



**HAL**  
open science

## Indicateurs de confort acoustique dans les bureaux ouverts

Krist Kostallari, Etienne Parizet, Patrick Chevret

► **To cite this version:**

Krist Kostallari, Etienne Parizet, Patrick Chevret. Indicateurs de confort acoustique dans les bureaux ouverts. Congès Français d'Acoustique 2016, Apr 2016, Le Mans, France. hal-01475334

**HAL Id: hal-01475334**

**<https://hal.science/hal-01475334>**

Submitted on 23 Feb 2017

**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.

# CFA/VISHNO 2016

## Indicateurs de confort acoustique dans les bureaux ouverts

K. Kostallari<sup>a</sup>, E. Parizet<sup>b</sup> et P. Chevret<sup>a</sup>

<sup>a</sup>INRS, 1 rue Morvan, 54519 Vandoeuvre Les Nancy, France

<sup>b</sup>Univ Lyon, Insa Lyon, LVA, 25 bis, avenue Jean Capelle, 69621 Villeurbanne, France  
krist.kostallari@inrs.fr



LE MANS

Les indicateurs acoustiques utilisés aujourd'hui pour qualifier la satisfaction des salariés vis-à-vis de l'ambiance sonore des bureaux ouverts ne sont pas totalement adaptés à la problématique car ils visent principalement à décrire l'acoustique du local vide (DL2, Tr, D2S).

Même si la norme française en préparation (NF S 31-199), constitue une avancée majeure sur ce sujet, dans la mesure où elle repose sur une analyse de l'activité (en dressant une typologie des bureaux ouverts) et dans la mesure où elle recommande l'utilisation du niveau moyen pondéré A pour caractériser le bruit en activité, elle reste insuffisante. En effet, les expériences en laboratoire montrent que pour les valeurs généralement mesurées dans les bureaux ouverts, le niveau global n'est pas le principal facteur contribuant à une baisse de performance ou une augmentation de la gêne ressentie. Les enquêtes de terrain semblent indiquer que ce sont les conversations intelligibles qui sont les plus délétères vis-à-vis du ressenti de l'environnement sonore.

Au cours du projet GABO, soutenu par l'ANSES, plusieurs expériences en laboratoire ont été menées afin de tester l'effet des fluctuations liées aux conversations sur la performance et la gêne ressentie par les personnes. Au cours de ces expériences, les sujets devaient réaliser une tâche cognitive tout en étant soumis à des bruits ambiants dont les fluctuations correspondaient à des valeurs différentes d'indicateurs d'intelligibilité ou de fluctuation de niveaux. Cette présentation décrira les conditions expérimentales (tâche, type de bruit ambiant, etc.), le choix des descripteurs physiques ainsi que les relations obtenues entre les indicateurs retenus, la performance et la perception de l'environnement sonore.

## 1 Introduction

Le bruit est un des facteurs les plus gênants dans les bureaux ouverts [1, 2]. Ce bruit peut affecter la performance et la santé des salariés, ce qu'ont montré différentes études initiées dès les premières utilisations massives des bureaux ouverts, aux milieux des années 70. Dans une étude de Sundstrom et al. [3], 54% des employés interrogés ont confirmé un mécontentement lié à la gêne que l'environnement sonore pouvait apporter, surtout si celui-ci provenait de la parole et des sonneries de téléphone. Plus récemment, dans le cadre du projet ANSES GABO (Gêne Acoustique dans les Bureaux Ouverts), une enquête a été réalisée au sein de plusieurs entreprises [4]. 58% des 237 personnes interrogées ont répondu que le bruit était le facteur le plus gênant et, plus précisément, celui provenant des conversations intelligibles. Concernant les effets de l'intelligibilité, Banbury et Berry [5], lors d'une expérience en laboratoire, ont montré qu'un signal de parole a un effet perturbateur sur la performance lors d'une tâche de mémoire sérielle. Plus récemment, une expérience de Zimmer et al. [6] a été réalisée pour évaluer la gêne ressentie avant, en cours, et après une tâche de mémoire sérielle pour 4 environnements sonores différents (silence, bruit blanc, parole et signal modulé). Les participants ont rapporté qu'ils sont finalement plus gênés par la parole que par les autres bruits.

Le phénomène décrivant la baisse de performance durant une tâche en présence du bruit de parole est appelé ISE (Irrelevant Speech Effect) [7]. L'ISE peut être mis en évidence avec des tests objectifs (performance) et des tests subjectifs (questionnaires d'autoévaluation de la gêne).

A partir d'études montrant le lien entre l'intelligibilité de la parole et l'ISE, Hongisto et al. [8] ont proposé un modèle prédictif de la performance basé sur l'indicateur *STI* (Speech Transmission Index) de Steeneken et Houtgast [9].

Afin de tester le modèle d'Hongisto, de nouvelles expériences ont été conduites dans le cadre du projet ANSES GABO [10-12]. Au cours de ces expériences, deux facteurs ont été évalués : (i) la gêne et (ii) la performance des sujets pour différents environnements sonores. Les résultats suggéraient que l'augmentation du *STI* avait un effet sur la gêne et sur la performance.

