the loudness function was done at the beginning of the session. After 20 training trials, each tone was presented 9 times in a "pseudo-random" order to reduce sequential effects (Cross, 1973), as it was done previously (see Ponsot *et al.*, 2015a). The experiment continued with the presentation of rising and falling ramps at various durations and intensities. A blocked-duration design was adopted, i.e., each block was made of sounds having the same duration. Each block consisted of interleaved rising and falling ramps of equal-duration presented at the four different intensity-regions, as mentioned in Sec. II A 2. Each stimulus was presented five times. Thus, a total of 200 stimuli (2 directions × 4 intensity-regions × 5 durations × 5 repetitions) were presented to the participants of group A and a total of 120 stimuli to the participants of group B (2 directions × 4 intensity-regions × 3 durations × 5 repetitions). The order of presentation of the blocks was randomly chosen for each participant.

B. Results

For each listener of each group, the average perceived global loudness of each stimulus was computed using the geometric mean of all his/her ratings. These mean loudness estimates were then normalized individually.[4]

The data were analyzed separately for each group. Repeated measures analyses of variance (rmANOVAs; direction × duration × intensity-region) with univariate approaches were performed on the logarithm of the normalized loudness ratings accorded to rising and falling ramps within each group, respectively. The statistical analyses were conducted using R (R Core Team, 2015). All the tests were two-tailed and used a probability level of 0.05 to test for significance. The Huynh-Feldt corrections for degrees of freedom were used where appropriate. Effect sizes are reported using partial eta-squared η_p^2.

The normalized magnitude estimates obtained in each group are presented in Fig. 2, on a *y*-log scale and as a function of the duration of the ramp. Overall, global loudness estimates of rising tones (△) appeared to be always greater than (or at least equal to) those given to their time-reversed versions, i.e., falling tones (▼). This was supported by significant effects of the direction obtained both for group A $[F(1{,}29) = 25.07,\ p < 0.001,\ \eta_p^2 = 0.464]$ and for group B $[F(1{,}14) = 12.14,\ p = 0.004,\ \eta_p^2 = 0.464]$. Furthermore, the averaged plots in Fig. 2 for both groups A and B showed that global loudness increased with duration for both ramp directions, at least until 6 s, but that the speed of this growth might differ between rising and falling tones. Beyond 6 s, global loudness tended to remain constant for rising-intensity tones, whereas it seemed that there might be a

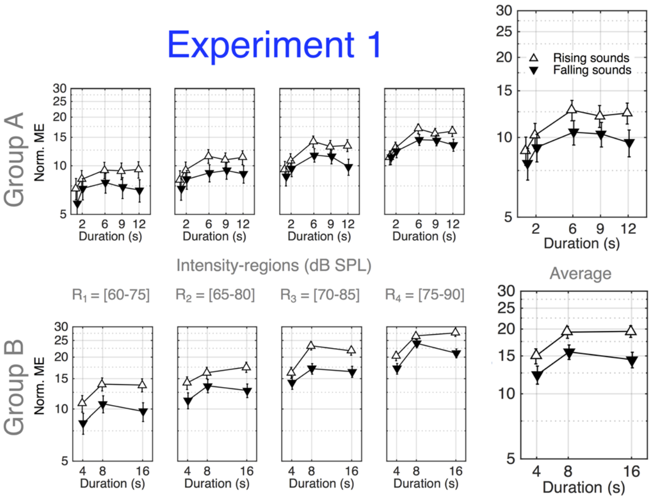


FIG. 2. (Color online) Normalized estimates of global loudness obtained in experiment 1 in the two conditions (group A on top, who received 15-dB ramps of 1, 2, 6, 9, and 12 s; group B at bottom, who received 15-dB ramps of 4, 8, and 16 s) for both rising (△) and falling sounds (▼). Results are plotted a *y*-log axis as a function of the duration of the ramp respective to each group for the different intensity-regions on the left panels (from R_1 to R_4) and averaged on the rightmost panels. Error bars show standard errors of the mean (SEM) in each configuration.

slight decrease for falling-intensity tones. The analyses in group A showed a significant effect of duration [$F(4,116) = 4.70$, $p = 0.013$, $\eta_p^2 = 0.139$, $\tilde{\varepsilon} = 0.49$] and a significant duration × direction interaction [$F(4, 116) = 4.60$, $p = 0.009$, $\eta_p^2 = 0.137$, $\tilde{\varepsilon} = 0.61$]. In group B, a significant effect of duration was found [$F(2,28) = 5.82$, $p = 0.013$, $\eta_p^2 = 0.294$, $\tilde{\varepsilon} = 0.81$] but the duration × direction interaction was only marginally significant [$F(2,28) = 3.20$, $p = 0.056$, $\eta_p^2 = 0.186$, $\tilde{\varepsilon} = 1.15$].

To gain further insight into the duration × direction interaction obtained in group A, multiple *post hoc* rmANOVAs were conducted on pairs of adjacent durations to determine the duration at which this interaction appeared. When the p-value threshold for significance was corrected using Bonferonni at the α level of 0.05 (0.05/4), the interaction between 2 and 6 s was significant [$F(1,29) = 13.47$, $p = 0.001$, $\eta_p^2 = 0.317$], but there were no significant interactions beyond 6 s, which make it impossible to statistically support the slight decrease that could be observed for falling tones.

Finally, the overall difference between the curves presented in each panel (i.e., the asymmetry between rising and falling tones) was observed to diminish at higher intensity-regions in group A, as revealed by a significant direction × intensity-region interaction [$F(3,87) = 5.94$, $p = 0.003$, $\eta_p^2 = 0.170$, $\tilde{\varepsilon} = 0.78$]. This interaction could also be observed in group B (see Fig. 2, bottom) but was not significant [$F(3,42) = 1.85$, $p = 0.160$, $\eta_p^2 = 0.117$, $\tilde{\varepsilon} = 0.90$].

C. Discussion

This first experiment examined how global loudness judgments of rising and falling ramps of a few seconds evolve when these sounds are stretched in time such that their dynamics is kept constant. Global loudness of both rising and falling sounds was found to increase with duration until about 6 s. For longer durations from 6 up to 16 s, global loudness of rising tones reached a constant plateau and a slight decrease was observed for falling tones. Significant or marginally significant direction×duration interactions were thus obtained between the patterns of each profile. It was, however, not possible to highlight the slight decrease of the curve observed for falling tones with *post hoc* analyses; only the interaction between 2 and 6 s was statistically reliable.

