

# canadian acoustics

# acoustique canadienne

Journal of the Canadian Acoustical Association - Journal de l'Association Canadienne d'Acoustique

JUNE 2013

JUIN 2013

Volume 41 -- Number 2

Volume 41 -- Numéro 2

GUEST EDITORIAL / EDITORIAL INVITÉ

1

TECHNICAL ARTICLES AND NOTES / ARTICLES ET NOTES TECHNIQUES

UN APERÇU DE L'EFFET DU TYPE DE STIMULUS, DE L'ÂGE ET DE LA DOMINANCE DE L'OREILLE SUR L'INTÉGRATION BINAURALE

*Caroline St-Germain, David Mayer, Marie-Ève Carrier, Élyse Joannette, Gwendoline Chénard et Benoît Jutras* 3

EVALUATION DE L'ÉCOUTE DICHOTIQUE CHEZ L'ENFANT DYSLEXIQUE

*Hung Thai-Van & Evelyne Veuillet* 13

EFFECTS OF INTRA-TALKER DIFFERENCES ON SPEECH UNDERSTANDING IN NOISE BY YOUNGER AND OLDER ADULTS

*Huiwen Goy, M. Kathleen Pichora-Fuller and Pascal van Lieshout* 23

OBJETS SONORES: UNE REPRÉSENTATION BIO-INSPIRÉE, HIÉRARCHIQUE, PARCIMONIEUSE À TRÈS GRANDES DIMENSIONS UTILISABLE EN RECONNAISSANCE

*Simon Brodeur et Jean Rouat* 33

EFFET DU TYPE DE BRUIT ET DE LA SÉPARATION SPATIALE SUR LA RECONNAISSANCE DE LA PAROLE MESURÉE À L'AIDE DU HINT

*Chantal Laroche, Jean-Grégoire Roveda, Christian Giguère, Victoria Milloy, Véronique Vaillancourt* 49

MISE À JOUR SUR LE DÉVELOPPEMENT DU TEST DE MOTS DANS LE BRUIT

*Josée Lagacé, Laudia LeBlanc, Véronique Boisvert, Marika Joëlle Arseneau, Stéphanie Breau-Godwin* 65

Other Features / Autres Rubriques

Minutes of CAA Directors' Meeting / Compte rendu de la réunion des directeurs de l'ACA 73

CAA Prizes Announcement / Annonce de Prix 79



# canadian acoustics

THE CANADIAN ACOUSTICAL ASSOCIATION  
P.O. BOX 1351, STATION "F"  
TORONTO, ONTARIO M4Y 2V9

CANADIAN ACOUSTICS publishes refereed articles and news items on all aspects of acoustics and vibration. Articles reporting new research or applications, as well as review or tutorial papers and shorter technical notes are welcomed, in English or in French. Submissions should be sent directly to the Editor-in-Chief. Complete instructions to authors concerning the required camera-ready copy are presented at the end of this issue.

CANADIAN ACOUSTICS is published four times a year - in March, June, September and December. Copyright on articles is held by the author(s), who should be contacted regarding reproduction. Annual subscription: \$40 (student); \$90 (individual, institution); \$400 (sustaining - see back cover). Back issues (when available) may be obtained from the CAA Secretary - price \$25 including postage. Advertisement prices: \$350 (full page - \$1200 for four issues); \$200 (half page - \$700 for four issues); \$150 (quarter page - \$500 for four issues). Contact the Associate Editor (advertising) to place advertisements. Canadian Publication Mail Product Sales Agreement No. 0557188.

# acoustique canadienne

L'ASSOCIATION CANADIENNE D'ACOUSTIQUE  
C.P. 1351, SUCCURSALE "F"  
TORONTO, ONTARIO M4Y 2V9

ACOUSTIQUE CANADIENNE publie des articles arbitrés et des informations sur tous les domaines de l'acoustique et des vibrations. On invite les auteurs à soumettre des manuscrits, rédigés en français ou en anglais, concernant des travaux inédits, des états de question ou des notes techniques. Les soumissions doivent être envoyées au rédacteur en chef. Les instructions pour la présentation des textes sont exposées à la fin de cette publication.

ACOUSTIQUE CANADIENNE est publiée quatre fois par année - en mars, juin, septembre et décembre. Les droits d'auteur d'un article appartiennent à (aux) auteur(s). Toute demande de reproduction doit leur être acheminée. Abonnement annuel: \$40 (étudiant); \$90 (individuel, société); \$400 (soutien - voir la couverture arrière). D'anciens numéros (non-épuisés) peuvent être obtenus du Secrétaire de l'ACA - prix: \$25 (affranchissement inclus). Prix d'annonces publicitaires: \$350 (page pleine - \$1200 pour quatre publications); \$200 (demi page - \$700 pour quatre publications); \$150 (quart de page - \$500 pour quatre publications). Contacter le rédacteur associé (publicité) afin de placer des annonces. Société canadienne des postes - Envois de publications canadiennes - Numéro de convention 0557188.

---

## Editor-in-Chief / Rédacteur en Chef

### Frank A. Russo

Department of Psychology  
Ryerson University  
350 Victoria Street  
Toronto, Ontario M5B 2K3  
Tel: (416) 979-5000; Ext: 2647  
Fax: (416) 979-5273  
E-mail: russo@ryerson.ca

## Editor / Rédacteur

### Josée Lagacé

Programme d'audiologie et d'orthophonie  
École des sciences de la réadaptation  
Université d'Ottawa  
451, chemin Smyth, pièce 3053  
Ottawa, Ontario K1H 8M5  
Tél: (613) 562-5800 # 8668; Fax: (613) 562-5428  
E-mail: jlagace@uottawa.ca

## Advisory Board / Publicité

### Jérémie Voix

École de technologie supérieure, Université de Québec

### Bryan Gick

University of British Columbia

### Ramani Ramakrishnan

Ryerson University

## Advertising / Publicité

### Clair W. Wakefield

Wakefield Acoustics Ltd.  
301 - 2250 Oak Bay Avenue  
Victoria, BC, V8R 1G5  
Tel: (250) 370-9302  
Fax: (250) 370-9309  
clair@wakefieldacoustics.com

Il me fait plaisir de vous présenter ce numéro spécial portant sur la reconnaissance de signaux sonores en présence de bruit. Comme vous serez à même de le constater, les efforts de recherche dans ce domaine ne sont pas au ralenti. Par exemple, une équipe de chercheurs de l'Université de Toronto a étudié l'intelligibilité de la parole produite en présence de différents niveaux et types de bruit. Dans une autre perspective, une équipe de l'Université de Sherbrooke propose un modèle conceptuel bio-inspiré qui devrait permettre la création d'un système présentant un taux de reconnaissance de la parole supérieur à ceux déjà existants lorsqu'utilisé dans des environnements bruyants. Ainsi, de la meilleure compréhension des habiletés de perception de la parole dans le bruit à la proposition de système de reconnaissance automatique de la parole plus efficace en milieux bruyants, force est de constater que les travaux dans le domaine sont en pleine effervescence aux quatre coins du pays, et même de l'autre côté de l'Atlantique!

Vous remarquerez que ce numéro a une portée clinique plus importante que ce qu'on retrouve habituellement dans la revue. Deux équipes de chercheurs s'intéressant au mécanisme d'écoute dichotique ont accepté notre invitation à soumettre un article dans ce numéro spécial. Une équipe de l'Université de Montréal a étudié le développement des habiletés d'écoute dichotique chez les enfants de six à 12 ans et rapporte les résultats obtenus à l'*Épreuve dichotique de chiffres (ÉPC)*. L'ÉPC est une épreuve récemment développée pour mesurer les habiletés d'écoute dichotique auprès de la population francophone du Canada. Ce type d'épreuve fait généralement partie des batteries de tests administrées en clinique auprès des personnes qui consultent en raison de difficultés à comprendre la parole dans le bruit. Pour leur part, des chercheurs du Centre de Recherche en Neurosciences de Lyon (France) rapportent les résultats de leur étude comparant les habiletés d'écoute dichotique d'enfants présentant une dyslexie à celles d'enfants normo-lecteurs, appariés selon l'âge chronologique. Une meilleure compréhension des habiletés d'écoute dichotique selon les différentes populations cliniques et groupes d'âge contribuera à l'avancement des connaissances sur la reconnaissance de signaux sonores en présence de bruit de fond, de même qu'à l'amélioration des pratiques clinique dans ce domaine.

J'aimerais aussi mentionner d'autres travaux de recherche qui pourront contribuer à l'amélioration des pratiques cliniques. Par exemple, l'équipe de chercheurs de l'Université d'Ottawa ayant travaillé à l'adaptation franco-canadienne du *Hearing in Noise Test (HINT)* continue leurs efforts pour maximiser son utilisation. Dans leur article, l'équipe fait part des données normatives pour le HINT lorsqu'administré avec un bruit intermittent de 16 Hz auprès d'adultes. Ce type de bruit est davantage présent dans les conditions acoustiques de tous les jours. De plus, une autre équipe de la même université continue ses travaux portant sur le développement de données normatives pour le *Test de Mots dans le Bruit (TMB)*.

It is with great pleasure that I present this special issue about sound recognition in noise. As you will realize, work in this area is not in the slow gear. For example, one research team from the University of Toronto reports on an investigation on the effects on intelligibility when listening to a talker who produced speech while listening to different types and levels of noise. In another project, a research team from the Université de Sherbrooke has proposed a bio-inspired conceptual model for the development of an automatic speech recognition system that should allow superior speech recognition performance compared to existing systems when used in noisy environments. Hence, from the perspective of better comprehension to improvements in automatic recognition systems, one has to acknowledge that a "buzz" exists in this area of research and this, in all parts of the country as well as on the other side of the Atlantic!

You will notice that the current issue includes more clinical articles than usual. Two research teams working on dichotic listening have accepted our invitation for this special issue. One team from the Université de Montréal presents a study on the developmental aspect of dichotic listening on six- to 12-year old children using the *Épreuve dichotique de chiffres (ÉPC – Dichotic Digit listening test)*. The ÉPC has recently been developed to measure dichotic listening skills in French Canadian populations. This type of test is generally included in a clinical test battery when an individual presents with speech perception in noise problems. Researchers from the Centre de Recherche en Neurosciences de Lyon (France) report on the results of their study comparing children with dyslexia on a dichotic task to age-matched children with normally developing reading abilities. A better understanding of the dichotic listening abilities among different age groups and clinical populations should contribute to advancing knowledge about acoustic signal perception in noise, as well as to improvements in clinical practice.

I would also like to make note of several other projects described in this issue that may further contribute to better clinical practice. For example, the team from the University of Ottawa has developed an adaptation of the *Hearing in Noise Test (HINT)* for French Canadians; they continue their work to maximize its potential use. In their article, they report adult normative data for the HINT test when used with an intermittent noise of 16 Hz. Intermittent noise is quite frequent in everyday situations. Another team from the same University has continued their work toward the development of normative data for the *Test de Mots dans le Bruit (TMB)*. They also report new data about the possible effect of the cultural linguistic variations on performance measured with the test. Audiologists working with Canadian French communities will soon have access to a variety of standardised tools to measure speech listening skills in noise.

I would like to thank the authors who contributed to what I consider 'my first issue', as well as the reviewers who helped

Cette équipe rapporte aussi leurs données quant à l'effet possible des variantes linguistiques culturelles sur les performances à cette épreuve. Les audiologistes qui travaillent auprès des communautés francophones du Canada auront bientôt accès à une variété d'outils cliniques standardisés pour mesurer les habiletés de perception de la parole dans le bruit.

Je tiens à remercier les auteurs qui ont contribué à ce que j'appelle 'mon premier numéro', de même que les réviseurs qui m'ont aidée dans le processus d'évaluation des articles soumis. Je veux aussi remercier Chantal Laroche pour ses précieux conseils en me passant la barre comme coéditrice de la revue, de même que mes collègues de l'équipe de rédaction pour leurs réponses rapides et efficaces à mes nombreux courriels.

Il me fait plaisir d'entreprendre cette fonction en tant que coéditrice. Je suis professeure adjointe à l'Université d'Ottawa depuis juillet 2009. Mes intérêts de recherche portent sur les difficultés de perception de la parole en milieu bruyant, plus particulièrement chez les enfants. Vous comprenez pourquoi le thème de ce numéro spécial me tenait à cœur.

Bonne lecture de ce numéro à saveur francophone et bel été à tous.

Josée Lagacé  
Co-rédactrice

me in the evaluation process. I also want to thank Chantal Laroche for her valued advice when she passed me the torch as the co-editor of the review, as well as my colleagues at the advisory board for their quick and efficient responses to my numerous emails.

It is a pleasure for me to uptake this role as a co-editor. Since July 2009, I am an Assistant Professor at the University of Ottawa. My research focuses on speech perception in noise, particularly in listeners who are children. You can therefore appreciate why the theme of this special issue is so important to me.

I hope you will enjoy reading this special issue with a French twist. I wish you all a wonderful summer.

Josée Lagacé  
Co-editor

## WHAT'S NEW in Canada ??

*Promotions  
Deaths  
New jobs  
Moves*

*Retirements  
Degrees awarded  
Distinctions  
Other news*

*Do you have any news that you would like to share with Canadian Acoustics readers? If so, send it to:*

## QUOI DE NEUF en Canada??

*Promotions  
Décès  
Offre d'emploi  
Déménagements*

*Retraites  
Obtention de diplômes  
Distinctions  
Autres nouvelles*

*Avez-vous des nouvelles que vous aimeriez partager avec les lecteurs de l'Acoustique Canadienne? Si oui, écrivez-les et envoyez à:*

**Jérémy Voix - Email: [voix@caa-aca.ca](mailto:voix@caa-aca.ca)**

# UN APERÇU DE L'EFFET DU TYPE DE STIMULUS, DE L'ÂGE ET DE LA DOMINANCE DE L'OREILLE SUR L'INTÉGRATION BINAURALE

Caroline St-Germain, David Mayer, Marie-Ève Carrier, Élyse Joannette, Gwendoline Chénard et Benoît Jutras  
École d'orthophonie et d'audiologie, Université de Montréal, 7077 avenue du Parc, Montréal, Québec H3N 1X7

## SOMMAIRE

**Objectifs :** Cette étude a pour but d'évaluer l'effet du type de stimulus, de l'âge et de la dominance de l'oreille sur l'intégration binaurale d'enfants d'âge scolaire au moyen de tests d'écoute dichotique en français. **Méthodologie :** Trente-huit enfants francophones de six, huit, neuf et douze ans, dont 28 provenaient de l'étude de Jutras & coll. (2012) présentant une acuité auditive normale ont participé à cette étude. Ils ont complété quatre versions d'un test d'écoute dichotique de chiffres (EDC) nouvellement créé en français, présentées sous écouteurs intra-auriculaires à niveau normal de conversation. Les trois premières versions étaient respectivement constituées d'une, deux et trois paire-s différente-s de chiffres compris entre un et neuf. La quatrième version du test consistait en la présentation de quatre chiffres, dont seuls les deuxième et troisième chiffres étaient entendus simultanément. Les participants ont également été soumis à un test d'écoute dichotique de mots ayant la même structure que la quatrième version du test EDC, soit l'adaptation en français du test *Staggered Spondaic Word* (Rudmin & Normandin, 1983). **Résultats :** Les données des 28 enfants de l'étude de Jutras & coll. (2012) ont été incluses dans les analyses en plus de celles des dix autres enfants retenus dans la présente étude. Les performances sont significativement supérieures au test d'écoute dichotique employant des chiffres plutôt que des mots pour l'ensemble des groupes. Les résultats montrent aussi que les performances s'améliorent de façon significative selon l'âge. De façon générale, les performances des enfants sont significativement supérieures lorsque les mots ou chiffres sont présentés à l'oreille droite comparativement aux performances obtenues lorsque les stimuli sont envoyés à l'oreille gauche. **Conclusion :** Le test d'écoute dichotique de chiffres en français est sensible à l'influence de l'utilisation de mots familiers et complexes dans les tests d'écoute dichotique, entre autres chez les enfants de huit ans. Le test est également sensible à l'effet de l'âge des participants qui y sont soumis ainsi qu'à un biais relatif à l'oreille testée.

## ABSTRACT

**Objective:** The aim of the present study is to assess stimulus type, age and ear dominance effect on the binaural integration capacity of school-aged children using dichotic listening tests. **Methods:** Thirty-eight French speaking children with normal hearing sensitivity of six, eight, nine and twelve years old (28 participated in Jutras & al., 2012' study) performed the four subtests of the newly developed test in French – *Écoute dichotique de chiffres (EDC)*, under earphones at normal conversational level. Three subtests were composed respectively of one, two and three different digit pairs. The fourth version comprised four different digits, but only the second and third were heard simultaneously. The participants were also assessed with a word dichotic listening test, having the same structure than the EDC fourth version, namely the French adaptation of the *Staggered Spondaic Word* test (Rudmin & Normandin, 1983). **Results:** Data from the 28 participants in Jutras & al. (2012)' study were included in the analysis in addition of the data from the ten children of the present study. Results on the dichotic listening tests were significantly higher with digits than with words and improved as a function of age. Also, in general, participants obtained higher performance when the stimuli were presented in the right ear then when it was sent to the left ear. **Conclusion:** The dichotic digit listening test in French is sensitive to stimulus type and to the age of the participants. Results showed also a bias towards the ear tested.

## 1. INTRODUCTION

Au quotidien, la majorité des personnes, enfants et adultes, se trouvent dans des situations de communication où inévitablement la parole entre en compétition avec le bruit présent dans l'environnement. Le bruit nuit généralement à la compréhension de la parole puisqu'il dissimule en partie

ou entièrement l'information transmise oralement. L'écoute dans le bruit nécessite donc de la part du système auditif de filtrer l'information pertinente parmi d'autres stimuli auditifs non pertinents, entendus de façon simultanée (Stuart, 2005). Dans ce type de situation, les capacités cognitives sont aussi mises à profit. En effet, l'attention

dirigée vers l'information pertinente est nécessaire afin de comprendre ce qui est dit et d'ignorer les stimuli indésirables (Martin, Jerger & Mehta, 2007). Il s'agit de la séparation figure/fond (ASHA, 2005). Toutefois, il existe aussi des situations d'écoute au quotidien où l'information pertinente est entendue simultanément aux deux oreilles, mais dont le contenu est différent dans chaque oreille. Par exemple, cette situation survient lorsqu'une personne écoute un locuteur au téléphone pendant qu'un autre locuteur s'adresse à elle dans la même pièce. Cette personne doit comprendre à la fois ce qui est dit au téléphone et ce qu'elle a entendu du locuteur situé près d'elle. Cette situation est un parfait exemple d'intégration binaurale, définie comme étant la capacité à traiter une information auditive différente soumise simultanément aux deux oreilles (Bellis, 2003). Elle est mesurée par des tests d'écoute dichotique. L'attention divisée, qui consiste en la capacité à percevoir deux stimuli ou plus entendus simultanément, est aussi mise à contribution dans ce processus (Martin & coll., 2007). Une capacité réduite d'intégration binaurale peut conduire à des difficultés d'écoute lorsque plusieurs personnes parlent simultanément (Bellis, 2003). En résumé, l'écoute dans le bruit et l'écoute dichotique sont des situations d'écoute compétitive où il faut percevoir des stimuli cibles parmi d'autres stimuli auditifs. Toutefois, ces situations se distinguent par le fait que l'écoute dichotique exige de percevoir et d'identifier tous les stimuli entendus, alors que l'écoute dans le bruit demande d'ignorer volontairement les signaux jugés indésirables pour identifier l'information pertinente.

Plusieurs tests existent en anglais afin d'évaluer l'intégration binaurale. Les tests d'écoute dichotique construits à l'aide de mots incluent le *Competing Sentences Test* (CST; Willeford & Burleigh, 1994), le *Staggered Spondaic Word Test* (SSW; Katz, Basil, & Smith, 1963), le *Dichotic Consonant-Vowel Test* (Berlin, Hughes, Lowe-Bell, & Berlin, 1973) et le *Dichotic Digit Test* (Musiek, 1983a; Strouse & Wilson, 1999). Ces différents tests d'écoute dichotique s'avèrent sensibles aux lésions corticales, sous-corticales et du corps calleux (Champoux, Paiement, Mercier, Lepore, Lasonde, & Gagné, 2007; Gadea, Arana, Espert, Salvador, & Casanova, 2009; Hannay, Walker, Dennis, Kramer, Blaser, & Fletcher, 2008; Jacobson, Deppe, & Murray, 1983; Musiek, 1983b; Musiek, Gollegly, Kibbe, & Verkest-Lenz, 1991; Musiek, Reeves, & Baran, 1985). L'audiologie clinique au Canada fait communément usage d'un test d'écoute dichotique, soit l'adaptation en français du test *SSW* (Rudmin & Normandin, 1983).

Le *SSW* contient 40 séries de quatre différents mots monosyllabiques. Chaque série est présentée au niveau normal de conversation selon quatre conditions d'écoute : (a) un mot à l'oreille droite uniquement (condition droite non compétitive, DNC); (b) un mot à l'oreille droite (condition droite compétitive, DC) simultanément avec (c) un mot à l'oreille gauche (condition gauche compétitive, GC); et (d) un mot à l'oreille gauche uniquement (condition gauche non compétitive, GNC). La personne évaluée doit répéter chacun des mots entendus selon l'ordre de

présentation. Les données normalisées par Bérard (1990-1993) pour la version adaptée en français du *SSW* ont été obtenues auprès de 68 enfants (38 filles, 30 garçons), divisés selon l'âge en cinq groupes, soit de 6 à 11 ans. Sensibles à la maturation, les données sous-tendent une amélioration des performances au test avec l'âge, tendance également remarquée selon les normes du *SSW* établies par Katz (1992).

Les performances aux tests d'écoute dichotique peuvent être parfois biaisées en soumettant des enfants à des épreuves contenant des mots peu fréquemment utilisés dans la langue, ceux étant peu familiers et complexes au plan linguistique (Techentin & Voyer, 2011). Plusieurs mots utilisés dans les tests, notamment dans le *SSW*, demeurent non familiers pour bon nombre d'enfants ou paraissent plus complexes et nuisent à leur identification. Ce contexte d'évaluation peut conduire à une interprétation erronée, particulièrement chez les enfants jeunes ou ceux aux prises avec des troubles langagiers. Dans le but de réduire l'influence de tels facteurs sur les performances des enfants aux différents tests d'écoute dichotique, un nouveau test utilisant des chiffres plutôt que des mots a été développé en français : l'écoute dichotique de chiffres, ou EDC (Jutras & Mayer, 2004). Ce test compte quatre versions dont les trois premières comprennent respectivement une, deux et trois paires de chiffres entre un et neuf. La quatrième version du test est constituée de quatre chiffres, dont seuls les deuxième et troisième chiffres sont entendus simultanément. Tous les chiffres présentés en écoute dichotique sont différents d'une oreille à l'autre. Selon l'étude de Jutras & coll. (2012), le type de stimulus utilisé dans les tests d'écoute dichotique influence les performances des enfants de six, neuf et douze ans. Ces derniers obtiennent des résultats significativement supérieurs à l'EDC qu'au *SSW*, c'est-à-dire lorsque les stimuli entendus sont des chiffres plutôt que des mots. De plus, les résultats de l'étude ont montré, entre autres, que les données du groupe d'enfants de six ans étaient significativement inférieures à celles du groupe d'enfants de neuf ans, suggérant un effet de développement sur la capacité d'intégration binaurale de chiffres. La présente étude examine la trajectoire développementale de cette capacité auditive en soumettant des enfants de huit ans aux tests d'écoute dichotique de mots et de chiffres et en comparant leurs résultats à ceux d'enfants de six, neuf et douze ans, provenant majoritairement de l'étude de Jutras & coll. (2012). L'étude vise également à situer les performances des enfants de huit ans par rapport à celles des enfants appartenant aux groupes d'âge susmentionnés aux quatre versions du test EDC. Ce groupe d'âge a été considéré puisque Jutras & coll. (2012) rapportaient des différences significatives entre les résultats des enfants de six ans et ceux de neuf ans, d'où l'intérêt d'examiner le comportement des enfants de huit ans par rapport à ceux-ci.

L'écoute dichotique de stimuli verbaux entraîne par ailleurs un phénomène non négligeable remarqué dans l'analyse des résultats aux différents tests. La majorité des individus soumis à ces tests obtiennent des performances supérieures lorsque les stimuli sont présentés à l'oreille

droite comparativement à la condition où ils le sont à l'oreille gauche. Ce phénomène est appelé la dominance de l'oreille droite, ou DOD (Berlin & coll., 1973; Kimura, 1961a, 1961b; Martin & coll., 2007; Moncrieff, Jerger, Wambaq, Greenwald, & Black, 2004). La DOD peut s'expliquer entre autres par la notion de dominance hémisphérique, le langage (traitement des stimuli verbaux) étant, pour la majorité de la population, latéralisé dans l'hémisphère gauche (Meyer, 1957; Milner, 1958). Selon Kimura (1961a), la DOD est également associée à l'avantage de la voie controlatérale sur la voie ipsilatérale dans les divers relais du système auditif central, qui serait plus efficace dans la transmission de l'information auditive. Les voies controlatérales seraient dominantes et plus nombreuses que les voies ipsilatérales et, en situation d'écoute dichotique, ces dernières sont réprimées par les voies controlatérales, d'où la dominance de l'oreille droite (Kimura, 1961a). Lors de tâches d'écoute dichotique de chiffres, une DOD est présente dans les performances d'enfants de quatre à neuf ans, mais elle tend à s'amenuiser avec l'âge (Kimura, 1963). Certaines contradictions subsistent dans les études, puisque ce phénomène n'est pas noté chez d'autres groupes d'enfants lorsque les stimuli utilisés constituent des syllabes sans sens (Berlin & coll., 1973). La DOD demeure équivalente pour l'ensemble des enfants âgés de cinq à treize ans (Berlin & coll., 1973). Plus récemment, le Dichotic Digit Test en malaisien et constitué de paires simples et doubles de chiffres, a été administré à 120 enfants malaisiens âgés entre six et onze ans (Mukari, Keith, Tharpe, & Johnson, 2006). Les résultats démontrent une DOD au moins trois fois plus importante chez les enfants de six à huit ans comparativement aux enfants de dix et onze ans pour les paires doubles de chiffres que pour les paires simples (Mukari & coll., 2006). Les résultats de l'étude de Jutras & coll. (2012) montrent une variabilité en ce qui a trait à la prévalence de la DOD. Elle varie selon les versions de l'EDC. Ainsi, la DOD est moins importante pour la condition de test la plus difficile – *Paires triples*, que pour les trois autres versions du test – *Paires simples*, *Paires doubles* et *Paires ressemblant à la structure du test SSW*, lesquelles impliquent dans une moindre mesure l'attention et la mémoire. Le troisième objectif de cette étude était donc d'évaluer la différence entre les performances obtenues à chaque oreille au test EDC en tenant compte de l'effet de l'âge, en situant plus spécifiquement les performances d'enfants de huit ans dans la trajectoire développementale déjà dessinée par les résultats d'enfants de six, neuf et douze ans de l'étude de Jutras & coll. (2012). Une DOD supérieure à la normale en fonction de l'âge peut constituer un indice d'un trouble de traitement auditif (Moncrieff & coll., 2004).

## 2. MÉTHODOLOGIE

Le comité scientifique du comité d'éthique du CHU Sainte-Justine a approuvé le projet de recherche, tandis que les membres du comité d'éthique du CHU Sainte-Justine ont sanctionné le formulaire d'information et de consentement. Le consentement écrit du parent de même que l'assentiment

de l'enfant ont été obtenus avant de procéder à l'expérimentation.

### 2.1 Participants

La présente étude compte des enfants ayant été recrutés en deux temps. Dans un premier temps, 28 enfants avec une acuité auditive et une capacité d'intégration binaurale normales ont participé à l'étude de Jutras & coll. (2012). Ils avaient été divisés en trois groupes : (a) neuf âgés de six ans (moyenne d'âge = 6;6 [années; mois]; étendue = 6;3 à 6;11; cinq garçons et quatre filles), (b) neuf âgés de 9 ans (moyenne d'âge = 9;3; étendue = 9;0 à 9;11; cinq garçons et quatre filles), et (c) dix âgés de douze ans (moyenne d'âge = 12;4; étendue = 12;0 à 12;11; cinq garçons et cinq filles). Dans un deuxième temps, dix enfants âgés entre 8 ans 1 mois et 8 ans 9 mois (moyenne d'âge de 8 ans 4 mois) ont été recrutés pour participer à l'étude, dont cinq garçons et cinq filles. Les critères d'inclusion et d'exclusion spécifiaient que l'enfant devait fréquenter une école francophone, présenter des résultats scolaires dans la moyenne ou supérieurs à celle-ci, parler le français à la maison, réussir le test de dépistage auditif (seuils auditifs équivalents ou inférieurs à 15 dB HL entre 250 et 8 000 Hz aux deux oreilles) et ne pas avoir eu une otite au moins un mois précédent l'expérimentation. De plus, à l'adaptation en français du test *Staggered Spondaic Word (SSW)* (Rudmin & Normandin, 1983), les participants devaient obtenir des résultats se situant à l'intérieur de deux écarts-types de la moyenne selon les normes de Bérard (1990-1993), sans quoi ils étaient exclus de l'étude.

En raison de difficultés scolaires soupçonnées et de complications survenues lors de la réalisation des différents tests de l'expérimentation, un enfant a dû être exclu de l'étude. Au total, quatre garçons et quatre filles de huit ans ont été retenus pour l'étude, dont la moyenne d'âge se situait à 8 ans 4 mois (étendue : 8 ans 1 mois à 8 ans 9 mois). De plus, nous avons ajouté dans l'échantillon un participant de six ans et un de neuf ans qui avaient été initialement évalués lors de l'expérimentation de l'étude de Jutras & coll. (2012). Ces deux participants ont obtenu un résultat au test SSW compris entre deux et trois écarts-types de la moyenne dans la condition d'écoute gauche compétitive. Cependant, puisqu'ils remplissaient les critères pour les autres conditions d'écoute au SSW, ces deux enfants ont été inclus dans la présente expérimentation.

### 2.2 Stimuli

Les chiffres de un à neuf utilisés dans cette étude ont été prononcés par une voix masculine (Jutras & Mayer, 2004). L'enregistrement de ces stimuli a été réalisé à l'aide du logiciel SoundEdit (1999) utilisant un microphone Labtec (Verse 704) connecté à un ordinateur Mac Power Book G4. La durée de chacun des stimuli présentés variait de 311 à 403 ms. Dans le but de calibrer l'intensité des chiffres entendus afin que celle-ci soit équivalente pour chacun d'eux, le niveau de pression sonore a pu être mesuré par l'envoi de ces stimuli depuis un ordinateur Mac Power Book G4 à un audiomètre Madsen Orbiter 902, puis via un

écouteur supra-auriculaire TDH-39 placé sur une oreille artificielle Larson-Davis AE 100, connectée à un sonomètre Larson-Davis 800 B.

### 2.3 Matériel

L'EDC se décompose en quatre versions (voir Tableau 1). La première version consiste à présenter un chiffre différent simultanément à chaque oreille, l'un entendu du côté gauche et l'autre du côté droit (P1) – *Paires simples*. Les versions deux et trois de ce test consistent respectivement à présenter de façon simultanée aux deux oreilles deux séries de deux chiffres (P1, P2) – *Paires doubles* et de trois chiffres (P1, P2, P3) – *Paires triples*. La quatrième version prend appui sur le modèle du *SSW – Modèle SSW* et comporte quatre chiffres différents : un chiffre est envoyé à une oreille seulement (condition non compétitive pour l'oreille stimulée), deux autres chiffres sont par la suite entendus simultanément aux deux oreilles (P1) (condition compétitive à droite et à gauche), puis le dernier chiffre est envoyé à l'autre oreille seulement (condition non compétitive). La présentation de chaque série de quatre chiffres est alternée chaque fois entre l'oreille droite qui débute et l'oreille gauche qui débute. Enfin, pour les versions deux, trois et quatre de l'EDC, un intervalle de silence de 500 ms est introduit entre chaque chiffre. Chaque version, composée de vingt groupes de stimuli se trouve également précédée d'une liste de cinq items de familiarisation. Chaque chiffre est considéré comme étant un stimulus à répéter. Le test SSW a été décrit dans la section Introduction.

**Tableau 1.** Illustration du test d'écoute dichotique de chiffres (EDC) par un exemple pour une paire (P1), deux paires (P2) ou trois paires (P3) de chiffres étant en compétition ou non aux deux oreilles (conditions d'écoute droite compétitive – DC, gauche compétitive – GC, droite non compétitive – DNC, gauche non compétitive – GNC).

Version Paires simples			Version Paires doubles	
			DC	GC
	DC	GC	P1	6-----8
P1	2-----1		P2	2-----7
Version Paires triples			Version Modèle SSW	
			DNC	
	DC	GC	9	
P1	8-----5		DC	GC
P2	4-----2		P1	3-----8
P3	3-----9			GNC
				1

### 2.4 Procédure

Une première sélection des enfants s'est effectuée par l'entremise d'entrevues téléphoniques avec les parents. Les renseignements obtenus concernant la langue parlée à la maison et à l'école fréquentée, les performances scolaires et l'histoire passée ou présente d'otites ont permis de déterminer l'admissibilité de l'enfant à l'expérimentation. Le cas échéant, le formulaire d'information et de consentement a brièvement été discuté avec le parent et un rendez-vous a été désigné à l'École d'orthophonie et d'audiologie de l'Université de Montréal ou au CHU Sainte-

Justine. Pour les participants sélectionnés à la suite de cette entrevue, le formulaire de consentement a d'abord été expliqué au parent et à l'enfant avant d'être signé au début de la rencontre. L'enfant devait donner son assentiment. Un test de dépistage auditif à 15 dB HL de 250 à 8 000 Hz, selon la méthode Hughson-Westlake modifiée (Carhart & Jerger, 1959), a été administré aux enfants, incluant l'évaluation du seuil de réception de la parole, afin de s'assurer que chacun présente une acuité auditive dans les limites de la normale. Ces derniers ont par la suite été soumis au test *SSW* (version en français) dans le but de confirmer une bonne habileté d'intégration binaurale. Puis, une étape de familiarisation contenant les neuf chiffres désignés selon un mode aléatoire prédéterminé permettait à l'examineur de s'assurer que chaque enfant était en mesure d'identifier correctement les chiffres de 1 à 9. Les quatre versions ont été administrées de façon randomisée à chaque participant, qui devait répéter l'ensemble des chiffres entendus après chaque série de présentation. Connectés à un audiomètre AC-40 ou Madsen Itera, un lecteur de disques compacts TASCAM CD-A500 ou Panasonic SL-SX330 et des écouteurs intra-auriculaires EAR-Tone 3A ont servi à la transmission des stimuli. Les deux tests (*SSW* et EDC) ont été administrés à un niveau de présentation de 55 dB HL aux deux oreilles, lequel était préalablement calibré avec un son de 1 000 Hz. Variant entre 90 et 120 minutes, l'expérimentation se concluait par la remise d'une compensation financière au parent du participant pour les frais de déplacement occasionnés.

### 2.5 Analyses statistiques

L'analyse des résultats a été effectuée à partir des données de huit enfants de huit ans, vus dans la présente étude, celles des 28 enfants de six, neuf et douze ans de l'étude de Jutras & coll. (2012) ainsi que celles d'un enfant de six ans et d'un enfant de neuf ans évalués lors de l'expérimentation de Jutras & coll. (2012), pour un total de 38 enfants. Pour les versions 1, 2 et 3 du EDC, un point a été accordé à chaque chiffre correctement répété pour chaque oreille. Le nombre de bonnes réponses a été converti en pourcentage pour l'oreille droite et pour l'oreille gauche. Pour la version 4 du EDC et pour le SSW, un point a été alloué à chaque bonne réponse donnée pour chaque chiffre présenté dans les quatre conditions d'écoute. L'accumulation des points a été transformée en pourcentage pour chaque condition. Pour l'ensemble des données issues de diverses analyses statistiques et rapportées dans la présente étude, une transformation arcsinus des données brutes a été réalisée, considérant que les résultats obtenus dans certaines conditions d'écoute ou certaines versions de l'EDC se trouvaient au maximum des performances ou près des valeurs maximales (effet plafond). Cette transformation permet donc d'étaler l'écart se trouvant entre les valeurs extrêmes et de le rendre plus important que celui se rapportant aux valeurs plus près de la moyenne (Studebaker, 1985). De cette façon, les probabilités de détecter la présence de différences significatives et de pouvoir les mesurer sont augmentées. Des ANOVA à mesures répétées



ont également été utilisées pour l'ensemble des analyses effectuées dans le cadre de la présente étude. Lors de la réalisation de tests post-hoc, le facteur de correction de Bonferroni a été appliqué systématiquement aux résultats, avec un niveau de  $p < .008$ .

### 3. RÉSULTATS

#### 3.1 Écoute dichotique de chiffres et de mots

Dans le but de vérifier l'influence du type de stimuli (mots et chiffres) sur les performances des enfants, les données provenant des quatre groupes d'enfants dans les quatre conditions d'écoute du test *SSW* et de la version *Modèle SSW* du test EDC ont été analysées. Une ANOVA à trois facteurs a été réalisée en utilisant les variables condition d'écoute (DNC, DC, GC et GNC – quatre niveaux) et type de stimulus (mot et chiffre – deux niveaux). Les résultats révèlent un effet significatif pour les trois facteurs principaux (voir Figure 1) : les performances des groupes se trouvent généralement supérieures au test d'écoute dichotique de chiffres comparativement à celles obtenues au test d'écoute dichotique de mots [Stimulus,  $F(1, 34) = 26.35, p < .01, \eta^2 = .437$ ]; les performances des groupes se trouvent généralement plus élevées en conditions d'écoute non compétitives que celles retrouvées en conditions d'écoute compétitives [Condition,  $F(3, 102) = 72.00, p < .01, \eta^2 = .679$ ]; les performances entre les groupes d'enfants sont significativement différentes [Groupe,  $F(3, 34) = 8.01, p < .01, \eta^2 = .414$ ]. Une seule interaction double s'est avérée significative (voir Tableau 2 et Figure 1) : Condition x Groupe,  $F(9, 102) = 4.93, p < .01, \eta^2 = .303$ . L'interaction triple, Stimulus x Condition x Groupe, était significative,  $F(9, 102) = 1.98, p < .05, \eta^2 = .149$ .

L'interaction triple, Stimulus x Condition x Groupe, a donc été décomposée. Une ANOVA (Condition x Groupe) a été réalisée pour chacun des deux types de stimulus. Les résultats s'avèrent significatifs pour les mots,  $F(9, 102) = 2.24, p < .05, \eta^2 = .165$ , et pour les chiffres,  $F(9, 102) = 4.39, p < .01, \eta^2 = .279$ . Une analyse de variance univariée a été réalisée pour chacune des quatre conditions du test, en fonction des stimuli. Ainsi, avec les mots, une différence significative entre les groupes a été remarquée pour les conditions : DC, (3, 34) = 6.51,  $p < .01, \eta^2 = .365$ , et GC, (3, 34) = 5.05,  $p < .01, \eta^2 = .308$ , mais pas pour les conditions d'écoute non compétitives : DNC, (3, 34) = .792,  $p > .05, \eta^2 = .065$ , et GNC, (3, 34) = 2.56,  $p > .05, \eta^2 = .184$ . Les données significatives ont ensuite été analysées par l'exécution de tests *t*, afin de comparer les résultats obtenus entre chacun des groupes. Les résultats révèlent que pour les mots, les performances du groupe d'enfants de six ans dans la condition DC,  $t(15) = 4.30, p = .001$ , et du groupe d'enfants de neuf ans dans la condition GC,  $t(16) = 3.74, p < .008$ , sont significativement inférieures à celles du groupe d'enfants de douze ans. Aucune différence significative pour les deux conditions n'est remarquée entre le groupe d'enfants plus jeunes et les groupes de huit ans et de neuf ans, ni entre le groupe d'enfants de huit ans et ceux de neuf ans et de douze ans (voir Tableau 3).

**Tableau 2.** Anova effectuée auprès de 38 participants divisés en quatre groupes (G) avec deux facteurs répétés : quatre conditions d'écoute (C) et deux types de stimulus (S).

	<i>ddl</i> facteur	<i>ddl</i> erreur	<i>F</i>	<i>p</i>	$\eta^2$
Groupe (G)	3	34	8.01	<.001*	.414
Condition (C)	3	102	72.00	<.001*	.679
Stimulus (S)	1	34	26.35	<.001*	.437
G x C	9	102	4.93	<.001*	.303
G x S	3	34	1.07	>.05	.086
C x S	3	102	2.28	>.05	.063
G x C x S	9	102	1.98	<.05*	.149

\*significatif

Pour les chiffres, une différence significative entre les groupes a été remarquée pour les conditions : DC, (3, 34) = 7.87,  $p < .01, \eta^2 = .410$ , et GC, (3, 34) = 6.13,  $p < .01, \eta^2 = .351$ , mais pas pour les conditions d'écoute non compétitives : DNC, (3, 34) = .408,  $p > .05, \eta^2 = .035$ , et GNC, (3, 34) = .060,  $p > .05, \eta^2 = .005$ . Ainsi, des tests *t* ont été réalisés. Les résultats révèlent que dans les deux conditions d'écoute DC et GC, les performances des groupes d'enfants de neuf ans [DC,  $t(17) = 4.63, p < .001$ , GC,  $t(18) = 2.97, p = .008$  – à la limite], et de douze ans [DC,  $t(18) = 3.87, p = .001$ , GC,  $t(18) = 4.29, p < .001$ ] sont significativement supérieures à celles des enfants de six ans. Aucune différence significative pour les deux conditions n'est remarquée entre les deux groupes d'enfants plus jeunes; entre le groupe d'enfants de huit ans et ceux de neuf ans et de douze ans; entre les enfants de neuf ans et ceux de douze ans (voir Tableau 4).