Par ailleurs, des mesures de fluctuations du bruit ambiant et des enquêtes sur la gêne ressentie ont été menées dans 25 entreprises : dans certains bureaux l'amplitude des fluctuations pouvait dépasser 10 dB(A). Dans ces conditions, les employés exprimaient « une gêne

importante » liée au bruit ambiant. Aucune différence significative n'a été trouvée pour des valeurs plus faibles de fluctuations. Un indicateur a été développé pour la mesure des fluctuations, défini comme la différence entre le niveau moyen et l'indice fractile  $L_{90}$  pouvant être associé au bruit. L'avantage est qu'il s'agit d'un réel indicateur d'exposition (à l'ambiance globale), alors que le *STI* décrit la gêne poste-à-poste. De plus, cet indicateur est simple à mesurer, contrairement au *STI* qui nécessite de calculer un rapport signal sur bruit à partir d'un bruit artificiel généré dans le local.

L'objet de cet article est d'éclaircir la relation qui existe entre l'ISE et le *STI* ou les fluctuations du bruit ambiant en s'appuyant sur les expériences réalisées en laboratoire dans les études précédentes.

Dans un premier temps, les indicateurs physiques mentionnés seront présentés, pouvant être potentiellement utilisés comme descripteur de l'ISE.

Dans un second temps, les expériences en laboratoire seront décrites et analysées en vue de l'évaluation de l'ISE par le *STI* et les fluctuations du bruit ambiant.

Pour finir, les nouvelles pistes de réflexions et perspectives seront analysées.

## 2 Indicateurs pour l'évaluation de l'ISE

### Intelligibilité de la parole

Les enquêtes de terrain rapportent que le bruit des conversations est considéré comme la source la plus gênante pour les personnes travaillant dans les bureaux ouverts [12]. Un indicateur d'intelligibilité très utilisé dans les problématiques de bruit dans les locaux est le *STI* (Speech Transmission Index) développé par Steeneken et Houtgast [9]. Cet indicateur a été à l'origine mis au point pour évaluer l'intelligibilité de mots ou de phrases en présence de bruit de fond stationnaire. Il est défini comme [13] :

$$STI = \frac{\overline{SNR'} + 15}{30}, \quad (1)$$

avec

$$\overline{SNR'} = \sum_{i=1}^7 w_i SNR'_i, \quad (2)$$

$SNR'_i$  est un rapport signal sur bruit apparent par bande d'octave pondéré. Suivant la façon dont il est calculé, le

*STI* permet également de tenir compte de la réponse impulsionnelle du local, du seuil de réception moyen de l'auditeur et du masquage fréquentiel d'une bande sur ses voisines supérieures.

Pour les bruits de fond non stationnaires, une extension du *STI* a été développée par Ebissou [10] appelé *STI<sub>t</sub>* et basée sur le principe de fenêtre glissante moyennée de la même manière que le *ESII* de Rhebergen [14], version non stationnaire du *SII* de la norme ANSI S3.5-1997[15]. Une autre approche a été également mise en œuvre par Chevret [16] qui utilise les modulations à 4 Hz du bruit ambiant pour déterminer une valeur d'intelligibilité dans un bruit ambiant fluctuant (*STI<sub>m</sub>*).

Quelle que soit la formulation pour bruits de fond stationnaires ou non, le *STI* (ou *STI<sub>t</sub>* ou *STI<sub>m</sub>*) prend des valeurs entre 0 et 1, 0 signifiant que l'intelligibilité de la phrase est très mauvaise et 1 qu'elle est excellente. Plusieurs auteurs ont réalisé des expériences en laboratoire afin d'établir une relation entre l'ISE et le *STI*. Une compilation des principaux résultats tirés de la littérature est présentée sur la figure 1, sur laquelle a été ajouté un modèle de prédiction développé par Hongisto en 2005 [8].

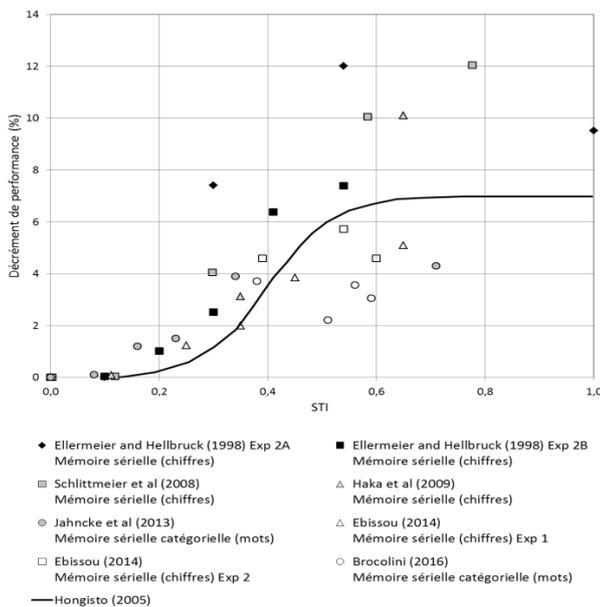


Figure 1 : Résultats des différentes études portant sur le décretement de performance en fonction du *STI*.