Overall, these results are different to those reported by Susini *et al.* (2007), where sequences of 1-kHz tones with time-varying intensity profiles were employed. In their study, the stimuli sequences were made of rising or falling ramps of different durations (from 2 to 20 s) having a fixed dynamics equal to 20 dB, combined with 3-s constant-intensity plateau presented either before the rising ramp or after falling ramp. Concerning (plateau-rising) sequences, they showed that global loudness increased by a fixed amount for each doubling of duration, whereas we found that the global loudness of simple rising ramps reached a constant value at 6 s and then remained constant for longer durations, at least until 16 s. Concerning (falling-plateau) sequences, they showed that global loudness did not vary significantly with duration, whereas in the present study, global loudness of falling tones increased significantly with duration until 6 s. Whether the differences between the present results and the results of Susini *et al.* (2007) can be attributed to the absence vs presence of a plateau before/after the ramp or whether they stem from procedural differences cannot be determined.

In the present experiment, we were expecting that the integration mechanism would increase global loudness of rising tones and its association with the decay mechanism would make global loudness of falling tones grow less rapidly. Our results only partially supported these hypotheses. Global loudness of rising tones indeed increased with duration, as if there was some kind of loudness integration, but beyond 6 s appeared to "saturate"; a result that cannot—at first sight—be explained by this integration mechanism. Indeed, if one considers the area contained in a fixed window located under the level profile of a linear ramp, this area should grow logarithmically as a function of the duration of the "time-stretching" until a certain point corresponding to the level of a constant-intensity sound. Concerning falling tones, global loudness also increased with duration, but at a slower rate, as revealed by a significant direction × duration interaction found between 2 and 6 s for group A. This might support our hypothesis that two mechanisms, an integration mechanism and a decay mechanism, might add up. However, our results thus indicate that the decay mechanism plays a significant role only between 2 s and 6 s (i.e., where the direction × duration interaction was found), whereas we were rather expecting a somewhat gradual effect as a function of time. We have no clear explanation for this result, but it might be possible that the integration mechanism has no noticeable effect if the slope is small.

It can also be noted that there were differences between the rating patterns in groups A and B. Although normalized in the same fashion, the ratings obtained in group B were overall 63% higher than those collected in group A, showing that observers of group B used higher numbers. Next, the size of the effects related to ramp duration was also different between the two groups (although they had overlapping loudness functions, on average). We do not have any clear explanation of these differences, but they might be related to the contextual differences, as ramp durations were higher in experiment 1B (4–16 s) than in experiment 1 A (1–12 s).

Another outcome of the present experiment concerns the asymmetry between rising and falling tones: greater global loudness judgments were obtained for rising ramps compared to falling ramps at all durations and, as expected, the size of this asymmetry increased with sound duration. This increase, which is assumed to be due to the decay mechanism was visible at all durations above 6–8 s, but it was only statistically significant between 2 and 6 s. Last, the asymmetry was found to depend on the intensity-region of the ramps in Group A. This decrease of the asymmetry in high intensity-regions was already observed in other studies on this topic (Ponsot *et al.*, 2015a; Ponsot *et al.*, 2015b), but its causes still remain undetermined.

The second experiment was designed to further assess the plausibility of the two proposed candidate mechanisms using the other experimental design presented in the introduction, i.e., using a time-stretching manipulation where the *slope* of the ramp was preserved. In addition to the ramp stimuli, we added constant-intensity tones presented at different levels and durations (corresponding to the maximum levels and the durations of the ramps), in order to compare their loudness with the global loudness of the ramps. This also allows us to control that listeners do not deviate[5] with duration in their loudness evaluations for sounds lasting several seconds.

III. EXPERIMENT 2

A. Materials and method

1. Participants

Twenty-nine subjects took part in this experiment (13 women, 16 men; age 19–34 years old). All reported normal hearing. They gave their informed written consent according to the Declaration of Helsinki prior to the experiment and were paid for their participation. The participants were naive with respect to the hypotheses under test.

2. Stimuli

All the stimuli were 1-kHz pure tones. As in experiment 1, 500-ms constant-intensity tones were used to measure the loudness function of each participant prior to the experiment (the same levels of presentation were used, equally spaced between 45 and 90 dB SPL). In the main part of the experiment, tones with constant, rising-, or falling-intensity profiles were employed. The constant-intensity tones were presented at four durations (2, 4, 6, and 12 s) and four levels ($M_1 = 75$, $M_2 = 80$, $M_3 = 85$, and $M_4 = 90$ dB SPL). Different combinations of duration and slope were used to create an appropriate set of rising- and falling-intensity tones. Their slope (i.e., absolute rate of change) was either 2.5 dB/s or 5 dB/s. The ramps varying at 2.5 dB/s were presented at four durations (2, 4, 6, and 12 s) and the ramps varying at 5 dB/s were presented at three durations only (2, 4, and 6 s) in order to avoid too low (start or end) levels that would have been induced if the 12-s duration had also been used. All the ramps were presented with four different maximum levels (75, 80, 85, 90 dB SPL). Their minimum levels and, consequently, their dynamics, resulted from the combinations of duration and slope.

3. Apparatus

The apparatus were the same as described in experiment 1.

4. Procedure

The procedure employed in this experiment was similar to the one described in experiment 1, i.e., an AME procedure with a blocked-duration design. After the preliminary loudness function measurement (similar to experiment 1), participants were presented with longer constant and ramp tones. Each block consisted of interleaved constant, rising, and falling ramps of equal duration presented at different levels, as mentioned in Sec. III A 2. Each stimulus was repeated three times. The order of presentation of the four blocks (one for each duration) was randomly varied between participants. A total of 48 constant-intensity tones (4 levels × 4 durations × 3 repetitions), 96 ramps varying at 2.5 dB/s (4 maximun levels × 2 directions × 4 durations × 3 repetitions) and 72 ramps varying at 5 dB/s (4 maximum levels × 2 directions × 3 durations × 3 repetitions) were thus presented to the participants.

B. Results

The same normalization as in experiment 1 was applied to the loudness ratings given by each listener. Different rmANOVAs were conducted to analyze the results in different ways because the ramps varying at 2.5 dB/s and those varying at 5 dB/s did not exactly share the same set of durations. Since these multiple analyses were planned prior to the experiment, uncorrected p-values are reported. To provide the reader a clear picture of the results obtained with these different analyses, the normalized loudness ratings are presented in Fig. 3 separately for ramps having a slope of 2.5 dB/s (top) and 5 dB/s (bottom), and in Fig. 4 for rising (top) and falling (bottom) ramps. Loudness estimates for constant-intensity tones, which had the same maximum levels and durations as the ramps, are superimposed in each panel of Figs. 3 and 4.