#### 3.2 Écoute dichotique de chiffres en quatre versions

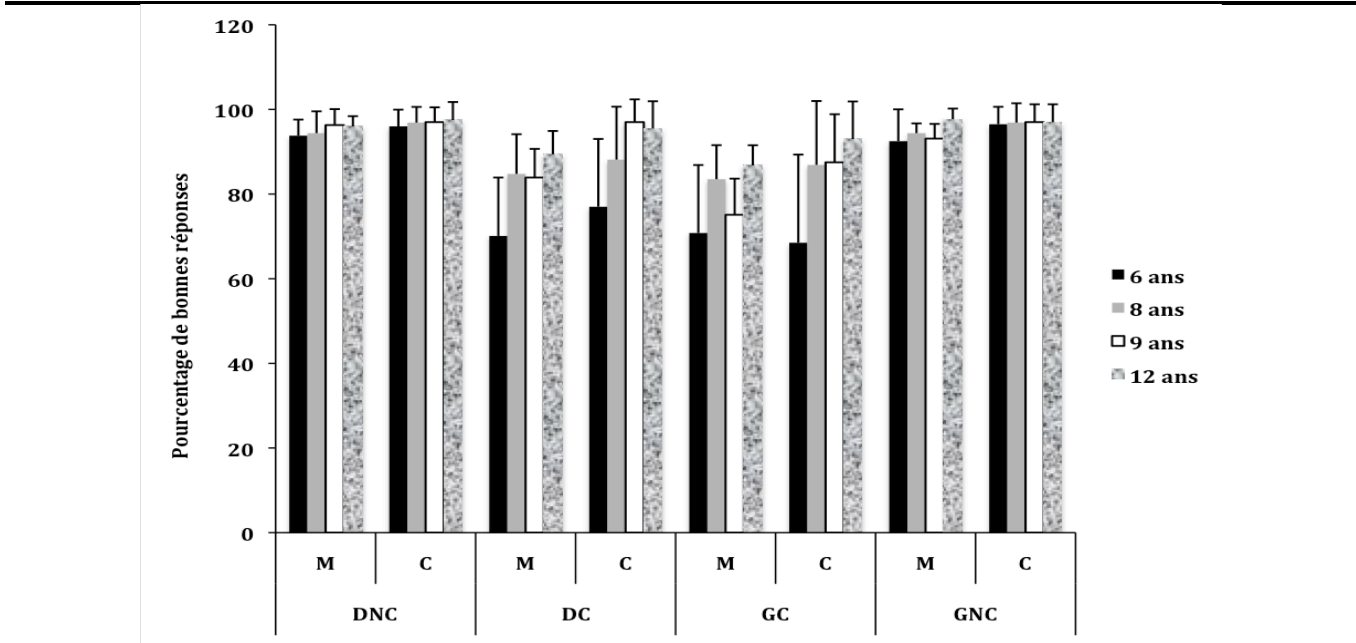
La variable dépendante à l'étude correspondait au pourcentage de bonnes réponses au test EDC. Trois variables indépendantes intra-sujets ont été considérées dans cette analyse : l'oreille stimulée (oreille droite vs oreille gauche – deux niveaux), l'ordre de présentation des stimuli pour chaque sous-test (dix premiers stimuli vs dix derniers stimuli pour vérifier les effets d'apprentissage et de fatigue – deux niveaux) et la version du test présentée (*Paires*

**Tableau 3.** Comparaison des résultats obtenus (tests *t*) entre les quatre groupes de participants – six, huit, neuf et douze ans – selon les conditions d'écoute – droite compétitive (DC) et gauche compétitive (GC) – pour le test d'écoute dichotique de mots.

Condition d'écoute	Comparaison entre les groupes	<i>ddl</i>	<i>t</i>	<i>p</i>
DC	6 ans vs 8 ans	15	2.56	>.008
	6 ans vs 9 ans	17	2.69	>.008
	6 ans vs 12 ans	15	4.30	.001*
	8 ans vs 9 ans	13	0.32	>.05
	8 ans vs 12 ans	11	1.00	>.05
	9 ans vs 12 ans	17	1.74	>.05
GC	6 ans vs 8 ans	15	2.12	>.05
	6 ans vs 9 ans	15	0.56	>.05
	6 ans vs 12 ans	12	3.09	>.008
	8 ans vs 9 ans	15	2.13	>.05
	8 ans vs 12 ans	12	0.94	>.05
	9 ans vs 12 ans	16	3.74	<.008*

\*significatif ( $p < .008$ )

**Figure 1.** Pourcentage de bonnes réponses obtenues auprès de 38 participants divisés en quatre groupes – six, huit, neuf et douze ans – en fonction du type de stimulus – mots (M) et chiffres (C) – selon les quatre conditions des tests d’écoute dichotique – droite non compétitive (DNC), droite compétitive (DC), gauche compétitive (GC) et gauche non compétitive (GNC).



simples vs Paires doubles vs Paires triples vs Modèle SSW – quatre niveaux). La variable indépendante inter-sujet correspond au groupe (enfants de six, huit, neuf et douze ans – quatre niveaux).

Une ANOVA à quatre facteurs a été effectuée. Les résultats révèlent un effet significatif pour trois des quatre facteurs principaux (voir Tableau 5 et Figure 2) : les performances des groupes se trouvent généralement supérieures lorsque les stimuli sont perçus à l’oreille droite plutôt qu’à l’oreille gauche [Oreille,  $F(1, 34) = 13.40, p < .01, \eta^2 = .283$ ]; les performances des groupes pour les versions *Paires simples* et *Modèle SSW* du test sont généralement supérieures à celles pour les versions *Paires doubles* et *Paires triples* [Version,  $F(3, 102) = 225.04, p <$

.01,  $\eta^2 = .869$ ]; et les performances entre les groupes d’enfants sont significativement différentes [Groupe,  $F(3, 34) = 7.53, p < .01, \eta^2 = .399$ ]. Seulement une des six interactions doubles existantes s’est avérée significative (voir Tableau 4) : Version x Groupe,  $F(9, 102) = 4.55, p < .01, \eta^2 = .287$ . L’ANOVA réalisée a également permis de démontrer qu’aucune interaction triple ni quadruple n’est significative (voir Tableau 5 et Figure 2).

Une analyse de variance univariée a été réalisée pour chacune des versions du test, afin de décomposer l’interaction double Version x Groupe. Une différence significative entre les groupes a été remarquée pour les versions suivantes : *Paires doubles*,  $(3, 34) = 8.52, p < .01, \eta^2 = .429$ ; *Paires triples*,  $(3, 34) = 12.18, p < .01, \eta^2 = .518$

**Tableau 4.** Comparaison des résultats obtenus (tests t) entre les quatre groupes de participants – six, huit, neuf et douze ans – selon les conditions d’écoute – droite compétitive (DC) et gauche compétitive (GC) – pour le test d’écoute dichotique de chiffres.

Condition d’écoute	Comparaison entre les groupes	ddl	t	p
DC	6 ans vs 8 ans	14	1.74	>.05
	6 ans vs 9 ans	17	4.63	<.001*
	6 ans vs 12 ans	18	3.87	.001*
	8 ans vs 9 ans	12	2.08	>.05
	8 ans vs 12 ans	13	1.49	>.05
	9 ans vs 12 ans	18	0.71	>.05
GC	6 ans vs 8 ans	13	2.56	>.008
	6 ans vs 9 ans	18	2.97	.008*
	6 ans vs 12 ans	18	4.29	<.001*
	8 ans vs 9 ans	13	0.18	>.05
	8 ans vs 12 ans	14	1.09	>.05
	9 ans vs 12 ans	18	1.54	>.05

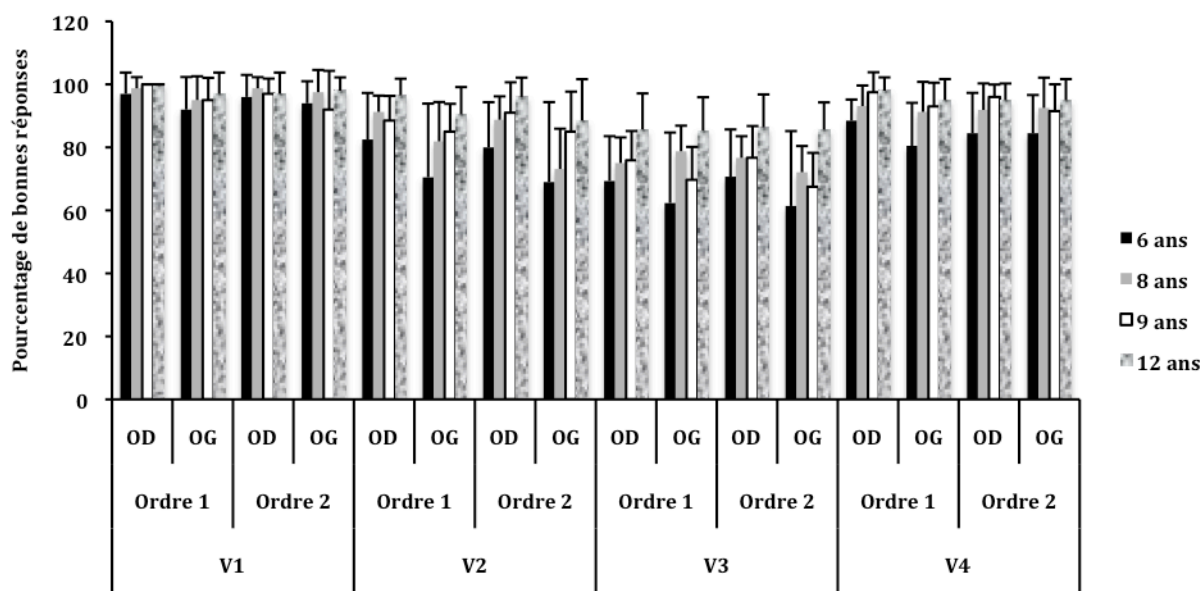
\*significatif ( $p < .008$ )

**Tableau 5.** Anova effectuée auprès de 38 participants divisés en quatre groupes (G) avec trois facteurs répétés : oreille stimulée (A), ordre de présentation des stimuli (O) et version (V) du test EDC.

	ddl facteur	ddl erreur	F	p	$\eta^2$
Oreille (A)	1	34	13.40	.001*	.283
Ordre (O)	1	34	2.79	.104	.076
Version (V)	3	102	225.04	<.001*	.869
Groupe (G)	3	34	7.53	.001*	.399
A x O	1	34	0.82	>.05	.024
A x V	2	77	2.89	>.05	.078
A x G	3	34	0.38	>.05	.032
O x V	3	102	0.17	>.05	.005
O x G	3	34	0.33	>.05	.028
V x G	9	102	4.55	<.001*	.287
A x O x V	3	102	2.64	>.05	.072
A x V x G	9	102	0.65	>.05	.054
A x O x G	3	34	0.52	>.05	.044
O x V x G	9	102	1.50	>.05	.117
A x O x V x G	9	102	0.13	>.05	.011

\*significatif

**Figure 2.** Pourcentage de bonnes réponses obtenues auprès de 38 participants divisés en quatre groupes – six, huit, neuf et douze ans en fonction de l’oreille stimulée – oreille droite (OD) et oreille gauche (OG), de l’ordre de présentation des stimuli – ordre 1 (10 premiers items) et ordre 2 (10 derniers items) – selon la version du test – *Paires simples* (V1), *Paires doubles* (V2), *Paires triples* (V3) et *Modèle SSW* (V4).



et *Modèle SSW*, (3, 34) = 5.59,  $p < .01$ ,  $\eta^2 = .330$ . Toutefois, aucune différence significative n’a été obtenue pour la version *Paires simples*, (3, 34) = 1.59,  $p > .05$ ,  $\eta^2 = .123$ . Les données significatives ont ensuite été analysées par l’exécution de tests t, afin de comparer les résultats obtenus entre chacun des groupes. Les résultats révèlent que pour les trois versions du test, les performances du groupe d’enfants de six ans, [*Paires doubles*,  $t(17) = 4.63$ ,  $p < .001$ , *Paires triples*,  $t(17) = 4.96$ ,  $p < .001$ , *Modèle SSW*,  $t(17) = 4.03$ ,  $p < .008$ ] sont significativement inférieures à celles du groupe d’enfants de douze ans. Aucune différence significative pour les trois versions n’est remarquée entre les deux groupes d’enfants plus jeunes ni entre les groupes d’enfants de huit ans et de neuf ans. Pour la version *Paires triples* du test, les enfants de huit ans,  $t(13) = 3.24$ ,  $p < .008$ , et de neuf ans,  $t(15) = 3.83$ ,  $p < .008$ , obtiennent des performances significativement inférieures à celles des enfants de douze ans, mais aucune différence n’est notée entre les groupes d’enfants de six ans et de neuf ans. Pour les versions *Paires doubles* et *Modèle SSW*, les performances du groupe d’enfants de six ans, [*Paires doubles*,  $t(18) = 3.10$ ,  $p < .008$ , *Modèle SSW*,  $t(18) = 3.48$ ,  $p < .008$ ] sont significativement inférieures à celles du groupe d’enfants de neuf ans, mais aucune différence n’est remarquée entre les enfants de douze ans comparés aux groupes d’enfants de huit ans et de neuf ans.

#### 4. DISCUSSION

Cette étude avait pour objectif, entre autres, d’évaluer la trajectoire développementale pour la capacité d’intégration binaurale d’enfants d’âge scolaire. Comme les données

provenant de l’étude de Jutras & coll. (2012) (données d’enfants de six, neuf et douze ans) n’étaient que partielles pour dessiner cette trajectoire, des enfants de huit ans ont été évalués pour mieux la définir. La moyenne des résultats obtenus par les enfants âgés de huit ans occupe une position intermédiaire entre celle des groupes d’enfants de six et neuf ans. En effet, les résultats n’ont pas montré de différence significative entre les performances des enfants de huit ans et celles des groupes d’enfants de six ans et de neuf ans dans les conditions d’écoute de mots ou de chiffres en compétition. Bien que les résultats des enfants de huit ans se situent parfaitement dans la trajectoire développementale, l’effet de la maturation du système auditif pourrait ne pas être suffisamment important pour se manifester à travers les résultats des enfants d’âge rapproché, contrairement aux groupes d’enfants plus éloignés en âge, comme ce fut le cas pour les performances des enfants de six ans et celles des enfants de douze ans. Un faible échantillon (huit participants), de même que la présence d’un écart-type important chez le groupe des huit ans, comptent aussi parmi les facteurs pouvant expliquer l’absence de résultats significatifs entre les groupes d’enfants plus jeunes. De plus, la nature du stimulus employé influence les résultats de certains groupes d’enfants aux tests d’écoute dichotique. De façon générale, l’utilisation de chiffres conduit à de meilleures performances que celle de mots. La répétition de chiffres constitue une tâche plus facile à réaliser que la répétition de mots, les chiffres étant des stimuli plus familiers et moins complexes à traiter au plan langagier que les mots. Ces résultats concordent avec ceux de l’étude de Hulme, Maughan, & Brown (1991) démontrant que les adultes parviennent à se rappeler avec davantage de

précision les mots qui leur sont familiers en comparaison avec les mots non familiers. De plus, le test EDC a l'avantage d'avoir un ensemble limité de réponses, puisqu'il est constitué des mêmes neuf chiffres, limitant ainsi le nombre d'erreurs possibles. À l'inverse, dans le test de mots, le choix de réponses est très varié (nombreux mots potentiels), ce qui augmente les risques de commettre plus d'erreurs qu'au test EDC. Aussi, la vitesse d'articulation mérite d'être prise en compte en ce qui concerne le rappel d'unités linguistiques, celle-ci étant plus rapide pour les mots familiers conservés dans la mémoire à long terme que pour les mots non familiers (Hulme & coll., 1991). Les enfants plus âgés possèdent également un plus grand bagage langagier, de même qu'une expérience auditive plus vaste et diversifiée, ce qui constitue également des facteurs pouvant être pris en considération dans ce phénomène (Gathercole, 1999).

Par ailleurs, dans la condition d'écoute dichotique la plus difficile (GC avec des mots), les enfants de six ans ont obtenu des performances similaires aux enfants de neuf ans et de douze ans, alors que ceux de neuf ans ont plutôt obtenu des résultats inférieurs à ceux de douze ans. L'absence de différence significative entre les deux groupes d'enfants d'âge éloigné pourrait s'expliquer, entre autres, par un important écart-type à la moyenne, étant près de quatre fois plus grand chez le groupe des six ans comparativement au groupe des douze ans, et deux fois plus grand par rapport au groupe des neuf ans. Dans ce dernier groupe, la variabilité des valeurs tend à diminuer, ce qui peut contribuer à montrer des différences significatives entre leurs résultats et ceux des enfants de douze ans.

Les performances obtenues à l'EDC montrent que les enfants de neuf et douze ans ont obtenu, de façon générale, de meilleures performances aux différentes tâches que les enfants de six ans. Les enfants de neuf et douze ans réussissent donc mieux aux versions *Paires doubles*, *Paires triples* et *Modèle SSW* de l'EDC que ceux de six ans. La capacité de la mémoire auditive pourrait expliquer, en partie, les différences observées. Cette capacité s'améliore avec l'âge (Gathercole, Pickering, & Ambridge, 2004; Roncadin, Pascual-Leone, Rich, & Dennis, 2007; Visu-Petra, Miclea, Cheie, & Benga, 2009).

Un autre but poursuivi par la présente étude était de vérifier si l'oreille à laquelle était perçu le stimulus modulait les performances des enfants au test EDC. Ainsi, cet objectif visait une fois de plus à situer les performances des enfants de huit ans par rapport à celles d'enfants d'autres groupes d'âge pour ce qui est des résultats obtenus à une oreille par rapport à ceux de l'autre oreille. Les performances des enfants de huit ans ne se sont pas distinguées de celles des enfants de six et de neuf ans. L'effet de maturation du système auditif n'est possiblement pas assez important entre les groupes d'âge rapproché pour montrer des différences significatives, tel qu'illustré par les normes des tests auditifs centraux rapportées dans Bellis (2003). De façon générale, indépendamment des groupes, les performances étaient meilleures lorsque les stimuli ont été présentés à l'oreille droite plutôt qu'à l'oreille gauche, ce qui suggère une dominance de l'oreille droite (DOD). Ces résultats abondent

dans le même sens que ceux d'études antérieures (Bellis, 2003; Fennell, Satz, & Morris, 1983; Kimura, 1963; Moncrieff & Wilson, 2009; Mukari & coll., 2006).

Enfin, le test EDC est également sensible à l'influence du nombre d'éléments présentés. Les performances des enfants ont diminué de façon plus marquée pour la version *Paires triples* qui contient le plus grand nombre d'items à répéter, ce qui concorde avec les résultats d'autres études (Mukari & coll., 2006; Musiek, 1983a, 1983b). En plus du traitement auditif, cette tâche à six chiffres nécessiterait davantage l'implication de la mémoire de travail que les épreuves comprenant moins de chiffres. Ce phénomène s'observe aussi en comparant les performances des enfants à la version du test *Paires doubles* et à la version *Modèle SSW*. Elles étaient légèrement inférieures pour la version *Paires doubles* comparativement à celles obtenues à la version *Modèle SSW*. Les quatre chiffres de la version *Paires doubles* étaient présentés en situation d'écoute dichotique, alors que seulement deux des quatre chiffres l'étaient dans la version *Modèle SSW*. Cette disparité entre les résultats pourrait être reliée au fait que le traitement associé à l'écoute dichotique exigerait davantage de ressources que la condition où les stimuli entendus ne sont pas en compétition. Selon le modèle d'utilisation de ressources employé par Pichora-Fuller & coll. (1995), un nombre limité de ressources demeure disponible pour le traitement de l'information, à la fois perceptif et cognitif. Ainsi, plus le traitement perceptif de l'information est exigeant, plus il nécessiterait un nombre important de ressources, ce qui laisserait moins de ressources disponibles pour le traitement cognitif. Par exemple, dans une tâche d'écoute dichotique de deux paires de chiffres, le traitement perceptif demanderait davantage de ressources qu'une tâche à quatre chiffres où seulement deux sont en compétition. Il resterait alors moins de ressources pour mémoriser et récupérer les informations de la mémoire afin de répéter les chiffres, diminuant ainsi les chances de réussite.

## 5. CONCLUSION

Les performances des enfants de huit ans soumis aux tests d'écoute dichotique de chiffres ou de mots ne se distinguent pas significativement de celles des enfants de six ou de neuf ans. Toutefois, la moyenne des performances des enfants de huit ans se situe entre celle des performances du groupe d'enfants de six et celle du groupe d'enfants de neuf ans, montrant une amélioration des performances en fonction de l'âge. Les résultats étaient significativement différents lorsque les données des enfants de huit ans sont comparées à celles des enfants de douze ans. Les performances des enfants de huit ans demeurent de façon générale inférieures à celles des enfants de douze ans. Globalement, les enfants ont des performances significativement supérieures lorsque les stimuli sont des chiffres plutôt que des mots. L'EDC met en évidence une dominance des résultats obtenus lorsque les stimuli étaient présentés à l'oreille droite par rapport à ceux envoyés à l'oreille gauche, indépendamment de l'âge et de la version de l'EDC. Enfin, le nombre de chiffres à répéter influence les performances des enfants en situation d'écoute

dichotique puisque celles-ci tendent à diminuer lorsque le nombre de chiffres augmente.

L'utilisation de l'EDC permettra d'évaluer la capacité d'enfants francophones d'intégrer une information différente provenant de chaque oreille. À partir de résultats anormaux, il sera possible d'intervenir afin d'aider les enfants ayant un problème d'intégration binaurale à mieux comprendre dans des situations où la parole est en compétition. Le but de cette intervention vise surtout à améliorer la compréhension de la parole pour atteindre une communication efficace et fluide, principalement en condition d'écoute compétitive.

## REMERCIEMENTS

Les auteurs voudraient remercier les enfants ainsi que leurs parents pour leur disponibilité, de même que les intervenants de la Commission Scolaire Marguerite-Bourgeoys pour leur contribution au recrutement des participants.

## RÉFÉRENCES

- ASHA – American Speech-Language-Hearing Association. (2005). *(Central) Auditory Processing Disorders*. Repéré à: <http://www.asha.org/docs/html/tr2005-00043.html>.
- Bellis, T. J. (2003). *Assessment and management of central auditory processing disorders in the educational setting : From science to practice* (2<sup>nd</sup> ed.). Toronto, Canada: Thomson Delmar Learning, 532 p.
- Bérard, C. (1990-1993). *Normes du SSW et du SSI-ICM en français*. Montréal, Canada : Hôpital Rivière-des-Prairies.
- Berlin, C. I., Hughes, L. F., Lowe-Bell, S. S., & Berlin, H. L. (1973). Dichotic right ear advantage in children 5 to 13. *Cortex*, 9, 392-402.
- Carhart, R., & Jerger, J. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech and Hearing Disorders*, 16, 340-345.
- Champoux, F., Paiement, P., Mercier, F., Lepore, F., Lassonde, M., & Gagné, J.-P. (2007). Auditory processing in a patient with a unilateral lesion of the inferior colliculus. *European Journal of Neuroscience*, 25, 291-297.
- Fennell, E. B., Satz, P., & Morris, R. (1983). The development of handedness and dichotic listening asymmetries in relation to school achievement : A longitudinal study. *Journal of Experimental Child Psychology*, 35, 248-262.
- Gadea, M., Marti-Bonmatí, L., Arana, E., Espert, R., Salvador, A., & Casanova, B. (2009). Corpus callosum function in verbal dichotic listening: Inferences from a longitudinal follow-up of relapsing-remitting multiple sclerosis patients. *Brain & Language*, 110, 101-105.
- Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*, 3(11), 410-419
- Gathercole, S. E., Pickering, S. J., & Ambridge, B. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40, 177-190.
- Hannay, H. J., Walker, A., Dennis, M., Kramer, L., Blaser, S., & Fletcher, J. M. (2008). Auditory interhemispheric transfer in relation to patterns of partial agenesis and hypoplasia of the corpus callosum in spina bifida meningomyelocele. *Journal of the International Neuropsychological Society*, 14, 771-781.
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, 30, 685-701.
- Jacobson, J. T., Deppe, U., & Murray, T. J. (1983). Dichotic paradigms in multiple sclerosis. *Ear & Hearing*, 4, 311-317.
- Jutras, B., & Mayer, D. (2004). *Normalisation du test d'écoute dichotique de chiffres en français : Étude préliminaire*. (Rapport scientifique). Montréal : Université de Montréal, Ecole d'orthophonie et d'audiologie, Faculté de médecine.
- Jutras, B., Mayer, D., Joannette, É., Carrier, M.-È., & Chénard, G. (2012). Assessing the Development of Binaural Integration Ability With the French Dichotic Digit Test : Écoute Dichotique de Chiffres. *American Journal of Audiology*, 21, 51-59.
- Katz, J. (1992). Classification of auditory processing disorders. In J. Katz, N. Stecker, & D. Henderson (Eds.), *Central auditory processing: A transdisciplinary view* (pp. 81-91). Baltimore, MD: Mosby-Yearbook.
- Katz, J., Basil, R. A., & Smith, J. M. (1963). A Staggered spondaic word test for detecting central auditory lesions. *Annals of Otolaryngology & Laryngology*, 72, 908-917.
- Kimura, D. (1961a). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 15(3), 166-171.
- Kimura, D. (1961b). Some effects of temporal-lobe damage on auditory perception. *Canadian Journal of Psychology*, 15(3), 156-165.
- Kimura, D. (1963). Speech lateralization in young children as determined by an auditory test. *Journal of Comparative & Physiological Psychology*, 56, 899-902.
- Martin, J., Jerger, J., & Mehta, J. (2007). Divided-attention and directed-attention listening modes in children with dichotic deficits: An event related potential study. *Journal of the American Academy of Audiology*, 18, 34-53.
- Meyer, V. (1957). *Cognitive changes following temporal lobectomy for the relief of focal temporal lobe epilepsy*. Unpublished doctoral thesis, University of London.
- Milner, B. (1958). Psychological defects produced by temporal lobe excision. *Research Publications - Association for Research in Nervous and Mental Disease*, 36, 244-257.
- Moncrieff, D., Jerger, J., Wambacq, I., Greenwald, R., & Black, J. (2004). ERP evidence of a dichotic left-ear deficit in some dyslexic

children. *Journal of the American Academy of Audiology*, 15, 518-534.

Moncrieff, D. W., & Wilson, R. H. (2009). Recognition of randomly presented one-, two-, and three-pair dichotic digits by children and young adults. *Journal of the American Academy of Audiology*, 20, 58-71.

Mukari, S. Z., Keith, R. W., Tharpe, A. M., & Johnson, C. D. (2006). Development and standardization of single and double dichotic digit tests in the Malay language. *International Journal of Audiology*, 45, 344-352.

Musiek, F. E. (1983a). Assessment of central auditory dysfunction: The Dichotic Digit Test revisited. *Ear & Hearing*, 4, 79-83.

Musiek, F. E. (1983b). Results of the three dichotic speech tests on subjects with intracranial lesions. *Ear & Hearing*, 4, 318-323.

Musiek, F. E., Gollegly, K. M., Kibbe, K. S., & Verkest-Lenz, S. B. (1991). Proposed screening test for central auditory disorders: Follow-up on the dichotic digits test. *The American Journal of Otology*, 12, 109-113.

Musiek, F. E., Reeves, A. G., & Baran, J. A. (1985). Release from central auditory competition in the split-brain patient. *Neurology*, 35, 983-987.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of Acoustical Society of America*, 97(1), 593-608.

Roncardin, C., Pascual-Leone, J., Rich, J. B., & Dennis, M. (2007). Developmental relations between working memory and inhibitory control. *Journal of the International Neuropsychological Society*, 13, 59-67.

Rudmin, F., & Normandin, N. (1983). Experimental dichotic tests in French modeled on SSW design. *Human Communication*, 7, 348-360.

Strouse, A., & Wilson, R. H. (1999). Recognition of one-, two-, and three-pair dichotic digits under free and directed recall. *Journal of the American Academy of Audiology*, 10, 557-571.

Stuart, A. (2005). Development of Auditory Temporal Resolution in School-Age Children Revealed by Word Recognition in Continuous and Interrupted Noise. *Ear & Hearing*, 26(1), 78-88.

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research*, 28(3), 455-462.

Techentin, C., & Voyer, D. (2011). Word frequency, familiarity, and laterality effects in a dichotic listening task. *Laterality*, 16, 313-332.

Visu-Petra, L., Miclea, M., Cheie, L., & Benga, O. (2009). Processing efficiency in preschoolers' memory span: Individual differences related to age and anxiety. *Journal of Experimental Child Psychology*, 103, 30-48.

Willeford, J. A., & Burleigh, J. M. (1994). Sentence procedures in central testing. In J. Katz (Ed.), *Handbook of clinical audiology* (4<sup>th</sup> ed., pp. 256-268). Baltimore, MD: Williams & Wilkins

Enhancing where people  
live, work and play  
through the application  
of the principles of  
acoustical engineering.



Consulting Engineers specializing in  
**Acoustics, Noise and Vibration**

HOWE GASTMEIER CHAPNIK LIMITED  
Mississauga, Ontario  
P: 905-826-4044 F: 905-826-4940  
[www.hgcengineering.com](http://www.hgcengineering.com)

# ÉVALUATION DE L'ÉCOUTE DICHOTIQUE CHEZ L'ENFANT DYSLEXIQUE

Hung Thai-Van et Evelyne Veuillet

Service d'Audiologie & d'Explorations Orofaciales, Hospices Civils de Lyon  
& Centre de Recherche en Neurosciences de Lyon, Inserm U1028, CNRS UMR5292

## SOMMAIRE

Parmi les tests permettant de mesurer la reconnaissance des sons de parole en situation d'écoute compétitive, le test d'écoute dichotique est de plus en plus utilisé y compris chez l'enfant à partir de l'âge de huit ans. Il consiste à délivrer simultanément aux deux oreilles des informations verbales de même nature (substantifs, chiffres ou adjectifs) mais différents, puis de demander au sujet de répéter ce qu'il a entendu dans différentes conditions attentionnelles: attention partagée ou dirigée successivement sur chaque oreille. Ce test permet de mesurer les capacités d'intégration et de séparation binaurale qui sont cruciales pour sélectionner l'information auditive d'intérêt lorsqu'elle est délivrée en présence d'informations concurrentes. L'objectif de cette étude est de préciser les mécanismes sous-jacents à la faiblesse relative des aptitudes dichotiques décrite chez l'enfant dyslexique en se focalisant sur les compétences mises en jeu pour séparer binauralement les informations auditives. Pour ce faire, nous avons comparé les aptitudes dichotiques d'un groupe d'enfants normo-lecteurs et d'un groupe d'enfants dyslexiques appariés en genre, latéralité manuelle et âge chronologique. Les résultats confirment l'existence d'une dégradation des aptitudes dichotiques chez les enfants dyslexiques et montrent que cette dégradation n'est pas oreille-dépendante. Ils montrent également une dégradation des capacités d'attention dirigée et du déplacement de l'attention d'une oreille vers l'autre. Cette étude corrobore l'hypothèse d'une co-morbidité entre les troubles spécifiques d'apprentissage de la lecture et la présence d'un trouble du traitement auditif central.

## ABSTRACT

Among the various tests used for assessing speech sounds recognition in competing speech perception, there is a growing interest in using the dichotic listening test also for children above eight years of age. Dichotic listening consists in presenting different speech material (spondees, digits or adjectives) to both ears simultaneously; the subject then has to verbally reproduce either all the perceived words (free recall) or only those perceived in one ear (directed attention). This behavioral technique measures the ability to binaurally separate and integrate information, a crucial skill for speech sound feature extraction in the presence of concurrent auditory information. Here we aimed at investigating further the relative weakness of dichotic listening skills in dyslexic children more particularly the skills allowing the binaural separation. For this, dichotic aptitudes in a group of dyslexic children were compared with those of gender-, manual laterality- and age-matched controls. In line with literature data, dyslexic children were found to have deteriorated dichotic listening skills. The deterioration was found to equally affect the right ear and the left ear. In addition, directed auditory attention and attentional shift index were also found to be impaired. The results support the hypothesis that comorbidity between developmental reading disorder and central auditory processing disorder might occur.

## 1. INTRODUCTION

L'évaluation des compétences d'écoute dichotique repose sur la présentation simultanée de sons de parole différents aux deux oreilles (Broadbent, 1954). Pour le sujet testé, la tâche consiste à répéter soit l'ensemble des mots perçus (rappel libre) soit seulement ceux présentés à une des deux oreilles (attention dirigée). Ce test comportemental non-invasif permet de mesurer un aspect fondamental du traitement des sons de parole par le système nerveux auditif central, à savoir la capacité de ce dernier à séparer ou

intégrer l'information verbale délivrée binauralement. Il permet également de déterminer, pour un sujet donné, quelle est son oreille privilégiée pour l'écoute, Kimura (1967) expliquant l'avantage en faveur de l'oreille droite par une prédominance des voies auditives contralatérales et la spécialisation des aires corticales temporales gauches dans le traitement de la parole. Ce degré d'asymétrie auditive représente un indicateur indirect des différences de latéralité entre les deux hémisphères lorsqu'ils sont activés simultanément par les deux oreilles (Bryden, 1988; Hugdahl, 1995). Le type de matériel verbal utilisé est varié:

mots, chiffres, syllabes de type Consonne-Voyelle (CV) ou CVC, voire phrases. De plus, ce test dichotique trouve sa richesse dans la variété des conditions de passation qui permettent d'examiner soit l'attention partagée (rappel libre) soit l'attention sélective (attention dirigée vers une des deux oreilles). D'après Jerger & Martin (2006), c'est la combinaison de ces deux types de tâches (rappel libre et attention dirigée) qui est la plus à même de dissocier ce qui revient aux processus auditifs de ce qui revient aux processus cognitifs et aux fonctions exécutives (notamment aux capacités attentionnelles, mais également à la mémoire verbale à court-terme et à la vitesse de traitement de l'information délivrée). Cela est essentiel puisque depuis maintenant plusieurs années, les sujets pour qui un trouble du traitement auditif (TTA) est suspecté sont soumis à des procédures d'écoute dichotique (Jerger & Musiek, 2000). De même, du fait d'une forte comorbidité entre TTA et dyslexie (McArthur & Bishop, 2001), il peut s'avérer informatif d'administrer à des enfants présentant un trouble spécifique du langage, ce type de test évaluant certains processus auditifs centraux. Mais les premières études ont conduit les auteurs à s'interroger sur la fiabilité des résultats obtenus au moyen de cette procédure dichotique (Hiscock & Kinsbourne, 1982 ; 1995). Puis, après contrôle des facteurs cognitifs de biais, le test dichotique s'est révélé utile pour identifier, voire même catégoriser les enfants présentant un trouble du langage (Asbjornen et al., 2003). Les aptitudes dichotiques, mesurées au moyen d'épreuves d'intégration et de séparation binaurale, sont significativement altérées chez des enfants dyslexiques (Demanez et al., 2003a). Plus récemment, Moncrieff & Black (2008) ont confirmé l'existence d'un déficit d'intégration binaurale chez les enfants dyslexiques en montrant que leurs performances dichotiques sont plus faibles que celles d'enfants normo-lecteurs de même âge. Nous nous demandons à présent si les dyslexiques présentent également un déficit de séparation binaurale. Cette étude a donc pour objectif de comparer les capacités de séparation binaurale entre des enfants dyslexiques et normo-lecteurs, sachant qu'elles sont mises en jeu lorsque la procédure dichotique inclut la présentation de listes pour lesquelles on demande à l'enfant de déplacer son attention alternativement sur l'oreille droite et gauche (20 paires d'items constitués de chiffres ou d'adjectifs). Une comparaison des capacités de déplacement attentionnel a été réalisée. Enfin, pour le groupe des dyslexiques, nous avons mis en corrélation ce déplacement avec les capacités d'attention auditive.

## 2. MATÉRIEL ET MÉTHODES

### 2.1 Participants

Quatre-vingt enfants âgés de 7 à 14 ans ont participé à cette étude. Ils étaient répartis en deux groupes de 40 enfants chacun appariés en âge: un groupe d'enfants normo-lecteurs (19 filles – 21 garçons) et un groupe d'enfants dyslexiques (17 filles – 23 garçons). Aucun de ces enfants ne présentait de déficit auditif, chacun présentant sur chaque oreille des seuils audiométriques en conduction aérienne inférieurs ou égaux à 15 dBHL entre 0.25 et 8 kHz. A l'issue de

l'interrogatoire et de l'examen clinique, aucun n'avait de troubles de l'attention avec ou sans hyperactivité, d'affection chronique psychologique ou neurologique. Tous avaient un niveau intellectuel normal évalué selon les critères du WISC 4 chez l'enfant dyslexique (Weschler, 2005) et au moyen des Matrices Progressives de Raven (PM47-édition 98, Raven et al., 1984) chez l'enfant normo-lecteur. Tous étaient droitiers ayant un score compris entre 71 et 100% au questionnaire de latéralité d'Edinburgh (Oldfield, 1971) et monolingues avec le français comme langue maternelle. Les enfants normo-lecteurs n'avaient jamais redoublé à l'école. Ils présentaient des performances de lecture correspondant à leur âge chronologique selon les critères du test de lecture « L'Alouette » (Lefavrais, 1965). Ce test évalue le niveau de lecture en se basant à la fois sur le décodage des mots et non-mots et sur la vitesse de lecture. L'anamnèse des enfants dyslexiques révélait des difficultés persistantes et spécifiques d'acquisition de la lecture : ils accusaient un retard de plus de 18 mois de l'âge de lecture par rapport à leurs pairs.

Ces caractéristiques sont résumées dans le Tableau 1.

**Tableau 1.** Caractéristiques des sujets normo-lecteurs et dyslexiques (Moyennes ± erreurs standards, minimum - maximum et valeurs des t-tests).

	Normo-lecteurs (n=40)	Dyslexiques (n=40)	T ou T, p (ddl=78)
Sexe	19 filles 21 garçons	17 filles 23 garçons	
Age chronologique (en mois)	125,1 ± 3 [92-168]	126,8 ± 2,3 [97-155]	t= 0,441, p=0,66
Age de lecture <sup>#</sup> (en mois)	125,8 ± 3 [97-169]	90,8 ± 1,4 [75-120]	T = 2378, p<0,001
Retard de lecture <sup>##</sup> (en mois)	0,7 ± 1,5 [-23-15]	36,1 ± 1,8 [19-59]	t= 15,569, p<0,001
Quotient intellectuel non verbal	41,7 ± 1,2 <sup>†</sup> [-25-54]	96,0 ± 2,3 <sup>‡</sup> [71-128]	---

<sup>#</sup>Test de lecture (Alouette, Lefavrais, 1967)

<sup>##</sup>Décalage entre l'âge chronologique et l'âge de lecture

<sup>†</sup>Matrices de Raven PM47: À noter qu'aucun des enfants normo-lecteurs ne présente un score de Raven en dessous du percentile 10 pour l'âge

<sup>‡</sup>Indice de raisonnement perceptif (WISC4, Weschler, 2005): considéré comme normal s'il est en-dessous de 70.

### 2.2 Procédure d'évaluation de l'écoute dichotique

L'écoute dichotique étant une écoute « à deux oreilles », on présentait simultanément à chaque oreille de l'enfant un stimulus différent. Les stimuli ainsi groupés en paires étaient de nature verbale (mots, chiffres), équilibrés acoustiquement, phonétiquement, linguistiquement et sémantiquement. L'écoute dichotique était évaluée dans deux conditions d'attention : la tâche de rappel libre (Oreille non désignée (OrNonDés)) et la tâche d'attention dirigée (Oreille désignée (OrDés)) mesurant respectivement l'intégration et la séparation binaurale.

Ce test est inclus dans le Bilan Auditif Central (BAC) de Demanez et al. (2003b). Il est constitué de cinq listes de 10 paires d'items (substantifs, adjectifs, chiffres) chacune utilisable à partir de l'âge de huit ans. L'épreuve complète



porte ainsi sur 50 paires d'items présentées soit en condition OrNonDés (« répéter les deux items perçus respectivement à droite et à gauche »), soit en condition OrDés (« répéter uniquement l'item perçu à droite ou à gauche »).

Les listes ont toujours été présentées dans l'ordre suivant :

- Liste de un substantif bissyllabique en condition OrNonDés (exemple : « maison » et « radis »);
- Liste de trois chiffres en condition OrDés (exemple : « 5-6-4 » et « 8-10-3 »);
- Liste de deux chiffres en condition OrNonDés (exemple « 5-6 » et « 8-10 »);
- Liste de deux adjectifs monosyllabiques en condition OrDés (exemple « brun-doux » et « grand-beau »);
- Liste de deux adjectifs monosyllabiques en condition OrNonDés (exemple « brun-doux » et « grand-beau »).

En condition OrNonDés (rappel libre), l'enfant devait répéter tous les items présentés à la fois à l'oreille droite et à l'oreille gauche (10 paires de 1 substantif, 10 paires de 2 chiffres et 10 paires de 2 adjectifs).

En condition OrDés (rappel dirigé), l'attention était dirigée soit :

- vers l'oreille droite (OrDrDés): on demandait à l'enfant de se concentrer sur l'oreille droite. Il devait ignorer les informations délivrées à l'oreille gauche et répéter les items entendus à droite ;
- vers l'oreille gauche (OrGaDés): on demandait à l'enfant de se concentrer sur l'oreille gauche. Il devait ignorer les informations délivrées à l'oreille droite et répéter les items entendus à gauche.

Cette condition OrDés a concerné deux listes d'items (trois chiffres et deux d'adjectifs), l'enfant devant d'abord répéter les cinq premiers items entendus dans l'oreille droite, puis les cinq suivants entendus dans l'oreille gauche. Ces 10 mêmes items ont alors été présentés une seconde fois, l'enfant commençant par porter son attention sur l'oreille gauche pour les cinq premiers items puis sur l'oreille droite pour les cinq suivants.

