



HAL
open science

Evaluation of speech intelligibility in open-plan offices

Patrick Chevret, Ange Ebissou, Etienne Parizet

► **To cite this version:**

Patrick Chevret, Ange Ebissou, Etienne Parizet. Evaluation of speech intelligibility in open-plan offices. InterNoise 2012, 2012, New-York, United States. pp.1. hal-00994393

HAL Id: hal-00994393

<https://hal.science/hal-00994393>

Submitted on 21 May 2014

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.



Evaluation of speech intelligibility in open-plan offices

Patrick Chevret ^{a)} and Ange Ebissou ^{b)},
Institut National de Recherche et de Sécurité
1, rue du Morvan
CS60027, Vandoeuvre cedex, France.

Etienne Parizet ^{c)}
Laboratoire Vibrations Acoustique, INSA Lyon,
25 bis avenue Jean Capelle,
Villeurbanne, F-69621 Lyon, France.

In open-plan offices, ambient noise made of intelligible conversations is generally perceived as one of the most important annoyance for tasks requiring concentration efforts. This annoyance has been proved to lead to a decrease of task performance and to health troubles for people in the mean and long term (tiredness, stress, etc.) Consequently, the improvement of working conditions should pass by the evaluation of speech annoyance that could give rise to recommendations on the arrangement of the open-plan office and on acoustical treatment of walls, ceiling and furniture. In that context, a standard SRT (Speech Reception Threshold) measurement campaign was set up, in order to evaluate the effect of speech ambient noise on intelligibility. For the trial purpose, a large set of syllabically-balanced French sentences has been constituted from Combesure, Fourier and Hint sentences. The ambient noise was made of both synthetic (modulated white noise, speech like modulated white noise, inversed speech noise) and real recordings in open-plan offices. In parallel, a second measurement campaign with subjects was conducted whose aim was to study the influence of intelligibility on the performance and effort in the realization of a seriation task. Results showing some links between intelligibility, performance and effort are presented in the paper.

1 INTRODUCTION

Noise is reported in many studies as the most disturbing physical nuisance in open-plan offices. It has been noted that speech noise is particularly annoying for office workers¹. A need has thus emerged amongst occupational health specialists for a way to assess the acoustical quality of an office with regards to this problem. At INRS (French National Institute for

^{a)} email: patrick.chevret@inrs.fr

^{b)} email: ange.ebissou@inrs.fr

^{c)} email: etienne.parizet@insa-lyon.fr

Occupational Health), this concern was integrated in the research topics several years ago, starting with studies on the impact of continuous phone conversations in call centers. More recently, projects have been initiated on the effects of intelligibility on performance and workload in open-plan offices. This is the topic of the paper, which mainly focuses on the description of experimental campaigns and the associated procedures, the results and the orientation for the next years.

2 SPEECH INTELLIGIBILITY

Speech intelligibility is the main factor used to describe acoustic comfort in open-plan offices. It is commonly accepted, supported by that the speech intelligibility may be impacted by several factors:

1. The levels of the speech and the background noises (BGN), the latter being a mixture of equipment noise and speech noise. Indirectly, the ratio (SNR) between the two components (speech and BGN) is a main effect.
2. The arrangement of the offices and the quality of the materials used for the walls, the ceiling and the furniture inside the working site. These elements have an effect on the sound attenuation and the sound transmission between the speakers and the listener.
3. The type of activity and the number of persons that work in the space which determine the signal structure of the background noise, in particular its modulation rate.

In this section, we describe the set up of an experiment, whose objective is to evaluate the effect of the background noise modulation on sentence intelligibility. This is done by the determination of the Speech Reception Threshold (SRT) which is defined as the signal to noise SNR value for which 50 % of the sentences are 100% understood. Recently, such experiences have been conducted by Rhebergen² on simulated and natural signals (true voices). In this paper, we show that the use of the Plomp and Mimpen³ procedure is adequate for the ranking of open-spaces in terms of intelligibility.

2.1 The experiment

The test procedure

The test is based on the SRT procedure developed by Plomp and Mimpen. Basically, it consists in determining the SNR value for which 50% of the sentences are totally understood by the listeners. The value also corresponds to an intelligibility score of 50%, and determines the inflection point of the psychometric function.