1. Analysis A: Loudness of constant tones

A first rmANOVA was conducted on the estimates given to constant tones only. A small increase of loudness estimates with duration until 6 s could be observed (see Fig. 3), but the analysis revealed that the effect of duration was not significant ($p > 0.05$). There was no significant duration×level interaction ($p > 0.05$). The constant tones are not taken into account in the analyses that follow, which focus on the effects related to rising and falling ramps. However, it is important to observe that the loudness of these constant tones was always greater or at

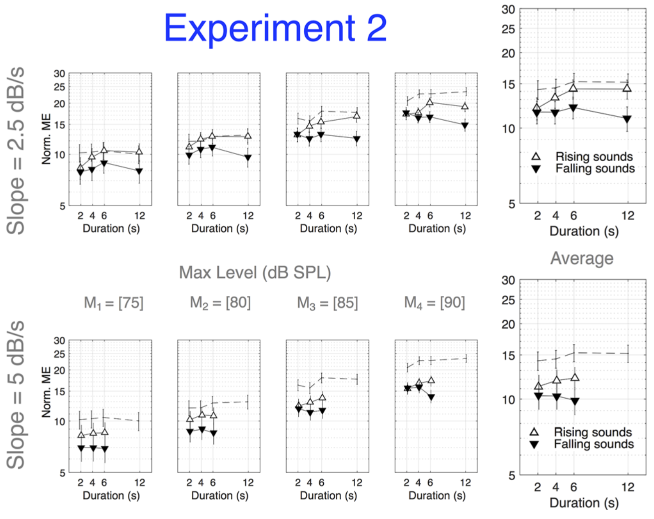


FIG. 3. (Color online) Normalized estimates of global loudness obtained in experiment 2 for rising (△) and falling ramps (▼) whose slope was equal to 2.5 dB/s (top) or 5 dB/s (bottom), plotted as a function of their duration. Same layout as in previous figures; the data are presented for the different maximum level of the ramps on the left panels (from M_1 to M_4), and after an averaging over these different levels on the rightmost panels. The loudness estimates obtained for constant-intensity tones having the same level as the maximum level of the ramp are superimposed in each panel (dashed lines). Error bars correspond to SEM.

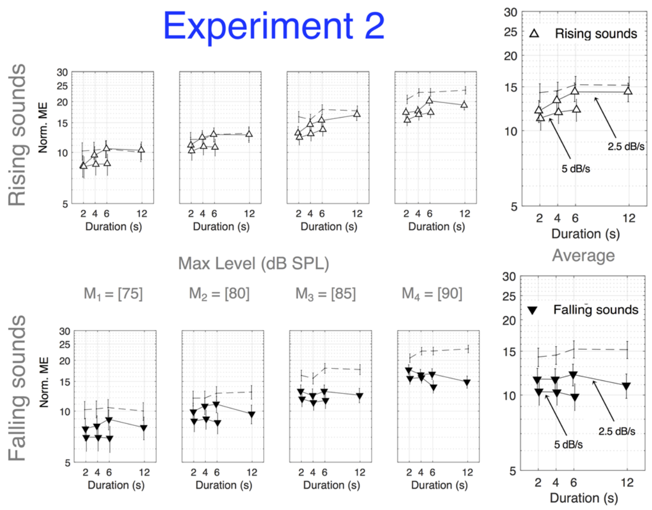


FIG. 4. (Color online) Normalized estimates of global loudness collected in experiment 2 for rising sounds (upper panels) and falling sounds (lower panels). This figure presents, in a different way, the results plotted in Fig. 3 to provide a clearer picture of the influence of the slope on rising and falling ramps separately, and should be seen as a visual support to analysis D. On each panel, highest triangles correspond to the estimates given to 2.5 dB/s ramps (Fig. 3, higher panels) or to 5 dB/s ramps (Fig. 3, lower panels). Otherwise, the plotting convention is the same as in Fig. 3.

least equal to the global loudness estimates given to ramps with the same maximum level (see Fig. 3).

2. Analysis B: Global loudness of rising and falling ramps varying at 2.5 dB/s

A second analysis was performed to specifically compare rising and falling ramps varying at 2.5 dB/s, for which global loudness estimates are presented on the top of Fig. 3. Greater estimates were overall obtained for rising ramps compared to falling ramps, as supported by a significant effect of the direction [$F(1,28) = 11.39$, $p = 0.002$, $\eta_p^2 = 0.289$]. However, as it can be observed in Fig. 3, the size of this difference appeared to increase with the duration because the curves took somewhat different directions. Global loudness judgments of rising tones slightly increased with duration until 6 s and then reached a plateau, whereas global loudness judgments of falling tones remained fairly constant with duration and even appeared to decrease slightly between 6 and 12 s. This was supported by the fact that there was no significant main effect of the duration ($p > 0.05$) but a significant direction × duration interaction [$F(3,84) = 3.94$, $p = 0.034$, $\eta_p^2 = 0.123$, $\tilde{\varepsilon} = 0.54$]. All these effects appeared to be similar at the different maximum levels tested, as supported by no significant interactions between the maximum level of the ramps and other factors ($p > 0.05$). We conducted *post hoc* tests to determine whether the changes observed with duration for each profile separately (rising/falling) were significant or not. We found no significant effects of duration neither for rising tones, nor for falling tones ($p > 0.05$).

3. Analysis C: Global loudness of rising and falling ramps varying at 5 dB/s

This analysis was concerned with global loudness estimates of ramps varying at 5 dB/s. These data are presented in Fig. 3 (bottom). Overall, similar conclusions to those obtained with ramps varying at 2.5 dB/s were reached. A significant effect of the direction was found [$F(1,28) = 10.92$, $p = 0.003$, $\eta_p^2 = 0.281$]. As on the top of Fig. 3, we could observe a slight increase of the judgments of rising tones as a function of duration, and a slight decrease of the judgments of falling tones with duration. However, neither the effect of the duration nor the duration × direction interaction were significant ($p > 0.05$). Last, the overall difference between rising and falling tones estimates was slightly decreased as the maximum level of ramp increased; there was a significant direction × maximum level interaction [$F(3,84) = 3.14$, $p = 0.031$, $\eta_p^2 = 0.101$, $\tilde{\varepsilon} = 0.96$].

4. Analysis D: Global loudness of rising and falling ramps separately

In order to specifically examine the influence of the duration on each ramp direction and assess the effect of the slope on their global loudness judgments, we conducted additional analyses on rising and falling ramps separately. Ramps varying at 2.5 dB/s and 5 dB/s were combined for the set of durations shared between these two groups, i.e., 2, 4, and 6 s (the 12-s ramps varying at 2.5 dB/s were, thus, not considered in these analyses). The data separated for rising and falling ramps are presented in Fig. 4 (upper panels, rising tones; lower panels, falling tones).

