

COVER PAGE

Document downloaded by @DAEL

Thu Nov 5 01:49:46 2026
For personal use

When automatic English translation is provided, only the original document is authentic.
The EAA cannot be held responsible of any translation error

Bibliographical reference

Perceptual Impact of Distance on High-Speed Train Sound Quality, Jonathan Terroir and Catherine Lavandier, Acta Acustica **vol. 100** (Number 2), 2014, pp. 328-340

DOI
https://doi.org/10.3813/AAA.918712

Acta Acustica united with Acustica
Vol. 100 (2014) 328 – 340

DOI 10.3813/AAA.918712

Perceptual Impact of Distance on High-Speed Train Sound Quality

Jonathan Terroir, Catherine Lavandier
Laboratoire Mobilités, Réseaux, Territoires et Environnement, Université de Cergy Pontoise,
95000 Cergy Pontoise, France. [jterroir, catherine.lavandier]@u-cergy.fr

Summary
This work is part of a French-German PREDIT/DEUFRAKO project conducted between 2006 and 2009, entitled "*Transport noise, influences of the temporal structure on annoyance, cognitive performances and sleep*." The present study focuses on the sound quality of six high-speed train pass-bys recorded at several distances from the track. A traditional psychophysical approach was used to carry out perceptual experiments, which involved twenty participants in listening tests, based on a paired comparison method. For each pair, dissimilarity was assessed on an eleven-point scale, combined to a preference judgment and a written explanation. Data were analyzed using an INDividual multidimensional SCALing (INDSCAL) for perceptual characterization, a statistical analysis for preferences and a linguistic analysis for explanations. Fluctuation strength and roughness appeared to be related to the first discriminating axis, whereas the second one appeared to be linked to the pass-by duration, the spectrum and the perceived level. Besides, when analyzing the preference scale, the spectrum and duration properties did not appear to be relevant. This is presumably due to changes in strategy (intra and inter subjects), concerning the temporal envelope and the spectrum, leading to a cancellation of their influence on the preference scale.

PACS no. 43.50.Lj

1. Introduction

When talking about quality of life, noise (and more specifically transport noise) appears as a huge concern of European households [1]. Among all noise sources, railway is considered as less annoying than aircraft and road traffic (see for example [2, 3, 4]), with a limited impact on air pollution. Therefore, as part of the efforts towards sustainable development, politicians propose to develop this type of transport for passengers as well as for goods, in an attempt to also control noise annoyance. Railway traffic is consequently bound to increase in the next years as well as the number of people affected. Among the different types of trains, high-speed trains (French TGV) have to be studied because (1) in the near future their number is supposed to increase more than the others, (2) new TGV tracks will be created to expand the covered network and, (3) high-speed trains seem to be more annoying than conventional trains [5].

A lot of works have been dedicated to the reduction of train-related noise, such as aerodynamic noise (e.g.: [6, 7, 8]), track-wheel interaction noise (e.g.: [9, 10, 11]), rolling noise (e.g.: [12, 8]), mechanical noise sources (e.g.: [13]), and all these technical improvements are systematically evaluated at an international standard [14] location (7.5 m from the railway track). However, a recent study has shown that for a given noise level, the distance from the track and the rise speed of the temporal evolution of the sound level had a significant impact on sound perception [15]. For all these reasons, SNCF (Société Nationale des Chemins de fer Français / the French National Railway Company) and DB (Deutsche Bahn / the German National Railway Company) decided to have a look at sound features, focusing on the temporal effects (more than on spectral ones) of distance, in a large PREDIT/DEUFRAKO project entitled "*Transport noise, influences of the temporal structure on annoyance, cognitive performances and sleep*." The results presented in this study are part of this project and deal with single high-speed train pass-bys with an approach based on sound quality. Blauert and Jekosch [16, 17] define sound quality as "the result of an assessment of the perceived auditory nature of a sound with respect to its desired nature." [17] The perceived nature depends on the perceived quality features of the sounds. These features are the recognizable characteristics that can be observed or detected and denominated by listeners. The desired nature of a sound corresponds to the "totality of features as projected by individual expectations." [17] This sound-quality approach has already been used for road traffic and railway studies e.g.: [18, 19, 20, 21], as well as for aircraft flyovers [22]. In Barbot *et al.*'s paper [22] the authors developed a protocol where the perceived nature of aircraft sounds was studied through a dissimilarity test

Received 5 March 2012,
accepted 27 October 2013.

© S. Hirzel Verlag · EAA

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

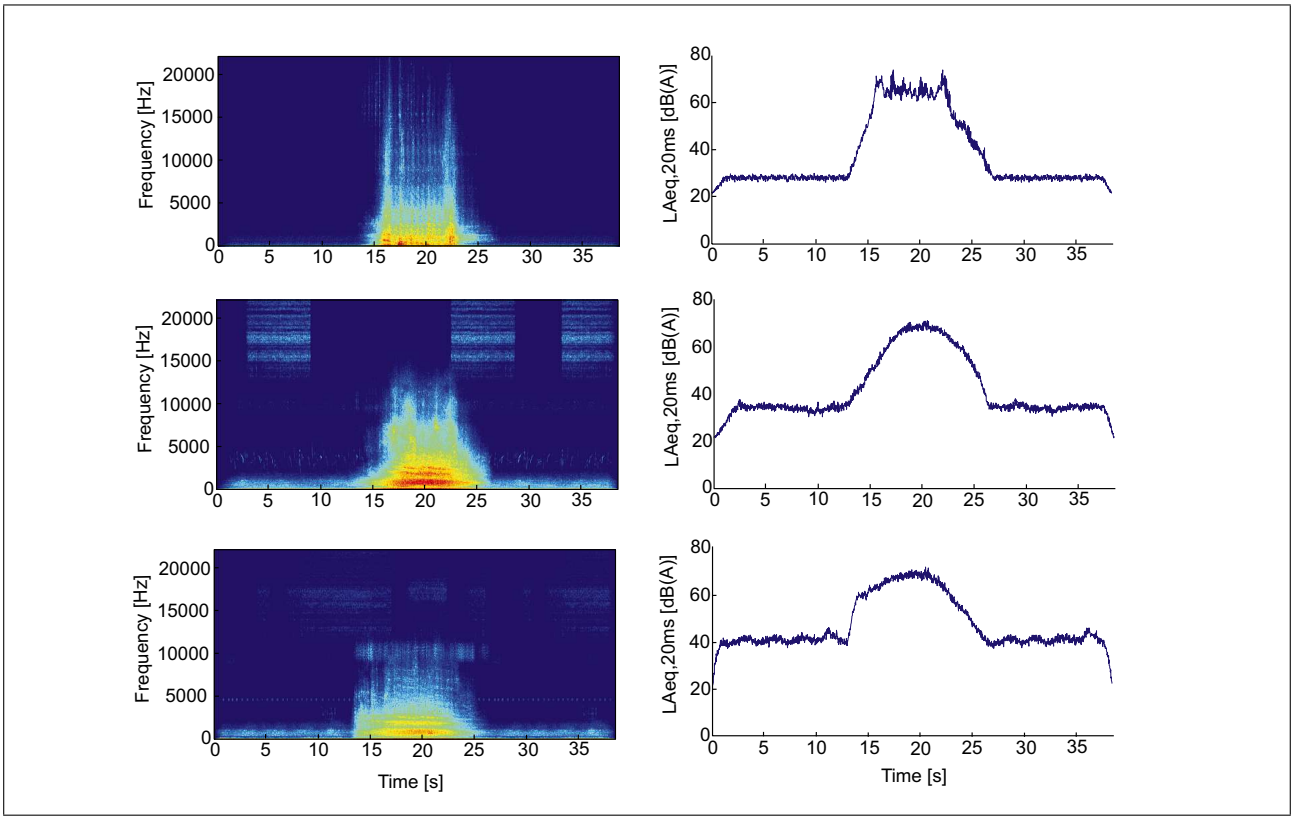


Figure 1. Spectrogram (left) and temporal shape (right) of pass-bys recorded at 7.50 m (top), 50 m (center) and 100 m (bottom) from the track after equalization in relation to the $L_{Aeq,38s}$ level. The background level differences are linked to the $L_{Aeq,38s}$ equalization. Spectrogram and temporal shapes were generated with LEA® software (Genesis S.A.).

and the desired nature was studied through a preference test. The features of the sounds were then denominated with the help of verbal descriptions. A similar method has been used in the present study.