Sur cette courbe prédictive, il apparaît que le décretement de performance varie fortement pour les valeurs de *STI* entre 0,3 et 0,7, pour atteindre un plateau autour de 7%. Même si les auteurs ([10, 11, 17-20]) ne s'accordent pas toujours sur la valeur du début de l'augmentation de la perte de performance et même si la courbe de Hongisto ne constitue pas forcément une bonne prédiction pour les fortes valeurs de *STI*, on observe globalement une diminution des performances à mesure que l'intelligibilité augmente.

#### Indicateur d'émergence

Le *STI* est un indicateur dédié à la mesure de l'intelligibilité qui nécessite, en plus de la mesure du bruit de fond, l'utilisation d'une source de bruit. La procédure de mesure, relativement complexe, est décrite dans la norme IEC 60268-16 [21]. Puisque l'une des propriétés d'un signal de parole est sa modulation d'énergie à des fréquences inférieures à 16 Hz, correspondant au débit syllabique d'une conversation non contrainte, une

alternative intéressante au *STI* consiste à évaluer directement le degré de modulation d'un bruit ambiant.

Sur ce principe, l'idée développée dans cet article est de calculer le niveau équivalent pondéré A à la fréquence d'échantillonnage de 8 Hz (pas temporel de 125 ms) et de déterminer un taux d'émergence en faisant la différence entre le niveau moyen  $L_{eq,A}$  sur la durée du bruit ambiant considéré et l'indice fractile  $L_{90}$  pondéré A de ce bruit ambiant :

$$M_{eqA,125ms} = L_{eq,A} - L_{90A,125ms}, \quad (3)$$

où  $L_{90A,125ms}$  est la valeur de  $L_{eqA,125ms}$  (une fonction du temps représentant les valeurs de  $L_{eq,A}$  toutes les 125 ms) atteinte ou dépassée 90 % du temps. Cette valeur est a priori représentative du niveau de bruit de fond stationnaire dans le bruit ambiant. Cet indicateur s'exprime en dB(A) et il est calculé à partir d'une simple mesure microphonique. Le  $M_{eqA,125ms}$  a été utilisé lors des mesures en entreprise présentées dans le paragraphe suivant.

### 3 Mesures en entreprise

Depuis 2012, une grande campagne de mesure a été conduite en entreprise dans le cadre de projets soutenus par l'ANSES (GABO et MEPAS) [10-12, 22]. La base de données constituée à l'issue de cette campagne regroupe à la fois des mesures physiques du bruit (temps de réverbération, décroissance spatiale, bruit ambiant) et la connaissance de la perception de l'environnement sonore par les salariés. 25 entreprises ont été visitées et 619 personnes ont été interrogées.

La répartition des entreprises, établie en fonction du type d'espace ouvert qu'elles utilisent, est présentée dans le tableau 1. Cette répartition est conforme à celle qui est proposée dans la norme NF S 31-199 [23].

Type	Centre d'appel	Plateau collaboratif	Plateau administratif	Accueil du public
Nombre d'entreprises	4	12	2	7
Nombre de questionnaires	150	259	83	127

Tableau 1 : Répartition en fonction du type d'espace ouvert des entreprises figurant dans la base de données utilisée pour cette étude et des questionnaires recueillis.

Le questionnaire utilisé a été mis au point au cours du projet ANSES GABO (il est maintenant intégré dans la norme NF S 31-199 [23]). Le questionnaire comprend une centaine de questions qui portent à la fois sur l'environnement physique en général, sur l'environnement sonore, sur la fatigue et la gêne ressentie et sur la santé. Les réponses retenues portent surtout sur l'environnement sonore et la gêne ressentie par rapport aux différents types de bruit. Le nombre de questionnaires recueillis par type d'entreprise est donné dans le tableau 1.

La figure 2 fait une synthèse des résultats des 12 entreprises pour lesquelles il y a eu à la fois des mesures subjectives et objectives. Elle présente le niveau subjectif de gêne (différentes couleurs) en fonction des facteurs  $L_{eq,A}$  (abscisses) et  $M_{eqA,125ms}$  (ordonnées). Sur ce schéma, la significativité du facteur, obtenue par l'analyse statistique, est symbolisée par une séparation nette entre les ensembles de différentes couleurs. On peut y lire par exemple que pour des niveaux de bruit ambiant inférieurs à 60 dB(A), l'effet du  $M_{eqA,125ms}$  sur la gêne est

significativement différent entre les catégories 10 dB(A) et 15 dB(A), que ce soit pour les plateaux collaboratifs ou les salles d'accueil du public (pour les deux autres types d'espaces, le  $M_{eqA,125ms}$  n'a jamais atteint 15 dB(A)).

