



HAL
open science

Gêne sonore en bureaux ouverts : Impact de la présence de plusieurs voix intelligibles

Ange Ebissou

► **To cite this version:**

Ange Ebissou. Gêne sonore en bureaux ouverts : Impact de la présence de plusieurs voix intelligibles. Acoustique [physics.class-ph]. INSA de Lyon, 2013. Français. NNT : 2013ISAL0119 . tel-01127491

HAL Id: tel-01127491

<https://theses.hal.science/tel-01127491>

Submitted on 7 Mar 2015

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.

Année 2013



Thèse



Gêne sonore en bureaux ouverts : Impact de la présence de plusieurs voix intelligibles

Présentée devant
L'Institut National des Sciences Appliquées de Lyon



Pour obtenir
le grade de Docteur

Ecole Doctorale : Mécanique, Energétique, Génie civil, Acoustique
Spécialité : Acoustique

Par

Ange Ebissou
Ingénieur Ecole Centrale de Lyon

Soutenue devant la Commission d'Examen

Jury

| | | |
|-------------------------|------------------------------|------------|
| M. Didier Baptiste | INRS (Vandoeuvre-lès-Nancy) | |
| M. Michel Bérengier | IFSTTAR (Bouguenais) | |
| M. Patrick Chevret | INRS (Vandoeuvre-lès-Nancy) | |
| M. Wolfgang Ellermeier | TU Darmstadt (Allemagne) | Rapporteur |
| Mme Catherine Lavandier | Université de Cergy-Pontoise | Rapporteur |
| M. Etienne Parizet | INSA (Lyon) | Directeur |

INSA Direction de la Recherche - Ecoles Doctorales – Quinquennal 2011-2015

| SIGLE | ECOLE DOCTORALE | NOM ET COORDONNEES DU RESPONSABLE |
|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Chimie | CHIMIE DE LYON http://www.edchimie-lyon.fr Insa : R. GOURDON | M. Jean Marc LANCELIN Université de Lyon – Collège Doctoral Bât ESCPE 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cedex Tél : 04.72.43 13 95 directeur@edchimie-lyon.fr |
| E.E.A | ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE http://edeea.ec-lyon.fr Secrétariat : M.C. HAVGOUDOUKIAN eea@ec-lyon.fr | M. Gérard SCORLETTI Ecole Centrale de Lyon 36 avenue Guy de Collongue 69134 ECULLY Tél : 04.72.18 65 55 Fax : 04 78 43 37 17 Gerard.scorletti@ec-lyon.fr |
| E2M2 | EVOLUTION, ECOSYSTEME, MICROBIOLOGIE, MODELISATION http://e2m2.universite-lyon.fr Insa : H. CHARLES | Mme Gudrun BORNETTE CNRS UMR 5023 LEHNA Université Claude Bernard Lyon 1 Bât Forel 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cédex Tél : 06.07.53.89.13 e2m2@univ-lyon1.fr |
| EDISS | INTERDISCIPLINAIRE SCIENCES-SANTE http://www.ediss-lyon.fr Sec : Samia VUILLERMOZ Insa : M. LAGARDE | M. Didier REVEL Hôpital Louis Pradel Bâtiment Central 28 Avenue Doyen Lépine 69677 BRON Tél : 04.72.68.49.09 Fax : 04 72 68 49 16 Didier.revel@creatis.uni-lyon1.fr |
| INFOMATHS | INFORMATIQUE ET MATHEMATIQUES http://infomaths.univ-lyon1.fr Sec : Renée EL MELHEM | Mme Sylvie CALABRETTO Université Claude Bernard Lyon 1 INFOMATHS Bâtiment Braconnier 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cedex Tél : 04.72. 44.82.94 Fax : 04 72 43 16 87 infomaths@univ-lyon1.fr |
| MATERIAUX | MATERIAUX DE LYON http://ed34.universite-lyon.fr Secrétariat : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Saint Exupéry Ed.materiaux@insa-lyon.fr | M. Jean-Yves BUFFIERE INSA de Lyon MATEIS Bâtiment Saint Exupéry 7 avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél : 04.72.43 83 18 Fax 04 72 43 85 28 Jean-yves.buffiere@insa-lyon.fr |
| MEGA | MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE http://mega.ec-lyon.fr Secrétariat : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Saint Exupéry mega@insa-lyon.fr | M. Philippe BOISSE INSA de Lyon Laboratoire LAMCOS Bâtiment Jacquard 25 bis avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél : 04.72 .43.71.70 Fax : 04 72 43 72 37 Philippe.boisse@insa-lyon.fr |
| ScSo | ScSo* http://recherche.univ-lyon2.fr/scso/ Sec : Viviane POLSINELLI Brigitte DUBOIS Insa : J.Y. TOUSSAINT | M. OBADIA Lionel Université Lyon 2 86 rue Pasteur 69365 LYON Cedex 07 Tél : 04.78.77.23.86 Fax : 04.37.28.04.48 Lionel.Obadia@univ-lyon2.fr |

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Remerciements

Je tiens en premier lieu à remercier Etienne Parizet, mon directeur de thèse, pour avoir accepté de suivre mon travail pendant ces trois années et pour m'avoir accueilli lors de mes passages au LVA. Son expérience des tests psychoacoustiques, de la confection du protocole à l'analyse des résultats, aura été précieuse, tout autant que son sens du détail en période de rédaction.

Je remercie également Patrick Chevret pour avoir encadré mon travail avec soin et avoir fait preuve d'une grande disponibilité, au jour le jour. J'ai énormément apprécié de pouvoir, à tout moment, entrer dans son bureau et commencer une discussion qui, si elle ne devait durer que quelques instants au départ, finissait régulièrement par prendre bien plus de temps.

Je suis reconnaissant envers M. Wolfgang Ellermeier et Mme Catherine Lavandier pour avoir consenti à être rapporteurs de ce mémoire, ainsi qu'envers MM. Didier Baptiste et Michel Bérengier pour avoir accepté de participer à ce jury.

Je remercie l'INRS pour l'excellent soutien matériel dont j'ai bénéficié pendant cette période. J'ai été parfaitement accueilli au sein du laboratoire RBT ; pour cela, je remercie tout d'abord Jacques Chatillon et Isabelle Manet, qui ont toujours su répondre à la moindre de mes questions et résoudre bien des petits problèmes : grâce à eux, je n'ai jamais été perdu au milieu de cet institut. La bonne ambiance présente constamment dans les locaux du laboratoire en aura fait un lieu parfait pour accomplir ce travail de thèse : je remercie donc Nicolas, Julien, Armand, Jean-Pierre, Alain, Joël, Adil, Daniel et Léon pour toutes ces discussions et ces moments de bonne humeur. J'exprime une gratitude toute particulière envers Geneviève, pour l'aide considérable qu'elle m'a apporté sur ce travail. Je suis heureux d'avoir pu m'appuyer sur son enthousiasme et son sens de l'organisation durant le déroulement des expériences. Enfin, j'adresse un grand merci à Adel, mon voisin de bureau. Ses conseils avisés et les si nombreuses conversations totalement impromptues que nous avons menées ont rendu les premières années de mon passage à l'INRS bien plus plaisantes.

Je remercie mes proches pour avoir suivi avec un intérêt parfois amusé mes tribulations nancéiennes. Je remercie mes parents, grâce à qui je n'ai jamais craint de ne pas pouvoir atteindre mes buts, quels qu'ils soient. Merci enfin à Marine, qui sait mieux que je ne saurais l'écrire à quel point sa présence m'aura été, et continue de m'être, indispensable.

Résumé En bureaux ouverts, le bruit de parole des autres occupants représente un facteur de gêne important. L'objectif de ce travail est d'évaluer les outils de prédiction du pouvoir de nuisance de la parole et de proposer des voies d'amélioration. Ceci nécessite de détailler l'influence des caractéristiques d'un bruit de parole qui peuvent rendre le travail plus difficile. Pour ce faire, plusieurs expériences basées sur des exercices mentaux classiques sont menées. La surcharge cognitive engendrée par le bruit de parole ambiant y est mesurée de façon objective et subjective. Dans un premier temps, l'étude est centrée sur l'influence de l'intelligibilité de la parole. En effet, certains auteurs proposent d'utiliser un indicateur d'intelligibilité, le *Speech Transmission Index* (STI), pour prédire la nuisance dans un bureau ouvert. Les résultats obtenus confirment cette influence de l'intelligibilité. Cependant, les bruits de fond rencontrés en bureaux ouverts peuvent être fortement fluctuants. Ceci a un effet sur la nuisance dont le STI ne peut rendre compte. Dans un second temps, une alternative permettant de rendre compte de la variabilité temporelle de réelles ambiances sonores de bureaux est donc proposée. Le pouvoir prédictif de cet indicateur sur la perturbation au travail est évalué pour plusieurs types de tâches. Si les effets observés sont faibles, en particulier au regard des différences de sensibilité au bruit entre individus, la prise en compte des fluctuations de niveau pour un environnement sonore donné semble bel et bien essentielle à la prédiction chiffrée de la gêne sonore causée. De plus, l'évaluation subjective de la gêne s'avère être un complément essentiel à l'évaluation objective pour rendre compte de l'influence de certains facteurs.

Mots-clés Acoustique, Bruit, Gêne sonore, Parole, Intelligibilité, Tertiaire

Abstract Speech noise is a major cause of annoyance in open-plan offices. The purpose of this research was to assess the quality of existing tools for predicting the nuisance potential of speech and highlight possibilities for improvement. The ways in which speech sounds influence noise annoyance were thus investigated. To this end, experiments were carried out based on classical mental tasks, with concurrent objective and subjective measurements of the increase in mental workload caused by ambient speech. First, the effect of speech intelligibility was studied. Indeed, some argue for the use of a speech intelligibility measure, the *Speech Transmission Index* (STI), to predict noise annoyance in open-plan offices. Results confirmed the increase in disruption caused by speech made more intelligible. However, the STI was incapable of predicting noise annoyance for fluctuating background sounds, which are usual in open-plan offices. An alternative was then provided that accounted for the temporal variability seen in actual office sound environments. The predictive ability of this measure was assessed for varied tasks. The resulting effects were small when compared with individual differences in susceptibility to speech noise. However, they indicated that allowing for level fluctuations in the sound environment is essential to quantifying the induced noise annoyance. Moreover, it appeared that subjective measurements of noise annoyance were a critical complement to objective measurements for identifying certain causes of noise annoyance.

Keywords Acoustics, Noise, Sound annoyance, Speech, Intelligibility, Service industries

Table des matières

| | | |
|-------|------------------------------------------------------------------------|----|
| 1 | Introduction | 14 |
| 2 | Etat de l'art..... | 18 |
| 2.1 | Nuisance sonore au travail..... | 18 |
| 2.1.1 | Différents aspects de la nuisance sonore | 18 |
| 2.1.2 | Perturbation d'une activité : la surcharge cognitive..... | 20 |
| 2.1.3 | Mesures de la surcharge cognitive..... | 20 |
| 2.2 | Intelligibilité de la parole..... | 25 |
| 2.2.1 | L'Articulation Index et le Speech Intelligibility Index..... | 25 |
| 2.2.2 | Importance de la conservation des modulations du signal | 27 |
| 2.2.3 | Le Speech Transmission Index | 30 |
| 2.2.4 | Importance des fluctuations du bruit de fond..... | 33 |
| 2.2.5 | Autres facteurs | 35 |
| 2.2.6 | Conclusion..... | 35 |
| 2.3 | Impact de l'intelligibilité sur la nuisance sonore..... | 36 |
| 2.3.1 | Performance : cas particulier de la mémoire sérielle..... | 36 |
| 2.3.2 | Bruit et mémoire à court terme | 37 |
| 2.3.3 | Le STI comme indicateur du potentiel de nuisance..... | 38 |
| 2.3.4 | Autres tâches : apport des mesures indirectes de charge cognitive..... | 40 |
| 2.4 | Approche suivie - Objectifs du présent travail | 41 |
| 3 | Influence du STI du discours ambiant sur la gêne..... | 44 |
| 3.1 | Cas d'un bruit de fond stationnaire..... | 44 |
| 3.1.1 | Protocole expérimental | 45 |
| 3.1.2 | Résultats..... | 53 |
| 3.1.3 | Discussion..... | 58 |
| 3.2 | Cas d'un bruit de fond fluctuant..... | 62 |
| 3.2.1 | Protocole expérimental | 64 |
| 3.2.2 | Résultats..... | 66 |
| 3.2.3 | Discussion..... | 69 |
| 3.3 | Conclusion..... | 71 |
| 4 | Influence des fluctuations de niveau pour des ambiances réelles | 72 |
| 4.1 | Quantification des variations de niveau avec le STIt | 74 |
| 4.2 | Perturbation de la mémoire sérielle..... | 77 |

| | | |
|-------|----------------------------------------------------------------------------------------------------------------|-----|
| 4.2.1 | Stimuli sonores..... | 77 |
| 4.2.2 | Résultats..... | 78 |
| 4.2.3 | Discussion..... | 81 |
| 4.3 | Perturbation d'autres activités | 83 |
| 4.3.1 | Exercice de raisonnement grammatical | 83 |
| 4.3.2 | Exercice de fluence verbale..... | 84 |
| 4.3.3 | Protocole expérimental | 85 |
| 4.3.4 | Résultats..... | 86 |
| 4.3.5 | Discussion..... | 91 |
| 4.4 | Conclusion..... | 93 |
| 5 | Conclusion..... | 94 |
| | Récapitulatif des résultats | 95 |
| | Implications pratiques et perspectives | 96 |
| | Bibliographie | 100 |
| | Annexe A : Items du questionnaire NASA-TLX (NASA, 1986) | 108 |
| | Annexe B : Présentation du modèle de prédiction du STI en bureaux ouverts (Hongisto <i>et al.</i> , 2004)..... | 110 |
| | Annexe C : Préviation de la puissance statistique du protocole de la première expérience | |

Liste des figures

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1 : Schéma de calcul de l'Articulation Index..... | 25 |
| Figure 2 : Schéma de calcul du STI..... | 30 |
| Figure 3 : Relation entre la valeur du STI (en abscisses) et l'intelligibilité de différents types d'énoncés (en ordonnées). Courbe PB : listes de mots phonétiquement équilibrées. Courbe CVC : syllabes consonne-voyelle-consonne. Courbe Sentences : phrases entières. Tiré de (Hongisto, 2005) ; données de (Steeneken et Houtgast, 2002)..... | 32 |
| Figure 4 : Comparaison des coefficients de pondération par bande pour le STI et le SII..... | 32 |
| Figure 5 : Schéma de calcul du STIt. | 34 |
| Figure 6 : Modèle de prédiction du décrétement de performance en fonction de la valeur du STI dans le bureau ouvert. Ici, $DP_{max} = 7\%$ | 39 |
| Figure 7 : Spectres moyens originaux des signaux de parole et de bruit. Pour les deux spectres, le niveau global est fixé à 60 dB(A)..... | 47 |
| Figure 8 : Spectres moyens des signaux de parole et de bruit pour chacune des conditions bruyantes..... | 51 |
| Figure 9 : Intelligibilité de phrases pour les quatre conditions sonores étudiées. Les valeurs médianes sont accompagnées de leurs écarts interquartiles respectifs. En continu, la courbe de tendance associée aux résultats. En pointillés, l'expression (2.19) proposée par Hongisto (2005). | 52 |
| Figure 10: Décréments de performance pour les valeurs de STI étudiées. Les moyennes sont accompagnées de leur intervalle de confiance. | 54 |
| Figure 11 : Décréments de performance pour les valeurs de STI étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance. | 55 |
| Figure 12 : RTLX obtenus pour les valeurs de STI étudiées. Les moyennes sont accompagnées de leur intervalle de confiance. | 57 |
| Figure 13 : RTLX obtenus pour les valeurs de STI étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance. | 57 |
| Figure 14 : Comparaison des performances obtenues par les sujets du second groupe avec le modèle de Hongisto (2005). En continu, courbe de tendance des résultats d'intelligibilité, mise à l'échelle (voir Figure 9). | 60 |
| Figure 15 : Evolution des performances suivant l'ambiance sonore dans l'expérience de Jones et Macken (1995). En abscisse, le nombre de voix constituant le bruit ambiant. En ordonnée, le nombre moyen d'erreurs pour une série de 7 lettres à restituer. | 63 |
| Figure 16 : Décréments de performance pour les bruits de fond étudiés. Les moyennes sont accompagnées de leur intervalle de confiance. | 67 |

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 17 : RTLX obtenus pour les bruits de fond étudiés. Les moyennes sont accompagnées de leur intervalle de confiance..... | 68 |
| Figure 18 : Valeur du STIt pour un bruit de fond modulé à 4 Hz masquant un bruit de parole de même spectre en fonction de la valeur de l'indice de modulation. | 75 |
| Figure 19 : Enveloppes de la courbe d'intensité de signaux modulés en intensité à 4 Hz, pour différentes valeurs de l'indice de modulation. Les valeurs de STIt relatives à chaque courbe sont indiquées en légende. | 75 |
| Figure 20 : Décréments de performance pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance..... | 79 |
| Figure 21 : Décréments de performance pour les valeurs de STIt étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance. | 79 |
| Figure 22 : RTLX obtenus pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance..... | 80 |
| Figure 23 : RTLX obtenus pour les valeurs de STIt étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance..... | 81 |
| Figure 24 : Comparaisons des performances pour les expériences 2 et 3 en fonction du STIt. | 82 |
| Figure 25 : Décréments de performance pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance..... | 88 |
| Figure 26 : Décréments de performance pour les valeurs de STIt étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance. | 88 |
| Figure 27 : RTLX obtenus pour les valeurs de STIt. Les moyennes sont accompagnées de leur intervalle de confiance. | 89 |
| Figure 28 : RTLX obtenus pour les valeurs de STIt étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance..... | 90 |
| Figure 29 : Fluence verbale pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance..... | 91 |
| Figure 30 : RTLX obtenus pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance..... | 91 |
| Figure 31 : Représentation schématique des postes de travail étudiés..... | 110 |

Liste des tableaux

| | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Tableau 1 : Récapitulatif des coefficients d'absorption par bande d'octave relatifs aux éléments constitutifs du bureau ouvert modélisé | 49 |
| Tableau 2 : Configurations retenues pour la modélisation d'environnements sonores en bureaux ouverts. | 50 |
| Tableau 3 : Rapports signal-à-bruit pour les valeurs de STI retenues. | 50 |
| Tableau 4 : Récapitulatif des mesures de performance et de charge cognitive. | 53 |
| Tableau 5 : Comparaison des résultats en performances pour plusieurs protocoles | 59 |
| Tableau 6 : Récapitulatif des mesures de performance et de charge cognitive. | 66 |
| Tableau 7 : Valeurs de STIt pour les quatre conditions de test. Pour rappel, les valeurs moyennes de décrétement de performance et de RTLX sont présentées. | 70 |
| Tableau 8 : Valeurs moyennes du STIt pour les ambiances sonores étudiées, considérées comme masquant un signal de parole. | 76 |
| Tableau 9 : Récapitulatif des mesures de performance et de charge cognitive. | 78 |
| Tableau 10 : Catégories utilisées dans l'exercice de fluence verbale. Les fluences moyennes indiquées sont tirées de (Bueno et Megherbi, 2009). | 86 |
| Tableau 11 : Récapitulatif des mesures de performance et de charge cognitive pour le raisonnement grammatical. | 87 |
| Tableau 12 : Récapitulatif des mesures de performance et de charge cognitive pour la fluence verbale. | 87 |
| Tableau 13 : Texte des questions formant le questionnaire NASA-TLX. | 109 |

1 Introduction

L'aménagement en bureaux ouverts est une pratique qui consiste à utiliser un espace de travail ne comportant pas de cloison sur toute la hauteur. Ce type de bureaux, apparu à la fin du XIXème siècle aux Etats-Unis, s'inscrivait à l'origine dans une perspective directement inspirée du taylorisme. Ces « pools » de personnels administratifs, conjuguant « bureaux serrés en rangs compacts, standardisation des équipements et concentration du personnel dans un espace totalement banalisé et transparent » (Fischer, 1990) répondaient à un objectif de rationalisation de l'activité comparable à celle recherchée, au même moment, dans les usines (Labardin, 2011). L'absence de murs permettait alors une supervision du travail plus aisée et une exploitation maximale de l'espace disponible. Plus tard, les exigences spécifiques d'un secteur tertiaire en plein développement ont, petit à petit, conduit à placer l'échange d'information et la communication entre salariés au cœur de la conception des espaces de travail ouverts. Le premier bureau paysager a ainsi vu le jour en Allemagne en 1960. Le bureau est organisé en groupes de postes, répartis suivant la structure de travail de l'entreprise, et non suivant sa structure hiérarchique. Les cloisons disparaissent afin d'assurer une plus grande flexibilité : c'est l'aménagement du mobilier qui délimite les zones fonctionnelles. Les conversations entre collègues y sont vues comme bénéfiques dans la mesure où elles contribuent à la bonne marche de l'entreprise et au bien-être des employés (Evette et Fenker, 1994).

Dès les années 1970-1980, la vérification de la présence de ces effets positifs est devenue un sujet de préoccupation. Il est alors apparu que le passage d'un bureau conventionnel à un bureau ouvert entraîne bel et bien, du point de vue des salariés, une amélioration de l'esthétique de l'espace de travail, une plus grande facilité dans la communication et un sentiment de sociabilité plus grand (Brookes et Kaplan, 1972). Ces avantages sont néanmoins contrebalancés par le sentiment d'être plus à l'étroit, de manquer d'intimité et d'être de plus en plus distrait par le bruit de fond (Brookes et Kaplan, 1972 ; Boyce, 1974 ; Sundstrom *et al.*, 1980 ; Hedge, 1982 ; Zalesny et Farace, 1987). Il s'en suit une perte d'enthousiasme vis-à-vis du travail et des conditions de travail (Oldham et Brass, 1979) ainsi qu'une baisse des performances de travail (Becker *et al.*, 1983). Dans ce contexte, le bruit est cité comme facteur de gêne majeur. La présence de bruit de parole apparaît alors, à l'instar des sonneries de téléphones et des bruits d'équipements (imprimantes, fax, photocopieuses...), comme étant particulièrement gênante (Nemecek et Grandjean, 1973 ; Jensen *et al.*, 2005). Il s'avère de plus que l'intelligibilité des discours environnants a de l'importance : le brouhaha ambiant est d'autant plus perturbant qu'il est possible d'en comprendre des extraits (Boyce, 1974).

Le bruit de parole, parmi l'ensemble des nuisances sonores, semble donc jouer un rôle particulier dans la sensation de gêne chez les salariés en bureaux ouverts. L'intelligibilité du discours semble être, de ce point de vue, un facteur prépondérant. Hongisto (2005) propose ainsi un modèle prédictif de la baisse de performance observée lors de la réalisation d'une tâche dans un environnement sonore donné. Son postulat de départ est le suivant : dans une situation de travail en bureau, des mots ou syllabes seuls ne seront que rarement prononcés. Il faut donc s'attacher à l'intelligibilité des phrases pour caractériser l'ambiance sonore et son impact sur la performance. La représentation de cette fonction présente ainsi une forme sigmoïdale : ceci permet de se rapprocher de la relation entre la valeur d'un indicateur d'intelligibilité classique, le STI, et le score observé lors d'un test d'intelligibilité de phrases. Ce modèle a entraîné la mise en place de recommandations concernant l'aménagement des bureaux ouverts. Dans un article, Virjonen *et al.* (2009) mettent en évidence la possibilité offerte par le modèle de caractériser spatialement le pouvoir perturbateur d'une source de parole. En effet, l'intelligibilité décroît lorsque la distance entre le locuteur et l'auditeur augmente. Lors de la conception d'un espace de travail ouvert, il convient dès lors de chercher à rendre cette décroissance aussi rapide que possible. Ces recommandations font désormais l'objet de la norme internationale ISO 3382-3 (ISO, 2012).

Cette approche basée sur le calcul d'un indicateur d'intelligibilité classique peut cependant être contestée sur plusieurs points. Premièrement, la référence à l'intelligibilité de phrases est à l'origine de la forme sigmoïdale retenue pour la relation entre intelligibilité et diminution des performances. Le modèle prévoit ainsi l'existence d'une zone intermédiaire de valeurs de l'indicateur STI, dans laquelle les variations de performances devraient être importantes. Or, peu d'études présentes dans la littérature fournissent des informations sur l'adéquation du modèle dans cette zone critique : les travaux existants semblent donner des résultats mitigés. Dans un deuxième temps, il convient de remarquer que le choix de l'indicateur en lui-même peut être remis en question. En effet, le STI est efficace lorsqu'il s'agit de tenir compte de situations où la compréhension d'un discours-cible est perturbée par un bruit de fond masquant stationnaire ou par la réverbération. Cependant, dans le cadre de bureaux ouverts, plusieurs locuteurs peuvent simultanément perturber le travail d'un individu. Dès lors se présente la difficulté liée au choix de la voix à considérer comme « cible ». En outre, le calcul de l'intelligibilité doit tenir compte du fait que le masque, constitué de bruit de parole, peut ne pas être stationnaire. Enfin, cette approche est basée sur une hypothèse sous-jacente : celle suivant laquelle l'étude de la performance d'un individu suffit à déterminer le niveau de gêne dont il fait l'expérience. Pourtant, il est possible d'envisager une situation où l'exécutant anticipe l'effet perturbateur de ses conditions de travail et, en conséquence, porte une plus grande attention à la tâche effectuée pour maintenir ses performances à un niveau adéquat. Dans ce cas, l'évaluation de la performance tendrait à sous-estimer l'effet de l'environnement. Un tel

exemple prouve la nécessité de compléter l'attention portée aux performances en accédant à la charge de travail supportée par les exécutants.

Le travail présenté tout au long de ce document se propose d'améliorer les connaissances sur la gêne engendrée par le bruit de parole environnant dans le contexte d'un bureau ouvert en s'intéressant particulièrement à la place prise par l'intelligibilité d'un discours perturbateur. Pour ce faire, il est tout d'abord nécessaire d'explicitier la notion de nuisance sonore dans un contexte de travail et de se donner les moyens de l'évaluer. Il faut de plus comprendre quelles sont les propriétés physiques d'un bruit de parole permettant une évaluation objective de son intelligibilité. Il est enfin nécessaire de faire le lien entre l'influence qu'ont respectivement ces différents facteurs sur l'intelligibilité d'un discours et sur le pouvoir perturbateur de celui-ci. L'état des connaissances actuelles sur le sujet est recensé dans le chapitre 2. La suite de ce travail présentera les dispositifs expérimentaux mis en place afin de répondre à l'objectif fixé ainsi que les résultats obtenus. Le chapitre 3 est ainsi porté sur l'étude de l'effet de deux facteurs agissant sur l'intelligibilité : le contenu énergétique du bruit masquant et les variations de son niveau dans le temps. Les deux expériences présentées permettent de fournir des indications sur l'adéquation d'indicateurs d'intelligibilité classiques tels que le STI à la problématique de l'évaluation de la gêne sonore en bureaux ouverts. Le chapitre 4, tenant compte des résultats précédemment obtenus, propose un indicateur a priori mieux adapté et, au travers de deux expériences, le confronte à des signaux réels enregistrés en bureaux ouverts. Le dernier chapitre récapitule l'ensemble des conclusions que ce travail aura permis de tirer et offre des perspectives pour la poursuite d'études sur le sujet.

2 Etat de l'art

2.1 Nuisance sonore au travail

Pour se donner les moyens de quantifier la nuisance sonore, il faut tout d'abord s'accorder sur une définition à lui donner. L'enquête menée par Guski et al. (1999) auprès de spécialistes du domaine offre un aperçu de la difficulté de la tâche. Les auteurs de cette étude ont interrogé 68 chercheurs travaillant dans le domaine de la nuisance sonore. Au travers de questionnaires, les experts étaient interrogés sur les significations qu'ils attachaient aux termes « nuisance sonore ». Il ressort de l'analyse que les réponses sont très variées ; toutefois, la nuisance sonore semble pouvoir être abordée depuis deux angles d'observation : l'aspect émotionnel d'une part et l'aspect comportemental de l'autre.

2.1.1 Différents aspects de la nuisance sonore

L'aspect émotionnel s'attache à l'évaluation de la qualité du bruit ambiant, de son caractère plus ou moins plaisant. A cet égard, un environnement sonore est gênant lorsqu'il est considéré comme irritant, désagréable à l'oreille. Il s'agit alors de déterminer quelles caractéristiques d'un son le rendent plus ou moins désagréable, telles que son niveau ou encore sa composition spectrale.

Le caractère d'un environnement sonore le plus aisément lié à une nuisance exprimée est son niveau sonore. Ainsi, le niveau de pression sonore équivalent pondéré A, $L_{A,eq}$, exprimé en décibels (dB(A)), est couramment employé pour juger de la qualité de l'ambiance sonore d'un espace de travail. La valeur de cet indicateur est obtenue en calculant le niveau qu'un bruit stationnaire devrait avoir pour contenir, sur la même période, la même énergie que le bruit variable mesuré :

$$L_{A,eq} = 10 \log_{10} \left(\frac{1}{T} \int_0^T 10^{L_A(t)/10} dt \right), \quad (2.1)$$

où T représente la durée de mesure et L_A le niveau de pression sonore pondéré A. Dans un bureau ouvert, au poste de travail, les niveaux observés sont autour de 46 dB(A) (Warnock et Chu, 2002 ; Bradley et Gover, 2004), avec des variations dépendant de la nature de l'activité. Ainsi, dans un centre d'appel, dans lequel un grand nombre d'opérateurs parlent simultanément, le niveau se situe plutôt autour de 55 dB(A) (Trompette et Chatillon, 2010). Le niveau de pression sonore équivalent est bien corrélé avec les jugements subjectifs d'occupants de bureaux ouverts en ce qui concerne leur seul confort acoustique (Ayr *et al.*, 2003).

Au-delà du niveau moyen, la présence d'émergences est également source de désagrément. Afin de mesurer leur prévalence, des indicateurs statistiques tels que le $L_{A,5}$ ou le $L_{A,10}$ sont employés : ils correspondent respectivement au niveau sonore dépassé pendant 5% et 10% du temps

de mesure. De même, le $L_{A,90}$ est utilisé pour rendre compte de l'impact du niveau du bruit de fond sur la nuisance exprimée : sa valeur est celle du niveau sonore dépassée pendant 90% du temps de mesure.

D'autres indicateurs sont basés sur l'idée qu'au-delà de la simple prise en compte de la sensibilité de l'oreille à différentes fréquences permise par la pondération A, certaines plages de fréquences contribuent plus que d'autres à la nuisance sonore ressentie dans les bureaux. De ce fait, le spectre mesuré est comparé à un ensemble de courbes de référence, représentant chacune un niveau de gêne. Sont ainsi utilisées les *noise criterion curves* (NC) et leurs dérivées : *preferred noise criterion* (PNC), *balanced noise criterion* (NCB), ou encore les *room criterion curves* (RC). En Europe, les *noise rating curves* (NR) sont plus usuelles (AFNOR, 1974).

Il est possible de se référer autrement à la perception du niveau sonore par un individu. Les variations de cette sensation peuvent en effet différer de celles du L_A , notamment à cause d'effets de masquage entre bandes de fréquence adjacentes ou entre événements sonores successifs. L'indicateur psychophysique adéquat est alors la sonie, dont le calcul est normalisé (ISO, 1975 ; 2007). En outre, pour tenir compte de l'influence de la présence de hautes fréquences dans le contenu spectral, l'acuité est introduite. La perception des variations lentes de la sonie est quantifiée par la force de fluctuation ; la rugosité rend compte de la modification de timbre engendrée par des variations plus rapides du niveau.

L'existence d'une nuisance « sans biais », indépendante du contexte, est cependant remise en cause par la seconde signification, correspondant aux implications d'ordre comportemental, qui se dégage de l'enquête de Guski *et al.* (1999) : la nuisance sonore peut aussi prendre le sens de dérangement, d'interférence avec une activité en cours dont la présence du bruit vient perturber la réalisation. A ce titre, Zimmer *et al.* (2008) ont montré que l'évaluation de la gêne par des auditeurs est modifiée s'il leur est demandé d'accomplir une tâche particulière pendant l'exposition au bruit. La notion d'interférence semble bel et bien avoir une importance prépondérante. Dans cette étude, il est de plus remarquable que pendant la réalisation de la tâche, les évaluations correspondant à la nuisance sonore sont en concordance avec les performances de mémorisation observées : l'environnement jugé le moins acceptable est celui pour lequel le plus d'erreurs sont commises. En témoignant de la gêne subie, les sujets semblent être en mesure de rendre compte de la façon dont le bruit augmente la difficulté de la tâche. Ceci prend une importance particulière dans les cas des bureaux ouverts. En effet, le bruit n'y a pas un niveau suffisamment élevé pour être nuisible en lui-même ; c'est du fait de sa présence pendant le travail que le fond sonore est gênant. Ainsi, pour mesurer la nuisance sonore au cours d'une action, il faut savoir quantifier l'accroissement de difficulté qui en résulte dans l'activité simultanée.

2.1.2 Perturbation d'une activité : la surcharge cognitive

A mesure qu'une tâche est rendue plus complexe, les ressources cognitives nécessaires à sa réalisation augmentent et l'exécutant tend alors à augmenter son investissement dans la mesure de ses capacités (Delignières, 1993). Cependant, dans un article précurseur, Miller (1955) avance que les capacités mentales de tout un chacun sont limitées par le chiffre « magique » 7 ± 2 . D'après les résultats présentés dans cet article, un individu moyen n'est en mesure de retenir ou de différencier qu'environ sept faisceaux d'informations. Aussi les ressources cognitives auxquelles une personne peut faire appel sont elles finies. La part de ces ressources mises en œuvre au cours de la réalisation d'une tâche donnée est la charge cognitive liée à cette activité (Paas et van Merriënboer, 1994a ; Paas *et al.*, 2003). L'élaboration d'une théorie de la charge cognitive prend sa source dans le domaine des sciences de l'éducation (Sweller, 1988). Il s'agit alors de comprendre comment concevoir au mieux les instructions d'un exercice afin de s'assurer que l'apprenant aura pu assimiler les schémas nécessaires à sa résolution. Pour cela, pour un exercice de difficulté donné, il est nécessaire de limiter au maximum la complexité supplémentaire ajoutée par les consignes données afin de laisser autant de ressources que possible à l'élève pour intégrer les bons mécanismes de réflexion.

L'environnement sonore dans lequel une activité doit être réalisée peut avoir un impact sur la charge cognitive supportée (Moreno et Mayer, 2000). Pour quantifier cette influence, il faut pouvoir mesurer la surcharge cognitive engendrée par l'ambiance sonore. Cette difficulté accrue peut se traduire directement par une baisse de la performance ou, plus subtilement, par une perception plus aigüe par l'exécutant lui-même de sa charge de travail.

2.1.3 Mesures de la surcharge cognitive

2.1.3.1 Mesures objectives

Le phénomène par lequel un individu voit ses performances diminuer en présence d'un bruit de fond n'ayant, *a priori*, aucun rapport avec l'activité en cours est appelé *irrelevant sound effect* (ISE). Cet effet a été souvent étudié dans le domaine de la psychologie cognitive (Hughes et Jones, 2001). Pour ce faire, il convient tout d'abord de déterminer les tâches proposées aux exécutants. Certaines études s'appuient sur des tâches s'approchant autant que possible d'un réel travail de bureau. Se retrouvent ainsi des exercices de tri de courrier (Venetjoki *et al.*, 2006), de calcul mental (Banbury et Berry, 1998), de raisonnement verbal (Schlittmeier *et al.*, 2008b), de compréhension écrite (Oswald *et al.*, 2000) ou encore de rédaction (Randsell et Gilroy, 2001) ou de correction de texte (Jones *et al.*, 1990). D'autres exercices, moins spécifiques, sont utilisés parce qu'ils font appel à des mécanismes courants dans l'exécution de nombreuses autres tâches, comme par exemple l'attention visuelle. Parmi celles-ci, les tâches les plus souvent employées dans les études sur l'ISE

testent la mémoire sérielle, c'est-à-dire la capacité à garder en mémoire non seulement une série d'éléments mais également l'ordre dans lequel ces éléments ont été présentés. Le choix de tâches de mémoire sérielle s'explique, d'une part, par l'importance de ce mécanisme dans un grand nombre d'activités mentales quotidiennes (Lashey, 1951 ; Rosenbaum *et al.*, 2007). D'autre part, ce type d'activité semble particulièrement sensible à la présence d'un bruit perturbateur. Ainsi, en présence de certains types de bruit, la performance dans une tâche de mémoire sérielle diminue significativement par rapport à ce qui est observé dans le silence (Banbury *et al.*, 2001).

Pour autant, les seules mesures de performances peuvent s'avérer trompeuses. En effet, de manière générale, l'ajout d'une difficulté entraîne généralement une augmentation de l'effort fourni par une personne devant effectuer une tâche, dans la mesure où il lui est encore possible de redoubler d'attention (Delignières, 1993). Suivant la nature de la tâche et de la perturbation, cet effort supplémentaire peut suffire à combler le déficit de performance engendré par l'interférence du perturbateur (Kahneman, 1973 ; Dornic *et al.*, 1974). L'observation de la seule performance ne rend donc pas systématiquement compte de l'intensité de l'interférence entre une tâche à réaliser et un bruit de parole perturbateur. Ce constat met en évidence la nécessité d'accéder, en complément de la performance d'un sujet lors d'une activité donnée, à la charge de travail qu'il pense supporter. Pour être en mesure de quantifier cette perception, trois catégories de méthodes peuvent être distinguées.

Lorsqu'un investissement supplémentaire en ressources cognitives est nécessaire, l'adaptation de l'exécutant aux nouvelles exigences peut s'observer directement au travers de certaines mesures physiologiques. Ainsi, une augmentation de la difficulté d'une tâche a pour conséquence d'augmenter la variabilité du rythme cardiaque, la pression artérielle ou encore le tonus musculaire (Hanson *et al.*, 1993). De fait, l'étude de ces réponses peut être envisagée pour la mesure de la charge cognitive. L'observation de la dilatation de la pupille fournit par exemple une mesure sensible et réactive de l'effort mental consenti (Beatty et Lucero-Wagoner, 2000 ; van Gerven *et al.*, 2004). Ces techniques présentent l'avantage de fournir des résultats aussi objectifs que possible ; cependant, un appareillage particulier est souvent nécessaire. Cela tend à rendre la mesure trop intrusive : le protocole de mesure de la charge cognitive risque alors d'avoir une trop grande influence sur le déroulement de l'exercice (Paas et van Merriënboer, 1994b).