The aim of this research is to provide answers to the following questions: (1) what is the impact of distance on the sound quality of high-speed train pass-bys? (2) What is the impact of distance on preference judgments? Therefore, sounds of passing trains were recorded at three different distances from the track (7.5 m, 50 m and 100 m). In order to compare the impact of distance variations to a global sound level variation, it was decided to artificially decorrelate the global level and the other acoustic features of the sounds. The three distance recordings were therefore equalized in perceived loudness, and – in order to conceal the purpose of the study – an additional level variation of 2 dB was introduced in the subjective tests. The tests were carried out using a traditional psychophysical approach, similarly to Barbot *et al.*'s protocol [22]. Participants were involved in listening tests, based on a paired comparison method. They were asked for dissimilarities and preference judgments and had to justify their preference choices. The factors extracted from the data analysis were perceptually and objectively characterized using the verbalizations and acoustic indicators (fluctuation strength, roughness, duration of the pass-by, rise speed, sound level and spectrum indicators). The role of each factor in the preference scale is discussed.

2. Operating method

2.1. Stimuli

2.1.1. Recordings

A campaign has been conducted to record French passenger high-speed trains (TGV), with the help of the project partner SNCF. This was first initiated between September and October 2006 and completed in June 2007 in Chagny (France). Recording level calibration was made with a B&K4188 microphone, while ambisonic B-format recordings (at a 44.1 kHz sampling frequency) were simultaneously made with a Soundfield ST250 microphone. A binaural conversion was then performed for the needs of the experiment [23, 24]. Stimuli were successively recorded at 7.5 m (control distance for regulations [14]), 50 m and 100 m from the track at a height of 1.60 m. Recordings were done perpendicularly to the track. During the recording sessions, the weather conditions were kept as similar as possible (temperature was between 15 °C and 20 °C with no wind). For each measurement distance one signal was selected. Figure 1 shows the spectrograms and the temporal shape of the selected signals after equalization in relation to the $L_{Aeq,38s}$ level. The high-frequency noise (between 15 kHz and 20 kHz) that can be observed on spectrograms is due to the presence of insects. The pulsed noise that can be observed around 5 kHz is due to the presence of birds. These natural elements remain discreet during the whole playback and are completely masked during the

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

train pass-bys. They have no impact on the acoustic measurements.

The recording distance of 100 m was selected because of the specifically hilly topography of the recording environment. In a flat landscape, the (dis)appearance of the train is less abrupt than in a hilly environment. The influence of the landscape's topography on the time evolution of the sound pressure level was also described by Lee *et al.* [25] In the present work, the 100 m recording was made in a hilly landscape. Thus, a hill was partially physically shielding the arrival of the train and the temporal shape shows a two-stage slope, as it can be clearly observed in the bottom right panel of Figure 1 (see also right panel of Figure 6). The shielding effect was then only perceived from a long distance from the tracks. This recording was purposely selected to enable the study of the influence of the temporal rise speed on sound quality assessment.

Each stimulus lasted around 38 seconds and contained only one single temporally centered pass-by. The French Railway Company (SNCF) provided valuable information concerning the pass-bys to ensure that every recorded train was similar and comparable in speed and size (2 powerheads +10 cars; 160 kph).

2.1.2. Stimulus levels

In order to decorrelate the global sound level from the other acoustic features, the three stimuli were first equalized to equal perceived loudness. Previous studies (e.g.: [26, 27, 22]) have led to various results when looking for the best way to equalize time-varying signals (N_5, $N_{\text{mean}} + N_5$, $L_{\text{AMax,1s}}$). More complex loudness models for time-varying sounds have also been developed (Chalupper and Fastl [28], Glasberg and Moore [29]), but as shown by Rennies *et al.* [30] no model seems to be able to provide accurate predictions of the overall perceived loudness for non- stationary complex sounds. Moreover, due to time and budget limitations (this study being part of a larger PREDIT/DEUFRAKO project), it was not possible to do an individual adaptive loudness matching for each subject. Therefore even if this was not the best way to equalize, the perceived loudness of the sounds was equalized on the basis of listening tests with four "expert" researchers of the laboratory. Six different types of equalization ($L_{\text{Aeq,38s}}$, $L_{\text{AMax,100ms}}$, L_{A5}, L_{N}, $L_{\text{NMax,100ms}}$ and L_{N5}) were evaluated through listening tests. The L_{Aeq}, $L_{\text{A,n}}$ and $L_{\text{AMax,100ms}}$ levels respectively correspond to the mean equivalent level, to the level exceeded n% of the time and to the maximum level in dB(A). The L_{N}, L_{N5} and $L_{\text{NMax,100ms}}$ levels respectively correspond to the equal loudness level, to the percentile loudness level exceeded 5% of the time and to the maximum loudness level in phons. Finally, all the experts selected the equalization in relation to the sound equivalent level $L_{\text{Aeq,38s}}$. This choice is homogeneous with European dB(A) standards and is in line with Fastl's and Ugolotti's results [31]. It is worth noting that the equalization level leads to considerable differences in background level (L_{A90}), up to 10 dB(A) (see Table I).

In order to be able to compare the influence of the level with other perceptual effects, three new stimuli were created, including a level variation. A 2 dB step was chosen to ensure that the level change was audible and that it did not affect the temporal and spectral variations. Moreover, and as it has already been shown, the sound level is widely linked to judgments [5, 15]; it was therefore decided to consider only two different sound levels in order to reduce the number of stimuli in the final corpus. The final corpus thus included 6 sounds (3 distances * 2 levels).

2.1.3. Acoustic and psychoacoustic properties of the stimuli

The values of the various indicators were calculated to subsequently correlate them to subjects' responses. These values correspond to the average of the calculations obtained from left and right channels [32].

Numerous levels indicators were included in the analysis: L_{Aeq}, L_{A5}, $L_{\text{AMax,100ms}}$, L_{N}, L_{N5} and $L_{\text{NMax,100ms}}$. The loudness level calculation is based on the loudness value calculation on the basis of the ISO532B standard [33, 34].

SCG_{Max} and SCG_{Mean} are the maximum and mean Spectral Centers of Gravity [26]. The fluctuation strength corresponds to the sound modulations below 20 Hz, while roughness is related to 20 Hz–300 Hz sound modulations. Sharpness is related to the timbre and to the spectral envelope shape, to the frequency distribution of the energy, as well as to loudness. Definitions for fluctuation strength, roughness and sharpness can be found in Zwicker and Fastl [26].

Two indicators were also included in order to temporally characterize train pass-bys: (1) the pass-by duration $T_{\text{LAMax,100ms-10}}$, which corresponds to the time during which the level of the signal is 10 dB below the maximum level $L_{\text{AMax,100ms}}$. The level range was chosen since, in the aircraft industry, an event corresponds to the time during which the level ranges from the maximum level minus 10 dB [35]; (2) the rise speed of the arrival V_{Rise} (in dB(A)/s) of the train (also used by De Coensel [15] and Barbot *et al.* [22] to respectively describe train and aircraft pass-bys). V_{Rise} was calculated between the $L_{\text{AMax,100ms}}$ maximum level of the signal and the $L_{\text{AMax,100ms}}$ minus 10 dB threshold for the considered section (corresponding to the approach of the train) [22].

Values were calculated with the help of LEA® software (Genesis S.A.) except V_{Rise} and $T_{\text{LAMax,100ms-10}}$ which were calculated using an own programme developed with MATLAB.

2.2. Participants

Twenty subjects (7 females, 13 males) were involved in the experiments. The mean age was 20 years old (minimum age: 18; maximum age: 23). All the subjects were students, naïve listeners who were paid for their services. None of them reported hearing difficulties.

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

Table I. Acoustic and psychoacoustic properties of the six stimuli. * denotes an increase of 2 dB.

