



HAL
open science

Effect of Pitch on the Asymmetry in Global Loudness Between Rising- and Falling-Intensity Sounds

Sabine Meunier, Jacques Chatron, Blandine Abs, Emmanuel Ponsot, Patrick
Susini

► **To cite this version:**

Sabine Meunier, Jacques Chatron, Blandine Abs, Emmanuel Ponsot, Patrick Susini. Effect of Pitch on the Asymmetry in Global Loudness Between Rising- and Falling-Intensity Sounds. *Acta Acustica united with Acustica*, 2018, 104 (5), pp.770-773. 10.3813/aaa.919220 . hal-01928467v1

HAL Id: hal-01928467

<https://hal.science/hal-01928467v1>

Submitted on 14 Feb 2019 (v1), last revised 20 Nov 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Effect of pitch on the asymmetry in global loudness between rising- and falling- intensity sounds

Sabine Meunier¹⁾, Jacques Chatron¹⁾, Blandine Abs¹⁾, Emmanuel Ponsot²⁾, Patrick Susini³⁾

¹⁾ Aix Marseille Univ, CNRS, Centrale Marseille, LMA,
13453 Marseille Cedex 13, France. meunier@lma.cnrs-mrs.fr

²⁾ Laboratoire des Systèmes Perceptifs (CNRS UMR 8248) and Département d'études cognitives,
Ecole Normale Supérieure, PSL Research University, Paris, France

³⁾ STMS (UMR9912), Ministère de la Culture, Ircam, CNRS, Sorbonne Université, Paris, France.

1 Summary

The global loudness of a varying intensity sound is greater when the intensity increases than when it decreases. This global loudness asymmetry was found to be larger for pure tones than for broadband noises. In this study, our aim was to determine whether this difference between pure tones and noises is due to the difference in bandwidth between sounds or to the difference in the strength of the sensation of pitch. The loudness asymmetry was measured for broadband and for narrow-band signals that do or do not elicit a sensation of pitch. The asymmetry was greater for sounds that elicit a sensation of pitch whatever their bandwidth. The loudness model for time varying sounds [1] predicted well the asymmetry for the broadband noise that does not elicit a sensation of pitch and for a multi-tonal sound. For the other sounds the asymmetry was greater than predicted. It is known that loudness and pitch interact. The difference in asymmetry between sounds that elicit pitch and sounds that do not elicit pitch might be due to this interaction.

1 Introduction

Perceptual differences have been observed between sounds that increase and sounds that decrease in intensity with identical energies and long-term spectra. Rising-intensity sounds are perceived longer [2, 3] and as changing more in loudness than falling-intensity sounds [4, 5]. Their global loudness (that is, the overall loudness of the sound over its entire duration) is also greater than the global loudness of a falling sound. At the point of subjective equality (PSE), a falling 1-kHz pure tone varying over 15-dB has a rms level that is 4 dB higher than that of its symmetrical rising version [6]. This global loudness difference is usually called loudness asymmetry. Various studies have attempted to explain and model the global loudness asymmetry [7, 6]. Ponsot et al. [7]

showed a larger loudness asymmetry for pure tones than for broadband noises. Neuhoff [5] found the same tendency when addressing loudness change. Ponsot et al. [8] showed that the asymmetry in loudness was not due to a different weighting of the loudest part of the signal when presented at the end rather than at the beginning of the signal. The aim of the present study was to determine whether the difference in loudness asymmetry between tones and noises is due to their difference in spectrum width or in pitch strength. Indeed, pure tones elicit strong perception of pitch whereas broadband noises do not elicit any pitch, which may explain the difference in asymmetry. It is known that the dimensions of loudness and pitch interact with each other (see [9] for example). For example, when listeners have to perform an intensity discrimination task, the reaction time is faster when both stimuli have the same pitch or when pitch and loudness are congruent (high pitch/high loudness, low pitch/low loudness) than when pitch and loudness are incongruent (high pitch/low loudness, low pitch/high loudness). Moreover, the judged pitch of rising-intensity sound (with constant frequency) increases and the judged pitch of falling-intensity sound decreases ([10, 11]). Thus combining loudness variation (due to intensity variation) and pitch variation (also due to intensity variation) may reinforce the difference between rising and falling ramps.

