ACOUSTICS 2012

Perceptual asymmetry in the subjective duration of ramped and damped sounds

M. Vannier[a], S. Meunier[a], P. Susini[b] and J. Chatron[a]

[a]Laboratoire de Mécanique et d'Acoustique, 31, Chemin Joseph Aiguier - 13402 Marseille Cedex 20
[b]Institut de Recherche et Coordination Acoustique/Musique, 1, place Igor Stravinsky 75004 Paris
ml.vannier@gmail.com

Time-varying-level sounds that increase or decrease in level are well established to induce auditory perceptual asymmetries, for loudness and subjective duration. Several studies revealed ramped sounds to be perceived louder than equivalent damped sounds using durations from few milliseconds to few seconds. In addition, other studies revealed ramped sounds to be perceived longer than damped sounds for durations from 10 ms to 500 ms. As a consequence, it could be hypothesized that the perceived duration asymmetry may be responsible for the loudness asymmetry. Thus, the aim of the present study was to extend the results about asymmetries in subjective duration for tones longer than 500 ms, in order to test the plausibility of the hypothesis under these conditions. Using a 2I-2AFC adaptive method, ramped and damped tones were matched in duration to the point of subjective equality. At equal subjective duration, short-damped sounds (< 0.5 s) were matched longer than short-ramped sounds, confirming previous results, whereas long-damped sounds (0.5 to 2 s) were matched to the same duration as long-ramped sounds, which question the hypothesis for durations over 500 ms.

1 Introduction

In the present article, the measurement of subjective duration of sounds that increase and decrease continuously in level is investigated. It was well established that perceptual asymmetries take place between ramped sound and its equivalent damped sound. Especially, even though ramped and damped sounds have the same overall levels, level changes, overall spectra and identical durations, ramped sounds are perceived louder and longer than damped sounds under specific conditions.

On the one hand a few studies deal with the comparison of global loudness (judgement on the total sound energy contained over the all duration of the sound) between ramped and damped sounds [1, 2, 3]. The different results are consistent with the fact that ramped sounds are perceived louder than damped sounds for durations from 250 ms [1] to durations longer than 2 s [2, 3], regardless of the method of measurement (two-interval comparison, magnitude estimation, analogical/categorical (A/C) scaling).

On the other hand, Susini et al. [4] observed similar results between the direct rating of overall loudness regarding ramped sounds, and the direct rating of loudness change [4, 5, 6]. Thus, when they are asked to rate the amount of change in loudness using a direct rating method, most listeners may have based their judgment on global loudness rather than on actual loudness change. For durations of 1.8 s and using direct rating methods (magnitude estimation [5] or analogical/categorical (A/C) scaling [6]), some authors have pointed out that loudness change is judged larger for ramped than for damped sounds. Their results are in concord with those of Stecker and Hafter [1], and Susini et al. [2, 3], on global loudness.

Concerning the subjective duration, ramped sounds were revealed to be perceived longer than damped sounds, for durations between 10 and 1000 ms. A substantial underestimation of the duration of damped sounds has been shown using either matching tasks [7, 8, 9, 10] or magnitude estimation [7].

Furthermore, asymmetries in subjective duration vary as a function of duration. According to Ries et al. [10], the differences between ramped and damped sounds are maximum at around 50-100 ms. At equal perceived duration, the ramped sounds are 1.7 to 1.9 longer than the equivalent damped sound. The differences decrease with increasing duration. Only one study investigated durations above 500 ms [8] where a significant reduction with increasing duration is also underlined by the authors. For durations of 1 s, the perceptual asymmetry is about 20% of duration.

The asymmetry founded for loudness as well as for perceived duration bring up questions about a possible common base in perception of loudness and of duration of time varying-level sounds. The question then arises on the influence of perceived duration on the loudness judgment. It could be hypothesized that the perceived duration asymmetry may be responsible for the loudness asymmetry. While major results have been established about the loudness asymmetry for durations of few seconds, the subjective duration asymmetry has not yet been the subject of specific studies for durations longer than 500 ms. Thus, the main question addressed by the present study is whether or not the disparities in subjective duration between ramped and damped sounds hold for sounds with duration of few seconds, in order to extend previous results, and test the validity of the hypothesis aforementioned, under these conditions.