Une autre façon d'accéder à l'effort mental peut être de mesurer la baisse de performance observée dans une tâche secondaire effectuée simultanément. Dans ce cadre, les performances lors des tâches principales et secondaires seront respectivement appelées performances primaires et secondaires. Il s'agit alors de considérer, d'une part, que les ressources cognitives d'un individu sont à tout moment limitées et, d'autre part, que cet individu est capable de les allouer en priorité à l'exécution d'une tâche qu'il aura jugée prioritaire (Pashler, 1994). Toute diminution des

performances secondaires par rapport à une situation originale sera alors considérée comme le signe que la tâche principale, prioritaire, nécessite de réquisitionner plus de ressources. Ceci signifie alors que la nouvelle situation a rendu la tâche principale plus difficile. Cette méthode, fiable et sensible, s'accompagne néanmoins d'importantes précautions. Il convient en effet de vérifier que la tâche secondaire choisie est insensible à la présence de bruit de parole lorsqu'elle est effectuée indépendamment de la tâche principale. Si cette condition n'est pas vérifiée, la comparaison entre une situation calme et une situation en présence de masques de parole est impossible. La dégradation des performances peut alors tout autant être causée par l'augmentation de l'effort sur la tâche primaire que par la seule présence du bruit de parole, affectant directement l'exécution de la tâche secondaire. De plus, lorsque la tâche principale est d'emblée très difficile, il existe le risque de soumettre le sujet à des conditions rapidement difficiles à supporter.

2.1.3.2 Mesures subjectives

Il est enfin possible de demander directement aux sujets de faire part de la manière dont ils ont vécu l'expérience. Il s'agit alors de présenter aux exécutants, à l'issue de leur travail, une ou des interrogations sur le déroulement de l'exercice passé. Parmi les mesures subjectives de la charge de travail les plus établies, il se trouve, d'une part, les mesures unidimensionnelles. La *Modified Cooper-Harper Scale* (MCH) et l'échelle *Overall Workload* (OW) en sont des exemples (Hill *et al.*, 1992). Il s'agit d'échelles uniques, grâce auxquelles les sujets peuvent rendre compte de leur charge de travail par l'intermédiaire d'une ou d'une succession de questions. Ces méthodes présentent l'avantage d'être simples et rapides à mettre en œuvre comme à compléter. Cependant, elles fournissent un score unique, qui ne peut pas être décomposé suivant les différentes composantes de la charge cognitive.

Ainsi, dans le but de permettre une analyse plus précise des sources de gêne, des mesures multidimensionnelles telles que le *Task Load Index* (NASA-TLX, (Hart et Staveland, 1988)) ou la *Subjective Workload Assessment Technique* (SWAT) sont couramment utilisées (Hill *et al.*, 1992). La mise en place de ces techniques est plus complexe que pour les mesures unidimensionnelles, mais il est alors possible d'établir un diagnostic quant aux composantes de la charge cognitive les plus sollicitées par une tâche ou les plus affectées par une modification de l'environnement. Dans les deux cas, un questionnaire est présenté aux sujets afin de décrire la charge mentale subie durant la tâche suivant plusieurs dimensions. Une procédure de pondération des dimensions permet enfin d'obtenir un score global. Cependant, cette dernière étape, la plus consommatrice de temps, n'est pas toujours nécessaire : une simple moyenne sur les composantes permet de parvenir à des résultats aussi bons, voire meilleurs, que les multiples comparaisons nécessaires prescrites par les protocoles originaux (Byers *et al.*, 1989 ; Luximon et Goonetilleke, 2001). Il est alors possible de

construire des échelles dérivées plus simples à utiliser : le *Raw Task Load Index* (NASA-RTLX) et le A_{SWAT} . Ces différentes échelles sont généralement comparées suivant leurs sensibilités respectives, c'est-à-dire leur capacité à différencier deux conditions de charge. A ce titre, l'échelle *Overall Workload* et le NASA-TLX sont considérées comme étant meilleures (Hill *et al.*, 1992). Du fait de leur caractère multidimensionnel, les capacités de diagnostic du NASA-TLX et de sa forme simplifiée, le NASA-RTLX, en font les méthodes les plus couramment utilisées (Hart, 2006).

2.1.3.3 Le questionnaire NASA-TLX

Le questionnaire permettant d'obtenir le NASA-TLX comporte six questions, chacune correspondant à une composante de la charge cognitive (NASA, 1986 ; Hart et Staveland, 1988). Les trois premières concernent les exigences intrinsèques de la tâche, c'est-à-dire les ressources requises par la tâche indépendamment de la personne qui la réalise. Les composantes concernées sont :

- l'exigence mentale : l'importance des processus mentaux et perceptuels appelés,
- l'exigence physique : la dépense physique nécessaire,
- l'exigence temporelle : la pression temporelle imposée.

Les trois autres questions relèvent de l'interaction entre la tâche et le sujet, suivant les composantes suivantes :

- la performance : l'accomplissement d'objectifs explicites ou implicites,
- l'effort fourni pour atteindre ce niveau de performance,
- la frustration : l'intensité, faible ou forte, de sentiments d'aise et de confiance.

Le sujet répond à chaque question à l'aide d'échelles comportant 21 graduations, ce qui permet d'obtenir six scores allant, de cinq en cinq, de 0 à 100. Le texte original des questions, en anglais, est présenté dans l'annexe A.

Il est envisageable que pour une tâche donnée, les différentes composantes n'aient pas le même poids dans la charge cognitive subie par le sujet. De même, tous les sujets ne sont peut-être pas sensibles de la même façon à chacune des composantes. En conséquence, à l'issue de la réalisation de l'activité, une série de comparaisons par paires des composantes est effectuée. Pour chacune des quinze paires de composantes qu'il est possible de former, il est demandé au sujet de désigner celle qui lui semble être la plus importante pour cette activité. Un poids, allant de 0 à 5, est attribué à chaque composante suivant le nombre de fois qu'elle a été choisie. Ces paramètres permettent finalement d'effectuer une moyenne pondérée des six scores afin d'obtenir un score global, le *Task Load Index*.

Ces comparaisons par paires allongent de façon non négligeable la durée de passage d'un sujet et complexifient la méthode. Byers *et al.* (1989) montrent pourtant que les avantages attendus en compensation de ces difficultés ne sont pas toujours suffisamment observables. Il est possible

d'effectuer une moyenne non pondérée des scores des six composantes et d'obtenir ainsi un indicateur aussi performant que l'original. Cette procédure simplifiée, suivant laquelle les six composantes ont le même poids relatif dans le calcul, donne le *Raw Task Load Index*, ou RTLX.

Il apparaît finalement que dans un contexte de travail en bureau ouvert, la gêne sonore doit être considérée comme une surcharge cognitive causée par le bruit ambiant. La gêne peut alors être caractérisée par deux méthodes complémentaires : l'une directe, porte sur la mesure des performances. L'autre, indirecte, repose sur une évaluation subjective par l'exécutant de sa charge de travail. Dans le cas de la gêne causée par un bruit de parole, l'intelligibilité du discours est une propriété influente. Le concept de gêne sonore vient d'être éclairci : il convient d'en faire de même pour celui d'intelligibilité.

2.2 Intelligibilité de la parole

Dans le langage courant, l'intelligibilité d'un discours est définie comme le degré de compréhension du contenu du message. La manière la plus directe de mesurer l'intelligibilité d'un discours est donc de recourir à des tests subjectifs où un auditeur doit reconnaître des syllabes de type CVC (consonne-voyelle-consonne) ou VCV, des mots ou même des phrases (tests dits "aux mots" ou "aux phrases") (Kryter, 1962b). Cependant, du fait du nombre de sujets nécessaire et du coût des tests en temps et en moyens, des essais de ce type ne sont pas adaptés en pratique à la caractérisation d'une ambiance sonore. À cette fin, divers indicateurs objectifs d'intelligibilité, se basant uniquement sur des caractéristiques physiques des signaux sonores, ont été développés. Leur étude, dans ce chapitre, mettra en évidence les différents déterminants de l'intelligibilité de la parole.

2.2.1 L'Articulation Index et le Speech Intelligibility Index

Dans un article étudiant les facteurs régissant l'intelligibilité des bruits de parole, French et Steinberg proposent l'*Articulation Index* (indice d'articulation, AI) (French et Steinberg, 1947). Cet indicateur prend ses valeurs entre 0 et 1 et croît avec l'intelligibilité du signal de parole étudié. Les paragraphes suivants exposent les principes qui sous-tendent le calcul de cet indice, synthétisés dans la Figure 1.

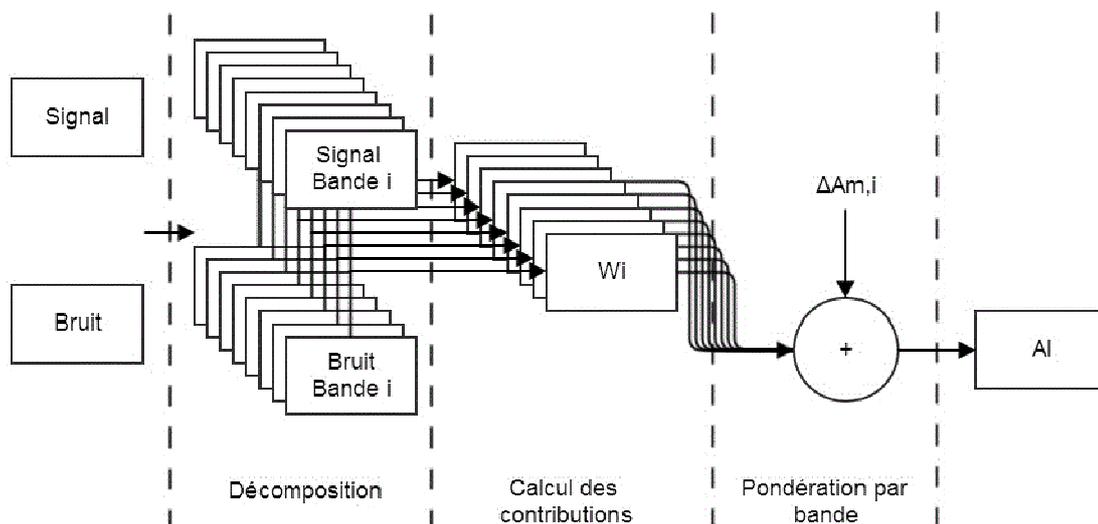


Figure 1 : Schéma de calcul de l'Articulation Index.

French et Steinberg avancent que l'information portée par le signal de parole est distribuée sur l'ensemble du spectre de la parole. Chaque bande de fréquence de ce spectre contient alors une

part de l'information à transmettre. Plus cette part est grande, plus la bande de fréquence étudiée contribue à l'intelligibilité globale du signal. Cette contribution est, d'une part, indépendante de celle d'une autre bande de fréquence dès lors que les deux intervalles considérés sont disjoints. D'autre part, la contribution d'un ensemble de bandes disjointes à l'intelligibilité du signal est égale à la somme des contributions respectives de chaque bande. Pour une partition donnée $(\Delta f_i)_{i=1..N}$ du spectre de la parole, ceci peut s'écrire comme suit :

$$AI = \sum_{i=1}^N \Delta A_{m,i} W_i . \quad (2.2)$$

Dans cette équation, la valeur d'une donnée W_i témoigne de la qualité de la communication pour la bande de fréquence i . En effet, dans des conditions de communication adverses, une bande de fréquence donnée peut ne contribuer que pour partie de sa contribution maximale. Si, dans cette bande de fréquence, l'information est parfaitement exploitable par l'auditeur, alors W_i vaut 1. A l'inverse, si l'information contenue dans cette bande de fréquence est de trop mauvaise qualité pour servir à la compréhension du message, W_i vaut 0. La valeur de W_i dépend, d'une part, du niveau de la parole à l'oreille de l'auditeur. En effet, si la voix du locuteur est trop faible ou trop atténuée par la chaîne de transmission, le discours peut devenir inintelligible. D'autre part, pour calculer W_i , il est nécessaire de déterminer l'influence des autres facteurs pouvant modifier la situation de communication. En ce sens, French et Steinberg mettent en évidence l'effet du masquage énergétique. Lorsque le niveau du bruit de fond dans une bande de fréquence devient trop élevé par rapport à celui du signal de parole, la transmission de l'information est dégradée. Dès lors, la connaissance des SNR dans chacune de ces bandes fournit une grandeur représentative de l'intelligibilité d'un bruit de parole masqué.

Les valeurs des poids utilisés dans cette moyenne pondérée, notés $\Delta A_{m,i}$, correspondent aux contributions maximales de chaque bande de fréquence à l'intelligibilité du signal. French et Steinberg, par exemple, proposent une partition du spectre pour laquelle la valeur des $\Delta A_{m,i}$ est constante suivant i : les bandes de fréquences contribuent alors toutes de la même façon à l'intelligibilité du signal. De manière plus générale, la partition du spectre de la parole employée détermine la valeur des poids à utiliser. Kryter récapitule ainsi un ensemble de méthodes de calcul de l'AI permettant de l'adapter à différentes partitions du spectre de la parole : bandes de poids égal mais aussi bandes d'octave et tiers de bandes d'octaves (Kryter, 1962a).

Une forme révisée de l'AI, le *Speech Intelligibility Index* (indice d'intelligibilité de la parole, SII), est normalisée (ANSI, 1997). Le calcul du SII est similaire à celui de l'AI : les valeurs des $\Delta A_{m,i}$ pour différentes partitions du spectre ont été actualisées. L'*Articulation Index* et le *Speech Intelligibility Index* sont parmi les indicateurs d'intelligibilité les plus couramment utilisés.

2.2.2 Importance de la conservation des modulations du signal

Il est cependant possible d'accéder à une meilleure compréhension de l'influence du masquage énergétique sur l'intelligibilité. En effet, comme énoncé plus haut, les signaux de parole sont naturellement fluctuants. Ces fluctuations, visibles à l'étude du spectre de modulation d'un signal de parole classique (Houtgast et Steeneken, 1972), correspondent à la succession des éléments constitutifs du langage (syllabes, voire phonèmes) porteurs de l'information que le locuteur cherche à transmettre. Ces variations rapides de niveau peuvent être étudiées par l'intermédiaire du spectre de modulation d'un signal de parole classique.

Le spectre utile de la parole est généralement considéré comme se trouvant dans les bandes d'octave centrées sur des fréquences allant de 125 Hz à 8000 Hz. Cependant, le spectre de modulation de la parole, c'est-à-dire le spectre de l'enveloppe de la courbe d'intensité du signal de parole, ne s'étend de façon significative que jusqu'à 20 Hz, environ, avec un maximum entre 3 Hz et 5 Hz (Houtgast et Steeneken, 1972 ; Greenberg et Kingsbury, 1997). Ces fréquences correspondent au rythme des mouvements combinés des articulateurs de la parole (lèvres, mâchoires, langue, voile du palais, glotte) lors de l'élocution (Browman et Goldstein, 1992 ; Smith *et al.*, 1993 ; Greenberg, 2004). Le pic, en particulier, peut être rapporté à la durée moyenne des syllabes dans le discours usuel (Greenberg *et al.*, 2006).

Il est intéressant de remarquer que ces modulations sont les mieux perçues par l'être humain. En effet, des protocoles centrés autour de la détection de modulation ont été utilisés dans de nombreuses études (Riesz, 1928 ; Viemeister, 1979). Il y est demandé à des sujets de détecter la présence d'un signal modulé en amplitude présenté en concurrence d'un masque stationnaire de même contenu spectral. La porteuse peut être une sinusoïde pure ou être plus ou moins à large bande. Le seuil de détection d'un signal à une fréquence de modulation donnée est caractérisé par l'indice de modulation m . Pour un signal modulant sinusoïdal, il est défini comme suit :

$$m = \frac{A_{max} - A_{min}}{A_{max} + A_{min}}, \quad (2.3)$$

où A_{max} et A_{min} représentent respectivement la valeur maximale (la « crête ») et la valeur minimale (le « creux ») de l'enveloppe de l'intensité du signal. Ainsi, lorsque m vaut zéro, la crête et le creux de l'enveloppe sont au même niveau : il n'y a pas de modulation effective. A mesure que l'indice de modulation augmente, tout se passe comme si le creux s'affirmait ; les crêtes sont de plus en plus distinctes. Lorsque m atteint une valeur de 1, le creux atteint le niveau zéro, la modulation est au plus fort. Il apparaît alors que le système auditif humain détecte particulièrement bien le signal variable pour des fréquences de modulations comprises entre 0 et 32 Hz (Chi *et al.*, 1999). Au-delà, la modulation est alors moins perçue comme une fluctuation de niveau que comme une

modification du timbre. La sensibilité aux modulations d'amplitude présente un maximum autour de 4 Hz.

De plus, ces modulations sont porteuses d'information. En effet, lorsqu'un signal de parole est altéré par la diminution de l'indice de modulation dans cette plage de fréquence, l'intelligibilité en diminue (Houtgast et Steeneken, 1973). Ces modifications constituent un « étalement » temporel du signal ; les limites entre syllabes sont brouillées (Drullman *et al.*, 1994a ; b). Ainsi, lorsqu'en diminuant son contenu en modulation, le signal de parole est rendu plus difficile à segmenter, son intelligibilité décroît. Or, comme précédemment évoqué, une diminution du SNR tend également à altérer l'intelligibilité d'un signal de parole. Il s'avère en effet que la présence d'un bruit de fond stationnaire revient à abaisser le spectre de modulation du signal à transmettre.

Cela peut être démontré en considérant comme signal un bruit blanc d'intensité moyenne I_p modulé par une sinusoïde de fréquence F , de façon à ce que l'indice de modulation vaille 1. L'expression analytique de l'enveloppe de l'intensité du signal est la suivante :

$$x(t) = I_p (1 + \cos 2\pi Ft) . \quad (2.4)$$

Du fait de l'ajout d'un bruit blanc stationnaire d'intensité moyenne I_b , l'enveloppe de l'intensité du signal à l'oreille de l'auditeur s'exprime analytiquement comme suit :

$$y(t) = I_p (1 + \cos 2\pi Ft) + I_b , \quad (2.5)$$

ce qui peut également s'écrire ainsi :

$$y(t) = (I_p + I_b) \left(1 + \frac{I_p}{I_p + I_b} \cos 2\pi Ft \right) . \quad (2.6)$$

La nouvelle valeur de l'indice de modulation est donc

$$m = I_p / (I_p + I_b) . \quad (2.7)$$

Cette valeur est plus petite que 1 : les modulations du signal d'entrée ont été atténuées. De plus, la relation (2.7) peut également s'écrire comme suit :

$$m = \frac{10^{SNR/10}}{1 + 10^{SNR/10}} . \quad (2.8)$$

Ainsi, m est une fonction monotone croissante du rapport signal-à-bruit. C'est ainsi que s'explique l'effet d'une diminution du SNR sur l'intelligibilité d'un signal de parole.

De manière plus générale, dans le cadre de l'acoustique des salles, lorsqu'un signal est transmis d'un locuteur à un auditeur au travers d'un local, l'enveloppe de la courbe d'intensité du signal reçu est une forme lissée de celle de l'enveloppe émise. L'effet de la salle sur la transmission est caractérisé par la fonction de transfert reliant le spectre des deux enveloppes, c'est-à-dire les spectres de modulation des signaux d'entrée et de sortie du système. Il s'agit de la fonction de transfert de modulation (MTF, pour *modulation transfer function*) (Houtgast et Steeneken, 1973).

Pour la démonstration suivante, le même signal d'entrée que précédemment est considéré :

$$x(t) = I_p (1 + \cos 2\pi Ft). \quad (2.9)$$

Du fait des transformations apportées au signal d'entrée par la salle, l'enveloppe de l'intensité du signal de sortie, perçu par l'auditeur, s'exprime comme suit :

$$y(t) = I_s [1 + m \cos 2\pi F(t + \tau)]. \quad (2.10)$$

Ici, I_s représente l'intensité moyenne du signal de sortie. Les quantités m et τ correspondent aux réponses en amplitude et en phase de la chaîne de transmission. En particulier, l'indice de modulation a changé : égal à l'unité en entrée, il prend désormais la valeur m , inférieure à 1. *A priori*, la valeur de l'indice de modulation peut dépendre de la fréquence de modulation à laquelle il est étudié. La MTF est alors définie comme le rapport entre les indices de modulation en sortie et en entrée. Dans le cas où l'indice vaut 1 en entrée, ceci s'exprime ainsi :

$$MTF : F \mapsto m(F). \quad (2.11)$$

Ainsi, comme démontré précédemment, dans le cas où seul un bruit stationnaire masquant est présent, l'atténuation est constante pour toutes les fréquences de modulation :

$$m(F) = \frac{10^{SNR/10}}{1 + 10^{SNR/10}}. \quad (2.12)$$

Ce n'est pas le cas pour une salle dont le temps de réverbération est égal à T ; l'atténuation de la modulation est alors décrite par la fonction suivante :

$$m(F) = \frac{1}{\sqrt{1 + \frac{1}{4}F^2T^2}}. \quad (2.13)$$

La réverbération agit comme un filtre passe-bas sur la modulation ; les variations les plus rapides de l'intensité du signal sont fortement atténuées. La compréhension des consonnes, en particulier, est fortement dégradée (Drullman *et al.*, 1994b). Dans le cas particuliers de bureaux ouverts, Virjonen *et al.*, à partir d'un échantillon de 16 bureaux finlandais, obtiennent un T_{20} moyen de 0,57 secondes (Virjonen *et al.*, 2009). Pour cette valeur, environ un tiers des modulations contenues dans le signal original autour de 4 Hz sont perdues.

Ainsi, étant donnée l'importance de la conservation des modulations dans la bonne transmission d'un message parlé, la connaissance de la MTF permet de quantifier l'influence d'une chaîne de transmission sur l'intelligibilité d'un bruit de parole. Il est possible de construire un indice d'intelligibilité basé sur la MTF : le plus couramment utilisé est le STI (*speech transmission index*, (Steeneken et Houtgast, 1980 ; Houtgast et Steeneken, 1985)).

2.2.3 Le Speech Transmission Index

Comme précédemment pour l'AI, le STI est calculé comme une combinaison linéaire des contributions de bandes de fréquence adjacentes à l'intelligibilité d'un signal de parole. L'ensemble de la procédure de calcul est schématisée dans la Figure 2.

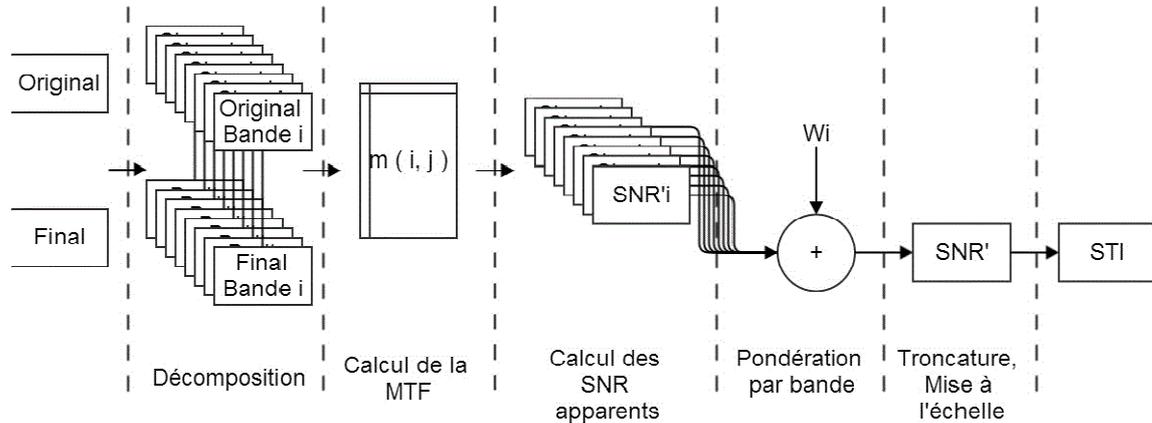


Figure 2 : Schéma de calcul du STI.

Le spectre de la parole est ainsi divisé en sept bandes d'octave allant de 125 Hz à 8 kHz. Dans chacune d'entre elles, la MTF est déterminée pour les 14 fréquences de modulation correspondant aux centres des bandes de tiers d'octaves allant de 0,63 Hz à 12,5 Hz. De cette façon, une matrice $(m_{i,j})$ est obtenue, pour $i = 1 \dots 7$ et $j = 1 \dots 14$.

L'étape suivante du calcul consiste donc à fournir pour chaque coefficient de la matrice un rapport signal à bruit apparent, noté $SNR'_{i,j}$, en s'inspirant de la formule (2.8) :

$$SNR'_{i,j} = 10 \log \left(\frac{m_{i,j}}{1 - m_{i,j}} \right). \quad (2.14)$$

Il s'agit de déterminer, quelle que soit l'altération réelle subie par le signal, de quelle valeur du rapport signal à bruit aurait pu résulter la même valeur de la MTF à cette fréquence de modulation, dans cette bande du spectre audible. Théoriquement, dans le cas d'une atténuation de la modulation, la valeur de m est comprise entre 0 et 1. Ainsi, la plage de variation des $SNR'_{i,j}$ s'étend de $-\infty$ à $+\infty$. Pour la suite du calcul, les valeurs possibles des $SNR'_{i,j}$ seront limitées entre -15 et +15 dB. Cette plage est choisie car dans le cas où la seule perturbation provient d'un bruit de fond stationnaire, lorsque le niveau du bruit masquant est supérieur de 15 dB à celui du signal de parole, il est d'ores et déjà parfaitement impossible de déchiffrer ce dernier ; une augmentation du bruit de fond ne pourrait donc pas détériorer la situation. Inversement, lorsque le niveau du bruit masquant est inférieur de 15 dB à celui du signal de parole, la compréhension est parfaite : une

diminution du bruit de fond ne pourrait pas rendre la situation meilleure. D'après Houtgast et Steeneken (1973), il est légitime de considérer que les contributions de chacune des fréquences de modulation sont égales. De ce fait, le rapport signal à bruit apparent global pour chaque bande de fréquence du spectre audible, noté SNR'_i , est défini comme la simple moyenne des $SNR'_{i,j}$ suivant la fréquence de modulation :

$$SNR'_i = \frac{1}{14} \sum_{j=1}^{14} SNR'_{i,j} \quad (2.15)$$

Une combinaison linéaire des SNR apparents par bande est ensuite établie :

$$\overline{SNR'} = \sum_{i=1}^7 w_i SNR'_i. \quad (2.16)$$

Les coefficients de pondération $w_i, i = 1 \dots 7$, de somme égale à 1, ont été déterminés à la suite de mesures d'intelligibilité. Il ressort de ce processus un indicateur d'intelligibilité basé sur la MTF, construit par combinaison linéaire des contributions des 7 bandes d'octaves principales du spectre de la parole. Du fait de la troncature effectuée précédemment, $\overline{SNR'}$ prend ses valeurs dans la plage ± 15 dB. Pour obtenir un indicateur prenant ses valeurs entre 0 et 1, tout en conservant une dépendance linéaire aux valeurs de rapport signal à bruit, le STI est finalement calculé comme suit :

$$STI = \frac{\overline{SNR'} + 15}{30}. \quad (2.17)$$

La valeur du STI se présente alors comme la proportion de segments acoustiques conservés au cours de la transmission du signal. En effet, la relation reliant le STI au score d'intelligibilité de mots-syllabes est proche de la linéarité, comme le montre la Figure 3. Pour un STI valant 0, l'information est complètement brouillée : le discours est inintelligible. Au contraire, lorsque le STI est égal à 1, le contenu est parfaitement restitué et le locuteur est entièrement compris par son auditeur. Pour autant, le discours courant est constitué de phrases cohérentes. Un auditeur attentif est généralement en mesure de comprendre le sens d'un énoncé sans pour autant en avoir entendu l'intégralité. Pour ce faire, il utilise le contexte formé par les éléments entendus correctement pour reconstituer l'information manquante. C'est pourquoi il n'est pas nécessaire d'avoir une transmission du signal parfaite pour avoir un discours d'intelligibilité satisfaisante. En pratique, l'intelligibilité d'un discours est considérée comme excellente au-delà de $STI = 0,7$. La version normalisée du STI, plus élaborée, permet de tenir compte de l'audibilité du signal ainsi que du masquage inter-bandes (IEC, 2003).

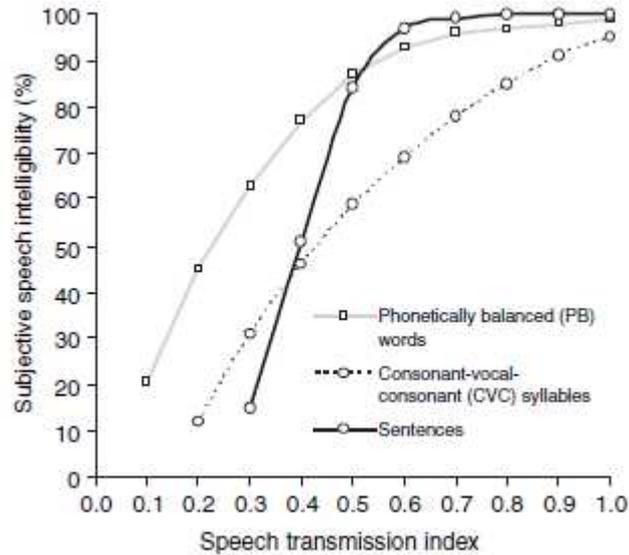


Figure 3 : Relation entre la valeur du STI (en abscisses) et l'intelligibilité de différents types d'énoncés (en ordonnées). Courbe PB : listes de mots phonétiquement équilibrées. Courbe CVC : syllabes consonne-voyelle-consonne. Courbe Sentences : phrases entières. Tiré de (Hongisto, 2005) ; données de (Steeneken et Houtgast, 2002).

Dans le cas où la seule perturbation provient de la présence d'un bruit de fond, le STI et le SII relèvent de la même démarche et produisent des résultats similaires. La Figure 4 présente une comparaison des coefficients de pondération pour ces deux indicateurs.

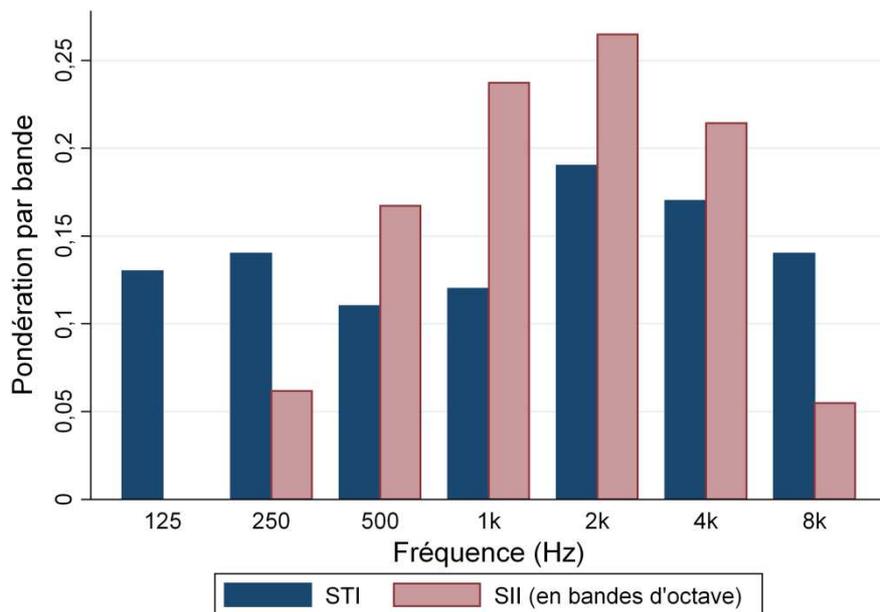


Figure 4 : Comparaison des coefficients de pondération par bande pour le STI et le SII.

2.2.4 Importance des fluctuations du bruit de fond

Le STI convient pour l'évaluation de l'effet d'un masque stationnaire ou de la réverbération sur l'intelligibilité d'un signal de parole. Cependant, lorsque le niveau du bruit de fond est variable dans le temps, cet indicateur devient moins performant. En effet, le STI est souvent directement calculé à partir de la connaissance du SNR. Or, le SNR est évalué à partir des spectres moyens. Pour autant, en pratique, le bruit de fond peut présenter des fluctuations de niveau, dont la valeur moyenne du niveau ne rendra pas compte. Dans ce cas, l'auditeur a la possibilité « d'écouter dans les blancs », c'est-à-dire de profiter des instants où le niveau du bruit de fond est à un creux pour comprendre le discours plus aisément : comme évoqué plus haut, il lui sera possible de reconstituer une part de l'information perdue en s'appuyant sur l'information bien reçue (Miller et Licklider, 1950). La variabilité temporelle du bruit masquant est un facteur important dans des situations où le bruit de fond est un signal de parole, naturellement fluctuant. Le phénomène peut se retrouver dans des bureaux ouverts, lorsque plusieurs voix sont audibles en même temps. Le terme fréquemment utilisé pour parler de tels environnements sonore est « *cocktail party* », en référence à ces situations sociales dans lesquelles chaque discussion est menée au milieu de nombreuses autres conduites à proximité (Bronkhorst, 2000).

Le STI ne permet donc pas de traiter de l'intelligibilité d'un discours masqué lorsque « l'écoute dans les blancs » devient importante. Comme évoqué précédemment, cela peut être le cas dans un bureau ouvert. Il convient donc, pour qualifier de telles situations de communication, d'améliorer la façon dont l'indice est calculé. Rhebergen *et al.* décrivent ainsi une procédure d'amélioration du SII (Rhebergen et Versfeld, 2005). L'indicateur proposé, le *Extended SII* (eSII) est calculé en effectuant une moyenne glissante du SII évalué sur des fenêtres temporelles suffisamment courtes pour permettre de rendre compte aussi fidèlement que possible de la dynamique de signaux étudiés.

Dans un premier temps, le signal et le bruit sont tous deux décomposés par un banc de 21 filtres. Ces filtres correspondent aux bandes critiques telles que précisées dans la norme relative au calcul du SII. Pour chaque bande de fréquence, l'intensité instantanée des deux échantillons est calculée. Le temps d'intégration employé dépend de la bande de fréquence étudiée, étant donné que la sensibilité du système auditif aux modulations en dépend également. Ainsi, pour la bande critique la plus basse, les segments de calcul durent 35 ms ; pour la bande la plus élevée, l'intégration est réalisée sur 9,4 ms. A partir des intensités instantanées du signal et du bruit, un SII « instantané » est calculé avec un pas de calcul de 9,4 ms. La valeur finale du eSII correspond à la moyenne des valeurs instantanées du SII. Ce eSII est alors mieux à même de traiter des cas où le bruit masquant est de niveau variable (Rhebergen *et al.*, 2006).

Etant donné les proximités entre le SII et le STI, une approche similaire peut être envisagée pour ce dernier indice. Premièrement, Rhebergen et Versfeld remarquent que de bons résultats peuvent être obtenus avec une largeur de fenêtre commune à toutes les bandes de fréquences, égale à 12.5 ms (Rhebergen et Versfeld, 2005). De plus, le banc de filtres utilisé devrait être en bandes d'octaves, et non en bandes critiques. En effet, comme évoqué dans la section 2.2.3, le calcul du STI s'effectue en bandes d'octaves.

Le eSII peut ainsi être adapté afin d'obtenir un indicateur proche du STI, mais permettant de rendre compte du caractère fluctuant de certains bruits de fond. Un tel indice, le STIt (« t » pour « temporel » ou « transitoire ») peut ainsi être calculé en suivant la procédure suivante :

- Décomposer le signal et le bruit à l'aide d'un banc de filtres à bande d'octave,
- Dans chaque bande, calculer les niveaux moyens respectifs des échantillons dans une fenêtre rectangulaire de 12.5 ms,
- Pour chaque bande, calculer le rapport signal-à-bruit dans la fenêtre temporelle considérée,
- A partir des valeurs de niveau et de SNR ainsi obtenues sur l'ensemble des bandes, calculer le STI instantané pour cette fenêtre,
- Répéter continuellement l'opération en avançant à chaque fois la fenêtre de calcul de 6.25 ms ; ceci correspond à un taux de recouvrement de 50%,
- Calculer la moyenne des valeurs instantanées du STI pour obtenir le STIt.

La procédure de calcul est présentée de façon schématique dans la Figure 5. La pertinence de cet indicateur a été validée à l'aide de tests d'intelligibilités employant, comme bruit masquant, divers brouhahas de parole enregistrés en bureaux ouverts. Le STIt rend donc mieux compte de l'impact des fluctuations du bruit de fond sur l'intelligibilité d'un signal de parole masqué.

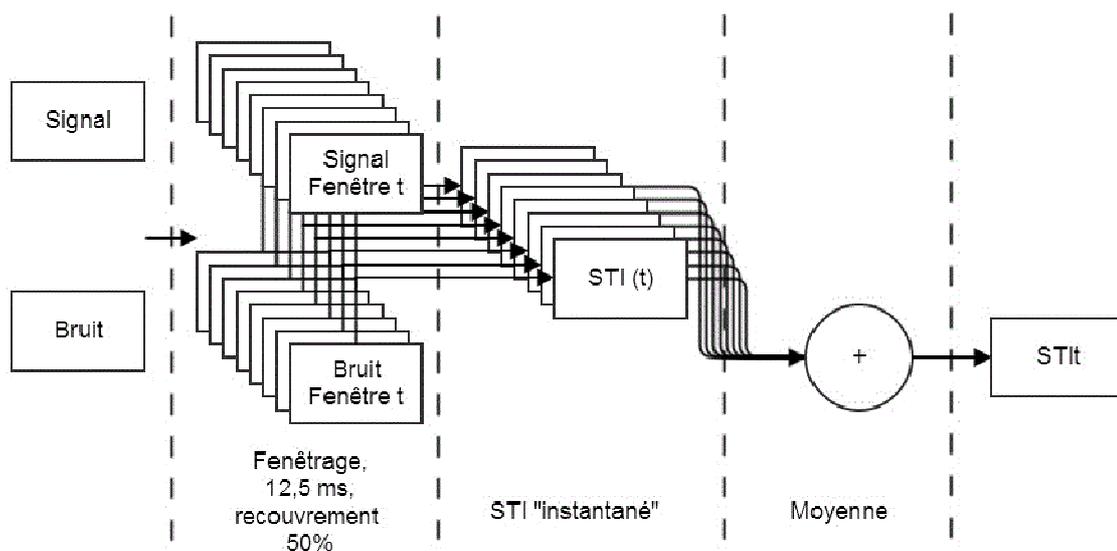


Figure 5 : Schéma de calcul du STIt.

