Effect of the interaural time difference on the loudness of pure tones as a function of the frequency

Vincent Koehl, Mathieu Paquier, Etienne Hendrickx

Effect of the interaural time difference on the loudness of pure tones as a function of the frequency

Vincent Koehl[1)], Mathieu Paquier[1)], Etienne Hendrickx[1)]
[1)] Lab-STICC (UMR CNRS 6285), University of Brest,
6 avenue Victor Le Gorgeu, 29200 Brest, France.
vincent.koehl@univ-brest.fr

Summary

Significant loudness variations with source azimuth (i.e. directional loudness) are generally accounted for by at-ear pressure modifications. An effect of the interaural time difference (ITD) was also reported in previous studies by the authors: the loudness of pure tones (200 and 400 Hz) significantly increased when the stimuli were presented with an ITD of 772 μs, corresponding to an azimuth of 90°. The present study aims at observing this effect for higher frequencies, including frequencies at which ITD is no longer useful as a localization cue. The effect of ITD on the loudness of pure tones was thus studied at 500, 707, 1000, 1404 and 2000 Hz. Results show that the effect of ITD on loudness is not significant above 500 Hz, even for frequencies where ITD is still a localization cue. The effect observed at 500 Hz is still in agreement with the results reported by previous studies as the loudness of a pure tone significantly increases when its ITD is 772 μs.

1 Introduction

Directional loudness (i.e. loudness variations with the direction of the sound source) has been highlighted by presenting bands of noise through loudspeakers located in various directions around a listener in an anechoic room [1, 2]. As an example, a third-octave noise band centered at 5000 Hz presented by a loudspeaker at an azimuth of 90° was perceived as being about 5 dB louder than when presented by a loudspeaker at an azimuth of 0°, in free field. The effect is indeed particularly salient in high frequencies when the sources are located in the horizontal plane. Physical modifications of the at-ear pressures caused by the acoustic shadow of the head largely account for this effect in this case. The directional loudness sensitivity (DLS) was even reported to be significantly different from zero at 400 Hz [2]. Below 500 Hz, shadowing effects are usually considered rather small [3] as variations of ILD with the azimuth do not exceed 5 dB at 500 Hz and decrease for lower frequencies. Whether or not these variations can be considered as small, it was suspected that at-ear pressure modifications could not be the only cause of directional loudness.

Recent studies have highlighted a significant effect of interaural time differences (ITD) on the loudness of low-frequency pure tones (200 and 400 Hz) at 40 phon (but not at 70 phon) [4]. Loudness was significantly increased by 1.25 dB when the ITD was increased from 0 to 772 μs. This effect was similarly observed when presenting the pure tones with an interaural level difference (ILD $\leq$ 5 dB) [5]. This was observed when ITD and ILD were congruent, but also when opposite (i.e. leading to opposite sides). As such opposite interaural differences should compensate for the lateralization induced by each other, it suggests that the loudness increase with ITD is not related to the perceived source lateralization but that ITD itself affects the loudness process.

Above 500 Hz, it is still assumed that the variation of DLS with source position is mainly caused by at-ear pressure modifications. However, at 5000 Hz, model predictions reported that pressure modifications did not totally account for the loudness variations observed when varying the azimuth [6]. This model was designed to predict binaural summation for sounds differing in level at the two ears by taking into account contralateral inhibitions, assuming that a strong input to one ear can inhibit a weaker input to the other one. A tendency for the predicted DLS to be slightly below the measured values was found and reached 1.7 dB for the largest deviation. Such a deviation could be caused by a contribution of ITD to directional loudness for frequencies above 500 Hz as the binaural summation might also be affected by time differences at the two ears. Therefore, the aim of the present study is to observe further the effect of ITD on loudness that was observed at 40 phon for 200 and 400 Hz, when considering frequency from 500 Hz to 2000 Hz by half-octave steps.

2 Experimental setup

2.1 Stimuli

Interaural time differences were applied to pure tones whose frequencies were 500, 707, 1000, 1404 and 2000 Hz. As ITD may slightly vary with frequency, it was computed according to two different models; Kuhn's model [7]:

$$\mathrm{ITD} = \frac{3a}{c_0}\sin\theta_{inc} \qquad (1)$$

and Woodworth's model [8]:

$$\mathrm{ITD} = \frac{a}{c_0}(\theta_{inc} + \sin\theta_{inc}) \qquad (2)$$

where $a = 8.75$ cm is the standard head radius, $c_0 = 340\ \mathrm{m \cdot s^{-1}}$ is the speed of sound in the air and θ_{inc} is the incidence angle for a sound source in the horizontal plane (i.e. azimuth). Eq. (1) provides a better estimate of the ITD than Eq. (2) below 800 Hz, whereas the opposite trend is observed above 1500 Hz [9].