Recording	7.5 m	50 m	100 m	7.5 m*	50 m*	100 m*
Fluctuation strength (complete stimulus) [vacils]	2.3	0	0.3	2.4	0.1	0.4
Fluctuation strength (central part only 17 s-22 s) [vacils]	1.9	0.1	0.6	1.9	0.1	0.7
Roughness (complete stimulus) [aspers]	3.4	1	1.1	3.4	1	1.1
Roughness (central part only 17 s-22 s) [aspers]	1.1	0.5	0.5	1.1	0.5	0.5
Rising slope of the approach ΔL [dB(A)/s]	15	3.7	2.9	15	3.7	2.9
Passby duration $T_{\mathrm{LAmax.100ms-10}}$ [s]	7	6.5	8.1	7	6.5	8.1
$L_{\mathrm{Aeq.38s}}$ [dB(A)]	60	59.8	58.9	62	61.8	60.9
L_{A5} [dB(A)]	67.3	68.3	67	69.7	70.3	69
$L_{\mathrm{AMax.100ms}}$ [dB(A)]	79.8	75.9	76	81.8	77.9	78
L_{N} [phons]	78.2	76.9	75.7	80	78.9	77.7
L_{N5} [phons]	82.7	81.9	79.7	83.1	82.9	81.5
$L_{\mathrm{NMax.100ms}}$ [phons]	87	83	82	88.7	84.7	84
Background noise level (L_{A90}) [dB(A)]	25.9	34.2	36.2	27.9	36.2	38.2
$\mathrm{SCG_{Mean}}$ (complete stimulus) [Hz]	419	325	163	419	325	163
$\mathrm{SCG_{Mean}}$ (central part only 17 s–22 s) [Hz]	937	967	673	937	967	673
$\mathrm{SCG_{Max}}$ (complete stimulus) [Hz]	1274	1096	794	1274	1096	794
$\mathrm{SCG_{Max}}$ (central part only 17 s–22 s) [Hz]	1196	1088	776	1196	1088	776
Sharpness (complete stimulus) [acum]	1.5	1.5	1.4	1.5	1.5	1.4
Sharpness (central part only 17 s–22 s) [acum]	1.4	1.5	1.4	1.4	1.4	1.4

2.3. Set up

The listening tests took place in a semi-anechoic room. Signals were presented in a binaural format through AKG K 501© high fidelity headphones at a $L_{\mathrm{Aeq,38s}} = 59.5 \pm 0.6$ dB(A) or $L_{\mathrm{Aeq,38s}} = 61.5 \pm 0.6$ dB(A) level via a MATLAB interface, running on a DELL LATITUDE D820 laptop with a M-Audio QUATTRO soundcard. The headphones calibration was made using the in-situ calibration level recordings and an artificial head Neumann KU 100. Pre-listening sessions suggested that no additional filtering was required to equalize the headphone transfer function.

2.4. Test procedure

The experiment was based on the traditional method of paired comparisons. This method has already been used in several previous studies on transportation noise: car sound quality [36, 37], road traffic [18] and aircraft sounds [22]. This experiment included 15 pairs of sounds (the comparison of a stimulus with itself was excluded), and lasted between 45 and 60 minutes. A graphical user interface was developed using MATLAB for both listening and answering purposes (see Figure 2). Once the test started, the subject was autonomous. Nevertheless, the experimenter was available during the whole test if any help was needed.

Pairs could be listened to as many times as desired. Each pair of stimuli was randomly presented to avoid any order effect. For each pair the subject successively (1) estimated the dissimilarity between the two stimuli on a 0 to 10 categorical scale (verbally fixed at the end from 'identical' to 'totally different'), (2) indicated his or her preference on a 'forced' scale (subjects had to indicate one preferred stimulus) and finally (3) verbally justified his or her preference choice. The verbal justification was not guided by the experimenter but left as open as possible (using a simple 'Why?' question). Subjects justified their preference choices with the help of a positive or negative appreciation, or, in some cases, judged both signals to be equally annoying or preferred. It was expressly mentioned before the experiment that the verbal justification was related to the preference choice. It can nevertheless be considered that the features used to justify the preference choices were also involved in the dissimilarity assessment.

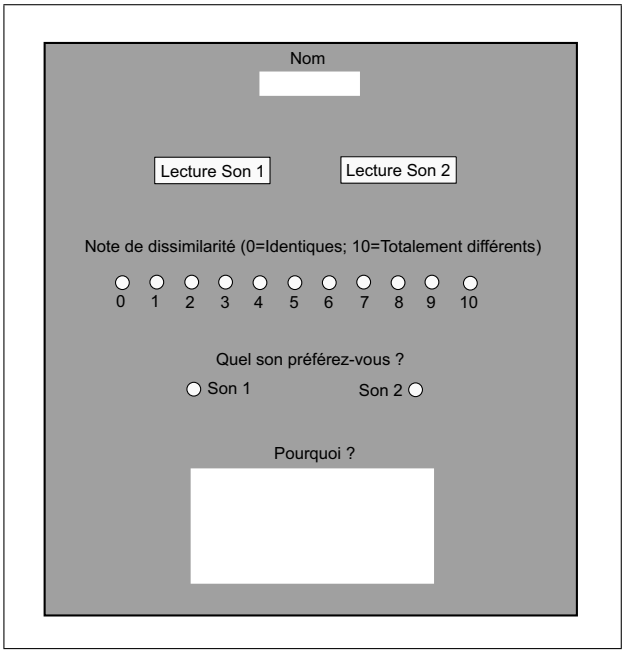


Figure 2. Graphical user interface conducted by the participants during the test. The English translation of the terms used in the interface are (from top to bottom): Name; Play sound 1; Play sound 2; Dissimilarity score (0=Identical; 10= Totally different); Which sound do you prefer ?; Sound 1 – Sound 2; Why?

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

Considering that the dissimilarity assessment can be difficult for untrained people, the subject was first familiarized with the ratings using two very dissimilar signals. These signals were extracted from the corpus of sounds to help subjects build their own scale. The selected stimuli were recordings at 7.5 m+2 dB and at 100 m and were always the same for each subject. Perceptual pretests showed that these recordings were the most dissimilar.

Prior to the test, instructions detailing the procedure (in French language) were given to subjects for them to know what was expected of them. Below is the English translation:

During this test, you will be asked to compare audio recordings of train pass-bys on the basis of paired comparisons. For each pair of sounds, you will have to indicate from your point of view:

1. *how similar or different the signals are on a 0 to 10 scale. 0 corresponding to identical sounds, 10 corresponding to totally different sounds;*
2. *which sound you prefer;*
3. *a justification of your preference choice in the dedicated input box.*

At the beginning of the test, two very dissimilar sounds will be presented. It will enable you to have in mind the dissimilarity that could exist between two very different sounds.

2.5. Evaluations

2.5.1. Dissimilarity data

An INDSCAL multidimensional analysis [38] was conducted. The goal of this method is the construction of a perceptual Euclidean space in which all stimuli are distributed. It is based on an iterative process: objects are placed in a multidimensional space which is adjusted with the dissimilarity data. The distance between two points then represents the degree of similarity between these two points. The multifactor method INDSCAL was chosen because axes of the perceptual space are fixed. This feature makes interpretation easier. The difficulty with this method, compared to more conventional methods such as MDS (Multi-Dimensional Scaling), is that the coordinates of the objects are normalized. Another specificity of this method is that axes are not necessarily orthogonal.

The d_i^{jk} distance between the j and k stimuli for the i subject is calculated with the help of the w_{it} weighting of the i subject on the t dimension and of the x_{kt} component (respectively x_{jt}) of the k stimulus (respectively j) on the t dimension, r being the number of dimensions:

$$d_i^{jk} = \sqrt{\sum_{t=1}^{r} w_{it}\left(x_{jt} - x_{kt}\right)^2}. \quad (1)$$

The INDSCAL analysis provides a spatial representation of factors that drive the perception, the weight of the subjects on these factors and the related explained variance.

2.5.2. Preferences

Preference scale: The Thurstone-Mosteller method [39] was used to classify sounds and to establish the subjects' preference scale. This method assumes that there is a continuous subjective scale representing the measured sensation. The use of this type of scale implies a transitive relationship between the stimuli.

Consistency: For each subject, the "consistency" of the answers was estimated [40]. Consistency implies a transitive relationship between sounds (if A is preferred to B and B to C, then A is preferred to C). A circular error appears when this relationship is not confirmed. The c number of circular errors made by each subject can be calculated as [41]:

$$c = \frac{N_S(N_S-1)(2N_S-1)}{12} - \frac{1}{2}\sum_{k=1}^{N_S} S_k^2. \quad (2)$$

N_S is the number of stimuli and S_k is the number of votes attributed by the subject for the k sound (i.e. the number of times the subject has preferred the k sound). The consistency coefficient is then [40]

$$\xi = 1 - \frac{c}{D}. \quad (3)$$

$D = (N_S^3 - N_S)/24$ if N_S is odd, $D = (N_S^3 - 4N_S)/24$ if N_S is even. Thus, the lower the number of circular errors (i.e. the more ξ is close to 1), the greater the subject's consistency.