2.2.5 Autres facteurs

Il existe d'autres caractéristiques du bruit ambiant intervenant sur l'intelligibilité d'un discours : celles-ci, cependant, ne peuvent pas être ramenées à des caractéristiques physiques du bruit masquant. Par exemple, les termes de masquage informationnel désignent les cas où le signal et le bruit peuvent être aisément confondus, de par leur contenu linguistique notamment. La confusion peut alors engendrer une incertitude sur le message reçu (Carhart *et al.*, 1969 ; Brungart, 2001). De même, les positions relatives des sources du signal et du bruit sont importantes. En effet, il est plus facile de séparer correctement deux signaux sonores lorsqu'ils émanent de sources spatialement nettement distinctes (Hirsh, 1950 ; Freyman *et al.*, 2001). De manière générale, lorsque l'analyse de scène auditive permet d'attacher à divers signaux incidents des caractéristiques (spatiaux, spectraux, rythmiques) distinctes, la ségrégation des flux d'information devient plus aisée (Bregman, 1990).

2.2.6 Conclusion

L'intelligibilité d'un discours dans le bruit dépend de plusieurs facteurs. Le masquage énergétique causé par le bruit de fond ainsi que des fluctuations de niveau tiennent un rôle particulier. Il s'agit en effet de caractéristiques physiques, facilement objectivables, du bruit ambiant. Pour évaluer l'intelligibilité, plusieurs indices existent. Le STI, en particulier, est basé sur le calcul des modulations du signal correctement transmises. Cet indicateur reflète correctement les effets du masquage énergétique. Il offre ainsi la possibilité de quantifier l'intelligibilité d'un signal de parole masqué, ce qui est essentiel pour étudier la relation entre intelligibilité et nuisance sonore. Cependant, dans sa forme classique, ne rend pas compte du phénomène « d'écoute dans les blancs ». Pour ce faire, il est nécessaire d'adapter son mode de calcul à la dynamique des signaux. Le STIt ici présenté rend ainsi mieux compte des éventuelles fluctuations de niveau du bruit de fond.

2.3 Impact de l'intelligibilité sur la nuisance sonore

2.3.1 Performance : cas particulier de la mémoire sérielle

Le bruit de parole a un effet particulièrement néfaste sur la bonne conduite d'une telle activité (Colle et Welsh, 1976 ; Banbury et Berry, 1998). Dans le cas d'un bureau ouvert, la configuration de l'espace, le nombre de locuteurs ainsi que la présence ou non de divers bruits masquants peuvent avoir une influence sur la façon dont le bruit de parole est perçu. Son intelligibilité, en particulier, peut être modifiée. L'intelligibilité des discours ambiants a déjà été présentée comme ayant une influence sur la sensation de gêne éprouvée par les salariés en bureaux ouverts. Cette gêne ressentie correspond-elle à un accroissement de la difficulté du travail quotidien ? Par suite, est-il possible d'observer des différences dans la bonne exécution d'une tâche de laboratoire du fait d'une modification de l'intelligibilité du bruit de parole environnant ?

Il faut ici remarquer qu'un bruit large bande stationnaire n'affecte pas la performance pour une tâche de mémoire sérielle (Ellermeier et Zimmer, 1997). De ce fait, un bruit de parole masqué par un bruit stationnaire, en plus de devenir moins intelligible, devrait également perdre de son potentiel de nuisance. Cette hypothèse a déjà été validée : par exemple, les sujets d'une expérience menée par Ellermeier et Hellbrück (1998) ont ainsi dû effectuer une tâche de mémoire sérielle dans cinq conditions sonores comportant un mélange de parole et de bruit rose à différents niveaux de masquage. Les rapports signal-à-bruit (SNR, pour *signal to noise ratio*) proposés étaient +4, 0, -4, -8 et -12 dB. Dans tous les cas, le niveau de présentation de la parole était fixé à 65 dB(A). Les résultats montrent une amélioration linéaire de la performance à mesure que le SNR diminue. En outre, l'effet ne se cantonne pas à la présence d'un bruit stationnaire masquant : des résultats similaires peuvent être obtenus en ajoutant un nombre croissant de voix concurrentes (Jones et Macken, 1995). Ainsi, en masquant un bruit de parole gênant, on en diminue le pouvoir perturbateur.

De la même façon, Schlittmeier *et al.* (Schlittmeier *et al.*, 2008b) ont directement altéré un signal de parole en le soumettant à un filtrage de type « passe-bas », dans le but de simuler la présence ou l'absence d'une paroi entre un locuteur et un opérateur. Pour confirmer l'effet de cette manipulation du signal sur son intelligibilité, les auteurs ont mené un test d'intelligibilité en parallèle de leur expérience principale. 30 phrases ont ainsi été diffusées, par haut-parleur, à un niveau sonore faible de 35 dB(A). En l'absence de filtrage, 98% des mots ont été compris par les auditeurs. L'ajout du filtrage a fait baisser ce score jusqu'à 75%. Les auteurs ont alors observé que les performances dans une tâche de mémoire sérielle étaient meilleures dans le cas où l'intelligibilité était la plus faible. Jones *et al.* (2000) montrent même qu'en augmentant l'intensité d'un tel filtrage (avec des filtres de pentes plus fortes après coupure), on améliorerait encore les performances des sujets.

Ces expériences mettent en évidence une corrélation entre une baisse de l'intelligibilité d'un discours environnant et une amélioration des performances dans une tâche type en laboratoire. Cependant, si la voix humaine semble être un perturbateur puissant pour les tâches de mémoire sérielle, il ne s'agit pourtant pas d'un phénomène lié au sens porté par le discours : le bruit de parole n'est pas moins perturbant lorsqu'il est renversé temporellement ou lorsqu'il est dans un langage qui n'est pas compris par l'exécutant (Jones *et al.*, 1990). Le fait que de la musique (Schlittmeier *et al.*, 2008a) ou même, simplement, une succession de courtes tonalités (Jones et Macken, 1993) suffise à engendrer un effet délétère, amène à penser que l'instationnarité du bruit est un déterminant essentiel de son pouvoir perturbateur.

2.3.2 Bruit et mémoire à court terme

Jones *et al.* (1993) proposent ainsi l'hypothèse d'instationnarité (*changing state hypothesis*, CSH) pour expliquer l'ISE dans le cas de tâches de mémoire sérielle : pour que les performances soient altérées, il est nécessaire que le son perturbateur soit, d'une part, segmentable, c'est-à-dire décomposable dans le temps en segments distincts. D'autre part, des segments successifs doivent être suffisamment différents les uns des autres. L'hypothèse est ensuite intégrée dans un modèle de la mémoire à court terme, nommé *object-oriented episodic record* (O-OER, (Macken *et al.*, 1999)) qui se propose de rendre compte de l'ISE. Suivant cette théorie, le stimulus sonore est analysé pour déterminer sa pertinence vis-à-vis, entre autres, de la tâche principale. L'analyse requiert de le traiter comme un flux d'information cohérent, un « objet auditif » dont il faut essayer de faire sens. Cet objet est organisé : les événements sonores qui le composent ne sont pas simultanés mais arrivent dans un certain ordre. Ces éléments sont donc pris en compte avec la nécessité de connaître leurs positions respectives dans la séquence. Aussi le traitement de l'information relève-t-il de la mémoire sérielle. Ceci se produit de manière automatique, éventuellement de façon inconsciente lorsque l'attention n'est pas portée sur le bruit environnant mais, par exemple, sur une activité prenante. Les auteurs avancent l'hypothèse que ce traitement obligatoire, faisant appel à la mémoire sérielle, interfère particulièrement avec les activités qui la sollicitent également. Cette interférence de processus a un impact négatif sur les performances. Ainsi, lorsque le signal sonore est, d'une part, aisément segmentable et, d'autre part, présente des segments successifs variés, le nombre d'éléments dont il faut tenir compte est important. La mémoire sérielle est donc particulièrement chargée et, par suite, est moins disponible pour d'autres tâches. De ce fait, à mesure qu'un bruit vérifie pleinement la CSH, son pouvoir handicapant augmente.

Il existe d'autres modélisations de la mémoire à court-terme comprenant une explication de l'ISE. Le *feature model* (Neath, 2000), par exemple, insiste plus sur l'interférence de contenu que sur

l'interférence de processus : la perturbation proviendrait de la similarité entre éléments à retenir et éléments du signal de bruit. La *temporal distinctiveness theory* (Glenberg et Swanson, 1986 ; LeCompte, 1996), en outre, souligne l'importance de la surcharge de la mémoire mais, pour autant, accorde moins de spécificité à la mémoire sérielle par rapport à d'autres processus de mémorisation. Ce qu'il est important de remarquer, c'est que tous ces modèles s'appuient sur la notion de segment sonore élémentaire dans le bruit de fond perturbateur. C'est ainsi que s'explique le lien entre l'intelligibilité d'un discours et son potentiel de nuisance : une des propriétés principales d'un bruit de fond perturbateur est son caractère segmentable ; or, la segmentabilité d'un signal de parole, telle qu'évaluée au travers de son contenu en modulation, est essentielle pour son intelligibilité : ce résultat a été présenté dans la section 2.2.2. Une conséquence importante de cette relation est qu'une mesure de l'intelligibilité d'un signal de parole basée sur son contenu en modulation devrait également pouvoir donner une indication sur son potentiel perturbateur. Le STI se présente alors comme un candidat prometteur. De nombreux travaux ont ainsi consisté à étudier la relation entre la valeur du STI et les performances observées (Venetjoki *et al.*, 2006 ; Haka *et al.*, 2009 ; Liebl *et al.*, 2011).

2.3.3 Le STI comme indicateur du potentiel de nuisance

Hongisto (2005) propose un modèle prédictif de la baisse de performance observée lors de la réalisation d'une tâche dans un environnement sonore à STI donné. Son postulat de départ est le suivant : dans une situation de travail en bureau, des mots ou syllabes seuls ne seront que rarement prononcés. C'est donc l'intelligibilité des phrases qui doit être liée à l'impact sur la performance de l'ambiance sonore. La présente section consiste en la description de ce modèle.

Afin de se munir d'une mesure de la perturbation lors de la réalisation d'une tâche, la notion de décrement de performance (DP) est introduite. Cette quantité, généralement exprimée en pourcentage, est définie comme la différence entre les taux d'erreurs observés, d'une part, dans la condition sonore étudiée et, d'autre part, dans une condition sonore de référence, généralement silencieuse :

$$DP = \frac{\text{erreurs dans le bruit}}{\text{quantité d'erreurs maximale}} - \frac{\text{erreurs dans le silence}}{\text{quantité d'erreurs maximale}}. \quad (2.18)$$

Dès lors, le modèle relie directement la performance à l'intelligibilité de phrases, telle que caractérisée par la valeur du STI :

$$DP = DP_{max} \left(1 - \frac{1}{1 + e^{\frac{STI-0.4}{0.06}}} \right). \quad (2.19)$$

Dans cette expression, DP_{max} représente la valeur maximale du décrement de performance, attendue lorsque l'intelligibilité est parfaite, c'est-à-dire pour un STI valant 1. Ce paramètre dépend de la tâche à réaliser. La représentation de cette fonction présente une forme sigmoïdale, clairement visible sur la Figure 6 : ceci permet de se rapprocher de la relation entre le STI et le score observé lors d'un test d'intelligibilité de phrases. Trois zones se dégagent clairement :

- la courbe présente un plateau initial lorsque le STI est inférieur à 0,3 : l'intelligibilité de phrases est mauvaise, la performance dans un tel environnement est peu altérée ;
- une zone de forte croissance du DP est observée pour un STI compris entre 0,3 et 0,5 : dans la même plage de valeurs, le nombre grandissant d'indices linguistiques intelligibles permettent une amélioration rapide de la compréhension de phrases ;
- la courbe atteint un plateau final à mesure que le STI dépasse 0,5 : la transmission du contenu de phrases est de très bonne qualité ; à l'inverse, un tel environnement de travail s'avèrera particulièrement handicapant.

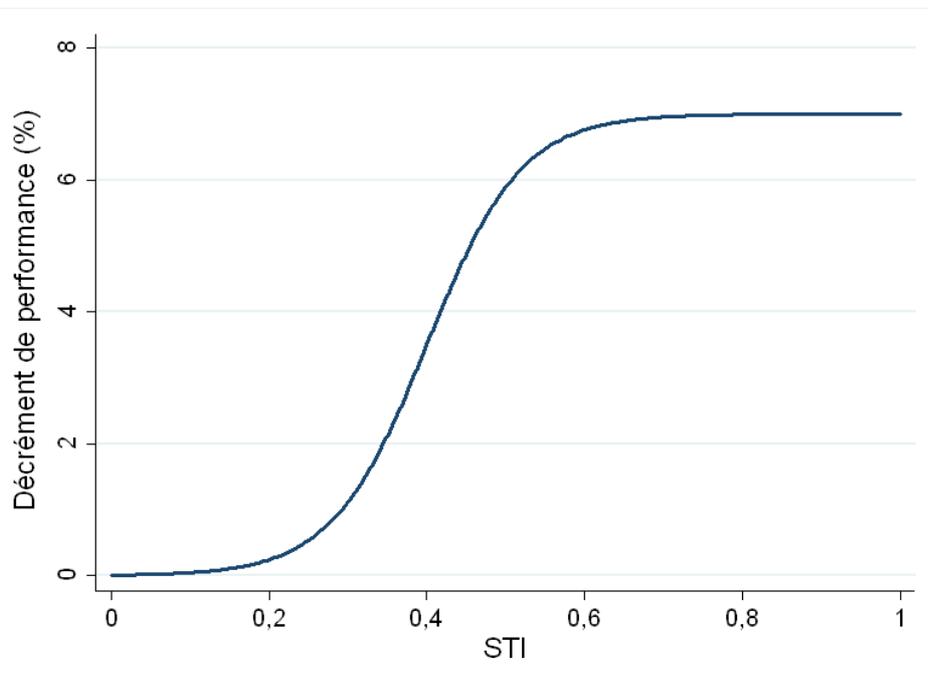


Figure 6 : Modèle de prédiction du décrement de performance en fonction de la valeur du STI dans le bureau ouvert. Ici, $DP_{max} = 7\%$.

Ce modèle a entraîné la mise en place de recommandations concernant l'aménagement des bureaux ouverts. Dans un article, Virjonen *et al.* (2009) mettent en évidence la possibilité offerte par le modèle de caractériser spatialement le pouvoir perturbateur d'une source de parole. Le rayon de distraction, r_D , est ainsi défini comme étant égal à la distance locuteur-auditeur pour laquelle le STI vaut 0,5. D'après le modèle, dans la zone comprise à l'intérieur d'un cercle de rayon r_D autour du locuteur, le travail est difficile car le bruit de parole est très perturbateur. Lors de la conception d'un

espace de travail ouvert, il convient dès lors de chercher à réduire la valeur de ce paramètre. Ainsi, pour un bureau ouvert d'excellente qualité (classe A), l'article cité ci-dessus spécifie qu'il est nécessaire de limiter r_D à 5 mètres.

2.3.4 Autres tâches : apport des mesures indirectes de charge cognitive

Lorsque l'activité ne nécessite pas de faire appel à un processus de mémoire sérielle de façon intensive, la sensibilité de la performance à l'intelligibilité du bruit de parole environnant est fondamentalement modifiée. Ainsi, le raisonnement grammatical (Schlittmeier *et al.*, 2008b ; Liebl *et al.*, 2011), le calcul mental (Banbury et Berry, 1998), la compréhension écrite (Venetjoki *et al.*, 2006) ou encore la correction de texte (Landström *et al.*, 2002 ; Haka *et al.*, 2009) ne sont pas affectées par l'intelligibilité. Face à un environnement de parole à intelligibilité variable, le comportement des sujets n'est pas affecté de la même façon pour ces quelques tâches que pour les exercices évoqués jusqu'ici. La seule analyse des performances des sujets tendrait à montrer qu'ils ne sont pas sensibles à l'intelligibilité du discours environnant.

Il ne faut cependant pas en conclure qu'une modification du paramètre d'intelligibilité est sans répercussion sur la gêne ressentie par les exécutants. Par exemple, dans les travaux menés par Haka *et al.* (2009), un ensemble de cinq activités sont proposées aux sujets. Trois conditions sonores sont présentées : constituées par addition d'un signal de parole et d'un bruit masquant, elles correspondent respectivement à des valeurs de STI égales à 0,10, 0,35 et 0,65. Un bloc expérimental correspond à la réalisation de l'ensemble des tâches dans une condition sonore donnée. A l'issue de chacun des trois blocs nécessaires au passage dans toutes les conditions, chaque exécutant doit rendre compte de son état de stress et de la perception qu'il a eu du caractère perturbateur de l'ambiance sonore par l'intermédiaire d'un questionnaire comportant onze questions. L'analyse des résultats montre que la sensibilité des performances des sujets face à l'intelligibilité variable n'était pas la même suivant la tâche étudiée. La tâche de mémoire sérielle, par exemple, était fortement touchée. Ce n'était pas du tout le cas de celle de correction de texte. Pour autant, les réponses aux questionnaires ont fait état d'une augmentation générale du mécontentement pour les valeurs plus élevées du STI. Schlittmeier *et al.* (2008b) avaient observé un phénomène similaire, tout en n'utilisant qu'une seule question sur le potentiel perçu de nuisance du bruit après chaque tâche au lieu d'un questionnaire plus complet. Ces résultats montrent que la gêne est aggravée par une amélioration de l'intelligibilité des discours environnants même lorsque la performance ne l'indique pas. L'étude de ces activités met en lumière l'intérêt de compléter systématiquement les mesures de performances par des mesures indirectes de la surcharge cognitive, telles que le NASA-TLX présenté plus tôt.

2.4 Approche suivie - Objectifs du présent travail

Ce chapitre a permis, dans un premier temps, de clarifier la notion complexe de nuisance sonore. Ce travail se concentre particulièrement sur la gêne dans les bureaux ouverts, en situation de travail. Pour cette raison, il semble pertinent de considérer le bruit non seulement comme plaisant ou déplaisant, mais comme perturbateur potentiel d'une activité en cours. La nuisance sonore se traduit alors comme un accroissement ponctuel de la difficulté d'une tâche, augmentation due à la présence d'une ambiance sonore particulière. Il en découle la nécessité de disposer d'outils pour mesurer cette surcharge cognitive. A cette fin, il convient d'effectuer conjointement des mesures objectives et subjectives.

Dans un second temps, le concept d'intelligibilité de la parole a été éclairé au travers de l'étude de plusieurs indicateurs objectifs servant à la quantifier. La construction du *Speech Transmission Index*, indice largement diffusé, permet de relier intelligibilité et modulations du bruit de parole. Ce dernier caractère est important. En effet, dans la troisième partie du chapitre, la nuisance sonore engendrée par un discours intelligible est reliée, entre autres, au caractère fluctuant de ce dernier. De ce fait, le STI apparaît comme un potentiel outil de prédiction de l'effet d'un bruit de parole sur la nuisance sonore subie en situation de travail. Cependant, pour mesurer cette gêne, la nature de l'activité perturbée est à prendre en compte.

Pour étudier l'impact de l'intelligibilité sur la nuisance sonore dans un contexte de bureaux ouverts, il convient, dans un premier temps, de mieux comprendre l'influence de différents déterminants de l'intelligibilité de la parole. Ceci permettra, par la suite, de juger de l'adéquation d'indicateurs, STI ou autres, à la problématique de prédiction du pouvoir de nuisance de la parole intelligible. C'est ce qui sera entrepris dans la suite de ce travail, à l'aide d'expériences sur sujets en conditions de laboratoire.

Une préoccupation constante tout au long de ce travail sera la grande attention prêtée au choix des ambiances sonores restituées. En particulier, toute manipulation de l'intelligibilité devra être effectuée d'une façon se rapprochant autant que possible de conditions réalistes. Ceci constitue un aspect novateur de ces travaux. En effet, dans la grande majorité des études portant sur la relation entre parole et gêne au travail, l'intelligibilité des signaux de parole n'est altérée que par l'ajout d'un bruit de fond stationnaire, de niveau plus ou moins élevé (Ellermeier et Hellbrück, 1998 ; Venetjoki *et al.*, 2006 ; Haka *et al.*, 2009 ; Liebl *et al.*, 2011). Il serait pourtant discutable d'affirmer que de tels bruits de fond puissent être couramment retrouvés en bureaux ouverts.

En effet, la propagation sonore à l'intérieur de tels espaces entraînent des modifications sur le spectre des signaux qui ne peuvent se résumer à des atténuations d'importance égale pour toutes les fréquences (Hongisto, 2005). Les effets respectifs de la transmission au travers des écrans, de la

diffraction autour de ceux-ci et de la réverbération sont tous dépendants de la fréquence considérée. Il conviendra d'en tenir compte : ce sera le cas en particulier dans la première expérience. De plus, il est possible que le masquage par un signal de type bruit blanc ou bruit rose ait un effet différent sur la gêne que celui d'un brouhaha de parole, plus plausible dans ces espaces de travail. Cette question sera, entre autres, abordée dans la deuxième expérience.

Ces deux expériences permettront ainsi de compléter les connaissances sur le lien entre STI du discours ambiant et gêne sonore au travail. Plus précisément, la première apportera des données expérimentales, pour des valeurs critiques du STI, à confronter au modèle de référence (Hongisto, 2005). La seconde permettra de mesurer l'impact d'une des faiblesses du STI en tant qu'indicateur d'intelligibilité, à savoir son incapacité à rendre compte des effets d'éventuelles fluctuations du bruit de fond, sur son utilisation en tant que mesure du potentiel de nuisance d'un signal de parole en bureaux ouverts.

De manière plus générale, l'utilisation de stimuli sonores représentatifs d'ambiances réellement rencontrées en bureaux ouverts apporterait un point de vue inédit sur la quantification de la gêne sonore dans ces lieux de travail. En effet, dans les quelques études utilisant du bruit enregistré en bureaux ouverts comme stimuli sonores, celui-ci n'est considéré que comme bruit masquant et est d'ailleurs délibérément sélectionné pour ne pas contenir de bruit de parole (Banbury et Berry, 1998 ; Haapakangas *et al.*, 2011). Pourtant, un bruit ambiant constitué par un ensemble de conversations plus ou moins lointaines ne peut sans doute pas être assimilé à une unique voix masquée par un ensemble de bruits divers, aucun n'étant constitué de parole. C'est suivant ce constat qu'ont été conçues les deux dernières expériences. Dans celles-ci, un nouvel indicateur sera utilisé pour quantifier le caractère fluctuant d'ambiances sonores recueillies dans plusieurs bureaux ouverts. La capacité de cette mesure à rendre compte d'éventuelles différences dans le pouvoir de nuisance d'ambiances sonores sera alors évaluée sur plusieurs tâches.

Dans chacune des expériences, des mesures objectives de performances seront réalisées. Pour les compléter, les exécutants évalueront eux-mêmes la charge cognitive supportée à l'aide du questionnaire NASA-TLX. L'intérêt de l'utilisation conjointe de mesures objectives et subjectives de la surcharge cognitive a déjà été mis en évidence dans la section 2.3.4. Parmi les études sur le lien entre parole intelligible et gêne sonore, il ne s'agit pourtant pas d'une pratique systématique (Liebl *et al.*, 2011 ; Jahncke *et al.*, 2013). De plus, les mesures subjectives sont souvent des compositions *ad hoc*, qu'il s'agisse d'une simple question (Schlittmeier *et al.*, 2008b) ou d'un questionnaire plus complet (Venetjoki *et al.*, 2006 ; Haka *et al.*, 2009). Il conviendrait pourtant d'encourager l'utilisation de mesures communes afin de faciliter les comparaisons entre expériences. Les expérimentateurs seraient ainsi assurés de mesurer correctement la même grandeur que leurs collègues. A cette fin, comme évoqué dans la section 2.1.3.3, le NASA-TLX présente l'avantage d'être d'ores et déjà

fréquemment utilisé dans le domaine de l'évaluation de la charge cognitive (Hart, 2006) et d'être reconnu comme étant un bon outil de mesure (Hill *et al.*, 1992). Pour ces raisons, l'utilisation du NASA-TLX dans cette étude constitue une avancée dans une direction qui mériterait d'être suivie plus généralement.

Enfin, les effets des facteurs étudiés ne pourront pas être considérés de façon indépendante. La variabilité interindividuelle, en particulier, devrait jouer un rôle non-négligeable. Ellermeier et Zimmer ont étudié ce phénomène sur 72 sujets (Ellermeier et Zimmer, 1997). Dans cette expérience, l'effet d'une parole parfaitement intelligible est comparé à celui d'un environnement silencieux. Le décrétement de performance moyen (tel que défini dans la section 2.3.3) est de 5%. Pour autant, autour de cette valeur, les comportements individuels sont très variés : les sujets les plus sensibles présentent un DP atteignant 17%. A l'opposé, pour environ un huitième du panel, l'*irrelevant sound effect* est nul, voire négatif : parmi les sujets avantagés par la présence de parole, le décrétement de performance minimal vaut -5%. Ces résultats montrent à quel point les différences entre individus peuvent être importantes dans l'analyse de l'ISE. Dans la suite de ce travail, l'importance relative des effets étudiés sera donc évaluée à chaque fois que cela sera possible.

3 Influence du STI du discours ambiant sur la gêne

3.1 Cas d'un bruit de fond stationnaire

Il est à noter que dans le modèle de référence proposé par Hongisto (2005), la forme de la relation entre le STI et le décrétement de performance, telle que présentée par la Figure 6, n'a pas été suggérée par des données expérimentales existantes. En effet, la forme sigmoïdale a été choisie pour reproduire celle de la courbe reliant le STI et l'intelligibilité de phrases. C'est ainsi qu'il se trouve, sur cette Figure 6, un plateau « bas » pour les faibles valeurs de STI. La baisse de performance attendue y est très peu prononcée. De la même façon, la courbe forme un plateau « haut » pour les valeurs de STI élevées. Les performances sont alors fortement altérées. Ces deux extrémités encadrent une zone intermédiaire caractérisée par une décroissance quasi-linéaire de la performance prédite suivant le STI. L'échelle de la courbe est paramétrable afin qu'il soit possible de l'adapter à tout tâche en fonction de sa sensibilité à la présence de parole.

Il n'existe que peu de travaux permettant de valider ce choix. La plupart des expériences sur l'impact du bruit de parole opposent simplement de la parole claire (dans ce cas, l'intelligibilité est parfaite) à du silence ou à un bruit quasiment stationnaire (l'intelligibilité est alors nulle). Ces deux conditions, correspondant respectivement à des valeurs de STI de 1 et de 0, ne permettent pas de juger de la pertinence du choix effectué. De plus, si Liebl *et al.* (2011) retrouvent bien une forte pente du décrétement de performance dans cette plage de valeurs de l'indicateur, entre 0,3 et 0,5, ce n'est pourtant pas le cas pour l'ensemble des études.

Dans les travaux menés par Haka *et al.* (2009), les signaux de paroles sont présentés suivant deux valeurs de STI : 0,35 et 0,65. La condition de contrôle est silencieuse. Le modèle de référence prédit que pour STI = 0,35, 30% du décrétement de performance maximal devraient déjà être visibles. Pourtant, il n'apparaît pas de différence significative entre cette condition sonore et la condition de contrôle. Ceci suggère que la forte baisse des performances suivant le STI (la « pente » de la sigmoïde) ne devrait débuter que nettement après cette valeur.

Pourtant, dans l'expérience menée par Jahncke *et al.* (2013), les conclusions sont opposées. Le minimum de performances est déjà atteint pour STI = 0,36 ; il n'y a pas de différence entre cette condition et la suivante, pour laquelle STI = 0,71. La « pente » de la sigmoïde se trouverait donc pour des valeurs de STI bien plus basses que prévues.

Au vu de l'incertitude présente dans la littérature, il semble intéressant de fournir des données expérimentales supplémentaires auxquelles confronter le modèle. Les auteurs des deux dernières études citées ci-dessus déplorent ainsi le manque de données pour des conditions situées entre STI = 0,3 et le plateau supérieur de la sigmoïde. Utiliser plusieurs valeurs de STI comprises entre

0,3 et 0,5 constituerait un apport important aux connaissances sur le lien entre STI et performance. Pour atteindre cet objectif, une première expérience a été réalisée.

3.1.1 Protocole expérimental

3.1.1.1 Déroulement de l'expérience

Pour les raisons citées en 2.2, une tâche de mémoire sérielle est choisie : les processus de mémoire sérielle sont, à plus ou moins grande échelle, au cœur d'un grand nombre d'actions du quotidien. De plus, les performances observées au cours de ce type de tâche présentent une forte sensibilité à la présence de parole. La tâche se déroule donc comme suit : une permutation de la série des entiers compris entre 1 et 9 est tout d'abord aléatoirement déterminée. La série de chiffres est présentée à l'exécutant, installé devant un écran d'ordinateur, à raison d'un chiffre par seconde. Chaque chiffre apparaît à l'écran pendant 0,7 s et est séparé du suivant par une pause de 0,3 s. Les chiffres sont affichés clairement, en noir sur un fond gris clair. La taille de chacun des chiffres à l'écran est de 3 cm de haut afin qu'ils soient aisément lisibles (fonte MS Sans Serif 85 points). L'interface graphique a été réalisée à l'aide du logiciel MATLAB. Après la disparition du dernier chiffre, il est demandé au sujet de retenir l'ordre dans lequel les chiffres lui ont été présentés pour une durée de 5 secondes. Dans ce laps de temps, rien n'est affiché à l'écran. Un pavé numérique, comportant un bouton pour chaque chiffre, apparaît finalement à l'écran. Dès lors, le sujet doit restituer la série en cliquant sur les boutons dans l'ordre qui, selon lui, correspond à celui de la série qu'il vient d'apercevoir. Le sujet ne rencontre pas de contrainte temporelle lors de cette phase de restitution : la série suivante n'est lancée qu'à partir du moment où le dernier chiffre est proposé, après une courte pause de 3 s. De plus, lorsque le sujet propose un chiffre, le bouton correspondant est retiré et il ne lui est pas possible de le sélectionner à nouveau. Le sujet doit donc proposer neuf chiffres, sans possibilité de corriger une erreur ou laisser une position incertaine vacante. Chaque chiffre mal placé est compté comme une erreur. Les séries s'enchaînent pendant une durée de 10 minutes, correspondant finalement à un bloc expérimental.

Les sujets travailleront dans différentes ambiances sonores. L'une d'entre elles est totalement silencieuse : il s'agit de la condition de contrôle, à laquelle les autres conditions seront comparées pour calculer le décrétement de performance. Les quatre autres correspondent à quatre valeurs différentes de STI. La première se situe dans le plateau bas de la courbe de la Figure 6, à $STI = 0,25$. Le modèle de Hongisto (2005) prévoit alors un décrétement de performance égal à 7,5 % du décrétement maximal. Les deux suivantes sont prises dans la zone intermédiaire de la même courbe, de part et d'autre du point d'inflexion situé à $STI = 0,4$. Ainsi, le décrétement attendu vaut 30 % de la valeur maximale pour la seconde valeur choisie, à $STI = 0,35$. Cette proportion atteint 70 % pour la

troisième valeur, à $STI = 0,45$. Il s'agit de la plus forte détérioration des performances du fait d'un incrément de 0.1 de la valeur du STI prévue par le modèle. Enfin, une valeur élevée de STI appartenant au plateau haut est choisie : 0,65. La perturbation devrait alors être presque maximale, à raison de 98,5 % de DP_{max} . Un bloc expérimental correspond à un ensemble de séries effectuées sur une durée de 10 minutes dans une condition sonore donnée. Chaque sujet effectue donc six blocs expérimentaux. Pour éviter de rendre l'expérience trop éprouvante pour les sujets, les blocs sont répartis sur deux journées consécutives. Au cours de chacune des journées, un bloc silencieux est présent. Les deux autres sont choisis parmi les blocs « bruyants », de façon à ce que toutes les conditions sonores soient rencontrées à l'issue des deux sessions. L'ordre de passage des conditions est autant que possible équilibré sur l'ensemble des sujets. Ceci signifie que pour exactement 60 sujets, la condition $STI = 0,45$, par exemple, sera rencontrée par 10 sujets lors du premier bloc de la première session, par 10 autres lors du second bloc de la première session et ainsi de suite. Si le nombre de sujets n'est pas un multiple de 6, un léger déséquilibre apparaîtra.

Pour calculer un décrement de performance, les résultats observés face à une condition bruyante doivent être comparés à ceux correspondant à une référence silencieuse. Il faut cependant s'attendre à ce qu'en s'habituant à la tâche, les sujets deviennent meilleurs et voient leurs performances s'améliorer de bloc en bloc. De ce fait, le nombre d'erreurs par série devrait être plus faible le second jour par rapport au premier pour toutes les conditions. Ainsi, par exemple, comparer un bloc bruyant rencontré au cours de la seconde session avec un bloc silencieux de la première risque d'engendrer une surestimation de l'effet délétère du bruit de parole pour cette valeur de STI. A l'inverse, si les résultats d'un bloc bruyant de la deuxième session étaient confrontés à ceux d'un bloc silencieux de la première, l'effet de la condition sonore serait masqué par l'apprentissage du sujet. Pour cette raison, les performances d'un sujet dans un bloc bruyant sont comparées à celles observées pour le même exécutant lors du bloc silencieux réalisé le même jour pour calculer un décrement de performance. Cependant, du fait de l'absence de contrainte sur le temps de réponse, les sujets n'effectuent pas nécessairement le même nombre de séries dans un même bloc. De ce fait, plutôt que la simple somme des erreurs au cours d'un bloc, c'est la moyenne du nombre d'erreurs par série dans un bloc qui sera utilisée pour effectuer la comparaison. Pour ce qui est du jugement subjectif de la charge cognitive, à l'issue de chaque bloc, le questionnaire NASA-TLX est proposé au sujet au travers de l'interface. Une fois le formulaire rempli, la procédure simplifiée est utilisée, c'est-à-dire que le RTLX est calculé par la moyenne des scores pour les six questions. Le texte des questions, en français, est présenté dans l'annexe A. C'est cette grandeur qui servira de variable de réponse.

Pour l'ensemble de l'expérience, les participants sont placés dans une cabine insonorisée, aux parois absorbantes, dans les locaux du laboratoire LVA de l'INSA de Lyon. Ils sont installés face à

l'écran sur lequel l'exercice est présenté. Cet écran, ainsi qu'un clavier et de la souris à l'aide de laquelle les sujets donnent leurs réponses, sont disposés sur un petit bureau. Les stimuli sonores sont diffusés par un haut-parleur situé face à l'exécutant, derrière l'écran. La réponse en fréquence de la salle est compensée par l'utilisation d'un banc de filtres en tiers d'octaves dans la chaîne de retransmission.

3.1.1.2 Stimuli sonores

Pour construire les environnements de parole nécessaires, il faut disposer d'échantillons de parole clairement intelligibles et d'un bruit masquant. Ces deux composantes sont constituées par des enregistrements disponibles auprès du Collège National d'Audioprothèse. Sur les pistes vocales, il est possible d'entendre un locuteur entraîné lire des listes de phrases cohérentes mais n'ayant pas de lien entre elles. Ces listes présentent la particularité d'être équilibrées phonétiquement. Ceci signifie qu'un ensemble de listes comporte l'ensemble des phonèmes de la langue française, aux fréquences auxquelles ils sont présents dans le langage parlé. Il s'agit plus précisément des phrases de Fournier, de Combescure ainsi que des listes du *Hearing in Noise Test* (HINT). Lors de tests d'audiométrie vocale, de tels enregistrements sont utilisés pour évaluer la capacité d'un patient à non seulement entendre, mais également comprendre un discours au milieu d'un brouhaha de voix concurrentes. La piste contenant ce bruit masquant, mélange de voix inintelligibles, est justement celle qui est utilisée lors de la présente expérience. Les spectres moyens des signaux de paroles et de bruit, pour un niveau global de 60 dB(A), sont présentés dans la Figure 7.

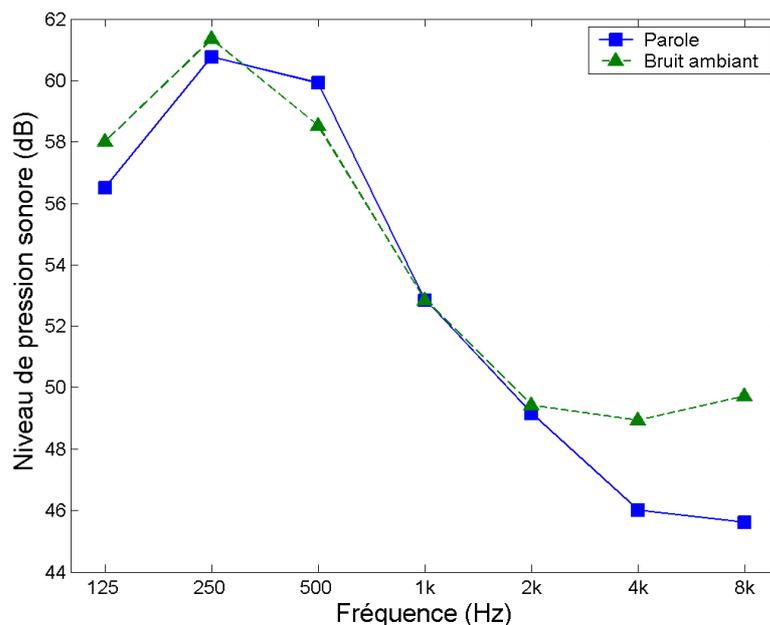


Figure 7 : Spectres moyens originaux des signaux de parole et de bruit. Pour les deux spectres, le niveau global est fixé à 60 dB(A).