First, we analyzed global loudness judgments of rising ramps lasting 2, 4, and 6 s and varying at 2.5 dB/s or 5 dB/s. Global loudness estimates of rising ramps varying at 2.5 dB/s were clearly higher than those given to rising ramps having the same maximum level and duration but varying at 5 dB/s; a large and significant effect of the slope supported this observation [$F(1,28) = 31.92$, $p < 0.001$, $\eta_p^2 = 0.533$]. There was no significant effect of the duration ($p > 0.05$), but a significant duration × slope interaction [$F(2,56) = 4.57$, $p = 0.015$, $\eta_p^2 = 0.140$, $\tilde{\varepsilon} = 1.06$], revealing that the effect of the duration, although not significant as a main factor, was different for the two slopes.

Second, we examined the judgments for falling ramps having the same parameters. Global loudness estimates of falling ramps varying at 2.5 dB/s were again clearly higher than those given to falling ramps having the same maximum level and duration but which varied at 5 dB/s, as supported by a large and significant effect of the slope [$F(1,28) = 5.94$, $p < 0.001$, $\eta_p^2 = 0.680$]. There was no significant effect of duration ($p > 0.05$) but again, a significant duration × slope interaction [$F(2,56) = 4.38$, $p = 0.024$, $\eta_p^2 = 0.135$, $\tilde{\varepsilon} = 0.83$].

C. Discussion

Analyses B and C showed that stretching ramps in time while maintaining their slope unchanged had no noticeable influence on global loudness: The effect of duration was not significant. Therefore, as compared to when the time-stretching was made at constant *dynamics* (experiment 1), where a large main effect of the duration was found, a time-stretching at constant *slope* did not strongly affect global loudness. As discussed in the introduction, this result is compatible with the integration mechanism. Moreover, analysis D showed that the ramps varying at 2.5 dB/s were perceived clearly louder than those varying at 5 dB/s, a result that also appears to be consistent, at least at a qualitative level, with the integration mechanism, given that more energy is contained in a similar integration window for ramps varying at 2.5 dB/s as compared to ramps varying at 5 dB/s. The fact that global loudness of ramps was always below or equal to the loudness of constant tones presented at their maximum level also provides support to the integration mechanism.

Although the effect of duration was not significant overall, a slight increase in the judgments of rising tones and a slight decrease in the judgments of falling tones as a function of duration could be noticed, respectively. This duration × direction interaction reached significance only for ramps varying at 2.5 dB/s. It cannot be excluded that other phenomena are operating when the ramps are stretched in time at constant slope. In particular, the significant duration × slope interactions revealed by analysis D suggests that the effect of the stretching is not the same whether the ramps vary at 2.5 or 5 dB/s, a result that it is not possible to account for by the two mechanisms considered in this study.

Overall, these data are not incompatible with our hypothesis of a decay mechanism involved in global loudness evaluations of long falling ramps, but we were not able to statistically support it. The fact that the interindividual variability was large and the range of tested durations restricted by the current design is undoubtedly a limiting factor when trying to capture a small and slow-acting effect as the one considered here.

IV. GENERAL DISCUSSION AND CONCLUSION

A. Summary of experimental findings

In two magnitude estimation experiments, we addressed whether two mechanisms, which had been proposed as potential candidates in previous studies, might indeed underlie the perceptual computation of global loudness for rising and falling ramps. The first mechanism under study, called the integration mechanism, relies on the assumption that the global loudness of a ramp could be determined by an integration of its loudest portion over a certain temporal window. The second mechanism considered here, called the decay mechanism, is based on the assumption that the output of this loudness integration is weakened as a function of the time lapse between the beginning and the end of a falling ramp, this mechanism being specific to falling ramps.

The plausibility of these two mechanisms was determined by looking at the extent to which global loudness judgments of rising and falling tones (with linear level changes) were influenced by different manipulations of their parameters, namely, their *slope*, their *dynamics*, and their *duration*. It should be noted that disentangling perceptual mechanisms in play with such stimuli by "stretching" their parameters (slope, duration, and dynamics) is a complex task because these parameters are not independent; the slope is indeed equal to the dynamics divided by the duration. In both experiments, several parameters varied simultaneously and might, thus, have tainted the results such that it is not possible to make "clear-cut" conclusions. However, the two experiments yield various results that allow us to further discuss the plausibility of the two proposed mechanisms.

Overall, the results obtained in this study provide significant support to the hypothesis that an integration mechanism might be involved. First, the stretching adopted in experiment 1 (i.e., stretch in duration keeping ramps' dynamics unchanged) caused global loudness increase with duration for both rising and falling ramps, which is consistent with the fact that the energy contained within a fixed temporal window located around the peak stimulus is increased. Second, the stretching adopted in experiment 2 (i.e., stretch in duration keeping ramps' slope unchanged) led to (1) non-significant effects with respect to duration, consistent with the fact that the energy contained within a fixed temporal window remains unchanged, and (2) a large and significant effect of the slope in experiment 2 both for rising and falling tones (ramps varying at 2.5 dB/s were significantly louder than ramps varying at 5 dB/s), consistent with the fact that less energy is contained in a window located under ramps having steeper slopes. There are, however, two departures from this mechanism that can be noticed: (1) the "saturation" of the estimates of rising tones beyond 6 s observed in experiment 1 and (2) the slope × duration interaction obtained in experiment 2. As discussed earlier, since the stretching adopted in experiment 1 induced both variations in duration and slope at the same time, it might be possible that another mechanism was involved in the evaluation of the ramps of long durations and that this mechanism was responsible for the "saturation" in the judgments beyond 6 s; the ramps had very small slope and could possibly be assimilated to constant tones. In that sense, the saturation would finally not be imputable to the integration mechanism itself. Note that we only examined *qualitatively* the extent to which our results agreed with an integration mechanism; all these results remain to be verified quantitatively, for instance, whether loudness integration over a fixed temporal window of which length is compatible with the rate of increase observed in experiment 1. The presence of the decay mechanism was assessed by comparing the estimates obtained for falling tones with those of rising sounds (for which only integration is involved). We found small but significant duration × direction interactions both in experiment 1 and experiment 2 (only for the ramps varying at 2.5 dB/s), consistent with our assumption that a decay mechanism might play a role. The estimates collected for falling ramps in experiment 2 presented a small decline as a function of duration after 6 s but the ramps varying at 2.5 dB/s also showed an increase between 2 and 6 s; there was no significant effect of duration except a significant duration × slope interaction. As a result, the data collected in the present study are not incompatible with the idea that a certain decay mechanism might underlie the processing of falling ramps of long durations, but its implication could not be statistically supported. More specifically, the data of experiment 2 suggest that the underlying machinery is not solely composed of the two mechanisms considered here,[6] but that other mechanisms might be acting or interacting with those, e.g., by modulating the decay mechanism as a function of the absolute slope of the ramps.