In this study, we attempted to measure the loudness asymmetry for broadband and narrow-band sounds that do and do not elicit pitch, in order to separate the effects of signal bandwidth from those of pitch.

2 Method

Seven women and nine men took part in the experiment. All listeners had hearing thresholds less than 15 dB HL. Their thresholds were determined using standard pure-tone audiometry in the frequency range between 0.125 and 8 kHz (with an AC40 audiometer). The participants were all naïve with respect to the

78 hypotheses under test. They were paid for their partic-
79 ipation.

80 The sounds were played via a RME UCX Fireface
81 sound-card and were presented diotically through
82 headphones (Sennheiser HDA 200). They were sam-
83 pled at a frequency of 44.1 kHz with 16 bits resolu-
84 tion. The headphones were calibrated using a Brüel
85 and Kjær Artificial Ear (type 4153) coupled with the
86 mounting plate provided for circumaural headphones
87 with no free-field equalization and a Brüel and Kjær
88 voltmeter (26209). The experiment took place in a
89 double-walled soundproof booth.

90 The level of the stimuli varied linearly over its du-
91 ration between 50 and 65 dB SPL, either increasing in
92 intensity for the rising-intensity ramps or decreasing
93 for the falling-intensity ramps. Four different spec-
94 tral contents were created in order to obtain a broad-
95 band sound with pitch, a broadband sound without
96 pitch, a narrow-band sound (less than 1 critical band)
97 with pitch and a narrow-band sound without a con-
98 stant pitch. The broadband sound with pitch was a
99 frozen white noise repeated each millisecond, which
100 created a pitch corresponding to a frequency of 1 kHz
101 (RWN). Its spectrum is that of a harmonic complex
102 tone with a fundamental frequency of 1 kHz, random
103 phases and amplitudes (see [12]). The broadband
104 sound without pitch was a white noise (20Hz-20kHz,
105 WN). The narrow-band sound with pitch was a pure
106 tone at 1 kHz (PT). Creating a narrow-band sound
107 that does not elicit any pitch is not straightforward.
108 Noise of bandwidth less than one critical band cannot
109 be considered as providing no pitch. The strength of
110 the pitch is weaker than that of a pure tone, but it
111 exists and corresponds to the central frequency of the
112 noise. Instead of looking for narrow-band noise with-
113 out pitch, we designed a signal with multiple pitches.
114 The idea was that listeners could not concentrate on
115 each pitch and thus would not perceive a global pitch
116 at the end of the sound. The stimulus was a pure
117 tone whose frequency varied over variable periods of
118 time. The frequency was randomly drawn on an uni-
119 form distribution (on the hertz scale) between 920 and
120 1080 Hz (width of one critical band). Segment dura-
121 tions were chosen randomly between 50 and 150 ms
122 (figure 1). Each segment was gated on and off (r/f
123 time=2 ms), preventing spectral splatter. The seg-
124 ments were not overlapping. Once the frequency and
125 the duration of each segment selected, the multi-tonal
126 (MT) sound was frozen throughout the experiment.
127 Affirming that MT does not elicit any pitch would be
128 incorrect. Its pitch changes very quickly, it can only
129 be said that it does not elicit a global pitch. Thus
130 its pitch strength might be weaker than that of PT
131 or RWN. The first three sounds lasted 2s. The multi-
132 tonal stimulus was slightly shorter (1930.5 ms) be-
133 cause of the variable duration segments. The rise/fall
134 times of all signals were 10 ms.

135 The loudness asymmetries were evaluated by mea-

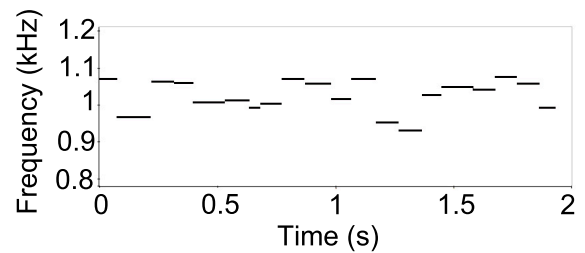


Figure 1: time-frequency representation of the multi-
tonal stimulus (MT).