Il existe plusieurs manières de parvenir à différentes valeurs de STI à partir de la connaissance du signal de parole cible et du bruit masquant. La plus simple consiste à ajuster les niveaux globaux des deux signaux afin d'atteindre l'intensité de masquage appropriée. Cette approche est celle généralement utilisée dans la littérature sur le sujet. Or, comme évoqué dans la section 2.4, un des objectifs de ce travail est d'utiliser des situations sonores correspondant à des bureaux ouverts réalistes. Pour cette raison, une autre méthode de variation du STI est choisie. Une situation de face-à-face est simulée à l'aide d'un modèle de prédiction de l'intelligibilité poste à poste en bureaux ouverts (Hongisto *et al.*, 2004). Le scénario considéré est le suivant : deux personnes, un locuteur et un auditeur sont installées à leur poste de part et d'autre d'une cloisonnette de séparation. Il s'agit alors de déterminer l'intelligibilité du discours du locuteur à partir de la connaissance du spectre du signal de parole, de celui du bruit ambiant et d'autres paramètres géométriques et acoustiques du local. La modélisation s'effectue en deux dimensions, dans le plan transverse à la cloisonnette, suivant la démarche présentée dans l'annexe B. Ceci permet d'obtenir les spectres des signaux de parole et de bruits transmis : dès lors, un calcul de STI est possible.

Dans le cadre de l'expérience présente, plusieurs hypothèses simplificatrices sont faites. D'une part, l'énergie transmise au travers de l'écran est négligeable par rapport celle réfléchi sur le plafond et à la diffraction au-dessus de l'écran. D'autre part, du fait de la courte distance entre le locuteur et l'auditeur, l'influence de la réverbération sur l'intelligibilité est négligeable par rapport à celle des réflexions précoces. La réverbération aura cependant un impact sur le bruit de fond. Ainsi, pour la détermination de l'énergie réfléchi par le plafond, il est nécessaire de connaître trois paramètres :

- la distance entre les deux individus, fixée à 2,5 m ;
- la hauteur de la tête des individus assis, fixée à 1,2 m ;
- la hauteur sous plafond h_{sp} ;
- l'absorption du plafond α_p .

Deux types de plafond sont utilisés : un absorbant et un réfléchissant : leurs caractéristiques en absorption sont détaillées dans le tableau 1. Pour connaître l'influence de la diffraction au-dessus de l'écran, il faut disposer des informations suivantes :

- la distance entre les deux individus (l'écran étant placé au milieu),
- la hauteur de tête des individus,
- la hauteur de l'écran h_e .

Enfin, le niveau du champ diffus réverbéré dépend des caractéristiques suivantes :

- la surface du bureau S et la hauteur sous plafond h_{sp}
- l'absorption du plafond,

- l'absorption des murs,
- l'absorption des sols,
- l'absorption équivalente du contenu du local.

Pour connaître l'absorption globale due au contenu du local, il faut spécifier le nombre de personnes présentes dans le bureau. Pour cela, un paramètre supplémentaire est introduit : la surface par opérateur, S_{op} . Il faut également connaître l'absorption représentée par l'opérateur et son équipement. Dans l'espace de travail, il est supposé que chaque personne dispose d'un bureau, d'une armoire de rangement et d'un « demi-écran », c'est-à-dire de la moitié de la cloison séparatrice qui lui fait face. A partir d'une estimation des caractéristiques acoustiques de l'ensemble de ces éléments, ainsi que de l'individu lui-même, un coefficient d'absorption est obtenu pour une personne et son équipement. La surface équivalente représentée par l'ensemble « opérateur » est estimée à 10 m². La surface de l'écran, estimée à part, est déterminée en fixant sa largeur à 1,8 m. Les valeurs de coefficient d'absorption par bande d'octave sont présentées dans le tableau 1.

L'absorption globale est également modulée par celle des murs. Le local est supposé parallélépipédique, deux fois plus long que large. Une des parois longues est vitrée ; les trois autres sont constituées d'un matériau parmi deux possibles. Les caractéristiques des parois réfléchissantes et absorbantes ainsi que de la paroi vitrée sont récapitulées dans le tableau 1. Il en est de même pour celles des sols, pour lesquels une seule possibilité a été retenue.

| Élément | | 125 Hz | 250 Hz | 500 Hz | 1 kHz | 2 kHz | 4 kHz | 8 kHz |
|------------------------|----------------|--------|--------|--------|-------|-------|-------|-------|
| Plafond | Absorbant | 0,35 | 0,4 | 0,6 | 0,95 | 0,95 | 0,85 | 0,68 |
| | Réfléchissant | 0,15 | 0,11 | 0,04 | 0,04 | 0,07 | 0,08 | 0,08 |
| Murs | Absorbants | 0,4 | 0,65 | 0,85 | 0,75 | 0,65 | 0,6 | 0,55 |
| | Réfléchissants | 0,29 | 0,1 | 0,06 | 0,05 | 0,04 | 0,04 | 0,04 |
| | Paroi vitrée | 0,18 | 0,06 | 0,04 | 0,03 | 0,02 | 0,02 | 0,02 |
| Sols | | 0,02 | 0,06 | 0,14 | 0,37 | 0,6 | 0,65 | 0,52 |
| Ensemble « Opérateur » | | 0,23 | 0,2 | 0,2 | 0,3 | 0,31 | 0,26 | 0,24 |
| Ecran | | 0,2 | 0,15 | 0,25 | 0,3 | 0,3 | 0,4 | 0,45 |

Tableau 1 : Récapitulatif des coefficients d'absorption par bande d'octave relatifs aux éléments constitutifs du bureau ouvert modélisé

Chaque combinaison de paramètres correspond alors à un ensemble de modifications spectrales sur les signaux et, par suite, à une intelligibilité de la parole donnée. Les calculs correspondant sont détaillés dans l'annexe B. Il en résulte donc à chaque fois une configuration caractérisée par une valeur de STI et de rapport signal à bruit. Une configuration doit enfin être retenue parmi plusieurs possibles pour chacune des quatre valeurs de STI étudiées. Comme évoqué plus haut, pour obtenir diverses valeurs de STI à partir d'un signal-cible et d'un bruit masquant, une approche répandue dans la littérature consiste à ne faire varier que le SNR. Ceci ne permet pas d'affirmer qu'une modification de STI correspond au passage d'un bureau ouvert réaliste à un autre. Ici, le choix des configurations est effectué de manière à minimiser l'écart entre les SNR obtenus. Le niveau de présentation final du mélange sonore obtenu est fixé à 46 dB(A). Ce niveau correspond, en bureau ouvert, à un environnement sonore considéré comme ni trop bruyant, ni trop calme (Bradley et Gover, 2004). Les valeurs des paramètres retenues sont présentées dans le tableau 2. La Figure 8 présente les spectres moyens des signaux au niveau de l'auditeur dans chacune des conditions. Le tableau 3 contient les valeurs de SNR correspondant aux conditions sonores étudiées.

| STI | 0,25 | 0,35 | 0,45 | 0,65 |
|----------------------------|------------|----------------|----------------|---------------|
| h_{sp} (m) | 3 | 3 | 2,6 | 2,7 |
| h_e (m) | 2 | 1,4 | 2 | 1,4 |
| S (m ²) | 400 | 400 | 700 | 2000 |
| s_{op} (m ²) | 7 | 12 | 12 | 9 |
| Plafond | Absorbant | Absorbant | Réfléchissant | Réfléchissant |
| Murs | Absorbants | Réfléchissants | Réfléchissants | Absorbants |

Tableau 2 : Configurations retenues pour la modélisation d'environnements sonores en bureaux ouverts.

| STI | 0,25 | 0,35 | 0,45 | 0,65 |
|-----------------|-------|-------|-------|------|
| SNR (dB) | - 4,8 | - 4,1 | - 3,9 | 3,1 |

Tableau 3 : Rapports signal-à-bruit pour les valeurs de STI retenues.

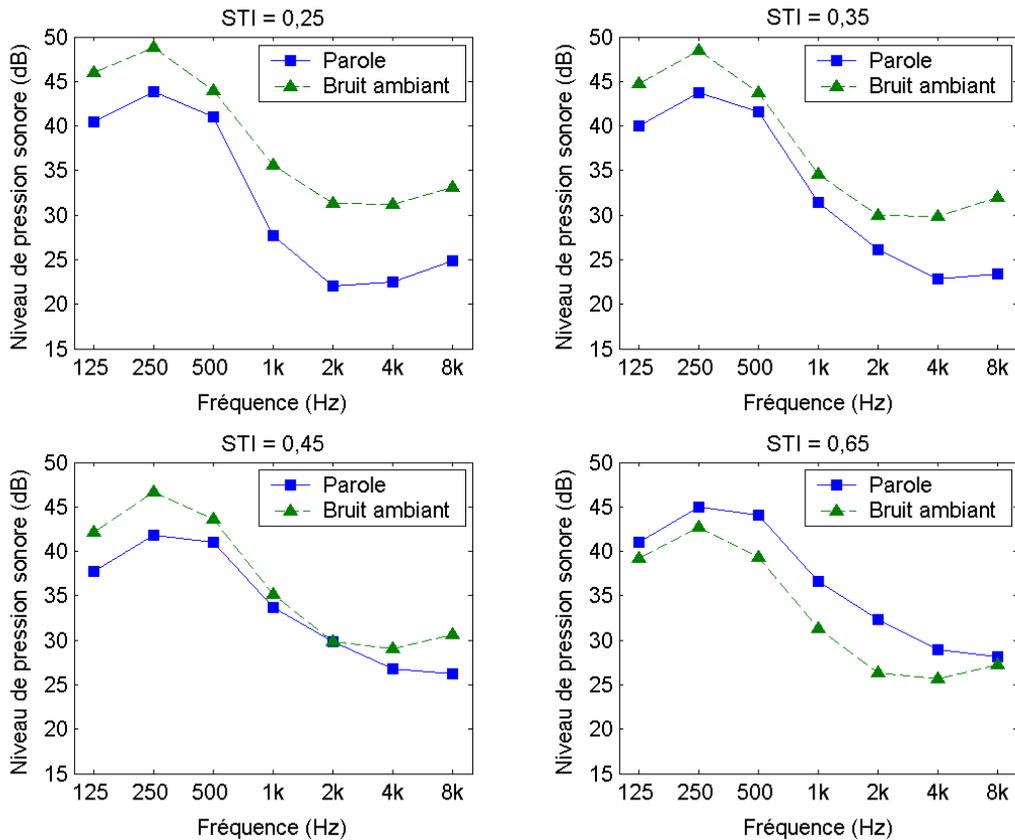


Figure 8 : Spectres moyens des signaux de parole et de bruit pour chacune des conditions bruyantes.

Un test d'intelligibilité a été mené pour chacune des conditions. Un groupe de 21 individus, différents de ceux recrutés pour la présente expérience, a été constitué. Pour chaque condition, un participant entendait 10 phrases. Après chaque énoncé, le sujet devait répéter la phrase aussi précisément que possible ; la proportion de mots correctement restitués constituait la variable réponse. Hongisto a proposé une expression sigmoïdale, l'équation (2.19), pour modéliser la relation entre le STI et l'intelligibilité de phrases. Comme la Figure 9 l'indique, les résultats des tests d'intelligibilité pour les présents échantillons prennent bel et bien la forme d'une sigmoïde. Cependant, celle-ci est légèrement décalée vers la droite par rapport à l'expression (2.19). La pente centrale de la sigmoïde est conforme à ce qui était attendu ; pour autant, les résultats suggèrent une valeur de STI égale à 0,36 pour une intelligibilité de 50 % : d'après l'équation (2.19), cette valeur centrale est de 0,4.

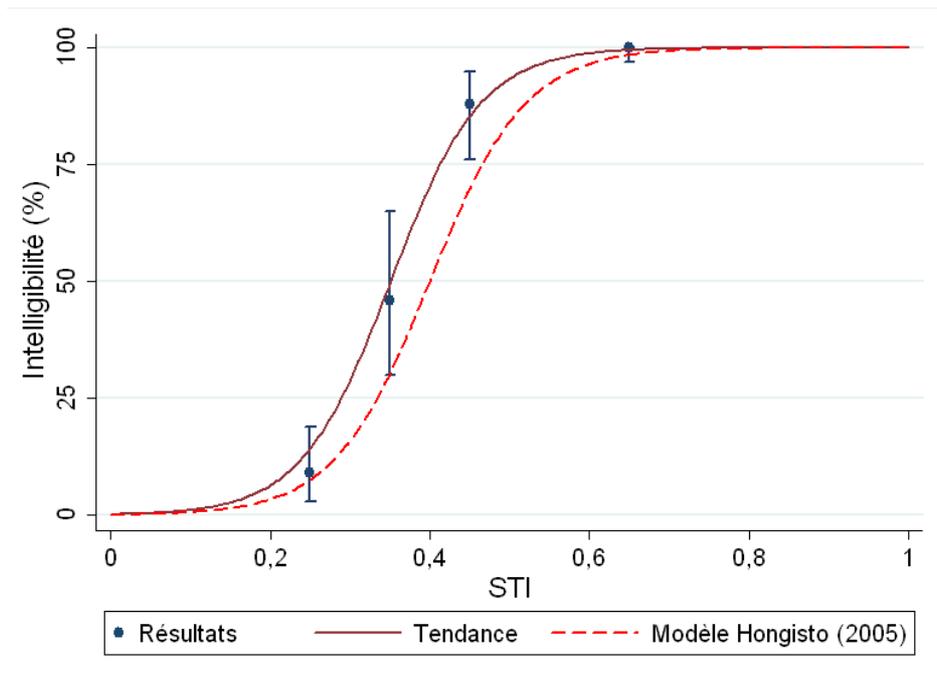


Figure 9 : Intelligibilité de phrases pour les quatre conditions sonores étudiées. Les valeurs médianes sont accompagnées de leurs écarts interquartiles respectifs. En continu, la courbe de tendance associée aux résultats. En pointillés, l'expression (2.19) proposée par Hongisto (2005).

3.1.1.3 Prédiction de la puissance statistique de l'expérience

Le modèle de Hongisto prévoit que les différences de performances observées soient importantes entre les valeurs de STI situées dans la zone intermédiaire, c'est-à-dire, par exemple, STI = 0,35 et STI = 0,45. Cependant, l'étude menée par Ellermeier et Zimmer (1997) rend compte d'une importante variabilité interindividuelle. Il convient alors de se demander combien de sujets il est nécessaire de recruter pour réduire le risque d'erreur de deuxième espèce, c'est-à-dire pour ne se laisser que peu de chances de ne pas observer de différences significative entre deux conditions alors même qu'un effet existe réellement. Il s'agit de calculer le nombre d'individus statistiques nécessaire pour atteindre *a priori* une puissance de test η prédéterminée. En général, une puissance égale à 0,8 est considérée comme suffisante. Un effet réel ne sera alors ignoré par erreur que pour 20% des échantillons d'individus possibles.

Pour effectuer ce calcul, il faut avoir une idée de l'amplitude de l'effet ainsi que de sa variabilité. La première nous est donnée par le modèle étudié. La seconde, elle, est estimée à l'aide des résultats de l'expérience d'Ellermeier et Zimmer (1997). Le calcul, détaillé en annexe C, montre que pour 20 séries effectuées par bloc, il faut recruter 62 sujets pour étudier la différence entre les conditions STI = 0,35 et STI = 0,45 avec une puissance de test suffisante. Il s'agit d'un nombre très élevé pour des études similaires : l'étude de Ellermeier et Zimmer, la plus complète sur le sujet de la variabilité interindividuelle pour une tâche de ce type, comporte 72 sujets (Ellermeier et Zimmer, 1997). Pour l'expérience présente, 57 sujets ont pu être recrutés. Ce contingent est constitué d'une

part de 27 étudiants de l'INSA de Lyon, tous âgés de moins de 30 ans, et, d'autre part, de 30 sujets, âgés de plus de 30 ans, recrutés par l'intermédiaire d'une société spécialisée dans la formation de panels pour des tests cliniques cosmétiques. Parmi les sujets se trouvaient 32 femmes et 25 hommes, âgés de 22 à 73 ans, avec une moyenne d'âge située à 36 ans. L'ensemble des sujets ont déclaré n'avoir aucun problème majeur d'audition. Tous ont été indemnisés pour leur participation. Pour cette taille d'échantillon, en supposant que chaque sujet réalise en moyenne 20 séries par bloc, la puissance de test attendue est de 0,77. Cette valeur n'est pas éloignée de la cible à 0,8 évoquée plus haut.

3.1.2 Résultats

Les analyses statistiques suivantes ont été menées avec le logiciel STATA v12.0. Dans cette expérience, les données semblent avoir une structure hiérarchique : chaque bloc expérimental est rencontré lors d'une session donnée par un sujet donné. Ainsi, les blocs réalisés par un même sujet risquent de présenter des similarités. La même remarque peut être émise au sujet des blocs réalisés le même jour par un sujet particulier. De ce constat provient le choix d'un modèle statistique multi-niveau. Une telle analyse est plus appropriée qu'une ANOVA classique lorsque les données présentent une structure hiérarchique naturelle. Le premier niveau est celui des unités expérimentales, les blocs. Le second est la session, comportant trois blocs. Le troisième est le sujet. Ceci permet de mieux tenir compte des corrélations attendues entre les données du fait de leur structure. Le seuil de signification pour les inférences statistiques est fixé à 5 %. Une correction de type Bonferroni est appliquée aux comparaisons par paires de conditions. Les décrets de performance moyens par bloc et les scores RTLX moyens, ainsi que les écarts-types associés, sont présentés dans le tableau 4.

| Condition | Décrément de performance (%) | | RTLX | |
|------------|------------------------------|------------|---------|------------|
| | Moyenne | Ecart-type | Moyenne | Ecart-type |
| Contrôle | - | - | 50,2 | 16,8 |
| STI = 0,25 | 1,39 | 9,23 | 59,9 | 15,9 |
| STI = 0,35 | 2,06 | 8,74 | 59,5 | 18,2 |
| STI = 0,45 | 3,47 | 9,31 | 61,3 | 19,3 |
| STI = 0,65 | 4,78 | 9,62 | 64,6 | 16,9 |

Tableau 4 : Récapitulatif des mesures de performance et de charge cognitive.

3.1.2.1 Performance

Dans la condition silencieuse, les sujets font, en moyenne, 2,25 erreurs par série. Il y a bel et bien un effet d'apprentissage : la moyenne du nombre d'erreurs dans le silence est de 2,60 le premier jour et de 1,90 le second. Ceci justifie le choix présenté dans la sous-section 3.1.1 : le décrétement de performance sera évalué en ne comparant que des blocs effectués au cours de la même session. La performance semble décroître avec l'amélioration de l'intelligibilité, comme le montre la Figure 10 : le nombre moyen d'erreurs par série atteint 2,71 lorsque le STI vaut 0,65. Il est à noter que ni l'âge ($\chi^2(1) = 0,11 : p = 0,74$), ni le sexe des sujets ($\chi^2(1) = 0,76 : p = 0,38$) n'ont un effet significatif sur le décrétement de performance.

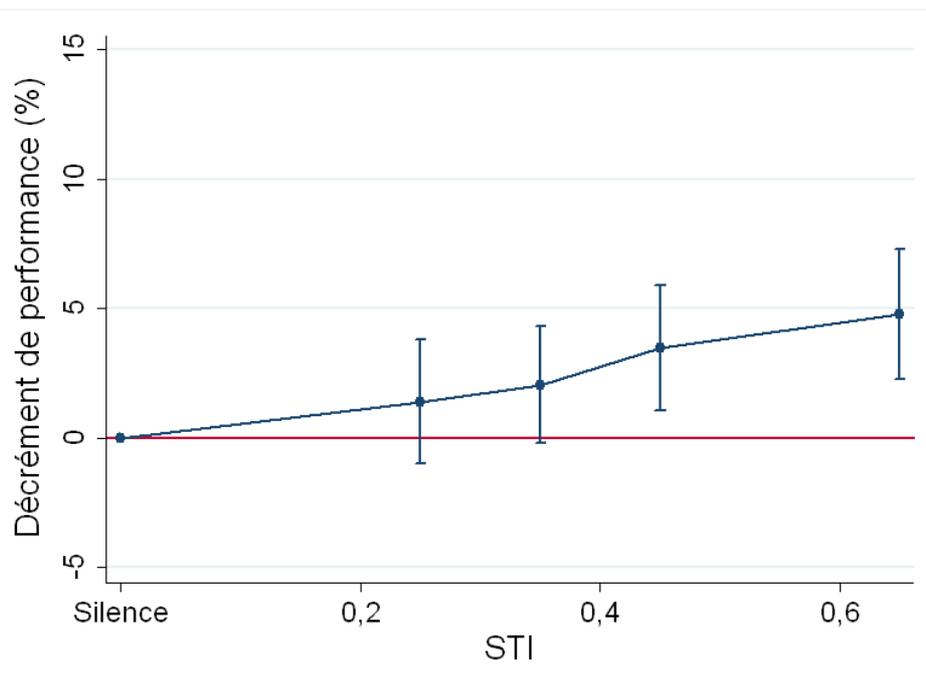


Figure 10: Décréments de performance pour les valeurs de STI étudiées. Les moyennes sont accompagnées de leur intervalle de confiance.

Il existe un effet global du facteur « STI » ($\chi^2(4) = 23,73 : p < 0,01$). En particulier, l'*irrelevant sound effect* est significatif : ceci signifie que les résultats moyens pour l'ensemble des conditions bruyantes sont significativement plus élevés que pour la condition de contrôle silencieuse ($\chi^2(1) = 15,78 : p < 0,01$). Le décrétement de performance moyen pour les conditions de test vaut 2,93 %. La différence de performance entre les conditions de test extrêmes (STI = 0,25 et STI = 0,65) semble élevée, ce qui paraît indiquer que le décrétement de performance croît avec la valeur du STI. Cependant, parmi les conditions de test, l'amplitude de l'effet du STI n'est pas suffisamment importante pour être significative au seuil de signification fixé ($\chi^2(3) = 6,95 : p = 0,07$). En d'autres termes, les variations du STI au dessus de la valeur nulle ne provoquent pas de changement significatif de la performance des sujets. Aucune des valeurs de DP pour les STI strictement positifs n'est distinguable de la valeur moyenne de 2,93 %.

L'étude des coefficients de partition de variance permet d'éclairer ce résultat. Le seul facteur STI n'explique que moins de 2 % de la variance totale des données. Ce chiffre doit être comparé à ce qui est observé pour les variables de niveau supérieur : session et sujet. D'après le modèle utilisé, l'effet de la session est négligeable. Pourtant, l'observation des données brutes montre que toutes conditions confondues, la différence entre condition bruyante et silence passe de 3,56 erreurs supplémentaires à 2,29 entre les deux jours. De plus, alors que la variable sujet seule n'explique que 17,6 % de la variance totale, la part expliquée conjointement par ces deux variables s'élève à 51,5 %. Ceci témoigne d'un fort effet d'interaction dans la modélisation. Une explication plausible provient du fait qu'il est possible qu'un plus grand nombre de sujets aient rencontré des conditions bruyantes moins acceptables le premier jour que le second plutôt que l'inverse. L'effet brut du jour correspondrait à un léger déséquilibre dans l'ordre de passage des conditions. Dans ce cas, il est possible d'attribuer l'ensemble de l'effet d'interaction à la seule variable « sujet », qui expliquerait donc 51,5 % de la variance totale. Cette valeur est importante et laisse penser qu'une part de l'effet de la condition peut être brouillée par la forte variabilité interindividuelle.

Pour explorer plus profondément les effets de la variabilité interindividuelle, une classification hiérarchique est effectuée. Les données utilisées pour calculer les distances entre sujets sont les quatre décrets de performance observés pour chacun d'entre eux, un par condition sonore bruyante. Cette analyse permet de diviser le panel de sujets en deux ensembles, dont les résultats sont visibles sur la Figure 11.

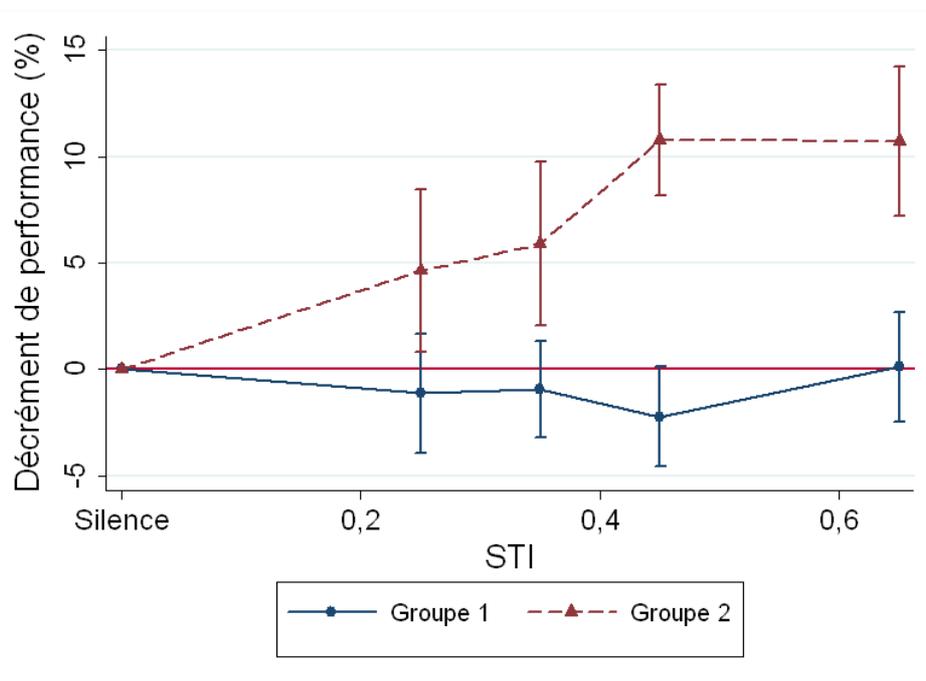


Figure 11 : Décréments de performance pour les valeurs de STI étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance.

Le premier groupe est constitué de 32 individus ayant obtenu de bons résultats dans l'ensemble : la moyenne d'erreurs par série y est de 1,8. Pour cette classe de sujets, la performance ne dépend clairement pas de la condition sonore : l'effet du STI sur le décrétement de performance n'est pas significatif ($\chi^2(4) = 4,26 : p = 0,37$). A l'inverse, les 25 sujets du second groupe présentent une sensibilité à l'intelligibilité de la parole ambiante assez élevée. L'effet du STI sur le décrétement de performance parmi les conditions bruyantes est significatif ($\chi^2(4) = 75,84 : p < 0,01$). La différence des taux d'erreurs par rapport au silence vaut 4,6 % pour STI = 0,25 : il s'agit d'une valeur comparable au maximum observé sur l'ensemble du panel. De l'autre côté de l'intervalle étudié, le décrétement de performance atteint 10,7 % lorsque le STI vaut 0,65. La part de la variance totale expliquée par la seule condition atteint 8,4 %. Cependant, même au sein de cette classe, les différences entre les décrétements observés dans les conditions intermédiaires ne sont pas significatives.

3.1.2.2 Charge cognitive

Les RTLX obtenus sont nettement plus élevés pour les conditions bruyantes que pour la condition de contrôle, comme la Figure 12 le montre. Il faut cependant se garder d'y voir le seul effet de l'intelligibilité du discours ambiant. En effet, s'il semble avéré que le niveau de bruit n'a pas d'impact significatif sur les performances lors d'une tâche de sériation (Tremblay et Jones, 1999), il n'est pas d'étude qui vienne corroborer ce résultat pour la charge cognitive. Comme pour le décrétement de performance, ni l'âge ($\chi^2(1) = 1,03 : p = 0,31$) ni le sexe des sujets ($\chi^2(1) = 1,67 : p = 0,20$) n'ont un effet significatif sur la valeur du RTLX.

Pour cette variable, l'effet de la condition est significatif ($\chi^2(4) = 131,71 : p < 0,01$). Cependant, comme précédemment, les différences entre les conditions intermédiaires ne sont pas statistiquement significatives ($\chi^2(1) = 1,92 : p = 0,38$). La condition extrême en STI = 0,65 est significativement différente des conditions intermédiaires prises dans leur ensemble ($\chi^2(1) = 8,48 : p < 0,01$). Il apparaît que la variabilité interindividuelle n'agit pas de la même manière sur la performance que sur le RTLX. En effet, si les sujets du groupe 1 défini plus haut témoignent d'une charge cognitive plus faible que ceux du groupe 2, les schémas de réponses, visibles sur la Figure 13, sont pour autant semblables : la sensibilité du RTLX à la valeur du STI ne dépend pas de l'appartenance d'un sujet à l'un des deux groupes.

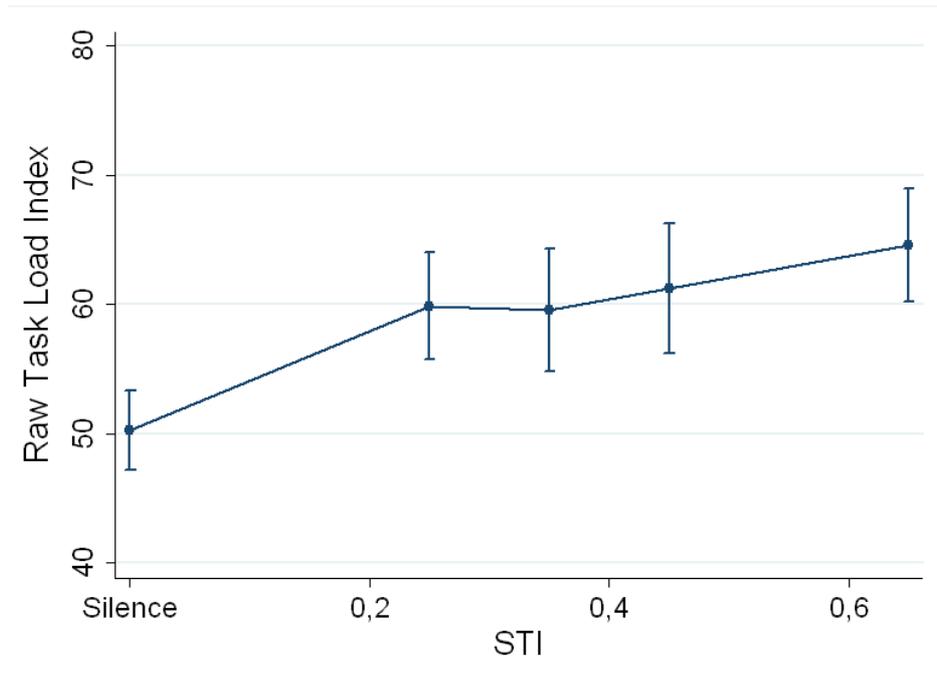


Figure 12 : RTLX obtenus pour les valeurs de STI étudiées. Les moyennes sont accompagnées de leur intervalle de confiance.

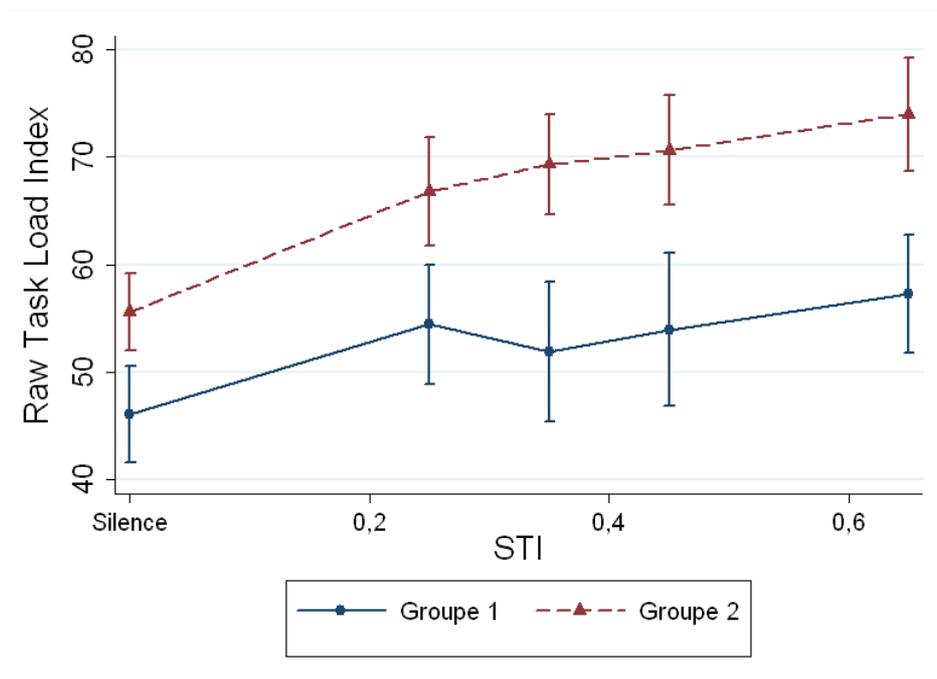


Figure 13 : RTLX obtenus pour les valeurs de STI étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance.

Pour l'ensemble des conditions, la part de variance du RTLX expliquée par le STI représente 9,1 % de la variance totale. Cependant, pour obtenir une valeur comparable à celle calculée pour le décrétement de performance, il convient de ne se référer qu'aux conditions de test, c'est-à-dire de ne pas considérer la part de variance correspondant simplement au passage de la condition silencieuse

aux conditions bruyantes. Sur ces quatre conditions, le STI explique, comme pour le décrétement de performance, un peu moins de 2% de la variance totale.

La méthode NASA-TLX a été conçue pour fournir un score global de charge cognitive. Cependant, l'analyse des réponses à chacune des questions peut constituer un outil de diagnostic utile afin d'identifier les déterminants de la difficulté et de l'effort. Il en ressort notamment qu'aucun effet significatif de l'intelligibilité n'est observé pour les trois premières questions, concernant la difficulté intrinsèque de la tâche. Il s'agit d'une surprise pour ce qui est du troisième item du questionnaire, relatif à l'exigence temporelle de la tâche. En effet, au cours de l'expérience, les sujets témoignaient souvent de leur étonnement face à la rapidité à laquelle les chiffres défilaient, ce qui laissait présager d'une sensation de pression temporelle accrue dans les conditions de travail les plus difficiles.

Néanmoins, une partie des membres du panel complétaient cette remarque en affirmant que la phase de restitution des chiffres était pour eux une occasion de reprendre le contrôle du rythme de la tâche. Ceci était possible, par exemple, en marquant une pause volontaire avant de cliquer sur le dernier chiffre disponible, se laissant ainsi quelques instants pour se préparer à la présentation de la série suivante. Cette possibilité d'ajustement tendrait à atténuer les différences de pression chronométrique ressentie entre les différentes conditions. Un effet significatif de la valeur du STI est visible pour les questions 4 et 5, relatives respectivement au jugement de sa propre performance et à l'effort investi dans la tâche. La conclusion à tirer est qu'une amélioration de l'intelligibilité du bruit de parole environnant contraignait les exécutants à fournir un effort supplémentaire ; malgré cela, l'impression d'une détérioration des performances était présente.

3.1.3 Discussion

Le modèle de prédiction de la performance en présence de bruit de parole proposé par Hongisto (2005) est basé sur une hypothèse forte. Selon cet auteur, le décrétement de performance varie avec le STI de la même façon que l'intelligibilité de phrases. Le modèle fournit dès lors deux prédictions principales. La première est la plus simple : pour les tâches sensibles au bruit de parole, les performances se détériorent lorsque l'intelligibilité de la parole ambiante, quantifiée par une valeur de STI, augmente. Les résultats de cette expérience vont dans le sens de cette affirmation, même si, pour le panel entier, il n'existe qu'une simple dichotomie entre la condition de contrôle et les autres. Les décrétements de performances obtenus avec ce protocole sont cependant faibles au regard de résultats présents dans la littérature. Par exemple, Liebl *et al.*, pour une tâche similaire, obtiennent un DP supérieur à 9% pour $STI = 0,61$ (Liebl *et al.*, 2011). Pour un STI valant 0,65, le décrétement de performance moyen n'atteint ici que 4,78%. En supposant, comme le suggère le

modèle, que cette valeur constitue 98,5% du décrement maximal, le taux d'erreurs dans un discours parfaitement intelligible devrait être supérieur de 4,85 points de pourcentage supérieur à celui obtenu dans le silence. En comparaison, Ellermeier et Zimmer mettent en évidence un décrement maximal de 6,3 % (Ellermeier et Zimmer, 1997); pour Schlittmeier *et al.*, cette valeur atteint 12 % (Schlittmeier *et al.*, 2008b), et même 13 % pour Zimmer *et al.* (2008). Il est difficile de trouver une explication à ces disparités, tant il existe de légères différences entre tous ces protocoles. Par exemple, les sons sont présentés par blocs, comme ici, pour Schlittmeier *et al.* ; à l'inverse, pour Liebl *et al.* et pour Ellermeier et Zimmer, la condition sonore change aléatoirement d'une série de chiffres à l'autre. Zimmer *et al.* utilisent les deux manières de procéder, sans différence significative sur la perturbation. De même, la durée de rétention des chiffres est de 10 secondes pour Schlittmeier *et al.* et Liebl *et al.* mais de 5 secondes pour Ellermeier et Zimmer, Zimmer *et al.* ainsi que dans cette étude. Ces disparités dans les protocoles sont récapitulées dans le tableau 5.

| Auteurs | Ordre des conditions | Temps de rétention | STI maximum | DP maximum |
|------------------------------------|-----------------------------|---------------------------|--------------------|--------------------|
| Ellermeier et Zimmer (1997) | Aléatoire, par série | 5 secondes | 1 | 6,3 % |
| Schlittmeier <i>et al.</i> (2008b) | 1 condition par bloc | 10 secondes | 1 | 13 % |
| Zimmer <i>et al.</i> (2008) | 2 expériences | 5 secondes | 1 | 13% (en aléatoire) |
| Liebl <i>et al.</i> (2011) | Aléatoire, par série | 10 secondes | 0,61 | 9 % |
| Expérience 1 | 1 condition par bloc | 10 secondes | 0,65 | 4,8 % |

Tableau 5 : Comparaison des résultats en performances pour plusieurs protocoles

La seconde forte prédiction du modèle de Hongisto, est la suivante : la relation entre la valeur du STI et le décrement de performance prend la forme d'une sigmoïde dont la pente est située entre les valeurs 0,3 et 0,5. Comme évoqué au début de la section 3.1, ce propos est remis en cause par les résultats de certaines études en laboratoire. En effet, pour Haka *et al.*, les performances des sujets décroissent fortement entre STI = 0,35 et STI = 0,65 mais pas entre STI = 0,10 et STI = 0,35

(Haka *et al.*, 2009). Au contraire, pour Jahncke *et al.*, sur l'ensemble des tâches proposées, le maximum du décrétement de performances est déjà atteint pour $STI = 0,34$ (Jahncke *et al.*, 2013).