In Experiment II, the physical durations of damped and ramped sounds were equalized to the point of subjective equality, via a 2I-2AFC adaptive method. However, when setting-up the experiment, it became necessary to know the discrimination threshold in duration for the stimuli used, in order to adjust the step size of the adaptive procedure. To our knowledge, there is no literature about the measurement of duration discrimination for ramped and damped sounds, with exception of one study for short duration stimuli (10 to 200 ms) [7]. As a consequence, duration discrimination thresholds were measured first in Experiment I.

2 Experiment I : Duration discrimination

This first experiment intends to measure the duration discrimination thresholds for ramped and damped stimuli. They were measured for six different durations from 50 ms to 2000 ms.

2.1 Listeners

Ten volunteered listeners (2 females and 8 males) aged from 23 to 54 years, participated in the experiment. The listeners were selected from staff and students of the laboratory. All of them reported normal hearing (thresholds ≤ 25 dB HL) at octave frequencies from 0.25

to 8.0 kHz, except one person that reported hearing loss between 30 and 60 dB HL at high frequencies ($\geq$ 6.0 kHz). Listeners were paid for their participation.

2.2 Stimuli

All the stimuli were 1 kHz pure tones. The six standard durations were 50, 100, 200, 500, 1000 and 2000 ms. The sounds had a ramped or a damped envelope that increase (65 to 80 dB SPL) or decrease (80 to 65 dB SPL) linearly in level, corresponding to a dynamic range of 15 dB. The sounds of 50, 100 and 200 ms had rise and fall times of 10 ms (linear ramps), for the longer sounds (500, 1000 and 2000 ms) the rise and fall times were of 20 ms (linear ramps).

2.3 Apparatus

Sounds were generated in real time by a processor DSP RP2 from the Tucker-Davis system III, using the graphical user interface RPVD. A custom-designed software also controlled the experiment. Tones were synthesized at a sample rate of 48828.125 kHz with 24 bits resolution. The output of the DAC converters of system III was amplified through the Edirol sound card (Audio Capture UA-5), and presented diotically through the Sennheiser HD280 Headphones. The calibration of sound levels from synthesis to reproduction was based on the headphone characteristic curves "at the eardrum". All the measurements were performed in a soundproof room. The listeners gave their answers by pushing a button on a response box. A screen on the box was used to provide temporal references as the interval currently played or the time to answer.

2.4 Procedure

Duration discrimination thresholds were measured with a 3I-3AFC procedure.

The test was divided into two parts: thresholds were first measured for long sounds (500, 1000 and 2000 ms), then, in a second part, for short sounds (20, 100 and 200 ms). In each part, 6 conditions were tested: 3 durations and 2 ramp directions (upramp and downramp). One threshold (for one condition) was obtained after a minimum of two runs. Thus, 12 runs per listeners were at least performed into each part of the experiment. The order in which the 12 runs were presented was randomized for each listener. A run was stopped after 12 reversals.

For each trial, two of the three intervals contained the test ramps with fixed standard durations (according to the drawn condition) and the remained interval contained the comparison ramp with variable duration. The interval containing the comparison ramp (1, 2 or 3) was randomized from one trial to another. The three ramps were presented alternately with a 500 ms interval between each one. Listeners were asked to select the interval containing the longest sound.

A 2-down/1-up rule was used from beginning to the 4th reversal in order to limit the number of trials at the beginning. Then, the rule became 3-down/1-up until the end of the run, it targeted 79.4% correct responses on the psychometric function [11]. For one run, the threshold was calculated as the geometric mean of the comparison ramp duration across the eight last reversals.

The starting duration of the comparison ramp was twice the standard duration. The step size varied within each run. It was 1/5th of the duration at the beginning, 1/10th of the duration after the second reversal, then 1/50th of the duration from the fourth reversal to the end. After each incorrect response, the duration of the increment was increased by a factor of 1.2, 1.1 (after the second reversal) and 1.02 (from the fourth reversal to the end) and three consecutive correct responses resulted in a decrease by a factor of 0.8, 0.9 and 0.98 respectively.

A measurement was accepted provided that the geometrical standard deviation of the duration values across the eight last reversals was lower than the multiplication factor to the power of three (1.02^3). A new run was performed and taken into account, until the consistency condition was satisfied. The same criterion was used between two measurements, considering the geometrical standard deviation between the two calculated thresholds.

The discrimination threshold in duration (or just-noticeable difference (JND)) was defined as the difference between the threshold (averaged across 2 repetitions or more if extra runs were performed) and the standard duration.