SRT procedure consists in presenting sentences in noise to listeners with an adaptative SNR value which depends on the listener response (correct or incorrect). First, the SNR value is high (typically 8 dB), so that the listener has no doubt for its answer. Then, the SNR value is decreased by step of 4 dB until the response is incorrect. From this step, the SNR value is adapted with a value of +2dB when the response is wrong and -2dB when it is correct. For each series, 20 phonetically balanced sentences are presented to the listener. The lasts 15 values of SNR are averaged to give the SRT value.

The test conditions

For practical reasons, measurements have been conducted in two separate places so that two groups of people are constituted. The first group is installed in a small silent room (26 dBA of background noise) of dimensions 4m x 4m x 2.5m. This group is composed of 23 students aged from 21 to 37 years. They all had a normal hearing, as measured using an audiometric calibration. The second one was performed in a large semi-anechoic room (22 dBA of background noise) of dimensions 10m x 7m x 7m with 18 persons aging from 14 to 61 years. No hearing test was performed for this group but none of the people declared any hearing impairment.

Except for the room and the loudspeaker, the test conditions were exactly the same for the two groups. During the test, the listener is placed at a distance of about 1.5m from of a unique loudspeaker. Samples of sound, composed of a mixture of the target sentence and background noise, are presented.

The sound level of the sample is always held at a level of 60 dB at 1m from the loudspeaker. The SNR between the target sentence and the background level is adjusted according to the answer of the listener (repetition of the sentence), following the SRT procedure of Plomp and Mimpen.

Listening material

Lists of sentences in French (299 sentences in total), designed for audiometry purposes, were used as the target speech signal. The recordings⁴ were provided by the Collège National d'Audiométrie (French Organization for Audiometry). The sentences are the so called Combescure, Fournier and HINTS sentences. They are pronounced by a male voice. Background noise was constituted either from the lists of sentences or synthetic signals according to the following description:

1. Stationary white noise (LTASS-WN),
2. White noise 100% sinusoidally intensity-modulated with 8Hz frequency modulation (LTASS-SIM8Hz),
3. Sentences pronounced by 1 female voice (LTASS-MOD1pers),
4. Mixture of sentences pronounced by 3 female voices (LTASS-MOD3pers),
5. White noise modulated by the envelop of sentences pronounced by 1 female voice (LTASS-WNMOD1pers),
6. White noise modulated by the envelop of a mixture of sentences pronounced by 3 female voice (LTASS-WNMOD3pers),
7. White noise modulated by the envelop of a mixture of sentences pronounced by 5 female voice (LTASS-WNMOD5pers),
8. Sentences pronounced by 1 female voice and time inversed (LTASS-MOD1pers-inv),
9. Ambient noise recorded in open-plan office 1 (LTASS-OPO1),
10. Ambient noise recorded in open-plan office 2 (LTASS-OPO2),

The reason for choosing a male voice for the target sentences and a female one for the background noise is to avoid confusion between the target sentences and the disturbing ones. The labeling LTASS means that all the samples are equalized according to the long term averaged

speech spectrum; the male one for the target samples and the female one for the background noise.

All of these background noises are presented to the listeners in a randomized order. For each BGN and each listener, one SRT value is obtained which is used for the ANOVA described in the next section. The duration of the test is about 30min, this value mainly depending on the reactivity of the tested person.

2.2 The results

In this paper, two outputs of the test are analyzed through a standard Analysis of variance (ANOVA). First, the consistency is studied by presenting to the listeners the same background noise type at the beginning and at the end of the test (the type is different from one listener to the other). Second, the effect of background noise modulation is analyzed.

Test consistency

In order to analyze the test consistency, the same type of background noise is presented to the listeners (but different from one listener to the other) at the beginning (first series) and the end (last series) of the test. For each group, we analyze the consistency of the responses with the ANOVA in which we have two statistical populations; the first one is composed of the responses of the listeners for the first series and the second of the responses of the last series. The consistency is verified thanks to the H_0 hypothesis. The respective probabilities for the equality of the mean values for the two groups are 97% and 68%.

Therefore, consistency is obtained for group1, meaning that this group is less sensitive to a learning effect and to a fatigue effect that could be linked to the duration of the test. This is not true with the second group as far as there is a non negligible probability (32%) that either the learning phase is useful for the listener to get use to the test or that the effect of fatigue is noticeable. The difference of age between the two groups may be the more simple explanation. In the next section, the two groups are merged for the analysis of background noise effect.