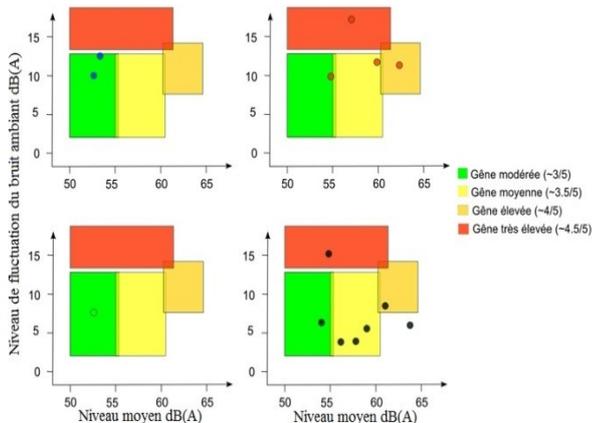


Figure 2 : Classification des espaces ouverts de la base de données MEPAS. (haut gauche) bleu : les centres d'appels, (haut droite) rouge : les plateaux collaboratifs, (bas gauche) vert : le plateau administratif, (bas droite) noir : les salles d'accueil du public

## 4 Expériences en laboratoire

Les effets de l'intelligibilité de la parole sur la gêne ont été étudiés au travers de deux campagnes d'essais en laboratoire dans les locaux du LVA de l'INSA de Lyon. La première campagne concerne les effets de bruit de fond stationnaire, la seconde des tests en présence de bruits de fond fluctuants.

### 4.1 ISE pour une voix en présence de bruits de fond stationnaires

#### 4.1.1 Méthode

Cette première expérience fait partie des travaux de thèse d'Ebissou [10]. Elle a été effectuée dans la salle d'écoute du LVA de l'INSA de Lyon. Cinquante-sept sujets ont participé à des mesures de performance et de charge cognitive dans des conditions de bruit de fond pour lesquelles un bruit de parole plus ou moins intelligible était présent. Une tâche de mémoire sérielle a été choisie pour cette expérience. Pour chaque série, associée chacune à une condition sonore différente, une permutation aléatoire de chiffres de 1 à 9 était présentée avec un temps d'affichage de 0,7 s pour chaque chiffre et un temps de pause de 0,3 s entre chaque chiffre. Après la disparition du dernier chiffre, un temps de rétention de 5 s était laissé au sujet. Ensuite, les sujets devaient restituer les chiffres présentés dans l'ordre de la présentation. Les sujets n'étaient soumis à aucune contrainte temporelle lors de la phase de restitution. La série suivante n'était lancée qu'à partir du moment où le dernier chiffre était saisi par le sujet, après une pause de 3 s. La durée totale de la tâche était de 10 min, les sujets réalisant plus ou moins de série suivant leur rapidité de réponse. Chaque chiffre mal placé était considéré comme une erreur. Chaque sujet devait répéter cette expérience six fois lors de deux sessions de 30 min chacune, ceci sur deux jours. Pour les deux sessions, une condition de contrôle (dans le silence) était prévue au démarrage puis quatre conditions sonores différentes, chacune caractérisée par un

STI donné. Les stimuli sonores étaient composés d'un bruit de parole (Phrases des listes de Fournier, Combesure et HINT) combiné à un bruit de fond stationnaire. Les valeurs de STI choisies étaient de 0,25, 0,35, 0,45 et 0,65. Ces valeurs décrivent la région de plus grande variation de performance du modèle de Hongisto (Figure 1). L'ordre de passage pour chaque condition était aléatoire.

La performance était mesurée au moyen du décrement de performance défini comme la différence entre le nombre d'erreurs pour la condition considérée et celui pour la condition de contrôle divisée par le nombre d'erreurs maximal :

$$DP(\%) = \frac{\text{erreurs}_{\text{bruit}} - \text{erreurs}_{\text{silence}}}{\text{quantité d'erreurs maximale}} \quad (4)$$

Pour mesurer la charge cognitive une forme simplifiée du NASA-TLX (Task Load Index) a été utilisée qui permet d'obtenir des résultats similaires en un temps réduit (Hart et Staveland, [24]) Ce questionnaire, nommé NASA-RTLX (Raw Task Load Index) repose sur 6 composantes de la charge cognitive : (1) l'exigence mentale, (2) l'exigence physique, (3) l'exigence temporelle, (4) la performance, (5) l'effort fourni et (6) la frustration ; ainsi qu'un score global de charge de travail. Sur le même principe que le décrement de performance, l'augmentation de la charge de travail ressentie est calculée comme suit :

$$ANRTLX(\%) = \frac{NRTLX_{\text{bruit}} - NRTLX_{\text{silence}}}{NRTLX_{\text{maximale}}} \quad (5)$$

#### 4.1.2 Résultats

La figure 3 présente la valeur moyenne et l'écart type du décrement de performance et du NASA-RTLX en fonction du STI. Les courbes montrent que l'intelligibilité des phrases a un effet sur le décrement de performance. Cet effet est significatif entre la première et la dernière condition sonore ( $\chi^2(4)=23,73 : p < 0,01$ ). Comme la charge cognitive, le décrement de performance augmente quand l'intelligibilité de la parole augmente. L'écart entre les deux conditions extrêmes est de 3,4 % pour le décrement de performance et de 4,5 % pour le NASA-RTLX.

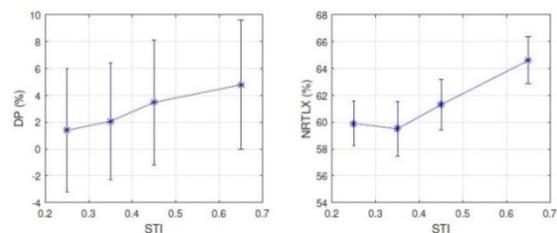


Figure 3: Le décrement de performance DP (à gauche) et le score de NASA-RTLX (à droite) en fonction du STI.