Last, global loudness judgments of ramps appear to be, in addition to their maximum level, primarily guided by their slope (strong effects of the slope in analysis D, which are obvious in Fig. 4), much more than by their duration or dynamics. Besides, the size of the asymmetry between rising and falling ramps depends on the duration of the ramps (cf. experiment 1, where asymmetries were increased with duration) and on their dynamics (cf. experiment 2, where the asymmetries were reduced for small dynamics). Therefore, the present study shows that the asymmetry between rising and falling tones is not specific to the 2-s, 15-dB ramps employed in previous studies (e.g., Ponsot *et al.*, 2015a; Ponsot *et al.*, 2015b); it occurs in many other conditions, but its magnitude depends on the parameters of the ramps.

B. An attempt to predict global loudness directly from the model of Glasberg and Moore

We evaluated, in particular, the extent to which the loudness model of Glasberg and Moore (2002) could account for the results exposed in experiment 1, where the time-stretching manipulation was assumed to trigger both the integration and the decay mechanisms. Three basic indicators directly based on STL and LTL time series of the model outputs were examined to see how well they could reproduce the patterns of observers' ratings. The first two indicators considered were the maxima of STL and LTL patterns, i.e., STL_{max} and LTL_{max}, respectively, which we already examined (Ponsot *et al.*, 2015a; Ponsot *et al.*, 2015b). The third indicator considered was inspired from the integration mechanism hypothesis: We introduced STL_{int}, which corresponds to the *average* of STL over a fixed temporal window located around its maximum. Since we had no specific assumptions concerning the shape of this temporal window, we used a simple rectangular temporal window whose length was arbitrarily chosen equal to 500 ms (in order to roughly account for the growth of global loudness estimates with duration obtained in experiment 1, group A). This window was located so that it started (ended) at the maximum of the STL pattern of falling (rising) tones. The way these three indicators are computed from STL and LTL outputs, respectively, is illustrated on the left part of Fig. 5. The global loudness predictions are presented for various ramp durations, ranging from 1 to 16 s, in the right panels of Fig. 5.

The rising vs falling asymmetries given by the indicators proposed to predict global loudness so far, i.e., the maximum of STL or LTL patterns, considerably underestimate what was measured by means of various psychophysical experiments. STL_{max} and LTL_{max} produce only small asymmetries in the desired direction at short durations[7] (i.e., rising louder than falling ramps); however, such asymmetries disappear at longer ramp durations because the influence of the temporal integration stages is weakened as ramp duration increases. The growth with duration obtained with STL_{max} is negligible, whereas, as expected, a substantial and logarithmic increase can be observed with STL_{int}. However, STL_{int} does not predict any asymmetries between rising and falling tones.

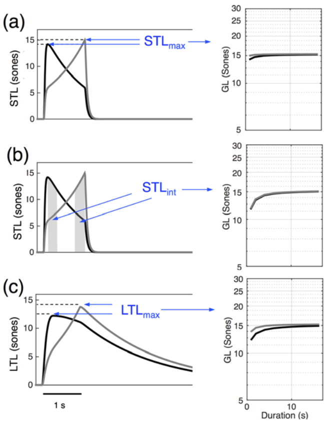


FIG. 5. (Color online) Different indicators introduced to estimate global loudness from the outputs of the model of Glasberg and Moore (2002). Both STL and LTL time series of 1-s ([65–80 dB SPL) rising-intensity (grey lines) and falling-intensity ramps (black lines) predicted by the model are considered. (a) STL_{max}: using the maximum of the STL. (b) STL_{int}: using an average of STL over a fixed arbitrary 500-ms rectangular integration window (note that the temporal windows are not to scale for clarity purpose) starting (ending) at the maximum value of the falling (rising) pattern. (c) LTL_{max}: using the maximum of the LTL. On the right part of the figure, global loudness predictions based on these different indicators for rising (grey) and falling (black) ramps of durations ranging from 1 to 16 s are shown.

STL_{int} probably provides the best reproduction of the main trend observed as a function of ramp duration in experiment 1 (group A): the increase between 1 and 2 s is comparable to what was measured and the logarithmic growth well approximates the fact that "saturation" was observed experimentally beyond 6 s. It is clear that STL_{max} is not appropriate to account for the increase with duration. The integration mechanism we are examining here is better accounted for by LTL_{max}, which shows an increase that would fit reasonably well the data for rising sounds; but, it does not do the job for falling sounds. With respect to the rising vs falling asymmetry, none of these indicators are, however, able to account for the magnitude of the effect observed experimentally.

This brief investigation with three simple indicators derived from the outputs of the model of Glasberg and Moore (2002) shows that none of them is able to account for the data collected in experiment 1, group A. The same outcomes would have been reached if one would have considered the data of experiment 1 (group B) or experiment 2, and also using the dynamic loudness model of Chalupper and Fastl (2002; see Ponsot *et al.*, 2015a). The predictions derived from these indicators show that a subsequent temporal integration stage induces the desired growth with duration observed experimentally. STL_{int} computed with a 500-ms constant is able to produce an increase with duration similar to what we obtained with group A. Note that this 500-ms value is about an order of magnitude higher than the time constant involved greater than "traditional" experiments on the temporal integration of loudness. This supports our hypothesis that the integration mechanism examined here operates at a much coarser, likely cognitive, time scale.

However, while STL_{int} could indeed account for the integration phenomenon, it cancels at the same time any loudness difference between up- and down-ramps produced by the automatic gain control (AGC), so that it is not possible to account both for the integration phenomenon *and* the asymmetry at the same time using STL. LTL (which AGC uses a time constant of 200 ms) predicts too small of an increase of global loudness with duration, but is able to amplify the asymmetries between rising and falling STL patterns. Motivated by the hypotheses of the mechanisms examined in this paper, we tested various modifications of the TVL model (Glasberg and Moore, 2002) for time-varying sounds that would best fit our data and identified two minimal changes. It appears necessary to (1) significantly increase the time-constant of the AGCs to mimic the integration phenomenon while reinforcing the asymmetry produced by the model at short ramp durations (such as LTL does, to a certain extent, but not enough), and (2) add another decay-like stage, not only to produce the small decline observed at longest durations for falling ramps but, most critically, to maintain the asymmetry at longer ramp durations. Regarding the first point, we observed (not detailed here) that it was necessary to increase the time constants of the AGCs to reproduce the growth with duration obtained experimentally at short durations, but that different values were required to fit with our different experimental conditions (experiment 1 A, experiment 1B, experiment 2—2.5 dB/s, experiment 2—5 dB/s). Regarding the second point, it is of note that a decay-like stage is required to downsize the increase of the falling pattern that would otherwise superimpose with the rising pattern at long ramp durations with any integration stage. Indeed, if the asymmetry is solely driven by the AGC stages, it will inevitably decreased toward zero as the duration of the ramp increases because the slope of the ramps will thus approach zero (i.e., the rising and falling patterns will become indistinguishable). Taken together, these computational analyses support our working hypothesis that the temporal integration and the decay mechanisms considered here do underlie the global loudness evaluation of rising and falling ramps. However, it is in our view too preliminary to propose any analytical expression of these stages before understanding the reasons why different time constants would be needed in different experimental contexts.