Effects of background conditions

A $10[\text{background condition BGN}] \times 41[\text{subject}]$ analysis of variance (ANOVA) was performed on the data. "Background condition" was found to be significant ($F[9,400] = 62.5, p < 0.001$). The result of the ANOVA is presented on figure 1 which represents the SRT values for all the background noise configurations together with the corresponding confidence interval. The first remark is that 4 categories of background noise can be considered according to a SRT classification.

1. The category of BGN which are responsible of a very high intelligibility. In this category, are the sentences pronounced by one person normally or reversely. The SRT values are very low, which corresponds to very high level of BGN (60 +13.3 dB and 60 + 12.9 for, respectively, the normal and reverse signals).
2. The category of high intelligibility. It contains 2 BGN: the LTASS-WNMOD1pers (with noise modulated by the envelop of one person's voice) and the open-space2.

3. The category of moderate intelligibility which contains 3 BGN: LTASS-WNMOD3pers, LTASS-MOD3pers and LTASS-SIM8Hz which SRT value is very close to the one obtained by Rhebergen⁴.
4. The category of low intelligibility which contains 3 BGN: the white noise (LTASS-WN), the open-space 1 and LTASS-WNMOD5pers.

These results confirm the literature in the sense that the more the BGN is modulated, the higher is the intelligibility. The first category seems to indicate that there is no effect of informational masking for this test. A major conclusion is that intelligibility is different between the two open-spaces, indicating that the procedure is relevant for establishing an intelligibility ranking of the open-spaces.

Confidence intervals are small for the largest values of SRT and increase with smaller values. This is due to a high variability of the listening conditions when only one speaker is making the background noise.

Figure 2 gives the SRT results for the two groups independently. We observe that the tendency is the same, except that a higher intelligibility is noticeable for the second group that performed the test in the semi-anechoic room (where the listening conditions are better).