Pour vérifier si les résultats de la présente expérience sont conformes à cette prédiction, il est possible de confronter le modèle proposé par Hongisto aux données expérimentales. Pour ce faire, l'expression (2.19) est paramétrée en utilisant la moyenne du décrétement de performance pour $STI = 0,65$. La Figure 14 présente une telle comparaison, en recalant le modèle sur les valeurs du second groupe. L'évolution des performances se rapproche plutôt de celle observée par Jahncke *et al.* : la forte détérioration des performances se produit pour des valeurs de STI plus basses que ne le suggère le modèle de référence. Cependant, à la fin de la section 3.1.2.2, il a été remarqué que la courbe de tendance correspondant aux résultats de tests d'intelligibilité pour les conditions étudiées différait légèrement de celle retenue par Hongisto. En comparant les résultats de performance avec cette courbe expérimentale, mise à l'échelle, une meilleure correspondance apparaît.

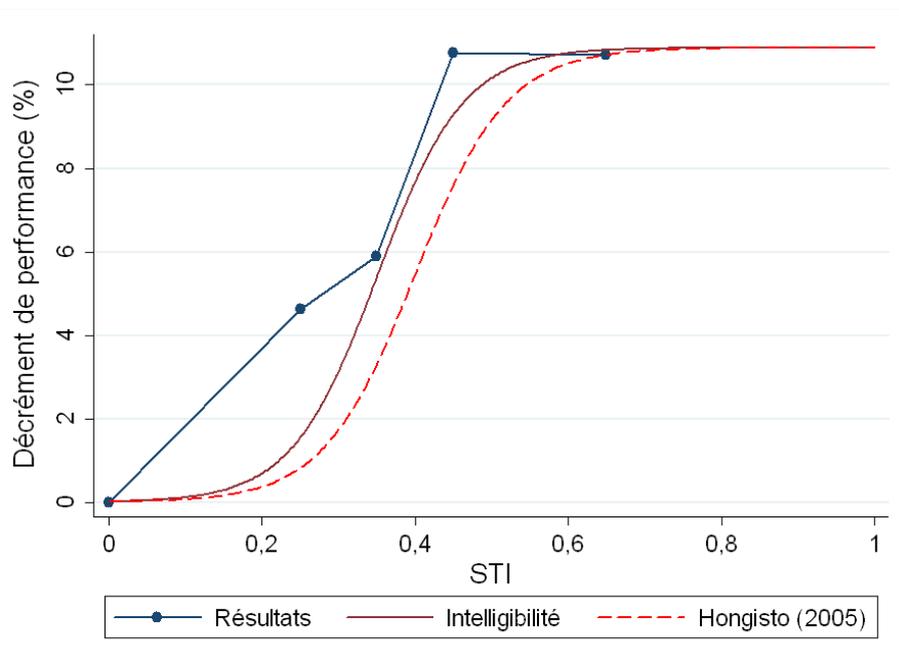


Figure 14 : Comparaison des performances obtenues par les sujets du second groupe avec le modèle de Hongisto (2005). En continu, courbe de tendance des résultats d'intelligibilité, mise à l'échelle (voir Figure 9).

Ainsi, tout du moins pour les sujets les plus sensibles, l'hypothèse de modélisation avancée par Hongisto semble confirmée : le décrétement de performance et l'intelligibilité de phrases évoluent de façon similaire avec le STI. Pour les personnes effectivement perturbées par la présence de bruit, le modèle de Hongisto (2005) rend compte correctement du pouvoir de nuisance d'un discours intelligible.

En ce qui concerne le jugement subjectif sur la perturbation causée par le bruit de parole, la valeur du RTLX augmente bien avec le STI. Ceci correspond aux résultats obtenus, avec d'autres

outils, par Schlittmeier *et al.* (2008b) tout comme Haka *et al.* (2009). Cependant, en termes de protocole expérimental, il s'est avéré que la question relative à l'exigence temporelle de la tâche n'avait pas été comprise et traitée de la même façon par l'ensemble des sujets. Il serait pertinent d'introduire une contrainte temporelle sur le temps de réponse afin de tenter de régulariser cette situation. En outre, il apparaît ici que le RTLX n'est pas sensible aux mêmes types de segmentations du panel de sujets que le décrétement de performance. Les variations de comportement entre sujets étaient ainsi plus marquées pour la mesure de performance. Ceci signifie que le RTLX est plus robuste face à la variabilité interindividuelle. Schlittmeier *et al.* tout comme Haka *et al.* avaient déjà montré que ce type de mesures est plus robuste que la seule observation des performances vis-à-vis d'un changement de tâche. Ces résultats complémentaires justifient clairement l'usage d'outils d'évaluation subjective de la charge cognitive.

3.2 Cas d'un bruit de fond fluctuant

Il est avéré qu'un signal de parole sera mieux compris lorsque le bruit masquant est une voix concurrente différente que lorsqu'il s'agit d'un bruit stationnaire de même spectre moyen. Ceci s'explique par un phénomène appelé « écoute dans les trous » : lorsque le bruit masquant présente des pauses suffisamment longues et fréquentes, comme c'est le cas d'un bruit de parole, l'auditeur est en mesure d'utiliser les indices contenus dans les portions non masquées du signal-cible pour en reconstituer le sens. Aussi la variabilité temporelle du bruit de fond a-t-elle une influence sur l'intelligibilité d'un signal-cible ; en outre, les valeurs d'un indicateur tel que le STI ne peuvent pas mettre en évidence ce phénomène (Bronkhorst, 2000). La première expérience a permis de confirmer l'utilité de la caractérisation d'une ambiance sonore par l'intelligibilité des discours qui la composent, et donc par une valeur de STI, pour évaluer son pouvoir perturbateur. Or, le STI ne permet pas de rendre compte de l'influence de la nature plus ou moins fluctuante d'un bruit masquant sur l'intelligibilité d'un signal-cible. Cette deuxième expérience doit permettre de savoir si cette lacune remet également en question l'utilisation du STI pour la prédiction du pouvoir perturbateur de ce type d'ambiance sonore.

Schlittmeier et Hellbrück (2009) montrent que pour un SNR global égal à 0, la présence d'un bruit large bande diminue le pouvoir de nuisance d'une ambiance sonore de bureau. Par contre, à SNR égal, masquer le bruit de bureau par de la musique ne conduit pas à une amélioration des performances ou de l'acceptabilité de l'environnement sonore. Une expérience du même type, menée par Haapakangas *et al.* (2011), compare l'efficacité de différents types de masques sonores pour la réduction du pouvoir de nuisance d'une voix perturbante. Les masques utilisés sont un bruit rose, un bruit de ventilation, de la musique instrumentale, de la musique vocale et un écoulement d'eau. A la lecture de leurs résultats, les sons musicaux, les plus fluctuants, sont plus perturbants que le bruit d'eau ou que le bruit rose. Pourtant, le masquage est tel que pour toutes ces conditions sonores, le STI de la voix parlée reste constant, à une valeur de 0,38. Ces deux travaux laissent penser que le STI peut effectivement être mis en défaut lorsque les bruits masquants sont fluctuants, ce qui peut être le cas en bureaux ouverts.

Pour autant, pour ces deux études, les bruits masquants sont envisagés en tant qu'éléments à ajouter à l'ambiance sonore originale. Leur but est de corriger les défauts d'environnements sonores riches en bruit de parole. Ces travaux ne permettent pas d'évaluer l'utilisation du STI pour les ambiances de bureau originales, sans ajout artificiel de bruit. Pour améliorer les connaissances sur le sujet, il faut donc étudier le cas où la voix perturbante est masquée par d'autres voix, celles qui constituent le reste de l'ambiance sonore.

Considérons le cas d'un locuteur privilégié dont la voix est masquée par un nombre croissant de voix concurrentes. Lorsque le nombre de voix contenues dans le masque augmente, le bruit résultant tend à se transformer en un brouhaha de plus en plus stationnaire. L'intelligibilité du signal de parole à comprendre s'en trouvera détériorée. Ceci devrait donc conduire à une diminution du potentiel de nuisance du bruit ambiant global qui ne serait pourtant pas prévue par le simple calcul du STI. L'utilisation de cet indicateur pour prédire le pouvoir de nuisance d'un signal de parole en bureaux ouverts s'en trouverait remise en cause.

Une expérience menée par Jones et Macken (1995) semble, à première vue, confirmer cette proposition. Les sujets y sont confrontés à une tâche de mémoire sérielle dans des ambiances sonores comportant une ou plusieurs voix (jusqu'à six locuteurs simultanément). Il en ressort que les performances des sujets, présentées dans la Figure 15, s'améliorent à mesure que le nombre de locuteurs augmente.

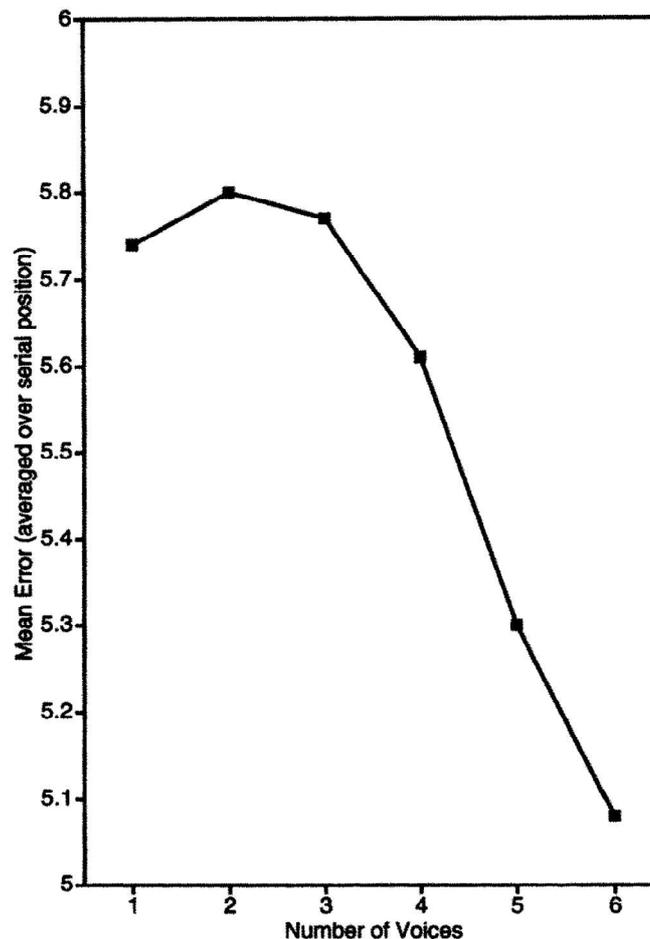


Figure 15 : Evolution des performances suivant l'ambiance sonore dans l'expérience de Jones et Macken (1995). En abscisse, le nombre de voix constituant le bruit ambiant. En ordonnée, le nombre moyen d'erreurs pour une série de 7 lettres à restituer.

Néanmoins, dans ce protocole, les voix supplémentaires sont ajoutées sans tenir compte du niveau global. De ce fait, le masquage s'effectue à niveau croissant : quelque soit la voix choisie comme référence, le rapport signal à bruit diminue lorsque le nombre de voix augmente. L'amélioration des performances pourrait alors être expliquée par l'action du SNR décroissant, ce dont le STI est en mesure de rendre compte. Par suite, pour évaluer la capacité du STI à tenir compte de l'influence de la variabilité du bruit de fond sur la gêne, il faut maintenir un niveau de masquage constant. C'est suivant ce principe qu'est conçue la seconde expérience menée au cours de ce travail.

Ainsi, du fait de la nature du bruit masquant (de la parole, non stationnaire) et de son niveau (masquage constant), les résultats de cette expérience doivent permettre d'avoir un jugement mieux fondé sur l'utilisation du STI pour la prédiction de la gêne sonore en bureaux ouverts.

3.2.1 Protocole expérimental

3.2.1.1 Déroulement de l'expérience

Le déroulement de l'expérience est analogue à celui de la précédente. Le panel de sujets est constitué de 31 individus, parmi lesquels se trouvent 8 hommes et 23 femmes âgés de 18 à 32 ans. Tous ont été indemnisés pour leur participation. La moyenne d'âge du groupe est de 23 ans. Le déséquilibre dans la composition de l'échantillon en termes de sexe est rendu acceptable par le fait qu'au cours de la première expérience, aucun effet significatif du sexe des sujets n'a pu être observé sur les variables de réponse. Le nombre total de participants n'a été déterminé que par la quantité de sujets qu'il a été possible de recruter et les disponibilités des équipements d'essai : la littérature ne fournissait pas de données permettant d'émettre *a priori* des hypothèses sur les différences de pouvoir perturbateur entre les conditions sonores testées.

L'activité de mémoire sérielle étudiée est presque inchangée. Toutefois, une contrainte temporelle a été ajoutée sur le temps de restitution. Ceci doit permettre de répondre aux interrogations soulevées autour du troisième item du questionnaire NASA-TLX, qui est également utilisé dans cette nouvelle expérience pour mesurer la charge cognitive à l'issue de chacun des blocs expérimentaux. Pour fixer cette limite, il convient de remarquer que dans l'expérience précédente, le temps de réponse était inférieur à 13 secondes pour 75 % des séries. Pour une restitution de cette durée, la présentation d'une série dure exactement 30 secondes, ce qui correspond à 20 séries effectuées au cours d'un bloc de 10 minutes. Cette valeur n'est pas éloignée du nombre moyen de séries effectuées par bloc lors de la première expérience, qui s'élevait à 21,2. C'est pourquoi la durée de la phase de restitution a été fixée à 13 secondes. Dans l'éventualité où un sujet n'aurait pas le temps de proposer une série entière, chaque position laissée vacante en fin de série compte pour

une erreur supplémentaire. Dans le cas contraire, si le sujet termine sa série avant le délai imparti, il devra attendre la fin des 13 secondes avant de voir la série suivante débiter.

L'expérience s'est déroulée dans un local insonorisé au centre de Lorraine de l'INRS. Le mode de présentation des stimuli visuels et sonores est similaire à celui de la première expérience.

3.2.1.2 Stimuli sonores

Les conditions sonores sont, elles, différentes. Comme précédemment, quatre environnements « bruyants » sont comparés à un environnement de contrôle. Cependant, si le signal de parole cible est toujours la voix d'un locuteur masculin, le bruit masquant est cette fois constitué dans un premier temps d'un nombre croissant de voix féminines. Ainsi, la première ambiance sonore ne comporte qu'une voix féminine concurrente. La seconde en comporte trois, et la troisième cinq. Au-delà de cinq locuteurs, l'impact d'un mélange de voix sur l'intelligibilité d'un discours-cible commence à se rapprocher de celui d'un bruit stationnaire (Bronkhorst, 2000). De ce fait, ajouter des voix supplémentaires ne serait pas utile. Pour cette raison, pour la quatrième condition, le bruit masquant est simplement un bruit stationnaire. Par la suite, ces conditions sonores seront notées comme suit :

- 1 voix masculine, 1 voix féminine : M + 1 F,
- 1 voix masculine, 3 voix féminines : M + 3 F,
- 1 voix masculine, 5 voix féminines : M + 5 F,
- 1 voix masculine, bruit stationnaire : M + BS.

Le masquage est effectué à niveau constant, c'est-à-dire que pour toutes les conditions sonores, le spectre et le niveau moyens du bruit masquant restent inchangés. Il en est de même pour la voix masculine cible. Il en résulte d'ailleurs que les spectres moyens des quatre mélanges sont également identiques. Les spectres retenus sont ceux des signaux de test spécifiés dans la norme IEC relative au calcul du STI (IEC, 2003). Le niveau de masquage retenu est tel qu'en prenant la voix masculine comme référence, la valeur du STI calculé sur la base des spectres moyens est de 0,4. Il s'agit ainsi de se placer dans la zone de valeurs de STI dans laquelle la performance est censée être la plus sensible à une variation de l'intelligibilité d'un discours perturbateur. Le niveau de présentation, comme précédemment, est fixé à 46 dB(A).

La condition de référence est également modifiée. L'analyse des réponses au questionnaire lors de l'expérience précédente montrait une grande différence entre les scores enregistrés pour la condition de contrôle et ceux issus des conditions bruyantes. Il était pourtant impossible d'affirmer que l'effet était uniquement dû à l'augmentation du STI et non au simple fait de l'augmentation drastique du niveau sonore ambiant : la condition de contrôle était silencieuse. Pour éviter cet écueil,

la condition de contrôle est, cette fois, constituée d'un bruit stationnaire présenté au même niveau que les conditions bruyantes. En outre, le spectre moyen de ce bruit correspond à l'addition des spectres du signal de parole et du bruit de fond des conditions bruyantes. Ainsi, les spectres et niveaux globaux de l'ambiance sonore dans sa globalité sont les mêmes dans chaque condition sonore. L'insensibilité des performances au niveau sonore pour des niveaux courants a déjà été mentionnée : par exemple, les sujets de l'expérience menée par Ellermeier et Zimmer (1997) ne font pas significativement plus d'erreurs dans un bruit rose (de niveau égal à 76 dB(A)) que dans le silence. Ce choix de condition de référence est donc adapté à la détermination du décrétement de performance. Comme pour les autres conditions sonores, le niveau de présentation de la condition de contrôle est 46 dB(A).

3.2.2 Résultats

Du fait de la similarité entre les deux protocoles, les modèles et analyses statistiques utilisés sont de même type. Les moyennes d'erreurs par bloc et les scores RTLX moyens, ainsi que les écarts-types associés, sont présentés dans le tableau 6.

| Condition | Décrétement de performance (%) | | RTLX | |
|-----------|--------------------------------|------------|---------|------------|
| | Moyenne | Ecart-type | Moyenne | Ecart-type |
| Contrôle | - | - | 46,4 | 12,3 |
| M + 1 F | 10,11 | 12,01 | 61,7 | 10,8 |
| M + 3 F | 4,59 | 10,25 | 59,2 | 12,6 |
| M + 5 F | 5,72 | 9,13 | 58,6 | 12,4 |
| M + BS | 4,35 | 8,97 | 52,3 | 13,7 |

Tableau 6 : Récapitulatif des mesures de performance et de charge cognitive.

3.2.2.1 Performance

Dans la condition de contrôle, les participants font, en moyenne, 2,10 erreurs par série. Cette valeur est légèrement inférieure à celle observée au cours de la première expérience, soit 2,25 erreurs par série. Cependant, la sensibilité des sujets au bruit est plus élevée : pour la condition M + BS, avec un STI égal à 0,4, le décrétement de performance observé vaut 4,35 %. Ce résultat se rapproche de celui obtenu précédemment pour un STI bien supérieur, valant 0,65. Il est envisageable

que cette sensibilité accrue soit liée au fait que cette fois, les sujets ont du donner leurs réponses pour chaque série dans un temps limité.

L'effet d'apprentissage est ici encore visible : de 2,42 erreurs par séries le premier jour, la moyenne d'erreurs tombe à 1,78 le second. Ainsi, pour chaque sujet, chaque bloc ne sera comparé qu'avec le bloc référence du même jour. La condition M + 1 F semble être particulièrement contraignante, comme le montre la Figure 16 : dans un tel environnement, 2,85 erreurs sont commises en moyenne.

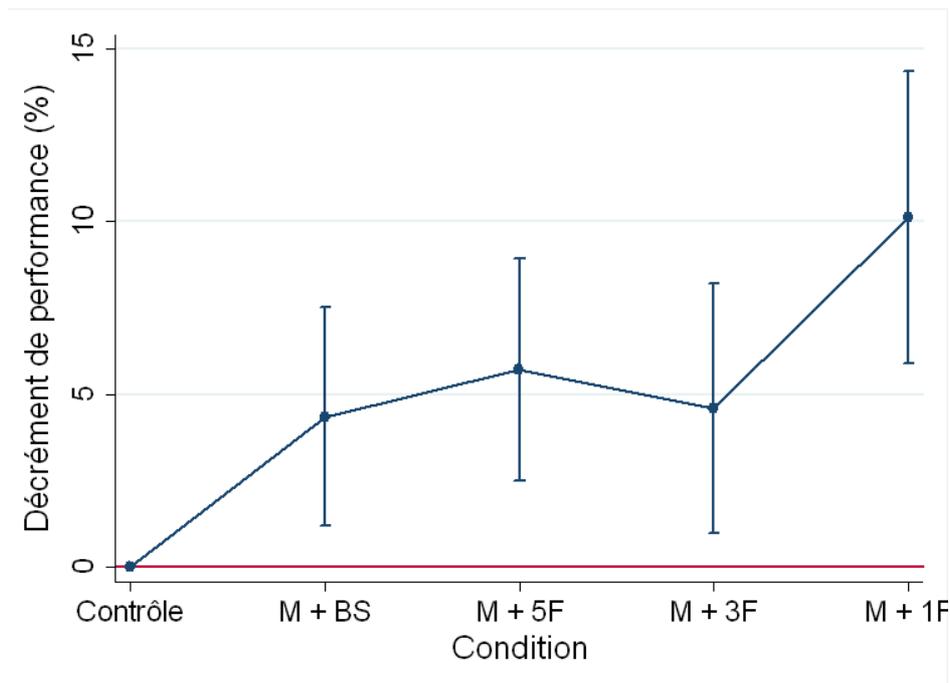


Figure 16 : Décréments de performance pour les bruits de fond étudiés. Les moyennes sont accompagnées de leur intervalle de confiance.

La modélisation du décretement de performance permet d'affirmer que par comparaison à un bruit de fond stationnaire, la présence de parole est défavorable : *irrelevant speech effect* est significatif ($\chi^2(1) = 33,14 : p < 0,01$). De plus, les conditions sonores bruyantes diffèrent dans leur pouvoir de nuisance ($\chi^2(3) = 11,29 : p = 0,01$). En particulier, la condition M + 1 F est significativement plus gênante que les trois autres conditions prises dans leur ensemble ($\chi^2(1) = 10,65 : p < 0,01$). Ces trois conditions, M + 3 F, M + 5 F et M + BS, ne sont pas différenciables l'une de l'autre ($\chi^2(2) = 0,64 : p = 0,73$). D'après l'observation des données brutes, les sujets présentent la même sensibilité à la parole au cours des deux sessions : pour le premier jour comme pour le second, la différence moyenne entre la condition de contrôle et les conditions bruyantes est de 4,13 erreurs. Ceci se retrouve dans les résultats du modèle statistique, suivant lequel l'effet du facteur « session » ne s'exprime qu'en interaction avec le facteur « sujet ». Il apparaît donc encore plus légitime qu'auparavant d'attribuer l'ensemble de cet effet à la seule variabilité interindividuelle,

qui explique ainsi 58,2 % de la variance observée sur les décrets de performance. La condition sonore, pour sa part, n'explique que 4,4 % de la variance totale. Pour explorer cette variabilité interindividuelle si importante, il convient de chercher à la décomposer au travers de l'existence de groupements de sujets aux réponses semblables. De ce fait, une classification hiérarchique ascendante a été menée. Les résultats mettent en évidence un groupe de 9 sujets se caractérisant par un très fort handicap observé pour la condition M + 1 F, le décrement de performance atteignant alors 20,7 %. Cependant, au vu de la faible taille de ce sous-échantillon, il semble peu prudent de tirer des conclusions de cette observation.

3.2.2.2 Charge cognitive

La valeur moyenne du RTLX dans la condition de contrôle est de 46,4. La Figure 17 présente l'évolution du RTLX suivant la condition sonore. Les conditions pour lesquelles une ou des voix féminines sont présentes dans le masque semblent être plus exigeantes que les autres.

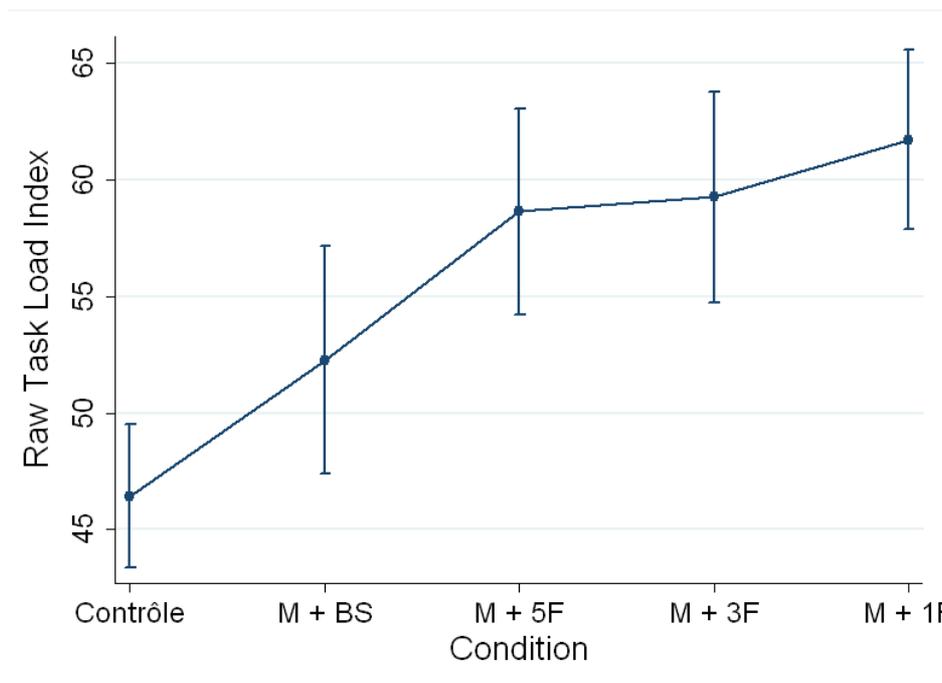


Figure 17 : RTLX obtenus pour les bruits de fond étudiés. Les moyennes sont accompagnées de leur intervalle de confiance.

La modélisation de l'effet de la condition sur le RTLX confirme cette impression en mettant en évidence une différence significative entre, d'une part, les conditions M + 1 F, M + 3 F et M + 5 F et, d'autre part, l'ensemble formé par les conditions M + BS et la condition de contrôle ($\chi^2(1) = 61,74 : p < 0,01$). Cependant, il est impossible de distinguer les différentes conditions à l'intérieur de ce premier groupement ($\chi^2(2) = 2,36 : p = 0,31$). De manière générale, la condition sonore a bien un effet significatif sur les réponses au questionnaire ($\chi^2(4) = 90,19 : p < 0,01$). En outre, le

simple fait d'ajouter un signal de parole suffit à générer un surplus de charge cognitive, comme en témoigne l'écart non négligeable entre la condition de contrôle et la condition M + BS ($\chi^2(1) = 8,18 : p < 0,01$). Pour les seules conditions de test, la part de la variance globale expliquée par le seul facteur « condition » s'élève à 7 %.

Le questionnaire NASA-TLX offre la possibilité d'obtenir un diagnostic sur la nature de la charge cognitive additionnelle engendrée par la présence du bruit de parole. Pour cela, l'analyse des scores obtenus question par question doit être menée. Il apparaît que pour toutes les questions, l'effet de la condition sonore sur le score observé est significatif. Cependant, pour les deux premiers items, traitant des demandes mentales et physiques imposées par l'exercice, cet effet se limite au seul impact de la présence de bruit de parole par opposition au bruit stationnaire. En effet, une fois retirée la condition de contrôle, seules les questions 3 à 6 (exigence temporelle, performance, effort, frustration) présentent un effet significatif de la condition sonore. La dichotomie observée sur le RTLX entre les conditions comprenant une ou des voix féminines et celles contenant du bruit stationnaire se retrouve sur chacune de ces quatre composantes.

Ainsi la présence de parole dans le bruit ambiant oblige-t-elle les sujets à agir plus vite qu'ils ne le feraient dans une situation confortable. De plus, les participants ont considéré que les environnements riches en parole les poussaient à faire plus d'efforts pour atteindre un niveau de performance adéquat, ce dernier étant alors tout de même abaissé. Enfin, la frustration ressentie était plus grande lorsque le bruit ambiant se rapprochait d'un mélange de parole. Cependant, de la même façon que pour le RTLX global, ces trois conditions, M + 1 F, M + 3 F et M + 5 F, ne sont pas différenciables entre elles au vu des scores pour ces questions.

3.2.3 Discussion

Dans cette expérience, il existe une différence significative entre les performances observées avec deux voix concurrentes (condition M + 1F) et quatre voix concurrentes (condition M + 3F). Au-delà de ce seuil, l'ajout de voix dans le masque est sans effet. Jones et Macken (1995) trouvent une tendance similaire mais des valeurs seuils différentes. En effet, d'après leurs résultats, il n'existe pas de différence de potentiel de nuisance entre une voix seule et trois voix concurrentes. Par contre, dès lors, l'ajout de voix supplémentaires une à une (jusqu'à 6 voix au total) réduit l'altération des performances. Cette différence est sûrement liée au niveau des voix supplémentaires. Comme évoqué au début de la section 3.2, pour Jones et Macken, l'ajout de voix tend à transformer petit à petit l'ensemble du mélange de parole en un brouhaha dont le niveau de fluctuation tend vers zéro. Ici, au contraire, l'ensemble des voix féminines garde toujours le même niveau face à la voix masculine. La situation limite n'est plus un brouhaha stationnaire, mais une voix claire masquée par

un brouhaha stationnaire. Ceci tend sans doute à compresser la plage de variation des performances suivant le nombre de voix présentes.

Suivant la mesure de performance, le masque ne contenant qu'une unique voix féminine se détache des autres masques ; suivant l'observation du RTLX, les masques contenant des voix se détachent d'un masque stationnaire. Cette hiérarchie des bruits de fond devrait correspondre à celle obtenue en s'attachant à la manière dont ils perturbent l'intelligibilité du signal de parole cible. Pour en juger, les valeurs de STIt relatives à la voix masculine ont été calculées, suivant la procédure présentée en 2.2.4, pour chacune des conditions sonore. Les valeurs récapitulées dans le tableau 7 montrent que dans l'ensemble, les conditions pour lesquelles l'intelligibilité est la meilleure sont celles pour lesquelles la nuisance sonore est la plus forte : à mesure que l'intelligibilité du discours avoisinant augmente, la nuisance sonore s'accroît. Cependant, le STI n'aurait pas permis de tirer une telle conclusion : il aura fallu utiliser un autre outil, dérivé du STI mais permettant de tenir compte des fluctuations du bruit masquant.

| Condition | M + 1 F | M + 3 F | M + 5 F | M + BS |
|------------------|----------------|----------------|----------------|---------------|
| STIt | 0,75 | 0,60 | 0,54 | 0,39 |
| DP (%) | 10,11 | 4,59 | 5,72 | 4,35 |
| RTLX | 61,7 | 59,2 | 58,6 | 52,3 |

Tableau 7 : Valeurs de STIt pour les quatre conditions de test. Pour rappel, les valeurs moyennes de décrémentation de performance et de RTLX sont présentées.

3.3 Conclusion

La première expérience, pour laquelle un bruit de parole est masqué par un bruit de fond stationnaire, confirme l'hypothèse suivante : l'ajout de bruit de parole intelligible dans une ambiance sonore augmente la gêne au cours de la réalisation d'une tâche faisant appel à la mémoire sérielle. Cette conclusion est atteinte en analysant les résultats de la mesure objective de la performance tout comme en étudiant les jugements subjectifs de la charge de travail. La difficulté de la tâche s'accroît à mesure que l'intelligibilité du discours environnant s'améliore. De plus, à l'intérieur du panel, les sujets sensibles à la présence de parole ont pu être identifiés. Pour ceux-ci, l'évolution de la performance en fonction du STI s'accordait avec celle de l'intelligibilité des phrases, indépendamment mesurée. Ces résultats justifient l'utilisation du modèle de prédiction du décrement de performance proposé par Hongisto (2005) pour les sujets perturbés par le bruit de parole intelligible, dans le cas particulier où le bruit masquant est stationnaire.

La seconde expérience, avec des voix fluctuantes servant cette fois de bruit masquant, permet d'affirmer que l'éventuelle instationnarité du bruit de fond doit être prise en compte pour déterminer le pouvoir de nuisance d'un discours intelligible. En tant qu'indicateur d'intelligibilité, le STI peine à correctement prendre en compte l'éventuel caractère fluctuant du bruit de fond. Au vu des résultats de cette expérience, le même constat peut être fait au sujet de sa capacité à prédire le potentiel de nuisance d'une ambiance sonore. En conclusion, le STI ne permet pas de quantifier le pouvoir perturbateur d'un bruit de parole masqué par un bruit lui-même instationnaire. Comme le montre l'évolution du STI pour les différentes conditions de masquage étudiées, le pouvoir de nuisance d'un signal de parole évolue bel et bien comme son intelligibilité : cependant, pour en rendre compte, il faut tenir compte de l'impact des fluctuations du bruit masquant.

Pour les deux expériences, la place tenue par la variabilité interindividuelle dans l'explication des performances et de la charge cognitive observées est plus importante que celle prise par les changements de condition sonore. Cependant, la mesure subjective de la perturbation par le RTLX s'est avérée être moins bruitée par la variabilité interindividuelle que le décrement de performance. L'apport d'un tel outil pour décrire l'accroissement de difficulté d'une tâche de mémoire sérielle est ainsi mis en évidence. Ceci confirme l'intérêt de compléter l'étude classique de ce type de tâche, habituellement menée au travers de la seule performance, par un jugement subjectif de la surcharge cognitive. A cette fin, la procédure NASA-TLX semble convenir.

4 Influence des fluctuations de niveau pour des ambiances réelles

Le chapitre précédent permet de conclure qu'un indicateur d'intelligibilité tel que le STI ne rend pas convenablement compte de l'effet d'un bruit de parole sur la nuisance sonore en bureau ouvert. Ceci est principalement lié au fait que le calcul du STI suppose un bruit de fond stationnaire (Houtgast et Steeneken, 1985), ce qui n'est pas le cas dans une situation où un nombre suffisamment restreint de locuteurs sont présents. En effet, la valeur du STI témoigne du fait qu'un bruit de fond stationnaire diminue le contenu en modulation du signal de parole. La voix du locuteur est alors moins segmentable, ce qui, d'après les théories de la mémoire de travail présentées dans la section 2.3.2, la rend moins perturbante. Cependant, un bruit de fond fluctuant apporte son propre contenu en modulation au bruit ambiant global, ce qui, au contraire, tendrait à rendre l'ambiance sonore moins propice au travail serein. Un indicateur d'intelligibilité adapté à la caractérisation du potentiel de nuisance d'un bruit de parole doit pouvoir tenir compte de ces variations de niveau du bruit de fond.

De manière plus générale, il serait avantageux de ne pas avoir à considérer séparément les contenus en modulations respectifs du signal de parole et du bruit de fond. Il ne serait ainsi plus nécessaire de réaliser l'opération suivante, aisée dans des conditions de laboratoire mais plus difficile en présence de véritables enregistrements d'ambiances sonores de travail : distinguer les contributions respectives d'un signal de parole principal et d'un bruit de fond masquant. Il faudrait alors disposer d'indices « non intrusifs », en ce sens qu'ils ne nécessiteraient pas une connaissance préalable de la composition interne du bruit ambiant à qualifier. A ce jour, la littérature ne comporte que peu d'exemples d'indicateurs d'intelligibilité non intrusifs. Le travail mené par Falk *et al.* est centré sur les effets de la réverbération et des algorithmes de déréverbération dans les systèmes de communication (Falk *et al.*, 2010). L'indicateur ainsi obtenu ne rend pas compte de l'effet pourtant réel de la présence d'un bruit de fond stationnaire sur l'intelligibilité. En outre, les performances de l'indice proposé par Chen *et al.* n'ont été spécifiquement étudiées que pour des auditeurs malentendants équipés d'implants cochléaires (Chen *et al.*, 2013). Aucun de ces deux protocoles ne semble se prêter aisément à la prédiction du pouvoir de nuisance d'un signal de parole en bureau ouvert dans le cas général.

En s'écartant de la notion d'indicateurs d'intelligibilité, la force de fluctuation, évoquée dans la section 2.1.1, paraît adaptée. C'est ainsi que Schlittmeier *et al.* ont proposé un algorithme permettant de prédire le décrétement de performance pour une tâche de mémoire sérielle causé par

un bruit ambiant de force de fluctuation donnée (Schlittmeier *et al.*, 2012). Pour valider ce modèle, les auteurs ont utilisé, entre autres, quelques ambiances sonores de bureau. Cependant, les différences de force de fluctuation entre les quatre échantillons concernés sont minimes. De plus, il n'existe pas de consensus sur le calcul de cet indice dans le cas de spectres de modulations non triviaux.

Ici, une autre approche sera retenue. Il s'agira de calculer le STIt que l'on obtiendrait si un bruit de parole était masqué par l'ambiance sonore étudiée. Dès lors, à niveau et spectre égaux, un bruit ambiant très fluctuant serait moins masquant qu'un bruit plus stationnaire : les valeurs de STIt ainsi obtenues témoigneraient de cette différence. Il est alors envisageable que le STIt puisse donner une indication sur le potentiel de nuisance d'une ambiance de travail. Cette hypothèse sera mise à l'épreuve dans la suite de ce chapitre au cours de deux expériences, de même type que les précédentes ; la différence étant que les ambiances sonores de test seront constituées à partir d'enregistrements sonores effectués dans de véritables bureaux ouverts.

L'utilisation d'un indicateur non-intrusif pour la quantification du pouvoir de nuisance d'une ambiance sonore constitue une approche encore très peu suivie. De même, les études expérimentales sur la nuisance sonore en bureaux ouverts n'utilisent pas de vraies ambiances de travail riches en bruit de parole (Banbury et Berry, 1998 ; Haapakangas *et al.*, 2011). Les résultats de ces deux expériences devraient donc mettre en évidence de nouvelles connaissances sur le sujet.

4.1 Quantification des variations de niveau avec le STIt

Dans l'expérience précédente, suivant les deux mesures, le bruit stationnaire est un perturbateur moins efficace que les autres ambiances sonores. A l'inverse, l'environnement sonore *a priori* le plus fluctuant, celui ne comportant que deux voix (M + 1F), semble particulièrement perturbant ; les valeurs du RTLX en témoignent. De manière générale, les résultats semblent indiquer que plus l'ambiance sonore est fluctuante, plus la nuisance sonore est forte. Un tel propos s'accorderait avec les résultats et théories évoquées en 2.3.1 et 2.3.2 (en particulier, la *changing state hypothesis*). Cependant, pour en juger, il faut être en mesure de quantifier les variations de niveau de l'ambiance sonore.