2.5 Results

Results for short duration signals (50, 100 and 200 ms) and long duration signals (500, 1000 and 2000 ms) are presented separately.

2.5.1 Short durations

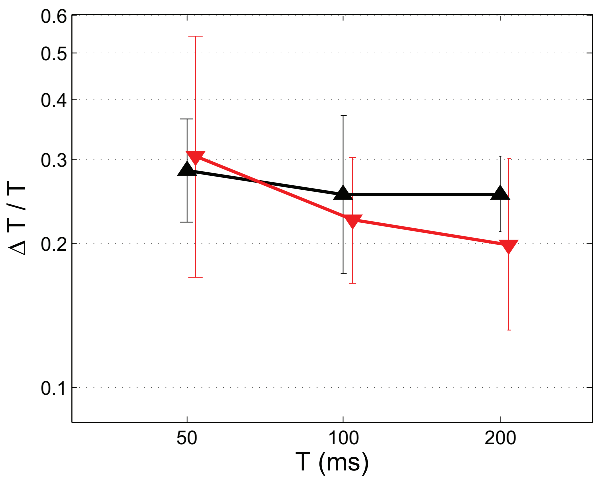

Figure 1: Duration Discrimination as a function of standard duration. The y-axis represents the Weber Fraction ($\Delta T/T$, where ΔT is the JND). Down-pointing triangles represent damped sounds and up-pointing triangles represent ramped sounds.

Figure 1 illustrates the JND averaged (geometric mean) across all the listeners for durations of 50, 100 and 200 ms. The figure represents the Weber fraction ($\Delta T/T$) as a function of duration (ms).

The averaged Weber fraction ranges from 0.28 at 50 ms and 0.25 at 200 ms for upramps and from 0.30 at 50 ms to 0.20 at 200 ms for downramps. The ratio decreases with increasing duration. Thus, the discrimination threshold is maximal at 50 ms as well for up- as for downramps, corresponding to about 30% of the standard ramp duration.

The Weber fractions across listeners were analyzed using a repeated-measure analysis of variance (ANOVA) with standard duration and ramp direction as factors. The analysis were done on the logarithm (base 10) of the data. There was a significant effect of standard duration [$F(2,18)=5.08$, $p<0.0178$]. However, there were no significant differences in the Weber fraction between ramped and damped sounds [$F(1,9)=2.73$, NS].

Moreover, substantial differences can be observed among the different listeners. According to the data, the Weber fraction can be larger by about a factor 6 from one listener to another under the same condition.

A principal component analysis (PCA) was completed on listeners, suggesting a reduction into 3 components, accounting for 42.6%, 32.0% and 13.7% of total variance. This analysis highlights some large inter-individual differences. These differences are represented by the standard deviations (geometric) plotted on Figure 1.

2.5.2 Long durations

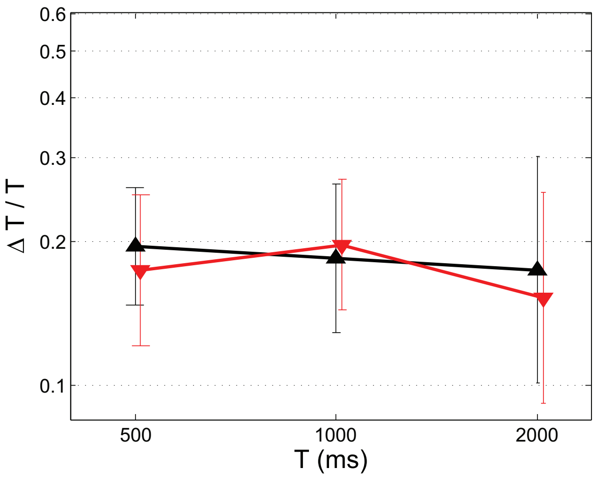

Figure 2: Duration Discrimination as a function of standard duration. The y-axis represents the Weber Fraction ($\Delta T/T$). Down-pointing triangles represent damped sounds and up-pointing triangles represent ramped sounds.

In the same way, the Figure 2 shows discrimination thresholds in duration averaged across all the listeners for long durations (500, 1000 and 2000 ms). The figure represents the Weber fraction ($\Delta T/T$) as a function of duration (ms).