Concordance: Kendall's W concordance coefficient [42, 43] enables to estimate the agreement between subjects,

$$W = \frac{12\sum_{j=1}^{N_S} R_j^2}{n_s^2 N_S(N_S-1)} - \frac{3(N_S-1)^2}{N_S^2-1}. \quad (4)$$

n_s is the number of subjects and R_j is the sum of the ranks assigned to the j stimuli [44]. The greater the W value, the greater the concordance. The calculation of Spearman's r_s coefficient on the basis of W makes its interpretation easier. Its calculation is based on the rankings of all possible pairs of subjects [45],

$$\bar{r}_S = \frac{n_s W - 1}{n_s - 1}. \quad (5)$$

The number of N_S stimuli is taken into account during the analysis of r_s, with the help of Spearman's table [42]. This table links critical values of r_s to a given P-value as a function of N_S.

2.5.3. Verbalizations

Subjects were asked to verbally justify their preference choice through the open-ended question: "Why?" (see part II.D). The user interface of the experiment provided a dedicated input box for the answers in a free text form (see Figure 2). The linguistic analysis was specifically focused on the analysis of the descriptive adjectives [46]. This enabled to determine which properties of the sounds were perceived by the subjects and influenced their choices.

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

Table II. Pearson's correlations between INDSCAL axes and acoustic and psychoacoustic properties of the stimuli. The value is followed by ** if the probability testing the statistical significance is below 0.01 and * if it is below 0.05.

ξ	1	0,875	0,75	0,625	0,5	0,375	0,25	0,125	0
c	0	1	2	3	4	5	6	7	8
p	0,02	0,05	0,12	0,21	0,40	0,51	0,77	0,92	1
Number of subjects	7	1	1	4	2	0	5	0	0

3. Results

3.1. Dissimilarities

The analysis of perceptual space and of stimuli properties makes it possible to determine which features had an influence on subjects' ratings. Kruskal and Wish [47] provided an empirical law stating that for a 6 stimuli corpus, the analysis should not exceed two dimensions. Consequently, in accordance with their guideline, the analysis was limited to two dimensions. Figure 3 shows the distribution of the stimuli in the INDSCAL dissimilarity space. The sum of explained variances shows that considering two-dimensions leads to an 11% improvement of the explained variance (62% for the X-axis and 14% for the Y-axis) compared to a one dimension analysis (65% for the single X-axis).

The dissociation between the 7.5 m and the 50 m pass-bys is parallel to the X-axis (which explains 62% of the variance) and the dissociation between the 50 m and the 100 m pass-bys is parallel to the Y-axis (which explains 14% of the variance): the perception of the distance variation needs two dimensions to be explained. Considering the properties of each stimulus (see Table I), the 7.5 m pass-bys can be distinguished from the 50 m ones through many aspects: fluctuation strength, roughness, rise speed V_{Rise}, maximum level or spectral parameters. Pearson's correlations between INDSCAL axes and acoustic and psychoacoustic properties of the stimuli are presented in Table II.

Considering the complete stimulus or only the central part of the pass-by had usually no impact on the significance of correlations between INDSCAL axis and quantities. The only exception was sharpness. But considering that sharpness only varied between 1.4 and 1.5 acum (for the complete sound or only its central part – see Table I), this feature can be considered as constant for all 6 stimuli and correlations related to this feature should be ignored. Thus, in the following, only correlations related to the whole stimuli will be quoted.

The X-axis seems to be linked to the fluctuation strength ($r = -0.99$; $p < 0.001$), the maximum loudness level $L_{\mathrm{NMax,100ms}}$ ($r = -0.86$; $p = 0.03$) and the roughness ($r = -0.99$; $p = 0.003$). Moreover temporal indicators such as the rise speed V_{Rise} ($r = -0.97$; $p = 0.002$) appears to be also significantly correlated to the X-axis distribution. Nevertheless, having a look at Figure 4 (which shows the distribution of fluctuation strength and roughness values for every sound stimulus along the X-axis) shows that the stimuli are not regularly distributed along the regression: two groups can be identified. This implies that with this limited number of stimuli (which is due to the limitation of the test duration), it is not possible to conclude that the X-axis corresponds to a continuous perceptual dimension.

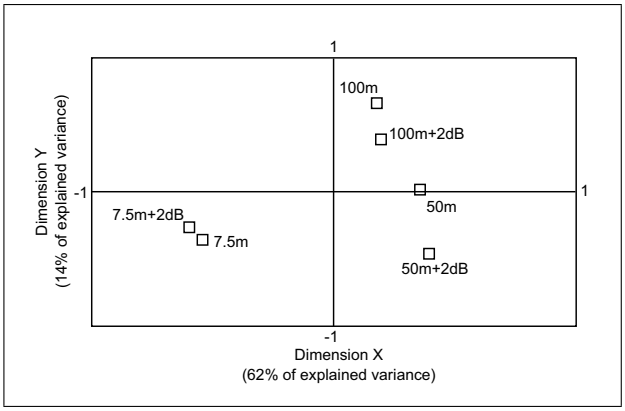


Figure 3. Distribution of the stimuli in the INDSCAL dissimilarity space. The percentage of explained variance for each axis is indicated between brackets.

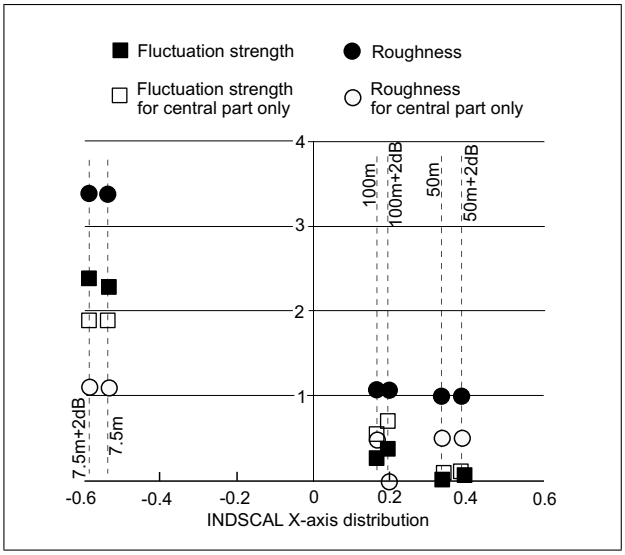


Figure 4. Distribution of fluctuation strength and roughness values for every sound stimulus along the X-axis.

The discrimination between the 50 m and 100 m signals was connected to the Y-axis (see Figure 3) (14% of the explained variance). Table II shows the correlations between this axis and spectral aspects such as the maximum spectral center of gravity $\mathrm{SCG_{Max}}$ ($r = -0.89$; $p = 0.019$), the mean one $\mathrm{SCG_{Mean}}$ ($r = -0.89$; $p = 0.019$) as well

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

Table III. Relationship between consistency coefficients ξ, associated number of circular errors c, limit value p for $N_s = 6$ and number of subjects.

p	0.5	0.2	0.1	0.05	0.02	0.01	0.05
$r_{s\ \text{critical}}$	0.37	0.66	0.83	0.89	0.94	1	1

Table IV. Critical values of r_s with a p probability to be exceeded in absolute terms (for $N_s = 6$) [40].

Positive		Negative	
Temporal structure			
Duration			
short (court)	26	lengthy (long)	8
fast (rapide)	10	slow (lent)	1
brief (bref)	2		
Fluctuations			
continuous (continu)	8	staccato (saccadé)	13
regular (régulier)	7	irregular (irrégulier)	3
fluid (fluide)	3	disordered (brouillon)	2
unchanging (constant)	2	changing (variable)	1
monotous (monotone)	1		
linear (linéaire)	1		
		reccuring (périodique)	1
		repetitive (répétitif)	1
Suddenness			
progressive (progressif)	2	abrupt (brusque)	1
		sudden (soudain)	1
		brutal (brutal)	1
62		33	
95 occurrences			
Intensity			
soft (doux)	11	loud (fort)	25
subtle (discret)	9	noisy (bruyant)	6
		piercing (strident)	4
quiet (calme)	5	intense (intense)	2
		high (élevé)	2
light (léger)	2	sonorous (sonore)	1
27		40	
67 occurrences			
Spectrum			
low-pitched (grave)	10	high-pitched (aigu)	8
harmonious (harmonieux)	2		
		strident (strident)	4
		muffled (sourd)	2
		metallic (métallique)	1
12		15	
27 occurrences			
Spatiality			
far away (loin)	5	close (proche)	2
away (lointain)	4		
distant (éloigné)	3		
12		2	
14 occurrences			

as the duration $T_{\text{LAMax,100ms-10}}$ ($r = 0.85$; $p = 0.03$) and the percentile loudness level L_{N5} (r=-0.94; p=0.005). But all these quantities are linked: the spectrum and temporal shape are inextricably changing with the measuring distance. Thus the Y-axis was linked to both temporal and spectral aspects. Moreover Figure 3 shows that the short segments (100 m; 100 m +2 dB) and (50 m; 50 m +2 dB) are parallel to the Y-axis, showing that this axis is also linked to the loudness variation.