Une analyse plus poussée des résultats, basée sur une classification hiérarchique, a permis de séparer les sujets en deux groupes [10]. Pour le premier, constitué de 32 sujets, ni la gêne ni la performance ne sont sensibles aux valeurs de STI, contrairement au deuxième où les 25 sujets présentent une sensibilité à l'intelligibilité de la parole assez élevée par rapport à la condition de contrôle, mais sans significativité entre les différentes conditions sonores.

Afin de tester les effets de l'indicateur utilisé en entreprise, le  $M_{eqA,125ms}$  a été calculé pour chacune des conditions de STI de l'expérience. Le tableau 2 montre la correspondance entre les deux indicateurs. Pour chaque condition sonore, le  $M_{eqA,125ms}$  a été calculé toutes les

20 s. L'écart-type des mesures est inférieur à 0,05 dB(A). On note sur ce tableau que lorsque le  $STI$  est inférieur ou égal à 0,45, les valeurs de  $M_{eqA,125ms}$  sont proches et clairement séparées de la valeur  $STI = 0,65$ . On note également que la plage couverte pendant ces tests en laboratoire ne couvre pas toute la dynamique observée en entreprise (sur certains plateaux les valeurs de  $M_{eqA,125ms}$  atteignaient 15 dB(A)) (figure 2).

STI	0,25	0,35	0,45	0,65
$M_{eqA,125}$	1,5	1,78	2,01	5,21

Tableau 2 : Valeurs du  $M_{eqA,125ms}$  et du  $STI$  pour chaque condition sonore.

La figure 4 présente la valeur moyenne et l'écart type du décrement de performance et NASA-RTLX en fonction du  $M_{eqA,125ms}$ . La dynamique de l'axe des abscisses est de 15 dB(A) pour correspondre à la dynamique des valeurs de  $M_{eqA,125ms}$  mesurées en entreprise.

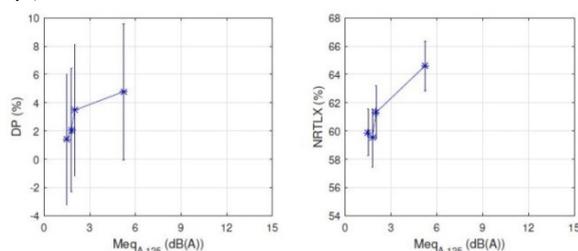


Figure 4: Décrément de performance DP (à gauche) et score de NASA-RTLX (à droite) en fonction du  $M_{eqA,125ms}$ .

Le  $M_{eqA,125ms}$  est donc très peu sensible à l'augmentation d'intelligibilité dans la zone d'étude habituelle de l'ISE, i.e.  $STI \in [0,3 - 0,5]$ . Sans amélioration (analyse différenciée par bandes d'octave par exemple), cet indicateur ne pourra donc pas être utilisé en remplacement du  $STI$  dans cette zone d'étude. En revanche, d'après ces premiers résultats, il semble être un bon descripteur de l'ISE pour des fortes valeurs d'intelligibilité. A cet égard, il pourrait d'ailleurs être intéressant de conduire des expériences qui couvrent la totalité de la dynamique observée en entreprise.

## 4.2 ISE pour une voix en présence de bruits de fond de parole

### 4.2.1 Méthode

Dans la seconde expérience [11], le décrement de performance et la charge cognitive sont étudiés en fonction d'un signal constitué d'une voix et d'un bruit de fond fluctuant. Cinquante-cinq personnes ont participé à l'expérience (vingt-neuf femmes et vingt-six hommes avec un âge moyen de 44 ans).

Lors de cette expérience une tâche de mémoire sérielle catégorielle a été utilisée: pour chaque condition sonore, chaque sujet devait mémoriser dix mots affichés sur un écran et les restituer oralement, l'un après l'autre. Les mots à mémoriser étaient issus des listes lexicales établies par Dubois et Poitou [25]. Les catégories retenues pour le test étaient les animaux, les fruits, les légumes, les oiseaux et les vêtements. Chaque catégorie comportait trente mots. Pour chaque sujet, une image annonçant le début de la série s'affichaient à l'écran pendant 5 s. Ensuite, sans pause temporelle, les mots s'affichaient un par un pendant 2 s

chacun. Sans pause, le sujet avait enfin 20 s pour restituer à haute voix les mots et 15 s de pause avant de passer à la série suivante. Cette procédure était répétée dix fois pour chaque condition sonore.