Seule la liste de paires de deux adjectifs monosyllabiques a été présentée dans les deux conditions d'écoute (OrDés et OrNonDés)

Les mesures ont été faites dans une cabine insonorisée. Les stimuli, administrés au moyen d'un lecteur CD relié à un audiomètre, étaient délivrés dans des écouteurs TDH39 à une intensité confortable de 60 dB SPL (correspondant à 100% d'intelligibilité pour chaque oreille). La durée totale de la procédure d'évaluation dichotique était de 20 minutes.

## 2.3 Paramètres dichotiques étudiés

L'évaluation de l'écoute dichotique a permis de calculer trois paramètres principaux :

- les aptitudes dichotiques qui mesurent les capacités du sujet à répéter tout ce qu'il a entendu
- la prévalence d'oreille qui permet de connaître l'oreille qui est privilégiée lors de l'écoute
- l'indice de déplacement attentionnel qui reflète les capacités de l'enfant à suivre les consignes en déplaçant son attention vers l'oreille désignée

Les aptitudes dichotiques (AD) s'évaluent par le nombre de paires d'items correctement et complètement répétés en consigne OrNonDés additionné du nombre d'items correctement reproduits en consigne OrDrDés et OrGaDés. Ce nombre final est multiplié par 2 pour obtenir une valeur exprimée en pourcentage.

La prévalence d'oreille (PO) est évaluée en dénombrant les réponses dites « exclusives » c'est-à-dire répétées justes (complètes) uniquement sur une oreille. En soustrayant ces réponses entre l'oreille droite et l'oreille gauche, en divisant ce nombre par la totalité des réponses exclusives et en le multipliant par 100, on obtient la prévalence d'oreille (PO). Une valeur positive indique un avantage en faveur de l'oreille droite alors qu'une valeur négative indique un avantage en faveur de l'oreille gauche.

Pour connaître le détail du mode de calcul des paramètres AD et PO, le lecteur peut se rapporter à la description qui en est faite par Demanez et al. (2003b).

L'indice de déplacement attentionnel (IDA) a été mesuré à partir des réponses obtenues en consigne OrDrDés et OrGaDés selon la procédure décrite par Asbjornsen & Hugdahl (1995). Ces auteurs ont proposé d'évaluer le facteur attentionnel en prenant en compte le nombre de réponses intrusives (c'est-à-dire le nombre de répétitions d'items présentés à l'oreille qui n'est pas désignée). Le calcul de l'IDA repose sur un ratio des réponses correctes et des réponses intrusives selon la formule suivante adaptée d'Asbjornsen et Bryden (1998) :

$$IDA = \ln \left[ \frac{(OD_{OrDrDés} * OG_{OrGaDés})}{(OD_{OrGaDés} * OG_{OrDrDés})} \right] \text{ avec}$$

- ✓  $OD_{OrDrDés}$  : nombre d'adjectif et chiffres répétés justes lorsqu'ils étaient présentés sur l'oreille droite et que l'enfant devait porter son attention sur l'oreille droite
- ✓  $OG_{OrGaDés}$  : nombre de mots et chiffres répétés justes lorsqu'ils étaient présentés sur l'oreille gauche et que l'enfant devait porter son attention sur l'oreille gauche

- ✓  $OD_{OrGaDés}$  : nombre de mots et de chiffres répétés justes lorsqu'ils étaient présentés sur l'oreille droite et que l'enfant devait porter son attention sur l'oreille gauche i.e. les réponses intrusives droites
- ✓  $OG_{OrDrDés}$  : nombre de mots et de chiffres répétés justes lorsqu'ils étaient présentés sur l'oreille gauche et que l'enfant devait porter son attention sur l'oreille droite i.e. les réponses intrusives gauches

## 2.4 Évaluation cognitive

Seuls les enfants dyslexiques avaient une évaluation neuropsychologique incluant pour certains d'entre eux (23 sur 40) la mesure de l'attention sélective auditive issue du sous-test de la batterie NEPSY proposée par Korkman et al. (1998).

## 2.5 Analyse des résultats

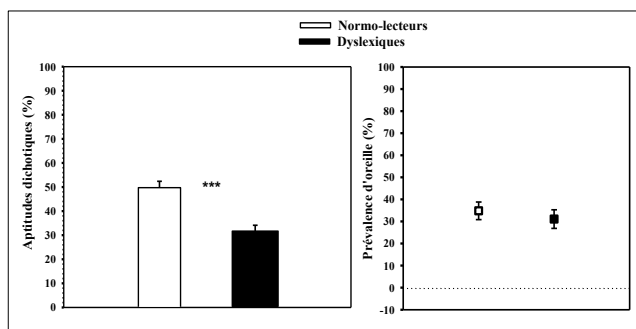
Toutes les variables ont été évaluées ou calculées séparément pour chaque enfant de l'étude. Des analyses statistiques paramétriques ou non paramétriques ont été conduites : test t de Student et T de Mann-Whitney comparant respectivement les âges chronologique et de lecture, tests t de Student pour comparer les AD, PO et IDA entre les deux groupes d'enfants (normo-lecteurs et dyslexiques) et Analyses de la Variance (ANOVA) à deux facteurs sur mesures répétées uniquement pour les facteurs Oreilles (deux niveaux) ou Listes (deux niveaux) et comme facteur non répété le Groupe (deux niveaux). En cas de significativité les ANOVA ont été suivies de tests de comparaison multiples des moyennes (test de Tukey). Pour les dyslexiques uniquement, un test de corrélation (coefficient de Pearson) a été effectué pour étudier le lien entre l'IDA et les capacités d'attention sélective (sous-test de la NEPSY).

La significativité est atteinte lorsque p est inférieur à 5%.

## 3. RÉSULTATS

### 3.1 Aptitudes dichotiques et prévalence d'oreille

La figure 1 compare les performances moyennes au test dichotique de chaque groupe d'enfants. La figure de gauche montre que les enfants dyslexiques présentaient des aptitudes dichotiques significativement plus faibles que celles des enfants normo-lecteurs ( $t(78) = 5.02, p < 0,001$ ). La figure à droite met en évidence que pour les deux groupes, la prévalence d'oreille correspondait à un avantage en faveur de l'oreille droite. Il n'y avait pas de différence significative de prévalence d'oreille entre les deux groupes d'enfants ( $t(78) = 0.645, p = 0,521$ ).



**Figure 1.** Comparaison des aptitudes dichotiques et de la prévalence d'oreille entre le groupe d'enfants dyslexiques et le groupe d'enfants normo-lecteurs ( $n=40$  dans chaque groupe). Le diagramme de gauche montre les aptitudes dichotiques, celui de droite la prévalence d'oreille. Les moyennes  $\pm$  erreur standard sont représentées. . \*\*\* :  $p < 0.001$ .

### 3.2 Performances dichotiques au sein de la condition d'oreille désignée (attention dirigée)

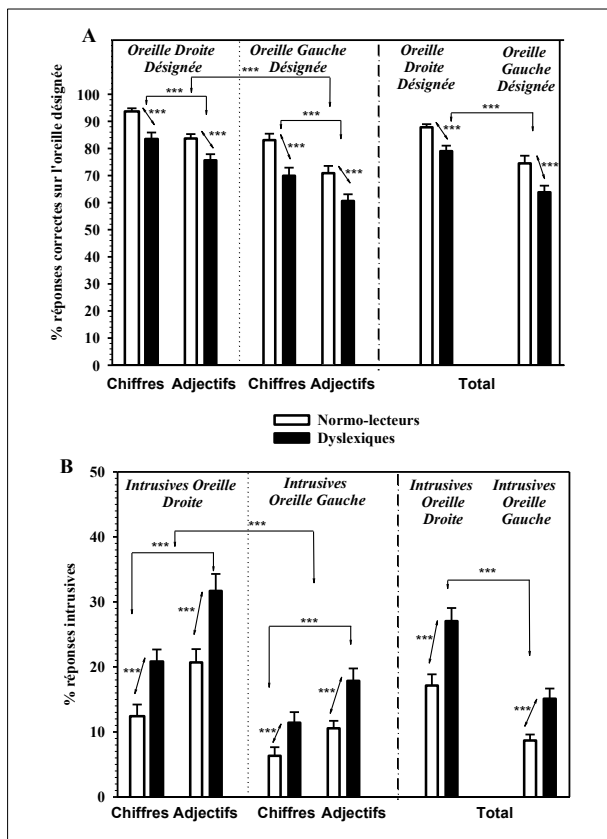
La figure 2 (gauche et centre) permet de comparer, entre les deux groupes d'enfants, les pourcentages moyens de réponses correctes (figure 2A) et de réponses intrusives (figure 2B) en fonction des conditions d'oreille désignée ( $OrDrDés$  et  $OrGaDés$ ) et du type de listes (chiffres et adjectifs). Les figures de droite correspondent aux données cumulées des scores obtenus après présentation des listes de chiffres et d'adjectifs.

L'observation de la figure 2A permet de constater que les items verbaux correctement répétés sur l'oreille désignée sont toujours moins nombreux chez les enfants dyslexiques. Cela est vrai aussi bien en présence de chiffres que d'adjectifs mais également en condition d'oreille désignée droite par rapport à gauche. Il y a ainsi toujours moins d'items correctement répétés en condition  $OrGaDés$  dans les deux groupes d'enfants. Des ANOVAs à 2 facteurs (groupe x oreilles) répétées sur le facteur « oreille » ont été réalisées pour chaque liste séparément. Pour la liste « adjectifs » il existe un effet significatif du facteur groupe ( $F(1,78)=14,51, p < 0.001$ ) et du facteur oreille ( $F(1,78)=41,02, p < 0.001$ ) avec une interaction non significative. Il en est de même pour les scores obtenus en réponse aux chiffres avec un effet significatif du facteur groupe ( $F(1,78)=20,46, p < 0.001$ ) et du facteur oreille ( $F(1,78)=45,02, p < 0.001$ ) en l'absence d'interaction significative. Les résultats ne diffèrent donc pas selon le matériel verbal utilisé (adjectifs ou chiffres).

Concernant les réponses intrusives, l'analyse détaillée par listes a montré que quelle que soit la liste présentée (figure 2B, gauche et centre), les enfants dyslexiques ont eu plus de difficultés que les normo-lecteurs pour ne pas répéter les items verbaux entendus dans l'oreille non-désignée commettant ainsi globalement plus de réponses intrusives. Toutefois les réponses intrusives droites étaient toujours plus nombreuses que les réponses intrusives gauches chez les normo-lecteurs comme chez les dyslexiques. Comme précédemment pour les réponses correctes, des ANOVAs à

deux facteurs (groupe x oreille) répétées sur le facteur oreille ont été réalisées sur les scores de réponses intrusives obtenus en réponse à chacune des deux listes (adjectifs et chiffres) séparément. En présence d'adjectifs, les facteurs groupe et oreille ont exercés tous deux un effet significatif (respectivement  $F(1,78)=17,58$ ,  $p<0.001$  et  $F(1,78)=44,75$ ,  $p<0.001$ ) sans interaction significative. Il en est de même lorsque des chiffres ont été présentés avec dans ce cas  $F(1,78)=12,69$ ,  $p<0.001$  pour le facteur groupe et  $F(1,78)=30,96$ ,  $p<0.001$  pour le facteur oreille en l'absence d'interaction significative entre les deux facteurs.

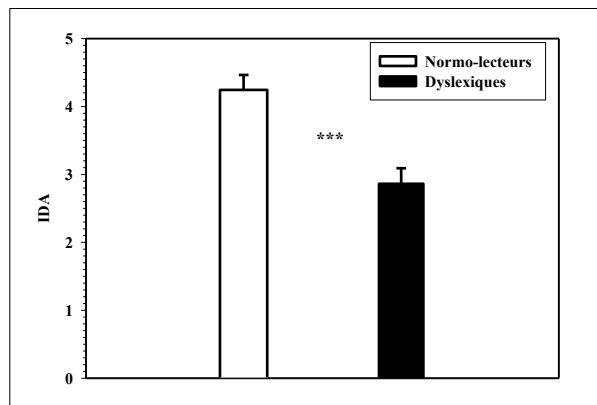
Les plus faibles pourcentages de réponses correctes et les plus fortes impossibilités d'inhiber l'écoute plus particulièrement sur l'oreille droite observés chez les enfants dyslexiques comparés à des normo-lecteurs ont suggéré l'existence de problèmes attentionnels dans la population dyslexique.



**Figure 2.** Comparaison du pourcentage moyen ( $\pm$  erreur standard) des réponses correctes (A) et intrusives (B) pour les deux listes présentées en condition d'oreille désignée (à gauche et au centre : listes séparées et à droite cumulé des deux listes) entre les deux groupes d'enfants. (\*\*\*) :  $p<0.001$ .

### 3.3 Déplacement Attentionnel

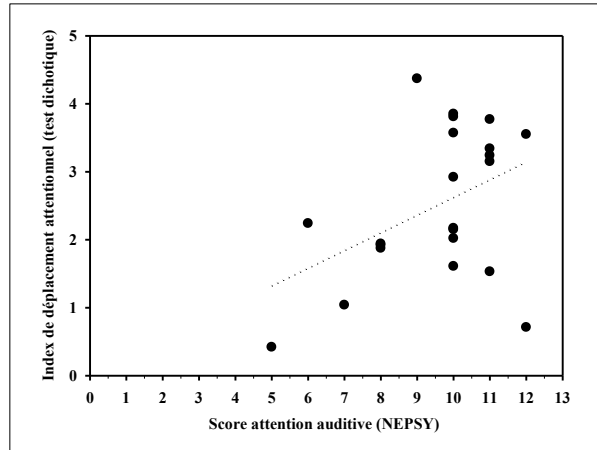
La figure 3 représente l'index de déplacement attentionnel (IDA) chez les enfants normo-lecteurs et dyslexiques. Il était significativement plus faibles pour le groupe des dyslexiques ( $t(78) = 4,38$ ,  $p < 0,001$ ).



**Figure 3.** Comparaison de l'index de déplacement attentionnel (IDA) entre les enfants normo-lecteurs et dyslexiques (moyenne  $\pm$  erreur standard, \*\*\* :  $p<0,001$ ).

### 3.4 Lien entre les capacités attentionnelles mesurées par l'évaluation neuropsychologique (NEPSY) et les compétences dichotiques

La figure 4 montre chez les enfants dyslexiques un lien significatif moyen ( $r=0,434$ ,  $p=0.04$ ) entre les capacités attentionnelles auditives et le déplacement attentionnel. Les enfants qui ont le plus de difficultés à déplacer leur attention (IDA faible) sont généralement ceux qui présentent les plus faibles scores neuropsychologiques d'attention.



**Figure 4.** Relation entre les scores attentionnels (NEPSY) et le déplacement attentionnel (test dichotique) chez les enfants dyslexiques.

## 4. DISCUSSION

Les résultats de cette présente étude mettent en évidence qu'en moyenne les aptitudes dichotiques sont réduites chez un groupe d'enfants dyslexiques droitiers alors que la prévalence moyenne d'oreille qui montre un avantage en faveur de l'oreille droite est comparable à celle observée dans un groupe d'enfants normo-lecteurs de même âge chronologique, latéralité manuelle et genre. Cette étude qui

visait à cibler plus particulièrement les performances mesurées en condition d'oreille désignée (rappel dirigé) met en évidence des déficits de séparation binaurale chez les enfants dyslexiques comparés au normo-lecteurs. Ils se caractérisent par des scores de répétition toujours plus faibles, et cela quel que soit le type de listes présentées (chiffres ou adjectifs) que ceux des enfants normo-lecteurs. Cela se vérifie que ce soit en condition d'écoute dirigée vers l'oreille droite ou gauche tout en présentant des pourcentages de réponses intrusives significativement plus élevés que les normo-lecteurs sur les deux oreilles. On a pu constater que chez les normo-lecteurs mais de manière encore plus marquée chez les dyslexiques, c'est sur l'oreille droite que les difficultés d'inhibition de l'écoute étaient les plus intenses.

Chez l'enfant dyslexique, malgré des résultats contradictoires attribuables à la diversité des procédures et du matériel verbal utilisés, les études antérieures concluent à un déficit de l'intégration binaurale, avec des performances significativement dégradées par rapport aux enfants normo-lecteurs appariés en âge chronologique (Moncrieff & Black, 2008). Ces enfants déficitaires dans l'acquisition de la lecture se caractérisent par des aptitudes dichotiques (réponses correctes sur les deux oreilles) significativement plus faibles que leurs pairs (Demanez et al., 2003a) ce que nous confirmons dans cette présente étude.

Concernant la prévalence d'oreille, deux principales configurations sont décrites chez les enfants dyslexiques dans la littérature. La première fait état d'asymétries non comparables à celles observées chez les enfants normo-lecteurs. Il a ainsi été décrit une forte asymétrie interaurale en faveur de l'oreille droite expliquée par des performances normales sur l'oreille droite mais significativement réduites sur l'oreille gauche (Lamm & Epstein, 1994 ; Morton & Siegel, 1991 ; Moncrieff & Musiek, 2002), comme cela est classiquement décrit chez les enfants présentant un trouble central de l'audition (Bellis & Ferre, 1999; Jerger et al., 1999; Musiek, 1999; Moncrieff, 2006). Cela laisse supposer l'existence d'un transfert inter hémisphérique altéré en accord avec le modèle de Kimura (1967). Des études anatomiques ont mis en évidence des différences de développement des fibres du corps calleux entre enfants dyslexiques et enfants normo-lecteurs (vonPlessen, 2002), venant ainsi corroborer l'hypothèse d'une moindre myélinisation dans un cerveau immature (Morton, 1994 ; Swanson & Cochran, 1991). D'autres études utilisant comme stimuli des syllabes de type CV, rapportent chez l'enfant dyslexique des performances symétriques avec un avantage en faveur de l'oreille droite nettement moins marqué que la normale (Boliek et al., 1988; Brunswick & Rippon, 1994; Kershner & Micaloff, 1992) voire même un déficit droit (Moncrieff & Black, 2008). Cette faiblesse de l'oreille droite persisterait durant l'adolescence et serait encore présente à l'âge adulte (Hugdahl et al., 1995). Cela a conduit à la formulation d'autres hypothèses et parmi elles celle d'un dysfonctionnement de l'hémisphère gauche ou

bien d'une hyper-activation concernant soit l'hémisphère droit seul soit les deux hémisphères (Kershner & Morton, 1990). Enfin d'autres études rapportent un avantage de l'oreille gauche chez l'enfant dyslexique lors de la tâche de rappel libre, suggérant la possibilité d'un planum temporel davantage développé à droite (Foster et al., 2002). Plus récemment, cette configuration de type « défaut d'asymétrie » se caractérisant par de plus forts avantages d'oreille a été interprétée selon le modèle de l'amblyopie dans la vision. Un enfant présentant ce phénomène apparenté à de « l'amblyaudio » (Moncrieff, 2011) présenterait des performances normales au niveau de l'oreille dominante et des performances significativement réduites à l'oreille non dominante assimilable à une « oreille paresseuse ». Il est intéressant de constater que cette déficience unilatérale, qui est potentiellement modifiable par l'entraînement (Moncrieff & Wertz, 2008), a été récemment reproduite au moyen d'un modèle animal de perte auditive de transmission (Popescu & Polley, 2010).

La seconde configuration est celle de performances réduites dans les deux oreilles. Dans ce cas, il est souvent bien difficile d'en départager les causes: déficit du langage, capacités intellectuelles réduites, manque de motivation, faiblesse de la mémoire de travail ou des processus attentionnels, fatigue. De plus, comme l'étude de Moncrieff & Black (2008) le montre, la détermination du côté déficitaire au test dichotique reste très tributaire de la procédure expérimentale utilisée avec des effets différents observés entre enfants dyslexiques et normo-lecteurs. La charge exercée par le facteur attentionnel est sans doute à l'origine de ces variabilités (Kinsbourne, 1970) et nos résultats viennent corroborer cette hypothèse. Les plus faibles capacités de séparation binaurale observées chez les enfants dyslexiques de cette étude s'accompagnent d'un déplacement attentionnel déficitaire, sachant que chez l'enfant dyslexique cet index est en lien avec ses capacités d'attention auditive. En comparaison aux normo-lecteurs, les performances des enfants dyslexiques sont significativement réduites sur les deux oreilles, laissant supposer, qu'entre les deux configurations précédemment décrites, ils correspondent plutôt à la seconde. Cependant, les réponses intrusives n'obéissent pas à ce schéma puisqu'elles sont beaucoup plus nombreuses chez les dyslexiques, plus particulièrement, même si aucune interaction ne s'est avérée être statistiquement significative, sur l'oreille droite. Leur déficit ne peut donc pas être imputé à un seul déficit verbal car dans ce cas toutes les réponses, qu'elles soient correctes ou intrusives auraient dû être moins nombreuses. Or nous observons qu'il existe des conditions attentionnelles faisant que l'enfant dyslexique peut répéter une plus grande quantité d'information de type verbal que l'enfant normo-lecteur. Lors d'une tâche de séparation binaurale, on peut expliquer ce fort taux de réponses intrusives par un défaut d'inhibition d'écoute sur l'oreille non désignée. Ainsi les résultats de cette étude sont à rapprocher avec l'absence de modulation de la prévalence de l'oreille droite au test dichotique lors du déplacement attentionnel rapportée chez l'enfant dyslexique (Hugdahl et al., 1998).

Bien qu'il soit spéculatif d'établir une relation de cause à effet entre anomalies des performances dichotiques et difficultés d'apprentissage de la lecture, deux études récentes permettent d'établir des passerelles. La première montre que parmi les enfants dyslexiques, ceux restant en échec scolaire en dépit d'une prise en charge scolaire et orthophonique et cela en dépit d'une prise en charge adaptée, sont plus enclins à présenter un déficit sur l'oreille droite lors d'une variante du test dichotique utilisant comme stimuli des syllabes de type CV où la consigne est de répéter la syllabe la mieux entendue afin de réduire l'effet de la mémoire de travail (Helland et al., 2008). Dans la seconde étude, des corrélations significatives ont été retrouvées chez l'enfant dyslexique entre les scores de reconnaissance de mots et les performances au test dichotique toujours avec des syllabes de type CV : les scores de reconnaissance de mots étaient d'autant plus faibles que les réponses de l'oreille gauche ainsi que les aptitudes dichotiques (réponses correctes sur les deux oreilles) étaient dégradées (Moncrieff & Black, 2008).

Les capacités d'intégration et de ségrégation binaurale sont très importantes dans une salle de classe où l'élève doit être capable d'extraire et surtout de privilégier le message oral délivré par l'enseignant. Toutefois, leurs mesures restent parfois encore difficiles à évaluer chez l'enfant dyslexique notamment lorsque certaines fonctions cognitives (attention mais aussi mémoire) sont altérées. Il a ainsi été observé que les aptitudes dichotiques, que ce soit pour l'attention dirigée vers l'oreille droite ou l'oreille gauche, étaient très dépendantes de l'empan mnésique de chiffres (Maerlender et al., 2004). A ce titre, la corrélation que nous avons pu mettre en évidence chez l'enfant dyslexique entre les mesures neuropsychologiques standards des capacités attentionnelles auditives et l'indice de déplacement attentionnel illustre parfaitement les interactions entre les voies auditives impliquées dans l'écoute dichotique et la maturation cognitive, d'autant plus qu'un lien significatif existait entre le score d'attention auditive et l'âge des enfants dyslexiques pour lesquels cette mesure était renseignée ( $r=0.536$ ,  $p=0.01$ ).

Ainsi, l'application en clinique du test dichotique chez des enfants, même lorsqu'il s'agit d'enfants en difficultés d'acquisition du langage écrit, permet de détecter d'éventuels déficits de traitement auditif. Ce test, dont l'utilisation remonte à plus de 50 ans et qui est tributaire de contrôle de type cognitif ou exécutif, est loin d'avoir livré tous ses secrets (Hugdahl, 2011). De futurs travaux permettront de mieux caractériser l'évolution des diverses performances dichotiques avec l'âge et l'acquisition du langage tout en comprenant mieux les bases neurales auditives sous-jacentes.

## RÉFÉRENCES

- Asbjornsen, A.E., and Bryden, M.P. (1998). "Auditory attentional shifts in reading-disabled students: quantification of attentional effectiveness by the attentional shift index," *Neuropsychologia* 36, 143-148.
- Asbjornsen, A.E., and Hugdahl, K. (1995). "Attentional effects in dichotic listening," *Brain Lang.* 49, 189-201.
- Asbjornsen, A.E., Helland T., Obrzut J.E., and Boliek, C.A. (2003). "The role of dichotic listening performance and tasks of executive functions in reading impairment: A discriminant function analysis," *Child Neuropsychol.* 9, 277-288.
- Bellis, T.J., and Ferre, J.M. (1999). "Multidimensional approach to the differential diagnosis of central auditory processing disorders in children," *J. Am. Acad. Audiol.* 10, 319-328.
- Boliek, C.A., Obrzut, J.E., and Shaw D. (1988). "The effects of hemispatial and asymmetrically focused attention on dichotic listening with normal and learning-disabled children," *Neuropsychologia* 26, 417-433.
- Broadbent, D.E. (1954). "The role of auditory localization in attention and memory span," *J. Exp. Psychol.* 47, 191-196.
- Brunswick N. and Rippon G. (1994). "Auditory event-related potentials, dichotic listening performance and handedness as indices of lateralization in dyslexic and normal readers," *Int J Psychophysiol.* 18, 265-275.
- Bryden, M.P. (1988). "An overview of the dichotic listening procedure and its relation to cerebral organization," In K. Hugdahl (Ed), *Handbook of dichotic listening: Theory, methods and research* (pp.1-43). Chichester, England: Wiley.
- Demanez, L., Boniver, V., Dony-Closon, B., Lhonneux-Ledoux, F., and Demanez, J.P. (2003a) "Central auditory disorders: some cohort studies," *Acta Oto-Rhino-Laryngol. Belgica* 57, 291-299.
- Demanez, L., Dony-Closon, B., Lhonneux-Ledoux, F., and Demanez, J.P. (2003b). "Central auditory processing assessment: a French speaking battery," *Acta Oto-Rhino-Laryngol. Belgica* 57, 275-290.
- Foster, L.M., Hynd, G.W, Morgan, A.E., and Hugdahl, K. (2002). "Planum temporale asymmetry and ear advantage in dichotic listening in developmental dyslexia and attention-deficit/hyperactivity disorder (ADHD)," *J. Int. Neuropsychol. Soc.* 8, 22-36.
- Helland, T., Asbjornsen, A.E., Hushovd, A.E., and Hugdahl, K. (2008). "Dichotic listening and school performance in dyslexia," *Dyslexia* 14, 42-53
- Hiscock, M., and Kinsbourne, M. (1982). "Laterality and dyslexia: A critical view," *Ann. Dyslexia* 32, 177-228.
- Hiscock, M., and Kinsbourne, M. (1995). "Progress in the measurement of laterality and implications for dyslexia research," *Ann. Dyslexia* 45, 249-268.
- Hugdahl, K. (1995). "Dichotic listening: probing temporal lobe functional integrity," In *Brain and Asymmetry*, ed. R.J. Davidson and K. Hugdahl. MIT Press, Cambridge, MA, 123-156.

- Hugdahl, K. (2011). "Fifty years of dichotic listening research," *Brain & Cognition* 76, 211-213.
- Hugdahl, K., Heiervang, E., Nordby, H., Smievoll, A.I., Steinmetz, H., Stevenson, J., and Lund, A. (1998). "Central auditory processing, MRI morphometry and brain laterality: applications to dyslexia," *Scand. Audiol. Suppl.* 49, 26-34.
- Hugdahl, K., Helland, T., Faerevaag, M.K., Lyssand, E.T., and Asbjornsen, A. (1995). "Absence of ear advantage on the consonant-vowel dichotic listening test in adolescent and adult dyslexics: Specific auditory-phonemic dysfunction," *J. Clin. Exp. Neuropsychol.* 17, 833-840.
- Jerger, J., and Martin, J. (2006). "Dichotic listening tests in the assessment of auditory processing disorders," *Audiol. Med.* 4, 25-34.
- Jerger, J., and Musiek, F.E. (2000). "Report on the consensus conference on the diagnosis of auditory processing disorders in school-aged children," *J. Am. Acad. Audiol.* 11, 467-474.
- Jerger, J., Chmiel, R., Tonini, R., Murphy, E., and Kent, M. (1999). "Twin study of central auditory processing disorder," *J. Am. Acad. Audiol.* 10, 521-528.
- Kershner, J., and Micaleff, J. (1992). "Consonant-vowel lateralization in dyslexic children: Deficit or compensatory development," *Brain & Lang.* 43, 66-82.
- Kershner, J., and Morton, L.L. (1990). "Directed attention dichotic listening in reading disabled children: A test of four models of maladaptative lateralization," *Neuropsychologia* 28, 181-198.
- Kimura, D. (1967). "Functional asymmetry of the brain in dichotic listening," *Cortex*, 3, 163-168.
- Kinsbourne, M. (1970). "The cerebral basis of lateral asymmetries in attention," *Acta Psychologica (Amst)* 33, 193-201.
- Korkman, M., Kirk, U., and Kemp, S. (1998). "NEPSY: a developmental neuropsychological assessment," San Antonio (TX): The Psychological Corporation.
- Lamm, O., and Epstein, R. (1994). "Dichotic listening performance under high and low lexical work load in subtypes of developmental dyslexia," *Neuropsychologia* 32, 757-785.
- Lefavrais, J. (1965). *Test de l'Alouette (révision 2005)*. Paris: ECPA.
- Maerlender, A.C., Wallis, D.J., and Isquith, P.K. (2004). "Psychometric and behavioural measures of central auditory function: the relationship between dichotic listening and digit span tasks," *Child Neuropsychol.* 10, 318-327.
- McArthur, G.M., and Bishop, D.V.M. (2001). "Auditory perceptual processing in people with reading and oral language impairments: Current issues and recommendations," *Dyslexia* 7, 150-170.
- Moncrieff, D.W. (2006). "Identification of binaural integration deficits in children with the competing words subset: standard score versus interaural asymmetry," *Int. J. Audiol.* 45, 545-558.
- Moncrieff, D.W. (2011) Dichotic listening in children: Age-related changes in direction and magnitude of ear advantage. *Brain & Cognit.* 76, 316-322.
- Moncrieff, D.H., and Black, J.R. (2008). "Dichotic listening deficits in children with dyslexia," *Dyslexia* 14, 54-75.
- Moncrieff, D.W., and Musiek, F. (2002). "Interaural asymmetries revealed by dichotic listening tests in normal and dyslexic children," *J Am. Acad. Audiol.* 13, 428-437.
- Moncrieff, D.W. and Wertz, (2008). Auditory rehabilitation for interaural asymmetry: preliminary evidence of improved dichotic listening performance following intensive training," *Int. J. Audiol.* 47, 84-97.
- Morton, L.L. (1994). "Interhemispheric balance patterns detected by selective phonemic dichotic laterality measures in four clinical subtypes of reading disabled children," *J. Clin Exp. Psychol.* 16, 556-567.
- Morton, L.L., and Siegel, S. (1991). "Left ear dichotic listening performance on consonant vowel combination and digits in subtypes of reading-disabled children," *Brain Lang.* 40, 162-180.
- Musiek, F.E. (1999). "Habilitation and management of auditory processing disorders: Overview of selected procedures," *J. Am. Acad. Audiol.* 10, 329-342.
- Oldfield, R.C. (1971). "The assessment and analysis of handedness: The Edinburgh inventory," *Neuropsychologia* 9, 97-113.
- Popescu, M.V. and Polley, D.B. "Deprivation disrupts development of binaural selectivity in auditory midbrain and cortex," *Neuron* 65, 718-731.
- Raven, J.C., Court, J.H., and Raven, J. (1984). "Manuel for Raven's progressive matrices and vocabulary scales," London: Lewis.
- Swanson, H.L. and Cochran, K.F. (1991). "Learning disabilities, distinctive encoding, and hemispheric resources," *Brain Lang.* 40, 202-230.
- von Plessen, K., Lundervold, A., Duta, N., Heiervang, E., Klauschen, F., Smievoll, A.I., Ersland, L., and Hugdahl, K. (2002). "Less developed corpus callosum in dyslexic subjects – a structural MRI study," *Neuropsychologia* 40, 1035-1044.
- Wechsler, D. (2005). *Manuel d'interprétation du WISC-IV, 4e ed.* Paris: ECPA.

For  
Digital Recorders

# Introducing

For  
USB A/D Systems

## PHANTOM POWER



### 7052PH

Measurement Mic System

7052H Type 1.5™

Titanium Diaphragm

3Hz to >20 kHz

<20 dBA > 140 dBSPL

MK224 (<14 dBA to >134 dBSPL) Optional

4048 Preamp



Superior  
IEC 1094 Type 1  
Long-term Stability  
Temperature and Humidity  
Performance

Now in Stock



Phantom  
to IEPE/ICP  
Adaptor  
Supplies 3-4 mA  
Power  
Accelerometers  
Microphones

ICP1248

A B  
C e  
O g  
u i  
S n  
t w  
i  
t  
h  
A  
C  
C  
O  
S O



MATT™  
Family

Mic  
Attenuator

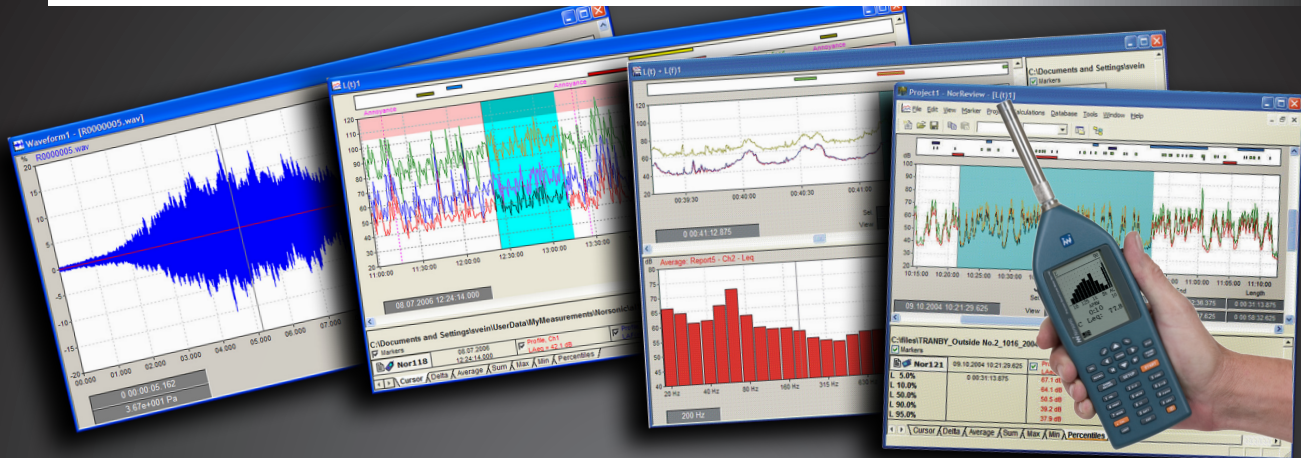
Handle Higher Sound Pressure Levels

ACO Pacific, Inc., 2604 Read Ave., Belmont, CA 94002

Tel: (650) 595-8588 FAX: (650) 591-2891 E-Mail: sales@acopacific.com

Web Site: www.acopacific.com

TM



The most complete Sound Level Meter on the market today! ➔

Specialists in Acoustic Measurement Instrumentation

**Integrated Solutions from World Leaders**

- Precision Measurement Microphones
- Intensity Probes
- Outdoor Microphones
- Sound Level Meters Type 1
- Ear Simulation Devices
- Speech Simulation Devices
- Calibrators
- Array Microphones
- Sound Quality
- Sound Intensity
- Sound Power
- Room Acoustics
- Noise Monitoring
- Dynamic Signal Analyzers
- Electro Dynamic Shaker Systems
- Doppler Laser Optical Transducers
- Laser Vibrometers
- Multi-Channel Dynamic Analyzer/Recorder



**Soundbook™**  
Designed for You:

- 2-4-8 Channel ✓
- IEC conform ✓
- PTB Approved ✓
- User friendly ✓
- General purpose ✓
- Tough (MIL) ✓
- Reliable ✓





# EFFECTS OF INTRA-TALKER DIFFERENCES ON SPEECH UNDERSTANDING IN NOISE BY YOUNGER AND OLDER ADULTS

Huiwen Goy, M. Kathleen Pichora-Fuller and Pascal van Lieshout

Department of Psychology, University of Toronto at Mississauga, 3359 Mississauga Rd N, Mississauga, Ontario L5L 1C6

## ABSTRACT

The purpose of the study was to investigate the effects on intelligibility when listening to a talker who produced speech while being exposed to different types and levels of noise. In addition, we also investigated how the word identification performance of younger and older listeners for these conditions was affected in different listening environments. A male young adult talker recorded the SPIN-R test while listening to different types and levels of noise and following different speaking instructions. Younger and older adults with clinically normal hearing from 250 to 3000 Hz were tested on these recordings in three signal-to-noise ratios. Significant word recognition differences were found between the standard talking condition and loud speech recorded in quiet, despite the fact that all sentences were equated for overall intensity. Older listeners found speech produced in noise beneficial when there was more background noise, but younger listeners did not. Clear speech in this study did not produce any benefit for listeners relative to the baseline condition. Acoustic analyses showed that intensity fluctuations within the sentences led to a higher intensity for target words in the loud speech condition relative to all other conditions, while target words in clear speech had a lower fundamental frequency compared to other conditions. Listener performance may change on a test of speech intelligibility when speech is produced under more ecologically valid conditions; however, these effects are small and may be more apparent in older adults and when task difficulty is greater due to lack of contextual support and higher levels of background noise.

## SOMMAIRE

Le but de cette étude était d'explorer l'effet sur l'intelligibilité de la parole lorsque celle-ci est produite en présence de différents niveaux et types de bruit. L'effet de ces conditions de production de la parole en présence de différents environnements d'écoute sur les performances à l'identification des mots a été examiné auprès de jeunes adultes et d'adultes plus âgés, présentant des seuils auditifs normaux entre 250 et 3000 Hz. Les phrases du test SPIN-R émises par un locuteur d'âge adulte ont été enregistrées pendant qu'il écoutait différents bruits à différents niveaux et suivait diverses instructions. Ces phrases ont été présentées à trois rapports signal-sur-bruit auprès des deux groupes de participants. Une différence significative a été notée entre le nombre de mots reconnus dans la condition de discours standard et la condition de discours enregistré à volume élevé dans un environnement de silence et ce, même en ayant normalisé le niveau de présentation de toutes les phrases. Les auditeurs plus âgés ont trouvé que le discours produit dans le bruit était plus facile à reconnaître lorsque les phrases étaient présentées avec le bruit de fond au niveau plus haut, mais ce n'était pas le cas pour les jeunes auditeurs. Pour tous les participants, le discours clairement prononcé n'a pas entraîné des performances significativement meilleures, par rapport au discours standard. Les analyses acoustiques ont montré que l'intensité était 2 dB plus haute sur les mots cibles des phrases produites en présence de bruit par rapport aux autres conditions d'enregistrement, alors que la fréquence fondamentale des mots cibles du discours clairement prononcé était plus basse que celle des autres conditions. Les performances aux mesures de reconnaissance de la parole peuvent varier en fonction des conditions dans lesquelles cette parole est produite. Ces effets sont peut-être minimes chez les jeunes adultes, mais peuvent être encore plus apparents chez les personnes âgées et lorsque la tâche de reconnaissance est plus difficile en raison du manque d'indices contextuels et des niveaux élevés de bruit de fond.

## 1. INTRODUCTION

The success of speech communication depends not only on factors related to the listener, but also on factors related to the talker. It is well-known that speech understanding in

noise is more challenging for older listeners than for younger listeners, even when older listeners have good audiograms (Dubno, Dirks, & Morgan, 1982). However, talkers may adjust their speech in an attempt to improve

communication when faced with a listener in a difficult listening situation (Smiljanić & Bradlow, 2009). Experiments have demonstrated that changes in speech production due to the talking environment may be beneficial to listeners. For example, speech produced in a noisy environment is more intelligible than speech produced in quiet, when speech is presented to listeners in a noisy environment (Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988; Pittman & Wiley, 2001). Such changes in speech may include an increase in intensity and fundamental frequency ( $F_0$ ; Letowski, Frank, & Caravella, 1993) as well as changes in articulation (Forrest, Abbas, & Zimmermann, 1986).

Little is known, however, about how younger and older listeners may differ in their word recognition performance when production of speech has been influenced by the environment. The few studies that have compared younger and older listeners have usually tested older adults with hearing loss (Picheny, Durlach, & Braidia, 1985; Schum, 1996), and older listeners with hearing loss do not always benefit from the acoustic information available to younger listeners (Ferguson & Kewley-Port, 2002). It is possible that normal-hearing older listeners may benefit more from additional cues compared to younger adults, much as they benefit more from supportive semantic context (Pichora-Fuller, 2008). Conversely, it is also possible that older adults will not benefit from additional acoustic information due to auditory processing deficits; e.g., one study showed that older adults were not able to use voice cues to reduce informational masking (Huang, Xu, Wu, & Li, 2010).

In this study, we investigated three main issues: 1) whether the word recognition performance of listeners would be affected when they heard speech produced in talking conditions other than the typical quiet conditions used for recording stimuli for speech intelligibility tests, 2) whether younger and older listeners would be differentially affected by these changes in talking conditions, and 3) which acoustic changes might underlie any changes in listener performance. We recorded a talker producing speech in different noise environments and under different speaking instructions, and presented those recordings as test sentences to listeners in various levels of background noise.