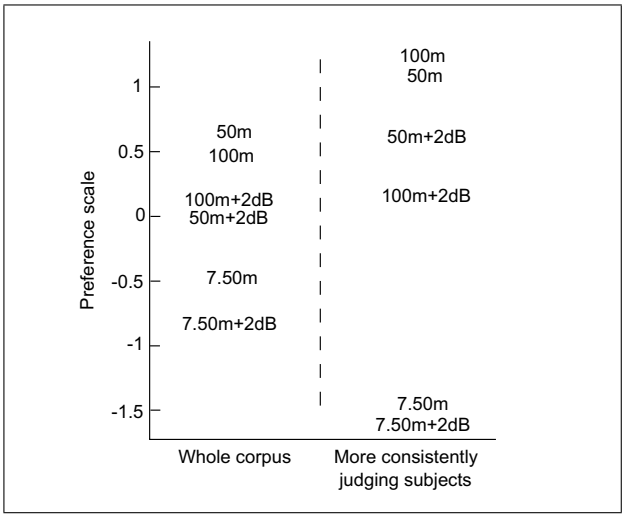


Figure 5. Thurstone preference scale for all subjects (left) and more consistently judging subjects (right).

3.2. Preferences

Consistency: Subjects' consistency and the number of circular errors have been calculated as described in paragraph 2.5.2 (see Table III). Depending on the number of c errors (and therefore on ξ) and on the number of N_s stimuli, Kendall and Babington Smith [40] propose limit values to assess the coherence of subjects' answers. For example, for $N_s = 6$, Table III shows the relationship between the number of circular errors and the corresponding probability.

Table III shows that only 8 subjects made 0 or 1 circular error (which corresponds to $p \leq 0.05$). This low consistency may result from changes in the evaluation policy during the test (the subject focuses successively on various properties of the sound) or from a randomness of the answers (e.g., the stimuli were similar or the test was tiring).

Concordance: Kendall's concordance coefficient [40] enables to estimate the degree of agreement between subjects. The values of R_j and r_s have been calculated (see paragraph 2.5.2). Referring to Spearman's table for $N_s = 6$ [40], Table IV shows the relationship between probability and r_s critical values.

In our case, r_s was 0.47 for the whole corpus and 0.64 when excluding the 5 most inconsistent subjects: in the best case, there is slightly more than a 20% risk of being wrong when saying that subjects are concordant. Therefore the subjects did not choose similar evaluation policies and there was no consensus.

Preference scale: Figure 5 presents the results of the preference task (see paragraph II.E.2) of the whole corpus (left) and the more consistently judging subjects

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

Table V. Pearson correlations between preference scores and the perceptual dimensions from the INDSCAL analysis (X-, Y-axis) as well as acoustic and psychoacoustic stimuli properties. The value is followed by ** if the probability testing the statistical significance is below 0.01 and * if it is below 0.05.

	X	Y
Fluctuation strength (complete stimulus)	-0,99**	-0,45
Fluctuation strength (central part only 17s-22s)	-0,99**	-0,29
Roughness (complete stimulus)	-0,99**	-0,51
Roughness (central part only 17s-22s)	-0,98**	-0,53
Rising slope of the approach ΔL	-0,97**	-0,58
Passby duration $T_{\mathrm{LAMax,100ms-10}}$	0,02	0,85*
$L_{\mathrm{Aeq,38s}}$	-0,21	-0,64
L_{A5}	0,14	-0,52
$L_{\mathrm{Amax,100ms}}$	-0,86*	-0,58
L_{N}	-0,53	-0,79
L_{N5}	-0,43	-0,94**
$L_{\mathrm{Nmax,100ms}}$	-0,86*	-0,7
Background noise level (LA90)	0,91*	0,59
$\mathrm{SCG_{Mean}}$ (complete stimulus)	-0,65	-0,89*
$\mathrm{SCG_{Mean}}$ (central part only 17s-22s)	-0,24	-0,91*
$\mathrm{SCG_{Max}}$ (complete stimulus)	-0,65	-0,89*
$\mathrm{SCG_{Max}}$ (central part only 17s-22s)	-0,55	-0,91*
Sharpness (complete stimulus)	-0,36	-0,92*
Sharpness (central part only 17s-22s)	0,39	0,02

Table VI. Classification of descriptive adjectives related to the temporal structure, spectrum, intensity and spatiality. Positive and negative categories have been included in the table depending on how the adjectives were used by the subjects when stating their preference choice. Adjectives are inscribed in English and in their original language (French). The number of occurrences for each word is indicated.

Temporal structure 17							
Fluctuations	11	Duration	6	Intensity	11	Spectrum	3
staccato	4	short	3	loud	4	low-pitched	2
irregular	2	fast	2	subtle	3	strident	1
regular	1	lengthy	1	quiet	2		
disordered	1			noisy	1		
unchanging	1			soft	1		
fluid	1						
reccuring	1						

(right). Scale values calculated with the Thurstone-Mosteller method [39] are also shown. As suggested by the dissimilarity spaces, 7.5 m signals are discriminated from 50 m and 100 m signals and designated as being the least preferred.

As shown in Table V, preferences are correlated to fluctuation strength, as well as roughness and the rise speed V_{Rise} for temporal aspects. Preferences are also correlated to the INDSCAL X-axis which is best related to the fluctuation strength (see paragraph III.A). Two groups appear to be related to the fluctuation strength values, discriminating 7.5 m signals from 50 m and 100 m ones. As far as the intensity aspects are concerned, preferences are best correlated to the maximum levels ($L_{\mathrm{AMax,100ms}}$, $L_{\mathrm{NMax,100ms}}$). Preferences are not correlated to the spectral indicators.

3.3. Linguistic analysis

Table VI summarizes the descriptive adjectives used to justify preference choices. English translations of the adjectives are also shown. Considering the fact that the participants' mother tongue is French, the original French adjectives are given in brackets. The reported number of occurrences corresponds to the absolute number of times an adjective was mentioned by all the subjects (i.e., the same adjective can be used several times by the same subject to describe several pairs of stimuli). Positive and negative categories were included in the table depending on how the adjectives were used by subjects when stating their preference choice: if the adjective was used to justify the selection of one stimulus, it was assigned to the positive category, otherwise it was assigned to the negative category.

Answers were most related to temporal structure (95 occurrences), followed by adjectives related to the intensity (67 occurrences). The timbre-related adjectives were represented by 27 occurrences. A "spatiality" category was also included: the adjectives "away"/"*lointain*", "distant"/"*éloigné*" and the adverb "far away"/"*loin*" respec-

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

Table VII. Classification of descriptive adjectives used while comparing the 7.5 m and 50 m stimuli or comparing stimuli 7.5 m+2 dB and 50 m+2 dB. These two pairs are only parallel to the INDSCAL X-axis (see Figure 3). The number of occurrences for each word or category is indicated.

Temporal structure 9							
Fluctuations	3	Duration	6	Intensity	11	Spectrum	3
fluid	1	fast	2	loud	3	high-pitched	2
continuous	1	lengthy	2	soft	3	harmonious	1
unchanging	1	brief	1	quiet	2		
		short	1	subtle	1		
				intense	1		
				noisy	1		

tively appeared 3, 4 and 5 times. This feature was taken into account during the sound evaluation but was always associated with other adjectives related to the level, spectrum, temporal slope and duration of the pass-by.

Many adjectives were related to the temporal fluctuation aspect of the pass-by (e.g., "regular"/"*régulier*", "fluid"/"*fluide*"), but the suddenness of the arrival was also perceived (e.g., "abrupt"/"*brusque*", "sudden"/"*soudain*") or the length of the pass-by. As far as spectral properties are concerned, verbalizations mostly focused on the "low-pitched"/ "*grave*" versus "high-pitched"/ "*aigu*" aspect of pass-bys, with a general preference for low frequencies.