All the Weber fractions are strictly ranged between 0.15 and 0.20. Data were analyzed by repeated-measure analysis of variance (ANOVA) with standard duration and ramp direction as factors. Results showed that neither duration [$F(2,18)=1.39$, NS] nor direction [$F(1,9)=1.94$, NS] had significant effect on the Weber ratio.

As for short durations, the Weber fraction substantially varies between the different listeners. It results in large standard deviations (geometric) on the Figure 2. Results from the Principal component analysis (PCA) completed on listeners, suggested to retain 3 components accounting for 43.2%, 28.6% and 18.4% of total variance. This analysis confirms the existence of large inter-individual differences.

2.6 Discussion

Results disclose important differences between listeners in the duration discrimination task of up- and downramps. However, according to the averaged data across all the listeners, the Weber fraction significantly decreases with duration for short durations. A decrease in the Weber fraction is also observed in Figure 7 (p. 2886) of Schlauch et al. [7], but it was not noticed by the authors. The Weber fraction was not constant for short durations, either for ramps or rectangular-gated sounds [7]. So, the subjective duration for such stimuli does not fit with the Weber's law ($\Delta T/T = k$ at point of subjective equality, where k is a constant). However, for long durations, the Weber fraction does not vary with duration for up- and downramps, so the Weber's law is verified.

Table 1: Averaged JND as a percentage of duration (($\Delta T/T) \times 100$).

	JND (%)			
	50 ms	**100 ms**	**200 ms**	**500-1000-2000 ms**
S1	22.59	15.81	15.9	8.78
S2	21.81	25.98	24.52	18.07
S3	20.04	17.62	17.67	12.58
S4	45.73	31.11	23.97	24.51
S5	27.38	18.98	22.02	15.79
S6	21.78	20.99	19.11	17.26
S7	51.98	43.7	40.14	25.25
S8	25.61	19.31	20.49	25.06
S9	30.42	24.26	22.67	27.85
S10	44.46	31.88	25.6	14.3

It can also be noticed, that in contrast with Schlauch et al. [7] who found significant differences in the Weber fraction between ramped and damped noises, no difference between up- and downramps was revealed in our study.

The aim of this experiment was to define an accurate and effective step size for the adaptive method used in Experiment II in order to equalize up- and downramps in subjective duration. The definition of the step size is then directly related to the measured discrimination thresholds. So, as for short duration sounds, the Weber fraction depends on the listener and on the duration, the step was individualized and depended on the duration. For long duration sounds, the step was only individualized, as neither the effect of duration nor the effect of direction were significant. The averaged values of JND, selected for each listener and each duration, are introduced in Table 1 as a percentage of duration.

3 Experiment II : Subjective duration

This experiment aims to measure the physical duration differences between ramped and damped sounds at point of subjective equality (PSE).

3.1 Listeners, Apparatus and Stimuli

The ten listeners who participated in Experiment I, also participated in Experiment II. The stimuli and the apparatus used are tightly the same as in Experiment I.

3.2 Procedure

A 2I-2AFC procedure, using a 2-down/2-up rule was used to match up- and downramps to equal subjective duration. This method targeted 70.7% correct responses on the psychometric function [11].

Each trial consisted in the presentation of one ramped sound and one damped sound. While the upramp always was the test sound with fixed standard duration, the downramp always was the sound of comparison with an incremented duration. The interval containing the comparison ramp (1 or 2) was randomized from one trial to another. 500 ms silent periods separated the two intervals. The listeners had to select the interval containing the longest sound, paying particular attention to consider duration only.

The step sizes in duration for the comparison ramp were individualized according to the discrimination thresholds obtained in Experiment I (Table 1). Moreover, the step size varied within each run. It was $round(2 \times (JND \times T))$ at the beginning, $round(JND \times T)$ after the second reversal, then $round((JND \times T)/1.2)$ from the fourth reversal to the end (T is the duration of the comparison ramp, JND is expressed as a proportion of duration (between 0 and 1) and $round$ represents the approximation to the closest integer). These rules were empirically determined on the occasion of this experiment. A run was stopped after 12 reversals. For each run, a PSE was calculated as the geometric mean of the damped ramp durations across the eight last reversals.

At the start of a run, two different durations were investigated for the ramp of comparison: a "long" starting duration set to a factor of 2 above the standard duration and a "short" starting duration set to a factor of 2 below the standard duration.