## 2. EXPERIMENT 1

### 2.1 Method

#### 2.1.1 Stimuli recording

The stimuli used were the eight equivalent sentence lists from the Revised Speech Perception in Noise Test (SPIN-R; Bilger et al., 1984). Each list consists of 25 high-context and 25 low-context sentences. High-context sentences contain a sentence-final target word that is highly predictable from the preceding phrase (e.g., Unlock the door and turn the *knob*), while low-context sentences contain an unpredictable sentence-final target word (e.g., We spoke about the *knob*).

For the present study, a new talker was selected to match the original SPIN-R talker as closely as possible on

average speaking fundamental frequency ( $F_0$ ) and speaking rate, in an effort to create a version of the test that was similar to the original except for changes in speech due to the talking environment or speaking instructions. The new talker was selected from a group of six young adult male talkers who spoke Canadian English as their first language. Candidate talkers heard the original SPIN-R sentences in List 1 presented at a level of 70 dB SPL (50 dB HL), using one loudspeaker placed at 225°; this level was chosen as the standard protocol is to present SPIN-R sentences at 50 dB SL (Bilger et al., 1984). Candidates were then asked to repeat each sentence after the talker in their normal voice, using the recordings' speaking rate and intonation as a guide when producing the sentences. An acoustic analysis of each candidate's sentences showed that talker DF was the closest match to the original talker in his average speaking  $F_0$  and speaking rate (121.7 Hz and 4.3 syllables/sec, respectively, compared to 120.5 Hz and 4.3 syllables/sec for the original talker; Kalikow, Stevens, & Elliott, 1977).

Talker DF recorded SPIN-R sentence lists 2 through 8 in different talking environments. The recordings were made while he was seated in a single-walled sound-attenuating International Acoustics Company (IAC) booth, with a Sennheiser Linear E825S microphone placed 18 cm from his lips. His speech was recorded using a Tucker-Davis Technologies System III and the Avaaz Time-Frequency Representation program running on a Dell Precision 360 computer. During the recording sessions, a hardcopy transcript of the sentences was made available to DF, while sentences were presented binaurally through Sennheiser HD265 headphones. All lists were recorded by DF while he wore headphones, so that any speech production changes due to occlusion of the ears were held constant across all conditions. After each sentence was presented, there was a two-second delay before a visual cue appeared on a computer screen to prompt DF to speak, and the rate of presentation of sound files was under the experimenter's control.

The speaking conditions used in this study are summarized in Table 1. SPIN-R lists 2, 7 and 8 were spoken in a quiet talking environment. For List 2, DF was instructed to repeat each sentence using the original talker's speaking rate and intonation (baseline condition). In List 7, he was instructed to "speak clearly" (clear speech condition), and in List 8, to "speak loudly" (loud speech condition). These two speaking styles were elicited for comparison with speech produced in a noisy environment, which may change in both articulation and intensity. Lists 3 and 4 were produced while DF listened to multi-talker babble from the original SPIN-R test, presented at 62 or 66 dB SPL, and Lists 5 and 6 were produced while he listened to speech spectrum noise matched to the babble from the original SPIN-R test, presented at 62 or 66 dB SPL. No specific speaking instructions were given in the noise conditions. Assuming that typical conversational level is about 70 dB SPL, the noise level of 62 dB SPL would match the +8 dB signal-to-noise ratio (SNR) used in the standard protocol for the

SPIN-R test (Bilger et al., 1984). The higher noise level of 66 dB SPL was chosen to create a more challenging condition which is still typical of everyday listening environments (Pearsons, Bennett, & Fidell, 1977). For lists that were spoken in noise, each SPIN-R sentence was presented to DF in quiet, followed by two seconds of noise. The noise continued while a visual cue was presented on a computer screen to prompt him to speak.

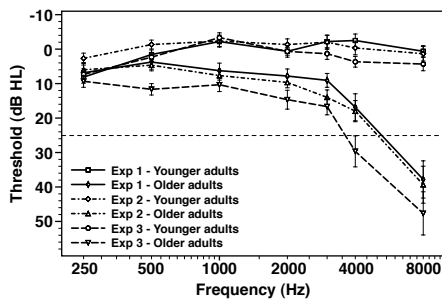
**Table 1.** Summary of speaking conditions (Babble = babble noise; SSN = speech spectrum noise).

SPIN-R List	Recording environment	Speaking instructions
2	Quiet	Follow original talker
7		Speak clearly
8		Speak loudly
3	Babble (62 dB SPL)	None
4	Babble (66 dB SPL)	
5	SSN (62 dB SPL)	
6	SSN (66 dB SPL)	

The average RMS energy of each sentence was equated to 0.05 Pa using a custom MATLAB program. Therefore, the overall sentence intensity level was controlled during the test, but local intensity fluctuations were preserved and could differ across sentences in different talking conditions. The new sentences were aligned with the original SPIN-R babble background, and a 0.6 sec 500-Hz warning tone preceded the presentation of each sentence by 1.25 sec.

### 2.1.2 Participants

Listeners were 16 younger adults who were undergraduate students from the university (mean age = 19.9 years, SD = 1.7) and 16 community-dwelling older adults who were recruited to participate in studies on healthy aging (mean age = 69.0 years, SD = 4.1). All listeners were native English speakers with pure-tone audiometric thresholds  $\leq 25$  dB HL from 250 to 3000 Hz in the test ear (Figure 1). None of the participants had heard the SPIN-R test previously. Participants all gave informed consent in compliance with the protocol approved by the institutional ethics review board, and were paid at an hourly rate.



**Figure 1.** Average pure-tone audiometric thresholds of younger and older listeners in three experiments. Error bars are standard error of the mean. The dotted line is considered to be the limit of clinically normal hearing.

### 2.1.3 Procedure

Listeners were tested on SPIN-R Lists 2 to 8 by talker DF. They also heard List 1 by the original talker for comparison with DF's baseline condition, but we did not include this list in the current analysis as we were interested in examining intra-talker differences in speech production rather than inter-talker differences. Participants listened to the eight sentence lists in two 1-hour sessions while seated in a double-walled sound-attenuating IAC booth. Sentences were presented monaurally over TDH-50P earphones at 70 dB above the participant's average pure-tone audiometric thresholds at 0.5, 1, and 2 kHz, mixed with SPIN-R babble at 0 dB SNR. Participants were instructed to report the last word of each sentence, and guessing was encouraged. There was no time limit on responding. Responses were scored by the experimenter as they were made, and participants' answers were audio-taped to enable later confirmation of the scoring. The order of talking conditions was counterbalanced across participants.

### 2.1.4 Data analysis

The data were transformed to satisfy the assumptions of the linear modelling procedure mentioned below. Scores were transformed from raw scores to rationalized arcsine units (RAU) using equations from Sherbecoe and Studebaker (2004)<sup>1</sup>.

We modelled listener performance in terms of RAU scores as a function of age group, talking condition, context condition and the interactions between these factors, using a random intercept model with a compound symmetry covariance structure.

## 2.2 Results

As shown in Figure 2, there was a significant main effect of age, with younger adults obtaining higher scores than older adults,  $F(1, 30) = 14.84, p < .001$ . There was also a significant main effect of context, with higher scores obtained for high-context sentences than for low-context sentences,  $F(1, 390) = 1249.60, p < .001$ , and a significant main effect of talking condition,  $F(6, 390) = 13.18, p < .001$ . Multiple *t*-tests with a Bonferroni correction showed that more words were correctly identified when speech was produced loudly (88%) or in the higher level of babble noise (87%) than when speech was produced in the baseline condition (83%). However, word recognition accuracy in the clear speech condition and the other three noise conditions did not differ from word recognition accuracy in the baseline speech condition.

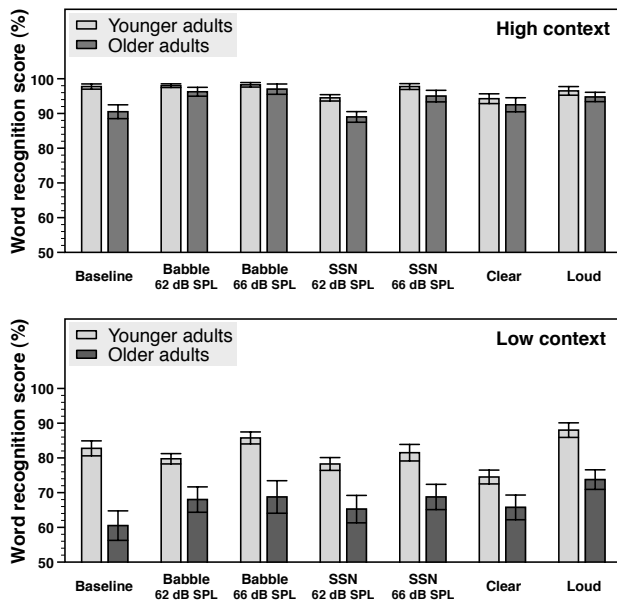
There were significant interactions between age and context,  $F(1, 390) = 34.21, p < .001$ , talking condition and

<sup>1</sup>  $\theta = \arcsin\sqrt{(X/N)} + \arcsin\sqrt{((X+1)/(N+1))}$

$$RAU = (146/\pi) \theta - 23$$

X denotes the number of correct items and N denotes the total number of items in the test set.

context,  $F(6, 390) = 3.73, p < .01$ , and age and talking condition,  $F(6, 390) = 3.54, p < .01$ . In the high-context condition, older adults performed similarly to younger adults ( $p > .05$ ), but in the low-context condition, the average score for older adults was 14 percentage points lower than the average score for younger adults ( $p < .001$ ). In the high-context condition, word identification scores for speech produced in louder babble were higher (98%) than for baseline speech (94%;  $p < .05$ ), while in the low-context condition, scores for loud speech (81%) were higher than scores for baseline speech (72%;  $p < .001$ ); word recognition scores in other talking conditions were not different from word recognition scores for baseline speech in either context. Relative to the word recognition scores for baseline speech (76%), older adults obtained higher scores when speech was produced loudly (84%) or in any of the noise conditions (82-83%) except for softer speech spectrum noise ( $p$ 's  $< .01$ ). However, word recognition for clear speech and baseline speech did not differ. In contrast to older adults, younger adults did not correctly recognize more words when speech was produced in noise or loudly, and younger adults correctly recognized fewer words in clear speech (84%) than in baseline speech (90%;  $p < .01$ ).



**Figure 2.** Mean word recognition scores for younger and older adults for high- and low-context sentences, while listening to different talking conditions in a noise environment of 0 dB SNR (SSN = speech spectrum noise). Error bars are standard error of the mean.

### 2.3 Discussion

Older adults obtained lower word recognition scores overall compared to younger adults, and there was a typical age-related difference when there was a lack of sentence contextual support.

Benefits from loud speech and speech produced in noise were only experienced by older listeners, which may

have been due in part to ceiling effects for younger listeners. Surprisingly, clear speech was not more intelligible than baseline speech for either age group, and it was actually less intelligible than baseline speech for younger listeners.

In Experiment 2, we examined whether the benefits of loud speech and speech produced in noise would be more apparent in younger adults in a more difficult listening environment. We tested a different group of younger and older listeners on five conditions (baseline speech, speech produced in babble at 66 dB SPL, clear speech and loud speech) in babble noise with an SNR of -2 dB.

## 3. EXPERIMENT 2

### 3.1 Method

#### 3.1.1 Stimuli and procedure

Participants heard five of the eight conditions from Experiment 1: baseline speech, speech in babble at 66 dB SPL, clear speech and loud speech by the new talker, and the original List 1 (which was later excluded from the analyses). Sentences were presented to participants using the same method as in Experiment 1, except that the SNR was -2 dB. The instructions, counterbalancing and scoring procedures were identical to those of Experiments 1, and data analysis was performed similarly as before.

#### 3.1.2 Participants

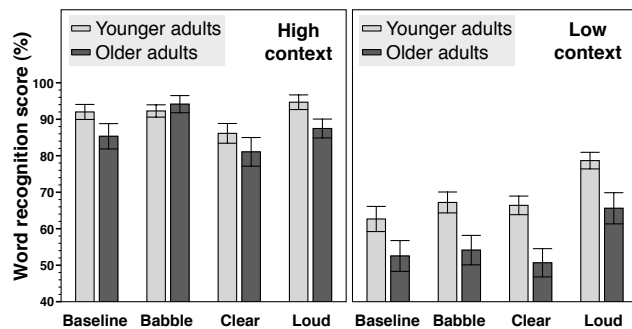
Participants were 15 younger adults who were undergraduate students (mean age = 21.5 years, SD = 2.4) and 15 community-dwelling older adults (mean age = 68.3 years, SD = 3.5). The criteria for participation were the same as in Experiment 1 (see Figure 1 for participants' average pure-tone audiometric thresholds), and participants had not participated in the previous experiment. Participants gave informed consent and were paid at an hourly rate.

### 3.2 Results

As shown in Figure 3, there was a significant main effect of age,  $F(1, 28) = 5.84, p < .05$ , with younger adults obtaining higher word recognition scores than older adults. There was a significant main effect of context,  $F(1, 196) = 692.49, p < .001$ , with higher scores for high-context sentences than for low-context sentences. There was also a significant main effect of talking condition,  $F(3, 196) = 20.36, p < .001$ . Multiple  $t$ -tests with a Bonferroni correction showed that word recognition scores were not significantly different for speech produced loudly (82%) and speech produced in babble (77%;  $p = .07$ ), while word recognition scores for these conditions were higher than scores for baseline (73%) and clear speech (71%;  $p$ 's  $< .05$ ). There was no significant difference between word recognition scores for the baseline and clear speech conditions.

There were significant interactions between age and context,  $F(1, 196) = 8.53, p < .01$ ; in the high-context condition, younger and older adults performed similarly, but

in the low-context condition, the average score for younger adults was 13 percentage points higher than for older adults ( $p < .01$ ). There was a significant interaction between talking condition and context,  $F(3, 196) = 7.36, p < .001$ . For high-context sentences, word recognition scores were higher for speech produced in babble (93%) than for baseline (89%) or clear speech (84%;  $p$ 's  $< .05$ ), but the scores were similar for loud speech and speech produced in babble. For low-context sentences, word recognition scores for loud speech (72%) were higher than scores for all other talking conditions by 11 to 15 percentage points ( $p$ 's  $< .001$ ), while scores for other conditions were not significantly different. There was a marginal interaction between age and talking condition,  $F(1, 196) = 2.62, p = .052$ . For younger adults, word recognition scores for loud speech (87%) were higher than scores for all other conditions by 7 to 10 percentage points ( $p$ 's  $< .01$ ), but scores for other conditions did not differ significantly. For older adults, word recognition scores did not differ significantly for loud speech (77%) and speech produced in noise (74%), but scores for both conditions were higher than for baseline (69%) and clear speech (66%;  $p$ 's  $< .05$ ). Word recognition did not differ for baseline and clear speech.



**Figure 3.** Mean word recognition scores for younger and older adults for high- and low-context sentences, while listening to different talking conditions in a noise environment of -2 dB SNR. (Base = baseline condition). Error bars are standard error of the mean.

### 3.3 Discussion

Similar to Experiment 1, younger adults obtained higher word recognition scores overall compared to older adults. In contrast to Experiment 1, younger adults benefited from loud speech, but clear speech was not different from the baseline condition. Older adults found both loud speech and speech produced in noise helpful.

Since listening difficulty was not equated between the two age groups in Experiments 1 and 2, the differences between age groups in their response to different talking conditions might have been due to a simple effect of listening difficulty, rather than differences in how younger and older adults used acoustic cues in different talking conditions. Therefore, in Experiment 3, we tested both age groups at their SPIN thresholds (i.e., the SNR at which

participants would obtain a 50% correct score on low-context sentences) using the same subset of conditions from Experiment 2.

## 4. EXPERIMENT 3

### 4.1 Method

#### 4.1.1 Stimuli and procedure

Participants were seated in a double-walled sound-attenuating IAC booth. To obtain the SPIN threshold of each participant, the original SPIN-R List 5 was presented at a high SNR (usually +6 dB for older adults and +4 dB for younger adults) and the original SPIN-R List 3 was presented at a low SNR (usually 0 dB for older adults and -1 dB for younger adults). List 6 was used for additional testing if the participant performed unexpectedly well or poorly in the first SNR condition, and linear interpolation was used to calculate the SPIN threshold. Younger adults had an average SPIN threshold of +2 dB SNR, whereas older adults had an average SPIN threshold of +4 dB SNR.

After each participant's SPIN threshold was obtained, the participant was tested on the same five lists as in Experiment 2 (with List 1 excluded from analyses). Sentences were presented to participants using the same method as in Experiments 1 and 2, except that the SNR was the participant's SPIN threshold. The instructions and counterbalance and scoring procedures were identical to those of Experiments 1 and 2, and data analysis was performed similarly as in Experiments 1 and 2.

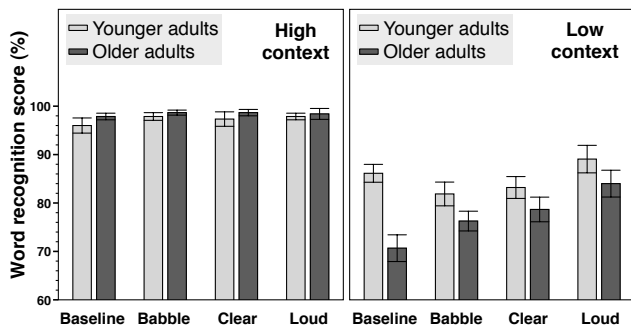
#### 4.1.2 Participants

Participants were 15 younger adults who were undergraduate students (mean age = 20.8 years, SD = 2.1) and 15 community-dwelling older adults (mean age = 71.3 years, SD = 5.8). The criteria for participation were the same as in Experiment 1 and 2 (see Figure 1 for participants' average pure-tone audiometric thresholds), and participants had not participated in either of the previous experiments. Participants gave informed consent and were paid at an hourly rate.

### 4.2 Results

As shown in Figure 4, there was no significant main effect of age,  $F(1, 28) = 2.55, p = .1$ . There was a significant main effect of context,  $F(1, 196) = 558.77, p < .001$ , with higher word recognition scores for high-context sentences than for low-context sentences. There was also a significant main effect of talking condition,  $F(3, 196) = 6.68, p < .001$ . Multiple  $t$ -tests with a Bonferroni correction showed that word recognition scores were higher for loud speech (92%) than for baseline speech (88%) and speech produced in babble (89%;  $p$ 's  $< .01$ ) and marginally more intelligible than clear speech (89%;  $p = .07$ ), but word recognition scores for other talking conditions did not differ significantly. There was a significant interaction between

age and context,  $F(1, 196) = 26.16, p < .001$ ; in the high-context condition, older adults performed similarly to younger adults, but in the low-context condition, the average word recognition score for older adults was lower than that of younger adults by 8 percentage points ( $p < .001$ ). There was a significant interaction between talking condition and context,  $F(3, 196) = 4.21, p < .01$ . For high-context sentences, there were no significant differences between talking conditions, but for low-context sentences, word recognition scores were higher for loud speech than for all other conditions by 6 to 8 percentage points ( $p$ 's  $< .01$ ), though word recognition did not differ in other talking conditions. There was no significant interaction between age and talking condition,  $F(3, 196) = 1.29$ .



**Figure 4.** Mean word recognition scores for younger and older adults for high- and low-context sentences, while listening to different talking conditions at their SPIN thresholds. (Base = baseline condition). Error bars are standard error of the mean.

### 4.3 Discussion

Unlike in Experiments 1 and 2, there was no overall age-related difference in word recognition scores after listening difficulty was equated for younger and older adults, although younger adults still outperformed older adults in the low-context condition. Both age groups benefited from loud speech but only in the low-context condition.

## 5. ACOUSTIC ANALYSIS

### 5.1 Procedure

Speech stimuli were measured using the PRAAT speech analysis program (Boersma & Weenink, 2012). The annotation function in PRAAT was used to mark the boundaries of the sentence, the final keyword and the location of the vowel or diphthong within the word. Measures of duration,  $F_0$  and intensity were extracted at each level using a custom script. Variation in intensity (intensity SD) and variation in speaking  $F_0$  ( $F_0$ SD) were measured for sentences, and F1 and F2 were measured for vowels. As the sentence lists were similar in their distributions of different vowels and diphthongs, values for each formant were averaged within each talking condition. A subset (14%) of the stimulus files was independently annotated by two researchers, and the differences between measures obtained

by the second researcher were within  $\pm 7\%$  of the absolute values of measures obtained by the first.

### 5.2 Data analysis

We focused our analyses on four main talking conditions that led to differences in word recognition: baseline speech, speech produced in babble at 66 dB SPL, clear speech and loud speech. Table 2 shows the acoustic measures taken for sentences and sentence-final target words. For both sentences and words, individual acoustic measures were modelled as a function of talking condition (four conditions), context and the interaction between these two factors, using a random intercept model with a compound symmetry covariance structure.

### 5.3 Results

For sentences, there was a significant main effect of talking condition on speaking rate,  $F(3, 192) = 17.42, p < .001$ , with clear speech having a slower speaking rate than baseline speech and loud speech ( $p$ 's  $< .001$ ). There was also a significant main effect of talking condition on mean  $F_0$ ,  $F(3, 192) = 271.56, p < .001$ ; clear speech had a lower  $F_0$  and loud speech had a higher  $F_0$  than any other talking condition ( $p$ 's  $< .001$ ). However,  $F_0$ SD did not differ significantly between conditions. There was a significant main effect of talking condition on mean intensity,  $F(3, 192) = 16.59, p < .001$ , an interaction of talking condition with context for intensity,  $F(3, 192) = 3.44, p < .05$ , and an interaction of talking condition with context for intensity SD,  $F(3, 192) = 2.78, p < .05$ . However, the largest overall intensity difference between conditions was 0.3 dB, and the largest difference in intensity SD was 0.7 dB; therefore, overall sentence-level differences in intensity are unlikely to be of practical significance. There was no main effect of context on any acoustic measure.

For target words, there was no significant effect of talking condition on word duration or formant values, but there was a significant effect of talking condition on  $F_0$ ,  $F(3, 192) = 24.37, p < .001$ , such that words spoken loudly had a higher  $F_0$  than words spoken clearly or in babble ( $p$ 's  $< .01$ ), while words spoken clearly had a lower  $F_0$  than in all other conditions ( $p$ 's  $< .001$ ). There was also a significant main effect of intensity,  $F(3, 192) = 11.36, p < .001$ ; words spoken loudly had a higher intensity than in all other conditions ( $p$ 's  $< .01$ ), but words did not differ on intensity among other talking conditions. There was no significant effect of context on any acoustic measure, and no interaction of talking condition with context.

**Table 2.** Mean acoustic measures of sentences and target words from four conditions, with standard deviations in parentheses (Base = baseline condition; Babble = babble noise at 66 dB SPL).

	Base	Babble	Clear	Loud
Sentences				
Rate (syl/s)	4.1 (0.4)	3.9 (0.4)	3.7 (0.4)	4.3 (0.5)
F <sub>0</sub> mean (Hz)	124 (9)	123 (7)	113 (5)	153 (7)
F <sub>0</sub> SD (Hz)	25.7 (19.5)	27.2 (15.2)	21.5 (13.5)	25.1 (6.8)
Intensity (dB)	68.8 (0.3)	69.0 (0.2)	68.9 (0.2)	69.1 (0.2)
Intensity SD (dB)	10.9 (1.2)	11.1 (1.3)	10.9 (0.9)	11.0 (1.2)
Target words				
Duration (s)	519 (94)	547 (129)	557 (88)	521 (94)
F <sub>0</sub> mean (Hz)	122 (25)	117 (27)	99 (12)	132 (12)
Intensity (dB)	66.8 (1.9)	66.9 (1.8)	66.2 (1.7)	68.2 (1.6)
F1 (Hz)	619 (124)	580 (132)	580 (139)	617 (135)
F2 (Hz)	1442 (358)	1522 (402)	1540 (391)	1544 (364)

#### 5.4 Discussion

Although sentences were equated on overall intensity and intensity variability was similar between talking conditions, the distribution of energy was different between talking conditions. In loud speech, the intensity of sentence-final target words was about 2 dB higher than in other conditions, which likely contributed to the higher word recognition scores in the loud speech condition. There were some benefits of speech produced in noise, but there were no clear acoustic differences between speech produced in noise and baseline speech. Listeners did not benefit from clear speech, which was characterized by a lower mean F<sub>0</sub> and a slower speaking rate compared to other talking conditions. In general, talkers produce clear speech using a slower speaking rate than in conversational speech; however, studies have suggested that speaking rate may not be the most important factor that affects intelligibility (Krause & Braida, 2002; Krause & Braida, 2004). Other properties such as F<sub>0</sub> and formant frequency may interact with the noise environment and listener characteristics to affect word recognition. One study found that F2 was raised in clear speech, which led to poorer vowel recognition by listeners with high-frequency hearing loss, presumably because the acoustic information became less audible (Ferguson & Kewley-Port, 2002). In our study, the mean F<sub>0</sub> of target words in clear speech was about 20 Hz lower than in other conditions. Since F<sub>0</sub> is an important cue for segregating a target from background noise (Oxenham, 2008), the lower F<sub>0</sub> of the target words may have enabled the background noise to mask them more effectively, resulting in lower word recognition scores.

## 6. SUMMARY OF FINDINGS

### 6.1 Effect of listener age

Younger and older adults listened to sentences that had been recorded in different talking environments, which were then played in different levels of background noise. Younger adults correctly identified more words than older adults in noise environments in which listening difficulty was not equated for the two age groups, but the two groups did not perform differently when tested at their respective SPIN thresholds. In all noise environments, the two age groups performed similarly for high-context sentences, but younger adults obtained higher scores than older adults for low-context sentences.

### 6.2 Effect of speaking condition

Loud speech was always more intelligible than baseline speech; however, in the two most difficult listening environments, speech produced in louder babble was also more intelligible than baseline speech. Clear speech did not lead to better word recognition in any of the tested conditions. For high-context sentences, there were minimal differences between talking conditions, but for low-context sentences, loud speech was always beneficial. When listening difficulty was not equated for younger and older adults, the two age groups differed on which talking conditions they found more intelligible in noise. Younger adults found loud speech more intelligible than baseline speech only in the most difficult listening environment, whereas speech produced in babble at 66 dB SPL was no different than baseline speech. In contrast, older adults found both speech produced in babble at 66 dB SPL and loud speech more intelligible than baseline speech. When tested at their respective SPIN-R thresholds, both younger and older adults found loud speech, but not speech produced in babble, more intelligible than baseline speech. The higher intensity of target words in loud speech and the lower F<sub>0</sub> of target words in clear speech may explain the benefit of loud speech over clear speech.

## 7. CONCLUSION

The purpose of the study was to investigate whether word recognition performance would improve on a standard test of word recognition when the condition in which the speech stimuli were produced were matched to the listening environment, and to investigate whether younger and older adults would benefit differently. Loud speech improved word recognition in noise but clear speech was not helpful; for listeners with normal or near-normal hearing, intensity changes in speech may be more important than changes in articulation when listening to speech in a noisy environment. The benefits of speech produced in noise were less consistent compared to the benefits of loud speech; it is possible that these benefits would have been more evident if the talker had been exposed to higher noise levels than those

used in this study. Older adults with good hearing benefited more than younger adults when listening to speech that was matched to the noise environment, but only under conditions in which listening difficulty was greater. Therefore, listener performance may change on a test of speech intelligibility when speech is produced under more ecologically valid circumstances. However, these effects may be more apparent in older adults than in younger adults and when task difficulty increases due to a lack of supportive contextual information or an increase in background noise.

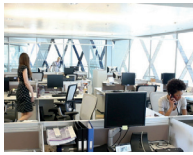
## ACKNOWLEDGMENTS

This study was funded by a Natural Sciences and Engineering Council of Canada grant awarded to M. Kathleen Pichora-Fuller. The authors wish to thank James Qi for technical assistance, and Sabrena Deonarain and Elizabeth Olsewska for their assistance in recruiting participants.

## REFERENCES

- Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research*, 27(1), 32-48.
- Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.35, retrieved 8 December 2012 from <http://www.praat.org/>
- Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of America*, 76(1), 87-96. doi:10.1121/1.391011
- Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 112(1), 259-271. doi:10.1121/1.1482078
- Forrest, K., Abbas, P. J., & Zimmermann, G. N. (1986). Effects of white noise masking and low pass filtering on speech kinematics. *Journal of Speech and Hearing Research*, 29(4), 549-562.
- Huang, Y., Xu, L., Wu, X., & Li, L. (2010). The effect of voice cuing on releasing speech from informational masking disappears in older adults. *Ear and Hearing*, 31(4), 579-583. doi:10.1097/AUD.0b013e3181db6dc2
- Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61(5), 1337-1351. doi:10.1121/1.381436
- Krause, J.C., and Braidia, L.D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America*, 112 (5), 2165-2172. doi:10.1121/1.1509432
- Krause, J.C., and Braidia, L.D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *Journal of the Acoustical Society of America*, 115 (1), 362-378. doi:10.1121/1.1635842
- Letowski, T., Frank, T., & Caravella, J. (1993). Acoustical properties of speech produced in noise presented through supra-aural earphones. *Ear and Hearing*, 14(5), 332-338. doi:10.1097/00003446-199310000-00004
- Oxenham, A. J. (2008). Pitch perception and auditory stream segregation: Implications for hearing loss and cochlear implants. *Trends in Amplification*, 12(4), 316-331. doi:10.1177/1084713808325881
- Picheny, M. A., Durlach, N. I., & Braidia, L. D. (1985). Speaking clearly for the hard of hearing. I. intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research*, 28(1), 96-103.
- Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology*, 47(SUPPL. 2), S72-S82. doi:10.1080/14992020802307404
- Pittman, A. L., & Wiley, T. L. (2001). Recognition of speech produced in noise. *Journal of Speech, Language, and Hearing Research*, 44(3), 487-496. doi:10.1044/1092-4388(2001/038)
- Schum, D. J. (1996). Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology*, 7(3), 212-218.
- Sherbecoe, R. L., & Studebaker, G. A. (2004). Supplementary formulas and tables for calculating and interconverting speech recognition scores in transformed arcsine units. *International Journal of Audiology*, 43(8), 442-448. doi:10.1080/14992020400050056
- Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Linguistics and Language Compass*, 3(1), 236-264. doi:10.1111/j.1749-818X.2008.00112.x
- Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. *Journal of the Acoustical Society of America*, 84(3), 917-928. doi:10.1121/1.396660





**Cadna R<sup>®</sup>**  
 Prediction of  
 Noise Levels inside Rooms

## New: Interior Noise calculation with CadnaR



**Cadna R<sup>®</sup>**  
 Prediction of  
 Noise Levels inside Rooms

**Out Now!**



**Cadna R<sup>®</sup>**  
 Prediction of  
 Noise Levels inside Rooms

Software & Manual

DataKustik

**Highlights:**

- Intuitive handling
- Efficient workflow
- Unique result display
- Detailed documentation
- Excellent support

### ❖ Intuitive Handling

The software is clearly arranged to enable you to build models and make simple calculations easily. At the same time you benefit from the sophisticated input possibilities as your analysis becomes more complex. Focus your time on the project and not on the software. All input and analysis features are easy and intuitive to handle.

### ❖ Efficient Workflow

Change your view from 2D to 3D within a second. Multiply the modeling speed by using various shortcuts and automation techniques. Many time-saving acceleration procedures enable fast calculations of your projects.

### ❖ Modern Analysis

CadnaR uses scientific and highly efficient calculation methods. Techniques like scenario analysis, grid arithmetic or the display of results within a 3D-grid enhance your analysis and support you during the whole planning and assessment process.

## ❖ Further informations at [www.Datakustik.com](http://www.Datakustik.com)



**Distributed (USA/Canada) by:**  
**Scantek, Inc.**  
 Sound and Vibration Instrumentation  
 and Engineering

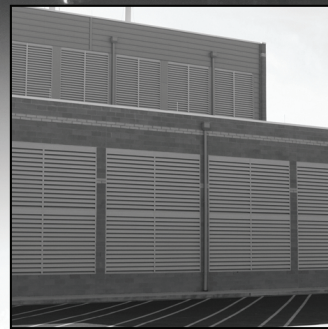
6430c Dobbin Rd Columbia, MD 21045  
 410-290-7726, 410-290-9167 fax  
 301-910-2813 cell PeppinR@ScantekInc.com  
[www.ScantekInc.com](http://www.ScantekInc.com)



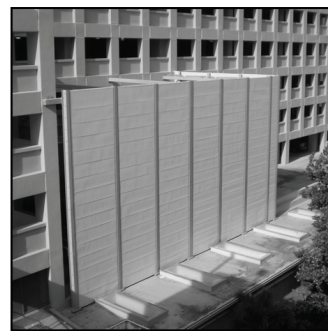
# Environmental Noise Control

*Community friendly solutions  
for chillers and cooling towers*

- Over 50 Years of Proven Design and Performance
- Independently Tested Products
- On Grade or Roof Top Applications
- Maximum Noise Reduction
- Low System Pressure Loss



**Central Energy  
Plant Louvers**



**Cooling Tower  
Barrier Wall System**



Kinetics Noise Control, Inc. | 1-800-684-2766 | [www.kineticsnoise.com](http://www.kineticsnoise.com) | [sales@kineticsnoise.com](mailto:sales@kineticsnoise.com)

Canadian Acoustics / Acoustique canadienne

Vol. 41 No. 2 (2013) - 32

# OBJETS SONORES: UNE REPRÉSENTATION BIO-INSPIRÉE, HIÉRARCHIQUE, PARCIMONIEUSE À TRÈS GRANDES DIMENSIONS UTILISABLE EN RECONNAISSANCE

Simon Brodeur et Jean Rouat

Simon.Brodeur@usherbrooke.ca, Jean.Rouat@usherbrooke.ca

Groupe de recherche en Neurosciences Computationnelle et Traitement Intelligent des Signaux (NECOTIS)  
Département génie électrique et génie informatique, Université de Sherbrooke, Sherbrooke QC Canada J1K 2R1

## ABSTRACT

The emphasis is put on the hierarchical structure, independence and sparseness aspects of auditory signal representations in high-dimensional spaces, so as to define the components of auditory objects. The concept of an auditory object and its neural representation is introduced. An illustrative application then follows, consisting in the analysis of various auditory signals : speech, music and natural outdoor environments. A new automatic speech recognition (ASR) system is then proposed and compared to a conventional statistical system. The proposed system clearly shows that an object-based analysis introduces a great flexibility and robustness for the task of speech recognition. The integration of knowledge from neuroscience and acoustic signal processing brings new ways of thinking to the field of classification of acoustic signals.

## SOMMAIRE

L'accent est placé dans cet article sur la structure hiérarchique, l'aspect parcimonieux de la représentation de l'information sonore, la très grande dimension des caractéristiques ainsi que sur l'indépendance des caractéristiques permettant de définir les composantes des objets sonores. Les notions d'objet sonore et de représentation neuronale sont d'abord introduites, puis illustrées avec une application en analyse de signaux sonores variés : parole, musique et environnements naturels extérieurs. Finalement, un nouveau système de reconnaissance automatique de parole est proposé. Celui-ci est comparé à un système statistique conventionnel. Il montre très clairement que l'analyse par objets sonores introduit une grande polyvalence et robustesse en reconnaissance de parole. Cette intégration des connaissances en neurosciences et traitement des signaux acoustiques ouvre de nouvelles perspectives dans le domaine de la reconnaissance de signaux acoustiques.

## 1 INTRODUCTION

L'organisation du système auditif reflète les structures des signaux sonores. L'accent est placé dans cet article sur une proposition de représentation par objets sonores qui vise à intégrer certaines connaissances de la physiologie et de la perception dans la conception des objets sonores.

### 1.1 Qu'est-ce qu'un objet sonore ?

Quels sont les structures et objets sonores à percevoir ? Comment ces structures ou objets sont-ils identifiés par le système auditif ? Comment mettre en oeuvre des systèmes de classification ou de reconnaissance capables d'extraire ou de reconnaître les objets sonores ? Toutes ces questions sont encore ouvertes et non résolues, cependant, plusieurs observations physiologiques et psycho-acoustiques permettent de circonscrire les réponses potentielles par l'élaboration de modèles. Une fois les réponses circonscrites, il est intéressant de confronter ces modèles avec les outils des traitements de signaux, de la théorie de l'information et de l'intelligence artificielle, afin d'affiner notre compréhension de la perception d'objets sonores et de leur représentation. Par cet article, nous répondons partiellement en proposant une repré-

sentation simple des objets sonores et nous indiquons comment il est possible d'utiliser les connaissances des neurosciences pour proposer un modèle de traitement des sons afin de réaliser une analyse des signaux qui conduit à une reconnaissance robuste et immunisée contre les perturbations extérieures. Nous posons l'hypothèse que l'architecture du système auditif est aussi liée à la structure des objets sonores, c'est-à-dire que l'évolution a fait en sorte que les structures respectives des objets sonores et l'architecture du système auditif sont intimement liées. Nous pouvons alors transférer une partie des connaissances des neurosciences vers le traitement du signal pour élaborer de meilleurs systèmes.

### 1.2 Tentative de définition d'un objet sonore

Un objet (ou entité) est une structure indépendante (autonome) dont les composantes sont liées ensemble comme faisant partie d'un même objet. Il est donc possible de manipuler les objets indépendamment les uns des autres. Cependant, une modification sur une composante aura un impact sur tous les objets élaborés à partir de cette composante. Nous considérons les objets sonores qui ne peuvent être observés directement dans le signal acoustique en raison de son caractère spatio-temporel multi-échelle. À priori, une segmentation di-

recte du signal continu introduira des erreurs d'estimation des objets sonores. En effet, les objets sonores ne sont séparables qu'une fois leurs caractéristiques perceptives obtenues. La séparation ne peut se faire que dans l'espace de la représentation auditive.

Notre environnement sonore est constitué d'éléments (composantes) que nous pouvons considérer comme indépendants, mais qui peuvent être assemblés (et donc liés) pour créer un objet, une partie d'objet ou une entité de niveau supérieur. Ces objets (ou entités) peuvent alors être manipulés de façon unitaire et indépendamment les uns des autres. Un environnement sonore est donc le résultat de la combinaison de composantes élémentaires (construites indépendamment les unes des autres, avec éventuellement des contraintes externes de conception) qui sont ensuite assemblées et liées pour créer des objets plus complexes. Au sein de ces objets, les composantes élémentaires sont liées tandis que les objets qui sont cette fois-ci plus complexes peuvent être indépendants.

Nous postulons aussi que ces parties élémentaires contribuent significativement à la perception sous forme d'événements acoustiques de type ONSET, OFFSET, clics, ou d'unités plus stables de type modulations en amplitude (AM) ou en fréquence (FM), etc. L'état des connaissances actuelles ne permet pas de connaître à priori de façon exacte ces parties élémentaires, donc nous tentons de les approcher. Un même événement sonore, c'est-à-dire dont les caractéristiques spatio-temporelles sont données, peut évoquer une perception différente suivant le bagage culturel et linguistique d'une population de personnes (e.g. occident, orient, langues tonales) ainsi que le contexte acoustique. Il est donc logique pour le concepteur de systèmes de reconnaissance de signaux acoustiques d'élaborer une stratégie permettant d'adapter la recherche de ces unités acoustiques élémentaires en fonction du contexte de l'utilisation du système. Une fois ces unités élémentaires trouvées il sera possible de les utiliser pour bâtir les représentations plus complexes des objets que nous souhaitons rechercher dans le signal.

Nous intégrons dans le système proposé une recherche automatisée des unités élémentaires, puis de leur organisation afin de générer des objets sonores composés d'une organisation hiérarchisée des unités élémentaires. Nous donnons ci-dessous la démarche qui nous permet de procéder ainsi.

### 1.3 Comment trouver les unités élémentaires acoustiques ?

Comment trouver ces unités ne connaissant pas à priori leurs formes ni leurs caractéristiques exactes, mais en disposant toutefois d'une quantité suffisante de signal acoustique ? Dans ce contexte, il n'est pas possible d'utiliser une technique d'apprentissage utilisant des données à priori étiquetées (puisque les caractéristiques exactes de ce que nous cherchons ne sont pas connues). De plus, le nombre à priori

de ces unités élémentaires n'est pas connu. Nous indiquons ci-dessous comment la connaissance du système auditif nous a orientés vers la solution choisie et qui permet de contourner ces difficultés. Pour cela nous exploitons des caractéristiques connues du cerveau :

1. L'accroissement de l'indépendance des activités neuronales lorsque l'on se déplace le long du chemin auditif, des noyaux périphériques vers le cortex auditif [1] ;
2. L'organisation hiérarchique du système auditif ;
3. La parcimonie neuronale et la très grande dimensionnalité des caractéristiques [2].

### Champs récepteurs et unités élémentaires

D'une certaine façon, le champ récepteur d'un neurone peut représenter une unité élémentaire acoustique (en réalité plusieurs neurones travaillant ensemble peuvent être interprétés en terme de champs récepteur). Plusieurs auteurs considèrent d'ailleurs qu'il y a analogie entre le filtre adapté (*matching filter*) et l'opération de filtrage et de reconnaissance réalisée par un neurone. On peut citer par exemple le travail de Daniel L. Alkon [3] qui propose en 1990 un modèle de la mémoire basé sur les caractéristiques physiologiques de neurones de l'escargot. Ce modèle considère que le champ récepteur d'un neurone s'adapte de façon à apprendre les relations de corrélations et d'anti-corrélations entre ses entrées et sa sortie. Le neurone est alors en mesure de ne répondre qu'à des configurations spécifiques de potentiels d'actions sur ses entrées. Ce thème a d'ailleurs orienté bon nombre de travaux de recherches actuels du domaine des neurosciences computationnelles. On peut citer au niveau du système auditif les travaux de l'équipe de Jos Eggermont [4] et de Shamma [5, 6], qui considèrent le lien de corrélation entre stimuli et champs récepteurs spatio-temporels de neurones. Bref, le champ récepteur d'un neurone auditif peut-être considéré comme étant équivalent en terme de traitement des signaux à une fonction dite de *base* [5]. Ceci est aussi appuyé par les travaux de Lewicki [7, 8] au début des années 2000, qui constate que la représentation optimale des sons naturels passe par une projection de ceux-ci sur des bases dont la forme se rapproche des réponses des filtres cochléaires de l'audition [9]. Il est d'ailleurs intéressant de constater que les psycho-acousticiens avaient observé ce type de réponse dès le milieu des années 1970 (e.g. [10]) et que ces mêmes réponses ont ensuite pu être obtenues uniquement à partir de critères d'optimisation de la représentation des sons [8].