3.4. Relationship between verbal responses and dissimilarity ratings

INDSCAL X axis: As shown in Table VII, the analysis of verbalizations dealing with the comparison for pairs 7.5 m/50 m and 7.5 m+2 dB/50 m+2 dB, which are parallel to the INDSCAL X-axis, highlights that this first dimension can be characterized by the perceived fluctuations (11 occurrences) and the perceived loudness (11 occurrences) of the sound. As far as the perceived fluctuations are concerned, this is in accordance with Table II which shows significant correlations between the X-axis distribution with the fluctuation strength and roughness. This is not the case for the perceived loudness. This may be due to the following reasons: 1) for naïve subjects, the intensity notion can be described by more than only the psychoacoustics loudness [22] and 2) the equalization with respect to the $L_{\mathrm{Aeq,38s}}$ results in differences in peak levels within the corpus. Consequently when subjects try to argue about their choices, the intensity notions can be related just as much to the mean level as to the peak level, the emergence from the background noise or its level. This can be observed in Table II with significant correlation between the X-axis distribution and maximum equal loudness levels $L_{\mathrm{NMax,100ms}}$ or the background noise level L_{A90}, which is presumably linked to the emergence.

INDSCAL Y axis: As shown in Table VIII, the analysis of verbalizations dealing with the comparison for pairs 50 m/100 m and 50 m+2 dB/100 m+2 dB, which are parallel to the INDSCAL Y-axis, indicates that this dimension can be characterized by the perceived loudness (11 occurrences) and the perceived duration (6 occurrences) of the sound. As far as the perceived duration is concerned, this is in line with Table II which shows significant correlations between the Y-axis distribution with the duration $T_{\mathrm{LAMax,100ms-10}}$. Concerning the perceived loudness, this result is further evidence supporting the hypothesis that the intensity notion can be described by more than only the psychoacoustics loudness. In this line of argument, one may for example ask whether the descriptive adjective "soft" is specifically linked to the intensity or if it could be related to the spectrum. It is interesting to notice that the adjectives related to the spectrum are less quoted than the temporal ones although both aspects are connected with the distance variation: even if the spectral indicators are correlated with the Y-axis distribution (see Table II), subjects focused more on the temporal aspect.

It can be noticed that the preferences are not correlated to the Y-axis as well as the duration of the pass-by $T_{\mathrm{LAMax,100ms-10}}$: it seems that the perception of the duration has no impact on the preferences, which is in contradiction with the preference verbalizations. Focusing on subjects' explanation for the 50 m/100 m and the 50 m+2 dB/100 m+2 dB pairs, there is no consensus or groups with given strategies: signals' properties are close and lead subjects to successively focus on different aspects of the sound. Overall, subjects do not appear homogeneous in their choices: only 9 subjects kept the same preference for both pairs and this cannot be related to their consistency score. Moreover these subjects do not usually indicate a similar strategy for these two pairs. No subject argues consistently about the spectrum. It is worth noticing that three consistent subjects always cite length as a discriminatory parameter: two subjects preferring the shorter sounds and one the longer ones. While differences between 50 m and 100 m pass-bys have been detected as well as the 2 dB level gain (see Figure 3), the differences between 50 m and 100 m signals appear of less importance for preference judgments (see Figure 5). Besides the impact of the 2 dB gain has a systematic effect on the preference judgments (all stimuli with a 2 dB increased level are less preferred than the original ones). The second temporal dimension that was involved in the dissimilarity analysis does not clearly appear in the preference choices. It can be explained by the fact that changes of strategy (intra and inter subjects) associated to the hypothesis of transitivity

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

Table VIII. Classification of descriptive adjectives used while comparing the 50 m and 100 m stimuli or comparing stimuli 50 m+2 dB and 100 m+2 dB. These two pairs are only parallel to the INDSCAL Y-axis (see Figure 3). The number of occurrences for each word or category is indicated. ALL: Preference scores for the whole corpus, CON: Preference scores for consistent subjects only.

	ALL	CON
INDSCAL X-axis	0,85*	0,94**
INDSCAL Y-axis	0,62	0,55
Fluctuation strength (complete stimulus)	-0,89*	-0,96**
Fluctuation strength (central part only 17s-22s)	-0,84*	-0,94**
Roughness (complete stimulus)	-0,87*	-0,95**
Roughness (central part only 17s-22s)	-0,87*	-0,95**
Rising slope of the approach ΔL	-0,87*	-0,95**
Passby duration $T_{\mathrm{LAMax,100ms-10}}$	0,18	0,14
$L_{\mathrm{Aeq,38s}}$	-0,66	-0,47
L_{A5}	-0,35	-0,12
$L_{\mathrm{Amax,100ms}}$	0,98**	-0,95**
L_{N}	-0,86	-0,74
L_{N5}	-0,71	-0,66
$L_{\mathrm{Nmax,100ms}}$	-0,97**	-0,95**
Background noise level (LA90)	0,73	0,85
$\mathrm{SCG_{Mean}}$ (complete stimulus)	-0,68	-0,70
$\mathrm{SCG_{Mean}}$ (central part only 17s-22s)	-0,36	-0,34
$\mathrm{SCG_{Max}}$ (complete stimulus)	-0,68	-0,70
$\mathrm{SCG_{Max}}$ (central part only 17s-22s)	-0,60	-0,62
Sharpness (complete stimulus)	-0,43	-0,42
Sharpness (central part only 17s-22s)	0,59	0,44

for the Thurstone-Mosteller method lead to a cancellation of their influence on the preference scale.

Rise speed: For the 100 m signals, verbalizations enabled to determine which rise speed corresponded to subjects' perception. It would be logical to observe a decrease of the rise speed with distance. But, as explained in paragraph 2.1.1, for the 100 m recordings a two-stage slope could be observed on the temporal shape on the pass-by during the approach. Figure 6 shows the temporal shape of the pass-by recorded at 50 m (upper part of the figure) and 100 m (lower part of the figure) from the track. For the 100 m recording the first part of the rise (corresponding to slope A on Figure 6) is more abrupt than the rise for the 50 m recording while the second part of the rise (corresponding to slope B on Figure 6) is softer than for the 50 m recording. A systematical analysis of verbal explanations comparing the 50 m and 100 m stimuli showed that slope B, which is the softest one, was the one perceived by the two subjects who specifically focused on this aspect. For example:

- Pair 50 m/100 m +2 dB: "its arrival [is] not too brutal (...): (the 100 m +2 dB sound is) progressive with only one single perceived noise peak" / "son arrivée [n'est] pas trop brutale (...): [le son à 100 m+2 dB est] progressif, avec un seul pic de bruit perçu".
- Pair 50 m/100 m +2 dB: "the approach [of the 100 m +2 dB sound] is less abrupt" / "l'approche [du son à 100 m +2 dB] est moins brusque".

This confirms that the choice of the 10 dB threshold used to calculate the rise speed V_{Rise} seems to be relevant with the feeling of suddenness verbally reported.

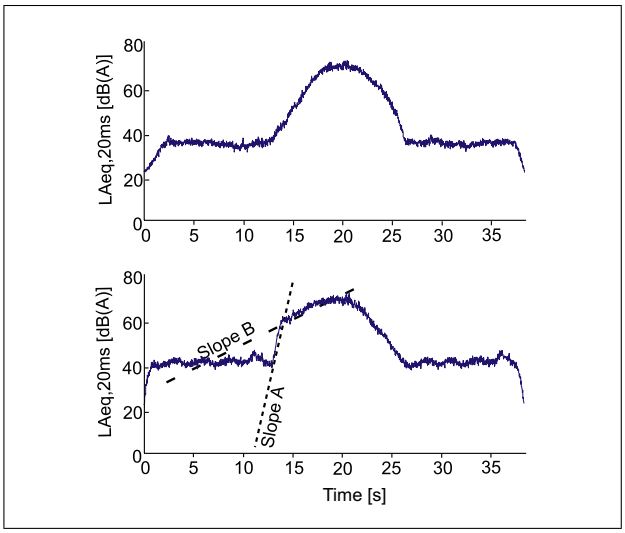


Figure 6. Temporal shape of the pass-by recorded at 50 m (upper part of the figure) and 100 m (lower part of the figure) from the track after equalization in relation to the $L_{\mathrm{Aeq,38s}}$ level. For the 100 m pass-by the two possibly heard slopes A and B being superimposed.