Les environnements sonore de la tâche correspondaient à des signaux permutés de façon aléatoire entre : (1) une voix masculine (M) et un bruit de fond stationnaire, (2) une voix M et une autre voix, (3) une voix M et 3 voix ou encore (4) une voix M et 5 voix (la voix M présentant un niveau plus fort que les autres voix). Pendant cette expérience une condition de contrôle (silence) a été réalisée. Les phrases utilisées pour les voix de masquage étaient du même type que dans l'expérience précédente (Fournier, Combescure et HINT). Les différentes conditions sonores étaient créées à partir de 40 échantillons de voix M (10 pour chaque condition) et 30 échantillons de voix de masquage (10 avec 1 voix masquante, 10 avec 3 voix et 10 avec 5 voix) utilisant 90 voix différentes. Tous les échantillons duraient 22 s. Ils étaient égalisés à un niveau de 47 dB(A). Le  $STI$  moyen pour la condition voix M en présence du bruit stationnaire était de 0,35 et pour les autres conditions de 0,36. Le  $STI_t$  ( $STI$  glissant moyenné) prenait les valeurs (1)  $STI_t = 0,38$  pour la condition M+bruit stationnaire, (2)  $STI_t = 0,69$  pour M+1 voix, (3)  $STI_t = 0,56$  pour M+3 voix et (4)  $STI_t = 0,51$  pour M+5 voix.

Le décrement de performance a été calculé de la même façon que dans la première expérience (expression 4), donc à la fin de chaque condition.

Le NASA-RTLX était proposé à chaque sujet à la fin de l'expérience en leur demandant de se prononcer sur chaque condition, incluant la condition de contrôle dans le silence. Puis le ANRTLX a été calculé à l'aide de l'expression 5.

### 4.2.2 Résultats

La figure 5 présente la valeur moyenne et l'écart type du décrement de performance et d'augmentation de NASA-RTLX en fonction du  $STI_t$  (haut) et du  $M_{eqA,125ms}$  (bas).

L'analyse de variance pour le décrement de performance ( $F(3,162)=1,3 : p = 0,293$ ) et le ANRTLX ( $F(3,102)=1,9 : p = 0,124$ ) en fonction du  $STI_t$  n'a pas montré d'effets significatifs de la modulation du bruit masquant sur l'ISE. Une classification hiérarchique a permis de distinguer deux groupes de sujets : le premier de 21 sujets insensibles au bruit de fond et le deuxième de 34 sujets pour lesquels le décrement de performance et le ANRTLX sont plus importants. C'est seulement pour la condition M+1 voix, qu'il y a une différence significative ( $F(3,159) = 3,5, p=0.018$ ) entre les valeurs du décrement de performance des deux groupes. Cette différence n'apparaît pas sur le taux d'augmentation de NASA-RTLX pour la même condition. D'ailleurs, l'analyse statistique ne montre pas d'effet significatif du  $STI_t$  sur le ANRTLX [11].

Bruit	$STI_t$	Ecart type ( $STI_t$ )	$M_{eqA,125ms}$	Ecart type ( $M_{eqA,125ms}$ )
M+bruit st.	0.38	0.003	0.44	0.035
M+5voix	0.51	0.013	3.74	0,61
M+3voix	0.56	0.018	4.75	0.85
M+1voix	0.69	0.032	7.89	1.77

Tableau 3: Valeurs moyennes du  $STI$  et du  $M_{eqA,125ms}$  pour l'expérience 2, accompagné de leurs écarts types.

Pour la même raison que dans l'expérience précédente, un calcul d'indicateur d'émergence a été réalisé a posteriori sur les échantillons sonores utilisés au cours de cette expérience. Les valeurs de  $M_{eqA,125ms}$  et de  $STI_t$  accompagnés de leurs écarts types sont donnés dans le tableau 3.

Les deux indicateurs varient dans le même sens. Cependant, l'utilisation du  $M_{eqA,125ms}$  n'apporte pas d'information supplémentaire. La dynamique utilisée est plus grande que celle de la première expérience mais elle n'atteint pas celle observée sur site. Globalement, l'ISE varie peu en fonction des indicateurs. Cela peut être dû au faible coût cognitif de tâche utilisée dans cette expérience, en particulier lié au fait qu'aucune période de rétention ne soit imposée au sujet entre la période de présentation des mots et la période de répétition.

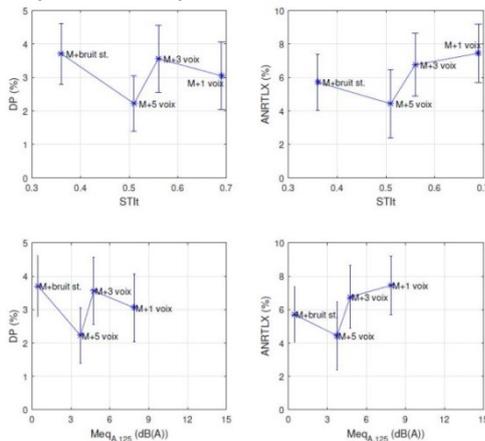


Figure 5: Décrément de performance DP (à gauche) et score de NASA-RTLX (à droite) en fonction du  $STI_t$  (haut) et  $M_{eqA,125ms}$  (bas).

## 5 Discussion et conclusion

Dans cet article, deux expériences ont été étudiées afin de corréliser l'ISE à différents indicateurs acoustiques. Ces mesures en laboratoire confirment l'effet perturbateur de l'intelligibilité sur l'ISE, notamment l'accroissement de la charge cognitive lorsque le  $STI$  varie entre 0,2 et 0,4 même si la significativité est discutable [8, 10, 11, 17-20]. Pour compléter l'analyse, un nouvel indicateur sensible aux fluctuations du bruit ambiant et mesuré sur le terrain a été rajouté au  $STI$  pour les expériences de laboratoire.