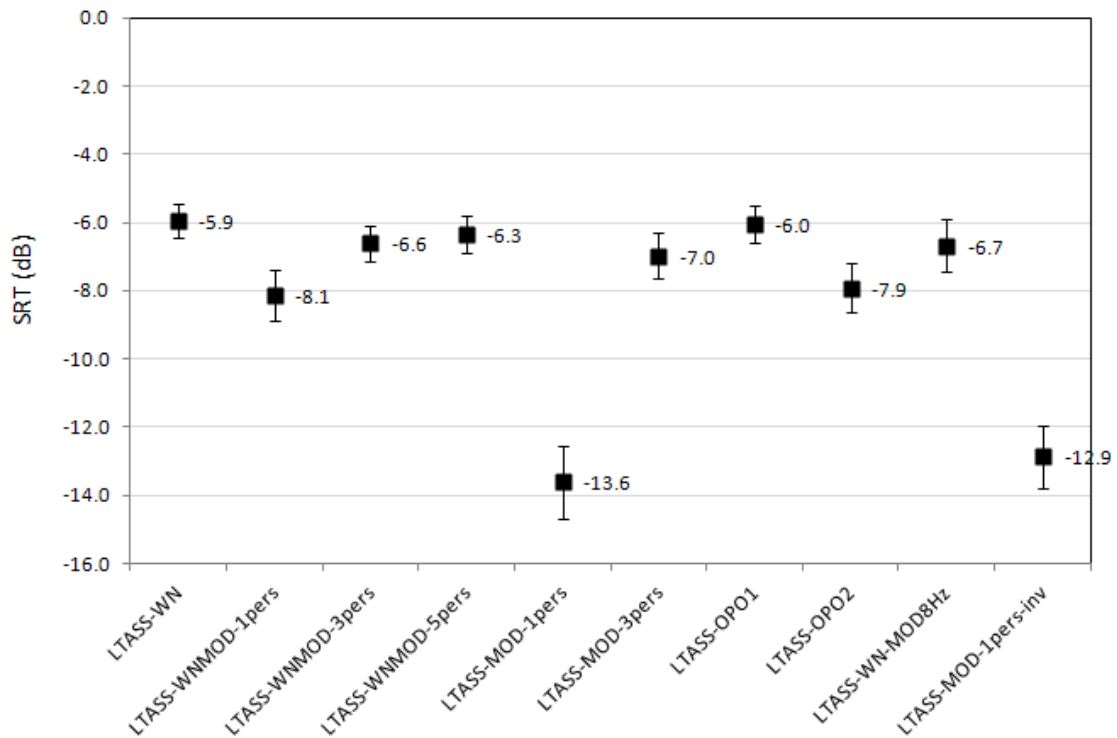


Fig. 1 – SRT as a function of background noise

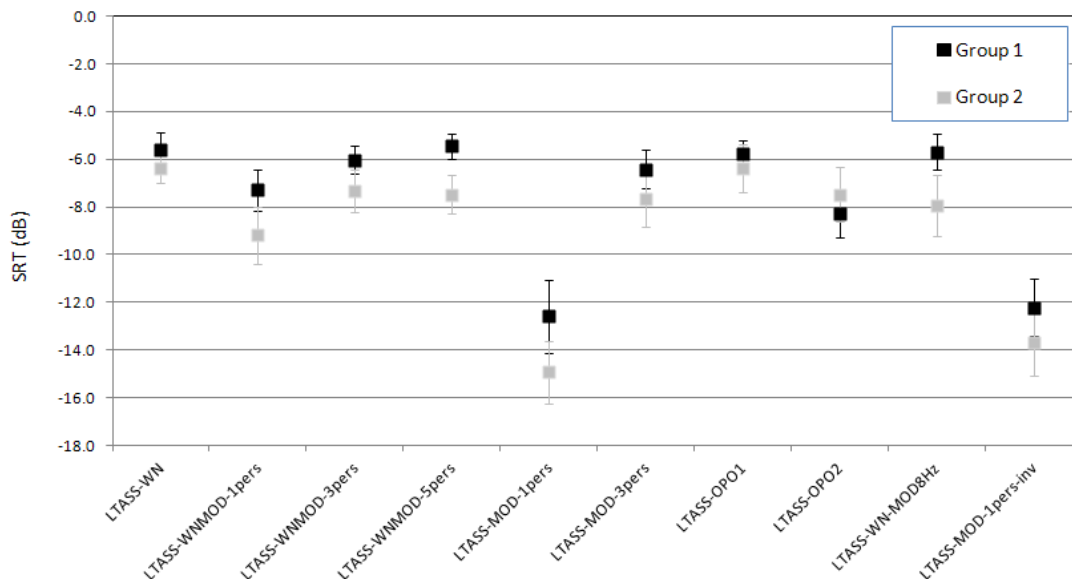


Fig. 2 – SRT as a function of background noise for the two groups

3 EVALUATION OF PERFORMANCE

For certain office-related tasks, notably those involving short-term memory, a drop-off in performance can be observed. This phenomenon is usually referred to as Irrelevant Speech Effect. In laboratory settings, it is often explored through a seriation task, where seven to nine elements (digits, letters or words) are presented in rapid succession to participants. After a short retention period, they are required to report back the series in the exact order of presentation. The ISE then seems to be related to speech intelligibility: the more the ambient speech is comprehensible, the harsher the decrease in performance is⁵. This is why an intelligibility index such as the STI has been used to propose target values for an open-plan office of good quality. In the same vein, Hongisto⁶ proposed a STI-performance curve to predict the decrement of performance induced by ambient speech of a given intelligibility. According to this model, for unintelligible speech, performance is the same as in silence. A growing impairment, following a sigmoid curve, occurs for intermediate STI values ranging from 0.3 to 0.5, where the deterioration of performance is already approaching its maximum.

For other common tasks, no ISE can be noticed⁷. As a consequence, when investigating the disturbing nature of ambient speech, one has to find alternative ways to estimate the difficulty of the task as perceived in a particular sound environment. In the ISE-related literature, questionnaires are the most frequently used method to obtain such an assessment. Their use is also relevant to complement performance measurements on short-term memory tasks. In that case, subjective assessments generally agree with objective ones. Participants express a higher level of disturbance when speech intelligibility increases. Nevertheless, both measurements do not seem to provide the same information with regard to the disturbance caused by the sound environment⁷. It should also be noted that the questions asked to participants vary across studies, which can make comparisons difficult. The use of a widespread workload measure could be useful for future reference.