4. Discussion

4.1. Impact of the distance on spectral features

From an objective point of view Lee *et al.* [25] have shown that acting on the measuring distance (45 m, 100 m and 140 m) (which was paired with topography variations in their experiment) for high-speed train sounds has an impact on high-frequencies (the closer from the railway, the

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

more important the high-frequencies) but this effect is not strong enough to impact the measures of sharpness. This is in accordance with the measured values found for the sharpness which varies in the present experiment only from 1.4 to 1.5 acum. Nevertheless the spectral center of gravity (SCG_{Mean}, SCG_{Max}) values seems more sensitive to the distance variations (cf. Table I). From a perceptual point of view the present study shows that the spectrum has no impact on the preference scale for distances between 7.5 m and 100 m. Even if preference is not exactly annoyance, it can be noticed that some previous works have found the same tendency for train pass-bys in the same range of distance. De Coensel *et al.* [15] who studied the effect of the distance (from 25 m to 200 m) of annoyance for high-speed trains did not present any spectral indicator as having an impact on the annoyance (although several temporal or level indicators were found as having a significant impact). From his side Vos [5] only found a small effect of high-frequencies for high-speed and MAGLEV trains (recorded from 25 m to 100 m), describing the "shrilly sound character of a few passages [...] as a cause of the annoyance".

4.2. Impact of the distance on temporal features

4.2.1. Fluctuations

In the present work fluctuations have been identified as being perceived negatively. This has been shown by the linguistic analysis or the correlation between preference scores and the fluctuation strength values. This is in accordance with the conclusions of Namba *et al.* [48] writing that (for train sounds) "large amplitude fluctuations are felt as annoying". Vos [5] also found that for high-level train sounds (from $ASEL = 85$ dB for a 45 seconds sound, i.e. L_{Aeq} around 70 dB(A)) the "heavily pounding" sound was related to annoyance. Concerning other type of transportation sources, the negative impact of fluctuations have already been reported for aircraft sounds (cf. Barbot *et al.* [22] and Schütte *et al.* [49]). The difference between trains and aircrafts are then that fluctuations are mostly noticed during the approach for aircrafts and during the central part of the pass-by for the trains. Similarly Ohshima and Yamada [50] found that the "impulsiveness" of helicopter sound during the approach was correlated to the annoyance. From an objective point of view this is confirmed by the values found for the fluctuation strength and the roughness (cf. Table I). Lee *et al.* [25] also found an impact of the measuring distance of high-speed trains (45 m, 100 m and 140 m) (which was paired with topography variations) on fluctuation strength and roughness.

4.2.2. Duration of the pass-by

The impact of the length of the pass-by on the preference has been found to be unclear: the second dimension of the INDSCAL analysis seems to be explained by the duration whereas this duration is not correlated to the preference scale. This seems related to the fact that the use of in- situ recordings implies relations between several attributes of the sounds. Consequently as it has been described in paragraph III.D changes of strategy associated to the hypothesis of transitivity for the Thurstone-Mosteller method lead to a cancellation of the influence of the duration of the pass-by on the preference scale.

With other psychoacoustic and statistical methods Ohshima and Yamada [50] and Barbot *et al.* [22] respectively found a correlation between the duration and the annoyance or the preference for helicopter and aircraft sounds. The test set up by Ohshima and Yamada [50] was based on magnitude estimation of the annoyance without reference stimulus. For every stimulus "the maximum sound pressure level was kept almost the same and the impulsiveness of sound waveforms was held constant" [50]. Results show a correlation between the duration and the average annoyance. On their side Barbot *et al.* [22] designed pair comparisons tests using systematically the same reference stimulus. Stimuli were equalized towards their maximum level in dB(A). Again a correlation was then found between the preference scale and the duration. The difference of the present work and these two studies raises the impact of equalization on the results. This difference seems due to the fact that for a maximum level equalization, the longer, the bigger the equivalent level although for an equivalent level equalization, the longer, the smaller the maximum level. This shows the impact of the equalization choice on the relation between the duration and the preference or the annoyance.

The question of the quality of the indicator used to describe the pass-by duration ($T_{LAMax,100ms-10}$) could be raised. Ohshima and Yamada [50] considered two methods: the time during which the level ranges 10 dB above the background noise or the time during which the level ranges from the maximum level minus 20 dB. For aircraft sounds Barbot [22] considered the time during which the level ranges from the maximum level minus 10 dB. This last indicator was chosen for this research as it is also used for the EPNL calculation and was then previously validated. It suggests that indicators supposed to characterize specific perceptive dimensions such as duration of a pass-by depend on the type of source (train, aircraft, helicopter, etc.).

4.2.3. Rise speed

The suddenness of the approach (described in the present paper by the rise speed V_{Rise}) was found to be negatively correlated to the preference scores (cf. Table V). This result appears to be in line with previous work dedicated to various sources. Indeed Vos [5] reported a negative impact of the "startle" effect for train sounds while De Coensel *et al.* [15] found a correlation between the rise speed and the annoyance. Similarly Morel *et al.* [21] found that the suddenness of the approach of cars and heavy-trucks has a negative impact on the annoyance. Concerning aircrafts Barbot *et al.* [22] also found a negative impact of the rise speed.

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

4.3. Issues for future research

According to Figure 3 and to the linguistic analysis the level variations for 7.5 m signals were not well perceived. This could be explained by the fact that the relationship between instantaneous and overall loudness for non-stationary sounds is complex and still not perfectly known [51] [52]. In parallel it has been raised in paragraph 3.4 that the psychoacoustic loudness can be inappropriate to describe the intensity notion. For naïve subjects this notion can include peak level, mean level, emergence level from the background noise, etc. It appears then necessary for future work to keep in mind that although the loudness is a useful and robust objective indicator, its relation with perception can be complex.

In the present work tests were based on a forced choice between stimulus A and stimulus B, but as Parizet [53] has shown, this can lead to a low consistency of the results when stimuli are too similar or too difficult to compare. In those cases, subjects give a random answer because of the design of the test. The consistency can be improved when a third "neutral" answer alternative is included. In the present work the linguistic analysis showed that most answers were justified with the help of appropriate arguments and only 22 answers on a total of 300 (i.e. around 7.5%) were explicitly due to the fact that subjects were forced to make a choice. It can therefore be concluded that answers cannot be considered as random but resulting from multiple changes in the evaluation policy. As it has been shown along the results section, for this type of high-speed trains, temporal features appear to be of major importance in the evaluation of the sound quality. But in this study, the perceptual aspects (temporal and spectral as well as suddenness and length) were linked by the structure of stimuli (in-situ recordings). These correlations between perceptual dimensions did not make it possible to clearly understand the effects of temporal structure. To better study the relationship between physical changes and perceptual effects, future studies may consider synthesized signals. The decorrelation of these aspects could help to clearly answer these perceptual questions. For example, a work involving time-compressed or time-stretched recordings was recently carried out on helicopter sounds to study the relationship between annoyance and duration without acting on other features usually correlated with the speed of the pass-by, such as the impulsiveness for example [50]. This study showed a high correlation between annoyance and duration. In a similar vein, it would be interesting to independently assess the influence of the duration, the suddenness, the spectrum or the temporal shape on the sound preference.

5. Summary and conclusion

In the present study the sound quality of six high-speed train pass-bys recorded at several distances from the track was studied using a traditional psychophysical approach based on a paired comparison method. It has shown that:

1. The perceived fluctuations of high-speed trains correspond to the first factor which differentiates pass-bys, recorded at different distances after $L_{\mathrm{Aeq.38s}}$ equalization;
2. For 7.5 m signals (i.e., the most fluctuating pass-bys), the 2 dB level difference is not clearly perceived;
3. The 2 dB level difference has a greater impact on preference than other features linked to the temporal shape (rise speed, duration) or the spectrum;
4. When comparing the 50 m and 100 m signals (after $L_{\mathrm{Aeq.38s}}$ equalization), there seems to be no clear discrimination strategy. Several aspects are then involved: there are several groups of subjects with given strategies which can change during the test. These changes can be similar or vary for every subject. This multiplicity of strategies then leads to a lack of consensus in preferences.

Acknowledgement

This research is part of a program funded by a PREDIT / DEUFRAKO contract. It is supported by the French Energy Agency ADEME. The authors would like to also thank Mrs. Zineb Bouizem for her help and proofreading of the paper.

References

[1] B. Berglund, T. Lindvall, D. H. Schwela: Guidelines for community noise. Technical report, World Health Organization, 2000.

[2] J. M. Fields, J. G. Walker: Comparing the relationship between noise level and annoyance in different surveys: a railway noise vs. aircraft and road traffic. Journal of Sound and Vibration **81** (1982) 51–80.

[3] H. M. E. Miedema, H. Vos: Exposure-response relationships for transportation noise. Journal of the Acoustical Society of America **104** (1998) 3432–3445.