Thus, the whole test consisted of 24 runs corresponding to 12 different conditions (two runs performed for each condition). Each condition was characterized by two parameters: one standard duration ($\times 6$) and one starting duration ($\times 2$). The order of presentation of the 24 runs was randomized.

A measurement was accepted provided that the geometrical standard deviation of the duration values across the eight last reversals was lower than the square of the multiplication factor $((round(1+JND/1,2))^2)$. Regarding the consistency between two runs, the geometrical standard deviation across the two thresholds had to be lower than the multiplication factor $(round(1 + JND/1,2))$. A new run was added until the consistency conditions were satisfied.

The duration of perceptual equality was defined as the geometric mean across the four thresholds (2 repetitions $\times$ 2 starting durations) or more (if extra runs were performed).

3.3 Results

Figure 3 illustrates the results expressed as the ratio of the damped duration to that of the ramped duration, at PSE.

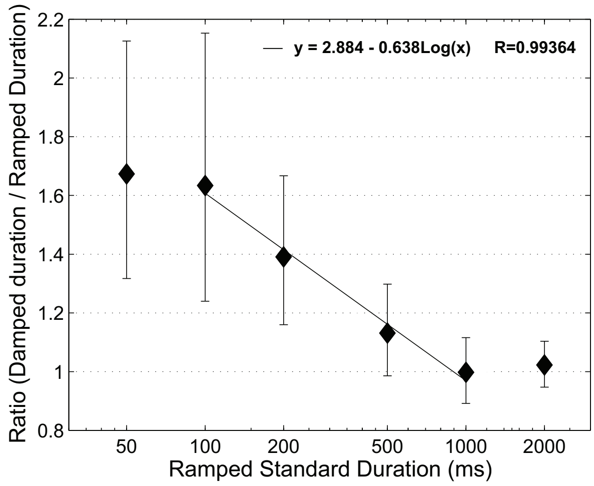

Figure 3: Damped/Ramped duration ratios at PSE as a function of standard duration. The black diamonds represent the averaged ratios across all the listeners with standard deviation of the mean. The line represent a logarithmic fit to the data across durations from 100 to 1000 ms.

A repeated-measure analysis of variance (ANOVA) completed with duration as factor revealed a significant decreasing of the ratio with increasing duration [$F(5,45)=33.14$, $p<0.01$]. The ratio ranges from 1.67 at 50 ms to 0.99 and 1.02 at 1000 ms and 2000 ms respectively. The decreasing is maximal between 100 ms and 1000 ms. For ramps longer than 1000 ms, the ratio stabilizes around the value of 1. Thus, for short durations, at PSE the downramps are longer than the upramps, showing that downramp are perceived shorter than upramps. The difference between the subjective durations of down- and upramps decreases down to zero for durations of 1 s and longer.

A logarithmic fit to the data (across durations from 100 to 1000 ms only) was plotted in Figure 3. The related correlation coefficient is $R = 0.99$, in relation to the mean data. Thus, the asymmetry is logarithmically reducing with increasing duration from 100 ms to 1000 ms.

Furthermore, the standard deviation (geometric) decreases with duration. This may be explained by the fact that discrimination thresholds are globally higher for short durations. Thus, the adaptive method may be less accurate for such durations and result in an increasing of inter-individual differences in the results. Generally, the feelings reported by listeners have revealed increasing difficulties to judge the overall duration of sounds with decreasing duration.

3.4 Discussion

This experiment allowed to observe the progression of asymmetries in perceived duration as a function of duration from 50 to 2000 ms.

Concerning short durations (until 500 ms), the results are consistent with the four main studies [10, 9, 8, 7] given that ramped sounds are perceived longer than damped sounds. For example, some results obtained for broadband noises via a matching procedure [10] (Fig. 2, p. 3774), show that the perceived duration of damped sounds is underestimated by a factor between 2 and 1.5 from 50 to 200 ms compared to ramped sounds. A similar underestimation by a factor of about 1.35 at 250 ms was noticed by Grassi and Darwin [8]. In DiGiovanni et al. [9], the ratio (damped duration / ramped duration) at equal subjective duration was 1.46 at 50 ms and 1.85 at 500 ms. Even though there are some differences between the results from the present experiment (ratio from 1.67 at 50 ms to 1.13 at 500 ms) and the results from other studies, a common trend stands out and is consistent with an underestimation of the duration of damped sounds. Differences in results may be due to methodological differences and to the nature of stimuli used [7]. A significant effect of instructions on subjective duration has also been shown [9].