C. Conclusion and perspectives

The data presented in this paper provide the first direct experimental investigation of the hypotheses raised in previous studies, that an integration mechanism and a decay mechanism might be involved in the global loudness evaluation process of rising and falling intensity tones. While these results qualitatively support the view that global loudness of intensity-ramps might partly be accounted for by a certain integration of their loudest portions, the presence of the decay mechanism could not be demonstrated. Further studies have to be conducted to directly tackle its implication with other experimental paradigms.

On a more practical basis, we computed independent indexes to assess the importance of the effects highlighted in the different experimental configuration of this study [experiment 1 A, experiment 1B, experiment 2 (ramps 2.5 dB/s), experiment 2 (ramps 5 dB/s)], in order to help the readers judge which of the reported effects could be relevant for the loudness of everyday sounds. These indexes (shown in Fig. 6) intend to reflect (a) the effects induced by a 5- and 10-dB change in ramp's intensity-region (or mean level), (b) the averaged asymmetry between up- and down-ramps, and (c) the largest change in loudness caused by a variation in duration. To allow linear comparisons, we took the log-value of the ratio between the ratings given, respectively, to ramps varying in intensity-regions 5-dB and 10-dB apart (black bars), to up- and down-ramps (red bars), and to ramps at the two durations that received most different ratings (blue bars). Note that these indexes make use of the same data so they are not independent; they should simply be taken as rough, first-order estimates. On average, this analysis shows that the size of the asymmetry between up- and down-ramps is comparable (or slightly smaller) to an increase of 5-dB in ramp level, consistent with previous studies (Ponsot *et al.*, 2015a; Ponsot *et al.*, 2015b). It can also be observed that the change in loudness caused by ramp duration depends on the context: In experiment 1 A, the effect laid between an effect caused by a 5 - and a 10-dB level increase; in experiment 1B, the effect was similar to a 5-dB increase in level; in experiment 2 (ramps varying at 2.5 dB/s), the effect was nearly two times smaller than a 5-dB increase; in experiment 2 (ramps varying at 5 dB/s), the effect was close to zero. These results show how the loudness of sounds ramping in level is affected by their parameters (direction of level change, duration), and which is worth taking into account when assessing the loudness of everyday sounds (many natural sounds have very similar level profiles, e.g., a sound source passing by). In most cases, the effects caused by these parametric changes are not small and, thus, have to be considered; most were comparable to a ∼5-dB increase in sound level.

The analyses of different indicators based on the outputs of the model of Glasberg and Moore (2002) show that the indicators most often used to predict global loudness, namely, STL$_{\text{max}}$, LTL$_{\text{max}}$, are not able to reproduce most of the trends we observed experimentally. In particular, the increase with duration obtained with the time stretching at constant dynamics employed in experiment 1 cannot directly be predicted by taking the maximum of STL or LTL time series provided by the model. We showed that STL$_{\text{int}}$ can fit,

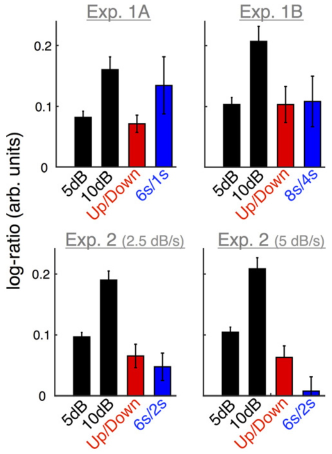


FIG. 6. (Color online) Mean values of the indexes reflecting the importance of the factors manipulated in this study, computed across the four main experimental conditions (different panels). For each index, all the factors (except the one the index was based on) were pooled together. Error bars show SEM across subjects.

overall, the increase obtained in experiment 1, indicating that global loudness could be compatible with a certain integration of STL over a window, which was around 500-ms in that case (i.e., experiment 1, group A). However, further analyses have to be undertaken to confirm that the integration mechanism is indeed involved, and if this is the case, to determine the shape and the length of such a temporal window, and to what extent it depends on the experimental context and conditions. Indeed, our results suggest that the size of this integration window would depend on the ramp parameters or experimental context. Nevertheless, none of the three indicators considered were able to account for the asymmetries observed between rising and falling tones and, consequently, for the fact this asymmetry depends on ramp parameters such as its dynamics. Therefore, even though an integration mechanism could be part of the global loudness evaluation process of rising and falling ramps, there are certainly other mechanisms involved that still remain to be determined.

According to Moore (2014), LTL is supposed to reflect "relatively high-level cortical processes and involves memory," although the location at which LTL is represented in the brain has not been examined yet (for a discussion, see Thwaites *et al.*, 2016). LTL computation involves another AGC with a long decreasing time constant aiming to reflect the fact that the overall impression "can persist for several seconds after a sound has ceased" while gently decaying when the sound is turned off. Here, we show that this is not the case and that the computation of global loudness from STL or LTL patterns is more complex than just considering their maximum value. It is very likely that the processes involved in global loudness evaluation of rising and falling tones of a few seconds are presumably part of high-level integration stages not yet reflected by LTL, which is simply based on a temporal integration of STL (reflecting the loudness consciously accessible at any instant). Further psychophysical studies have to be conducted with more complex time-varying sounds before a model covering the whole set of high-level processes underlying the computation of global loudness could be implemented.