The first goal of this study is to improve our understanding of the variations of performance with speech intelligibility, as measured by the STI. By exploring the [0.3-0.5] range, it aims to offer some material to further the work undertaken by Hongisto. By using an objective measure of performance and a subjective report of difficulty, the following experiment also seeks to compare these two types of assessments. Their respective evolutions with regard to the ambient speech intelligibility will be observed in order to compare their ability to discriminate between ambient speeches of different intelligibility levels, as measured by the STI.

3.1 The experiment

The experimental group consisted of 57 participants, one half being students and the other half recruited through an outsourcing company specialized in clinical trials. There were 32 women and 25 men, aged from 22 to 73 years, 36 years on average. All participants reported normal hearing and were paid for their participation.

The same lists of sentences as for the intelligibility tests were used as the speech signal as well as for the background noise rich with multi-talker babble. The total duration exceeded 12 minutes. Four values of the STI were implemented: 0.3, 0.4, 0.5 and 0.7. A model proposed by Hongisto⁸ predicts the STI between two nearby workstations. For a given office setting, the model provides octave-band attenuations which are applied to the spectrums of both speech and noise signals in order to obtain the corresponding STI value.

Participants worked over two sessions on consecutive days. Each session included three 10-minute blocks, separated by 5-minute pauses. One of them was carried out in silence, which served as a control condition for which $STI = 0$. For each of the other two, one of the four sound environments was used. At the end of the two sessions, each participant had been confronted to the four STI values. The order of the sentences was randomized before each block. The sound condition sequence was also balanced between subjects. The experiment took place in a sound-treated booth. Participants were alone in the booth, behind a desk. They faced a computer screen and were equipped with a mouse. Behind the screen was a loudspeaker, which created the sound environment. The A-weighted SPL was 46 dB(A) for all four noisy conditions.

Participants had to achieve a seriation task. A permutation of integers from 1 to 9 was presented, one number at a time (700 ms on, 300 ms off; MS Sans Serif font, 3 cm high). After the last number had disappeared, a five-second pause was observed. The individual numbers then came into view in a 3x3 response array. Participants were asked to reproduce the series in the exact order of presentation by clicking on the buttons. After clicking, the number disappeared and could not be selected again. The next series began 3 seconds after the ninth number was proposed. The response phase was self paced, which means that participants carried out as many series as they could in the 10-minute block. Performance was measured by first reporting the number of figures incorrectly placed in any series. The average number of errors per series in a block was then calculated.

After each block, participants were presented with a computer-based French version of the NASA-TLX questionnaire⁹. This workload assessment method comprises six questions, each of them addressing a particular component of the mental workload as detailed in Table 1.

Table 1: Overview of the NASA-TLX questionnaire

Question 1	Mental demands of the task
Question 2	Physical demands of the task
Question 3	Temporal demands of the task
Question 4	Self-rated performance
Question 5	Effort level
Question 6	Frustration

For each question, participants gave out a score comprised between 0 and 100 using a 21-point scale. The lower the score, the more comfortable the situation was reported to be. A Raw Task Load Index (RTLX) was eventually calculated by averaging the six scores¹⁰.

3.2 Results

Statistical analyses were led using Stata v12.0. A high degree of heteroscedasticity for both performance and RTLX excluded the use of a classical ANOVA. Moreover, the data presented a hierarchical structure: blocks were carried out on a given session by a given participant. Therefore, within-subject and within-session correlations had to be taken into account. These observations led to the use of a three-level mixed-effects regression model, with a dummy variable for each condition. After modeling, it was then possible to assess separately between-subject variability, the learning effect between sessions and the residual variance.

Performances

In the silent conditions, participants averaged 2.2 errors per series. A significant effect of the sound condition could be found ($\chi^2 = 22.84$, d.f. = 4, $p < 0.05$).

An agglomerative hierarchical clustering showed that the group of participants could be divided into two separate groups with distinctive sensitivities to the sound environment, as shown in Figure 3. Neither age nor sex had a significant influence on the classification.

The first group is composed with high-performing subjects (32 members, 1.8 errors per series on average in the silent condition). Their behavior depended very little on the sound condition, as no significant effect of STI value could be found.