Par ailleurs, il est connu depuis la fin des années 1980 que l'organisation des champs récepteurs est sur-complète et que les neurones sont organisés en couches de caractéristiques spécifiques dans le noyau cochléaire [11], pour le codage de l'information en modulation d'amplitude dans le colliculus inférieur [12, 13] et dans le cortex [14]. D'une certaine façon, les champs récepteurs peuvent être interprétés comme étant des bases parcimonieuses organisées de façon hiérarchique pour une représentation adaptée aux signaux auditifs.

Il est possible d'exploiter cette architecture pour proposer de nouvelles façons de représenter les signaux sonores [15]. Assez tôt, il a été proposé de trouver automatiquement des bases surcomplètes en tenant compte du fait qu'il devrait y avoir une indépendance statistique entre les bases [16].

En résumé, on peut poser l'hypothèse que les bases sont indépendantes et perceptivement significatives, et que l'analogie avec les champs récepteurs de micro-circuits neuronaux du système auditif est possible. En utilisant ces critères, il sera possible de représenter les éléments ou parties d'objets sonores et permettre la conception d'un algorithme de recherche automatique de ces bases. En effet, les techniques actuelles en traitement des images disposent de plusieurs outils utilisables pour la recherche et le traitement d'objets sonores. Par exemple l'analyse en composante indépendante (ICA) [17] permet de trouver des bases indépendantes et la factorisation en matrice non négative (NMF) [18, 19] des parties d'images qui sont interprétables visuellement.

### Hierarchie du système auditif

Un autre aspect très important à prendre en compte dans le traitement des sons est l'organisation très hiérarchisée du système auditif [20]. Il comprend de l'ordre de 7 à 8 noyaux nerveux traversés par l'information sonore avant de parvenir au cortex auditif. De plus, certains noyaux nerveux (e.g. le noyau cochléaire et le colliculus inférieur) sont aussi hiérarchisés. Cette organisation permet de prendre en compte le contexte acoustique [21], soit la distribution spatio-temporelle des objets sonores. Elle permet aussi d'analyser de façon multi-échelle le signal acoustique, d'accroître la robustesse aux interférences et d'introduire une invariance dans la "forme" de la représentation multi-échelle et spatio-temporelle des objets sonores. De plus, une représentation mentale d'un objet sonore pourrait être élaborée par appariement de différentes couches (ou groupes) de neurones. Cet appariement pourrait se faire par synchronie des décharges des neurones (*binding*). Ceci est d'ailleurs analogue à ce qui est observé dans le système visuel. L'hypothèse y est faite que la représentation mentale d'un objet puisse être le résultat de synchronie des décharges de sous groupes de neurones, chacun des sous-groupes représentant des caractéristiques ou parties différentes des objets [2, 22].

Le présent article présente une solution possible qui prend en compte la représentation objet des signaux telle que nous l'avons évoquée. Nous regardons maintenant comment la parcimonie et la très grande dimensionnalité des représentations peuvent être intégrées au traitement des signaux sonores.

## 1.4 La parcimonie et la très grande dimension

Peu de neurones sont actifs en même temps et leur réponse est parcimonieuse. Ceci est une conséquence indirecte du nombre considérable de neurones [2], qui conduit à une parcimonie spatiale. De plus, le premier neurone à répondre rapidement suite à la présentation d'un stimulus est celui qui encode et caractérise le mieux le stimulus. Il n'est alors pas nécessaire d'attendre la réponse des autres neurones [23] et la parcimonie est alors aussi temporelle [24]. En raison de cette parcimonie spatiale et temporelle, le codage de l'information repose sur des événements discrets (décharges des neurones) distribués dans un espace à très grande dimensionnalité (i.e. où chaque neurone correspond à une dimension). Cet aspect est aussi pris en considération dans le système de reconnaissance qui est proposé à l'aide d'un codage binaire des caractéristiques.

## 2 BREF RÉSUMÉ DE L'ÉTAT DE L'ART EN ANALYSE/RECONNAISSANCE DE SIGNAUX SONORES

La vaste majorité des systèmes contemporains de reconnaissance/classification des signaux utilise le même type de caractéristiques pour représenter les signaux. Il s'agit des coefficients cepstraux sur l'échelle de Mel (MFCC) [25]. Leur utilisation est motivée par le fait que l'échelle des fréquences y est transformée en échelle Mel afin de reproduire la distribution des bandes critiques de l'oreille. Par ailleurs, la transformation en cosinus appliquée sur le logarithme du spectre d'amplitude, pour obtenir ces coefficients, reproduit dans une moindre mesure les patrons de connectivité par inhibition latérale tels qu'ils pourraient exister au niveau du noyau cochléaire. L'opérateur de logarithme permet ici la séparation entre la source glottale et le conduit vocal. Mais le plus grand atout des MFCC est sans aucun doute le fait qu'en raison de la transformation en cosinus, ceux-ci sont relativement décorrélés, ce qui permet de meilleures performances lorsque le système de reconnaissance est basé sur les distributions statistiques des MFCC (e.g. les systèmes à base de chaînes de Markov [26]).

De nouvelles architectures ont émergé depuis les dernières années, avec des résultats prometteurs. Ces nouveaux systèmes sont organisés de façon hiérarchique et mettent l'accent sur l'obtention de caractéristiques plus proches des propriétés perceptives du signal acoustique. On peut donner comme exemple [27, 28, 29, 30]. Encore plus récemment, il a été démontré que la recherche de caractéristiques par l'utilisation d'architectures de réseaux de neurones à recherche profonde (Deep Belief Neural Networks, [31]) offre un potentiel supérieur aux systèmes de reconnaissance communément utilisés pour la parole spontanée. Le présent travail s'inscrit plutôt dans ce contexte.

### 3 APPROCHE PROPOSÉE

Nous proposons une nouvelle architecture qui prend en compte les points développés aux sections précédentes afin d'extraire une représentation des objets sonores qui respecte les propriétés importantes connues du système auditif. Nous le faisons en intégrant à la démarche des outils développés initialement pour le traitement des signaux et d'images afin de trouver, dans un premier temps, les bases surcomplètes qui seront équivalentes à des champs récepteurs placés à des niveaux hiérarchiques différents. Nous présentons une première expérience qui montre que la technique permet d'extraire effectivement des bases parcimonieuses caractéristiques des signaux utilisés durant l'apprentissage. Ensuite nous effectuons une reconnaissance de parole à l'aide de représentations parcimonieuses et de grandes dimensions. Dans notre cas, la grande dimension des données nous permet de travailler avec un codage binaire, ce qui permet d'accroître grandement la rapidité du traitement. Nous introduisons alors une reconnaissance à base de distributions de Bernoulli, adaptées à ce type de caractéristiques. Tous ces éléments permettent d'élaborer un système de reconnaissance de parole innovant.

Dans cette approche, les composantes élémentaires des objets sont les bases qui ont été trouvées pour le premier niveau hiérarchique. Les parties d'objets sonores (i.e. composantes complexes) sont représentées par les niveaux hiérarchiques supérieurs. Les objets ou parties d'objets sont représentés par des vecteurs dont les composantes sont binaires. Chaque coordonnée d'un vecteur objet (ou vecteur "partie d'objet") correspond à une composante particulière. Ces composantes sont soit élémentaires (premier niveau de la hiérarchie), soit complexes (dernier niveau). Une composante (ou partie d'objet) est considérée comme appartenant à un objet sonore si la coordonnée du vecteur qui lui est associée est différente de zéro.

Pour ce travail, l'accent est mis sur l'obtention de représentations objets à partir de caractéristiques dérivées des enveloppes des signaux cochléaires. Ceci permet de capturer de façon générique les formes les plus courantes de modulation en amplitude (AM) et en fréquence (FM), ainsi que les transitoires et les configurations de formants. La représentation d'entrée aux systèmes qui sont étudiés est le cochléogramme (illustré à la figure 1, page 4). On cherchera à exploiter l'information contenue dans les différents patrons de modulation spectro-temporelle locale pour caractériser les objets sonores. Dans la suite de l'article, un traitement par blocs/fenêtres et la prise en compte de l'aspect spatial référeront au fait que le cochléogramme sera interprété comme étant une image, par sa nature bidimensionnelle (i.e. représentation temps-fréquence).

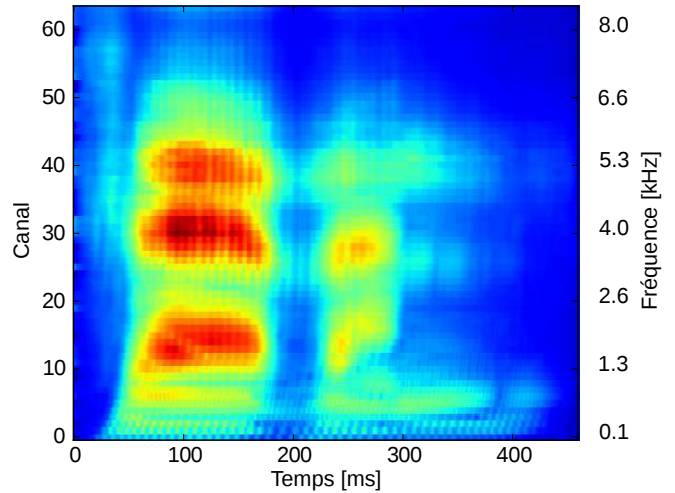


FIGURE 1 – Le cochléogramme est une représentation spectro-temporelle du signal acoustique. L'exemple correspond ici à la prononciation du mot anglais *seven* par un locuteur homme. Chaque canal correspond à la sortie d'un filtre cochléaire passe-bande et expose les caractéristiques de modulation temporelle propre à une bande de fréquence. Les paramètres des filtres sont dérivés d'observations neurophysiologiques de la cochlée et d'observations psychoacoustiques, favorisant une plus grande résolution temporelle que spectrale.

#### 3.1 Projection hiérarchique par dictionnaire

La formulation mathématique de la projection hiérarchique est la suivante : soit  $\mathbf{S}^{(h)}$  un ensemble de  $n$  signaux de dimension  $N$ , en entrée au niveau hiérarchique  $h$ , i.e.  $\mathbf{S}^{(h)} = [\mathbf{s}_1 \dots \mathbf{s}_n] \in \mathfrak{R}^{N \times n}$ . Soit  $\mathbf{D}^{(h)}$  un dictionnaire de  $K$  bases de dimension  $N$ , au niveau hiérarchique  $h$ , i.e.  $\mathbf{D}^{(h)} = [\mathbf{d}_1 \dots \mathbf{d}_K] \in \mathfrak{R}^{N \times K}$ . La dimension  $N$  peut varier selon le niveau de la hiérarchie pour effectuer une réduction de dimension, ou une expansion pour obtenir une représentation surcomplète. La projection des coefficients d'entrée de l'étage inférieur sur le dictionnaire produit un nouvel ensemble de coefficients  $\mathbf{C}^{(h)} = [\mathbf{c}_1 \dots \mathbf{c}_n] \in \mathfrak{R}^{K \times n}$ , comme montré à l'équation (1) et valide pour  $h > 0$ . La pseudo-inverse généralisée (Moore–Penrose) permet d'approximer  $(\mathbf{D}^{(h)})^{-1}$ , car  $\mathbf{D}^{(h)}$  est une matrice rectangulaire. Pour éviter les problèmes d'instabilité numérique, la pseudo-inverse  $(\mathbf{D}^{(h)})^+$  sera toutefois calculée par décomposition en valeurs singulières [32], plutôt que la forme directe d'optimisation au sens des moindres carrés. La transposée du dictionnaire est définie par  $\mathbf{D}^{(h)T}$ .

$$\mathbf{C}^{(h)} = (\mathbf{D}^{(h)})^+ \cdot \mathbf{S}^{(h-1)} = (\mathbf{D}^{(h)T} \mathbf{D}^{(h)})^{-1} \mathbf{D}^{(h)T} \cdot \mathbf{S}^{(h-1)}$$

pour tout niveau abstrait, où  $h > 0$  (1)

La projection hiérarchique possède un aspect spatial, où les fenêtres adjacentes sont concaténées et le vecteur résultant projeté sur l'étage supérieur. La reformulation selon les coordonnées  $(i, j)$  dans le référentiel local pour chaque étage est donnée à l'équation (2). Au premier niveau  $h = 0$ , les signaux d'entrée correspondent à des fenêtres  $\mathbf{W}^{(i,j)} \in \mathbb{R}^{L_C \times L_T}$  distribuées spatialement sur la représentation spectro-temporelle, et converties sous forme de vecteurs colonnes  $\mathbf{X}^{(i,j)} \in \mathbb{R}^{L_C \cdot L_T}$ . Les constantes  $L_C$  et  $L_T$  correspondent respectivement au nombre de canaux et d'échantillons temporels couverts par chaque fenêtre. Pour tout niveau  $h > 0$ , la projection est appliquée sur la concaténation des coefficients des projections du niveau inférieur  $h - 1$ . Les constantes  $M_{(h)}$  et  $N_{(h)}$  correspondent au nombre de fenêtres de projection adjacentes considérées respectivement sur l'axe fréquentiel et temporel.

$$\mathbf{C}_{(h)}^{(i,j)} = \begin{cases} \mathbf{D}_{(0)}^+ \cdot \mathbf{X}^{(i,j)} & \text{si } h = 0, \\ \mathbf{D}_{(h)}^+ \cdot \left\| \left\|_{i=0}^{M_{(h)}-1} \right\| \left\|_{j=0}^{N_{(h)}-1} \right\| \mathbf{C}_{(h-1)}^{(i,j)} & \text{si } h > 0. \end{cases} \quad (2)$$

Le symbole  $\left\| \left\| \right\|$  définit l'opération de concaténation sur  $P$  matrices de coefficients adjacentes :

$$\left\| \left\|_{i=0}^{P-1} \right\| \mathbf{C}^{(i)} = [\mathbf{C}^{(0)} \quad \mathbf{C}^{(1)} \quad \dots \quad \mathbf{C}^{(P-1)}] \quad (3)$$

La figure 2 montre plus intuitivement l'aspect spatial de cette projection. Par souci de simplicité, le chevauchement entre les fenêtres n'est pas illustré.

La projection hiérarchique permet aussi d'allouer un plus grand contexte temporel ou spectral selon le fenêtrage et le chevauchement choisi. Pour éviter les discontinuités causées par le fenêtrage initial  $\mathbf{W}^{(i,j)}$  au niveau du cochléogramme, il est possible d'introduire un chevauchement temporel et spectral. Ceci permet de mieux couvrir les caractéristiques du signal avec un nombre limité de bases, au profit d'une représentation de sortie contenant plus de coefficients et où certaines dimensions peuvent devenir fortement corrélées.

### 3.2 Analyse en composantes indépendantes

La qualité des bases obtenues par l'algorithme d'apprentissage non-supervisé du dictionnaire  $\mathbf{D}^{(h)}$  pour chacun des étages  $h$  est ultimement mesurée par l'effet sur la performance par exemple d'un système de classification. Toutefois, l'interprétation visuelle des bases (i.e. caractéristiques spatio-temporelles extraites) est pourtant importante pour valider que la projection hiérarchique permet bien l'extraction

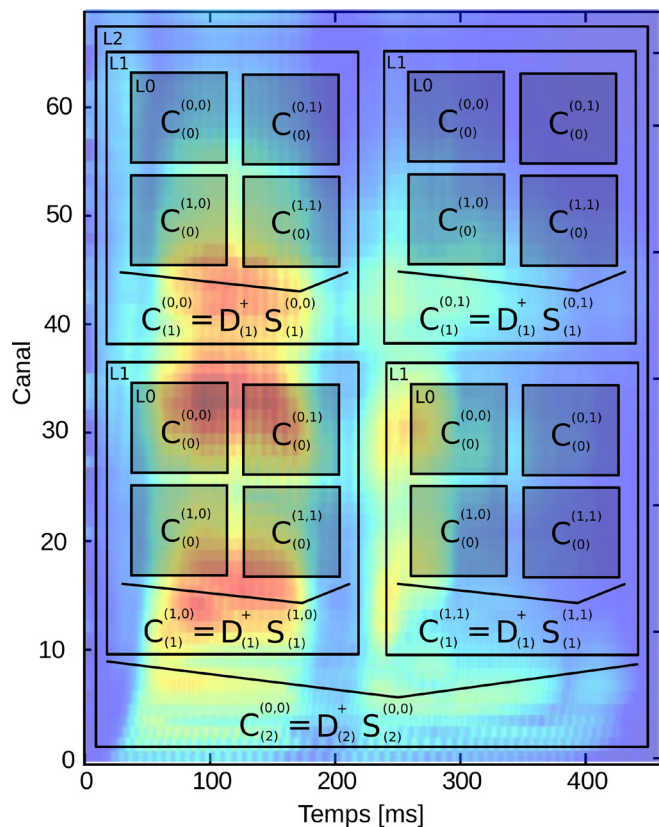


FIGURE 2 – Projection hiérarchique par projection successive sur des dictionnaires. L'aspect spatial est considéré par la concaténation des projections adjacentes du niveau inférieur, visant à extraire des caractéristiques couvrant de plus en plus de contexte fréquentiel et temporel. Dans l'exemple donné,  $M_{(h)} = N_{(h)} = 2$  pour  $h = \{1, 2\}$ , soit la concaténation de 4 fenêtres adjacentes à chaque niveau. Les contours figuratifs des fenêtres sont montrés sur le cochléogramme. Il est à noter qu'en réalité, les niveaux L1 et L2 ne couvriront jamais plus que l'aire définie par le niveau L0 (en ombragé).

de structures complexes dans le signal. L'analyse en composantes indépendantes (ICA) [33] est une méthode de décomposition linéaire (voir Équation 1) favorisant une représentation par objets d'un signal en forçant l'indépendance entre les composantes des objets. Cette décomposition définit des bases en se basant donc sur un critère de maximisation de leurs indépendances statistiques. Comparativement à l'analyse en composantes principales (PCA), il s'agit d'un critère plus strict que la décorrélation, car les moments d'ordres supérieurs à 2 sont considérés. Pour cette raison, il y a une contrainte que les composantes indépendantes doivent posséder une distribution non-gaussienne.

L'implémentation FastICA [34] a été utilisée pour les expériences. L'avantage principal est qu'il n'y a pas de méta-paramètres à choisir, comme une constante de régularisation définissant le compromis entre la parcimonie (coefficients ou

bases) et l'erreur de reconstruction. Seule une fonction de contraste pour l'approximation de la négentropie est requise, mais dont le choix n'est important que pour optimiser la performance de l'algorithme selon le type de non-gaussianité des composantes [34]. En fait, tant qu'il y a critère de minimisation de l'information mutuelle, les caractéristiques extraites seront localisées en temps et en fréquence si l'apprentissage est effectué sur des sons naturels ou des signaux de parole [7, 35].

Le problème d'estimation des sources par ICA devient plus complexe si le nombre de composantes indépendantes est supérieur au nombre de mélanges observés, car le processus de mélange est non-inversible en raison d'une perte d'information [17]. Il s'agit de la situation où la représentation est sur-complète. Pour la présente architecture, chaque étage  $h$  de la hiérarchie produit un vecteur de coefficients dont la dimension est inférieure à celle de la sortie de l'étage précédent. Ceci est principalement dû à la concaténation des projections adjacentes. Il s'agit alors du cas de représentation sous-complète. Un des avantages de cette situation est qu'en considérant l'aspect de parcimonie, il y a capacité inhérente à classifier les signaux d'entrée [36]. Pour une tâche d'extraction de caractéristiques, il est alors possible de compresser l'information sans explosion du nombre de dimensions. Alternativement, une réduction de dimension des vecteurs d'entrée est recommandée dans le cas de représentation sous-complète [37], mais l'utilisation de l'analyse par composantes principales (PCA) reste mitigée. Sachant l'ambiguïté de ICA au niveau de la variance des composantes (i.e. assumée fixe à 1), les composantes indépendantes ne sont pas forcément contraintes dans le sous-espace défini par PCA [38]. La projection hiérarchique proposée vise avant tout à décomposer le signal en composantes parcimonieuses, mais sans la contrainte de devoir reconstruire le signal. Les différents problèmes énoncés précédemment sont alors de moindre importance comparativement à une application en codage de la parole.

## 4 EXPÉRIENCE 1 : RECHERCHE NON-SUPERVISÉE DE BASES

L'apprentissage non-supervisé permet d'obtenir une représentation naturellement adaptée au signal à modéliser ou reconnaître. Peu d'effort manuel est alors requis dans le choix des paramètres optimaux du système. Dans cette expérience, les types de bases obtenues lorsque l'apprentissage est effectué sur des catégories de sons différentes (e.g. parole et musique) ont été comparés. La base de données TIMIT [39] consistant en 330 minutes de parole continue (sous forme de phrases) a été utilisée pour la catégorie de parole. Pour la catégorie de musique, 73 minutes de musique classique orchestrale [40] ont été utilisées. Enfin, pour la catégorie des sons naturels, 60 minutes d'enregistrement d'une plage tropicale [41] et 60 minutes d'enregistrement d'une forêt de mon-

tagne [42] ont composé la base de données d'apprentissage spécifique.

### 4.1 Description du système

Un filtre de pré-accentuation découlant de la réponse spectrale observée physiologiquement chez l'humain [43] est d'abord utilisé pour rehausser les moyennes fréquences. Une analyse spectro-temporelle est ensuite effectuée par un banc de 64 filtres Gammatone [44]. Le banc de filtres a été corrigé pour le décalage de phase entre les canaux, ce qui fait qu'une impulsion glottale produira une réponse instantanée sur tous les canaux. La plage de fréquence couverte est de  $[0, 8000 \text{ Hz}]$  et l'espacement des filtres est linéaire sur l'échelle de Mel [45]. La largeur de bande des filtres est grande pour privilégier la résolution temporelle à la sortie des canaux, qui favorisera une meilleure modélisation des transitoires tout en évitant la résolution de chacune des harmoniques lors des segments voisés. L'étape suivante dans le calcul du cochléogramme est une rectification simple-alternance suivie d'une compression par une racine cubique [46]. L'usage d'une compression fortement non-linéaire offre l'avantage d'augmenter radicalement le contraste entre les canaux de faibles amplitudes. Les modulations d'amplitude en moyenne et haute-fréquence (où l'énergie est souvent moindre) deviennent alors plus comparables à celles en basse-fréquence. Finalement, un filtre Butterworth passe-bas d'ordre 1, avec une fréquence de coupure à 40 Hz, permet le lissage du spectre et réduit ainsi l'effet des impulsions glottales sur les caractéristiques de modulation d'amplitude extraites par l'algorithme d'analyse en composante indépendante (ICA).

Une intégration du contexte spatio-temporel permet d'extraire des caractéristiques de plus en plus complexes. Sans poser d'hypothèses sur la nature des composantes des objets, mais en augmentant simplement le contexte spatial et temporel, des structures cohérentes et pertinentes sont apprises par les bases. Elles représentent des parties élémentaires d'objets sonores. La figure 3 montre en quoi une projection hiérarchique (sans chevauchement) exploite l'augmentation du niveau d'abstraction, ce qui permet la considération d'un plus large contexte spatio-temporel dans la représentation d'un signal de parole.

Pour cette expérience, les paramètres utilisés pour effectuer la projection hiérarchique sont les suivants : Le nombre d'étages de projection égal à 3, la taille  $K$  du dictionnaire à chaque étage étant respectivement de 128, 256, et 256. Il y a un fenêtrage initial  $\mathbf{W}^{(i,j)}$  de  $L_C = 16$  canaux par  $L_T = 40$  ms, sans chevauchement. Le cochléogramme comporte 64 canaux avec une fréquence d'échantillonnage de 1000Hz. Il y a concaténation de  $M_{(h)} = 2$  blocs au niveau spectral, et  $N_{(h)} = 3$  bloc au niveau temporel. Cette configuration fera en sorte que le dernier niveau de projection couvrira 64 canaux par 360 ms. L'apprentissage non-supervisé du diction-



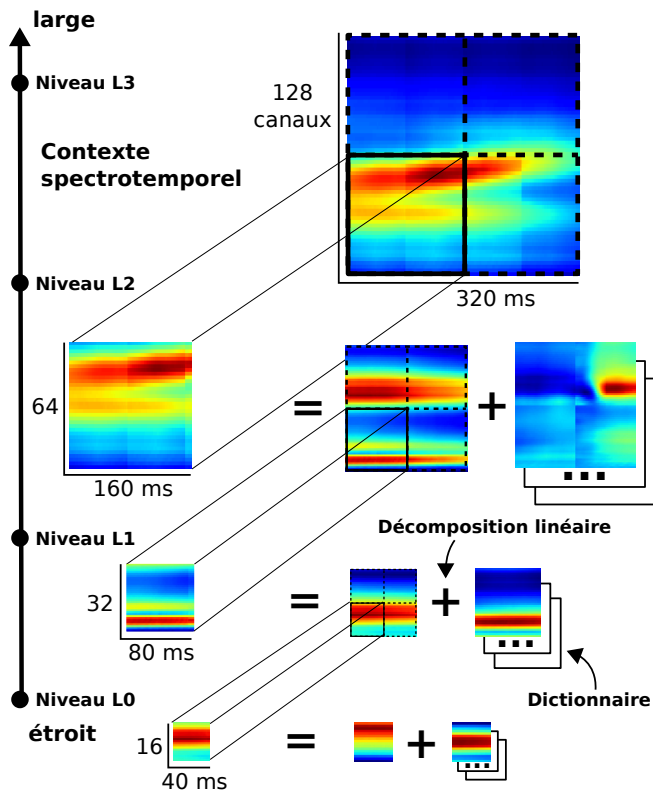


FIGURE 3 – Projection hiérarchique jointe dans l’axe temporel et fréquentiel pour des dictionnaires (ensembles de bases) de niveau L0, L1, L2 et L3. Par exemple, le dictionnaire L2 est basé sur les projections du niveau inférieur L1, lui-même basé sur les projections du niveau L0. Une grille spatiale de taille 2x2 est ici utilisée pour joindre les coefficients des 4 blocs de projection adjacents. Chaque bloc de projection est une composition linéaire (somme ou négation) des prototypes définis dans le dictionnaire spécifique à l’étage. Avec cette méthode de projection, le dictionnaire finira par définir des objets complexes et analogues à des entités syllabiques couvrant un contexte spectro-temporel de plus en plus large.

naire est effectué par analyse en composantes indépendantes (ICA). Les étages L0 et L1 sont entraînés avec 100,000 vecteurs de coefficients, tandis que pour l’étage L2, le nombre est limité entre 20,000 et 50,000 vecteurs d’apprentissage. Connaissant le grand contexte temporel des bases de haut-niveau, le nombre d’exemples pour l’entraînement est ultimement limité par la taille spécifique de la base de données.

L’entraînement des dictionnaires selon les étages de la hiérarchie est effectué de façon itérative, donc successivement du niveau L0 jusqu’au niveau L2. Ceci montre les contraintes de dépendance avec les étages inférieurs.

## 4.2 Résultats

Les bases obtenues après apprentissage modélisent bien les caractéristiques propres aux différentes catégories de sons,

comme il est illustré à la figure 4 (page 7). Peu importe le type de sons, il y a représentation parcimonieuse du signal à chaque étage (non montré). Cette propriété est évaluée objectivement par la mesure de kurtosis (moment centré d’ordre 4) du vecteur de coefficients. Peu importe le type de sons, il y a aussi représentation par objets du signal, comme révélée par une inspection visuelle et subjective des bases. Avec plus de données d’apprentissage pour chacune des catégories, des bases supplémentaires de plus haut-niveau (i.e. plusieurs secondes) auraient pu définir des patrons de modulation spectro-temporelle encore plus complexe.

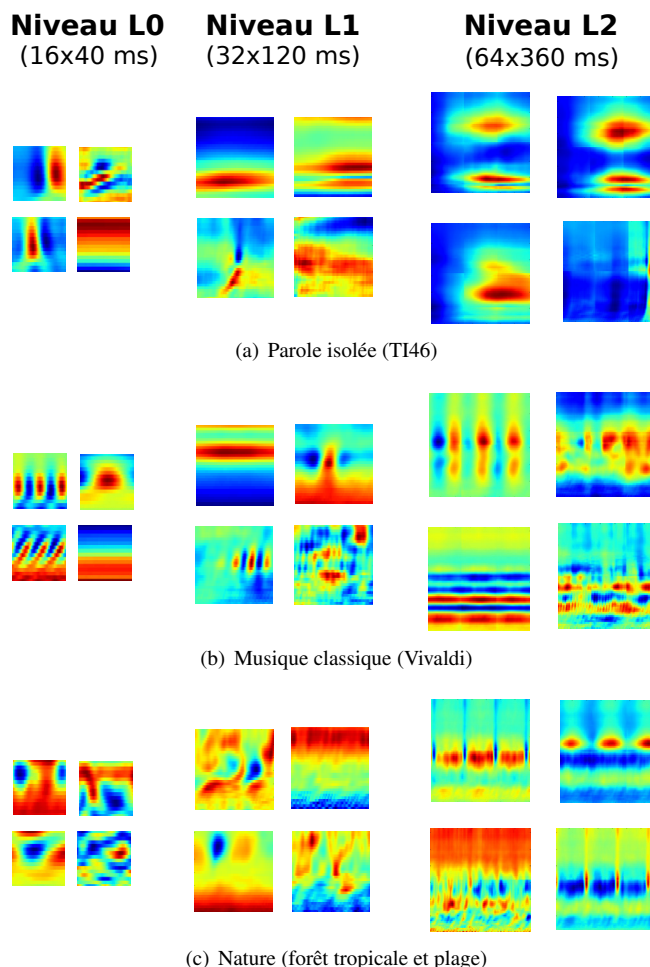


FIGURE 4 – Bases apprises pour différents types de sons naturels : (a) parole isolée, (b) musique classique et (c) sons de nature. On remarque les différences notables des bases apprises entre les catégories de sons, peu importe le niveau dans la projection hiérarchique. Les bases de haut-niveau apprises pour la parole font ressortir des configurations de formants, alors que pour la musique, il s’agit plutôt d’harmoniques constantes ou de séquences de tonalités brèves. Pour les sons de nature, les bases sont aussi très variées.

## 5 EXPÉRIENCE 2 : RECONNAISSANCE DE MOTS ISOLÉS

Pour un système de reconnaissance de forme, il devrait idéalement y avoir un passage graduel de l'apprentissage non-supervisé à un apprentissage supervisé. Alors qu'une certaine ambiguïté sur la nature du signal est permise pour l'étape d'extraction de caractéristiques, une décision stricte devra être prise à l'étape de classification. Dans cette expérience, une représentation dérivée de la projection hiérarchique est utilisée pour construire un système de reconnaissance de mots isolés basé sur une modélisation statistique par modèle de Markov caché (HMM). La base de données TI46 [47] est couramment utilisée pour tester les systèmes automatiques de reconnaissance de type mots isolés et dépendant du locuteur. La base de données contient seulement 46 classes (alphabet, chiffres et commandes). On dispose de 16 locuteurs (8 hommes et 8 femmes) et de 26 prononciations par locuteur par mot. Pour les ensembles d'entraînement et de test, on dispose respectivement de 10 et 16 prononciations par locuteur. Le partitionnement standard pour cette base de données a été utilisé. Pour tester la capacité de généralisation, le système de reconnaissance est entraîné de façon à être indépendant du locuteur.

Le HMM est un modèle statistique générateur d'une séquence d'observations basé sur un espace d'états et une topologie de transitions [26], communément utilisé en reconnaissance de la parole. Un modèle de mot entier et une structure gauche-droite à 16 états ont été utilisés pour chacune des classes. Lors de l'évaluation d'un signal acoustique contenant un mot (mais dont l'identité reste à déterminer), le modèle ayant le maximum de vraisemblance à avoir généré la séquence d'observations est retenu comme le mot prononcé le plus probable.

### 5.1 Description du système proposé

Les différents étages composant l'architecture proposée sont illustrés à la figure 5 (page 8), qui met l'emphase sur l'évolution de la dimension des représentations selon les étages. Il y a explosion des dimensions à la sortie d'une projection hiérarchique, mais la parcimonie de la représentation assure une activation restreinte à certaines dimensions seulement. Ceci permettra d'augmenter la robustesse au bruit du système, considérant de plus que les bases auront été adaptées aux caractéristiques de la parole.

Il doit être possible à chaque intervalle de temps  $\frac{1}{F_{sous}}$  dans le signal de parole de définir un vecteur de caractéristiques multi-échelle pour la modélisation statistique (e.g. modèle de Markov caché). La constante  $F_{sous}$  est la fréquence d'échantillonnage du vecteur de caractéristiques correspondant à une observation discrète. La séquence d'observations produite sert ensuite d'entrée à l'étape de classification. La figure 6 (page 9) illustre comment cette opération est effectuée

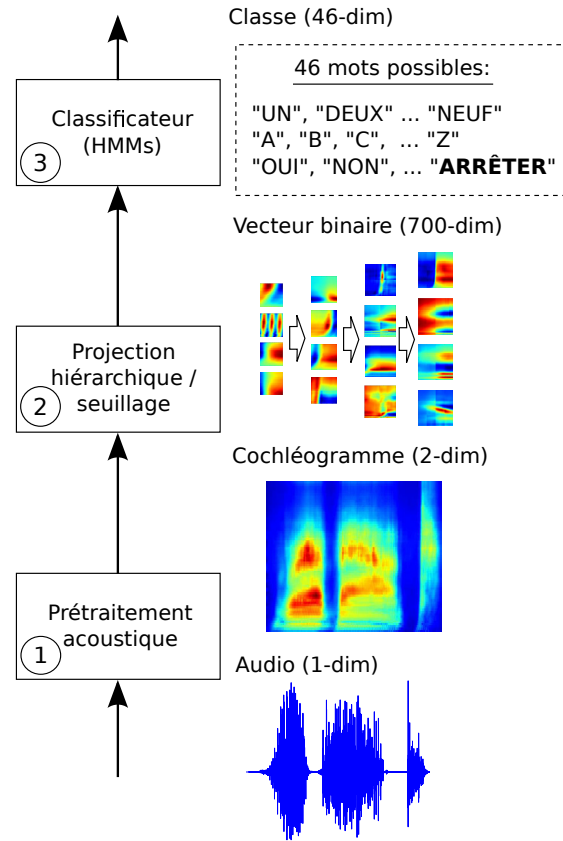


FIGURE 5 – Étapes de traitement dans l'architecture proposée, soit le prétraitement acoustique, la projection hiérarchique et la classification. La dimension des représentations augmente jusqu'à la sortie de l'étape de classification, où l'identité du mot dans le signal sonore est déterminée.

pour produire un vecteur binaire, parcimonieux et à grande dimension. La projection hiérarchique est efficace en termes de calculs dans le sens où des techniques de programmation dynamique permettent d'éviter le recalcul d'une même projection (bloc ou fenêtre) si les mêmes coefficients sont utilisés par plusieurs projections d'ordre supérieur.

L'expression générale d'une modélisation acoustique par modèle de mixture est donnée à l'équation (4). La densité de probabilité  $p(\mathbf{x})$  découle de la contribution de  $M$  mixtures, décrites par une probabilité a priori  $p(i)$  et une vraisemblance  $p(\mathbf{x}|i)$ .

$$p(\mathbf{x}) = \sum_{i=0}^{M-1} p(i) p(\mathbf{x}|i) \quad (4)$$

avec les contraintes  $\sum_{i=0}^{M-1} p(i) = 1$  et  $p(i) \in [0, 1]$

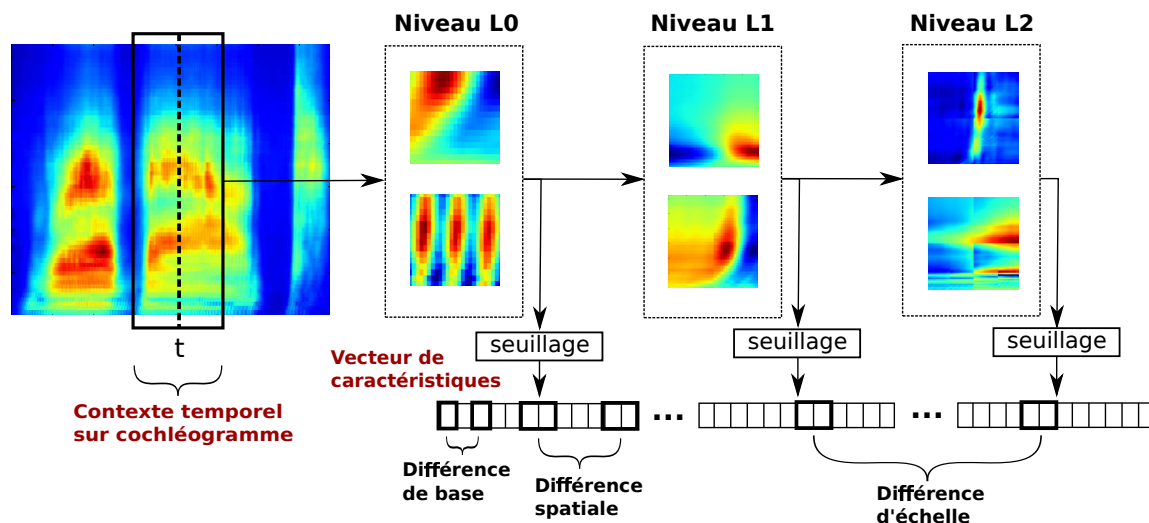


FIGURE 6 – Création du vecteur de caractéristiques à partir de la projection hiérarchique. À partir du cochléogramme (représentation temps-fréquence), il y a création d’un vecteur colonne représentant les caractéristiques de différentes bases, sur plusieurs positions spatiales et échelles spatio-temporelles d’analyse. Il est possible d’appliquer un traitement spécifique sur les représentations de sortie de chacun des étages, visant à favoriser la parcimonie et la production de coefficients binaires par des mécanismes de seuillage et de compétition.

Avec l’utilisation de représentations binaires parcimonieuses, les mixtures de gaussiennes (GMMs) ne sont plus adaptées pour la modélisation acoustique dans une architecture HMM, car elles modélisent des variables aléatoires continues et non discrètes. Le cas d’une mixture de fonctions gaussiennes multivariées à  $n$ -dimensions est défini à l’équation (5). Les moments de premier ordre  $\mu_i = E(X)$  et de deuxième ordre  $\Sigma_i = \text{Cov}(X)$  paramétrisent chacune des  $M$  mixtures.

$$p(\mathbf{x}|i) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mu_i)^T \Sigma_i^{-1} (\mathbf{x}-\mu_i)} \quad (5)$$

Les mixtures de Bernoulli (BMMs) offrent une alternative appropriée pour construire un modèle générateur d’observations binaires. Cette divergence par rapport aux systèmes conventionnels de reconnaissance de la parole permet de définir un espace de paramètre, au niveau du modèle statistique, qui disposera des mêmes propriétés que la représentation d’entrée : les paramètres des mixtures seront de grande dimension et parcimonieux.

Prenons le cas spécial d’une distribution de Bernoulli de dimension  $N$ , où les dimensions sont considérées comme statistiquement indépendantes. La fonction de probabilité est définie à l’équation (6), où  $x_n \in \{0, 1\}$  est l’élément à la dimension  $n$  du vecteur binaire  $\mathbf{x}$  de dimension  $N$ , et  $p_{i,n} \in [0, 1]$  est le paramètre de la dimension  $n$  associé à la mixture  $i$ . Ce dernier reflète intuitivement la probabilité moyenne d’avoir une activation positive (i.e. 1) sur une certaine dimension.

$$p(\mathbf{x}|i) = \prod_{n=1}^N p_{i,n}^{x_n} (1 - p_{i,n})^{1-x_n} \quad (6)$$

L’entraînement des paramètres des mixtures et du modèle markovien (e.g. probabilités de transitions) est fait par l’algorithme Expectation-Maximisation (EM). Comme les paramètres  $p_{i,n}$  correspondent à la probabilité moyenne d’apparition sur chacune des dimensions, la mise à jour est identique à celle utilisée pour le paramètre de moyenne des mixtures de gaussiennes. Pour un processus de Bernoulli multivarié, on fait l’hypothèse d’indépendance statistique de chacune des dimensions. Dans le cas où les caractéristiques d’entrée découlent d’une analyse par composantes indépendantes (ICA), cette hypothèse est réaliste.

L’utilisation de mixtures de Bernoulli est relativement rare en reconnaissance de forme, et se concentre sur les travaux de quelques chercheurs seulement. Les mixtures de Bernoulli ont d’abord été investiguées pour la reconnaissance d’images binaires [48, 49, 50]. L’effet des paramètres initiaux sur la convergence a ensuite été étudié [51]. Finalement, l’intégration avec un HMM a été effectuée [50, 52], mais toujours en reconnaissance d’image et non de parole.

Les avantages des mixtures de Bernoulli dans l’application présente sont nombreux. Premièrement, les paramètres d’une mixture de Bernoulli définissent essentiellement des prototypes dont la visualisation/interprétation est facilement concevable (surtout dans le cas du traitement d’image [e.g. 48]). Deuxièmement, un processus de Bernoulli considère seulement la moyenne, et ignore la variance. Il y a donc

une meilleure convergence, car nul besoin de techniques de seuillage des paramètres des mixtures pour éviter la surspécialisation (i.e. variance nulle ou concentration autour d'une seule observation). L'initialisation des paramètres est aussi moins susceptible à porter problème dans le cas des mixtures de Bernoulli, où quelques techniques simples permettent d'éviter les configurations pathologiques de paramètres [51]. Enfin, la parcimonie dans l'espace des paramètres peut être forcée si le nombre de dimensions ou variables latentes est large, dans le cas d'un modèle générateur par variables latentes binaires [53]. Il y a donc un potentiel intéressant pour conserver la similitude avec les représentations d'entrée, qui sont parcimonieuses et à haute-dimensionnalité.