136 suring the PSE in loudness between signals with rising
137 and falling ramps. This procedure has already been
138 employed in previous studies ([6], [7]) and showed
139 equivalent results to an absolute magnitude estima-
140 tion task. A two-interval, two-alternative forced-
141 choice paradigm (2I-2AFC) based on an interleaved
142 adaptive procedure ([13], [14]) was used to measure
143 the PSEs. On each trial, the listeners heard two
144 sounds, one with falling and one with rising intensity,
145 separated by 500 ms. Their task was to indicate which
146 sound was louder by pressing a key. The response ini-
147 tiated the next trial after a 1-s delay. The overall level
148 of one sound was fixed (test sound) all along the track
149 while the overall level of the other varied (comparison
150 sound) depending on the response of the listener to
151 the preceding trial. The minimum and maximum lev-
152 els of the comparison sound varied, while its dynamic
153 was kept constant at 15 dB, and were adjusted accord-
154 ing to a simple up-down procedure [15]. If the listener
155 indicated that the comparison sound was the louder
156 one, its maximum level (and thus its minimum level)
157 was reduced, otherwise it was increased. At the begin-
158 ning of a track the maximum (and minimum) level of
159 the comparison sound was 5 dB higher or lower than
160 the maximum (and minimum) level of the test sound.
161 Step-size variation was 6 dB until the second reversal
162 and was reduced by a ratio of 2 after every two re-
163 versals and was held constant after six reversals. A
164 track ended when 8 reversals were achieved. The level
165 corresponding to the matching loudness was defined
166 as the average of the maximum level of the compari-
167 son sound in the last two reversals of each track. This
168 procedure converges at the level corresponding to the
169 50% point on the psychometric function [15].

170 One block was composed of four interleaved tracks
171 (figure 2). In one track, the pair order was ris-
172 ing/falling and the second sound of the pair was the
173 comparison stimulus (figure 2a), in another track the
174 comparison stimulus was presented first (figure 2b).
175 In two other tracks the pair order was falling/rising
176 and the comparison stimulus was either presented first
177 (figure 2d) or second (figure 2c). Each type of stimu-
178 lus (RWN, PT, WN and MT) was tested in different
179 blocks. Block order was randomized across listeners.
180 Each listener completed two blocks. The asymmetry



Figure 2: Experimental configuration

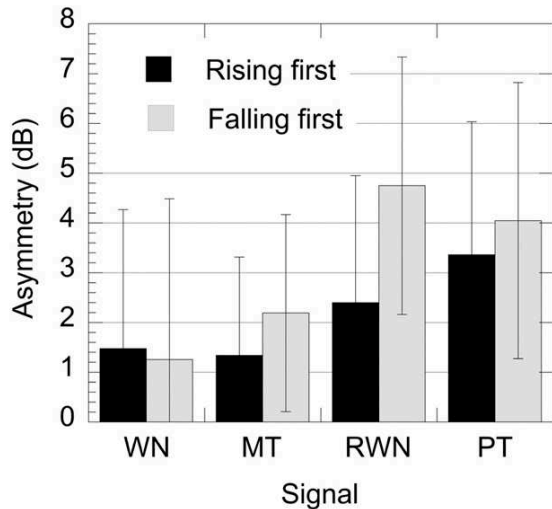


Figure 3: Loudness asymmetries between rising and falling ramps. Positive asymmetry indicates that the falling sound was matched higher in level than the rising sound in order to be perceived equally loud, which means that the falling sound would be perceived softer than the rising sound of the same level. Error bars correspond to standard deviation.

181 was calculated as the difference between the maxi-
 182 mum level of the falling sound and the maximum level
 183 of the rising sound at PSE (mean of the two values ob-
 184 tained in each block). The asymmetry was calculated
 185 for each pair order (rising/falling and falling/rising)
 186 as the average of the asymmetry obtained for the two
 187 tracks that differed by the position of the comparison
 188 stimulus (1st or 2nd interval). Rising first and falling
 189 first are shown separately, as previous studies have
 190 shown an effect of the order of presentation on the
 191 asymmetry ([3], [7]).