[4] U. Moehler, M. Liepert, R. Schuemer, B. Griefahn: Differences between railway and road traffic noise. Journal of Sound and Vibration **231** (2000) 853–864.

[5] J. Vos: Annoyance caused by the sounds of a magnetic levitation train. Journal of the Acoustical Society of America **115** (2004) 1597–1608.

[6] C. Talotte: Aerodynamic noise a critical survey. Journal of Sound and Vibration **231** (2000) 549–562.

[7] K. Nagakura: Localization of aerodynamic noise sources of shinkansen trains. Journal of Sound and Vibration **293** (2006) 547–556.

[8] C. Mellet, F. Létoruneaux, F. Poisson, C. Talotte: High speed train noise emission: latest investigation of the aerodynamic-rolling noise. Journal of Sound and Vibration **293** (2006) 535–546.

[9] D. J. Thompson, C. J. C. Jones: Review of modeling wheel/rail noise generation. Journal of Sound and Vibration **231** (2000) 519–536.

[10] T. X. Wu, D. J. Thompson: On the rolling noise generation due to wheel/track. Journal of Sound and Vibration **293** (2006) 566–574.

[11] D. J. Thompson, P. E. Gautier: Review of research into wheel/rail rolling noise reduction. Proc. of the Institution of Mechanical Engineers Part F, Journal of Rail and Rapid Transit, 220F, 2006, 385–408.

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only

[12] C. Talotte, P. E. Gautier, D. J. Thompson, C. Hanson: Identification, modeling and reduction potential of railway noise sources. Journal of Sound and Vibration **267** (2003) 447–468.

[13] C. Hanson, B. Barsikow: Noise sources on Amtrak's high speed train. Proc. Internoise, Vol. 2, Nice, France, 2000, 1057–1060.

[14] ISO3095-2005: Railway applications – Acoustics – Measurement of noise emitted by railbouned vehicule. 2005.

[15] B. De Coensel, D. Botteldooren, B. Berglund, M. E. Nilsson, T. De Muer, P. Lercher: Experimental investigation of noise annoyance caused by high-speed trains. Acta Acustica united with Acustica **93** (2007) 589–601.

[16] J. Blauert, U. Jekosch: Sound quality evaluation: A multilayered problem. Acta Acustica united with Acustica **83** (1997) 747–753.

[17] U. Jekosch: Basic concept and terms quality, reconsidered in the context of product-sound quality. Acta Acustica united with Acustica **90** (2004) 999–1006.

[18] G. W. Cermak, P. C. Cornillon: Multidimensional analysis of judgements about traffic noise. Journal of the Acoustical Society of America **59** (1976) 1412–1420.

[19] S. Meunier: Critères psycho-acoustiques des nuisances sonores: évaluation perceptive des bruits de circulation routière et ferroviaire. Acoustique et Techniques **26** (2001) 30–35.

[20] S. Kuwano, S. Namba, M. Florentine, Z. Da Rui, H. Fastl, A. Schick: A cross-cultural study of the factors of sound quality of environmental noise. Journal of the Acoustical Society of America **105** (1999) 1081–1081.

[21] J. Morel, C. Marquis-Favre, D. Dubois, M. Pierrette: Road traffic in urban areas: a perceptual and cognitive typology of pass-by noises. Acta Acustica united with Acustica **98** (2012) 166–178.

[22] B. Barbot, C. Lavandier, P. Cheminée: Perceptual representation of aircraft sounds. Applied Acoustics **69** (2008) 1003–1016.

[23] M. A. Gerzon: General metatheory of auditory localization. Audio Engineering Society Preprint, 3306, 1992.

[24] M. Jot, V. Larcher, J. M. Pernaux: A comparative study of 3-D audio encoding and rendering techniques. Proc. AES 16th International Conference: Spatial Sound Reproduction, Rovaniemi, Finland, 1999.

[25] P. J. Lee, J. Y. Hong, J. Y. Jeon, H. I. Koh: Perception on soundscape with rail road in Korea. Proc. Internoise, Lisbon, Portugal, 2010.

[26] E. Zwicker, H. Fastl: Psychoacoustics: facts and models. Second edition. Springer, 1999.

[27] S. Meunier, A. Marchioni: Loudness of sounds with temporal variable intensity. Proc. Forum Acusticum, Sevilla, Spain, 2002.

[28] J. Chalupper, H. Fastl: Dynamic loudness model (DLM) for normal and hearing-impaired listeners. Acta Acustica united Acustica **88** (2002) 378–386.

[29] B. R. Glasberg, B. C. J. Moore: A model of loudness applicable to time-varying sounds. J. Audio Eng. Soc. **50** (2002) 331–341.

[30] J. Rennies, I. Holube, J. L. Verhey: Loudness of speech and speech-like signals. Acta Acustica united with Acustica **99** (2013) 268–282.

[31] H. Fastl, E. Ugolotti: Technical note: Subjective and physical evaluation of aircraft noise. Noise Control Engeneering Journal **3** (1990) 61–63.

[32] A. Preis, F. Satoh, M. Ishibashi, H. Tachibana: Binaural loudness perception of the time varying environmental stimuli. Proc. Forum Acusticum and 137th ASA, Berlin, 1999, 627–628.

[33] ISO532-1975: Acoustics – Method for calculating loudness level. 1975.

[34] E. Zwicker, K. Deuter, W. Peisl: Loudness meters based on ISO 532B with large dynamic range. Proc. Inter'Noise 85: International Conference on Noise Control Engineering, Vol. II, 1985, 1119–1122.

[35] ICAO: Environmental protection. Vol. 1 aircraft noise - annex 16. International Civil Aviation Organisation, 2008.

[36] P. Susini, S. McAdams, S. Winsberg: A multidimensional technique for sound quality assessment. Acta Acustica united with Acustica **85** (1999) 650–656.

[37] R. P. Leite, S. Paul, S. N. Y. Gerges: A sound quality-based investigation of the HVAC system noise of an automobile model. Applied Acoustics **70** (2009) 636–645.

[38] J. D. Caroll, J. Chang: Analysis of individual differences in multidimensional scaling via n-way generalyzation of "ecart-young" decomposition. Psychometrika **3** (1970) 283–329.

[39] L. L. Thurstone: WA law of comparative judgment. Psychological Review **34** (1927) 273–286.

[40] M. G. Kendall, B. Babington Smith: On the method of paired comparisons. Biometrika **31** (1940) 324–345.

[41] S. I. Gass: Tournaments, transitivity and pairwise comparison matrices. Journal of the Operational Research Society **49** (1998) 616–624.

[42] M. G. Kendall: Rank correlation methods, 4th edition. Griffin, London, 1975.

[43] P. Legendre: Species associations: the Kendall coefficient of concordance revisited. Journal of Agricultural, Biological and Environmental Statistics **10** (2005) 226–245.

[44] D. C. Howell: Fundamental statistics for the behavioral sciences. Wadsworth Publishing, 2007.

[45] J. H. Zar: Biostatistical analysis (4th ed.). Prentice Hall, Upper Saddle River, New Jersey, 1999.

[46] C. Kerbrat-Orecchioni: L'Enonciation. Armand Colin, Coll. U, Paris, 2002.

[47] J. B. Kruskal, M. Wish: Multidimensional scaling. Sage University Paper series on Quantitative Applications in the Social Sciences, number 07-011. Sage Publications, Newbury Park, CA, 1978.

[48] S. Namba, S. Kuwano, T. Kato: Trade-off effect between the number of events and LAeq in the evaluation of train noise. Proc. Inter-Noise, Liverpool, England, 1996, 2199–2204.

[49] M. Schütte, U. Müller, S. Sandrock, B. Griefahn, C. Lavandier, B. Barbot: Perceived quality features of aircraft sounds: An analysis of the measurement characteristics of a newly created semantic differential. Applied Acoustics **70** (2009) 903–914.

[50] T. Ohshima, I. Yamada: Study on the effect of sound duration on the annoyance of helicopter noise by applying a technique of time compression and expansion of sound signals. Applied Acoustics **70** (2009) 1200–1211.

[51] P. Susini, S. McAdams, B. Smith: Loudness asymmetries for tones with increasing and decreasing levels. Proc. ICAD, Limerick, England, 2005.

[52] M. Florentine, A. Popper, R. F. (ed): Loudness. Springer-Hand Book of auditory research, Springer, 2011.

[53] E. Parizet: Paired comparison listening tests and circular error rates. Acta Acustica united with Acustica **88** (2002) 594–598.

Document downloaded by @DAEL, 2026-11-05 01:49:48 - personal use only