Pour l'extraction des caractéristiques, les paramètres sont légèrement différents de l'expérience précédente portant sur l'émergence de bases non-supervisées pour diverses catégories de sons. La taille  $K$  du dictionnaire pour chacun des 3 étages est respectivement de 64, 128, et 256. L'entraînement utilise maximalelement 25,000 exemples de vecteurs de coefficient lors de l'apprentissage non-supervisé des bases avec l'algorithme d'analyse en composantes indépendantes (ICA). La fréquence d'échantillonnage  $F_{\text{sous}}$  du vecteur de caractéristiques est de 100 Hz. Il y a un fenêtrage initial  $\mathbf{W}^{(i,j)}$  de  $L_C = 32$  canaux par  $L_T = 40$  ms, avec 50% de chevauchement temporel et spectral. Il y a concaténation de  $M_{(h)} = 2$  blocs au niveau spectral, et  $N_{(h)} = 2$  bloc au niveau temporel, avec un chevauchement de 25% au niveau des blocs abstraits (i.e. pour  $h > 0$ ). Cette configuration fera en sorte que le dernier niveau de projection couvrira 64 canaux par 160 ms, donc de l'ordre de grandeur suprasegmental ou syllabique. Pour l'apprentissage du modèle acoustique par mixture de Bernoulli à 8 composantes, 50 itérations de l'algorithme EM ont été effectuées. Chaque HMM possède une topologie gauche-droite à 16 états. L'influence de chaque méta-paramètre sur les taux de reconnaissance a été validée pour obtenir cette configuration optimale.

## 5.2 Description du système de référence

L'étage d'extraction des caractéristiques le plus commun en reconnaissance de parole et basé sur les coefficients cepstraux sur l'échelle fréquentielle de Mel (MFCC) et de leurs dérivées temporelles. Les paramètres utilisés ont été tirés du livre de référence de la suite HTK [54]. Après un filtre de préaccentuation ( $\alpha = 0.97$ ) et un fenêtrage par Hamming du signal d'entrée, une transformée de Fourier à court terme est appliquée. Les amplitudes spectrales sont alors projetées sur l'échelle de Mel, consistant en des filtres triangulaires chevauchants. Les logarithmes des amplitudes sont alors pris à la sortie du banc de filtres, puis une transformée en cosinus discrète est appliquée pour obtenir les coefficients cepstraux. Seuls les premiers coefficients sont considérés, car ils portent l'information grossière du profil spectral. Il y a rehaussement standard et normalisation de la moyenne des coefficients cepstraux, puis calcul des dérivés premières (delta)

et secondes (delta-delta). Le vecteur de caractéristiques possède en tout 39 dimensions : 1 log-énergie, 12 coefficients cepstraux, 13 delta, 13 delta-delta. Intuitivement, les MFCCs modélisent donc l'enveloppe spectrale et son évolution locale à chaque fenêtre de temps, par un vecteur dense à faible-dimensionnalité. Pour l'apprentissage du modèle acoustique par mixture de gaussiennes à 4 composantes, 50 itérations de l'algorithme EM ont été effectuées. Chaque HMM possède une topologie gauche-droite à 16 états. L'influence de chaque méta-paramètre sur les taux de reconnaissance a été validée pour obtenir cette configuration optimale.

La base de données NOISEX-92 [55] a été utilisée pour simuler l'effet indésirable d'un bruit additif durant l'entraînement et l'évaluation des systèmes de reconnaissance. Les bruits sont réalistes et non-stationnaires, sauf dans le cas du bruit blanc. L'entraînement en condition propre est effectué avec les données originales non-bruitées. Pour l'entraînement multi-condition, chaque fichier de parole de l'ensemble d'entraînement original a été mélangé avec un des bruits utilisés pour l'ensemble de test. Le bruit est choisi aléatoirement pour chacun des fichiers, et le rapport signal-à-bruit (RSB) a été fixé à 20 dB. Il ne s'agit donc que d'un faible niveau de bruit. Toutefois, l'ensemble d'entraînement contient maintenant de l'information sur tous les bruits qui seront présentés avec l'ensemble de test, ce qui devrait réduire la disparité entre l'ensemble d'entraînement et de test. Dans ces conditions, de meilleures performances à bas rapports signal-à-bruit (RSB) sont attendues, car la variabilité induite par le bruit pourra être modélisée durant l'entraînement. L'entraînement d'un système de reconnaissance de parole avec un ensemble d'entraînement bruité permet de voir en quoi le système tire avantage de cette information additionnelle.

Il n'existe pas à notre connaissance de travaux ayant utilisé toutes les 46 classes de la base de données TI46 pour évaluer les taux de performance de reconnaissance. Il est très commun que les sous-ensembles restreints TI-20 (20 classes), TI-ALPHA (26 classes), ou même seulement les chiffres (10 ou 11 classes) soient utilisés. Un avantage est que l'utilisation des 46 classes rend le problème de classification plus difficile, car la confusion entre les classes augmente. Une comparaison directe des performances avec la littérature est donc impossible. Toutefois, les configurations optimales ont été dérivées indépendamment pour le système proposé et le système de référence, ce qui assure une comparaison juste et sans biais. L'objectif est avant tout de démontrer la polyvalence du système proposé, en terme de robustesse, sur différents types d'entraînement (i.e. avec ou sans bruit). Plusieurs techniques d'optimisation (e.g. modèle de silence en début et fin de mot) pourraient être ajoutés aux deux systèmes dans le but d'augmenter les taux de reconnaissance absolue.

### 5.3 Résultats

Le système proposé (SPARSE) et le système de référence (MFCC) ont été évalués sur l'ensemble de données de test bruité après un entraînement en condition propre (voir Tableau 1). On remarque une dégradation des performances de près de 12% (relatif) pour le système MFCC utilisé avec le bruit de conversation, lorsqu'on compare la condition de test propre avec les différents bruits à un rapport signal-à-bruit (RSB) de 40 dB. Le système SPARSE est beaucoup plus robuste dans cette condition, avec une dégradation maximale de 1.2% (relatif). Pour l'intervalle de RSBs entre 10 dB et 40 dB, les performances de ce dernier sont en majorité supérieures au système de référence, sauf dans le cas du bruit blanc gaussien et du bruit de salle de machine. Ces types de bruits sont propices à poser problème avec les mécanismes de seuillage utilisés pour produire une représentation binaire. Les caractéristiques décrivant plutôt les zones de faible énergie (e.g. silences ou pause courte), qui étaient ignorées durant l'entraînement en condition propre, peuvent maintenant affecter les autres caractéristiques fiables. Ceci engendre une grande disparité avec le modèle appris, qui ne peut alors plus généraliser correctement. Néanmoins, pour les autres bruits réalistes, l'usage de caractéristiques parcimonieuses et à grandes dimensions est une alternative efficace pour améliorer la robustesse comparativement au système de référence.

TABLE 1 – Taux de reconnaissance avec entraînement en conditions propres et test en conditions adverses, pour différents types de bruits (babble, destroyerengine, volvo, white) et RSBs (-5 dB à 40 dB). Les résultats montrent que le système proposé SPARSE permet une meilleure généralisation à haut RSBs, comparativement au système de référence MFCC. Le bruit blanc gaussien et le bruit de salle de machine semblent toutefois poser problème au système.

(a) Bruit de conversation

RSB	-5 dB	0 dB	10 dB	20 dB	40 dB	Propre
MFCC	<b>4.1</b>	<b>6.1</b>	16.8	36.9	70.1	78.4
SPARSE	2.8	5.0	<b>19.5</b>	<b>47.6</b>	<b>92.0</b>	<b>93.1</b>

(b) Bruit de salle de machine d'un contre-torpilleur (bateau)

RSB	-5 dB	0 dB	10 dB	20 dB	40 dB	Propre
MFCC	<b>2.5</b>	<b>4.2</b>	<b>14.9</b>	<b>36.9</b>	68.3	78.4
SPARSE	2.5	3.3	8.0	21.8	<b>77.6</b>	<b>93.1</b>

(c) Bruit intérieur d'une voiture

RSB	-5 dB	0 dB	10 dB	20 dB	40 dB	Propre
MFCC	<b>25.6</b>	<b>36.5</b>	54.2	67.7	78.1	78.4
SPARSE	18.2	32.6	<b>63.6</b>	<b>84.9</b>	<b>93.2</b>	<b>93.1</b>

(d) Bruit blanc gaussien

RSB	-5 dB	0 dB	10 dB	20 dB	40 dB	Propre
MFCC	2.2	2.3	<b>15.2</b>	<b>50.3</b>	77.9	78.4
SPARSE	<b>2.2</b>	<b>2.5</b>	7.2	19.5	<b>90.8</b>	<b>93.1</b>

Enfin, le système proposé (SPARSE) et le système de référence (MFCC) ont été évalués sur l'ensemble de données de test bruité après un entraînement multi-condition (voir Tableau 2). Rappelons que les bruits ont tous été présentés durant l'entraînement à un RSB fixe de 20 dB. On remarque une dégradation considérable des performances pour le système MFCC, lorsqu'on compare avec l'entraînement en conditions de test propres. Cette dégradation affecte moins le système SPARSE, qui semble pouvoir mieux modéliser les caractéristiques de la parole mélangées dans le bruit. Pour l'intervalle de RSBs entre 10 dB et 40 dB, les performances de ce dernier sont presque qu'exclusivement supérieures comparativement au système de référence. Pour un RSB de 10 dB, les performances du système SPARSE sont nettement supérieures comparativement à l'entraînement en condition propre, et ce, pour tous les types de bruits. Ceci démontre qu'un entraînement avec bruit est bénéfique à bas-RSB, même s'il affecte légèrement les performances de reconnaissance à haut-RSB. L'usage de caractéristiques parcimonieuses et à grandes dimensions permet de limiter ce problème et d'améliorer la capacité à généraliser. En effet, une différence majeure est que le système MFCC modélise le bruit à même les dimensions utiles de la parole, car la représentation est dense. Le système SPARSE utilise des caractéristiques parcimonieuses et hiérarchiques, et effectue une séparation telle que le bruit sera distribué sur des dimensions différentes de celles de la parole. Dans l'espace des paramètres, le bruit sera alors modélisé avec ses propres fonctions de densité de probabilité, donc dans un sous-espace disjoint où la variabilité sera concentrée hors des dimensions utiles de la parole.

Un profilage du temps de calcul des différents étages de traitement a été réalisé pour le système de référence et le système proposé, sur un processeur Intel Xeon cadencé à 2.4 GHz (1 coeur). Le tableau 3 montre que selon le facteur temps-réel (TR), l'approche proposée SPARSE est beaucoup plus lente que l'approche standard MFCC. Le facteur temps-réel  $\tau_{TR} = d_{son}/d_{trait}$  correspond au rapport de la durée  $d_{son}$  du son d'entrée sur la durée  $d_{trait}$  de traitement (e.g. extraction de caractéristiques, classification). Le cas  $\tau_{TR} \geq 1$  indique l'atteinte d'une performance en temps-réel, donc le système peut traiter le son d'entrée en continu (e.g. venant d'un microphone). On cherche à obtenir un facteur temps-réel (TR) le plus haut possible. L'implémentation actuelle de l'architecture proposée ne permet actuellement que le traitement hors-ligne de la parole. Plusieurs améliorations au niveau de l'étage d'extraction des caractéristiques permettraient toutefois de réduire le temps de calcul. Par exemple, lorsque la fréquence d'échantillonnage  $F_{sous}$  du vecteur de caractéristiques  $\mathbf{Y}_t$  au temps  $t$  est choisie telle qu'il y a alignement avec les blocs de projection de premier niveau (i.e.  $h = 0$ ), la majorité des projections effectuées au temps  $t$  peuvent servir au calcul des bases de niveau  $h = 0$  et  $h = 1$  au temps  $t + 1$ . Ceci permet de tirer profit du chevauchement temporel existant entre les vecteurs  $\mathbf{Y}_t$  et  $\mathbf{Y}_{t+1}$ . Enfin, l'évaluation de mélanges de Bernoulli possédant des milliers de paramètres est ce qui alourdit le plus l'étage de classification. L'entraînement

TABLE 2 – Taux de reconnaissance avec entraînement et test en conditions adverses, pour différents types de bruits (babble, destroyerengine, volvo, white) et RSBs (-5 dB à 40 dB). Les résultats montrent que le système proposé SPARSE permet une meilleure généralisation pour une large gamme de RSBs, comparativement au système de référence MFCC. Il y a aussi une dégradation de base moindre, comparativement à l'entraînement en condition propre. Le bruit durant l'entraînement affecte donc peu les performances absolues.

(a) Bruit de conversation

RSB	-5 dB	0 dB	10 dB	20 dB	40 dB	Propre
MFCC	<b>3.9</b>	<b>10.3</b>	34.6	55.0	59.9	41.0
SPARSE	3.1	9.6	<b>63.3</b>	<b>88.8</b>	<b>88.1</b>	<b>87.0</b>

(b) Bruit de salle de machine d'un contre-torpilleur (bateau)

RSB	-5 dB	0 dB	10 dB	20 dB	40 dB	Propre
MFCC	<b>4.7</b>	<b>8.2</b>	<b>43.5</b>	62.9	61.5	41.0
SPARSE	2.4	3.5	37.5	<b>83.1</b>	<b>79.9</b>	<b>87.0</b>

(c) Bruit intérieur d'une voiture

RSB	-5 dB	0 dB	10 dB	20 dB	40 dB	Propre
MFCC	38.1	46.6	56.5	59.8	43.1	41.0
SPARSE	<b>42.0</b>	<b>67.2</b>	<b>87.6</b>	<b>88.6</b>	<b>87.8</b>	<b>87.0</b>

(d) Bruit blanc gaussien

RSB	-5 dB	0 dB	10 dB	20 dB	40 dB	Propre
MFCC	<b>4.6</b>	<b>7.6</b>	38.8	58.4	45.1	41.0
SPARSE	2.2	3.4	<b>43.0</b>	<b>84.6</b>	<b>87.0</b>	<b>87.0</b>

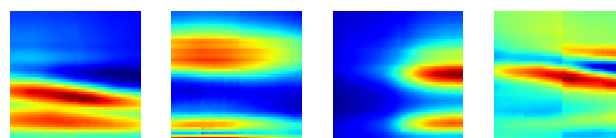
des systèmes SPARSE et MFCC requiert respectivement 190 minutes et 20 minutes de temps de calcul sur un processeur AMD Opteron cadencé à 2.2 GHz (16 coeurs). L'extraction de caractéristiques et l'évaluation de mélanges de Bernoulli à très grandes dimensions expliquent l'écart considérable entre les temps d'entraînement.

TABLE 3 – Facteur temps-réel (TR) pour l'exécution de l'étape d'extraction des caractéristiques et de classification, ainsi que de l'exécution globale. Les résultats montrent que la charge de calcul est significativement plus élevée pour l'approche SPARSE, comparativement au système de référence MFCC ayant l'avantage de pouvoir être utilisé en temps-réel.

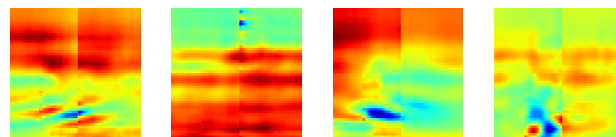
	Extraction (×TR)	Classification (×TR)	Global (×TR)
MFCC	<b>42.437</b>	<b>2.716</b>	<b>2.560</b>
SPARSE	0.792	0.518	0.314

## Effet du bruit sur l'apprentissage des bases

L'analyse en composantes indépendantes (ICA) est couramment utilisée pour la séparation aveugle de sources [e.g. 56], utile par exemple pour débruiter un signal de parole [e.g. 57, 58]. Il a déjà été remarqué que l'adaptation des bases en condition de bruit durant l'entraînement est bénéfique [59], car on tient alors compte de la variabilité introduite par le bruit. Avec des données d'entraînement bruitées reflétant mieux les conditions de tests, de meilleures performances de reconnaissance doivent être attendues, ce qui corrobore les résultats obtenus. La figure 7 montre qu'effectivement, lorsque les bases sont apprises en présence de bruit, il y a séparation des composantes du bruit de celles de la parole lors de la projection hiérarchique.



(a) Bases représentant la parole (64 x 160 ms)



(b) Bases représentant le bruit (64 x 160 ms)

FIGURE 7 – Exemple de bases apprises par analyse en composantes indépendantes (ICA), qui montre une séparation (a) des composantes de la parole et (b) des composantes du bruit. Le non-chevauchement du bruit et de la parole dans l'espace des paramètres du modèle acoustique statistique favorise une dégradation moindre des performances de reconnaissance en conditions adverses.

Les bases qui ont un contexte temporel plus grand sont plus aptes à extraire les composantes de bruit, car les régularités statistiques à long-terme peuvent être mieux capturées. Il s'agit d'une tâche plus difficile à petite échelle, ce qui explique la pertinence d'effectuer une projection hiérarchique pour l'obtention de bases de haut-niveau. Le choix de l'algorithme d'apprentissage des bases n'est qu'un des nombreux facteurs influençant les performances de reconnaissance, et pas nécessairement le plus important [60]. Dans le cas présent, l'analyse en composantes indépendantes (ICA) dans un contexte de projection hiérarchique s'est révélée l'algorithme d'apprentissage non-supervisé idéal. L'obtention du vecteur de caractéristiques parcimonieux à grandes dimensions est alors faite en s'assurant que les caractéristiques propres à la parole soient fidèlement extraites, même en présence de bruit dans les données d'entraînement. Ceci appuie bien l'aspect de puissance de discrimination des représentations parcimonieuses dans des espaces à grandes dimensions [61], avec un

compromis entre la séparabilité des classes et la qualité d'approximation. Dans le cas de la reconnaissance vocale, l'aspect de discrimination prime, car la reconstruction du signal est sans importance pour son identification.

## 6 DISCUSSION

L'un des objectifs était de comparer la robustesse en conditions difficiles d'un système de reconnaissance de mots isolés basé sur une représentation parcimonieuse à grande dimension. Plusieurs aspects bio-inspirés sont à la base des ces principes, et ont inspiré l'extraction de caractéristiques de modulation spectro-temporelle par décomposition linéaire et hiérarchique.

L'analyse en composantes indépendantes (ICA) semble être l'algorithme d'apprentissage non-supervisé idéal pour produire, avec l'intégration de ces aspects, des caractéristiques parcimonieuses à grandes dimensions où la parole et le bruit sont restreints dans des sous-espaces disjoints.

L'analyse du tableau 1 montre que le système SPARSE est supérieur lorsque l'apprentissage et la reconnaissance se font en conditions propres. Ceci valide l'intérêt de l'étude, puisque les résultats de départ sont au moins supérieurs à ceux d'un système conventionnel. Par contre, avec un apprentissage toujours en conditions propres, mais une reconnaissance en conditions difficiles (ou avec bruits), le système SPARSE proposé n'augmente que peu les taux de reconnaissance. Il est meilleur pour le bruit de voiture et de conversation, mais moins bon pour le bruit blanc et l'environnement sonore du bateau. Dans la situation de l'apprentissage avec des données propres, les deux systèmes possèdent des capacités similaires à généraliser en conditions de test difficiles – pour la majorité des bruits réalistes testés. Le système SPARSE est toutefois plus robuste à la présence de bruit dans les données d'entraînement (voir Tableau 2). On observe une dégradation moindre des taux de reconnaissance à des rapports signal-à-bruit entre 10 dB et 40 dB. En entraînement multi-condition, le système SPARSE permet donc de mieux généraliser, car les bases apprises lors de la projection hiérarchique créent une séparation des composantes de parole de celles du bruit. Cette séparation est conservée au niveau du modèle acoustique, soit dans les paramètres des mixtures de Bernoulli à grandes dimensions.

Les représentations parcimonieuses et par objets de la parole promettent donc d'améliorer les taux de reconnaissance des systèmes utilisés dans des environnements non-contrôlés et riches en bruits additifs, principalement là où les données d'entraînement sont de moindre qualité (i.e. aussi légèrement bruitées). Les travaux futurs porteront à valider l'approche proposée en reconnaissance de parole continue sur une tâche à large vocabulaire, ainsi qu'à augmenter la rapidité d'exécution.

## 7 CONCLUSION

La définition d'un objet sonore est différente selon qu'on se place du point de vue de la production ("génération") ou de la perception de l'objet. D'un point de vue de la production du son, on se plaît à concevoir qu'un objet sonore est composé de caractéristiques spécifiques qui se succèderaient dans le temps (e.g. ONSET ou transitoire suivi d'un cours silence, puis d'une voyelle). La synthèse de parole repose souvent sur cette conception [62]. En conséquence, la génération de certains objets sonores peut être réalisée par l'établissement d'une succession d'évènements acoustiques qui sont considérés comme étant distincts dans le temps (la qualité perceptive n'est d'ailleurs pas toujours excellente). En revanche la perception d'un objet sonore ne peut s'établir de la même façon. Un objet sonore simple n'est pas perçu comme étant la succession d'évènements discrets élémentaires temporels, mais plutôt comme étant un tout qui est le résultat de l'intégration spatio-temporelle des caractéristiques élémentaires par le système auditif. Par exemple le son /b/ est perçu comme étant un tout et non pas comme étant composé d'une barre de voisement en basse fréquence (activation des cordes vocales) suivie d'un transitoire puis de la voyelle. En reconnaissance automatique des sons, il semble donc plus logique de vouloir utiliser le point de vue perceptif d'un objet sonore que le point de vue de la génération de l'objet sonore pour pouvoir atteindre des performances qui soient le plus proche possible de celles de l'humain.

L'approche utilisée dans le présent travail considère un objet sonore comme étant la combinaison spatio-temporelle et hiérarchique d'unités élémentaires. Ces unités élémentaires ont été trouvées par apprentissage non-supervisé suivant l'analyse en composante indépendante (ICA) et une organisation hiérarchique. Le critère de recherche de ces unités élémentaires qui a été choisi repose essentiellement sur des données physiologiques qui montrent l'existence d'indépendance statistique entre certains neurones corticaux, permettant ainsi de maximiser la quantité d'information (entropie) par neurones ou groupes de neurones. Nous observons qu'avec le critère choisi, les parties d'objets sonores (ou bases) sont des combinaisons de patrons de transitoires, de modulations en amplitude (AM) et en fréquence (FM) des signaux d'enveloppes extraits à partir d'un banc de filtres cochléaires. Par ailleurs, la recherche automatique et non-supervisée de ces unités élémentaires permet une adaptation de l'approche à différents contextes acoustiques.

Toutefois, ce travail ne répond que très partiellement aux questions posées à l'introduction de cet article. En effet, la façon de combiner les caractéristiques entre elles est arbitraire (utilisation d'un seul vecteur) et seule l'information des enveloppes à la sortie d'un banc de filtres cochléaire a été exploitée. Il est reconnu que le système auditif est aussi en mesure de résoudre les harmoniques en basse fréquence, car un nombre important de fibres basses fréquences du nerf auditif (et donc de cellules ciliées de la cochlée) déchargent selon la

période de la fondamentale du signal et non pas selon l'enveloppe des sorties du banc de filtres [63]. Cet aspect n'a pas été pris en compte dans le présent travail, car l'objectif n'était pas d'intégrer les caractéristiques de suivi de fréquence fondamentale ou de prosodie dans la représentation des objets sonores utilisés en reconnaissance.

Il a été illustré comment l'intégration de principes inspirés des neurosciences permet de proposer une représentation objet de signaux sonores. Il aurait été possible de combiner différemment les composantes des objets sonores. En effet, la façon de combiner affecte la représentation des objets sonores et donc les résultats. Ici, il a été choisi d'utiliser une combinaison très simple : chaque dimension d'un vecteur est une caractéristique d'objet (ou de partie d'objet), peu importe sa position ou son rôle dans la structure de l'objet sonore en question. Malgré cette combinaison non ordonnée et non hiérarchisée des composantes des objets, les résultats sont très prometteurs.

Ce travail montre par ailleurs qu'il est possible de tirer profit des représentations à très grandes dimensions. Ces représentations sont rarement considérées comme intéressantes en modélisation statistique des signaux, notamment en raison de la complexité anticipée. Cependant, la complexité de calcul peut être grandement réduite, car l'extraction des caractéristiques reste simple et se prête bien au calcul distribué. De plus, les vecteurs des composantes objets sont à coordonnées binaires. Ceci permet aussi une grande simplicité de mise en oeuvre.

## REMERCIEMENTS

Calcul Canada, pour les ressources de calcul de haute-performance mises à disposition. Le conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), le fonds de recherche Québec en Nature et Technologies (FRQ-NT).

## RÉFÉRENCES

- [1] G. Chechik, M. J. Anderson, O. Bar-Yosef, E. D. Young, N. Tishby, I. Nelken. "Reduction of Information Redundancy in the Ascending Auditory Pathway." *Neuron*, **51**, 359–368 (2006).
- [2] S. Molotchnikoff, J. Rouat. "Brain at work : Time, Sparseness and Superposition Principles." *Frontiers in Bioscience (Landmark Edition)*, **17**, 583–606 (2012).
- [3] D. L. Alkon, K. T. Blackwell. "Pattern Recognition by an Artificial Network Derived from Biological Neuronal Systems." *Biological Cybernetics*, **62**, 363–376 (1990).
- [4] Z. Chen, S. Haykin, J. J. Eggermont, S. Becker. *Correlative Learning A Basis for Brain and Adaptive Systems* (Wiley-interscience, 2007).
- [5] D. Depireux, J. Simon, D. Klein, S. Shamma. "Spectrotemporal response field characterization with dynamic ripples in ferret primary auditory cortex." *Journal of Neurophysiology*, **85**, 1220–34 (2001).
- [6] D. N. Zotkin, T. Chi, S. Shamma, R. Duraiswami. "Neuromimetic Sound Representation for Percept Detection and Manipulation." *EURASIP Journal on Applied Signal Processing*, **9**, 1350–64 (2005).
- [7] M. S. Lewicki. "Efficient coding of natural sounds." *Nature neuroscience*, **5**, 356–363 (2002).
- [8] E. C. Smith, M. S. Lewicki. "Efficient auditory coding." *Nature*, **439**, 978–982 (2006).
- [9] T. Irino, R. D. Patterson. "A time-domain, level-dependent auditory filter : The gammachirp." *Journal of the Acoustical Society of America*, **101**, 412–419 (1997).
- [10] R. D. Patterson. "Auditory filter shapes derived with noise stimuli." *Journal of the Acoustical Society of America*, **59**, 640–654 (1976).
- [11] R. D. Frisina, R. L. Smith, S. C. Chamberlain. "Differential Encoding of Rapid Changes in Sound Amplitude by Second-Order Auditory Neurons." *Experimental Brain Research*, **60**, 417–422 (1985).
- [12] C. E. Schreiner, G. Langner. "Periodicity coding in the inferior colliculus of the cat. II. Topographical organization." *Journal of neurophysiology*, **60**, 1823–40 (1988).
- [13] J. Winer, C. Schreiner. *The inferior colliculus* (Springer Verlag, 2005).
- [14] C. E. Schreiner, J. V. Urbas. "Representation of amplitude modulation in the auditory cortex of the cat. I. The anterior auditory field (AAF)." *Hearing research*, **21**, 227–241 (1986).
- [15] D. J. Klein, P. König, K. P. Körding. "Sparse Spectrotemporal Coding of Sounds." *EURASIP Journal on Advances in Signal Processing*, **2003**, 659–667 (2003).
- [16] M. S. Lewicki, T. J. Sejnowski. "Learning Overcomplete Representations." *Neural Computation*, **12**, 337–365 (2000).
- [17] A. Hyvärinen, J. Karhunen, E. Oja. "ICA with Overcomplete Bases." In "Independent Component Analysis," volume 7, chapter 16, pp. 305–313 (John Wiley & Sons, Inc, 2001).
- [18] D. D. Lee, H. S. Seung. "Learning the parts of objects by non-negative matrix factorization." *Nature*, **401**, 788–791 (1999).
- [19] D. L. Donoho, V. C. Stodden. "When Does Non-Negative Matrix Factorization Give a Correct Decomposition into Parts ?" In "Proceedings of the 18th Annual Conference on Neural Information Processing System (NIPS)," (Vancouver, BC, Canada, 2004).



- [20] G. Hickok, D. Poeppel. "The cortical organization of speech processing." *Nature Reviews Neuroscience*, **8**, 393–402 (2007).
- [21] M. S. Lewicki, B. J. Arthur. "Hierarchical Organization of Auditory Temporal Context Sensitivity." *The Journal of Neuroscience*, **16**, 6987–98 (1996).
- [22] S. Molotchnikoff, J. Rouat. *Visual Cortex - Current Status and Perspectives*, (Intech, 2012).
- [23] R. Guyonneau, R. VanRullen, S. J. Thorpe. "Neurons tune to the earliest spikes through STDP." *Neural Computation*, **17**, 859–879 (2005).
- [24] R. Guyonneau, R. VanRullen, S. J. Thorpe. "Temporal codes and sparse representations : a key to understanding rapid processing in the visual system." *Journal of physiology, Paris*, **98**, 487–497 (2004).
- [25] S. B. Davis, P. Mermelstein. "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences." *IEEE Transactions on Acoustics, Speech and Signal Processing*, **28**, 357–366 (1980).
- [26] L. Rabiner. "A tutorial on hidden Markov models and selected applications in speech recognition." *Proceedings of the IEEE*, **77**, 257–286 (1989).
- [27] F. Valente, M. Magimai-Doss, C. Plahl, S. Ravuri. "Hierarchical Processing of the Modulation Spectrum for GALE Mandarin LVCSR system." In "Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech)," (Brighton, UK, 2009).
- [28] A. Mohamed, G. Dahl, G. Hinton. "Deep Belief Networks for phone recognition." In "Proceedings of the 24th Annual Conference on Neural Information Processing System (NIPS), Workshop on Deep Learning for Speech Recognition and Related Applications," (Whistler, BC, Canada, 2009).
- [29] A. Mohamed, D. Yu, L. Deng. "Investigation of full-sequence training of deep belief networks for speech recognition." In "Proceedings of the 11th Annual Conference of the International Speech Communication Association (Interspeech)," pp. 2846–49 (Makuhari, Chiba, Japan, 2010).
- [30] F. Seide, G. Li, D. Yu. "Conversational Speech Transcription Using Context-Dependent Deep Neural Networks." In "Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech)," August, pp. 437–440 (Interspeech, Florence, Italy, 2011).
- [31] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, B. Kingsbury. "Deep Neural Networks for Acoustic Modeling in Speech Recognition : The Shared Views of Four Research Groups." *IEEE Signal Processing Magazine*, **29**, 82–97 (2012).
- [32] C. Radhakrishna Rao, S. K. Mitra. "Generalized inverse of a matrix and its applications." In "Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability," pp. 601–620 (Berkeley, USA, 1972).
- [33] A. Hyvärinen, E. Oja. "Independent component analysis : algorithms and applications." *Neural networks*, **13**, 411–430 (2000).
- [34] A. Hyvärinen. "Fast and robust fixed-point algorithms for independent component analysis." *IEEE transactions on neural networks*, **10**, 626–634 (1999).
- [35] J. Lee, T. Lee, H. Jung, S. Lee. "On the efficient speech feature extraction based on independent component analysis." *Neural Processing Letters*, **15**, 235–245 (2002).
- [36] D. W. Wang, X. Ma, Y. Su. "Undercomplete Dictionary-Based Feature Extraction for Radar Target Identification." *Progress In Electromagnetics Research M*, **1**, 1–19 (2008).
- [37] G. R. Naik, D. K. Kumar. "An Overview of Independent Component Analysis and Its Applications." *Informatica*, **35**, 63–81 (2011).
- [38] J. Porrill, J. V. Stone. "Undercomplete Independent Component Analysis for Signal Separation and Dimension Reduction." Technical report, University of Sheffield, Department of Psychology (1998).
- [39] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, N. L. Dahlgren, V. Zue. "TIMIT Acoustic-Phonetic Continuous Speech Corpus." (1993).
- [40] I. Perlman. "Vivaldi - The Four Seasons and Violin Concertos [Audio CD]." (2002).
- [41] Nature Sound Series. "Tropical Ocean Surf (Nature Sounds Only version) [Audio CD]." (2007).
- [42] Nature Sound Series. "Mountain Stream (Nature sounds only version) [Audio CD]." (2006).
- [43] A. Huber, T. Linder, M. Ferrazzini, S. Schmid, N. Dillier, S. Stoeckli, U. Fisch. "Intraoperative assessment of stapes movement." *Annals of Otology, Rhinology & Laryngology*, **110**, 31–35 (2001).
- [44] V. Hohmann. "Frequency analysis and synthesis using a Gammatone filterbank." *Acta Acustica united with Acustica*, **88**, 433–442 (2002).
- [45] S. Stevens, J. Volkman, E. Newman. "A scale for the measurement of the psychological magnitude pitch." *Journal of the Acoustical Society of America*, **8**, 185–190 (1937).

- [46] C. Avendaño, L. Deng, H. Hermansky, B. Gold. "The analysis and representation of speech." *Speech processing in the auditory system*, **18**, 63–100 (2004).
- [47] M. Liberman et al. "TI 46-Word Linguistic Data Consortium, Philadelphia." (1993).
- [48] A. Juan, E. Vidal. "Bernoulli mixture models for binary images." In "Proceedings of the 17th International Conference on Pattern Recognition (ICPR)," volume 3, pp. 367–370 (IEEE, Cambridge, UK, 2004).
- [49] V. Romero, A. Giménez, A. Juan. "Explicit Modeling of Invariances in Bernoulli Mixtures for Binary Images." *Pattern Recognition and Image Analysis*, **4477**, 539–546 (2007).
- [50] A. Giménez, A. Juan. "Bernoulli HMMs at Subword Level for Handwritten Word Recognition." *Pattern Recognition and Image Analysis*, **5524**, 497–504 (2009).
- [51] A. Juan, J. García-Hernández, E. Vidal. "EM Initialisation for Bernoulli Mixture Learning." In "International Workshop on Structural and Syntactic Pattern Recognition (SSPR)," pp. 635–643 (Lisbon, Portugal, 2004).
- [52] A. Giménez, A. Juan. "Bernoulli HMMs at Subword Level for Handwritten Word Recognition." In "Proceedings of the 10th International Conference on Document Analysis and Recognition," volume 5524, pp. 497–504 (IEEE, Barcelona, Spain, 2009).
- [53] M. Henniges, G. Puertas, J. Eggert, J. Lücke. "Binary Sparse Coding." *Latent Variable Analysis and Signal Separation*, **6365**, 450–457 (2010).
- [54] S. Young, G. Evermann, D. Kershaw, G. Moore, J. Odell, D. Ollason, V. Valtchev, P. Woodland. *The HTK Book (for HTK Version 3.4)* (Cambridge University Engineering Department, 2009).
- [55] A. Varga, H. Steeneken. "Assessment for automatic speech recognition : II. NOISEX-92 : A database and an experiment to study the effect of additive noise on speech recognition systems." *Speech Communication*, **12**, 247–251 (1993).
- [56] D. Obradovic, G. Deco. "Blind signal separation revisited." In "Proceedings of the 36th IEEE Conference on Decision and Control," volume 2, pp. 1591–96 (IEEE, San Diego, USA, 1997).
- [57] J. Lee, H. Jung, T. Lee, S. Lee. "Speech enhancement with MAP estimation and ICA-based speech features." *Electronics Letters*, **36**, 1506–7 (2000).
- [58] L. Hongyan, R. Guanglong. "Blind separation of noisy mixed speech signals based Independent Component Analysis." In "Proceedings of the First International Conference on Pervasive Computing, Signal Processing and Applications," pp. 586–589 (IEEE, Harbin, China, 2010).
- [59] M. Heckmann, X. Domont, F. Joublin, C. Goerick. "A Hierarchical Framework for Spectro-Temporal Feature Extraction." *Speech Communication*, **53**, 736–752 (2011).
- [60] F. O'Donnell, F. Triefenbach, J.-P. Martens, B. Schrauwen. "Effects of Architecture Choices on Sparse Coding in Speech Recognition." In "Proceedings of the International Conference on Artificial Neural Networks (ICANN)," volume 7552 of *Lecture Notes in Computer Science*, pp. 629–636 (Springer, Lausanne, Switzerland, 2012).
- [61] I. Tasic, P. Frossard. "Dictionary Learning : What is the right representation for my signal?" *IEEE Signal Processing Magazine*, **28**, 27–38 (2011).
- [62] É. Keller. *Fundamentals of speech synthesis and speech recognition : basic concepts, state of the art, and future challenges* (Wiley, 1994).
- [63] M. I. Miller, M. B. Sachs. "Representation of voice pitch in discharge patterns of auditory-nerve fibers." *Hearing research*, **14**, 257–279 (1984).



## Lightweight, Attractive Enclosures

AIL Sound Walls are engineered to be the choice for today's rooftop and on-ground enclosures. Made from maintenance-free PVC, our Silent Protector® (Absorptive) and Tuf-Barrier® (Reflective) enclosures will save you time and money.

- ▶ Easy-to-install
- ▶ Durable and maintenance-free
- ▶ Lightweight, reduced loads
- ▶ Supports LEED points for green building

Professional  
education credit  
courses offered



1-866-231-7867

ailsoundwalls.com

# EFFET DU TYPE DE BRUIT ET DE LA SÉPARATION SPATIALE SUR LA RECONNAISSANCE DE LA PAROLE MESURÉE À L'AIDE DU HINT

**Chantal Laroche, Véronique Vaillancourt, Victoria Milloy, Christian Giguère, Jean-Grégoire Roveda**

Programme d'audiologie et d'orthophonie, École des sciences de la réadaptation,

Université d'Ottawa, 451, chemin Smyth, Ottawa, Ontario K1H 8M5

E-mail: [claroche@uottawa.ca](mailto:claroche@uottawa.ca)

## RÉSUMÉ

La perception de la parole dans le bruit est l'une des tâches auditives les plus importantes pour assurer une communication efficace dans la vie de tous les jours. Plusieurs tests cliniques ont été développés pour quantifier les difficultés de perception de la parole dans un bruit continu. Il est toutefois bien reconnu que les bruits rencontrés dans la vie de tous les jours sont davantage fluctuants et que les périodes moins intenses de bruit permettent à l'auditeur de mieux capter certaines parties du message verbal. Le présent projet a comme objectif d'évaluer la performance de jeunes adultes avec audition normale au HINT, sous écouteurs et en champ libre, en utilisant une version modifiée du bruit standard continu en y insérant un taux d'intermittence de 16 Hz. Des données normatives ont été établies avec ce bruit modifié pour les trois conditions de bruit du HINT, soit Bruit Avant, Bruit Droit et Bruit Gauche. Lorsque comparés aux normes avec le bruit continu, les seuils de réception de la parole s'améliorent de 13 dB et 10 dB sous écouteurs et de 9 dB et 6 dB en champ libre lorsque le bruit est positionné en frontal et sur le côté, respectivement. L'avantage binaural (différence de seuils entre bruit frontal et bruit côté) sous écouteurs est d'environ 8-9 dB et 5 dB pour le bruit continu et intermittent, respectivement, et de 6-7 dB et 3-4 dB en champ libre pour le bruit continu et intermittent, respectivement. Les données normatives maintenant disponibles pour un bruit intermittent de 16 Hz permettront d'ajouter au HINT clinique au moins une condition de bruit non-continu afin de mieux documenter la capacité des individus, particulièrement ceux avec perte auditive, à prendre avantage des périodes de « silence » pour reconnaître la parole.

## ABSTRACT

Speech perception in noise is one of the most important tasks for proper communication in everyday listening situations. Multiple clinical tests have been developed to quantify difficulties understanding speech in continuous noise. However, it is now well recognized that everyday noises are fluctuating in nature, and that segments with lower noise levels allow listeners to more easily capture speech elements. The objective of the current study is to evaluate the performance of young adults with normal hearing on the HINT, under headphones and in the sound field, using a modified 16-Hz interrupted version of the standard continuous noise. Normative data for the new intermittent noise was established in each of the three HINT noise conditions: Noise Front, Noise Right and Noise Left. When compared to performance with continuous noise, threshold improvements of 13 dB and 10 dB under headphones and of 9 dB and 6 dB in the sound field are noted in the Noise Front and Noise Side conditions, respectively. The binaural advantage (the threshold difference between the Noise Front and the Noise Side conditions) under headphones was approximately 8-9 dB and 5 dB for the continuous and intermittent noises, respectively, and 6-7 dB and 3-4 dB in the sound field for continuous and intermittent noises, respectively. Now that normative data is available for the modified 16-Hz intermittent noise, it could be useful to add at least one testing condition of intermittent noise to the clinical HINT. Such further testing would allow documentation of one's ability to take advantage of "dips" in noise to increase speech recognition, particularly in individuals with hearing loss.

## 1. INTRODUCTION

La perception de la parole dans le bruit est l'une des tâches auditives les plus importantes pour assurer une communication efficace dans la vie de tous les jours. Or, la plainte la plus fréquente des personnes avec perte auditive consiste en des difficultés de compréhension la parole dans des milieux bruyants (Wilson et coll., 2010; Moore et Gordon-Hickey, 2011). Malgré la nature de ces plaintes, l'audiométrie tonale et le seuil de réception de la parole dans le silence demeurent les outils les plus fréquemment utilisés par les cliniciens en audiologie pour quantifier les difficultés auditives. C'est d'ailleurs sur de tels résultats cliniques que les décisions en matière d'amplification sont souvent prises. Une étude menée par Mueller (2001) a démontré que 92% des cliniciens utilisaient le seuil de réception de la parole dans le silence en plus de l'audiométrie tonale, alors que 35% des cliniciens utilisaient des monosyllabiques dans le bruit et que seulement 6% des cliniciens évaluaient la perception de la parole dans le bruit à l'aide de phrases. Pourtant, il est de plus en plus reconnu que les deux premières mesures ne sont pas de bons prédicteurs des problèmes de communication dans le bruit (Killion et Niquette, 2000; Laroche et coll., 2005).