La première expérience montre que la variation du  $M_{eqA,125ms}$  est corrélée au décrément de performance et au NASA-RTLX et qu'elle couvre une plage plus importante que la variation du  $STI$ . Dans la seconde expérience, les trop faibles variations du DP et du NASA-RTLX ne permettent pas de confirmer l'intérêt du  $M_{eqA,125ms}$  par rapport au  $STI_t$ .

L'évolution du  $M_{eqA,125ms}$  en fonction du  $STI_t$  est tracée sur la figure 6 pour différentes configurations sonores. Pour chaque courbe, les indicateurs ont été calculés à partir de signaux constitués d'une voix cible et d'un signal masquant, en faisant varier le rapport signal sur bruit entre les deux. La plage de variation du SNR est [-10 ; 15 dB]. Les résultats de Brocolini (Exp. 2) [11] ont également été reportés. Il est d'abord intéressant de constater le  $M_{eqA,125ms}$  n'est pas corrélé à l'intelligibilité puisqu'une grande plage de valeurs de  $STI_t$  (entre 0.3 et 0.7)

correspond à une valeur constante du  $M_{eqA,125ms}$ . De plus, suivant le type de bruit de fond plusieurs valeurs de  $M_{eqA,125ms}$  peuvent être obtenues pour un même  $STI_t$ . On note également que le  $M_{eqA,125ms}$  commence à être sensible à la variation d'intelligibilité seulement quand le SNR est égal à 0 dB, soit un  $STI$  de 0.5. C'est à cette valeur que les émergences commencent réellement à apparaître sur le tracé du signal. Sur la figure, ce point correspond pour chaque courbe au changement de pente du  $M_{eqA,125ms}$  :  $STI_t = 0,5$  pour M+bruit stationnaire,  $STI_t = 0,63$  pour M+5voix,  $STI_t = 0,68$  pour M+3voix,  $STI_t = 0,78$  pour M+1voix. Au-delà de cette valeur les variations de  $STI_t$  engendrent de plus fortes variations de  $M_{eqA,125ms}$ . Le  $M_{eqA,125ms}$  n'est donc qu'un indicateur de fluctuation qui ne peut traduire qu'une certaine plage d'intelligibilité d'une voix cible. Par contre, à ce stade rien ne dit qu'il ne soit un plus mauvais estimateur de l'ISE sauf si l'on se focalise sur les signaux faiblement intelligibles. A contrario dans le cas des signaux très intelligibles ( $STI_t$  plus grand que 0,8), il pourra être utilisé avantageusement car sa plage de variation est plus importante que celle du  $STI_t$ .

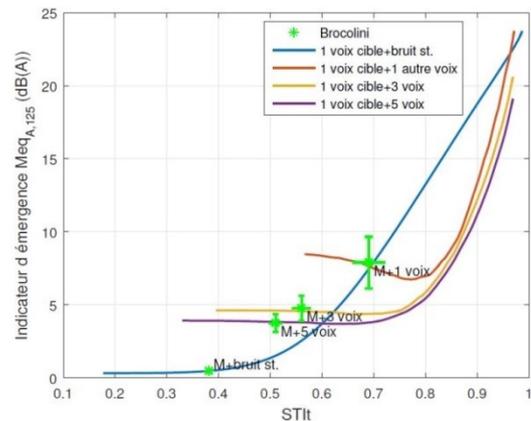


Figure 6:  $M_{eqA,125ms}$  en fonction du  $STI_t$  pour les signaux (1) M+ bruit st. (2) M+5 voix (3) M+3 voix (4) M+1 voix accompagné par les résultats de Brocolini [11] Exp 2 (Chaque courbe est obtenue en faisant varier le rapport signal sur bruit de -10 à 15 dB).

En conclusion, utiliser le  $STI$  pour évaluer l'ISE se base sur l'idée que c'est l'intelligibilité des conversations qui est à l'origine de la gêne sonore dans les bureaux ouverts. C'est aussi ce qui a conduit à se focaliser sur la zone 0,3-0,5 car c'est dans cette zone que la plus grande variation d'intelligibilité est observée, donc a priori et selon l'hypothèse d'Hongisto, c'est là qu'on aura la plus grande variation d'ISE. Cependant, les résultats obtenus dans les expériences relatées montrent des faibles variations d'ISE dans cette zone. De plus, pour des valeurs plus grandes de  $STI$ , l'ISE continue d'augmenter alors que l'intelligibilité n'augmente plus (pour  $STI > 0,7$ ). Il se peut donc que ce ne soit pas l'intelligibilité elle-même qui soit un facteur de la gêne mais seulement un des éléments de l'intelligibilité, comme les fluctuations temporelles du bruit ambiant. Dans ce cas, le  $M_{eqA,125ms}$  constituerait un indicateur complémentaire utile et plus simple à mettre en œuvre sur le terrain.

Une des perspectives de nos travaux est de faire des tests avec des valeurs plus élevées de  $STI$  pour mettre en évidence un effet sur l'ISE pour des expériences de courte durée. Une autre perspective consiste à faire des expériences avec des tâches plus difficiles en faisant varier le coût cognitif, par exemple en diminuant le temps de

réention ou le temps de réponse des sujets pendant la tâche. Il conviendrait aussi de faire des expériences de plus longues durées qui permettrait peut-être d'améliorer la significativité des résultats pour de faibles valeurs de *STI*.