The average performance in the second group was lower (2.8 errors per series on average in the silent condition, $\chi^2 = 7.89$, d.f. = 1, $p < 0.05$). Most noticeably, its members displayed a higher sensitivity with regards to the sound environment. The decrease in performance reached 11% for STI = 0.7. The effect of STI value on DP was significant ($\chi^2 = 93.36$, d.f. = 4, $p < 0.05$). Noise was significantly detrimental to performance when compared to the silence condition ($\chi^2 = 15.43$, d.f. = 1, $p < 0.05$). More errors are made for the two highest levels of STI (0.5 and 0.7) than for the two lowest ones (0.3 and 0.4; $\chi^2 = 11.46$, d.f. = 1, $p < 0.05$). The difference in DP was higher between the 0.4- and 0.7-levels (5.6 percentage points) than between the 0.3- and 0.5-levels (4.5 percentage points). This result does not comply with the shape of the curve in the prediction model.

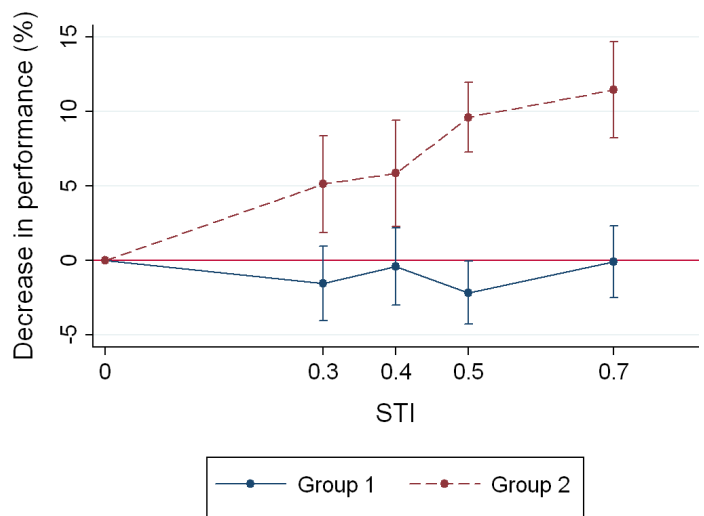


Fig .3 - Decrease in performance for the two groups of participants.

Adding speech noise affected negatively the good execution of a seriation task. This phenomenon grew stronger as the ambient speech noise intelligibility increased. Nevertheless, any analysis should allow for the high level of between-subject variability. A sizable portion of the panel exhibited no sensitivity to noise. The relationship between the decrease in performance and the value of the STI did not conform to the curve proposed by Hongisto. This sigmoid function was chosen because it mimics the shape of the curve describing the dependence of subjective intelligibility of sentences on the STI. This assumption suggests that the level of performance in the task at hand depends on the meaning of the ambient speech. However, it does not appear to be the case for many tasks, such as simple seriation tasks¹¹. A shape approaching the curve relative to CVC-syllables may be more adequate, as these elements do not carry any meaning.

Mental workload

There was a significant effect of the sound condition ($\chi^2 = 114.63$, d.f. = 4, $p < 0.05$). The reported RTLX was much higher in a noisy environment than in silence, as shown in Figure 4. This difference should not, however, be attributed solely to a change in STI. It may be a known fact that performance in a seriation task is not sensitive to the noise level¹¹, but it is not necessarily the case for mental workload assessments.

Once again, no significant difference could be made between the intermediate STI values. In order to show whether both objective and subjective assessments of disturbance vary in the same way with the STI, separate analysis were led with the same groups that were presented in previous section. The results are shown in Figure 5. Subjects drawn from the first group reported lower RTLX scores than their counterparts from the second group ($\chi^2 = 16.96$, d.f. = 1, $p < 0.05$). For both groups, the three intermediate conditions were not significantly different from one another. Although, the test is less discriminating for the first group (which is less sensitive to noise), it remains consistent with the second one (increase of workload with the intelligibility). Therefore, the test seems robust against a decreasing of the number of tested subjects.

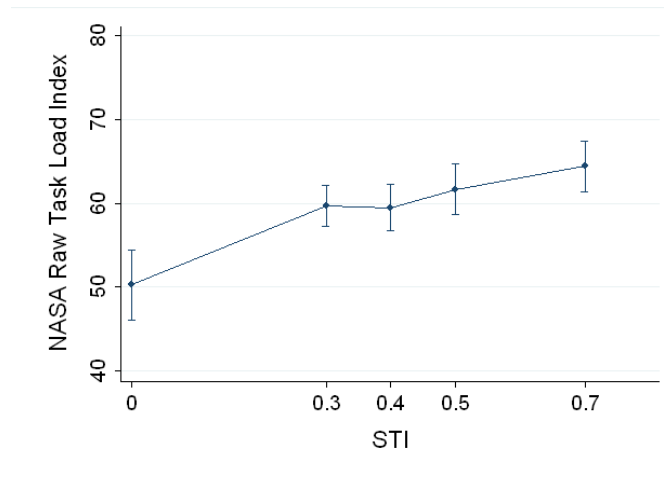


Fig. 4 - NASA-RTLX scores for the seriation task in sound environments of various STI values.