192 3 Results

193 Figure 3 shows the loudness asymmetries obtained
 194 with the different types of sounds and for the differ-
 195 ent pair orders. Repeated-measures ANOVAs showed
 196 a significant effect of the type of stimulus [$F(3,$
 197 $15)=15.21$, $p < 0.001$, $\eta_p^2 = 0.5$] and no effect of
 198 the pair order [$F(1,15)=3.02$, $p=0.1$]. The analysis
 199 showed a significant type-of-stimulus \times pair-order in-
 200 teraction [$F(3,45)=6.35$, $p=0.001$, $\eta_p^2 = 0.3$]. Post-hoc
 201 LSD tests showed no significant difference between
 202 WN and MT ($p=0.37$) and between PT and RWN
 203 ($p=0.76$). The asymmetry produced by WN was

204 significantly smaller than those produced by RWN
 205 and PT ($p < 0.001$). The asymmetry produced by
 206 MT was significantly smaller than those produced by
 207 RWN and PT ($p < 0.001$). The results show that the
 208 asymmetry is greater for the sounds producing a con-
 209 stant or global pitch (RWN and PT), whatever their
 210 frequency widths, than for the white noise.

211 4 Discussion

212 In the experiment reported here, sounds that elicit
 213 a constant or global pitch sensation produce larger
 214 loudness asymmetries than WN and MT, irrespective
 215 of bandwidth. It thus seems that pitch had an effect
 216 on the size of the loudness asymmetry, while band-
 217 width did not. To summarize the data, the asymme-
 218 tries were averaged across pair order. The loudness
 219 asymmetry induced by WN (asymmetry=1.36 dB)
 220 was smaller than that induced by MT (asymme-
 221 try=1.76 dB) which was smaller than that induced by
 222 RWN (asymmetry=3.56 dB) which was smaller than
 223 that induced by PT (asymmetry=3.68 dB).

224 In order to explain the asymmetries and the dif-
 225 ferences between RWN and PT on the one hand,
 226 and WN and MT on the other hand, the loudness
 227 model for time varying sounds of Glasberg and Moore
 228 [1] was used to predict loudness differences between
 229 the rising- and falling-intensity sounds used in our
 230 study. The difference in level between rising and
 231 falling sounds needed to obtain the same peak value of
 232 the Long Term Loudness (LTL) for both signals was
 233 used to estimate the asymmetry in dB. The predicted
 234 asymmetry was 1.3 dB for WN, 2.1 dB for MT, 1.5 dB
 235 for RWN and PT. The observed asymmetry was quite
 236 well explained by the model for the sound that does
 237 not elicit pitch (WN) whereas it was underestimated
 238 for the sounds that elicit pitch (RWN and PT). For
 239 MT, the loudness model slightly overestimated the
 240 asymmetry. Our results are in agreement with those
 241 of Ries et al. [3] who found that LTL predicted well
 242 the loudness asymmetry for white noise of 500 ms,
 243 and with those of Ponsot et al. [6], who predicted an
 244 asymmetry of 1.24 dB for a 1-kHz pure tone of 2 s
 245 varying between 55 and 70 dB, using the LTL model.

246 The long-term loudness is calculated from instan-
 247 taneous loudness using two stages of temporal inte-
 248 gration. It models the overall loudness impression
 249 which is probably formed at a central level of the
 250 auditory system. For the white noise, it seems that
 251 this model based on temporal integration is consistent
 252 with the behavioral data. However, for the repeated
 253 white noise and the 1-kHz pure tone the model un-
 254 derestimated the loudness asymmetry. Our assump-
 255 tion is that the loudness asymmetry is increased by
 256 the sensation of pitch. McBeath and Neuhoff [11]
 257 showed that tones with continuous intensity change
 258 and constant frequency were perceived as changing in