Pour ce faire, l'approche retenue ici consiste à déterminer, à l'aide du STIt, l'intelligibilité potentielle d'un signal de parole masqué par l'ambiance sonore étudiée. La procédure de calcul est donc la suivante :

- Décomposer le bruit à l'aide d'un banc de filtres à bande d'octave,
- Dans chaque bande, calculer le niveau moyen du bruit dans une fenêtre rectangulaire de 12.5 ms,
- Pour chaque bande, soustraire ce niveau au niveau moyen de l'échantillon entier dans cette bande ; ceci revient à effectuer un rapport signal-à-bruit en considérant un signal hypothétique de même spectre,
- A partir des valeurs de niveau et de SNR ainsi obtenues sur l'ensemble des bandes, calculer le STI instantané pour cette fenêtre,
- Répéter continuellement l'opération en avançant à chaque fois la fenêtre de calcul de 6.25 ms ; ceci correspond à un taux de recouvrement de 50%,
- Calculer la moyenne des valeurs instantanées du *STI* pour obtenir le STIt.

Le STIt ainsi calculé pour un bruit stationnaire vaut 0,5. A mesure que les variations de niveau du bruit augmentent, la valeur du STIt tendra vers 1. Pour illustrer ce propos, la Figure 18 montre l'évolution du STIt ainsi calculé pour un bruit de fond modulé en intensité à 4 Hz, en augmentant progressivement l'indice de modulation. Le STIt potentiel part bien de 0,5 pour un indice de modulation nul, c'est-à-dire pour un bruit stationnaire. A mesure que la modulation augmente, le signal devient plus fluctuant, comme le montre la Figure 19. Les valeurs du STIt croissent alors avec celles de l'indice de modulation. Le STIt vaut finalement 0,71 pour un indice de modulation de 1.

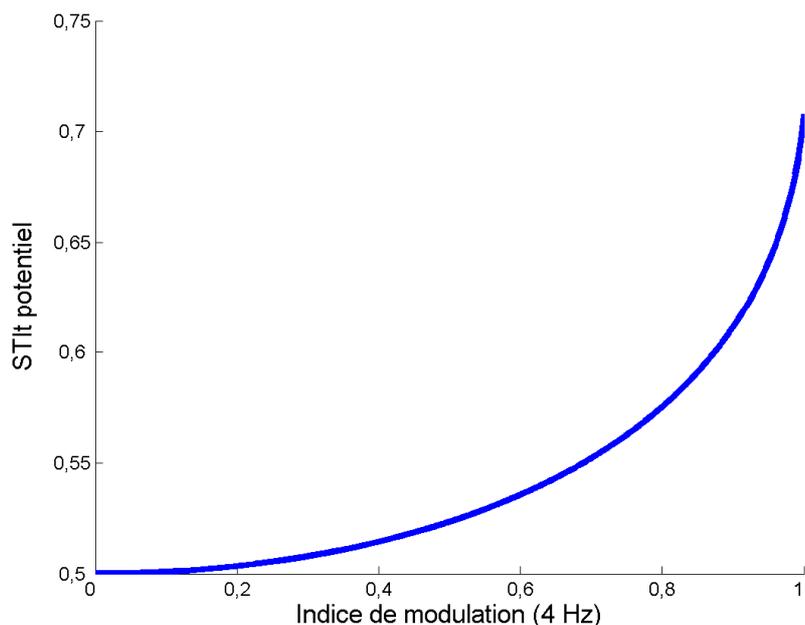


Figure 18 : Valeur du STIt pour un bruit de fond modulé à 4 Hz masquant un bruit de parole de même spectre en fonction de la valeur de l'indice de modulation.

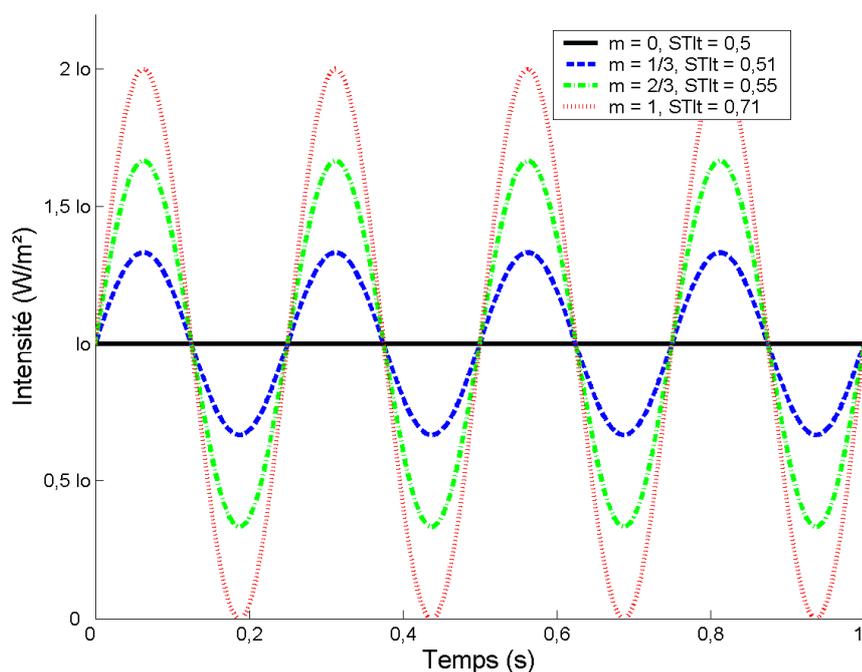


Figure 19 : Enveloppes de la courbe d'intensité de signaux modulés en intensité à 4 Hz, pour différentes valeurs de l'indice de modulation. Les valeurs de STIt relatives à chaque courbe sont indiquées en légende.

Les valeurs relatives aux ambiances sonores utilisées au cours de cette deuxième expérience sont consignées dans le tableau 8. Celles-ci confirment que les bruits les plus fluctuants étaient bel et bien les plus gênants. De plus, les valeurs obtenues pour les conditions M + 3F et M + 5F sont

extrêmement proches : ceci pourrait expliquer qu'elles aient été, suivant les deux mesures, impossibles à distinguer l'une de l'autre.

| Condition | M + 1 F | M + 3 F | M + 5 F | M + BS |
|-----------|---------|---------|---------|--------|
| STIt | 0,76 | 0,68 | 0,65 | 0,53 |

Tableau 8 : Valeurs moyennes du STIt pour les ambiances sonores étudiées, considérées comme masquant un signal de parole.

Ainsi utilisé, il semble donc que le STIt peut rendre compte de l'impact négatif des fluctuations de niveau des bruits de fond étudiés sur la gêne sonore. C'est pour cette raison que pour les deux expériences suivantes, cet indicateur sera utilisé pour quantifier les fluctuations de niveau contenues dans les ambiances sonores recueillies en bureaux ouverts. Il s'agira alors de savoir si un lien peut encore être établi entre les valeurs de STIt ainsi calculées et les potentiels de nuisance respectifs de ces bruits de fond.

4.2 Perturbation de la mémoire sérielle

Le déroulement de l'expérience est analogue à celui de la précédente, dans le même lieu et avec les mêmes conditions de présentation des stimuli sonores et visuels. Le panel de sujets est constitué de 45 individus, parmi lesquels se trouvent 22 hommes et 23 femmes âgés de 18 à 62 ans. Tous ont été indemnisés pour leur participation. La moyenne d'âge du groupe est de 24 ans. L'activité de mémoire sérielle est la même que celle réalisée par les sujets de la deuxième expérience, décrite dans la section 3.2.

4.2.1 Stimuli sonores

Les stimuli sonores sont constitués d'enregistrements de bruit ambiant recueillis au cours de plusieurs campagnes de mesure au sein d'entreprises où les employés travaillent en bureaux ouverts. Ces longues pistes sonores ont été découpées en échantillons durant chacun 5 secondes. Seuls ceux ne comportant que du bruit de parole ont été conservés. Ceux-ci ont alors été normalisés pour se conformer au spectre de voix masculine spécifié dans la norme IEC relative au calcul du STI (IEC, 2003). Ont finalement été retenus ceux dont la valeur potentielle de STI était égale, à $\pm 0,005$ près, à l'une des suivantes :

- 0,55 ;
- 0,60 ;
- 0,65 ;
- 0,70.

Au cours de la deuxième expérience, les performances des sujets n'ont pas été affectées par des variations de STI potentiel entre 0,53 (M + BS) et 0,68 (M + 3F). Au vu de ce constat, la plage de valeurs ici proposée, de 0,55 à 0,70, semble réduite. Le risque de ne pas observer de différence entre les conditions, tout du moins en termes de performance, est présent. Pour autant, il convient de remarquer que cette valeur maximale, 0,70, correspond au maximum qui ait pu être observé pour les échantillons utilisés. Il s'agit d'une valeur extrême parmi les enregistrements d'ambiances de travail recueillis.

Pour chaque valeur, les échantillons sélectionnés ont enfin été concaténés pour obtenir des pistes longues de 10 minutes. Ce sont ces morceaux qui ont été diffusés pendant les blocs expérimentaux. Comme précédemment, le niveau de présentation des sons était fixé à 46 dB(A). Pour la condition de contrôle, un bruit stationnaire de même spectre et de même niveau moyen que les sons de test a été fabriqué, correspondant ainsi à un STI potentiel de 0,50.

4.2.2 Résultats

Les modèles et analyses statistiques sont de même types que précédemment. Les moyennes d'erreurs par bloc et les scores RTLX moyens, ainsi que les écarts-types associés, sont présentés dans le tableau 9.

| Condition | Décrément de performance (%) | | RTLX | |
|-------------|------------------------------|------------|---------|------------|
| | Moyenne | Ecart-type | Moyenne | Ecart-type |
| Contrôle | - | - | 46,0 | 15,4 |
| STIt = 0,55 | 1,95 | 9,44 | 50,5 | 14,5 |
| STIt = 0,60 | 3,37 | 8,31 | 49,9 | 14,3 |
| STIt = 0,65 | 1,31 | 8,26 | 50,8 | 12,8 |
| STIt = 0,70 | 3,40 | 8,60 | 50,4 | 14,3 |

Tableau 9 : Récapitulatif des mesures de performance et de charge cognitive.

4.2.2.1 Performance

Dans la condition silencieuse, le nombre moyen d'erreurs par série est de 1,91. On retrouve un effet d'apprentissage : cette moyenne vaut 2,21 le premier jour pour redescendre à 1,61 le second. Pour l'ensemble des conditions, le décrément de performance est peu élevé et l'influence de la variable STIt semble faible, comme en témoigne la Figure 20. Ceci est confirmé par les résultats issus de la modélisation statistique. En effet, s'il s'avère que les ambiances de bureaux prises dans leur ensemble sont significativement plus gênantes que la condition de contrôle ($\chi^2(1) = 9,37 : p < 0,01$), le facteur « condition sonore » n'a aucun effet significatif sur le décrément de performance parmi les conditions de test ($\chi^2(3) = 2,52 : p < 0,47$).

Comme précédemment, afin d'explorer les effets éventuels de la variabilité interindividuelle, une classification hiérarchique des sujets a été effectuée en se basant sur les décréments de performance de chacun pour les quatre valeurs de STIt. Cette analyse fait ressortir deux groupes de participants, qui réagissent différemment aux ambiances sonores proposées. Les 27 membres du premier groupe constituent 60% du panel. Pour ces individus, le bruit stationnaire engendre significativement moins d'erreurs que les différentes ambiances de bureaux ($\chi^2(1) = 28,31 : p < 0,01$). Le décrément de performance moyen est de 5,92 %, contre 2,51% pour l'ensemble du panel. Ainsi, à l'inverse, les 18 membres du second groupe sont significativement avantagés par la présence de bruit de bureau ($\chi^2(1) = 8,94 : p < 0,01$). Pour cet ensemble de participants, le décrément de

performance moyen, négatif, vaut -2.61 %. La Figure 21 présente les résultats moyens des sujets pour chacun des groupes.

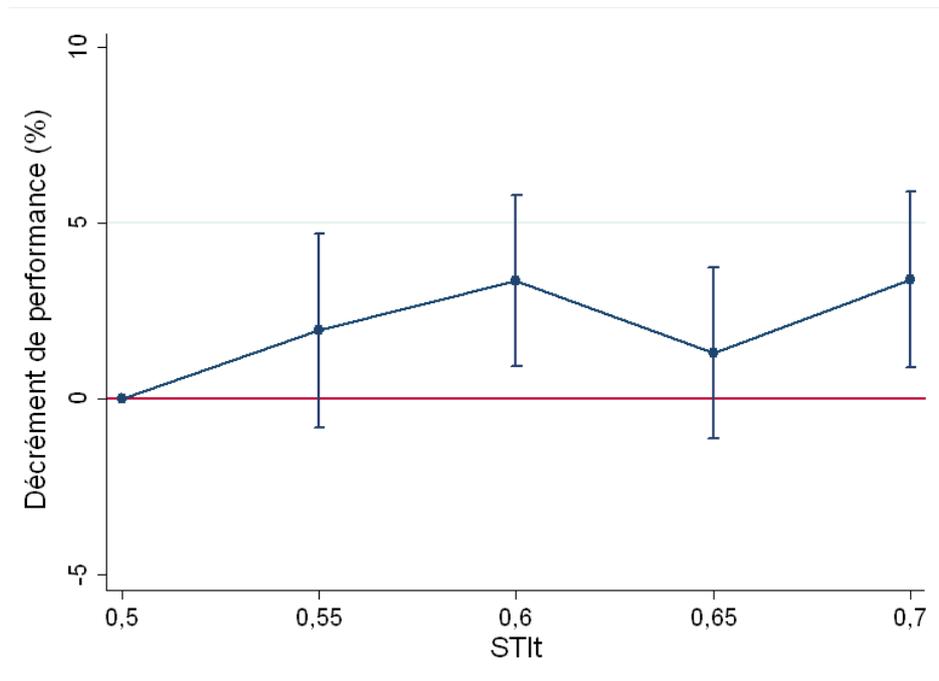


Figure 20 : Décréments de performance pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance.

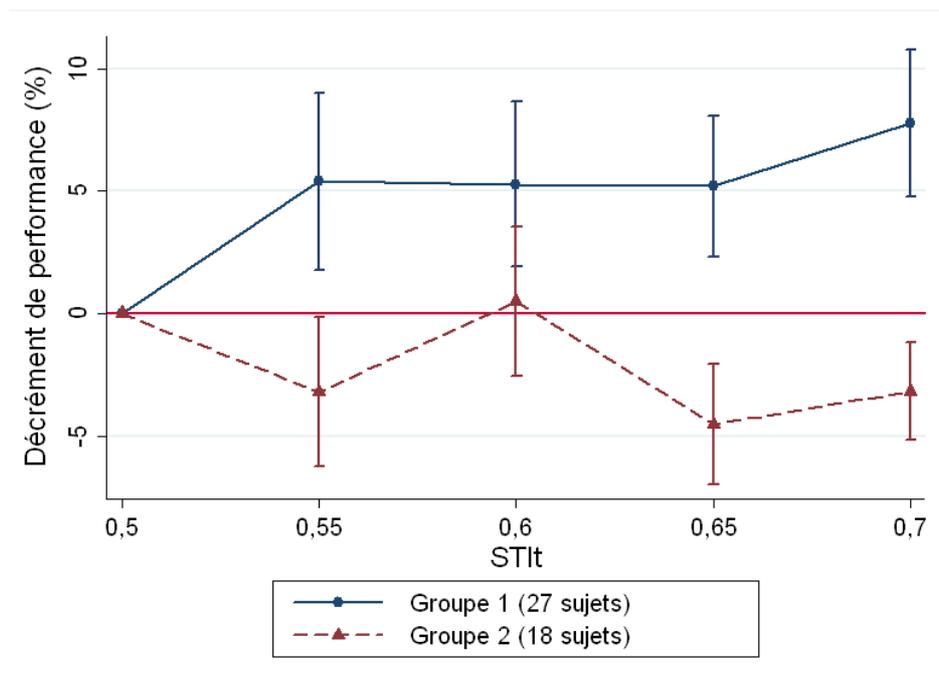


Figure 21 : Décréments de performance pour les valeurs de STIt étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance.

Cependant, en ne considérant que les conditions de test, la valeur du STIt n'a d'effet significatif sur le décrétement de performance pour aucun des deux groupes. Les deux groupes diffèrent vis-à-vis de l'effet global des bruits de bureaux, mais pas en termes de sensibilité aux fluctuations de niveau de ceux-ci.

4.2.2.2 Charge cognitive

En ce qui concerne les réponses au questionnaire d'auto-évaluation de la charge de travail, l'effet de la condition sonore est du même type. Ceci est visible sur la Figure 22.

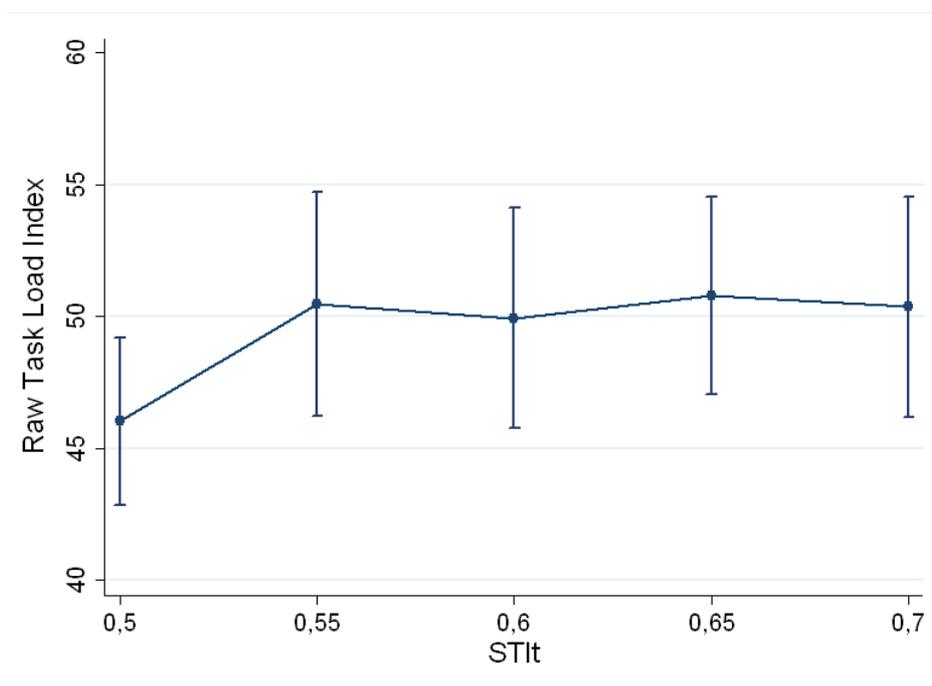


Figure 22 : RTLX obtenus pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance.

En effet, d'une part, les conditions de test sont significativement plus exigeantes que la condition de contrôle ($\chi^2(1) = 19,28 : p < 0,01$). Cependant, d'autre part, les ambiances sonores de bureaux ne sont pas différenciables en termes de RTLX ($\chi^2(1) = 0,18 : p = 0,98$). Il est à noter que comme pour l'expérience 1, les schémas de réponse des participants des deux groupes sont similaires pour le RTLX, comme le montre la Figure 23. La seule différence provient du fait que pour le second groupe, la différence entre bruit stationnaire et bruit de bureau devient elle-même négligeable ($p = 0,85$ pour l'effet global du STIt).

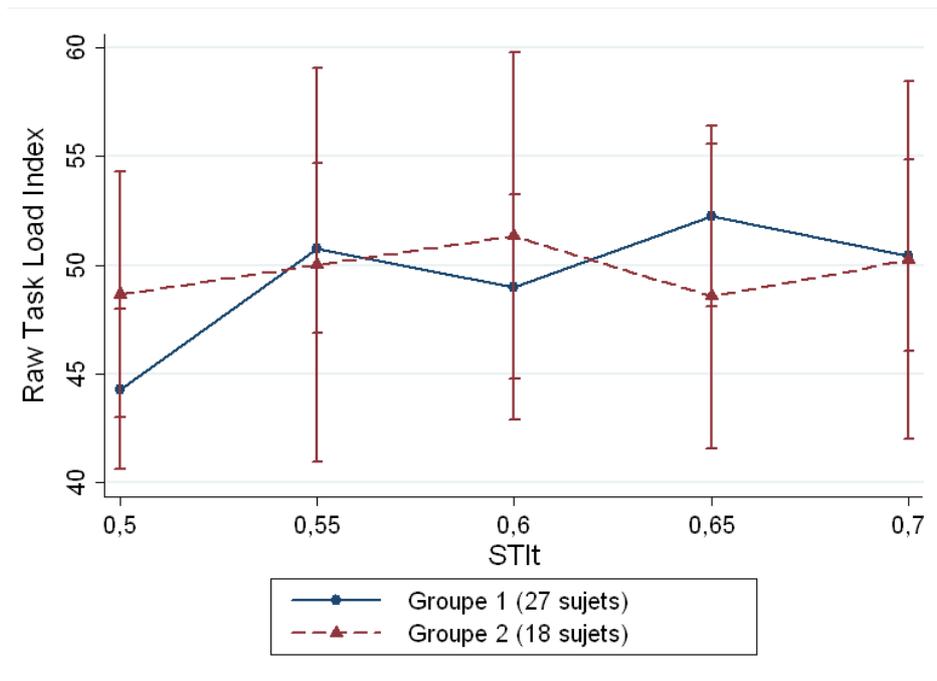


Figure 23 : RTLX obtenus pour les valeurs de STIt étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance.

En utilisant le questionnaire comme outil de diagnostic de la charge de travail, il apparaît que les scores pour les deux premières questions (exigence mentale, physique) ne sont pas sensibles à la condition sonore. L'effet de la valeur du STIt est significatif pour les quatre autres questions (traitant de l'exigence temporelle, de la performance, de l'effort et de la frustration). Cependant, à l'exception de la question 3, relative à l'exigence temporelle de la tâche, cet effet se limite à traduire l'avantage qu'il y a à travailler dans le bruit stationnaire. Pour cette question 3, les scores sont significativement plus élevés pour STIt = 0,65 ou 0,70 que pour STIt = 0,55 ou 0,60 ($\chi^2(1) = 8,18 : p < 0,01$).

Ceci signifie que dans une réelle ambiance sonore de bureau, les sujets doivent agir plus vite qu'ils ne le feraient dans une situation confortable. De plus, les participants ont considéré que ces environnements les poussaient à faire plus d'efforts pour atteindre un niveau de performance adéquat, ce dernier étant alors tout de même abaissé. Enfin, la frustration ressentie était plus grande lorsque le bruit ambiant n'était pas stationnaire. Cependant, seule l'exigence temporelle de la tâche dépend plus finement de la valeur du STIt.

4.2.3 Discussion

Cette expérience a été élaborée dans le but d'explorer les effets des fluctuations de niveau de réelles ambiances sonores de bureaux ouverts sur la nuisance sonore. Il apparaît que les effets sur la performance ainsi que sur la surcharge mentale perçue sont négligeables. A première vue, ceci semblerait indiquer que les fluctuations du bruit ambiant n'ont pas d'influence sur la nuisance

sonore pendant le travail. Cependant, comme expliqué dans la section 4.1.1, les valeurs de STIt sont situées dans une plage assez réduite en comparaison avec celle, par exemple, de la deuxième expérience. La Figure 24 permet, par exemple, de comparer les résultats obtenus dans les deux expériences.

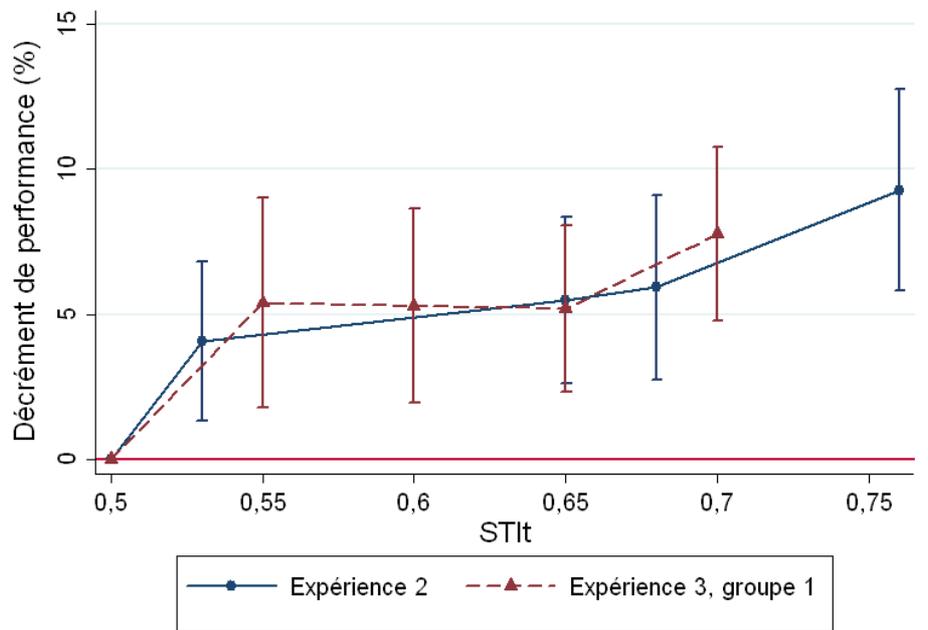


Figure 24 : Comparaisons des performances pour les expériences 2 et 3 en fonction du STIt.

Pour la troisième expérience, le premier groupe, pénalisé par la présence de bruit de fond fluctuant, est retenu. Les deux courbes sont fortement similaires. De plus, l'échelle des abscisses, exprimée en STIt, met en évidence le caractère extrême de la valeur de STIt correspondant à la condition sonore M + 1F de la deuxième expérience. Il semble donc que les conditions réelles étudiées ne contiennent pas de situations suffisamment extrêmes pour être remarquablement plus gênantes que d'autres. Ceci pourrait être une explication pour le manque d'impact du facteur « condition sonore » dans cette expérience.

Pour ces deux expériences, les valeurs de RTLX ne sont pas sensibles à la condition sonore pour les ambiances ne comportant que des bruits de parole. La tâche de sériation présente la particularité d'être, *a priori*, particulièrement sensible aux variations temporelles du bruit ambiant en termes de performance, comme expliqué dans la section 2.3.2. Cependant, rien dans la littérature n'indique que cette spécificité s'étende à des évaluations subjectives telles que celle permettant d'obtenir le RTLX. Il est possible que pour d'autres tâches, la sensibilité du RTLX à la condition sonore soit plus forte que pour un exercice de mémoire sérielle. Pour apporter des éléments de réponse à cette interrogation, une quatrième expérience a été menée.

4.3 Perturbation d'autres activités

Pour cette expérience, 41 sujets ont été recrutés. Le panel comportait 22 femmes et 19 hommes, âgés de 18 à 62 ans. La moyenne d'âge des sujets s'élevait à 25 ans. Les ambiances sonores dans lesquelles ils ont été amenés à travailler étaient les mêmes que celles retenues pour l'expérience précédente. Cependant, les activités proposées ont été modifiées. Une enquête menée en parallèle de ces travaux a permis de montrer que les occupants de bureaux ouverts étaient particulièrement gênés par le bruit de parole ambiant lorsqu'ils devaient réaliser une tâche liée à leurs capacités verbales, par exemple à la lecture ou à la rédaction de textes (Pierrette *et al.*, 2013). Ce constat a dirigé le choix des exercices choisis, présentés ci-dessous.

4.3.1 Exercice de raisonnement grammatical

Pour étudier l'acuité des processus de raisonnement logico-verbal des participants dans les environnements sonores proposés, une tâche de raisonnement grammatical a été utilisée. Cet exercice est inspiré d'un item de la batterie de tests AGARD STRES pour l'évaluation de la performance sous contrainte (ASGARD, 1989). Il s'agit d'un test dont la fiabilité, la validité et la sensibilité sont reconnues comme étant satisfaisantes. Au cours de cette expérience, les sujets ont été confrontés à une adaptation de cette tâche, présentée ci-dessous.

L'expérience est présentée au sujet à l'aide d'une interface graphique sur un écran d'ordinateur. Trois symboles, « # », « ! » et « & », sont affichés sur une ligne au milieu de l'interface, dans un ordre aléatoire. Au dessus de cette ligne, deux propositions relatives à l'ordre des symboles sont présentées. Les propositions utilisent les verbes « suivre » ou « précéder ». L'une des phrases est formulée de façon affirmative, l'autre étant à la forme négative. De même, l'une des phrases est écrite à la forme active, l'autre à la forme passive. Un exemple de combinaison possible est présenté ci-dessous :

- ! est suivi de #
- & ne précède pas !
- & # !

Il est alors demandé au sujet d'évaluer la valeur de chacune des propositions. Deux boutons lui sont proposés. Le premier comporte la mention « Identique » tandis que sur le second est inscrit « Différent ». Si les phrases sont toutes deux vraies ou toutes deux fausses, leurs valeurs sont identiques. Le sujet doit alors cliquer sur le bouton correspondant. Au contraire, si l'une des phrases est vraie tandis que l'autre est fausse, leurs valeurs sont différentes : le participant doit cliquer sur le second bouton. Dans l'exemple, ci-dessus, la première phrase est fausse : le symbole « ! » est situé

en troisième position, le symbole « # » en deuxième : c'est « # » qui est suivi de « ! », et non l'inverse. La seconde proposition est également fautive : le symbole « & » étant placé en première position, il précède nécessairement le troisième symbole, « ! ». Les deux phrases étant toutes deux fautes, elles ont la même valeur : le sujet doit cliquer sur le bouton « Identique ».

Cet exercice a déjà été utilisé par Schlittmeier *et al.* (2008b) dans une étude de l'effet de l'intelligibilité de la parole sur l'ISE. Dans cette expérience, les performances des sujets n'étaient pas influencées par la variation de l'intelligibilité d'un discours perturbant. Il en allait de même pour le simple jugement subjectif de l'effet de l'ambiance sonore. Par contre, d'après Liebl *et al.* (2012), cette activité est sensible aux variations d'intelligibilité d'une voix potentiellement perturbante en termes de performance.

4.3.2 Exercice de fluence verbale

La formalisation d'idées, préalable à l'écriture d'un texte, nécessite d'accéder efficacement à un ensemble de champs lexicaux. Pour étudier l'impact de conditions adverses sur ce processus, une tâche de fluence verbale catégorielle a été mise en place. Il s'agit d'un type d'exercice régulièrement utilisé pour le suivi de l'état de patients atteints de maladies dégénératives cérébrales, mais qui peut également être utilisé plus simplement comme outil de mesure classique de performance (Gierski et Ergis, 2004). La tâche consiste à demander au participant de donner, dans un laps de temps limité, le maximum de mots se rapportant à une catégorie donnée ; par exemple, si la catégorie est « animaux », le participant dispose du temps imparti pour donner le plus grand nombre d'animaux lui venant à l'esprit, comme « chien », « chat », « ours » ou « lion ». La tâche peut se dérouler à l'oral, auquel cas l'expérimentateur prend note des propositions avancées par le sujet, ou bien à l'écrit, sur papier ou au clavier dans le cas où le support de restitution est informatique. Ici encore, la difficulté des items proposés doit être prise en compte. Pour ce faire, il est possible de s'appuyer sur les travaux normatifs menés par Bueno et Megherbi (2009). Menés sur 218 sujets au total, ils fournissent une information sur le nombre moyen de réponses auquel il est raisonnable de s'attendre pour chaque catégorie. Les catégories pour lesquelles la fluence est la plus élevée peuvent donc être considérées comme les plus faciles et vice versa. La moyenne du total de mots, pour 30 secondes, va ainsi de 9,6 pour la catégorie « un membre de la famille » à 3,1 pour « un mandat politique ». Ces résultats permettent d'identifier des groupes de catégories de difficulté similaire.

Jahncke *et al.* (2013) utilisent cette tâche pour étudier la sensibilité des performances d'un individu à l'intelligibilité d'un signal de parole voisin. Dans leur expérience, aucun effet significatif de la condition sonore n'a pu être observé pour cet exercice.

4.3.3 Protocole expérimental

4.3.3.1 Déroulement de l'expérience

Pour chaque sujet, l'expérience se déroulait comme suit. Dans un premier temps, l'expérimentateur présentait au sujet les deux exercices puis le questionnaire NASA-TLX. Ensuite, le sujet, seul, devait effectuer six blocs expérimentaux. Le premier, dans le silence, correspondait à un entraînement ; les cinq autres étaient chacun effectués dans l'une des conditions sonores étudiées (les quatre conditions de test ainsi que la condition de contrôle). L'ordre de présentation des conditions a été équilibré sur les sujets. Chaque bloc expérimental, effectué dans une condition sonore donnée, se déroulait comme ci-dessous :

- l'un des deux exercices ;
- le questionnaire NASA-TLX, se rapportant au premier exercice ;
- l'autre exercice ;
- le questionnaire NASA-TLX, se rapportant au second exercice ;
- une pause de deux minutes avant le bloc suivant.

L'ordre de présentation des exercices était déterminé aléatoirement pour chaque bloc. Pour chaque bloc, quatre informations étaient donc enregistrées : la mesure de performance ainsi que le score RTLX, pour chacune des deux tâches. L'expérience s'est déroulée au même endroit et dans les mêmes conditions de présentation des stimuli que les expériences 2 et 3. Pour chaque sujet, l'ensemble de l'expérience se déroulait en une session d'environ 1h20.

4.3.3.2 Raisonnement grammatical

Dans l'expérience présente, après un compte à rebours de trois secondes, chaque combinaison apparaissait à l'écran pendant 10 secondes. Le sujet devait donner sa réponse durant ce laps de temps, avant le lancement de la combinaison suivante. Chaque bloc expérimental comportait 24 combinaisons. Il est à noter que les phrases négatives sont plus difficiles à interpréter que les phrases affirmatives ; de même, il est plus difficile de juger de la valeur des phrases passives que des phrases à la forme active (ASGARD, 1989). Pour cette raison, les blocs expérimentaux contenaient tous le même nombre de phrases de chaque type (active affirmative, active négative, passive affirmative, passive négative). Le nombre de bonnes réponses par bloc a été retenu comme mesure de performance. L'exercice pouvait donc durer jusqu'à environ 5 minutes par bloc.

4.3.3.3 Fluence verbale

Pour cette tâche, chaque bloc expérimental comportait 2 catégories, une de chaque groupe présenté dans le tableau 10. La présentation de chaque catégorie, au sein de l'interface de test, était

précédée d'un compte à rebours préparatoire. Les réponses devaient être données au clavier. Il a été précisé aux sujets que dans la mesure du possible, les réponses mal orthographiées seraient ramenées au mot correct correspondant. Le temps de réponse accordé était de 30 secondes, pour la première catégorie puis pour la deuxième. Par suite, la validité des mots relativement aux catégories proposées était évaluée. La somme du total de mots pour les deux catégories d'un bloc fournissait une mesure de performance. La durée de l'exercice était donc d'un peu plus d'une minute par bloc. Il faut remarquer que pour cette tâche, la notion de « taux d'erreur » n'existe pas : il n'est donc pas possible de définir un décrement de performance de la même façon que dans la section 2.3.3.

| Fluence moyenne : 4,9 mots / 30 s | Fluence moyenne : 4,6 mots / 30 s |
|------------------------------------------|------------------------------------------|
| Ce qui peut voler | Une arme |
| Ce qui fait du bruit | Un outil de bricolage |
| Un acte puni par la loi | Un arbre |
| Un paysage géographique naturel | Un poisson |
| Un objet de décoration | Un type de chaussure |

Tableau 10 : Catégories utilisées dans l'exercice de fluence verbale. Les fluences moyennes indiquées sont tirées de (Bueno et Megherbi, 2009).

4.3.4 Résultats

L'analyse statistique des résultats a été menée par une modélisation multi-niveau. L'effet de la variabilité interindividuelle, au niveau le plus élevé, est considéré comme un effet aléatoire. Au niveau le plus bas, la condition sonore mais aussi la position de bloc dans l'ordre de passage ont été considérées comme effets fixes. Ceci permet de tenir compte d'éventuels effets d'apprentissage. Les moyennes d'erreurs par bloc et les scores RTLX moyens, ainsi que les écarts-types associés, sont présentés dans les tableaux 11 et 12, relatifs respectivement à la tâche de raisonnement grammaticale et de fluence verbale.

| Condition | Décrément de performance (%) | | RTLX | |
|-------------|------------------------------|------------|---------|------------|
| | Moyenne | Ecart-type | Moyenne | Ecart-type |
| Contrôle | - | - | 50,6 | 17,3 |
| STIt = 0,55 | 1,22 | 17,2 | 48,0 | 17,0 |
| STIt = 0,60 | 0,51 | 14,0 | 53,6 | 16,4 |
| STIt = 0,65 | 2,95 | 19,4 | 53,7 | 16,3 |
| STIt = 0,70 | 4,27 | 19,3 | 54,0 | 17,3 |

Tableau 11 : Récapitulatif des mesures de performance et de charge cognitive pour le raisonnement grammatical.

| Condition | Performance (mots / 60 sec) | | RTLX | |
|-------------|-----------------------------|------------|---------|------------|
| | Moyenne | Ecart-type | Moyenne | Ecart-type |
| Contrôle | 13,6 | 3,51 | 41,7 | 17,0 |
| STIt = 0,55 | 13,1 | 3,43 | 44,1 | 13,8 |
| STIt = 0,60 | 13,3 | 4,17 | 42,7 | 16,3 |
| STIt = 0,65 | 12,8 | 3,46 | 44,8 | 16,3 |
| STIt = 0,70 | 13,2 | 3,89 | 44,7 | 14,7 |

Tableau 12 : Récapitulatif des mesures de performance et de charge cognitive pour la fluence verbale.

4.3.4.1 Raisonnement grammatical

Pour la tâche de raisonnement grammatical, les participants parviennent à donner 17,0 bonnes réponses en moyennes pour la condition de contrôle. Ce chiffre descend jusqu'à 16,0 pour STIt = 0,70. Un effet d'apprentissage est visible : de 14,7 bonnes réponses lors du premier bloc, la performance croît au fil des blocs pour atteindre 18,2 bonnes réponses au cours du cinquième. La modélisation statistique montre que les effets « condition sonore » et « ordre de passage », pris conjointement, agissent significativement sur la performance ($\chi^2(8) = 46,79 : p < 0,01$). Cependant, une analyse plus fine des résultats montre que seul l'effet d'apprentissage est significatif. En effet, comme le laisse deviner la Figure 25, le décrément de performance n'est significativement différent de zéro pour aucune des conditions sonores ($\chi^2(1) = 4,97 : p = 0,29$).