For $\theta_{inc} = 90°$, ITD = 772 μs according to Kuhn's model and ITD = 662 μs according to Woodworth's model. Smaller (607 μs), intermediate (717 μs) and larger (827 μs) ITD values were also selected according to a 55 μs step. Five non-zero ITD values were then under study, as well as a zero ITD for comparison. The six possible ITD values were introduced between the left and right channels of initially diotic pure tones by delaying one ear with respect to the other one. Stimuli were presented to the subject via headphones (Sennheiser HD 650, circumaural, open) and the possible head movements were not compensated. For low-frequency pure tones, ITD is an unambiguous information as long as the period of the sound is less than twice the maximum possible ITD, which corresponds to a frequency of about 725 Hz. The ambiguity can be resolved up to about 1500 Hz if head or source movements are possible [3]. So in the present study, ITD cannot provide localization information for $f = 2000$ Hz and may be ambiguous for $f = 1000$ Hz and $f = 1404$ Hz.

These stimuli were to be matched in loudness to diotic references at a loudness level of 40 phon. The reference level was set by placing the test headphones on a dummy head (Neumann KU 100) whose microphones are located at the entrance of the blocked ear canal. Firstly, the sound pressure level was adjusted to 94 dB SPL at 1000 Hz on each ear. Secondly, starting from this point, the sound pressure level was subsequently adjusted so as to produce 40 phon at the entrance of the ear canal for each of the five frequencies under test. The relationship between dB SPL and phon is defined by ISO 226 standard at the position where the center of the listener's head would be, but in the absence of the listener [10]. Corresponding sound pressure levels at the entrance of the blocked ear canal were derived using KU 100 HRTF measurements that account for the pressure transformation from free field to the entrance of the ear canal [11]. Sound pressure levels related to 40 phon at the center of the head in its absence and at the entrance of the blocked ear canal are indicated in Table 1. The duration of each stimulus was 1.6 s. Its onset and offset were smoothed by 100-ms-long raised-cosine functions. Similar stimuli proved to be well lateralized on the basis of interaural time differences [12].

Table 1: Sound pressure levels at the center of the head in its absence according to ISO 226 and at the entrance of the blocked ear canal of the KU 100 dummy head, as a function of the frequency for 40 phon.

f (Hz)	L (dB SPL) ISO 226	L (dB SPL) KU 100
500	43.0	41.5
707	40.6	39.9
1000	40.0	40.8
1414	42.6	42.7
2000	39.2	38.9

2.2 Procedure

Loudness matches were obtained by using a two-interval two-alternative forced choice (2I2AFC) paradigm following a 1-up-1-down rule converging on the point of subjective equality (PSE). This procedure is similar to that used in the previous studies that revealed a significant effect of ITD on loudness [4, 5]. In each trial, a test sound (stimulus including interaural differences) and a reference sound (diotic stimulus at 40 phon) lasting 1.6 s each were consecutively presented in random order with a 500-ms pause in between. The subject's task was to indicate whether the first or the second sound was perceived as louder, regardless of any other perceived difference. The instructions were given both orally and in written form. The subject responded by clicking a button on a MATLAB graphical user interface.

The starting level of each test sound was randomly set 10 dB above or below the level of the reference sound (defined at the entrance of the blocked ear canal in Table 1) to provide a clearly noticeable loudness difference at the beginning of the matching process. The sound pressure level of the test sound was stepwise varied from trial to trial depending on the subject's response. It was lowered when the subject judged it to be louder, whereas it was increased when the subject judged the reference to be louder. The step size was initially set to 4 dB and was decreased to 1 dB after two reversals (a reversal denotes a change in direction in the matching process). For each test sound, the adaptative sequence was ended at the eighth reversal. The arithmetic mean of the levels at the last six reversals was used to derive the PSE of the test sound with respect to its reference.

The 30 adaptive sequences related to the experimental conditions (5 frequencies, 6 ITDs) were randomly reordered. From the subject's point of view, each test appeared thus as a succession of unrelated paired comparisons of loudness. The subject sat in an audiometric booth and was asked to place the test headphones comfortably over his ears and to not modify this position once the test had started. The test lasted approximately 1 h and was preceded by a 3-min pretest to familiarize the listener with the task and the answering interface. Twenty sound engineering students (Bachelor's and Master's degree) from the University of Brest took part in this experiment and were remunerated for their participation. The subjects (six women and fourteen men, with ages ranging from 20 to 22 years) had normal hearing thresholds ($\leq$ 10 dB HL) based on an audiogram taken in the month preceding this test. None of them had particular experience in laboratory listening tests.