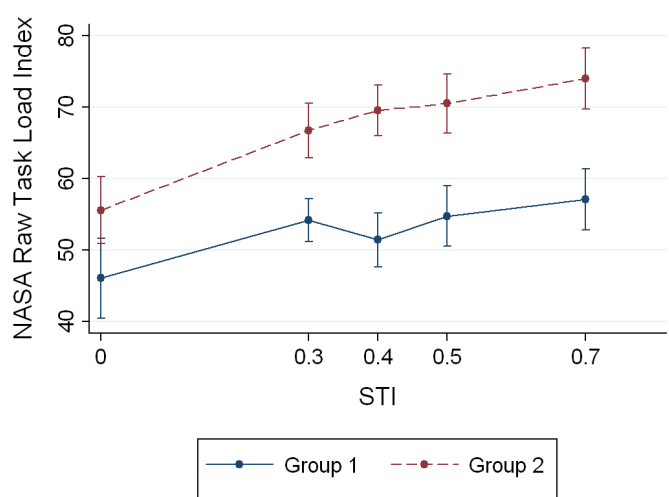


Fig. 5 - NASA-RTLX scores for the two groups of participants.

The NASA-TLX was conceived and validated as a global object, taking in consideration all questions. This makes any study of responses to an individual question theoretically unsound. Nevertheless, it can be a useful diagnostic tool to understand which questions bear the most importance in the variations of the global rating. The range of predicted scores for all 6 questions in noisy conditions, expressed as a relative increase from the lowest predicted score for each question, is presented in figure 6.

It appears that questions 2 (related to the physical demand of the task) was the one for which the range of predicted scores was the widest. This is quite surprising because the task does not seem to be physically strenuous, which should make the question irrelevant to our experiment. The study of variance partition coefficients (VPC) shines a light on that fact: 70% of the total variance of the scores for question 2 could be explained by between-subject variability and only 2% by the sound condition. These numbers tend to put the previous observation regarding the

high relative increase into perspective. As a matter of fact, despite the seemingly great differentiation, the residual error was too important for the effect of STI level to be significant. No significant differences between noisy conditions could either be noticed for questions 1 and 3. These results indicate an increase in STI did not change the way participants perceive the difficulty of the task.

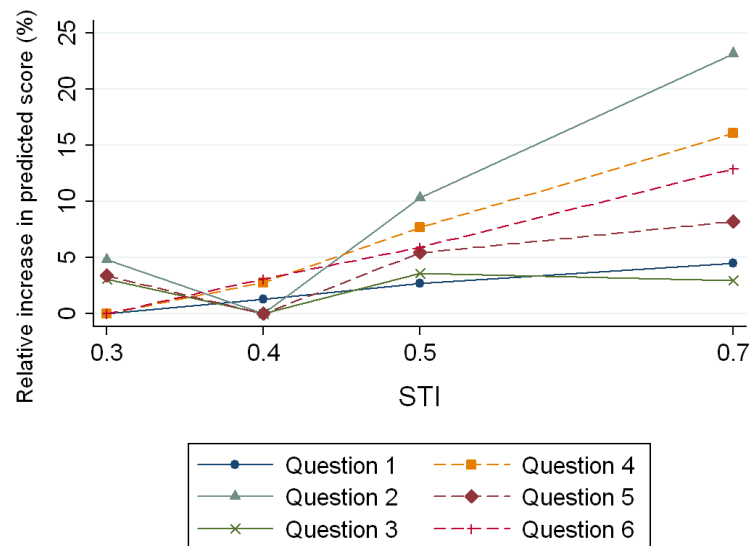


Fig. 6 - Increase in predicted score relative to the minimum for noisy conditions for the 6 questions of the NASA-TLX questionnaire.