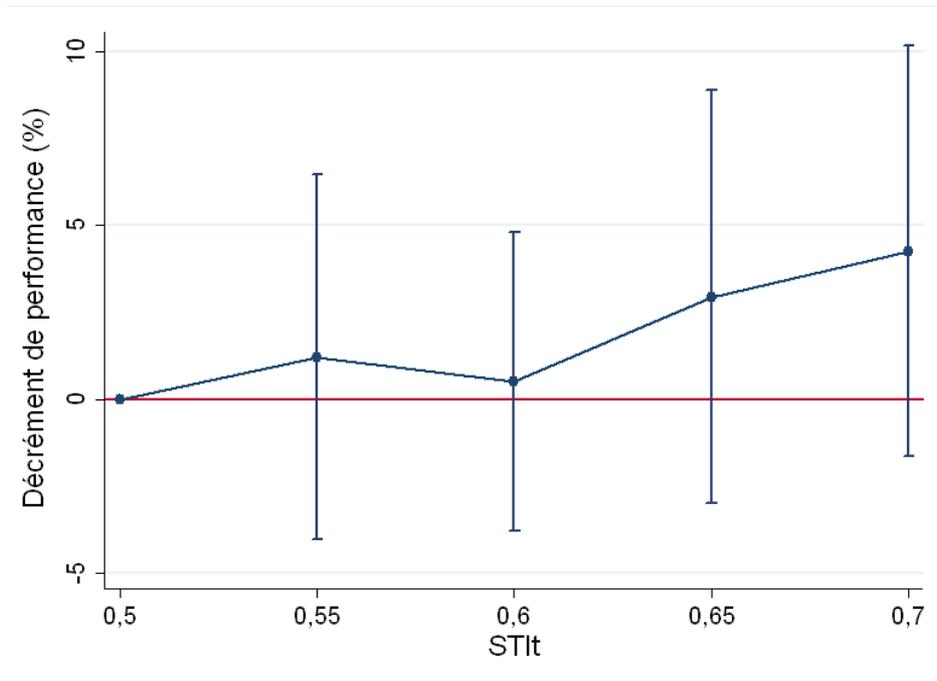


Figure 25 : Décréments de performance pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance.

Pour détailler l'éventuel effet de la variabilité interindividuelle, une classification hiérarchique des sujets a été réalisée. Cette analyse permet de mettre en évidence deux groupes de sujets aux comportements distincts, visibles sur la Figure 26.

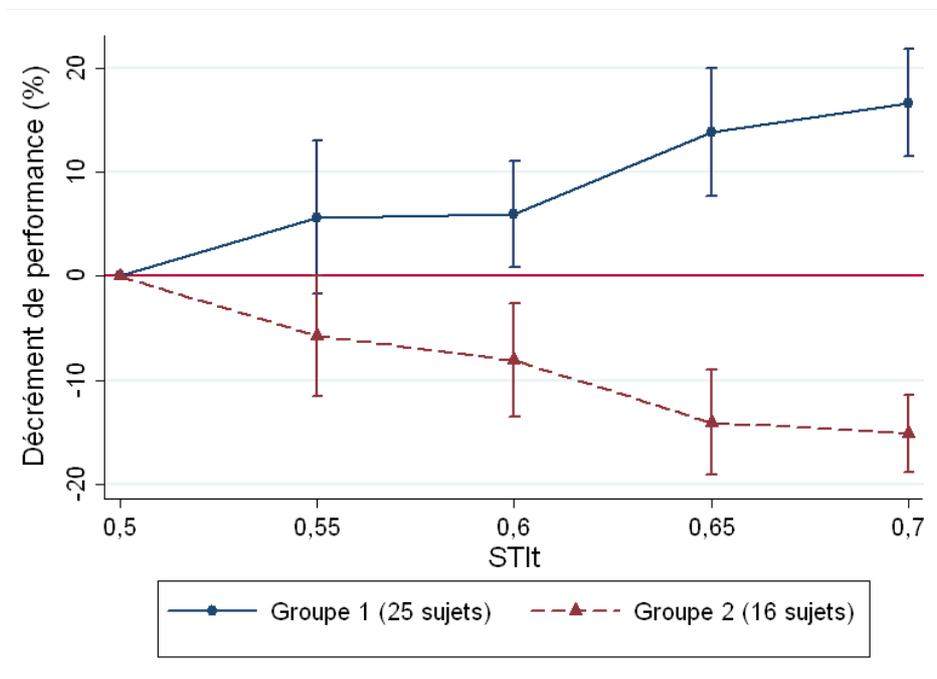


Figure 26 : Décréments de performance pour les valeurs de STIt étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance.

Le premier groupe est composé de 25 sujets, soit 61% du panel. Les performances des membres de ce groupe sont affectées par la présence de bruit de bureau : l'ISE est significatif ($\chi^2(1) = 17,57 : p < 0,01$). De plus, le nombre d'erreurs augmente significativement lorsque la valeur de l'indicateur atteint 0,65 ($\chi^2(1) = 8,81 : p < 0,01$). Les scores sont ainsi significativement plus faibles pour STIt = 0,65 ou 0,70 que pour STIt = 0,55 ou 0,60. Les 16 participants constituant le second groupe sont, eux, significativement moins performants dans la condition de contrôle : l'ISE est négatif ($\chi^2(1) = 21,40 : p < 0,01$). Les conditions de test, par contre, ne sont pas différenciables entre elles. Le comportement de ces sujets est symétrique de celui observé dans le premier groupe : les scores sont significativement meilleures pour STIt = 0,65 ou 0,70 que pour STIt = 0,55 ou 0,60 ($\chi^2(1) = 17,93 : p < 0,01$).

L'étude du RTLX montre que le facteur « condition sonore » a un effet significatif : les conditions pour lesquelles STIt > 0,55 ont été perçues comme plus exigeantes que les autres. En d'autres termes, les valeurs de RTLX sont significativement plus élevées pour STIt = 0,60, 0,65 ou 0,70 que pour STIt = 0,50 ou 0,55 ($\chi^2(1) = 18,23 : p < 0,01$). La Figure 27 permet d'entrevoir ces deux groupes de conditions sonores.

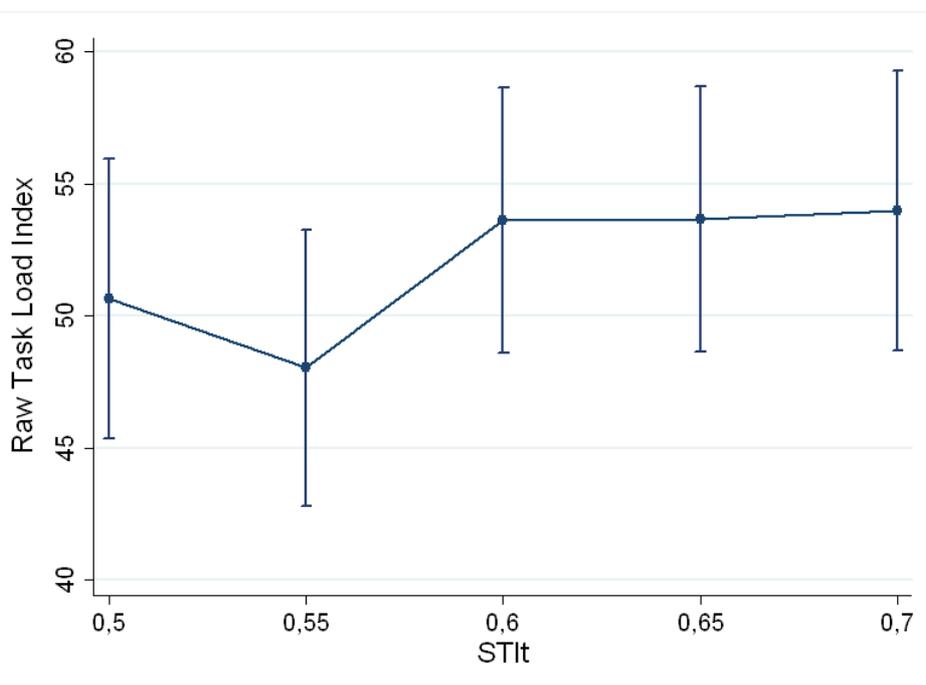


Figure 27 : RTLX obtenus pour les valeurs de STIt. Les moyennes sont accompagnées de leur intervalle de confiance.

L'effet de l'ordre de passage n'est, par contre, pas significatif. L'usage du questionnaire comme outil de diagnostic de la charge de travail montre que seules les réponses aux questions 3 et 4 (exigence temporelle, performance) sont affectées. Il apparaît ainsi que la pression temporelle et l'atteinte perçue à la performance sont plus fortes dès lors que le STIt atteint la valeur de 0,60.

Les réponses des sujets diffèrent suivant le groupe de performance auquel ils appartiennent, ce que la Figure 28 tend à confirmer. En effet, le facteur « condition sonore » n’a pas d’effet significatif sur les sujets du second groupe défini plus haut ($\chi^2(1) = 7,40 : p = 0,11$). Au contraire, pour les membres du groupe 1, le RTLX est significativement plus élevé pour les valeurs de STIt supérieures ou égales à 0,55 ($\chi^2(1) = 18,23 : p < 0,01$). Par ces scores, ces participants indiquent que selon eux, les bruits de fond les plus fluctuants entraînent une surcharge cognitive.

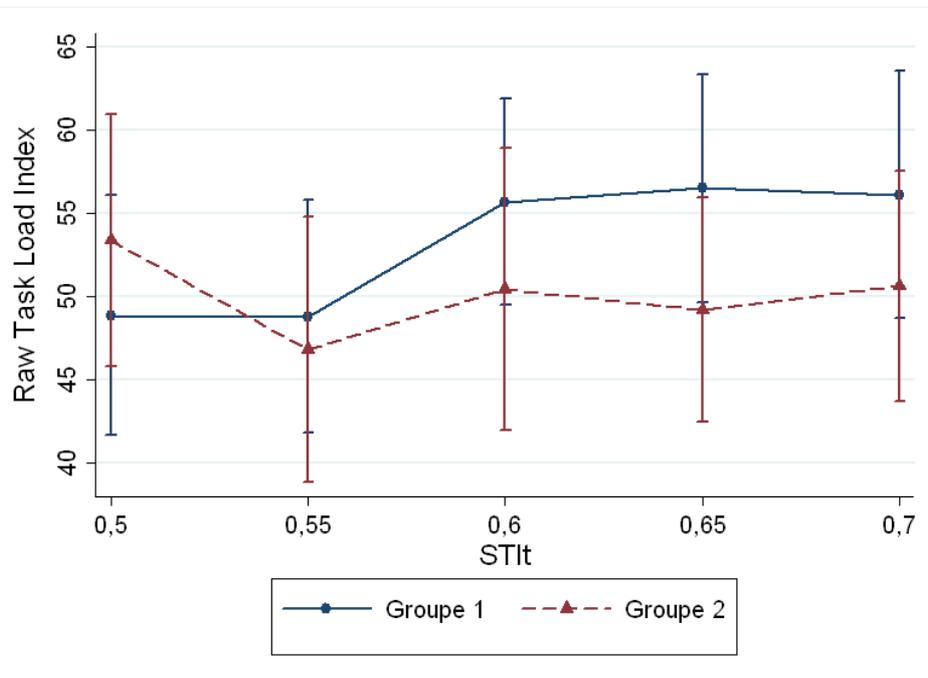


Figure 28 : RTLX obtenus pour les valeurs de STIt étudiées et pour chacun des deux groupes de sujets. Les moyennes sont accompagnées de leur intervalle de confiance.

4.3.4.2 Fluence verbale

Pour la tâche de fluence verbale, sur deux catégories proposées, les sujets restituent une moyenne de 13,6 mots dans la condition de contrôle. La modélisation statistique confirme l’impression donnée par l’observation directe des résultats et des Figure 29 et Figure 30 : ni l’effet d’apprentissage ni celui du STIt ne sont significatifs en ce qui concerne la performance ($\chi^2(8) = 10,48 : p = 0,23$). Le même constat vaut pour l’étude du RTLX (pour le facteur condition, $\chi^2(4) = 5,43 : p = 0,25$), de même que pour chacune de ses composantes. Une classification hiérarchique des sujets n’a pas permis de mettre en évidence des groupes distincts de taille suffisante. L’exercice de fluence verbale se sera avéré totalement insensible aux changements de condition sonore.

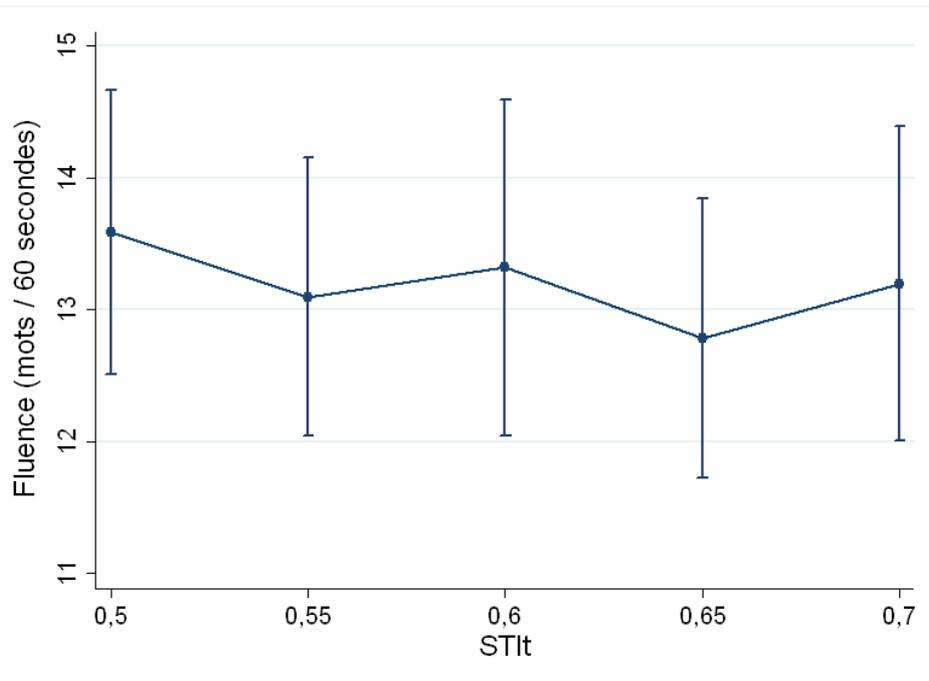


Figure 29 : Fluence verbale pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance.

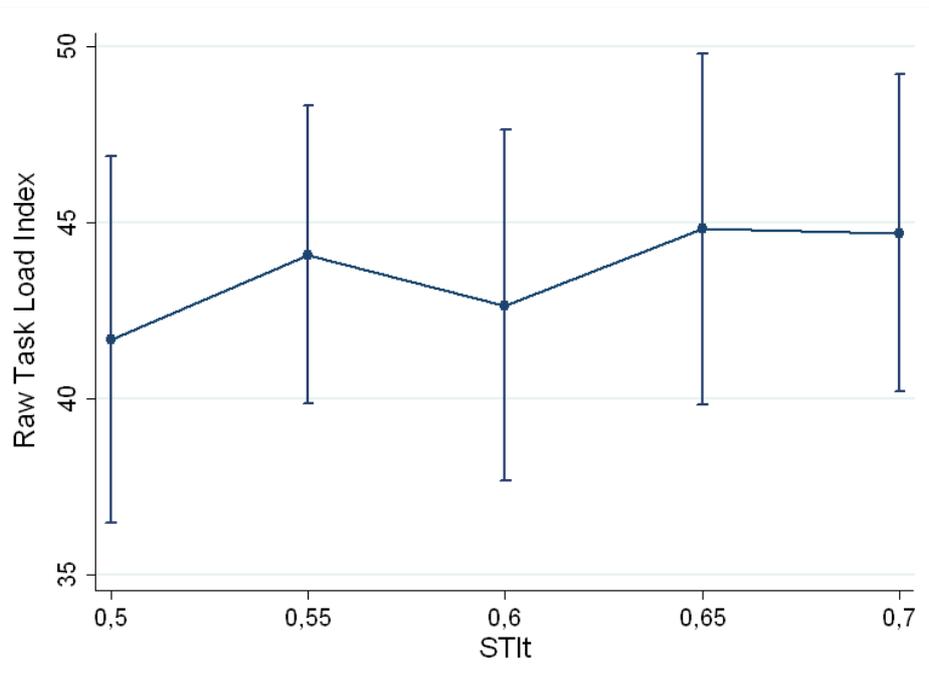


Figure 30 : RTLX obtenus pour les valeurs de STIt étudiées. Les moyennes sont accompagnées de leur intervalle de confiance.

4.3.5 Discussion

Lorsque le panel de sujets est pris dans son ensemble, la performance n'a été affectée par la nature des fluctuations de niveau du bruit de fond pour aucune des deux tâches. Ce résultat n'est pas déconcertant. En effet, *a priori*, ce facteur devait avoir moins grande influence sur la réalisation de

tâches de mémoire sérielle que pour ces exercices. Etant donné que la plage de valeurs choisie pour le *STIt* n'avait pas été suffisamment large pour mettre à jour, au cours de la troisième expérience, des différences de performances en mémoire sérielle, ne pas retrouver non plus d'effet pour d'autres exercices avec les mêmes ambiances sonores était prévisible. Le fait qu'il existe une catégorisation sous-jacente du panel met en avant, encore une fois, la place prépondérante prise par la variabilité interindividuelle dans l'étude de la nuisance sonore.

Il est cependant remarquable que pour la tâche de raisonnement grammatical, les réponses au questionnaire fournissent une information importante : le renforcement des variations de niveau du bruit de fond entraîne bel et bien une surcharge cognitive. Ce constat n'avait pas été retrouvé à l'étude du RTLX au cours de l'expérience précédente. Ce résultat confirme ainsi l'importance de la prise en compte des fluctuations du bruit ambiant tout en attirant l'attention sur le caractère déterminant du choix de la tâche à effectuer.

4.4 Conclusion

L'objectif de ce chapitre était d'étudier l'effet des fluctuations du bruit ambiant sur la nuisance engendrée par une réelle ambiance sonore de bureau riche en bruit de parole. A cette fin, un indicateur, le STIt, a été proposé. Cet outil permet bien de quantifier l'ampleur des variations de niveau du bruit de fond. Pour autant, la troisième expérience, traitant de la mémoire sérielle dans des ambiances sonores réelles, n'a pas permis d'observer d'effet important des fluctuations du bruit de fond. Le chapitre 3 avait pourtant mis en évidence ce facteur comme prépondérant. La tâche de mémoire sérielle, sensible à l'ampleur de ces fluctuations au cours de la deuxième expérience (section 3.2, parole dans un bruit de fond fluctuant), ne l'a pas été dans la troisième. Ceci montre que les stimuli sonores fabriqués en laboratoire, comme dans la plupart des expériences dans le domaine, peuvent ne pas refléter correctement la distribution des conditions sonores réellement rencontrées sur le terrain. Utiliser, pour bruits perturbateurs, des enregistrements en bureaux ouverts est une pratique très peu courante dans la littérature, encore moins avec du bruit de parole. Ceci s'explique par le fait que ces bruits « naturels » offrent moins de possibilité de contrôle pour la définition des variables indépendantes des protocoles expérimentaux. Pour autant, les résultats de cette expérience doivent encourager à la diffusion de cette pratique. A tout le moins, une plus grande attention doit être apportée à la fidélité des conditions de laboratoire aux conditions de terrain.

La quatrième expérience a tout de même permis de retrouver l'effet des fluctuations du bruit de fond sur la gêne pour un autre exercice que celui de mémoire sérielle. Le fait que cet impact soit visible sur le RTLX et non sur le décrétement de performance confirme l'importance déjà établie dans le chapitre 3 d'utiliser, de façon complémentaire, des mesures directes et indirectes de la surcharge cognitive. D'autre part, à la vue de ce résultat, un propos déjà tenu au début de ce document est de nouveau mis en lumière : la nuisance sonore dépend de l'activité perturbée. La présence d'un bruit de fond fluctuant aura bien rendu le travail d'un individu plus difficile et ce, d'autant plus que les fluctuations étaient fortes. Pour autant, il apparaît que l'évaluation de cette perturbation ne peut être effectuée de manière générale sans considérer la nature de la tâche à réaliser.

5 Conclusion

Dans les bureaux ouverts, l'échange libre d'informations entre collègues par le biais de conversations impromptues est censé participer au bien-être des employés et à l'efficacité de leur travail. Cependant, pour les personnes ne prenant pas part à une discussion donnée, la présence de bruit de parole dans l'environnement sonore est reconnue comme étant un important facteur de gêne. D'après les résultats de plusieurs enquêtes de terrain, ceci est d'autant plus vrai que les discours sont compréhensibles. Le but de ce travail était ainsi d'améliorer les connaissances sur la quantification du pouvoir de nuisance d'une ambiance sonore riche en bruit de parole intelligible.

Il s'agit d'un sujet de recherche très actuel, comme a permis de le montrer l'état de l'art établi dans le chapitre 2. Un modèle reliant l'intelligibilité de discours ambiants, mesurée à l'aide du STI, et la performance de travail d'individus sert par exemple de référence à des travaux normatifs récents sur l'acoustique des bureaux ouverts. Cependant, ces recherches bibliographiques ont mis en évidence certaines lacunes dans les connaissances actuelles. Dans un premier temps, l'utilisation du STI pour faire le lien entre intelligibilité des discours environnants et gêne sonore, suscitait quelques questions. En particulier, le modèle de référence souffrait d'un manque de données expérimentales nécessaires à sa validation. Les expériences présentées dans le chapitre 3 ont ainsi consisté à confronter le STI à des environnements sonores divers, mettant en jeu des phénomènes de masquages propres aux bureaux ouverts. Dans un second temps, il convenait de remarquer que pour de véritables ambiances sonores de bureaux ouverts, isoler une voix-cible pour effectuer un calcul d'intelligibilité semblait difficile. Pourtant, la littérature semblait particulièrement centrée sur des environnements sonores synthétiques, plus facilement contrôlables mais risquant d'être moins représentatifs de situations sonores réelles. Pour cette raison, dans le chapitre 4, un effort particulier a été effectué pour quantifier les déterminants de la nuisance sonore d'ambiances sonores enregistrées en bureaux ouverts, sans avoir connaissance *a priori* des natures distinctes du bruit de parole et du bruit masquant.

A ces préoccupations s'en ajoutaient d'autres, concernant les mesures utilisées pour quantifier la nuisance sonore. En effet, l'étude de l'effet du bruit sur le travail d'un individu permet d'insister sur la nécessité d'utiliser conjointement des mesures objectives et subjectives de la surcharge cognitive induite par un environnement sonore perturbant. Or, si les mesures objectives de performance sont systématiques dans l'étude de l'*irrelevant sound effect*, ce n'est pas le cas des mesures subjectives, telles que les questionnaires d'auto-évaluation par le participant lui-même. Qui plus est, lorsque ces mesures sont présentes, il s'agit généralement de créations *ad hoc*, rendant ainsi plus difficile les comparaisons entre protocoles. C'est pour cela que pour l'ensemble des expériences réalisées au cours de ce travail, les mesures de performances ont été accompagnées des

scores recueillis à l'aide du questionnaire NASA-TLX. De même, plusieurs effets peuvent intervenir sur la gêne sonore, l'ambiance sonore elle-même n'étant d'ailleurs pas nécessairement le plus influent. C'est ainsi que différents phénomènes influents auront été étudiés pendant chacune des expériences, en particulier les effets de l'apprentissage et de la variabilité interindividuelle.

Récapitulatif des résultats

Dans le chapitre 3, des environnements sonores contrôlés ont été utilisés. Le protocole de la première expérience (section 3.1) était relativement classique. En effet, la tâche employée, un exercice de mémoire sérielle, était reconnue pour être particulièrement sensible aux variations d'intelligibilité de discours environnants. De plus, les stimuli sonores étaient constitués de bruit de parole masqués par un brouhaha stationnaire. Dans cette situation, le STI est un indicateur performant de l'intelligibilité. Sur l'ensemble du panel, en termes de performance, aucun effet de la variable STI n'a pu être observé parmi les conditions de test. Par contre, pour les sujets sensibles au bruit, la correspondance entre intelligibilité et décrétement de performance a été retrouvée : le STI était alors un bon indicateur du pouvoir de nuisance de l'ambiance sonore. Cependant, la littérature faisait état de doutes sur l'utilisation du STI dès lors que les bruits masquants considérés n'étaient plus stationnaires. En effet, le calcul du STI présuppose la stationnarité du bruit de fond. Or, en bureaux ouverts, d'autres types de bruits peuvent être rencontrés. La seconde expérience s'est concentrée sur le cas où le bruit de fond masquant est un mélange de parole, plus ou moins stationnaire suivant le nombre de voix qui le composait. Ce type d'environnement semblait en effet assez représentatif de ce qui pouvait réellement être rencontré en bureaux ouverts. Il s'est alors avéré qu'à niveau de masquage constant, les situations comportant les masques les plus fluctuants étaient significativement plus perturbantes que les autres. Ceci signifie que pour une même valeur de STI, des environnements sonores pouvaient avoir des potentiels de nuisance différents. Le STI ne permettait donc pas de rendre compte d'un effet potentiellement important dans les bureaux ouverts, celui de la variabilité temporelle du bruit de fond.

Dans le chapitre 4, une solution alternative, le STIt, est ainsi proposée pour pallier cette lacune du STI. Ce nouvel indicateur est construit comme un STI dynamique, permettant de rendre compte de l'éventuel pouvoir masquant d'un bruit de fond en admettant qu'il puisse ne pas être stationnaire. Lorsqu'utilisée pour quantifier l'impact des variations de niveau de l'ambiance sonore, cette mesure présente l'avantage de ne pas être intrusive. Ceci signifie qu'il n'est nul besoin de connaître *a priori* les natures respectives d'un signal-cible et d'un bruit masquant pour en déterminer la valeur. Il a ainsi été possible d'étudier l'effet d'ambiances sonores réelles, aussi représentatives que possible d'environnements sonores en bureaux ouverts. L'efficacité du STIt pour prédire la nuisance sonore a ainsi été tout d'abord évaluée, dans la troisième expérience, pour la tâche

classique de mémoire sérielle utilisée jusqu'alors. Il s'est alors avéré que les conditions sonores utilisées ne différaient pas suffisamment dans leurs pouvoirs de nuisance respectifs. Ce constat est à mettre en rapport avec le fait que la valeur maximale de STIt pour ces ambiances réalistes était bien inférieure à celle de l'expérience précédente, aux ambiances synthétiques. Cependant, dans la quatrième expérience, l'effet de la condition sonore sur la mesure subjective de charge cognitive s'est avéré plus visible pour d'autres tâches que la mémoire sérielle. Ce résultat permet d'insister sur le fait que la nuisance causée par un bruit ambiant n'est jamais totalement découplée de l'activité qu'il perturbe.

Implications pratiques et perspectives

Il ressort de ce travail que le STI ne devrait pas être utilisé pour déterminer le pouvoir de nuisance d'une ambiance sonore donnée dans un bureau ouvert. En présence d'une voix « cible » et d'un bruit de fond stationnaire, le STI rend bel et bien compte de l'impact d'une voix intelligible sur la nuisance sonore. Cet indicateur peut donc être utilisé pour savoir si un locuteur plus ou moins distant, au milieu d'un hypothétique bruit de fond, perturbe le travail d'un exécutant. C'est d'ailleurs de cette façon, au travers de la notion de rayon de distraction évoquée dans la section 2.3.3, que le STI est utilisé dans la norme ISO 3382-3 (2012). Cependant, pour déterminer la gêne sonore subie par un individu dans son contexte de travail actuel, les fluctuations globales du niveau de l'ambiance sonore doivent être prises en compte. Pour cette problématique, le seul STI ne convient pas.

Il est donc nécessaire d'orienter la recherche future dans le domaine vers l'élaboration d'indices mieux adaptés pour cette fonction. Pour ce faire, la prise en compte de l'impact des fluctuations de niveau du bruit de fond dans son ensemble est essentielle. Il s'agit d'une caractéristique importante de l'ambiance sonore en ce qui concerne la nuisance sonore ressentie par les personnes amenées à travailler dans de tels bruits. A ce titre, le STI ici présenté constitue une solution acceptable : cet indicateur rend en effet compte de la dynamique du signal de bruit étudié. De plus, son caractère non-intrusif permet de travailler directement avec le bruit ambiant présent dans l'espace étudié, sans devoir supposer l'existence d'une voix particulière masquée par le bruit de fond.

Cependant, en pratique, la plage de variation du pouvoir de nuisance d'une ambiance sonore de bureau ouvert semble assez restreinte. Au vu des résultats ici présentés, l'amélioration de l'acoustique d'un bureau ouvert consisterait tout d'abord à éviter de générer des situations extrêmes en termes de nuisance sonore. Au delà, les marges de progrès semblent plus minces. Enfin, l'évaluation de l'ambiance sonore d'un bureau ouvert doit dépendre du travail qui y est effectué. Il s'avère nécessaire de classer intelligemment ces espaces de travail suivant les activités principales s'y déroulant en préalable de toute édicition de recommandations sur leur acoustique.

Ce travail ouvre de nombreuses perspectives de recherches. Tout d'abord, il faut remarquer que l'importance du rôle tenu par les mesures subjectives pour l'évaluation de la charge cognitive a été confirmée au cours de chacune des expériences réalisées. L'utilisation d'indices tels que le NASA-TLX doit être généralisée pour favoriser les comparaisons de résultats entre protocoles. De plus, il serait intéressant de pouvoir disposer d'une mesure combinée de surcharge cognitive, combinant les résultats de performance aux scores de RTLX. Ensuite, la dernière expérience a remis en évidence le lien fort entre la nuisance sonore et la tâche étudiée. Pour de futures expériences, il faudrait donc poursuivre la recherche de tâches les plus proches possible de celles réellement effectuées en bureaux ouverts.

De plus, il s'est avéré que les environnements synthétiques largement utilisés dans la recherche sur l'effet perturbateur du bruit ambiant dans les bureaux ne correspondent pas assez à des situations qui pourraient réellement être rencontrées dans ces espaces. Dans un protocole expérimental, il est plus difficile de contrôler les valeurs des grandeurs étudiées lorsque les échantillons utilisés ne sont pas synthétisés par l'expérimentateur mais, au contraire, proviennent de situations réelles. Pour cette raison, les expériences étudiant l'ISE s'appuient souvent sur des enregistrements de parole de très bonne qualité (émissions radio, enregistrement pour l'audiométrie) plus ou moins masqués par des bruits également maîtrisés (bruits blancs, roses ou plus généralement large bande). Or, au vu des résultats obtenus ici, les ambiances synthétiques les plus gênantes le sont trop pour être réalistes. Ce constat doit encourager les expérimentateurs, pour continuer de faire progresser les connaissances sur le sujet, à utiliser plus de sons « naturels » dans leurs protocoles.

A ce sujet, il faut cependant noter qu'un facteur important pour l'intelligibilité de la parole en milieu bruyant n'a pas été étudié ici : la spatialisation. Comme évoqué dans la section 2.2.5, lorsque plusieurs voix sont présentées simultanément, leurs intelligibilités respectives sont meilleures lorsque les sources sont spatialement séparées. Ainsi, des signaux de paroles diffusés par l'intermédiaire d'un haut-parleur seront moins intelligibles que si les voix provenaient d'endroits différents. Dans ce cas, il est envisageable que la nuisance sonore soit également accentuée ; Jones et Macken (1995) valident cette hypothèse. Dans la dernière expérience de cet article, ils montrent que le pouvoir de nuisance d'un brouhaha de six locuteurs est plus fort lorsque chacun des six haut-parleurs utilisés ne diffuse qu'une seule voix, plutôt que les six en même temps. Or, pour les quatre expériences, la présentation des stimuli sonores était identique : un unique haut-parleur frontal. Il est donc possible que le pouvoir de nuisance réel des ambiances recueillies ait été atténué par la présentation monophonique privilégiée ici. Ceci pourrait d'ailleurs expliquer les variations réduites de pouvoirs de nuisance observées pour les ambiances sonores réelles du chapitre 4. L'effet de la

spatialisation des sources de parole sur la perturbation subie par l'exécutant mérite donc d'être étudié avec soin.

En ce qui concerne l'outil de prédiction de la nuisance sonore, le STIt proposé dans ce travail semble être une bonne piste de départ. Il présente l'avantage d'être non-intrusif, donc d'utilisation pratique en bureaux ouverts. De plus, il permet de rendre compte du caractère fluctuant du bruit ambiant, propriété que les deux premières expériences ont permis d'identifier comme étant décisive. Cependant, des pistes d'améliorations sont envisageables. En particulier, la *changing state hypothesis*, évoquée dans la section 2.3.2 comme cause possible du fort potentiel de nuisance du bruit de parole, stipule qu'un environnement sonore est perturbant s'il remplit les deux conditions suivantes : le bruit doit être segmentable et les segments successifs doivent être suffisamment variés les uns par rapport aux autres. Le STIt, tel que construit actuellement, peut mettre en évidence les modulations temporelles contenues dans un bruit de fond. En ceci, il permet de savoir si la première partie de l'hypothèse est vérifiée. Pour ce qui est de la seconde, cependant, le STIt n'est pas adapté. Pour construire un meilleur indicateur, il conviendrait donc de tenir également compte des variations temporelles du spectre du bruit de fond. A ce titre, des travaux comme ceux de Chi *et al.* (1999), ou encore ceux de Elliot et Theunissen (2009), sur l'importance de la perception des modulations spectro-temporelles de la parole pour son intelligibilité pourraient servir d'inspiration.

Enfin, il est à noter que les exercices étaient effectués sur des durées très courtes, allant d'une minute pour la tâche de fluence verbale à 10 minutes pour celle de mémoire sérielle. Il est cependant envisageable que certains effets du bruit de fond ne soient visibles qu'après une exposition suffisamment longue, de plusieurs heures voire de plusieurs jours. La durée d'exposition au bruit des occupants de bureaux ouverts est en effet sans doute à rapprocher du temps passé sur leur lieu de travail : en comparaison, la durée des tests menés au cours de ce travail est minime. Comme précédemment pour la sélection d'ambiances sonores, le choix de mener des expériences sur des créneaux relativement courts relève pour l'expérimentateur de considérations pratiques. En l'occurrence, il devient difficile et onéreux de recruter des sujets et d'organiser leur passage à mesure que leur temps de présence sur le site de l'expérience augmente. Un protocole centré sur l'influence de la « dose » de bruit reçu, ainsi que des éventuels effets de fatigue et d'habituation, fournirait sans doute des résultats intéressants.

Bibliographie

- AFNOR (1974). Acoustique - Courbes NR d'évaluation du bruit. **NF S 30-010**.
- ANSI (1997). Methods for calculation of the Speech Intelligibility Index.
- ANSI (2007). Procedure for the computation of loudness of steady sounds. **ANSI S3.4-2007**.
- ASGARD (1989). Human Performance assessment methods. Neuilly-sur-Seine, France, Advisory Group for Aerospace Research and Development - NATO.
- Ayr, U., Cirillo, E., Fato, I. et Martellotta, F. (2003). "A new approach to assessing the performance of noise indices in buildings." Applied Acoustics **64**: 129-145.
- Banbury, S. et Berry, D. C. (1998). "Disruption of office-related tasks by speech and office noise." British Journal of Psychology **89**: 499-517.
- Banbury, S. P., Macken, W. J., Tremblay, S. et Jones, D. M. (2001). "Auditory distraction and short-term memory: Phenomena and practical implications." Human Factors: The Journal of the Human Factors and Ergonomics Society **43**: 12.
- Beatty, J. et Lucero-Wagoner, B. (2000). The pupillary system. Handbook of Psychophysiology, Second Edition. Cacioppo, J. T., Tassinari, L. G. et Berntson, G. G. Cambridge, Cambridge University Press.
- Becker, F. D., Gield, B., Gaylin, K. et Sayer, S. (1983). "Office design in a community college: Effect on work and communication patterns." Environment and Behavior **15**(6): 699-726.
- Boyce, P. R. (1974). "Users' assessment of a landscaped office." Journal of Architectural Research **3**(3): 44-62.
- Bradley, J. S. et Gover, B. N. (2004). Criteria for acoustic comfort in open-plan offices. Inter-noise 2004: The 33rd International Congress and Exposition on Noise Control Engineering. Prague, République Tchèque: 1-6.
- Bregman, A. S. (1990). Auditory Scene Analysis - The perceptual organization of sound. Cambridge, MA, MIT Press.
- Bronkhorst, A. W. (2000). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions." Acustica - Acta Acustica **86**: 117-128.
- Brookes, M. J. et Kaplan, A. (1972). "The office environment: space planning and affective behavior." Human Factors: The Journal of the Human Factors and Ergonomics Society **14**(5): 373-391.
- Browman, C. P. et Goldstein, L. (1992). "Articulatory phonology : an overview." Phonetica **49**: 155-180.
- Brungart, D. S. (2001). "Informational and energetic masking effects in the perception of two simultaneous maskers." Journal of the Acoustical Society of America **109**(3): 1101-1109.