Are such mechanisms specific to loudness evaluation? There are some studies in the literature indicating that the processes examined here in the case of global loudness evaluation might actually be involved in overall judgments of other types of sensory information. For example, although the time scale and the amount of perceptual variation are not the same, the results obtained for the overall evaluation of increasing and decreasing sequences of pain yielded similar trends to those observed in the present study (Ariely and Carmon, 2000). In particular, increasing sequences are judged as more painful that decreasing sequences and the slopes of the sequences play a significant role. Works addressing overall annoyance evaluation of aversive sounds (e.g., Kahneman *et al.*, 1993; Västfjäll, 2004) or the overall evaluation of image quality (Hamberg and de Ridder, 1999) have brought to light important principles governing overall evaluation, such as "peak-end" rules and recency effects, which are often observed in global loudness judgments of time-varying sounds (e.g., Dittrich and Oberfeld, 2009). Could an integration mechanism be the basis of any overall evaluation of every rising or falling pattern? Whether the mechanisms underlying global loudness evaluation are also involved in the overall evaluations of other sensory attributes is an aspect that deserves to be specifically addressed in future studies. It could be particularly fruitful to take a step back from the specific case of rising and falling level patterns, and investigate the processing of more complex contours specifically dedicated to test the decay and the integration mechanisms. One could think of using contours that increase to a fixed amplitude (or have multiple maximums) and finish at different values so that the maximum does not occur at the end. One could also vary the position of the maximum in the contour. If the integration mechanism applies, sounds of various contours but identical maxima should receive similar loudness judgments. In a similar fashion, one might test the decay mechanism using V-shaped level contours so that the minimum does not occur at the end. Such investigations should undoubtedly provide important elements for a better understanding of the machinery underlying global loudness processing in a general context.

As pointed out by Ariely (1998): "[…] *although this work examines only one domain of experience (namely pain), one can speculate that the relationship between momentary and overall evaluations will apply to other domains as well*." Reinforcing our knowledge of psychoacoustics with the investigation of higher-level integration mechanisms related to general principles of time-varying information processing would provide important information concerning the mechanisms underlying overall evaluation of dynamic loudness and, more generally, the mechanisms underlying overall evaluation of dynamic sensory events over long time scales.

ACKNOWLEDGMENTS

The authors would like to thank Anne-Laure Verneil for her help in conducting the experiments. We are particularly grateful to Robert Teghtsoonian for his comments on an earlier version of this document. We would also like to thank Daniel Oberfeld and Stephen Handel for insightful comments when reviewing this paper. This work was supported by the project *LoudNat* funded by the French National Research Agency (Grant No. ANR-11-BS09-016-01). Some of the results of this paper were presented at the 30th Meeting of the International Society for Psychophysics, *Fechner Day*, in Lund, Sweden, August 2014.

---

[1] One might ask whether taking the average values of STL or LTL outputs is a better option; this is not the case as this would return even poorer estimates because the averaging operation would dilute the small effects produced by the asymmetric temporal integration stages (see the discussion in Ponsot, 2015).

[2] One may ask why this decay mechanism would affect falling ramps stronger than rising ramps. One might compare the decay mechanism to the process behind time-order errors (TOEs), i.e., the fact that two stimuli compared in a pair receive different weightings (e.g., Hellström, 1985). The idea behind TOEs in successive paired comparison is that the second stimulus is compared with the trace or the memory image of the first one. This trace is inherently fading or disintegrating over time, a phenomenon that is accounted for by, for example, the sensation weighting model (Hellström, 1985). In practice, as the time interval between the two stimuli increases, the weight attributed to the first stimulus in the judgment decreases given that the uncertainty of its trace increases (Hellström and Rammsayer, 2004). Recent results on temporal loudness weighting of rising and falling ramps showed that observers exclusively focus on the loudest portions of the sounds for judging their global loudness, i.e., on the end of rising ramps and the start falling ramps

(Ponsot *et al.*, 2013). In this context, it is reasonable to assume that the disruption of the trace is weaker for up-ramps than for down-ramps: While no sensory stimulation separates the end of a rising ramp from its judgment, the trace of the starting portion of a falling ramp is disrupted by its continuing decreasing level. As a result, TOEs (or such related effects) are likely affecting more the judgment of a falling ramp (i) because we integrate a portion located further in time, and (ii) because this integration is disrupted by the following part of the stimulus.

[3] This parameter manipulation was inspired from studies examining the influence of these factors to investigate the mechanisms underlying *loudness change* judgments (Canévet *et al.*, 2003; Teghtsoonian *et al.*, 2005).

[4] All the ratings given by a listener to both constant tones and ramps were divided by the mean rating assigned to the 60-dB SPL constant-intensity tone and multiplied by four. As in previous papers (Susini *et al.*, 2010; Ponsot *et al.*, 2015a), this normalization intended to match the estimate attributed to the 60-dB SPL, 1-kHz pure-tone to a value of 4 sones. However, magnitude estimates normalized in that way are not directly interpretable as sones (according to the standard sone definition) because the sound level corresponds to the one of a monaural source presented at the input of the ear canal, not to a sound source frontally presented in free field.

[5] The durations employed in this experiment being equal or greater than 1 s, we are well beyond the durations at which "temporal integration of loudness" occurs, which is generally assumed to be fully completed at 300 ms (see Rennies *et al.*, 2010; Hots *et al.*, 2014). As a result, the loudness of constant tones should not be affected by any change of their duration.

[6] It is important to note that the analyses and discussions of the present paper are based on aggregate data. Because substantial interindividual differences were observed in the different experiments (not detailed in this paper), it remains to explore to what extent individual behaviors can be captured by the mechanisms inferred from the "aggregate observer." Indeed, even if the mean trends captured in these experiments comply to some extent to what one would expect from a given mechanism (e.g., an integration mechanism), this does not yet constitute a "proof" that observers were indeed behaving according to this mechanism. Further studies are necessary to demonstrate that the mechanisms indeed reflect the processing of every observer.

[7] The asymmetries produced by the model are the consequence of the two temporal integration stages employed to derive, first, STL and, second, LTL from the instantaneous loudness (IL) pattern (see Ponsot *et al.*, 2015a).

Ariely, D. (**1998**). "Combining experiences over time: The effects of duration, intensity changes and on-line measurements on retrospective pain evaluations," J. Behav. Decis. Making **11**(1), 19–45.

Ariely, D., and Carmon, Z. (**2000**). "Gestalt characteristics of experiences: The defining features of summarized events," J. Behav. Decis. Making **13**, 191–201.

Buus, S., Florentine, M., and Poulsen, T. (**1997**). "Temporal integration of loudness, loudness discrimination, and the form of the loudness function," J. Acoust. Soc. Am. **101**(2), 669–680.

Canévet, G., Teghtsoonian, R., and Teghtsoonian, M. (**2003**). "A comparison of loudness change in signals that continuously rise or fall in amplitude," Acta Acust. Acust. **89**(2), 339–345.

Chalupper, J., and Fastl, H. (**2002**). "Dynamic loudness model (DLM) for normal and hearing-impaired listeners," Acta Acust. Acust. **88**(3), 378–386.

Cross, D. V. (**1973**). "Sequential dependencies and regression in psychophysical judgments," Percept. Psychophys. **14**(3), 547–552.