2.3 Results and discussion

For each frequency, a repeated-measures analysis of variance was carried out to assess the effect of ITD on loudness matches (results are collected in Table 2). The PSE is presented as the difference between the matched level and the reference level (see Table 1 for corresponding SPL values). The PSE is then expected to be 0 dB when ITD = 0 μs as the test and reference sounds are identical in this case.

Table 2: Results of repeated-measures analysis of variance, as a function of the frequency.

f (Hz)	$F(5, 95)$	p-value
500	3.041	**0.014**
707	1.592	0.170
1000	0.656	0.657
1414	1.613	0.164
2000	0.435	0.823

At 500 Hz, ITD had a significant effect on loudness ($F(5, 95) = 3.041$; $p = 0.014$). As can be noted from Fig. 1, the PSE appears negative when the ITD is different from zero. It can be inferred that related stimuli would have been perceived as louder than their respective reference if presented at the same physical level. The PSE obtained for 772 μs is significantly different from the one obtained for a zero ITD ($p < 0.001$ according to a two-tailed Fisher's LSD test). It decreases here by around 1 dB which is in agreement with the decrease of 1.25 dB that was observed at 200 and 400 Hz for the same ITD in previous studies [4, 5]. The effect is small but still superior to the minimum perceptible change in sound pressure level which is around 0.5 dB for a 500-Hz pure tone at 40 phon [13]. Previous results [4] indicated that the PSE tended to decrease with increasing ITD and that it became significantly lower than the reference (zero ITD) for 669 μs and 772 μs, corresponding respectively to $\theta_{inc} = 60°$ and $\theta_{inc} = 90°$ according to Eq. (1). This is partially confirmed here as the PSEs obtained for 607 μs and 717 μs are significantly

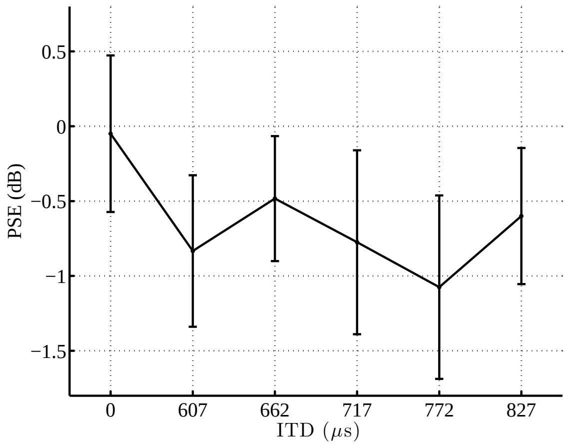

Figure 1: Mean PSE as a function of ITD at 500 Hz, with 95% confidence intervals.

different from the one obtained for a zero ITD (with $p = 0.007$ and $p = 0.012$ respectively), but this does not hold for 662 μs ($p = 0.132$). No significant difference was either found between zero and 827 μs ($p = 0.057$). The latter ITD value is notably higher than the maximum possible value at this frequency according to Eq. (1). It has been shown that when the ITD is higher than a quarter period [14], or close to a half period [15], a diffuse image or even two distinct images may be perceived. At 500 Hz, 827 μs is higher than $T/4$ and close to $T/2$. As this could lead to the perception of diffuse or multiple images delayed in time, other complex mechanisms related to auditory organization processes (auditory scene analysis) might take place prior to loudness computation [16].

As indicated in Table 2, ITD had no significant effect on loudness at any of the four other frequencies under study. This was observed for frequencies where ITD could respectively provide unambiguous (f = 707 Hz), ambiguous (f = 1000 Hz and f = 1404 Hz) or unusable localization information (f = 2000 Hz). As can be noted from Fig. 2, where PSEs are depicted in the same range as in Fig. 1, the differences at these frequencies are much lower than those reported at 500 Hz. The maximum difference that can be observed between a PSE obtained for a non-zero ITD and a PSE obtained for a zero ITD falls between 0.4 and 0.6 dB. It was previously hypothesized that ITD could help separate the signal from internal noise at low loudness levels [4, 5]. These results seem to indicate that it would only hold up to 500 Hz as the differences observed beyond this frequency are not significant and not likely to improve the separability. Therefore, the contribution of ITD to directional loudness appears significant only up to 500 Hz and significant loudness variations that would be observed at higher frequencies would rather be caused by at-ear pressure modifications, which significantly increase above 500 Hz [3]. The fact that directional