Enfin, dans le cas d'une bonne intelligibilité, l'indicateur  $M_{eqA,125ms}$  pourrait représenter une alternative intéressante au *STI* : il varie plus rapidement et est simple à obtenir car basé sur une mesure microphonique du bruit ambiant.

## Références

1. Boyce, P., *Users' assessments of a landscaped office*. Journal of Architectural Research, 1974: p. 44-62.
2. Klitzman, S. and J.M. Stellman, *The impact of the physical environment on the psychological well-being of office workers*. Social Science & Medicine, 1989. **29**(6): p. 733-742.
3. Sundstrom, E., et al., *Office Noise, Satisfaction, and Performance*. Environment and Behavior, 1994. **26**(2): p. 195-222.
4. Pierrette, M., et al., *Noise effect on comfort in open-space offices: development of an assessment questionnaire*. Ergonomics, 2015. **58**(1): p. 96-106.
5. Banbury, S. and D.C. Berry, *Disruption of office-related tasks by speech and office noise*. British Journal of Psychology, 1998. **89**(3): p. 499-517.
6. Zimmer, K., J. Ghani, and W. Ellermeier, *The role of task interference and exposure duration in judging noise annoyance*. Journal of sound and vibration, 2008. **311**(3-5): p. 1039-1051.
7. Salame, P. and A. Baddeley, *Disruption of short-term memory by unattended speech: Implications for the structure of working memory*. Journal of verbal learning and verbal behavior, 1982. **21**(2): p. 150-164.
8. Hongisto, V., *A model predicting the effect of speech of varying intelligibility on work performance*. Indoor air, 2005. **15**(6): p. 458-468.
9. Houtgast, T., H. Steeneken, and R. Plomp, *Predicting speech intelligibility in rooms from the modulation transfer function. I. General room acoustics*. Acta Acustica united with Acustica, 1980. **46**(1): p. 60-72.
10. Ebissou, A., *Gêne sonore en bureaux ouverts: Impact de la présence de plusieurs voix intelligibles (thèse de doctorat)*, in *Mécanique, Energétique, Génie civil, Acoustique*, 2013, INSA de Lyon: Lyon. p. 118.
11. Brocolini, L., E. Parizet, and P. Chevret, *Noise annoyance in open plan offices : influence of source type and evaluation of an intelligibility index adapted to a fluctuating background noise*. Applied Acoustics, 2016.
12. Parizet, E., *Gêne Acoustique dans les Bureaux Ouverts (GABO)*, 2014, INSA-Lyon: Rapport.
13. Goldsworthy, R.L. and J.E. Greenberg, *Analysis of speech-based speech transmission index methods with implications for nonlinear operations*. The Journal of the Acoustical Society of America, 2004. **116**(6): p. 3679-3689.
14. Rhebergen, K.S. and N.J. Versfeld, *A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners*. The Journal of the Acoustical Society of America, 2005. **117**(4): p. 2181-2192.
15. ANSI, A., *S3. 5-1997, Methods for the calculation of the speech intelligibility index*. New York: American National Standards Institute, 1997. **19**: p. 90-119.
16. Chevret, P., *Release from masking of speech intelligibility due to fluctuating ambient noise in open-plan offices*. Applied Acoustics, 2016. **101**(1): p. 156-167.
17. Ellermeier, W. and J. Hellbrück, *Is level irrelevant in "irrelevant speech"? Effects of loudness, signal-to-noise ratio, and binaural unmasking*. Journal of Experimental Psychology: Human Perception and Performance, 1998. **24**(5): p. 1406.
18. Schlittmeier, S.J., et al., *The impact of background speech varying in intelligibility: Effects on cognitive performance and perceived disturbance*. Ergonomics, 2008. **51**(5): p. 719-736.
19. Haka, M., et al., *Performance effects and subjective disturbance of speech in acoustically different office types – a laboratory experiment*. Indoor air, 2009. **19**(6): p. 454-467.
20. Jahncke, H., V. Hongisto, and P. Virjonen, *Cognitive performance during irrelevant speech: Effects of speech intelligibility and office-task characteristics*. Applied Acoustics, 2013. **74**(3): p. 307-316.
21. Commission, I.E., *IEC 60268-16: Sound system equipment-Part 16: Objective rating of speech intelligibility by speech transmission index*. IEC, Switzerland, 2003.
22. Chevret, P., *Etude des liens possible entre des paramètres d'acoustiques physiques et des grandeurs subjectives de la perception sonore à partir des données mesurées en entreprise (MEPAS)*, 2014, INRS: Rapport.
23. AFNOR, *NF S 31-199, Acoustique - Bureaux ouverts : programmation conception et usage / utilisation*, 2015.
24. Hart, S.G., *Nasa-Task Load Index (NASA-TLX); 20 Years Later*. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2006. **50**(9): p. 904-908.
25. Dubois, D. and J. Poitou, *Normes catégorielles pour vingt-deux catégories sémantiques en français et dix catégories en allemand*. Cahiers du LCPE, 2002. **5**: p. 31-104.