pitch. When the intensity increased (respectively decreased), the pitch increased (respectively decreased). Moreover, the pitch change of rising-intensity sounds was larger than the pitch change of falling-intensity sounds. The combined effect of loudness variation (due to intensity variation) and pitch variation (also due to intensity variation) might reinforce the loudness difference between sounds with rising and falling ramps. Thus, for RWN and PT the asymmetry would result from temporal integration of instantaneous loudness and a cognitive mechanism due to the interaction between the dimensions of loudness and pitch. For MT, we could emit the same hypothesis as for WN, but its peculiarity (several pitches) leads us to remain cautious. It should also be noted that the difference between WN and RWN might be due to the fact that WN has more energy in the low frequency region than RWN. The asymmetries were assessed with the same participants for the different sounds. This allows calculation of interindividual correlations between these conditions; significant correlations between conditions might indicate shared mechanisms [16] [17]. Non-parametric correlations between the 6 possible combinations (Kendall's tau) were all significant at $p < 0.05$ (after Bonferonni correction with $\alpha = 6$), except those with the MT condition. Thus subjects with greater asymmetries, e.g. for PT, also exhibited stronger asymmetries for WN and RWN. This might indicate that the mechanism(s) at the origin of the asymmetries for the white noise, the repeated white noise as well as the 1-kHz pure tone is (are) shared (at least partially), whereas other mechanism(s) might contribute to those observed for the multi-tonal sound, reinforcing its peculiarity. These assumptions could be tested in further studies in which the loudness asymmetry would be measured for sounds with parameterizable pitch strength (like iterated rippled noises or pure tone in noise) and the correlation between pitch strength and loudness asymmetry would be assessed.

References

- [1] B. R. Glasberg and B. C. J. Moore. A model of loudness applicable to time-varying sounds. *J. Audio Eng. Soc.*, 50(5):331–342, 2002.
- [2] M. Grassi and C. J. Darwin. The subjective duration of ramped and damped sounds. *Perception and Psychophysics*, 68(8):1382–1392, 2006.
- [3] D. T. Ries, R. S. Schlauch, and J. 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(6):3772–3783, 2008.
- [4] K. N. Olsen, C. J. Stevens, and J. Tardieu. Loudness Change in Response to Dynamic Acoustic Intensity. *J. Exp. Psychol.-Hum. Percept. Perform.*, 36(6):1631–1644, 2010.
- [5] J. G. Neuhoff. Perceptual bias for rising tones. *Nature*, 395(6698):123–124, 1998.
- [6] E. Ponsot, P. Susini, and S. Meunier. A robust asymmetry in loudness between rising- and falling-intensity tones. *Attention, Perception, & Psychophysics*, 77(3):907–920, 2015.
- [7] E. Ponsot, S. Meunier, A. Kacem, J. Chatron, and P. Susini. Are Rising Sounds Always Louder? Influences of Spectral Structure and Intensity-Region on Loudness Sensitivity to Intensity-Change Direction. *Acta Acustica united with Acustica*, 101(6):1083–1093, 2015.
- [8] E. Ponsot, P. Susini, G. Saint Pierre, and S. Meunier. Temporal loudness weights for sounds with increasing and decreasing intensity profiles. *J. Acoust. Soc. Am.*, 134(4):EL321–EL326, 2013.
- [9] R. D. Melara and L. E. Marks. Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology*, 16(2):398–414, 1990.
- [10] J. G. Neuhoff and M. K. McBeath. The Doppler illusion: The influence of dynamic intensity change on perceived pitch. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4):970–985, 1996.
- [11] M. K. McBeath and J. G. Neuhoff. The Doppler effect is not what you think it is: Dramatic pitch change due to dynamic intensity change. *Psychon Bull Rev*, 9(2):306–313, 2002.
- [12] R. M. Warren and J. A. Bashford. Production of white tone from white noise and voiced speech from whisper. *Bull. Psychon. Soc.*, 11(5):327–329, 1978.
- [13] W. Jesteadt. An adaptive procedure for subjective judgments. *Perception & Psychophysics*, 28(1):85–88, 1980.
- [14] M. Florentine, S. Buus, and T. Poulsen. Temporal integration of loudness as a function of level. *J. Acoust. Soc. Am.*, 99(3):1633–1644, 1996.
- [15] H. Levitt. Transformed Up/Down Methods in Psychoacoustics. *J. Acoust. Soc. Am.*, 49:467–477, 1971.
- [16] G. Canévet, B. Scharf, and M.-C. Botte. Simple and Induced Loudness Adaptation. *Audiology*, 24(6):430–436, 1985.
- [17] D. Pressnitzer and J.-M. Hupé. Temporal Dynamics of Auditory and Visual Bistability Reveal Common Principles of Perceptual Organization. *Current Biology*, 16(13):1351–1357, 2006.