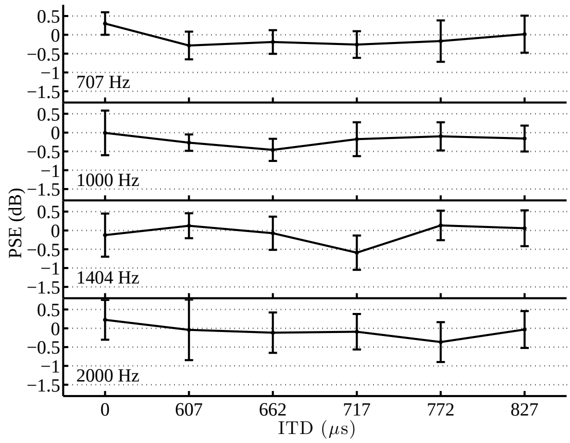

Figure 2: Mean PSE as a function of ITD at 707 Hz, 1000 Hz, 1404 Hz and 2000 Hz, with 95% confidence intervals.

loudness sensitivity at higher frequencies is underestimated by model predictions [6] may not be explained by an effect of time differences on binaural loudness summation (as hypothesized in section 1) but more probably by a different binaural interaction (i.e. contralateral inhibition) for stimuli exhibiting different levels at the two ears. As a result, it should be emphasized that ITD has no significant effect on loudness for frequencies above 500 Hz even if it may provide useful information about the source localization (e.g. at 707 Hz). This finding goes in line with the results establishing that the effect of ITD remained the same when the lateralization was compensated by an opposite ILD [5] and support the assumption that the effect of ITD on loudness is not related to the localization process.

3 Conclusion

The results of the present study confirm the effect of ITD on pure-tone loudness previously observed at 200 and 400 Hz, at a low loudness level. This effect proved to be significant at 500 Hz but not at higher frequencies. At 40 phon, ITD may thus contribute to the phenomenon of directional loudness only up to this frequency and loudness variations with the source azimuth that might be observed above would then be due to modifications of at-ear pressures. Moreover, the fact that ITD has no effect on loudness at 707 Hz, where it is still an unambiguous localization cue, confirms that the effect on loudness is caused by the ITD itself rather than by the related localization. These statements both indicate that directional loudness is not likely to be caused by the direction itself but rather by the modifications (pressure and time) that affect a stimulus coming from a given direction.

Acknowledgement

The authors wish to thank the staff and students from the “Image & Son” department from the University of Brest for participating in this experiment.

References

[1] D. W. Robinson, L. S. Whittle: The loudness of directional sound fields. Acta Acust united Ac **10** (1960) 74–80.

[2] V. P. Sivonen, W. Ellermeier: Directional loudness in an anechoic sound field, head-related transfer functions, and binaural summation. J Acoust Soc Am **119** (2006) 2965–2980.

[3] B. C. J. Moore: Space perception. – In: An introduction to the psychology of hearing. Sixth edition. Brill, Leiden, The Netherlands, 2012, 245-250.

[4] V. Koehl, M. Paquier: Loudness of low-frequency pure tones lateralized by interaural time differences. J Acoust Soc Am **137** (2015) 1040–1043.

[5] V. Koehl, M. Paquier, E. Hendrickx: Effects of interaural differences on the loudness of low-frequency pure tones. Acta Acust united Ac **101** (2016) 1168–1173.

[6] B. C. J. Moore, B. R. Glasberg: Modeling binaural loudness. J Acoust Soc Am **121** (2007) 1604–1612.

[7] G. F. Kuhn: Model for the interaural time differences in the azimuthal plane. J Acoust Soc Am **62** (1977) 157–167.

[8] R. S. Woodworth: Hearing. – In: Experimental psychology. Holt, New York City, NY, USA (1938) 501–539.

[9] N. L. Aaronson, W. M. Hartmann: Testing, correcting, and extending the Woodworth model for interaural time difference. J Acoust Soc Am **135** (2014) 817–823.

[10] ISO 226: Acoustics – Normal equal-loudness-level contours. International Organization for Standardization, Geneva, Switzerland, 2003.

[11] B. Bernschütz: A spherical far field HRIR/HRTF compilation of the Neumann KU 100. Proceedings of AIA-DAGA joint Conference, Merano, Italy, 2013, 592–59.

[12] P. X. Zhang, W. M. Hartmann: Lateralization of sine tones–interaural time vs phase. J Acoust Soc Am **120** (2006) 3471–3474.

[13] H. Fletcher: Minimum perceptible changes in frequency and sound pressure level. – In: Speech and hearing in communication. Second edition. Van Nostrand, Princeton, NJ, USA, 1953, 144–152.

[14] W. A. Yost: Lateral position of sinusoids presented with interaural intensive and temporal differences. J Acoust Soc Am **70** (1981) 397–409.

[15] B. M. Sayers: Acoustic-image lateralization judgments with binaural tones. J Acoust Soc Am **36** (1964) 923–926.

[16] S. McAdams, M. C. Botte, C. Drake: Auditory continuity and loudness computation. J Acoust Soc Am **103** (1998) 1580–1591.