It should be noted that question 3 was treated very differently from subject to subject. While some considered primarily the time pressure they felt during the presentation of the numbers, others expressed that, as the response phase was self-paced, there was no particular temporal demand to report. This could account for the limited variations of the reported scores for this item. On a related issue, response time was also measured. No significant effect of the sound condition could be found.

When considering the noisy conditions only, the effect of the STI value on reported scores reached significance for questions 4 ($\chi^2 = 9.61$, d.f. = 3, $p < 0.05$) and 5 ($\chi^2 = 9.30$, d.f. = 3, $p < 0.05$). These questions refer respectively to self-rated performance and effort level. Participants reported that an increase in STI forced them to work harder but still altered the quality of their output.

Adding speech noise increased the mental workload of a seriation task as reported by participants. The workload was stronger for ambient speech of high intelligibility. Noise affected less the perception of difficulty by the subjects than the appreciation of their work during the execution of the task. The information given by these subjective reports was not redundant with what the objective measurement of performance could bring by itself.

4 CONCLUSIONS

The objective of the paper was to report the experimental studies which have been done recently at INRS in order to evaluate the disturbance induced by ambient noise in open-plan offices.

The main goal was to create relevant set up and procedures in order to test both the intelligibility of speech, performances and mental workload in open-space conditions. This was done and the

results presented indicate that, on the axis of intelligibility, the SRT procedure seems to be relevant for a ranking of the open-spaces: different open-spaces, different SRT values. The test has now to be extended to other in-situ measurements of ambient noise and an objective quantification of it (through a modulation rate for instance) has to be performed. As concern the evaluation of performance and workload, the measurement results are encouraging as far as they clearly highlight the detrimental effect of the ambient noise conditions with a good intelligibility. This result is consistent with the literature, but more has to be done regarding the shape of the general STI-performance curve. For instance, one should be aware of the high level of between-subject variability. This problem may be mitigated by enlisting a greater number of participants. Further experiments are planned and a few remarks pertaining to the making up of an adequate protocol can be made.

5 ACKNOWLEDGEMENTS

We acknowledge Professeur Joël Ducourneau, from the Laboratoire de Biomathématiques et Audioprothèse of the Université de Pharmacie of Nancy for his help in setting up the tests and providing the facilities and students.

6 REFERENCES

1. S.P. Banbury, D.C. Berry, "Office noise and employee concentration: Identifying causes of disruption and potential improvements", *Ergonomics* 48, 25-37 (2005).
2. K.S. Rhebergen, N.J. Versfeld, W.A. Dreschler, "Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise. *J. Acoust. Soc. Am.* 120(6), 3988-3997, (2006).
3. R. Plomp, and A.M. Mimpen, Improving the reliability of testing the speech reception threshold for sentences. *Audiology*, 18, pp. 43-52 (1979).
4. <http://www.college-nat-audio.fr/accueil.html> (as seen on February 29th, 2012).
5. S.P. Banbury, W.J. Macken, S. Tremblay, D.M. Jones, "Auditory distraction and short-term memory: Phenomena and practical implications", *Human Factors: The Journal of the Human Factors and Ergonomics Society* 43, 12-28 (2001)
6. V. Hongisto, "A model predicting the effect of speech of varying intelligibility on work performance", *Indoor Air* 15, 458-468 (2005)
7. S.J. Schlittmeier, J. Hellbrück, R. Thaden, M. Vorländer, "The impact of background speech varying in intelligibility: Effects on cognitive performance and perceived disturbance", *Ergonomics* 51, 719-736 (2008)
8. V. Hongisto, J. Keränen, P. Larm, "Simple model for the acoustical design of open-plan offices", *Acta Acustica united with Acustica* 90, 481-495 (2004)
9. S.G. Hart, L.E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research". In P.A. Hancock, N. Meshkati (Eds.), *Human Mental Workload*, North Holland Press, Amsterdam (1988)
10. J.C. Byers, A.C. Bittner, S.G. Hill, "Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary?" In A. Mital (Ed.), *Advances in Industrial Ergonomics and Safety I*, Taylor and Francis (1989)
11. S. Tremblay, D.M. Jones, "Change of intensity fails to produce an irrelevant sound effect: Implications for the representation of unattended sound", *Journal of Experimental Psychology: Human Perception and Performance* 25, 1005-1015 (1999)