Plusieurs tests ont été développés pour quantifier les difficultés de perception de la parole dans le bruit, incluant le Connected Speech test (Cox et coll., 1987), le Speech Perception In Noise Test (Kalikow et coll., 1977; Bilger et coll., 1984), le Quick Speech-in-Noise Test (Killion et coll., 2004), le Test de phrases dans le bruit (Lagacé et coll., 2010) ainsi que le Hearing in Noise Test (HINT) (Soli et Wong, 2008). Les quatre premiers tests rapportent les résultats en pourcentage d'intelligibilité de la parole et utilisent des niveaux fixes de parole ou de bruit. Ils peuvent être ainsi limités par les effets de plafond et de plancher. Le HINT utilise une approche différente en déterminant, à l'aide d'une méthode adaptative, un seuil de réception de la parole (phrases) (SRP) dans le silence et dans différentes conditions de bruit. Les résultats sont donc exprimés en dBA (silence) ou en rapport signal-bruit (S/B). Un avantage du SRP comparativement au pourcentage d'identification de la parole est qu'une méthode identique est utilisée indépendamment du statut auditif (nul besoin d'ajuster le niveau de présentation des stimuli en fonction du degré de la perte auditive). L'approche est aussi facile à utiliser et élimine à la fois les effets de plafond/plancher et le besoin de calculer des pourcentages de mots/phrases correctement identifiés.

En utilisant le protocole standard du HINT, une liste de 20 phrases est utilisée pour mesurer un SRP dans chacune des 4 conditions du test, soit: 1) Silence (phrases présentées de l'avant), 2) Bruit Avant (phrases et bruit de l'avant), 3) Bruit Gauche (phrases de l'avant et bruit du côté gauche), et 4) Bruit Droit (phrases de l'avant et bruit de la droite). Le score composite  $[(2*BA) + BG + BD]/4$  permet de quantifier de façon globale la performance

d'écoute dans le bruit chez un individu. Le HINT permet également d'étudier l'effet de la séparation spatiale des sources de parole et de bruit, et donc de documenter la contribution du démasquage binaural. Il s'agit de la différence entre le SRP mesuré dans la condition Bruit Avant et celui mesuré dans les conditions de bruit provenant des côtés.

Des données normatives ont été établies pour le HINT clinique dans plusieurs langues (Soli et Wong, 2008), incluant le HINT en français canadien. L'écart par rapport aux données normatives permet de quantifier les difficultés de perception de la parole dans le bruit chez un individu en particulier. Cet écart est souvent nommé « perte en rapport S/B » ou « SNR loss » en anglais (Killion et Niquette, 2000). Si cette perte s'élève à 4 dB, par exemple, cela indique que cet individu nécessite 4 dB de plus en rapport S/B pour atteindre le même niveau de performance que la moyenne des individus avec audition normale. Les résultats au HINT peuvent d'ailleurs guider les cliniciens dans l'établissement d'un plan de réadaptation qui adresse plus spécifiquement les besoins et les difficultés de la personne. La perte en rapport S/B est par ailleurs un paramètre très utile pour établir des attentes réalistes lors du counseling.

Tous les tests cités précédemment sont réalisés avec des bruits continus à spectre de la parole ou du bruit de verbiage (« babble noise »). Les caractéristiques du bruit utilisé peuvent toutefois fortement influencer les SRPs. Il est par ailleurs bien reconnu que les bruits rencontrés dans la vie de tous les jours sont davantage fluctuants et que les périodes moins intenses dans le bruit permettent à l'auditeur de mieux capter certaines parties du message verbal, parfois suffisamment pour décoder le message dans son entier. Chez des auditeurs normaux, la compréhension de la parole est supérieure dans des bruits avec fluctuations temporelles que dans des bruits continus (Miller, 1947; Miller et Licklider, 1950; Wilson et Carhart, 1969; Duquesnoy, 1983; Festen et Plomp, 1990; Takahashi et Bacon, 1992; Howard-Jones et Rosen, 1993; Baer et Moore, 1994; Gustafsson et Arlinger, 1994; Peter et coll., 1998; Moore et coll., 1999; Cooke, 2006; Füllgrabe et coll., 2006; Rhebergen et coll., 2006).

Cette capacité à écouter dans les « trous » ou à capter une certaine partie de l'information dans les segments plus faibles du bruit ambiant, se traduisant par une amélioration de la perception de la parole dans le bruit, se nomme démasquage. Comparativement au bruit continu, on note une amélioration du SRP d'environ 6-10 dB chez des auditeurs normaux dans les bruits avec fluctuations temporelles, une valeur qui dépend fortement des conditions expérimentales et qui peut excéder 20 dB dans certains cas (Wilson et Carhart, 1969; Festen et Plomp, 1990; George et coll., 2006; Rhebergen et coll., 2006).

Quoiqu'il soit reconnu qu'une perte auditive neurosensorielle puisse nuire à la perception de la parole dans le bruit, la différence de performance entre les auditeurs normaux et ceux avec perte auditive semble

d'autant plus accrue pour les bruits intermittents comparativement aux bruits continus. En effet, les gens avec perte neurosensorielle semblent moins aptes à prendre avantage des moments de silence dans le bruit et démontrent ainsi un démasquage réduit comparativement à leurs pairs avec audition normale (Wilson et Carhart, 1969; Festen et Plomp, 1990; Takahashi et Bacon, 1992; Gustafsson et Arlinger, 1994; Eisenberg et coll., 1995; Moore et coll., 1995; Bacon et coll., 1998; Peters et coll., 1998; Dubno et coll., 2003; George et coll., 2006; Jin et Nelson, 2006; Lorenzi et coll., 2006b; Bernstein et Grant, 2009; Strelcyk et Dau, 2009), même lorsque la perte est légère ou limitée aux hautes fréquences (Middelweerd et coll., 1990; Takahashi et Bacon, 1992). Un démasquage réduit est également noté, du moins en partie, chez des individus avec perte auditive lorsque les sons présents dans les segments moins intenses de bruit sont amplifiés à des niveaux supra-seuils (Peters et coll., 1998; Moore et coll., 1999) ou quand l'audibilité du signal est assurée (Takahashi et Bacon, 1992; Eisenberg et coll., 1995; Bacon et coll., 1998).

Comparativement à la performance dans un bruit continu, Wilson et Carhart (1969) ont démontré une amélioration du SRP pour des spondées présentés dans un bruit intermittent (un bruit continu multiplié par une onde carrée) pouvant atteindre 28 dB chez des auditeurs normaux et 11 dB chez des individus avec perte auditive neurosensorielle (otosclérose cochléaire; SRP-silence < 60 dB SPL). Des études récentes menées par le même groupe de recherche (Wilson et coll., 2010) avec des taux d'interruption de 5, 10 et 20 Hz ont démontré un démasquage pouvant atteindre 34 dB chez des individus avec audition normale comparativement à seulement 2.1-2.4 dB pour des auditeurs avec une perte auditive.

Des facteurs autres que la perte auditive peuvent également influencer l'ampleur du démasquage, soit le matériel utilisé (spondées, phrases, chiffres), l'âge (Dubno et coll., 2003; George et coll., 2006), le rapport signal-sur-bruit (RSB), le seuil de réception de la parole dans un bruit continu, et les caractéristiques du bruit masquant telles que son enveloppe temporelle, le taux de fluctuation et la profondeur des fluctuations. La quantité de démasquage est directement proportionnelle à la profondeur des fluctuations (Miller et Licklider, 1950; Wilson et Carhart, 1969; Bronkhorst et Plomp, 1992; Howard-Jones et Rosen, 1993; Gustafsson et Arlinger, 1994; Cooke, 2003, 2006; Simpson et Cooke, 2005; George et coll., 2006; Cooke, 2006; Gnansia et coll., 2008) et est supérieure pour les bruits masquant dont l'enveloppe temporelle contient des chutes importantes dans le niveau sonore, tels que pour les ondes carrées (Bacon et coll., 1998; Rhebergen et coll., 2006). De plus, les taux de modulation faibles génèrent un démasquage plus important (Miller et Licklider, 1950; Bacon et coll., 1998; Howard-Jones et Rosen, 1993; Gustafsson et Arlinger, 1994; Kwon et Turner, 2001; Buss et coll., 2003; Nelson et coll., 2003; Füllgrabe et coll., 2006; Lorenzi et coll., 2006a; Rhebergen et coll., 2006), la

grandeur de l'effet étant sous l'influence du matériel vocal et des conditions expérimentales. Finalement, l'ampleur du démasquage varie avec le RSB et semble dépendre, du moins en partie, du SRP dans un bruit continu, avec un démasquage nul ou faible observé chez les individus qui ont un SRP élevé (Bernstein et Grant, 2009; Oxenham et Simonson, 2009; Bernstein et Brungart, 2011; Freyman et coll., 2012; Christiansen et Dau, 2012).

Miller et Licklider (1950) ont mesuré le taux d'identification de monosyllabes dans divers bruits intermittents auprès d'une population avec audition normale. Ils ont noté une amélioration de la perception de la parole dans des bruits dont le taux de modulation variait de 0.1 à 10 kHz par rapport à un bruit continu. Les performances s'amélioraient avec une diminution du taux de modulation. Pour des taux supérieurs à 200 Hz, aucune amélioration significative attribuable aux fluctuations n'a été notée. Smits et Houtgast (2007) ont mesuré le SRP pour une série de chiffres présentés dans un bruit continu et deux bruits intermittents (16 Hz et 32 Hz), et ont démontré que le bruit de 16 Hz était plus discriminatoire pour distinguer les auditeurs avec perte auditive de ceux avec une audition normale. Il semble donc qu'un bruit intermittent de 16 Hz soit un choix de stimulus adéquat pour séparer, dans un contexte clinique, les deux populations sur la base d'une mesure du SRP. Par ailleurs, Smits et coll. (2004) ont trouvé une corrélation de 0.87 entre la performance avec le matériel utilisé (séquence de 3 chiffres) et celle obtenue avec un test de perception de phrases. Il n'en demeure pas moins que les phrases sont plus représentatives des situations de communication rencontrées au quotidien que les mots isolés, les spondées ou les chiffres. En 2006, Rhebergen et coll. ont obtenu un démasquage de l'ordre de 10 dB chez des individus avec audition normale lors de la présentation de phrases dans un bruit intermittent de 16 Hz. Ils n'ont toutefois pas effectué de mesures chez des auditeurs avec une perte auditive.

Les résultats de quelques études additionnelles ayant utilisé des phrases et une méthode adaptative pour mesurer le SRP sont présentées sommairement au tableau 1. Les trois premières études ont utilisé le matériel vocal du HINT en anglais américain (Nilsson et al. 1994). Dans tous les cas répertoriés au tableau 1, le SRP mesuré dans un bruit continu de spectre de parole (BSP) est comparé à celui obtenu dans un bruit avec fluctuations temporelles pour obtenir la grandeur du démasquage. L'ampleur du démasquage varie de 4 à 15 dB chez les individus avec audition normale comparativement à seulement 1 à 5 dB chez des auditeurs avec une perte auditive, et dépend des conditions expérimentales à l'étude. Tel que noté dans la dernière colonne du tableau, le bruit fluctuant permet une différenciation plus importante des deux populations à l'étude (audition normale et perte auditive) que le bruit continu. Il semble donc que la perception de la parole dans un bruit fluctuant soit une mesure plus sensible pour distinguer les individus avec perte auditive de ceux dont l'audition est normale.

Quoiqu'un démasquage réduit chez les individus avec perte auditive comparativement à ceux avec audition normale soit clairement documenté dans la littérature, son explication demeure incertaine (Bernstein et Grant, 2009; Hopkins et Moore, 2009; Freyman et coll., 2012; Kwon et coll., 2012) et plusieurs mécanismes ont été proposés, incluant: 1) une audibilité réduite de la parole dans les segments plus faibles du bruit, 2) une résolution spectrale réduite, 3) une résolution temporelle réduite ou un masquage temporel plus important, 4) des caractéristiques compressives réduites au niveau périphérique, 5) une différence dans le RSB pour atteindre le SRP dans un bruit continu, et plus récemment 6) des capacités de résolution temporelle fine réduites (« reduced temporal fine structure processing »).

Comme il est reconnu que la différence entre les auditeurs normaux et ceux avec perte auditive s'accroît en présence d'un bruit intermittent comparativement à un bruit continu, il serait primordial de documenter cette

capacité d'écoute lors des mesures de perception de la parole dans le bruit. Un bruit intermittent dont le spectre est identique au bruit standard continu du protocole clinique du HINT permettra de documenter les différences entre ces deux populations et d'étudier l'effet du type de bruit en interaction avec la séparation spatiale. La plupart des études sur l'effet du taux d'intermittence du bruit sur la perception de la parole ont utilisé une présentation monaurale ou diotique sous écouteurs, ne permettant pas de documenter du même coup l'avantage fourni par la séparation spatiale de la parole et du bruit. Un bruit intermittent avec un taux d'interruption de 16 Hz est proposé, basé sur les travaux de Smits et Houtgast (2007) et Rhebergen et coll. (2006) dans lesquelles les stimuli de parole étaient des séries de chiffres ou des phrases, plutôt que des mots isolés. Les résultats cliniques issus d'un tel ajout permettront de documenter les capacités d'écoute des individus dans divers bruits et de mieux séparer les individus avec perte auditive entre eux ainsi que des auditeurs avec audition normale.

**Tableau 1** : Quantité de démasquage rapporté dans quelques études (BSP: Bruit continu de spectre de parole; DM: démasquage; AN: auditeurs normaux; PA: auditeurs avec perte auditive).

Étude	Bruit intermittent ou fluctuant	Niveaux de présentation	DM moyen (dB)	Différence moyenne dans le SRP des AN et PA (dB)
Bacon et coll. 1998	BSP multiplié par une onde carrée de 10 Hz	70 dBSPL	AN: 15 PA: 5 (étendue: 1-11)	BSP: 4 Bruit intermittent: 13
Peters et coll. 1998	BSP multiplié par les variations contenues dans l'enveloppe temporelle d'un interlocuteur	65 dBSPL	AN: 6.2 (jeunes); 4.3-4.7 (plus âgés) PA: 1-3	BSP: 2.4-6.3 pour PA jeunes et âgés comparativement aux jeunes AN Bruit fluctuant: 5.9-11.5 pour PA jeunes et âgés comparativement aux jeunes AN
Desloge et coll. 2010	BSP multiplié par une onde carrée de 10 Hz	65 et 80 dBSPL	AN: 9 et 14 (65 et 80 dBSPL) PA: jusqu'à 5 et 5.3 (65 et 80 dBSPL)	Données fournies non suffisantes pour effectuer le calcul
Hall et coll. 2012	BSP multiplié par une onde sinusoïdale de 10 Hz	86 dBA	AN: 4.9 PA: 1.2	BSP: 2.6 Bruit intermittent: 6.3
Bernstein et Grant, 2009	BSP multiplié par les variations contenues dans l'enveloppe temporelle d'un interlocuteur	Parole à 57 dBSPL (AN) ou à 87 dBSPL (PA)	AN: 6 PA: -1	BSP: 6 Bruit fluctuant: 13
Francart et coll. 2011	ICRA5-250 (variations contenues dans l'enveloppe temporelle d'un interlocuteur)	60 dBA	AN: 4.3-7.6 PA: 1.2-2.1	Données fournies non suffisantes pour effectuer le calcul
Strelyk et Dau, 2009	BSP multiplié par une onde sinusoïdale de 8 Hz	65 dB SPL	AN: 10.8 PA: 5	BSP: 4 Bruit intermittent: 10
Wagener et Brand, 2005	ICRA5 (variations contenues dans l'enveloppe temporelle d'un interlocuteur)	Divers	AN: 14.2-15.4 PA: 4.7-5	BSP: 2.9-4.4 Bruit fluctuant: 13.6

Une connaissance des capacités d'écoute d'un individu dans divers bruits comparativement à celle d'auditeurs normaux permet non seulement de guider les efforts de réadaptation, mais aussi d'instaurer des attentes réalistes. Aux fins d'évaluations audiologiques cliniques, des normes doivent d'abord être établies auprès de jeunes adultes avec audition normale. Le protocole clinique du HINT est normalement mené sous écouteurs, mais dans les cas où les personnes avec perte auditive portent des prothèses auditives, il peut être intéressant d'évaluer si celles-ci améliorent la perception de la parole dans les deux types de bruit, i.e. continu et intermittent. Comme le port de prothèses auditives est incompatible avec les écouteurs et que les données normatives sont différentes en champ libre, les mesures seront donc aussi effectuées en présence de 2 haut-parleurs disposés dans une cabine audiométrique, selon le protocole clinique associé au HINT, afin d'établir des données normatives utilisables avec les porteurs de prothèses auditives.

Le présent projet a donc comme objectif principal d'évaluer la performance de jeunes adultes avec audition normale aux différentes configurations spatiales du HINT, en utilisant le bruit standard modifié pour y inclure un taux d'intermittence de 16 Hz. Une première expérimentation permettra d'établir les normes sous écouteurs alors qu'une deuxième portera sur les normes en champ libre. Les données normatives seront établies avec le bruit modifié pour les trois conditions spatiales de bruit du HINT (bruit avant, bruit droit et bruit gauche).

## 2. PARTICIPANTS

Au total, 48 jeunes adultes avec audition normale ont participé à l'étude, soit 24 participants âgés entre 18 et 30 ans (moyenne = 24 ans; écart-type = 3.0 ans) pour la première expérimentation sous écouteurs et 24 participants âgés entre 18 et 35 ans (moyenne = 26 ans; écart-type = 3.6 ans) pour la deuxième expérimentation en champ libre. Tous rencontraient les critères d'inclusion suivants: 1) avoir comme langue maternelle et utiliser au quotidien le français canadien, 2) seuils auditifs  $\leq 15$  dB HL de 250 à 8000 Hz, 3) absence d'asymétrie entre les deux oreilles (définie comme une différence de seuils de 15 dB ou plus à trois fréquences ou une différence de 25 dB ou plus à une fréquence), 4) aucune différence plus grande que 20 dB dans les seuils aux fréquences voisines pour une même oreille, 5) résultats normaux à l'otoscopie et à la tympanométrie (tympanogramme de type A; compliance entre 0.3 et 1.7  $\text{cm}^3$ , pression entre -150 et +50 daPa et un volume du conduit auditif entre 0.6 et 2.6  $\text{cm}^3$ ) (Martin et Clark, 2003) et 6) histoire otologique négative (à l'exception de tubes de ventilation et infections d'oreille durant l'enfance). Des participants peu familiers avec des tests de perception de la parole dans le bruit ont été recrutés dans le milieu universitaire par l'entremise d'affiches de recrutement. Le projet a reçu l'approbation du comité

d'éthique de l'Université d'Ottawa.

## 3. PROCÉDURES

### 3.1 Dépistage audiométrique

En utilisant la méthode Hughston-Westlake (Hall et Mueller, 1997), les seuils auditifs ont été mesurés à l'aide d'un audiomètre portatif (InterAcoustics AC40) et des écouteurs TDH-39P, dans une cabine insonorisée IAC. L'intégrité de l'oreille externe et de l'oreille moyenne a été évaluée à l'aide d'un otoscope (WelshAllyn) et d'un tympanomètre (InterAcoustics GSI 38). Avant de participer à l'étude, les participants devaient lire une lettre d'information décrivant le projet, signer un formulaire de consentement et remplir un questionnaire d'histoire auditive.

### 3.2 Hearing in Noise Test (HINT)

Le HINT (installé sur un ordinateur IBM Pentium III couplé à la boîte HTD du HINT) a été administré dans une cabine insonorisée IAC, sous écouteurs (TDH-39P), et par l'entremise de deux haut-parleurs (FOSTEX PM0.5 MKII). Le système a été calibré selon les directives du fabricant à l'aide d'une oreille artificielle 6  $\text{cm}^3$  (Brüel et Kjaer Type 4152), d'un microphone de 1 po (Brüel et Kjaer Type 4144) et d'un sonomètre (Brüel et Kjaer Type 2235).

La version du HINT pour adultes francophones canadiens (Vaillancourt et coll., 2005) a été utilisée dans cette étude. Les 12 listes du test ont été utilisées afin de documenter les performances avec le bruit continu standard et avec le bruit modifié (intermittence de 16 Hz), selon le protocole suivant: 1) silence (2 listes), 2) bruit standard – continu: bruit avant (3 listes), bruit droit (1 liste), bruit gauche (1 liste) et 3) bruit modifié – intermittent: bruit avant (3 listes), bruit droit (1 liste), bruit gauche (1 liste). Les répétitions dans la condition de bruit avant ont permis de documenter la variabilité intra-individuelle. L'évaluation débutait toujours par une liste de pratique, suivie des 2 listes dans le silence. Par la suite, les participants ont été évalués dans le bruit continu ou le bruit intermittent en premier, ou vice-versa. L'ordre de présentation des conditions et des listes a été contrebalancé entre les participants. Le protocole et les paramètres standards du HINT ont été utilisés pour administrer le test.

Tel que spécifié plus haut, le bruit standard continu du HINT correspond au spectre moyen long terme de l'ensemble des 240 phrases qui composent le test. A l'aide du logiciel MatLab, ce bruit continu a été modifié en le multipliant par une onde carrée de 16 Hz, afin de générer un bruit intermittent. Afin de passer d'un bruit à l'autre lors de l'évaluation, les fichiers sonores des bruits continu et intermittent ont été interchangés

dans le dossier contenant les stimuli du test HINT. Lors des tests, les deux bruits ont été présentés à un niveau sonore à long terme de 65 dBA, alors que le niveau des phrases variait selon une méthode adaptative pour atteindre le SRP. En somme, les deux bruits ont un même spectre et niveau sonore à long terme, seul le décours temporel de l'enveloppe énergétique du signal est différent.

Les mêmes protocoles ont été utilisés dans les deux expérimentations sous écouteurs et en champ libre. Seule la méthode de présentation des stimuli aux participants a été différente. Dans la première expérimentation sous écouteurs, les signaux de parole et du bruit ont été filtrés par les fonctions de transfert azimutales du mannequin acoustique KEMAR ([www.gras.dk](http://www.gras.dk)) afin de produire des stimuli aux oreilles gauche et droite simulant une source virtuelle toujours vers l'avant pour la parole et une source virtuelle soit vers l'avant, le côté gauche ou le côté droit pour le bruit. Dans la deuxième expérimentation en champ libre, les signaux de parole et du bruit ont été transmis par l'entremise de deux haut-parleurs placés à 1 mètre du centre de la tête du participant, séparés par un angle de 90 degrés pour produire les mêmes configurations spatiales que pour la première expérimentation. Les stimuli en champ libre ont préalablement été filtrés pour réduire les effets spectraux dus à la réponse de la salle (p. ex. réflexions sonores) et des haut-parleurs. Les procédures ci-dessus sont conformes au protocole expérimental du HINT et décrites en plus de détail dans Soli et Wong (2008) et dans Vaillancourt et coll. (2005).

Finalement, puisque l'évaluation en champ libre est sensible aux mouvements de la tête, les participants ont été avisés de garder leur tête droite et de ne pas déplacer leur corps vers l'avant.

## 4. RÉSULTATS

Les performances moyennes et les écart-types sont présentées au tableau 2 pour chacun des 2 types de bruit dans les 3 conditions de bruit du HINT. On remarque que les résultats pour les trois essais de la condition de bruit avant sont relativement stables, qu'il y a une amélioration du seuil lorsque le bruit est déplacé de l'avant vers le côté gauche ou droit ainsi qu'une amélioration du seuil lorsque le bruit est intermittent plutôt que continu. La différence entre les 2 types de bruit pour chacune des conditions apparaît à la dernière colonne. Sous écouteurs, on note que l'intermittence produit une amélioration du seuil de l'ordre de 13 à 14 dB pour la condition bruit avant et d'environ 10 dB pour les conditions bruit côté. En champ libre, l'amélioration est moindre, soit d'environ 9 dB pour la condition bruit avant et 6 dB pour les conditions bruit côté.

La fiabilité test-retest des seuils obtenus peut être quantifiée en calculant l'écart-type intra-sujet pour les conditions du test qui ont été administrées à plusieurs

reprises, soit bruit avant continu et intermittent. Des écart-types intra-sujets de 1.1 dB et 2.0 dB ont été calculés pour le bruit continu et le bruit intermittent, respectivement, pour l'expérimentation sous écouteurs. En champ libre, les écart-types intra-sujets s'élèvent à 1.0 et 1.5 dB pour le bruit continu et le bruit intermittent, respectivement. Ainsi, dans les deux expérimentations, l'écart-type intra-sujet est plus élevé en présence du bruit intermittent.

**Tableau 2** : SRP moyen et écart-type pour le test HINT avec bruit continu et intermittent chez 24 jeunes adultes francophones avec audition normale.

Expérimentation 1 - écouteurs			
Condition de bruit	Continu – Moyenne en dB S/B (écart- type)	Intermittent – Moyenne en dB S/B (écart- type)	Différence en dB (continu – intermittent)
Avant 1	-3.1 (1.4)	-16.3 (3.0)	13.2
Avant 2	-3.0 (1.3)	-16.9 (3.1)	13.9
Avant 3	-3.0 (1.0)	-17.3 (2.8)	14.3
Gauche	-11.6 (1.4)	-21.1 (2.9)	9.5
Droit	-11.4 (1.9)	-21.8 (2.1)	10.4
Expérimentation 2 – champ libre			
Condition de bruit	Continu – Moyenne en dB S/B (écart- type)	Intermittent – Moyenne en dB S/B (écart- type)	Différence en dB (continu – intermittent)
Avant 1	-2.9 (1.3)	-11.4 (1.7)	8.5
Avant 2	-3.3 (0.9)	-11.7 (1.3)	8.4
Avant 3	-3.3 (0.6)	-12.2 (1.6)	8.9
Gauche	-8.6 (1.3)	-14.4 (2.3)	5.8
Droit	-9.4 (1.5)	-15.5 (2.0)	6.1

## 5. DISCUSSION

### 5.1 Comparaison des données normatives

Les performances au HINT avec bruit continu (bruit standard du test) obtenues sous écouteurs peuvent être comparées avec les données normatives du test publiées par Vaillancourt et coll. (2005), afin de vérifier la fiabilité des données expérimentales. Dans l'étude de Vaillancourt et coll. (2005), un protocole similaire avait été utilisé pour établir des données normatives sous écouteurs auprès de 36 jeunes (18-45 ans) canadiens francophones avec audition normale. Une telle comparaison, effectuée au tableau 3, témoigne de la similarité des deux séries de données. Pour toutes les conditions de bruit (bruit avant, droit et gauche) ainsi que pour le score composite [moyenne pondérée des 3 conditions de bruit selon la formule suivante: (2 x bruit



avant + bruit gauche + bruit droit)/4], on note des valeurs moyennes similaires aux données normatives, quoique les écarts-types soient légèrement supérieurs.

**Tableau 3** : SRP moyen et écart-type pour le test HINT avec bruit continu (standard) sous écouteurs chez 24 jeunes adultes francophones avec audition normale et aux données normatives établies par Vaillancourt et coll. (2005)

Condition de bruit	Moyenne en dB S/B (écart-type)	
	Bruit continu (présente étude)	Vaillancourt et coll. 2005
Avant	-3.1 (1.4)	-3.0 (1.1)
Gauche	-11.6 (1.4)	-11.4 (1.2)
Droit	-11.4 (1.9)	-11.5 (1.1)
Score composite	-7.2 (1.0)	-7.2 (0.8)

Une comparaison des données de la présente étude obtenues avec les deux méthodes de présentation des stimuli (tableau 4) témoigne de meilleurs résultats (seuils plus faibles) sous écouteurs, dans toutes les conditions du HINT. De telles différences sont en partie attribuables aux effets de la salle (p. ex. réflexions) qui affectent les mesures en champ libre et qui ne sont que partiellement éliminés par la méthode d'égalisation spectrale du HINT (Soli et Wong, 2008). On remarque que la différence entre les deux méthodes de présentation est plus prononcée pour le bruit intermittent que pour le bruit continu, un effet qui semble être lié à un moindre avantage de l'intermittence du bruit en champ libre que sous écouteurs (Section 5.4). Finalement, il est aussi à noter au tableau 3 que la différence entre les résultats sous écouteurs et en champ libre est légèrement plus prononcée dans les conditions de bruit côté que pour la condition de bruit avant, et ce pour les deux types de bruit. Ce phénomène est généralement observé avec le HINT, les fonctions de transfert du mannequin KEMAR utilisées pour traiter les stimuli sous écouteurs produisant un démasquage binaural virtuel souvent égal ou supérieur au démasquage binaural avec des sources physiques en champ libre (Soli et Wong, 2008). Tel qu'observée dans la présente étude, cette différence en démasquage binaural entre les conditions sous écouteurs et en champ libre est davantage marquée dans le cas du bruit intermittent que dans le cas du bruit continu. Il est fort possible que les réflexions sonores résiduelles dans la salle dégradent davantage le démasquage binaural en bruit intermittent qu'en bruit continu, étant donné qu'elles affecteront davantage les pauses sonores nécessaires à la reconnaissance de la parole en condition de bruit intermittent.

**Tableau 4** : SRP moyen pour le test HINT sous écouteurs et en champ libre avec bruit continu (standard) et bruit intermittent chez 24 jeunes adultes francophones avec audition normale

Type de Bruit	Condition	Moyenne en dB S/B		Différence en dB
		Sous écouteurs	En champ libre	
Continu	Avant	-3.1	-2.9	0.2
	Gauche	-11.6	-8.6	3.0
	Droit	-11.4	-9.4	2.0
Intermittent	Avant	-16.3	-11.4	4.9
	Gauche	-21.1	-14.4	6.7
	Droit	-21.8	-15.5	6.3

## 5.2 Écart-type intra-sujet

Puisque la condition bruit avant a été répétée à trois reprises avec les deux types de bruit, on a pu calculer la fiabilité test-retest ou l'écart-type intra-sujet. Sous écouteurs, la valeur de 2 dB obtenue pour le bruit intermittent est presque deux fois plus élevée que celle pour le bruit continu standard (1.1 dB). Cette dernière valeur est identique à celle obtenue lors de la publication des données normatives (Vaillancourt et coll. 2005). En champ libre, une plus grande valeur d'écart-type intra-sujets est également obtenue pour le bruit intermittent (1.5 dB) que pour le bruit continu (1.0 dB). Il semblerait donc que la mesure avec un bruit intermittent est plus variable que celle effectuée avec le bruit standard, une observation également documentée à plusieurs reprises (p. ex. Rhebergen et coll., 2008; Rhebergen et coll., 2006; Festen et Plomb, 1990; de Laat et Plomb, 1983; Middelweerd et coll., 1990; Versfeld et Dreschler, 2002; Peter et coll., 1998; Rhebergen and Versfeld, 2005).

## 5.3 Effet d'apprentissage

Pour la condition de bruit intermittent avant, on peut noter au tableau 1 une tendance à une légère amélioration du seuil pour chaque liste présentée (pour un total de trois listes). Un effet d'apprentissage semble donc être présent lorsqu'un bruit intermittent est utilisé. Une telle amélioration, à la suite de la répétition du test, n'est cependant pas retrouvée pour le bruit standard continu (voir tableau 1). Ce phénomène est comparable à celui mis en évidence par Rhebergen et coll. (2008), à savoir que le seuil s'améliore avec chaque liste subséquente pour se stabiliser après l'administration de 2-3 listes. Une telle amélioration des seuils peut atteindre 0.9 dB selon Rhebergen et coll. (2006) pour un bruit intermittent mais non pour un bruit continu. Les auteurs ont tenté de relier cette amélioration à un besoin d'entraînement pour maximiser la performance en

présence des périodes de silence dans le bruit intermittent. Une amélioration moyenne du même ordre de grandeur, soit 1.0 dB sous écouteurs et 0.8 dB en champ libre, a été observée dans la présente étude pour le bruit intermittent mais pas pour le bruit continu. Cet effet d'apprentissage est toutefois inférieur à l'écart-type intra-sujet (2 dB sous écouteurs et 1.5 dB en champ libre) établi pour cette condition (test-retest) avec le bruit intermittent et ne cause donc pas de problème particulier en pratique clinique.

## 5.4 Effet du type de bruit

L'étude de Wilson et Carhart (1969) a mis en évidence une amélioration de 28.3 dB dans un bruit intermittent comparativement à un bruit continu. Plus récemment, Wilson et coll. (2010) ont noté une amélioration du seuil de 34.0 dB, 30.2 dB et 28.4 dB pour les taux de 5 Hz, 10 Hz et 20 Hz, respectivement. Le matériel vocal (des mots) était présenté en monaural (oreille gauche ou droite) avec du bruit présenté à la même oreille (en ipsilatéral).

L'étude menée par Rhebergen et coll. (2006) est plus étroitement en lien avec la présente étude puisque le matériel vocal utilisé consiste en des phrases et le taux d'intermittence du bruit est de 16 Hz dans les deux cas. Une amélioration du SRP d'environ 10 dB a été notée. Dans la première expérimentation de la présente étude, une différence de 13 dB a été obtenue dans la condition bruit avant (voir tableau 1), un résultat qui se rapproche du 10 dB noté par Rhebergen et coll. (2006). La différence de 3 dB qui sépare les deux valeurs peut possiblement être expliquée par l'utilisation d'indices binauraux (en raison d'une administration binaurale des stimuli du test) dans la présente étude alors que les stimuli étaient présentés uniquement à la meilleure oreille (monaurale) dans Rhebergen et coll. (2006). Il est intéressant de noter que pour les conditions de côté, l'amélioration du SRP sous écouteurs est de 10 dB dans la présente étude.

Dans la deuxième expérimentation de la présente étude (champ libre), une différence entre les deux bruits de 9 dB et de 6 dB a été obtenue dans la condition bruit avant et bruit côté, respectivement. Encore une fois, il est important de noter que l'effet du type de bruit est légèrement moindre en champ libre que sous écouteurs, et ce, dans chacune des conditions d'écoute dans le bruit du HINT. Quoiqu'une différence dans la grandeur de l'effet du type de bruit soit notée entre les deux méthodes de présentation des stimuli, cette différence semble être équivalente pour les trois conditions d'écoute dans le bruit (soit environ 3 dB). La cause précise du moindre effet d'intermittence en champ libre demeure inconnue. Cependant, le phénomène pourrait fort bien être relié aux effets acoustiques de la salle tels que les réflexions sur les murs, le plancher et le plafond dans la cabine audiométrique. La réverbération

résiduelle du bruit dans la pièce, même faible, aura tendance à réduire l'ampleur de la modulation du bruit en dégradant les pauses sonores nécessaires à la reconnaissance de la parole en condition de bruit intermittent.

## 5.5 Effet de la séparation spatiale

Les données du tableau 1 permettent également de quantifier l'effet de la séparation spatiale, ou avantage binaural, soit l'amélioration des seuils lorsque la source de bruit est spatialement séparée de celle du bruit (bruit droit ou bruit gauche) comparativement au seuil obtenu pour une condition où les deux signaux proviennent de la même source (bruit avant).

Dans la première expérimentation (sous écouteurs), l'amélioration du seuil en présence d'un bruit continu s'élève à 8.3 dB pour la condition de bruit droit et 8.5 dB pour la condition de bruit gauche. Ces valeurs sont similaires à l'avantage binaural de 8 dB rapporté pour les valeurs normatives de l'étude de Vaillancourt et coll. (2005).

Pour le bruit intermittent, l'avantage binaural noté sous écouteurs dans la présente étude est d'environ 5 dB, soit près de la moitié moindre qu'en bruit continu. Le plus petit avantage binaural avec le bruit intermittent pourrait être expliqué par la nature même de ce bruit. Mis à part les effets de masquage antérieur et postérieur, aucun indice binaural n'est présent pendant les périodes de silence du bruit intermittent, et donc les individus peuvent moins profiter de la séparation spatiale. Les individus profitent seulement d'un avantage binaural lors des périodes où le bruit est présent (soit la moitié du temps). En général, les tests utilisent une présentation monaurale, avec le bruit présenté à l'oreille ipsilatérale (du même côté que la parole) (Wilson et coll., 2010; Wilson et Carhart, 1969; Rhebergen et coll., 2006). Puisque le HINT est administré en présentant les signaux aux deux oreilles (utilisation de l'audition binaurale), il est difficile de comparer les résultats.

Dans la seconde expérimentation, soit les tests d'écoute en champ libre, l'amélioration du seuil s'élève à 6.5 dB pour la condition de bruit droit et 5.7 dB pour la condition de bruit gauche en présence d'un bruit continu. Dans le bruit intermittent, l'avantage est de 4.1 dB pour la condition de bruit droit et 3.0 dB pour la condition de bruit gauche. Encore une fois, l'amélioration du seuil est moindre dans le bruit intermittent que dans le bruit continu.

Il est important de noter que l'effet de la séparation spatiale est légèrement moindre en champ libre que sous écouteurs, et ce, pour les deux bruits. Sous écouteurs, on obtient un effet d'environ 8 dB pour le bruit continu et 5 dB pour le bruit intermittent, comparativement aux valeurs en champ libre de l'ordre de 6 dB pour le bruit continu et 3-4 dB pour le bruit intermittent. Quoiqu'une différence dans la grandeur de

l'effet soit notée entre les deux méthodes de présentation, cette différence semble être équivalente pour les deux types de bruit (soit environ 2 dB). Tel que noté auparavant (Section 5.2), l'effet semble lié aux différences de fonctions de transferts binaurales dans les deux méthodes de présentation : celles du KEMAR pour tous les participants pour l'expérimentation sous écouteurs et les fonctions naturelles des participants pour l'expérimentation en champ libre.

## 5.6 Données normatives

Le tableau 5 fait état des données normatives du HINT sous écouteurs et en champ libre chez 2 groupes indépendants de 24 jeunes adultes francophones avec audition normale lorsque le bruit intermittent est utilisé. On y retrouve non seulement la moyenne et l'écart-type, mais également les limites inférieures et supérieures qui délimitent l'intervalle de confiance à 90% (soit  $\pm 1.65$  fois l'écart-type).

**Tableau 5** : Valeurs normatives sous écouteurs et en champ libre pour le HINT avec un bruit intermittent avec cadence de 16 Hz, établies auprès de deux groupes indépendants de 24 jeunes adultes francophones avec audition normale.

Expérimentation 1 - écouteurs			
Condition de bruit	Moyenne en dB S/B (écart-type)	Limite supérieure – 95 <sup>e</sup> percentile (en dB S/B)	Limite inférieure – 5 <sup>e</sup> percentile (en dB S/B)
Avant	-16.3 (3.0)	-21.3	-11.4
Gauche	-21.1 (2.9)	-25.9	-16.3
Droit	-21.8 (2.1)	-25.3	-18.3
Score composite	-18.9 (2.3)	-22.7	-15.1
Expérimentation 2 – champ libre			
Condition de bruit	Moyenne en dB S/B (écart-type)	Limite supérieure – 95 <sup>e</sup> percentile (en dB S/B)	Limite inférieure – 5 <sup>e</sup> percentile (en dB S/B)
Avant	-11.4 (1.7)	-14.2	-8.6
Gauche	-14.4 (2.3)	-18.2	-10.6
Droit	-15.5 (2.0)	-18.8	-12.2
Score composite	-13.3 (1.0)	-15.0	-11.7

Il est important de mentionner que les données normatives en champ libre ne sont strictement valides que pour la cabine audiométrique à l'étude. Ceci est tout particulièrement important dans le cas des valeurs normatives avec le bruit intermittent, lesquelles semblent être plus sensibles aux effets de réflexions résiduelles dans la salle qu'avec le bruit continu. En cas de bruit

intermittent, il est aussi important de s'assurer que le bruit de fond dans la pièce ne nuise pas à la reconnaissance de la parole durant les pauses de silence.

Quoique les résultats obtenus sous écouteurs sont sujets à moins de facteurs d'influence que ceux obtenus en champ libre, l'administration du HINT en champ libre est nécessaire lors de l'évaluation d'individus qui utilisent des prothèses auditives ou des protecteurs auditifs. Dans un tel cas, la position de l'individu dans le champ sonore par rapport aux haut-parleurs est critique car tout mouvement de la tête ou du corps peut influencer les niveaux sonores de la parole et du bruit atteignant l'oreille, et ainsi les résultats. Ce constat est d'autant plus vrai dans les conditions où les sources de parole et de bruit sont séparées spatialement.

## 6. CONCLUSION

Le but de ce présent projet était d'étudier l'effet du type de bruit (intermittent vs continu) sur le seuil de réception de phrases tel que mesuré au HINT sous écouteurs et en champ libre dans différentes configurations spatiales et d'établir des normes pour un bruit intermittent de 16 Hz avec ces deux méthodes de présentation des stimuli. Lorsque le HINT est effectué sous écouteurs, une amélioration du seuil d'environ 10 à 13 dB a été observée en présence du bruit intermittent comparativement à la performance dans le bruit continu standard. En champ libre, une telle amélioration du seuil est d'environ 6 à 9 dB. Aussi, le démasquage binaural est plus prononcé sous écouteurs qu'en champ libre par environ 2 dB et en bruit continu plutôt qu'en bruit intermittent par environ 3 dB.

Un léger effet d'apprentissage a également été noté lors de mesures répétées du seuil dans la condition de bruit avant, uniquement pour le bruit intermittent, un résultat similaire à celui rapporté dans l'étude de Rhebergen et coll. (2006). L'amélioration du seuil de 1.0 dB sous écouteurs et de 0.8 dB en champ libre observée dans la présente étude est toutefois inférieure à l'écart-type intra-sujet (2 dB sous écouteurs et 1.5 dB en champ libre) établi pour cette condition (test-retest) avec le bruit intermittent et ne cause donc pas de problème particulier en pratique clinique.