Besides, according to the present results, the perceptual asymmetries tend to reduce with duration, starting at 100 ms. Two studies also revealed a reduction of the asymmetries with duration [8, 10] starting at 250 and 200 ms, respectively. The present study, and those of Grassi and Darwin [8] and Ries et al. [10], show that asymmetries are reduced with duration, beginning at 100 ms. Moreover, asymmetries have been shown to disappear for durations above 1000 ms. The previous authors were unable to observe this phenomenon because of using durations shorter than 500 ms, except Grassi and Darwin [8] who investigated a longer duration. Asymmetries around 20% of the duration were observed by the authors [8] at 1000 ms for a complex tone, whereas no asymmetries were obtained in the present study for the same duration. A stationary sound was used as a reference to match the ramped and damped stimuli in Grassi and Darwin [8]. So, methodological differences may explain the differences with the present results.

Asymmetries were observed for short durations only, maybe below the temporal integration duration of the hearing system, ranged between 100 and 400 ms according to the study [12, 13]. The question then arises to know if listeners effectively judge duration or other parameters of the stimulus as loudness, pitch or timbre. Besides, at times some listeners had the feeling to base their judgment on other criteria than duration.

4 Conclusion

It was hypothesized that perceptual asymmetries in loudness between ramped and damped sounds may be linked to asymmetries in duration, and that the perceived duration asymmetries may be responsible for the loudness asymmetry. The aim of this study was to match the duration of up- and downramps to the point of subjective equality, in order to show the progression of asymmetries in subjective duration for stimuli longer than 500 ms, and so, tie in with durations of few seconds usually used in loudness investigations.

On the one hand, our results confirm already existing results in literature, showing perceptual asymmetries for short durations ($<$ 500 ms) and that asymmetries are reduced with increasing duration. On the other hand, the asymmetries disappear for longer durations ($>$ 500 ms). This constitute a novel result in literature about the subjective duration of non-stationary sounds. For future work, it could be interesting to repeat the experiment for durations longer than 2 s, in order to check if the absence of perceptual asymmetries is maintained with increasing duration.

According to the present study, asymmetries in subjective duration disappear for durations longer than 1 s whereas the loudness asymmetry is still observed for these durations [2, 3]. As a consequence, the perceptual asymmetry in duration could partially explain the loudness asymmetry for short durations, but could not explain it for long durations. One perspective is the measurement of loudness for ramps previously equalized in duration (at PSE).

References

[1] G. Stecker, E. Hafter, "An effect of temporal asymetry on loudness", *J. Acoust. Soc. Am.* **107**, 3358-3368 (2000)

[2] P. Susini, S. McAdams, B. Smith, "Loudness asymmetries for tones with increasing and decreasing levels", *Proceedings of ICAD 05* (2005)

[3] P. Susini, S. McAdams, B. Smith, "Loudness asymmetries for tones with increasing and decreasing levels using continuous and global ratings", *Acta Acustica united withh Acustica* **93**, 623-631 (2007)

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

[5] R. Teghtsoonian, M. Teghtsoonian, G. Canévet, "Sweep-induced acceleration in loudness change and the bias for rising intensities", *Attention, Perception, & Psychophysics* **67**(4), 699-712 (2005)

[6] J. Neuhoff, "Perceptual bias for rising tones", *Nature* **395**, 123-124 (1998)

[7] R. Schlauch, D. Ries, J. DiGiovanni, "Duration discrimination and subjective duration for ramped and damped sounds", *J. Acoust. Soc. Am.* **109**, 2880-2887 (2001)

[8] M. Grassi, C. Darwin, "The subjective duration of ramped and damped sounds", *Perception & psychophysics, Psychonomic Society Publications* **68**, 1382-1392 (2006)

[9] J. DiGiovanni, R. Schlauch, "Mechanisms responsible for differences in perceived duration for rising-intensity and falling-intensity sounds", *Ecological Psychology* **19**, 239-264 (2007)

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

[11] H. Levitt, "Transformed up-down methods in psychoacoustics", *J. Acoust. Soc. Am.* **49**, 467-477 (1971)

[12] I. Boullet, "La sonie des sons impulsionnels : Perception et mesure", *Les cahiers de l'audition* **21**, (2007)

[13] B. Scharf, *Handbook of Perception (Chapter "Loudness")* Academic Press, NY (1978)