- Bueno, S. et Megherbi, H. (2009). "French categorization norms for 70 semantic categories and comparison with van Overschelde et al.'s (2004) English norms." Behavior Research Methods **41**(4): 1018-1028.
- Byers, J. C., Bittner, A. C., Jr. et Hill, S. G. (1989). Traditional and raw Task Load Index (TLX) correlations: Are paired comparisons necessary? Advances in Industrial Ergonomics and Safety. Mital, A. London, Taylor & Francis. **1**: 481-485.
- Carhart, R., Tillman, T. W. et Greetis, E. S. (1969). "Perceptual masking in multiple sound backgrounds." Journal of the Acoustical Society of America **45**(3): 694-703.
- Chen, F., Hazrati, O. et Loizou, P. C. (2013). "Predicting the intelligibility of reverberant speech for cochlear implant listeners with a non-intrusive intelligibility measure." Biomedical Signal Processing and Control **8**: 311-314.
- Chi, T., Gao, Y., Guyton, M. C., Ru, P. et Shamma, S. (1999). "Spectro-temporal modulation transfer functions and speech intelligibility." Journal of the Acoustical Society of America **106**(5): 2719-2732.
- Colle, H. A. et Welsh, A. (1976). "Acoustic masking in primary memory." Journal of Verbal Learning and Verbal Behavior **15**(1): 17-31.
- Coutu, M.-F., Nastasia, I., Durand, M.-J., Corbière, M., Loisel, P., Lemieux, P., Labrecque, M.-E. et Pettigrew, S. (2011). Une approche systémique d'identification des déterminants de l'incapacité liés au travail et à la santé psychologique dans un secteur ciblé. Réadaptation au travail, IRSST.
- Delignières, D. (1993). La perception de l'effort et de la difficulté. Cognition et Performance. Famosé, J. P. Paris, INSEP: 183-218.
- Dornic, S., Sarnecki, M. M., Larsson, T. J. et Svensson, J. C. (1974). Performance and perceived difficulty: The effect of noise and distraction. Reports from the Institute of Applied Psychology, University of Stockholm. **51**.
- Drullman, R., Festen, J. M. et Plomp, R. (1994a). "Effect of reducing slow temporal modulations on speech reception." Journal of the Acoustical Society of America **95**(5): 2670-2680.
- Drullman, R., Festen, J. M. et Plomp, R. (1994b). "Effect of temporal envelope smearing on speech reception." Journal of the Acoustical Society of America **95**(2): 1053-1064.
- Ellermeier, W. et Hellbrück, J. (1998). "Is level irrelevant in "irrelevant speech" ? Effects of loudness, signal-to-noise ratio, and binaural unmasking." Journal of Experimental Psychology: Human Perception and Performance **24**(5): 1406-1414.
- Ellermeier, W. et Zimmer, K. (1997). "Individual differences in susceptibility to the "irrelevant speech effect"." Journal of the Acoustical Society of America **102**(4): 2191-2199.
- Elliot, T. M. et Theunissen, F. E. (2009). "The modulation transfer function for speech intelligibility." PLoS Computational biology **5**(3).
- Evette, T. et Fenker, M. (1994). A chacun ses bureaux, le pluralisme des modèles. De l'atelier au territoire : Le travail en quête d'espaces. Evette, T. et Lautier, F. Paris, L'Harmattan: 203-213.

- Falk, T. H., Zheng, C. et Chan, W. (2010). "A non-intrusive quality and intelligibility measure of reverberant and dereverberated speech." IEEE Transactions on Audio, Speech and Language Processing **18**(7): 1766-1774.
- Fischer, G.-N. (1990). Espace, identité et organisation. L'individu dans l'organisation. Chanlat, J.-F. Québec, Presses de l'Université de Laval et Editions Eska: 165-183.
- French, N. R. et Steinberg, J. C. (1947). "Factors governing the intelligibility of speech sounds." Journal of the Acoustical Society of America **19**(1): 90-119.
- Freyman, R. L., Balakrishnan, U. et Helfer, K. S. (2001). "Spatial release from informational masking in speech recognition." Journal of the Acoustical Society of America **109**(5): 2112-2122.
- Gierski, F. et Ergis, A. M. (2004). "Les fluences verbales : aspects théoriques et nouvelles approches." L'année psychologique **104**(2): 331-359.
- Glenberg, A. M. et Swanson, N. G. (1986). "A temporal distinctiveness theory of recency and modality effects." Journal of Experimental Psychology: Learning, Memory and Cognition **12**(1): 3-15.
- Greenberg, S. (2004). Temporal properties of spoken language. 18th International Congress on Acoustics. Kyoto, Japon.
- Greenberg, S., Arai, T. et Grant, K. W. (2006). The role of temporal dynamics in understanding spoken language. Dynamics of Speech Production and Perception. Divenyi, P., Greenberg, S. et Meyer, G. Amsterdam, IOS Press.
- Greenberg, S. et Kingsbury, B. E. D. (1997). The modulation spectrogram: in pursuit of an invariant representation of speech. ICASSP-97, 1997 IEEE International Conference on Acoustics, Speech and Signal Processing. Munich, Allemagne: 1647-1650.
- Guski, R., Felscher-Suhr, U. et Schuemer, R. (1999). "The concept of noise annoyance: How international experts see it." Journal of Sound and Vibration **223**(4): 513-527.
- Haapakangas, A., Kankkunen, E., Hongisto, V., Virjonen, P., Oliva, D. et Keskinen, E. (2011). "Effects of five speech masking sounds on performance and acoustic satisfaction. Implications for open-plan offices." Acta Acustica united with Acustica **97**: 641-655.
- Haka, M., Haapakangas, A., Keränen, J., Hakala, J., Keskinen, E. et Hongisto, V. (2009). "Performance effects and subjective disturbance of speech in acoustically different office types - a laboratory experiment." Indoor Air **19**: 454-467.
- Hanson, E. K. S., Schellekens, J. M. H., Veldman, J. B. P. et Mulder, L. J. M. (1993). "Psychomotor and cardiovascular consequences of mental effort and noise." Human Movement Science **12**(6): 607-626.
- Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later. The Human Factors and Ergonomics Society 50th Annual Meeting. Santa Monica, Etats-Unis d'Amérique: 904-908.
- Hart, S. G. et Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Human Mental Workload. Hancock, P. et Meshkati, N. Amsterdam, North Holland: 139-183.

- Hedge, A. (1982). "The open-plan office: A systematic investigation of employee reactions to their work environment." Environment and Behavior **14**(5): 519-542.
- Hill, S. G., Iavecchia, H. P., Bittner, A. C., Jr., Byers, J. C., Zaklad, A. L. et Christ, R. E. (1992). "Comparison of four subjective workload rating scales." Human Factors: The Journal of the Human Factors and Ergonomics Society **34**(4): 429-439.
- Hirsh, I. J. (1950). "The relation between localization and intelligibility." Journal of the Acoustical Society of America **22**(2): 196-200.
- Hongisto, V. (2005). "A model predicting the effect of speech of varying intelligibility on work performance." Indoor Air **15**: 458-468.
- Hongisto, V., Keranen, J. et Larm, P. (2004). "Simple model for the acoustical design of open-plan offices." Acta Acustica united with Acustica **90**: 481-495.
- Houtgast, T. et Steeneken, H. J. M. (1972). Envelope spectrum and intelligibility of speech in enclosures. IEEE the 1972 Conference on Speech Communication and Processing: 392-395.
- Houtgast, T. et Steeneken, H. J. M. (1973). "The modulation transfer function in room acoustics as a predictor of speech intelligibility." Acustica **28**: 66-73.
- Houtgast, T. et Steeneken, H. J. M. (1985). "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria." Journal of the Acoustical Society of America **77**(3): 1069-1077.
- Hughes, R. et Jones, D. M. (2001). "The intrusiveness of sound : Laboratory findings and their implications for noise abatement." Noise & Health **4**(13): 51-70.
- IEC (2003). Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index, International Electrotechnical Commission. **IEC 60268-16**.
- ISO (1975). Acoustics - Method for calculating loudness level. **ISO 532-1975 (E)**.
- ISO (2012). Acoustics - Measurement of room acoustic parameters - Part 3: Open plan offices, International Organization for Standardization. **ISO 3382-3**.
- Jahncke, H., Hongisto, V. et Virjonen, P. (2013). "Cognitive performance during irrelevant speech: effects of speech intelligibility and office-task characteristics." Applied Acoustics **74**: 307-316.
- Jensen, K. L., Arens, E. et Zagreus, L. (2005). Acoustical quality in office workstations, as assessed by occupant surveys. Indoor Air 2005, The 10th International Conference on Indoor Air Quality and Climate. Pékin, Chine.
- Jones, D. M., Alford, D., Macken, W. J. et Banbury, S. P. (2000). "Interference from degraded auditory stimuli: Linear effects of changing-state in the irrelevant sequence." Journal of the Acoustical Society of America **108**(3): 1082-1088.
- Jones, D. M. et Macken, W. J. (1993). "Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory." Journal of Experimental Psychology: Learning, Memory and Cognition **19**(2): 369-381.

- Jones, D. M. et Macken, W. J. (1995). "Auditory babble and cognitive efficiency: Role of number of voices and their location." Journal of Experimental Psychology: Applied **1**(3): 216-226.
- Jones, D. M., Macken, W. J. et Murray, A. C. (1993). "Disruption of visual short-term memory by changing-state auditory stimuli: The role of segmentation." Memory and Cognition **21**(3): 318-328.
- Jones, D. M., Miles, C. et Page, J. (1990). "Disruption of proof-reading by irrelevant speech: Effects of attention, arousal or memory ?" Applied Cognitive Psychology **4**: 89-108.
- Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, Prentice-Hall.
- Kryter, K. D. (1962a). "Methods for the calculation and use of the Articulation Index." Journal of the Acoustical Society of America **34**(11): 1689-1697.
- Kryter, K. D. (1962b). "Validation of the Articulation Index." Journal of the Acoustical Society of America **34**(11): 1698-1702.
- Labardin, P. (2011). L'espace, un outil de contrôle ? Le cas de Pont-à-Mousson (1921-1930). XXe conférence de l'Association Internationale de Management Stratégique. Nantes, France.
- Landström, U., Söderberg, L., Kjellberg, A. et Nordström, B. (2002). "Annoyance and performance effects of nearby speech." Acta Acustica united with Acustica **88**: 549-553.
- Lashey, K. S. (1951). The problem of serial order in behavior. Cerebral mechanisms in behavior. Jeffress, L. A. New York, Wiley: 112-131.
- LeCompte, D. C. (1996). "Irrelevant speech, serial rehearsal, and temporal distinctiveness: A new approach to the irrelevant speech effect." Journal of Experimental Psychology: Learning, Memory and Cognition **22**(5): 1154-1165.
- Liebl, A., Haller, J., Jödicke, B., Baumgartner, H., Schlittmeier, S. J. et Hellbrück, J. (2012). "Combined effects of acoustic and visual distraction on cognitive performance and well-being." Applied Ergonomics **43**: 424-434.
- Liebl, A., Hawighorst, M., Drotleff, H., Leistner, M., Wack, R. et Schlittmeier, S. J. (2011). The relationship between the Speech Transmission Index and measures of cognitive performance. Forum Acusticum. Aalborg, danemark: 1631-1636.
- Luximon, A. et Goonetilleke, R. S. (2001). "Simplified subjective workload assessment technique." Ergonomics **44**(3): 229-243.
- Macken, W., Tremblay, S., Alford, D. et Jones, D. (1999). "Attentional selectivity in short-term memory: Similarity of process, not similarity of content, determines disruption." International Journal of Psychology **34**(5/6): 322-327.
- Miller, G. A. (1955). "The magical number seven, plus or minus two: Some limits on our capacity for processing information." Psychological Review **101**(2): 343-352.
- Miller, G. A. et Licklider, J. C. R. (1950). "The intelligibility of interrupted speech." Journal of the Acoustical Society of America **22**(2): 167-173.

- Moreno, R. et Mayer, R. E. (2000). "A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages." Journal of Educational Psychology **92**(1): 117-125.
- NASA (1986). NASA Task Load Index (TLX) v1.0 Manual, NASA.
- Neath, I. (2000). "Modeling the effects of irrelevant speech on memory." Psychonomic Bulletin & Review **7**(3): 403-423.
- Nemecek, J. et Grandjean, E. (1973). "Noise in landscaped offices." Applied Ergonomics **4**(1): 19-22.
- Oldham, G. R. et Brass, D. J. (1979). "Employee reactions to an open-plan office: A naturally occurring quasi-experiment." Administrative Science Quarterly **24**: 267-284.
- Oswald, C. J. P., Tremblay, S. et Jones, D. M. (2000). "Disruption of comprehension by the meaning of irrelevant sound." Memory **8**(5): 345-350.
- Paas, F. G. W. C., Tuovinen, J. E., Tabbers, H. et van Gerven, P. W. M. (2003). "Cognitive load measurement as a means to advance cognitive load theory." Educational Psychologist **38**(1): 63-71.
- Paas, F. G. W. C. et van Merriënboer, J. J. G. (1994a). "Instructional control of the cognitive load in the training of complex cognitive tasks." Educational Psychology Review **6**(4): 351-371.
- Paas, F. G. W. C. et van Merriënboer, J. J. G. (1994b). "Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive load approach." Journal of Educational Psychology **86**(1): 122.
- Pashler, H. (1994). "Dual-task interference in simple tasks: Data and theory." Psychological Bulletin **116**(2): 220-244.
- Pierrette, M., Parizet, E. et Chevret, P. (2013). Perception and evaluation of noise sources in open plan offices. ICA 2013 - 21st International Congress on Acoustics. Montréal, Canada.
- Randsell, S. E. et Gilroy, L. (2001). "The effects of background music on word processed writing." Computers in Human Behaviour **17**: 141-148.
- Rhebergen, K. S. et Versfeld, N. J. (2005). "A Speech Intelligibility Index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners." Journal of the Acoustical Society of America **117**(4): 2181-2192.
- Rhebergen, K. S., Versfeld, N. J. et Dreschler, W. A. (2006). "Extended Speech Intelligibility Index for the prediction of the speech reception threshold in fluctuating noise." Journal of the Acoustical Society of America **120**(6): 3988-3997
- Riesz, R. R. (1928). "Differential intensity sensitivity of the ear for pure tones." Physical Review **31**(5): 867-875.
- Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J. et van der Wel, R. (2007). "The problem of serial order in behavior: Lashey's legacy." Human Movement Science **26**: 525-554.

- Schlittmeier, S. J. et Hellbrück, J. (2009). "Background music as noise abatement in open-plan offices : a laboratory study on performance effects and subjective preferences." Applied Cognitive Psychology **23**: 684-697.
- Schlittmeier, S. J., Hellbrück, J. et Klatt, M. (2008a). "Does irrelevant music cause an irrelevant sound effect for auditory items ?" European Journal of Cognitive Psychology **20**(2): 252-271.
- Schlittmeier, S. J., Hellbrück, J., Thaden, R. et Vörländer, M. (2008b). "The impact of background speech on intelligibility: Effect on cognitive performance and perceived disturbance." Ergonomics **51**(5): 719-736.
- Schlittmeier, S. J., Weißgerber, T., Kerber, S., Fastl, H. et Hellbrück, J. (2012). "Algorithmic modeling of the irrelevant sound effect (ISE) by the hearing sensation fluctuation strength." Attention, Perception & Psychophysics **74**(1): 194-203.
- Smith, C. L., Browman, C. P., McGowan, R. S. et Kay, B. (1993). "Extracting dynamic parameters from speech movement data." Journal of the Acoustical Society of America **93**(3): 1580-1588.
- Steeneken, H. J. M. et Houtgast, T. (1980). "A physical method for measuring speech-transmission quality." Journal of the Acoustical Society of America **67**(1): 318-326.
- Steeneken, H. J. M. et Houtgast, T. (2002). "Validation of the revised STI_r method." Speech Communication **38**: 413-425.
- Sundstrom, E., Burt, R. E. et Kamp, D. (1980). "Privacy at work: Architectural correlates of job satisfaction and job performance." Academy of Management Journal **23**(1): 101-117.
- Sweller, J. (1988). "Cognitive load during problem solving: Effects on learning." Cognitive Science **12**: 257-285.
- Tremblay, S. et Jones, D. M. (1999). "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**(4): 1005-1015.
- Trompette, N. et Chatillon, J. (2010). Evaluation des risques pour l'audition des opérateurs des centres d'appels téléphoniques et solutions de prévention. Note scientifique et technique, Institut National de Recherche et de Sécurité.
- van Gerven, P. W. M., Paas, F. G. W. C., van Merriënboer, J. J. G. et Schmidt, H. G. (2004). "Memory load and the cognitive pupillary response in aging." Psychophysiology **41**: 167-174.
- Venetjoki, N., Kaarlela-Tuomaala, A., Keskinen, E. et Hongisto, V. (2006). "The effect of speech and speech intelligibility on task performance." Ergonomics **49**(11): 1068-1091.
- Viemeister, N. F. (1979). "Temporal modulation transfer function based upon modulation thresholds." Journal of the Acoustical Society of America **66**(5): 1364-1380.
- Virjonen, P., Keränen, J. et Hongisto, V. (2009). "Determination of acoustical conditions in open-plan offices: Proposal of new measurement method and target values." Acta Acustica united with Acustica **95**: 279-290.

- Warnock, A. C. C. et Chu, W. T. (2002). Voice and background levels measured in open offices. IRC Internal Report, National Research Council - Conseil National de Recherches Canada.
- Zalesny, M. D. et Farace, R. V. (1987). "Traditional versus open-plan offices: A comparison of sociotechnical, social relations, and symbolic meaning perspectives." Academy of Management Journal **30**(2): 240-259.
- Zimmer, K., Ghani, J. et Ellermeier, W. (2008). "The role of task interference and exposure duration in judging noise annoyance." Journal of Sound and Vibration **311**: 1039-1051.

Annexe A : Items du questionnaire NASA-TLX (NASA, 1986)

Le NASA-TLX est une mesure de la charge de travail basé sur les réponses à un questionnaire. En répondant à six questions, les personnes interrogées fournissent une auto-évaluation de la charge supportée pendant l'exécution d'une tâche. Chaque question fait référence à une composante de la charge de travail : les participants répondent en indiquant, pour chacune, un score allant de 0 à 100. Plus le score est élevé, plus la charge induite par la tâche et les conditions de travail est importante. Cet outil a été développé par le *Human Performance Research Group* de la NASA, aux Etats-Unis. De ce fait, les questions ont été initialement rédigées en anglais. Pour l'ensemble des expériences menées au cours de ce travail, il a donc été nécessaire de traduire ces questions en français. Pour cet effort, la traduction utilisée en 2011 par une équipe de l'IRSST au Québec a servi de point d'appui (Coutu *et al.*, 2011). Le tableau 13 présente chacun des items du questionnaire, dans leur forme originale en anglais ainsi que la traduction en français utilisée au cours de ce travail.

| Composante (anglais/français) | Texte original (NASA, 1986) | Traduction utilisée |
|---------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Mental Demand <i>Exigence mentale</i> | How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving? | Jusqu'à quel point des activités mentales et perceptives étaient requises pour faire votre travail (penser, décider, calculer, se rappeler, regarder, rechercher, etc.) ? Votre travail était-il facile ou exigeant, simple ou complexe ? |
| Physical demand <i>Exigence physique</i> | How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious? | Jusqu'à quel point des activités physiques étaient requises pour faire votre travail (pousser, tirer, tourner, contrôler, activer, etc.) ? Votre travail était-il facile ou difficile, lent ou rapide, relâché ou énergique ? |
| Temporal Demand <i>Exigence temporelle</i> | How much time pressure did you feel due to the rate or pace at which the tasks or tasks elements occurred? Was the pace slow and leisurely or rapid and frantic? | Jusqu'à quel point vous êtes-vous senti pressé(e) par le temps du fait du rythme ou de la vitesse à laquelle votre travail ou les éléments de votre travail s'enchaînaient ? La vitesse était-elle lente et paisible ou rapide et frénétique ? |

| | | |
|----------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Performance</p> <p><i>Performance</i></p> | <p>How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?</p> | <p>Jusqu'à quel point pensez-vous avoir réussi à atteindre les objectifs de travail fixés par l'expérimentateur (ou par vous-même) ? Êtes-vous satisfait(e) de votre performance vis-à-vis de ces objectifs ?</p> |
| <p>Effort</p> <p><i>Effort</i></p> | <p>How hard did you have to work (mentally and physically) to accomplish your level of performance?</p> | <p>Jusqu'à quel point avez-vous eu à travailler dur (mentalement ou physiquement) pour atteindre votre niveau de performance ?</p> |
| <p>Frustration Level</p> <p><i>Frustration</i></p> | <p>How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?</p> | <p>Jusqu'à quel point vous sentiez-vous mal à l'aise, découragé(e), irrité(e), stressé(e) et contrarié(e) par opposition à confiant(e), à l'aise, content(e), relaxé(e) et satisfait(e) de vous durant votre travail ?</p> |

Tableau 13 : Texte des questions formant le questionnaire NASA-TLX.

Annexe B : Présentation du modèle de prédiction du STI en bureaux ouverts (Hongisto *et al.*, 2004)

Ce modèle se propose de prédire la valeur du STI entre deux postes adjacents d'un bureau ouvert en tenant compte de sa géométrie, de l'absorption apportée par ses composants et de la nature des signaux cible et masquants. Il a été utilisé au cours de la première expérience dans le but de recréer des environnements sonores correspondant à la fois aux valeurs de STI désirées et à des situations de bureaux réalistes.

La géométrie du bureau schématisé étudié est présentée dans la Figure 31. Les deux cercles représentent la tête de chacun des auditeurs. Ils sont séparés par une cloisonnette. Dans cette modélisation en 2D, seul le plafond est pris en compte : les deux postes sont supposés ne pas être à proximité des murs du bureau. Pour l'expérience 1, certaines distances sont fixées dans le paragraphe 3.1.2. En particulier, $d = 2x_0 = 2,5$ m et $h_0 = 1,2$ m.

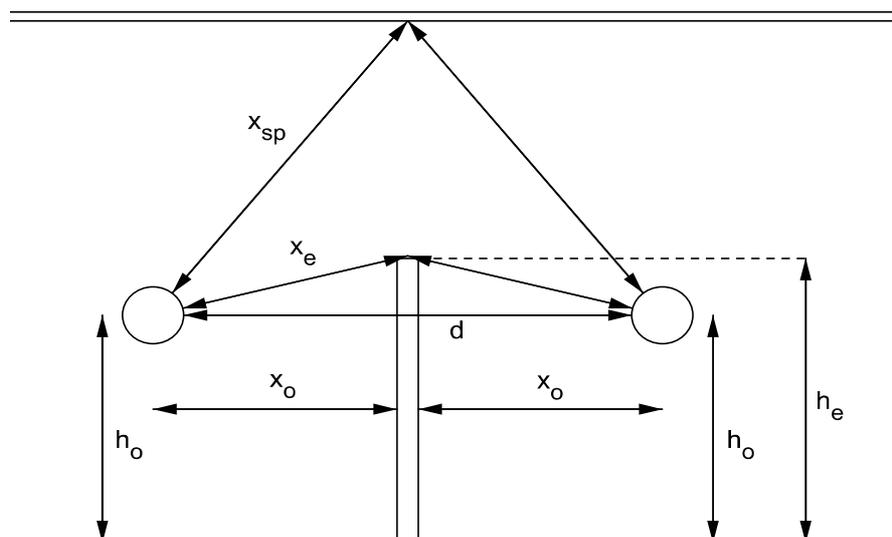


Figure 31 : Représentation schématique des postes de travail étudiés

Soient L_{wp} le niveau de puissance de la source de parole cible et L_{wb} celui de l'ensemble des sources de parole formant le brouhaha masquant, chacun étant fourni par bande d'octave. Pour cette expérience, en début de calcul, les niveaux globaux de la cible et du masques sont fixés à une

valeur arbitraire identique en tenant compte de la pondération A. En champ libre, le niveau du signal-cible à l'oreille de l'auditeur serait le suivant :

$$L_{p,0} = L_{wp} - 10 \log(4\pi d^2). \quad (\text{B.1})$$

Cependant, dans la présente situation, l'impact des premières réflexions est à prévoir. La présence du plafond est, en particulier, à ne pas négliger. A partir de la connaissance de l'absorption du plafond, α_p , l'atténuation liée à la réflexion se détermine comme suit :

$$\Delta L = 10 \log \frac{1}{1 - \alpha_p}. \quad (\text{B.2})$$

La contribution de la réflexion sur le plafond vaut alors :

$$L_{p,1} = L_{wp} - 10 \log(4\pi l^2) - \Delta L, \quad (\text{B.3})$$

avec $l = 2x_{sp}$. De plus, il faut tenir compte de la perte par insertion de la cloisonnette :

$$IL = 10 \log \left(3 + 40 \frac{z}{\lambda} \right), \quad (\text{B.4})$$

où $z = 2x_e - d$ représente la différence de marche et $\lambda = c_0/f$ la longueur d'onde pour la fréquence centrale f de la bande d'octave considérée. L'écran est supposé suffisamment isolant pour que l'énergie transmise par ce dernier soit négligeable devant les autres sources d'énergie : d'après les auteurs du modèle, cette hypothèse est généralement vérifiée. Dès lors, le chemin de transmission directe est remplacé par celui diffracté sur l'arête supérieure de l'écran :

$$L_{p,2} = L_{p,0} - IL. \quad (\text{B.5})$$

Au vu de la courte distance séparant le locuteur et l'auditeur dans cette expérience, l'effet de la réverbération est supposé négligeable devant celui des précédentes contributions énergétiques. Dès lors,

$$L_p = L_{p,1} + L_{p,2} \quad (\text{B.6})$$

représente le niveau de pression à l'oreille de l'auditeur correspondant au signal de parole. Pour les voix plus éloignées, celles qui constituent le brouhaha masquant, l'énergie réverbérée parvenant à l'auditeur s'exprime comme suit :

$$L_b = L_{wb} - 10 \log \left(\frac{4(1 - \alpha_r)}{S_t \alpha_r} \right). \quad (\text{B.7})$$

Dans cette expression, S_t représente l'aire de l'ensemble des surfaces absorbantes présentes dans le local. Pour rappel, le bureau est parallélépipédique et deux fois plus long que large. De plus, chaque opérateur dispose d'une surface au sol s_{op} . L'aire de la surface absorbante équivalente correspondant à un opérateur et à l'ensemble de son matériel est fixée à 10 m². Enfin, la largeur des cloisonnettes est de 1,8 mètre ; à chaque occupant est associée une face d'écran. Il en découle l'expression suivante :

$$S_t = 2S + 3h_{sp}\sqrt{2S} + (10 + 1,8h_e)\frac{S}{s_{op}}. \quad (\text{B.8})$$

L'absorption globale du local α_r est déterminée en calculant la moyenne des coefficients d'absorption des surfaces étudiées pondérée par leurs aires respectives.

Les niveaux de parole et de bruit L_p et L_b sont ensuite conjointement ajustés pour que le niveau global de la somme des signaux à l'oreille de l'auditeur vaille 46 dB(A). Le STI est finalement calculé à partir de la connaissance de ces niveaux suivant la démarche présentée dans la norme IEC 60268-16 (2003).

Annexe C : Prévision de la puissance statistique du protocole de la première expérience

Soient N sujets soumis à I séries. Les variables aléatoires suivantes sont introduites :

- $A_{n,i}$, le nombre d'erreurs du sujet n lors de la série i dans le silence,
- $B_{n,i,k}$, le nombre d'erreurs du sujet n lors de la série i dans la condition de bruit k .

La grandeur $C_{n,i,k} = A_{n,i} - B_{n,i,k}$ est alors définie comme étant la chute des performances du sujet n lors de la série i dans la condition de bruit k . Dès lors,

$$X_{n,k} = \sum_i C_{n,i,k} \quad (\text{C.1})$$

représente la somme des chute des performances du sujet n sur l'ensemble des séries dans la condition de bruit k ; autrement dit, la chute des performances du sujet n sur l'ensemble du test.

Avoir un test de puissance statistique η signifie que l'on ne se laisse qu'une probabilité $1 - \eta$ de ne pas voir une différence pourtant réellement présente entre les performances moyennes observées dans deux conditions de bruit données. Soient deux conditions de bruit, k_1 et k_2 , telles que les chutes des performances des sujets dans celles-ci soient notées $X_{n,1}$ et $X_{n,2}$.

La relation $D = \overline{X_2} - \overline{X_1}$ présente la différence entre les performances moyennes dans les conditions de bruit k_1 et k_2 . Son écart-type sera noté σ_D .

Première hypothèse : $X_{n,1}$ et $X_{n,2}$ sont de distribution normale suivant l'échantillon choisi. C'est ainsi également le cas pour D . Cette hypothèse n'est pas très contraignante : la procédure est robuste vis-à-vis de celle-ci.

Pour avoir un test de puissance statistique η , il faut vérifier cette équation :

$$u_\eta = \frac{D}{\sigma_D} - u_{1-\frac{\alpha}{2}}, \quad (\text{C.2})$$

où u est la fonction réciproque inverse (fonction quantile) de la loi normale et α le niveau de confiance de l'estimation de D (généralement choisi à 5%). Il faut alors une estimation de D et de σ_D .

Pour obtenir une estimation de D , le modèle de $\overline{X_k}$ proposé par Hongisto (2005), correspondant à l'équation (2.19), est adapté :

$$\overline{X_{STI}} = \frac{\overline{X_{STI=0}} - \overline{X_{STI=1}}}{1 + e^{\frac{STI-0,4}{0,06}}} + \overline{X_{STI=1}}. \quad (\text{C.3})$$

La littérature indique que $\overline{X_{STI=0}} = 0$: les moyennes des performances des sujets sont les mêmes dans le silence ou dans un bruit parfaitement inintelligible. De plus, suivant la définition (C.1), la chute maximale des performances peut s'écrire ainsi :

$$\begin{aligned}\overline{X_{STI=1}} &= \frac{1}{N} \sum_n X_{n,STI=1} \\ &= \frac{1}{N} \sum_n \sum_i C_{n,i,STI=1} \\ &= \sum_i \overline{C_{i,STI=1}}.\end{aligned}\tag{C.4}$$

Deuxième hypothèse : suivant i , les variables $C_{n,i}$ suivent toutes la même loi. En d'autres termes, un sujet donné se comporte de la même façon devant l'ensemble des séries qui lui sont présentées. Cette hypothèse n'implique pas pour autant que les résultats des séries sont corrélés ; cette précision aura de l'importance par la suite.

Dans le cadre de cette hypothèse, il est possible d'écrire :

$$\overline{X_{STI=1}} = I \overline{C_{STI=1}},\tag{C.5}$$

ce qui transforme l'équation (C.3) en :

$$\overline{X_{STI}} = I \overline{C_{STI=1}} \left(1 - \frac{1}{1 + e^{\frac{STI-0,4}{0,06}}} \right).\tag{C.6}$$

La connaissance de des valeurs de STI correspondant aux conditions de bruit k_1 et k_2 permet alors d'obtenir une valeur de D .

Pour connaître la variance de D vis-à-vis des échantillons de population choisis, il faut accéder à la variance des $\overline{X_k}$. D'ores et déjà, il est possible d'écrire :

$$Var \overline{X_k} = \frac{1}{N} Var X_{n,k} = \frac{1}{N} Var \sum_i C_{n,i,k}.\tag{C.7}$$

Troisième hypothèse : suivant i , les variables $C_{n,i}$ ne sont pas corrélées. Il faut donc voir les performances observées pour un sujet n donné sur les séries présentées comme un tirage aléatoire de I échantillons d'une population pour laquelle la valeur de la variable $C_{n,i}$ suit, d'après la deuxième hypothèse, une loi donnée. Il en résulte que :

$$Var \sum_i C_{n,i,k} = \sum_i Var C_{n,i,k}.\tag{C.8}$$

La deuxième hypothèse permettait déjà d'écrire :

$$\sum_i Var C_{n,i,k} = I \sigma_{C,k}^2.\tag{C.9}$$

Il apparaît alors que :

$$\text{Var } X_{n,k} = \text{Var} \sum_i C_{n,i,k} = I \sigma_{c,k}^2. \quad (\text{C.10})$$

L'expression (C.7) s'écrit alors comme suit :

$$\text{Var } \overline{X}_k = \frac{I}{N} \sigma_{c,k}^2. \quad (\text{C.11})$$

Quatrième hypothèse : les \overline{X}_k sont indépendants. Ceci signifie qu'il n'existe pas de corrélation entre les décrets de performance moyens des sujets entre deux conditions sonores données. Il est alors possible d'écrire :

$$\sigma_D^2 = \text{Var } \overline{X}_1 + \text{Var } \overline{X}_2. \quad (\text{C.12})$$

Cette hypothèse est contestable : peut-on s'assurer de cette indépendance alors même que, d'une part, le même groupe de sujets est confronté aux deux conditions et que, d'autre part, les \overline{X}_k correspondant à deux conditions différenciées par leur STI sont *a priori* liés par l'équation (2.19) proposée par Hongisto (2005) ? Il faut cependant remarquer que sans cette hypothèse, la variance de D s'écrirait comme suit :

$$\sigma_D^2 = \text{Var } \overline{X}_1 + \text{Var } \overline{X}_2 - 2 \frac{\text{Var } X_{n,1} \text{Var } X_{n,2}}{N} \text{corr}(X_{n,1}, X_{n,2}). \quad (\text{C.13})$$

L'ajout de ce dernier terme tend à diminuer la valeur de σ_D lorsqu'une corrélation positive existe entre les deux variables sommées : c'est, selon toute vraisemblance, le cas ici. Aussi la quatrième hypothèse permet-elle d'obtenir une sur estimation de la variance de D et, par suite, une approximation par défaut de la puissance du test tout en simplifiant le calcul. Il est donc acceptable de s'en contenter. Par suite, en combinant les expressions (C.11) et (C.12), il vient :

$$\sigma_D^2 = \frac{I}{N} (\sigma_{c,1}^2 + \sigma_{c,2}^2). \quad (\text{C.14})$$

Cinquième hypothèse : la variance de $\overline{B}_{n,k}$ est constante suivant k . Il en découle que :

$$\sigma_{c,1}^2 = \sigma_{c,2}^2 = \sigma_c^2 \quad (\text{C.15})$$

et donc, suivant l'équation (C.14), que

$$\sigma_D^2 = \frac{2 I \sigma_c^2}{N}. \quad (\text{C.16})$$

Ici encore, il est possible de contester la validité de l'hypothèse : les résultats fournis par Ellermeier et Zimmer (1997) montrent que la variance de $\overline{B}_{n,k}$ semble plus grande dans un bruit de parole parfaitement intelligible que dans un bruit de fond parfaitement inintelligible. En supposant que la variation de cette variance suivant la valeur de STI est monotone, le fait de prendre la valeur correspondant à un STI égal à 1 permet de surestimer la variance de D et, là encore, de sous-estimer la puissance du test. Cette simplification est donc, là aussi, acceptable.

D'après Ellermeier et Zimmer, pour 20 séries, $Var X_{n,k} = 340$. Il résulte alors de l'expression (C.10) que :

$$\sigma_C^2 = \frac{340}{20} = 17. \quad (C.17)$$

Par suite, suivant l'équation (C.15),

$$\sigma_D^2 = 34 \frac{I}{N} \quad (C.18)$$

ou encore,

$$\sigma_D = 5.8 \sqrt{\frac{I}{N}}. \quad (C.19)$$

D'après le modèle de Hongisto, les plus fortes variations du décrement de performance sont à attendre dans une plage de valeurs de STI comprises entre 0,3 et 0,5. L'équation (C.6) montre que si les conditions de bruit 1 et 2 correspondent respectivement à des valeurs de STI égales à 0.35 et 0.45, la différence entre les performances dans les deux conditions vaut :

$$D = 0,39 \times I \overline{C_{STI=1}}. \quad (C.20)$$

D'après Ellermeier et Zimmer, pour 20 séries, $\overline{X_{STI=1}} = 24$. On en déduit alors que :

$$\overline{C_{STI=1}} = \frac{24}{20} = 1,2. \quad (C.21)$$

Par suite,

$$D = 0,47 I. \quad (C.22)$$

Dés lors :

$$\frac{D}{\sigma_D} = 0,08 \sqrt{NI}. \quad (C.23)$$

Pour obtenir un test de puissance statistique égale à 80%, c'est-à-dire ne se laisser que 20% de chances de ne pas voir un effet qui existe pourtant réellement, il faut donc vérifier l'égalité suivante :

$$0,84 = 0,08 \sqrt{NI} - 1,96. \quad (C.24)$$

C'est le cas pour $\sqrt{NI} = 35$, c'est-à-dire pour $NI = 1225$. Ainsi, pour 20 séries effectuées, il faut 62 sujets pour obtenir un test de puissance statistique égale, *a priori*, à 80 %.

FOLIO ADMINISTRATIF

THESE SOUTENUE DEVANT L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE LYON

NOM : Ebissou

DATE de SOUTENANCE :

Prénoms : Ange Corneille

TITRE : Gêne sonore en bureaux ouverts : influence de la présence de plusieurs voix intelligibles

NATURE : Doctorat

Numéro d'ordre : AAAAISALXXXX

Ecole doctorale : MEGA – Mécanique, Energétique, Génie Civil, Acoustique

Spécialité : Acoustique

RESUME :

En bureaux ouverts, le bruit de parole des autres occupants représente un facteur de gêne important. L'objectif de ce travail est d'évaluer les outils de prédiction du pouvoir de nuisance de la parole et de proposer des voies d'amélioration. Ceci nécessite de détailler l'influence des caractéristiques d'un bruit de parole qui peuvent rendre le travail plus difficile. Pour ce faire, plusieurs expériences basées sur des exercices mentaux classiques sont menées. La surcharge cognitive engendrée par le bruit de parole ambiant y est mesurée de façon objective et subjective. Dans un premier temps, l'étude est centrée sur l'influence de l'intelligibilité de la parole. En effet, certains auteurs proposent d'utiliser un indicateur d'intelligibilité, le *Speech Transmission Index* (STI), pour prédire la nuisance dans un bureau ouvert. Les résultats obtenus confirment cette influence de l'intelligibilité. Cependant, les bruits de fond rencontrés en bureaux ouverts peuvent être fortement fluctuants. Ceci a un effet sur la nuisance dont le STI ne peut rendre compte. Dans un second temps, une alternative permettant de rendre compte de la variabilité temporelle de réelles ambiances sonores de bureaux est donc proposée. Le pouvoir prédictif de cet indicateur sur la perturbation au travail est évalué pour plusieurs types de tâches. Si les effets observés sont faibles, en particulier au regard des différences de sensibilité au bruit entre individus, la prise en compte des fluctuations de niveau pour un environnement sonore donné semble bel et bien essentielle à la prédiction chiffrée de la gêne sonore causée. De plus, l'évaluation subjective de la gêne s'avère être un complément essentiel à l'évaluation objective pour rendre compte de l'influence de certains facteurs.

MOTS-CLES : Acoustique, Bruit, Gêne sonore, Parole, Intelligibilité, Tertiaire

Laboratoire (s) de recherche : LVA (Laboratoire Vibrations Acoustique), INSA de Lyon
RBT (Réduction du Bruit au Travail), INRS, Vandoeuvre-lès-Nancy

Directeur de thèse : M. Etienne Parizet

Président de jury :

Composition du jury : M. Didier Baptiste
M. Michel Bérengier
M. Patrick Chevret
M. Wolfgang Ellermeier
Mme Catherine Lavandier