Dittrich, K., and Oberfeld, D. (**2009**). "A comparison of the temporal weighting of annoyance and loudness," J. Acoust. Soc. Am. **126**(6), 3168–3178.

EBU–Recommendation R 128 (**2011**). Loudness normalisation and permitted maximum level of audio signals (European Broadcasting Union, Geneva, Switzerland).

Glasberg, B. R., and Moore, B. C. (**2002**). "A model of loudness applicable to time-varying sounds," J. Audio Eng. Soc. **50**(5), 331–342.

Gottschling, G. (**1999**). "On the relations of instantaneous and overall loudness," Acta Acust. Acust. **85**(3), 427–429.

Hamberg, R., and de Ridder, H. (**1999**). "Time-varying image quality: Modeling the relation between instantaneous and overall quality," SMPTE J. **108**(11), 802–811.

Hellbrück, J. (**2000**). "Memory effects in loudness scaling of traffic noise. How overall loudness of short-term and long-term sounds depends on memory," J. Acoust. Soc. Jpn. (E) **21**(6), 329–332.

Hellman, R. P. (**1982**). "Loudness, annoyance, and noisiness produced by single-tone-noise complexes," J. Acoust. Soc. Am. **72**(1), 62–73.

Hellström, Å. (**1985**). "The time-order error and its relatives: Mirrors of cognitive processes in comparing," Psychol. Bull. **97**(1), 35–61.

Hellström, Å., and Rammsayer, T. H. (**2004**). "Effects of time-order, interstimulus interval, and feedback in duration discrimination of noise bursts in the 50- and 1000-ms ranges," Acta Psychol. **116**(1), 1–20.

Hots, J., Rennies, J., and Verhey, J. L. (**2014**). "Modeling temporal integration of loudness," Acta Acust. Acust. **100**(1), 184–187.

ITU-R BS.1770. (**2006**). "Algorithms to measure audio programme loudness and true-peak audio level," International Telecommunication Union (Geneva, Switzerland), available online at https://www.itu.int/rec/R-REC-BS.1770/fr.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., and Redelmeier, D. A. (**1993**). "When more pain is preferred to less: Adding a better end," Psychol. Sci. **4**(6), 401–405.

Kuwano, S., and Namba, S. (**1985**). "Continuous judgment of level-fluctuating sounds and the relationship between overall loudness and instantaneous loudness," Psychol. Res. **47**(1), 27–37.

Kuwano, S., Namba, S., Kato, T., and Hellbrück, J. (**2003**). "Memory of the loudness of sounds in relation to overall impression," Acoust. Sci. Technol. **24**(4), 194–196.

Meunier, S., Susini, P., Trapeau, R., and Chatron, J. (**2010**). "Global loudness of ramped and damped sounds," in *20th International Congress on Acoustics*.

Moore, B. C. (**2014**). "Development and current status of the 'Cambridge' loudness models," Trends Hear. **18**, 2331216514550620.

Moore, B. C., Glasberg, B. R., Varathanathan, A., and Schlittenlacher, J. (**2016**). "A loudness model for time-varying sounds incorporating binaural inhibition," Trends Hear. **20**, 2331216516682698.

Neuhoff, J. G. (**1998**). "Perceptual bias for rising tones," Nature **395**(6698), 123–124.

Oberfeld, D., and Plank, T. (**2011**). "The temporal weighting of loudness: Effects of the level profile," Atten. Percept. Psychophys. **73**(1), 189–208.

Olsen, K. N. (**2014**). "Intensity dynamics and loudness change: A review of methods and perceptual processes," Acoust. Aust. **42**(3), 159–165.

Ponsot, E. (**2015**). "Global loudness processing of time-varying sounds," Doctoral dissertation, UPMC, p. 39.

Ponsot, E., Meunier, S., Kacem, A., Chatron, J., and Susini, P. (**2015b**)." Are rising sounds always louder? Influences of spectral structure and intensity-region on loudness sensitivity to intensity-change direction," Acta Acust. Acust. **101**, 1083–1093.

Ponsot, E., Susini, P., and Meunier, S. (**2015a**). "A robust asymmetry in loudness between rising- and falling-intensity tones," Atten. Percept. Psychophys. **77**(3), 907–920.

Ponsot, E., Susini, P., and Meunier, S. (**2016**). "Loudness processing of time-varying sounds: Recent advances in psychophysics and challenges for future research," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Vol. 253, No. 2, pp. 6437–6442.

Ponsot, E., Susini, P., Saint Pierre, G., and Meunier, S. (**2013**). "Temporal loudness weights for sounds with increasing and decreasing intensity profiles," J. Acoust. Soc. Am. **134**(4), EL321–EL326.

R Core Team (**2015**). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, available at http://www.R-project.org/ (Last viewed July 6, 2017).

Rennies, J., Verhey, J. L., and Fastl, H. (**2010**). "Comparison of loudness models for time-varying sounds," Acta Acust. Acust. **96**(2), 383–396.

Ries, D. T., Schlauch, R. S., and DiGiovanni, J. J. (**2008**). "The role of temporal-masking patterns in the determination of subjective duration and loudness for ramped and damped sounds," J. Acoust. Soc. Am. **124**(6), 3772–3783.

Schreiber, C. A., and Kahneman, D. (**2000**). "Determinants of the remembered utility of aversive sounds," J. Exp. Psychol.: General **129**(1), 27–42.

Susini, P., McAdams, S., and Smith, B. K. (**2002**). "Global and continuous loudness estimation of time-varying levels," Acta Acust. Acust. **88**(4), 536–548.

Susini, P., McAdams, S., and Smith, B. K. (**2007**). "Loudness asymmetries for tones with increasing and decreasing levels using continuous and global ratings," Acta Acust. Acust. **93**(4), 623–631.

Susini, P., Meunier, S., Trapeau, R., and Chatron, J. (**2010**). "End level bias on direct loudness ratings of increasing sounds," J. Acoust. Soc. Am. **128**(4), EL163–EL168.

Teghtsoonian, R., Teghtsoonian, M., and Canévet, G. (**2005**). "Sweep-induced acceleration in loudness change and the 'bias for rising intensities,'" Percept. Psychophys. **67**(4), 699–712.

Thwaites, A., Glasberg, B. R., Nimmo-Smith, I., Marslen-Wilson, W. D., and Moore, B. C. (**2016**). "Representation of instantaneous and short-term loudness in the human cortex," Front. Neurosci. **10**, 183.

Västfjäll, D. (**2004**). "The 'end effect' in retrospective sound quality evaluation," Acoust. Sci. Technol. **25**(2), 170–172.

Zwicker, E., and Fastl, H. (**1999**). *Psychoacoustics: Facts and Models* (Springer Science and Business Media, Berlin), Vol. 22.