Les nouvelles données normatives maintenant disponibles pour le bruit intermittent de 16 Hz permettront d'évaluer des individus avec perte auditive dans les mêmes conditions expérimentales. Des données normatives en champ sonore doivent être établies pour chaque cabine audiométrique, en raison des effets de salle, alors que les normes sous écouteurs sont génériques et peuvent s'appliquer à toutes les cabines pourvu que les écouteurs soient bien calibrés. Sur la base des études citées précédemment, on peut poser l'hypothèse que les individus avec perte auditive ne profiteront pas autant de l'intermittence que les individus avec audition normale. Si cette hypothèse est vérifiée, on

pourra recommander d'ajouter au moins une autre condition au HINT, soit un bruit intermittent de 16 Hz provenant de l'avant, afin de documenter la capacité des individus à prendre avantage des périodes de silence dans le bruit pour reconnaître la parole. Un tel ajout nécessiterait environ 5 minutes supplémentaires à l'administration du test HINT et permettrait de mieux documenter les difficultés vécues au quotidien par les individus avec perte auditive. Afin de documenter l'avantage de la séparation spatiale avec le bruit intermittent, il faudrait toutefois inclure les conditions de bruit droit et gauche.

Finalement, il est important de préciser que les données normatives ont été établies chez de jeunes adultes. Par contre, tel que mentionné précédemment, des facteurs autres que la perte de sensibilité auditive peuvent également réduire les performances, incluant l'âge (Fitzgibbons et Gordon-Salant, 1996). L'établissement de normes auprès de personnes âgées serait alors souhaitable afin de quantifier la contribution de l'âge à la perception de la parole dans le bruit.

## RÉFÉRENCES

- Bacon SP, Opie JM, Montoya DY (1998). The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *J Speech Lang Hear Res*, 41:549-563.
- Baer T, Moore BCJ (1994). Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *J Acoust Soc Am*, 95:2277-2280.
- Bernstein JG, Brungart DS (2011). The effect of spectral and temporal-fine structure distortions on the fluctuating-masker benefit for speech at fixed signal-to-noise ratio. *J Acoust Soc Am*, 130:473-488.
- Bernstein JGW, Grant KW (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, 125:3358-3372.
- Bilger RC, Nuentzeq JM, Rabinowitz WM, Rzeczkowski C (1984). Standardization of a test of speech perception in noise. *J Speech Hear Res*, 27, 32-48.
- Bronkhorst AW, Plomp R (1992). Effect of multiple speech like maskers on binaural speech recognition in normal and impaired hearing. *J Acoust Soc Am*, 92:3132-3139.
- Buss E, Hall JW, Grose JH (2003). Effect of amplitude modulation coherence for masked speech signals filtered into narrow bands. *J Acoust Soc Am*, 113:462-467.
- Christiansen C, Dau T (2012). Relationship between masking release in fluctuating maskers and speech reception thresholds in stationary noise. *J Acoust Soc Am*, 132:1655-1666.
- Cooke MP (2003). Glimpsing speech. *J Phonetics*, 31:579-584.
- Cooke M (2006). A glimpsing model of speech perception in noise, *J Acoust Soc Am*, 119:1562-1573.
- Cox RM, Alexander GC, Gilmore C (1987). Development of the Connected Speech Test (CST). *Ear Hear*, 8:119S-125S.
- de Laat JAPM, Plomp R (1983). The reception threshold of interrupted speech for hearing-impaired listeners. In: R Klinke, R Hartman (Eds.). *Hearing: Physiological Bases and Psychophysics* (pp. 359-363), Berlin: Springer Verlag.
- Desloge JG, Reed CM, Braida LD, Perez ZD, Delhorne LA (2010). Speech reception by listeners with real and simulated hearing impairment: Effects of continuous and interrupted noise. *J Acoust Soc Am*, 128:342-359.
- Dubno JR, Horwitz AR, Ahistrom JB (2003). Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing. *J Acoust Soc Am*, 113:2084-2094.
- Duquesnoy AJ (1983). Effect of a single interfering noise or speech source on the binaural sentence intelligibility of aged persons, *J Acoust Soc Am*, 74:739-743.
- Eisenberg LS, Dirks DD, Bell TS (1995). Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *J Speech Hear Res*, 38:222-233.
- Festen JM, Plomp R (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am*, 88:1725-1736.
- Fitzgibbons PJ, Gordon-Salant S (1996). Auditory temporal processing in elderly listeners. *J Am Acad Audiol*, 7:183-189.
- Francart T, van Wieringen A, Wouters J (2011). Comparison of fluctuating maskers for speech recognition tests. *Int J Audiol*, 50:2-13.
- Freyman RL, Griffin AM, Oxenham AJ (2012). Intelligibility of whispered speech in stationary and modulated maskers, *J Acoust Soc Am*, 132(4):2514-2523.
- Füllgrabe C, Berthommier F, Lorenzi C (2006) Masking release for consonant features in temporally fluctuating background noise. *Hear Res* 211:74-84.
- George LJ, Festen J, Houtgast T (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing impaired listeners. *J Acoust Soc Am*, 120:2295-2311.
- Gnansia D, Jourdes V, Lorenzi C (2008). Effect of masker modulation depth on speech masking release. *Hear Res*, 239:60-68.
- G.R.A.S. Sound and Vibration. <http://www.gras.dk>
- Gustafsson HA, Arlinger SD (1994). Masking of speech by amplitude-modulated noise. *J Acoust Soc Am*, 95:518-529.
- Hall JW, Mueller GH (1997). *Audiologist Desk Reference Volume 1. Diagnostic Audiology*

- Principles, Procedures and Practices*. San Diego, CA: Singular Publishing Group Inc.
- Hall JW, Buss E, Grose JH, Roush PA (2012). Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. *Ear Hear*, 33(3):340–348.
- Hopkins K, Moore BCJ (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *J Acoust Soc Am*, 125:442–446.
- Howard-Jones PA, Rosen S (1993). The perception of speech in fluctuating noise. *Acustica*, 78:258–272.
- Jin SH, Nelson PB (2006). Speech perception in gated noise: The effects of temporal resolution. *J Acoust Soc Am*, 119:3097–3108.
- Kalikow DN, Stevens KN, Elliot LL (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J Acoust Soc Am*, 61:1337-1351.
- Killion MC, Niquette PA (2000). What can the pure-tone audiogram tell us about a patient's SNR loss? *Hear J*, 53:46–53.
- Killion, MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, 116:2395–2405.
- Kwon BJ, Perry TT, Wilhelm CL (2012). Sentence recognition in noise promoting or suppressing masking release by normal-hearing and cochlear-implant listeners. *J Acoust Soc Am*, 131:3111-3119.
- Kwon BJ, Turner CW (2001). Consonant identification under maskers with sinusoidal modulation: masking release or modulation interference? *J Acoust Soc Am*, 110:1130–1140.
- Lagacé J, Jutras B, Giguère C, Gagné J-P (2010). Development of the Test de phrases dans le bruit (TPB). *Can J Speech Lang Pathol*, 34:261-270.
- Laroche C, Giguère C, Vaillancourt V, Soli S (2005). Development and validation of hearing standards for Canadian Coast Guard Seagoing Personnel and C&P and land-based personnel. Phase II, Final report to Department of Fisheries and Oceans under Contract No. F7053-000009.
- Lorenzi C, Husson M, Ardoint M, Debrulle X (2006a). Speech masking release in listeners with flat hearing loss: Effects of masker fluctuation rate on identification scores and phonetic feature reception. *Int J Audiol*, 45:487–495.
- Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BCJ (2006b). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc Natl Acad Sci, USA* 103:18866–18869.
- Martin N, Clark SG (2003). *Introduction to Audiology*. 8<sup>th</sup> Edition. Boston: Allyn et Bacon.
- Miller GA (1947). The masking of speech. *Psychological Bulletin*, 44:105–129.
- Miller GA, Licklider JCR (1950). The intelligibility of interrupted speech. *J Acoust Soc Am*, 22:167-173.
- Middelweerd MJ, Festen JM, Plomp R (1990). Difficulties with speech intelligibility in noise in spite of a normal pure-tone audiogram. *Audiol* 29:1-7.
- Moore BCJ, Glasberg BR, Vickers DA (1995). Simulation of the effects of loudness recruitment on the intelligibility of speech in noise. *Br J Audiol*, 29:131-143.
- Moore R, Gordon-Hickey S (2011). Can behavioral speech-in-noise tests improve the quality of hearing aid fittings? *Perspect Audiol*, 7:8-14.
- Moore BCJ, Peters RW, Stone MA (1999). Benefits of linear amplification and multi-channel compression for speech comprehension in backgrounds with spectral and temporal dips. *J Acoust Soc Am*, 105:400–411.
- Mueller GH (2001) Speech audiometry and hearing aid fittings: going steady or casual acquaintances? *Hear J*, 54:19-29.
- Nelson PB, Jin S-H, Carney AE, Nelson DA (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *J Acoust Soc Am*, 113:961–968.
- Nilsson MJ, Soli SD, Sullivan J (1994). Development of a hearing in noise test for the measurement of speech reception threshold. *J Acoust Soc Am*, 9:1085–1099.
- Oxenham AJ, Simonson AM (2009). Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference. *J Acoust Soc Am*, 125:457–468.
- Peters RW, Moore BCJ, Baer T (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J Acoust Soc Am*, 103:577-587.
- Rhebergen KS, Versfeld NJ (2005). A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners. *J Acoust Soc Am*, 117:2181-2192.
- Rhebergen KS, Versfeld NJ, Dreschler WA (2006). Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise. *J Acoust Soc Am*, 106:3988-3997.
- Rhebergen KS, Versfeld NJ, Dreschler WA (2008). Learning effect observed for the speech reception threshold in interrupted noise with normal hearing listeners *Int J Audiol*, 47:185-188.
- Simpson S, Cooke MP (2005). Consonant identification in N-talker babble is a nonmonotonic function of N. *J Acoust Soc Am*, 118:2775–2778.
- Smits C, Kapteyn TS, Houtgast T (2004). Development and validation of an automatic speech-in-noise screening test by telephone. *Int J Audiol*, 43:15–28.
- Smits C, Houtgast T (2007). Recognition of digits in different types of noise by normal-hearing and

hearing-impaired listeners. *Int J Audiol*, 46:134-144.

Soli SD, Wong LLN (2008). Assessment of speech intelligibility in noise with the hearing in noise test. *Int J Audiol*, 47:356-361.

Strelcyk O, Dau T (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J Acoust Soc Am*, 125:3328-3345.

Takahashi GA, Bacon SP (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J Speech Hear Res*, 35:1410-1421.

Vaillancourt V, Laroche C, Mayer C, Basque C, Nali M, Eriks-Brophy A, Soli SD, Giguère C (2005). Adaptation of the HINT (hearing in noise test) for adult Canadian Francophone populations. *Int J Audiol*, 44:358-369.

Versfeld NJ, Dreschler WA (2002). The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners. *J Acoust Soc Am*, 111:401-408.

Wagener KC, Brand T (2005). Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters. *Int J Audiol*, 44:144-156

Wilson RH, Carhart R (1969). Influence of pulsed masking on the threshold for spondees. *J Acoust Soc Am*, 46:998-1010.

Wilson RH, McArdle R, Betancourt MB, Herring K, Lipton T, Chilsom TH (2010). Word-recognition performance in interrupted noise by young listeners with normal hearing and older listeners with hearing loss. *J Am Acad Audiol*, 7:183-189.



# Freedom Step

Convert a standard floor to a superior floor with the Freedom Step Acoustical & Impact Isolation Subfloor

**AcustiFloat®**  
Acoustical & Impact Subfloor Systems

**WILREP LTD.**

Tel. (905) 625-8944 Toll Free 1-888-625-8944

[www.acustifloat.com](http://www.acustifloat.com)

**Gym Rooms Playrooms Home Theaters Dance Floors**

AcustiFloat is a registered Trademark of WILREP LTD.

## High Quality CALIBRATION is a MUST When Accuracy is Critical!

Scantek provides:

- Quick calibration of ALL BRANDS of sound and vibration instruments and transducers:
  - ▶ Microphones
  - ▶ Preamplifiers
  - ▶ Sound level and vibration meters
  - ▶ Acoustical calibrators
  - ▶ Accelerometers & exciters
  - ▶ Windscreen characterization
- ISO 17025 accredited by NVLAP (NIST)
- Price Competitive
- Before & After data provided at no additional cost
- 48-hr turnaround accommodated

### Scantek, Inc.

Sound & Vibration Instrumentation and Engineering

[www.scantekinc.com](http://www.scantekinc.com)

[CalLab@ScantekInc.com](mailto:CalLab@ScantekInc.com)

**800-224-3813**

## When "BUY" does not apply, give RENTAL a try!

At Scantek, Inc. we specialize in **Sound and Vibration Instrument Rental** with *expert assistance*, and fully calibrated instruments for:

### Applications

- Building acoustics
- Sound power measurement
- Community noise
- Building vibration
- Industrial noise
- Human body vibration
- Machine diagnostics
- Vibration measurement

### Instruments

- analyzers •  
*FFT and real-time*
- *1/3 and 1/1 octave bands*
- noise and vibration dosimeters •
- vibration meters •
- human body dose/vibration •
- A-weighted sound level meters •
- rangefinders •
- GPS •
- windscreens •
- wide range of microphones •
- and accelerometers

### Scantek, Inc.

Sound & Vibration Instrumentation  
and Engineering

[www.scantekinc.com](http://www.scantekinc.com)

[info@scantekinc.com](mailto:info@scantekinc.com)

**800-224-3813**

## TAPPING just got easier!

The rugged brand new Norsonic N-277 Tapping Machine is ideal for making structureborne impact noise tests for floor/ceiling combination in the field and in the laboratory. This third-generation unit meets all international and US standards.

- Impact sound transmission testing according to ISO140 part VI, VII and VIII, ASTM E-492 and ASTM E-1007.
- Remote operation from hand switch or PC; Mains or battery operation.
- Low weight 10 kg (22 lb) incl. battery and wireless remote option.
- Built in self check of hammer fall speed, and tapping sequence for automatic calibration of major components.
- Retractable feet and compact size provide easy transportation and storage.



**Scantek, Inc.**  
Sound & Vibration Instrumentation  
and Engineering

[www.scantekinc.com](http://www.scantekinc.com)  
[info@scantekinc.com](mailto:info@scantekinc.com)  
**800-224-3813**

**NEW** TYPE 4448 PERSONAL NOISE DOSE METER

# Damaged hearing costs you dearly Preventing it doesn't



## TYPE 4448 FROM BRÜEL & KJÆR

Home of the world's  
best sound and vibration  
instrumentation

### Type 4448 – Helping to improve workplace noise assessment

#### Simple reliability

No cables, no connectors

#### Forget it is there

Secure shoulder mount with pin or clip attachment

#### Ready when you are

Long 28 hour battery-life

#### Verify your Standards compliance

HML option – verify hearing protection requirements

#### Works with Protector PC software – for intuitive analysis and reporting



[www.bksv.com/Type4448](http://www.bksv.com/Type4448)

BN 1054-11



# Concerto

4-Channel Multi-Function Acoustic Measuring System

All you need in one system:

- 4-channel SLM Class 1
- RT60, EDT, C80, D50 & Ts
- 4-channel Data Logger
- 4-channel Spectrum Analyzer
- Building and Human Vibration
- Monitoring Station with Remote Access

Custom Modules Available

See demo : [www.softdb.com/concerto.php](http://www.softdb.com/concerto.php)



Remote access via  
iPhone or PC

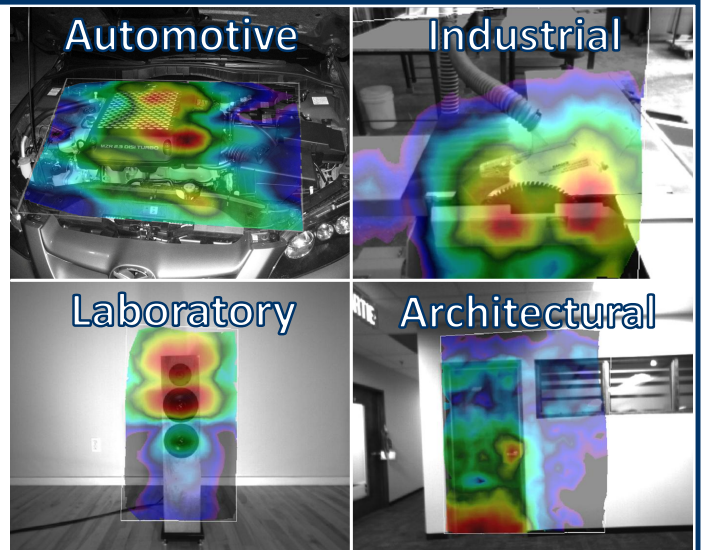


## I-Track

Automatic Real-Time Sound Mapping



See demo : [www.softdb.com/itrack.php](http://www.softdb.com/itrack.php)



**5-Minute Mapping**  
Freehand Scanning Without Grid

Efficient and Innovative Sound & Vibration  
Measurement Systems at a Competitive Price

**Soft dB**

[www.softdb.com](http://www.softdb.com)  
Toll free : 1 (866) 686-0993

# Better testing... better products.

## The Blachford Acoustics Laboratory

Bringing you superior acoustical products from the most advanced testing facilities available.



Our newest resource offers an unprecedented means of better understanding acoustical make-up and the impact of noise sources. The result? Better differentiation and value-added products for our customers.



### Blachford Acoustics Laboratory features

- Hemi-anechoic room and dynamometer for testing heavy trucks and large vehicles or machines.
- Reverberation room for the testing of acoustical materials and components in one place.
- Jury room for sound quality development.



### Blachford acoustical products

- Design and production of simple and complex laminates in various shapes, thicknesses and weights.
- Provide customers with everything from custom-engineered rolls and diecuts to molded and cast-in-place materials.

**Blachford** **QS 9000**  
REGISTERED

[www.blachford.com](http://www.blachford.com) | Ontario 905.823.3200 | Illinois 630.231.8300



# MISE À JOUR SUR LE DÉVELOPPEMENT DU TEST DE MOTS DANS LE BRUIT

Josée Lagacé<sup>1</sup>, Laudia LeBlanc<sup>2</sup>, Véronique Boisvert<sup>1</sup>, Marika Joëlle Arseneau<sup>1</sup>, Stéphanie Breau-Godwin<sup>1</sup>

<sup>1</sup>Programme d'audiologie et d'orthophonie, École des sciences de la réadaptation  
Pavillon Roger Guindon, Université d'Ottawa  
451, chemin Smyth, Ottawa, Ontario K1H 8M5

<sup>2</sup>Département d'audiologie, Centre hospitalier universitaire Dr-Georges-L.-Dumont  
330, avenue Université, Moncton, Nouveau-Brunswick E1C 2Z3

## RÉSUMÉ

Le *Test de Mots dans le Bruit* (TMB) est une épreuve franco-canadienne de reconnaissance de mots dans le bruit enregistrée sur cédérom. Les résultats d'une étude préliminaire (Lagacé, 2010) ont montré un effet d'âge sur les performances mesurées au TMB. Cet article fait état de la progression dans le développement de données normatives au service d'audiologie du Centre hospitalier universitaire Dr-Georges-L.-Dumont (Moncton, NB), auprès d'enfants (6-12 ans) et d'adultes (21-45 ans) francophones. En raison de l'effet possible des différences linguistiques régionales sur ce type d'épreuve, les performances mesurées auprès des adultes de la région de Moncton (NB) ont été comparées à celles d'adultes de la région d'Ottawa (ON). Les résultats suggèrent que le dialecte d'une population, aussi bien que l'âge, a un impact sur la capacité à reconnaître des monosyllabes présentés dans un bruit de fond. Le développement de normes pour le TMB auprès de différents groupes d'âge et de différentes communautés francophones du Canada contribue à réduire la pénurie d'outils cliniques standardisés disponibles pour cette population.

## ABSTRACT

The *Test de Mots dans le Bruit* (TMB) is a Canadian-French word recognition in noise test recorded on CD-ROM. Preliminary data (Lagacé, 2010) revealed an effect of age on TMB performance. This article presents a summary of the progress in the development of normative data at the Audiology department of the Dr. Georges-L.-Dumont University Hospital Centre (Moncton, NB) for both children (6-12 years old) and adults (21-45 years old). As regional linguistic differences can influence scores on this type of test, performances measured with adults from Moncton (NB) were compared to those obtained in adults from Ottawa (ON). Results suggest that, in addition to age, regional linguistic characteristics can influence one's ability to recognize monosyllabic words presented in background noise. The establishment of normative data applicable to different age groups and Francophone communities across Canada contributes to improving the lack of available standardized clinical tools for this population.

## INTRODUCTION

Il est bien connu que la reconnaissance de la parole est plus difficile en présence de bruit de fond qu'en condition silencieuse. Le traitement auditif, cognitif et linguistique requis pour la perception de la parole dans du bruit exigerait un effort d'écoute supplémentaire comparativement à ce qui est nécessaire dans des conditions acoustiques favorables (Pichora-Fuller, Schneider et Daneman, 1995). Grâce à cet effort d'écoute et aux habiletés de traitement cognitif, linguistique et auditif de l'information, l'auditeur ayant une sensibilité auditive normale arrive à percevoir la parole malgré l'effet masquant du bruit (Cole et Jakimik, 1980). Or, dans ces mêmes conditions sonores bruyantes, certains auditeurs présentent plus de difficultés à percevoir la parole que d'autres. Par exemple, les enfants obtiennent généralement de moins bonnes performances que les adultes dans diverses tâches de reconnaissance de la parole en présence de bruit (Elliott et coll., 1979; Fallon, Trehub, et Schneider, 2000; Johnson, 2000; Nittrouer et Boothroyd, 1990; Picard et Bradley, 2001). Pour expliquer ce phénomène, on invoque l'immaturation du système auditif (Fallon et coll., 2000), de même que celle des compétences cognitives et linguistiques (Elliott et coll., 1979; Nittrouer et Boothroyd, 1990).

D'autres études ont permis de documenter que certains groupes d'enfants ont plus de difficultés à percevoir la parole dans le bruit que leurs pairs du même âge. Par exemple, les enfants présentant une dyslexie (Ahissar, Lubin, Putter-Katz, et Banai, 2006; Brady, Shankweiler, et Mann, 1983), un trouble de traitement auditif (Bamiou, Musiek, et Luxon, 2001; Johnston et coll., 2009; Krishnamurti, 2007; Muchnik et coll., 2004), ou un trouble de langage (Stollman et coll., 2003; Ziegler et coll., 2005) ont des performances plus faibles que celles des groupes témoin à la mesure de reconnaissance de la parole dans le bruit, malgré des performances similaires dans le silence.

Le niveau de bilinguisme est un autre facteur d'influence. Par exemple, Tabri et coll. (2011) ont démontré que des adultes bilingues (depuis l'âge de cinq ans ou moins) présentaient des performances similaires à ceux d'adultes monolingues à la reconnaissance de mots en finale des phrases du test R-SPIN (Bilger, Nuetzel, Rabinowitz et Rzeczkowski, 1984) sans la présence bruit compétitif. Par contre, pour la même tâche d'écoute en présence d'un bruit de verbiage compétitif, les adultes bilingues ont eu des scores statistiquement inférieurs à ceux des adultes monolingues. Par exemple, au rapport signal-sur-bruit (s/b) de + 5 dB, la performance moyenne les adultes monolingues était de 71% alors qu'elle était de 56% pour les adultes bilingues (Tabri et coll., 2010).

Lorsqu'une personne éprouve des difficultés à comprendre la parole en présence de bruit, ceci peut avoir des répercussions importantes sur ses apprentissages, sa vie et celle de son entourage. L'audiologiste doit donc disposer d'outils permettant de mesurer et d'identifier ces difficultés

afin d'offrir une aide adéquate et personnalisée aux personnes et familles qui consultent en raison de difficultés à percevoir la parole dans du bruit. Malheureusement, il existe présentement une pénurie d'outils standardisés disponibles en français.

Le *Test de Mots dans le Bruit* (TMB) a été développé pour répondre à un besoin identifié par les audiologistes qui travaillent auprès des communautés francophones du Canada. Le test comporte quatre listes de 35 monosyllabes, dont le degré de difficulté est équivalent lorsque présentées dans un bruit de verbiage (*babble*) (Lagacé, 2010). L'enregistrement a été effectué par une locutrice, qui s'exprime en français canadien et qui avait déjà collaboré à d'autres enregistrements d'épreuves vocales. Chaque mot est précédé d'un signal avertisseur (son pur). Conformément aux procédures généralement utilisées en clinique pour ce type d'épreuve (DeBow & Green, 2000), il n'y a aucune étape de familiarisation lors de l'administration du TMB. Chaque mot est présenté qu'une seule fois (aucune répétition permise) au rapport s/b de +5 dB. D'après des données pilotes obtenues auprès de cinq participants ne faisant pas partie de l'étude, il a été déterminé que ce rapport s/b permettait d'éviter les effets de plancher et de plafond. Le bruit de verbiage francophone produit par Perrin et Grimault (2005) a été choisi en raison de ses caractéristiques similaires aux bruits de fond retrouvés de la vie courante comme dans les cafétérias, les restaurants, etc. Ce bruit de verbiage est préenregistré en boucle et est composé de quatre voix de femmes et de quatre voix d'homme. Les quatre listes de monosyllabes du TMB et le bruit de verbiage ont été enregistrés sur cédérom.

Puisque la familiarité des stimuli est un facteur important dans la reconnaissance de la parole dans du bruit (Epstein, Giolas, & Owens, 1968), celle des mots du corpus du TMB a été validée auprès d'enfants lors de la conception du test (Lagacé, 2010). Les listes de mots ont été présentées à 40 enfants de cinq à sept ans ayant une sensibilité auditive normale, recrutés auprès d'une école primaire d'Ottawa (NB) et de Montréal (QC). Le choix d'enfants aussi jeunes que cinq ans visait à assurer que le matériel soit adéquat pour évaluer les performances d'enfants de six ans et plus en clinique. Le recrutement a été effectué auprès de deux communautés francophones canadiennes pour faire en sorte que les mots sélectionnés soient représentatifs de l'ensemble de la population francophone au pays.

Comme il est important de tenir compte de l'effet d'âge lors de l'établissement de données normatives pour des épreuves de reconnaissance de la parole dans du bruit, une collecte de données a été amorcée au Centre hospitalier universitaire (CHU) Dr-Georges-L-Dumont (Moncton, NB). Les résultats d'une étude préliminaire ont montré un effet d'âge (Lagacé, 2010). Le taux de bonnes réponses des enfants de six ans était significativement inférieur à celui de tous les autres groupes à l'étude, soit les enfants de sept à 12 ans.

En plus de l'effet d'âge, on doit aussi tenir compte des variantes linguistiques culturelles lors du développement de

données normatives pour des épreuves de reconnaissance de la parole dans du bruit. Par exemple, Clopper et Bradlow (2008) ont comparé le taux de bonnes réponses à une tâche de reconnaissance du dernier mot de phrases auprès de trois groupes de 14 adultes anglophones unilingues des États-Unis. Le groupe du «Nord» était composé d'adultes ayant vécu dans le nord du pays depuis leur naissance. Le groupe d'«Américains moyens» était composé d'adultes ayant passé leur vie dans les régions dont le dialecte correspond à celui parlé en Nouvelle Angleterre et les états de l'ouest du pays (excluant les états du sud). Enfin, le groupe «Mobile» était composé d'adultes ayant vécu dans différentes régions des États-Unis depuis leur naissance. Quatre enregistrements ont été utilisés pour cette étude, soit un enregistrement produit par 1) un locuteur s'exprimant dans le dialecte typique du sud des États-Unis, 2) par un locuteur s'exprimant dans le dialecte typique du nord du pays, 3) par un locuteur s'exprimant dans le dialecte typique de l'est du pays et 4) un dernier enregistrement produit par un locuteur s'exprimant dans un anglais américain général. Ces quatre enregistrements ont été présentés avec un bruit à spectre vocal aux rapports s/b de -2 et -6 dB. De façon générale, le taux de reconnaissance était plus élevé au rapport s/b de -2 dB qu'au rapport s/b de -6 dB. Dans cette condition d'écoute difficile, le taux de reconnaissance le plus élevé a été noté pour l'enregistrement en anglais américain général et ce pour les trois groupes de participants. Pour le groupe du Nord, le taux de bonnes réponses étaient similaires pour l'enregistrement produit par le locuteur du Nord et celui du sud. Pour les trois groupes à l'étude, le plus faible taux de bonnes réponses a été observé pour l'enregistrement produit par le locuteur s'exprimant dans le dialecte de l'est du pays (pour plus de détails, voir Clopper et Bradlow, 2008). D'autres études ont permis de constater qu'en l'absence de bruit compétitif, le taux de reconnaissance de la parole est relativement similaire pour des enregistrements de dialectes différents, alors que ce taux est moins élevé lorsque des enregistrements sont présentée avec un bruit compétitif (Mason, 1946; Labov et Ash, 1997).

Il est déjà bien connu que le français parlé au Canada présente des variantes régionales importantes qui portent, entre autres, sur le vocabulaire usuel et sur l'articulation de certains phonèmes. Par exemple, le français couramment parlé dans la région du grand Moncton (NB), le «chiac», est particulièrement reconnu pour son mélange de français traditionnel et d'anglais (Boudreau, communication personnelle, janvier 2012). On y entend beaucoup d'expressions maritimes telles *abriér* (couvrir), *amarrier* (attacher), *caler* (s'enfoncer), etc. (Chevalier et Rodrigue, 2009). Par ailleurs, l'affrication des occlusives /t/ et /d/, si typique en français québécois, n'est presque pas observée dans le «chiac». De l'autre côté du Québec, chez les francophones de l'Est Ontarien, le français est aussi empreint d'anglais, la langue majoritaire de la province. Le français parlé par les francophones de l'Est de l'Ontario est relativement proche du franco-québécois en ce qui concerne l'accent et le vocabulaire. Bien que sommaire, cette

explication des variantes linguistiques du français parlé au Canada appuie l'importance de tenir compte de ce facteur pour l'établissement de données normatives pour le TMB.

Le présent article présente un compte rendu de la collecte de données normatives effectuée au service d'audiologie du CHU Dr-Georges-L-Dumont, à Moncton (NB) en date du premier janvier 2013, auprès de sept groupes d'enfants d'âge différent. Une comparaison des performances mesurées auprès des adultes de Moncton et d'Ottawa est également présentée pour mieux comprendre l'effet des différences linguistiques régionales sur ce type d'épreuve.

Le projet a été préalablement approuvé par le comité d'éthique en recherche du centre hospitalier où ces mesures ont été effectuées, de même que par celui de l'Université d'Ottawa. Le recrutement des enfants a principalement été effectué par le biais d'annonces affichées dans des centres de garde de la région, alors que les adultes et certains des enfants provenaient essentiellement de l'entourage des collaboratrices à ce projet.

## MESURES AVEC LE TMB

### 2.1. Collecte de données normatives auprès des enfants de six à 12 ans.

Les données normatives décrites dans cette section ont été établies suivant l'étude préliminaire ayant montré un effet d'âge sur les performances au TMB (Lagacé, 2010).

### Participants

En date du premier janvier 2013, 122 enfants (63 filles, 59 garçons), âgés entre six et 12 ans, ont participé à cette étude. Les caractéristiques de chaque groupe d'âge sont indiquées au Tableau 1. Pour fins d'inclusion dans l'étude, le français devait être la langue maternelle et celle d'usage quotidien, et le participant ne devait présenter aucun antécédent otologique, aucun trouble de développement du langage, d'apprentissage ou de développement général, ni de trouble cognitif, neurologique ou d'attention. Un questionnaire-maison a été employé pour documenter les informations au sujet de ces critères d'exclusion. Un dépistage audiométrique à 15 dB HL à 250, 500, 1000, 2000, 4000 et 8000 Hz a été effectué bilatéralement pour confirmer que l'acuité auditive était normale. De même, des mesures tympanométriques ont été effectuées pour assurer l'intégrité de l'oreille moyenne et des émissions otoacoustiques par produit de distorsion, pour assurer une fonction cochléaire normale.

### Procédure

La collecte de données a eu lieu dans une des cabines du service d'audiologie du CHU Dr-Georges-L-Dumont (Moncton, NB), à l'aide d'un audiomètre relié à un lecteur de disques compacts. Les quatre listes de 35 monosyllabes du TMB ont été présentées à tous les participants avec le

bruit de verbiage au rapport s/b de +5 dB (les monosyllabes à 60 dB HL et le bruit de verbiage à 55 dB HL) en condition ipsilatérale. La tâche consistait à répéter chaque mot, en devinant au besoin. Aucune familiarisation à la tâche n'a été effectuée et chaque mot n'était présenté qu'une seule fois. L'ordre de présentation des listes a été contrebalancé entre les participants, en suivant autant que possible le principe du carré latin. Les deux premières listes ont été entendues à l'oreille droite pour la moitié des participants et à l'oreille gauche pour l'autre moitié. Le pourcentage de bonnes réponses a été calculé pour chaque liste.

**Tableau 1.** Composition des groupes d'enfants en date du premier janvier 2013 et leur score moyen (pourcentage de réponses correctes) au TMB.

Groupe	Sexe	Âge moyen	Score moyen
6 ans	9 filles, 11 garçons	6 ans 6 mois	59.8% ( $\pm$ 8.7)
7 ans	11 filles, 9 garçons	7 ans 7 mois	63.5% ( $\pm$ 8.2)
8 ans	10 filles, 10 garçons	8 ans 7 mois	65.3% ( $\pm$ 8.5)
9 ans	12 filles, 7 garçons	9 ans 5 mois	65.9% ( $\pm$ 9.2)
10 ans	5 filles, 10 garçons	10 ans 6 mois	67.2% ( $\pm$ 7.3)
11 ans	8 filles, 6 garçons	11 ans 4 mois	69.5% ( $\pm$ 8.5)
12 ans	8 filles, 6 garçons	12 ans 6 mois	71.3% ( $\pm$ 9.4)

## Résultats

Le pourcentage moyen de reconnaissance de mots, compilé sur l'ensemble des listes, est présenté au Tableau 1 pour chacun des sept groupes d'enfants à l'étude.

Une ANOVA mixte pour mesures répétées à deux facteurs, dont le facteur GROUPE (7 niveaux) et le facteur LISTE (4 niveaux), a été effectuée. Les résultats démontrent un effet significatif du facteur LISTE [ $\Lambda$  de Wilks = .90,  $F_{(3,113)} = 4.31$ ,  $p = .01$ ,  $\eta^2 = .10$ ] et du facteur GROUPE [ $F_{(6,115)} = 6.92$ ,  $p = .00$ ,  $\eta^2 = .27$ ]. L'interaction entre les deux facteurs n'est pas significative [ $\Lambda$  de Wilks = .91,  $F_{(18,320)} = .59$ ,  $p = .91$ ,  $\eta^2 = .30$ ].

Bien que l'effet du facteur LISTE soit significatif, la différence notée entre la performance moyenne la plus élevée (Liste 1 = 67%) et la plus faible (Liste 4 = 64%) n'est que de 3%, ou approximativement un mot. En fait, un pourcentage de 2.86% est accordé à chaque mot d'une liste du TMB dans le calcul du pourcentage de réponses correctes. Cet effet significatif noté sur le plan de

l'équivalence des listes n'apparaît donc pas comme un facteur pertinent au plan clinique.

Des comparaisons multiples post-hoc (test *t*) ont été effectuées sur la base du pourcentage de réussite moyen calculé sur l'ensemble des quatre listes au sein de chaque groupe (données du Tableau 1), pour vérifier l'effet d'âge. Les résultats de ces analyses apparaissent au Tableau 2 et les différences statistiquement significatives sont représentées par un astérisque. Aucun ajustement de Bonferroni n'a été effectué pour les comparaisons afin d'observer les différences subtiles. Les résultats démontrent que les performances mesurées auprès des enfants de six ans (59.8%) sont significativement plus pauvres que celles de tous les autres groupes d'âge, à l'exception des enfants de sept ans (63.5%). À l'autre extrême, les performances mesurées chez les enfants de 12 ans (71.3%) sont significativement meilleures que celles des enfants de tous les groupes d'âge, à l'exception des enfants de 10 et 11 ans (67.2% et 69.5% respectivement). Par ailleurs, les performances des enfants de sept ans sont statistiquement différentes de celles des enfants de 10 et 11 ans. Selon les analyses, il n'y a pas de différence significative entre les groupes d'âge mitoyens, soit les enfants de huit, neuf et 10 ans. Par contre, les enfants de huit ans diffèrent des enfants de 11 ans.

**Tableau 2.** Résultats des comparaisons post-hoc entre les différents groupes. L'astérisque indique que la différence du pourcentage de réussite moyen est statistiquement significative.

Groupes	6 ans	7 ans	8 ans	9 ans	10 ans	11 ans
7 ans	$p = .199$					
8 ans	$p = .007^*$	$p = .117$				
9 ans	$p = .008^*$	$p = .099$	$p = .745$			
10 ans	$p = .001^*$	$p = .019^*$	$p = .282$	$p = .549$		
11 ans	$p = .000^*$	$p = .002^*$	$p = .029^*$	$p = .115$	$p = .245$	
12 ans	$p = .000^*$	$p = .001^*$	$p = .009^*$	$p = .038^*$	$p = .088$	$p = .423$

Selon les pratiques cliniques recommandées pour l'identification des difficultés auditives (ASHA, 2005; OOAQ, 2007; CISG, 2012), l'audiologiste doit appliquer un critère clair et stable pour faire la discrimination entre une capacité normale et une limitation. Ainsi, dans la plupart des cas vus en clinique, les performances mesurées sont comparées aux données normatives du test. Une performance considérée normale se situe à moins de deux écarts-types autour de la moyenne normative, alors qu'une

performance dépassant ce critère suggère une limitation. Les balises présentées au Tableau 3 ont été calculées en soustrayant la valeur de deux écarts-types de la moyenne obtenue pour chaque groupe d'âge. D'après ces calculs, il apparaît que les résultats obtenus par les enfants de six ans se démarquent de ceux de tous les autres groupes d'âge. L'étendue des résultats attendus est similaire pour les enfants de sept à neuf ans. Par exemple, la performance minimale qui peut être considérée normale pour un enfant de sept ans est de 48.9%, alors que celle des enfants de neuf ans est de 47.5%. De même, l'étendue des résultats des enfants de 10 à 12 ans est essentiellement la même. Par exemple, la performance minimale qui peut être considérée normale pour un enfant de 10 ans est de 52.6%, alors que celle des enfants de 12 ans est de 52.5%. Ainsi, lors de l'interprétation clinique, les résultats obtenus auprès des enfants de sept à neuf ans pourront être comparés à une balise, de même que celles obtenues auprès des enfants de 10 à 12 ans.

**Tableau 3.** Balises pour l'interprétation des données obtenues au TMB à partir des résultats obtenus en date du premier janvier 2013.

Groupe	Taux de réussite moyen	Écart -type	Taux de réussite moyen - 2 écarts-types
6 ans	59.8%	±8.7	42.4%
7 ans	63.5%	±8.2	48.9%
8 ans	65.3%	±8.5	48.3%
9 ans	65.9%	±9.2	47.5%
10 ans	67.2%	±7.3	52.6%
11 ans	69.5 %	±8.5	52.5%
12 ans	71.3%	±9.4	52.5%

## 2.2. Vérification de l'effet du dialecte sur les performances au TMB.

L'objectif de cette deuxième partie de l'étude était de vérifier l'effet du dialecte régional sur les performances au TMB chez les adultes. Pour ce faire, les données obtenues auprès d'adultes au CHU Dr.-Georges-L.-Dumont (Moncton, NB) ont été comparées à celles obtenues auprès d'adultes testés au Laboratoire de recherche en audiologie de l'Université d'Ottawa (Ottawa, ON).

### Participants

En date du premier janvier 2013, 24 adultes de la région de Moncton (NB), âgés entre 18 et 45 ans (moyenne = 33 ans) ont participé à cette étude, de même que 19 adultes de la région d'Ottawa (ON), dont l'âge moyen est de 23 ans. Les mesures ont été effectuées dans les deux régions concernées, les critères d'inclusion étant les mêmes que ceux présentés à la section précédente. Bien qu'aucune mesure formelle n'a été effectuée pour documenter le niveau du dialecte, l'expérimentatrice devait s'assurer que chaque participant s'exprimait en français typique de ces deux populations, soit le franco-acadien pour les participants du CHU Dr.-Georges-L.-Dumont, et le franco-

ontarien pour ceux du Laboratoire d'audiologie de l'Université d'Ottawa.

### Procédure

La même procédure que celle utilisée pour la collecte de données auprès des enfants, présentée dans la section précédente, a été suivie. Pour les participants vus à Moncton, il s'agissait de la même investigatrice et du même équipement que pour la collecte de données normatives effectuée auprès des enfants. Une autre investigatrice était responsable de la collecte de données effectuée auprès des adultes vus à Ottawa. Ces dernières ont été recueillies avec des équipements similaires à ceux de Moncton, dans une des cabines du Laboratoire de recherche en audiologie de l'Université d'Ottawa.

### Résultats

Pour évaluer l'effet des différences linguistiques régionales sur les performances au TMB, le pourcentage moyen obtenu sur l'ensemble des quatre listes chez ces deux populations d'adultes a été calculé (Tableau 4). La différence de performance entre les deux échantillons est de 8.5%, soit l'équivalent d'approximativement trois mots, et s'avère significative selon un test  $t$  [ $t(41) = 5.20, p = .000$  (bilatéral)].

**Tableau 4.** Pourcentage moyen et écart-type obtenus au TMB chez deux populations de francophones canadiens.

Groupe	Moyenne (écart-type)
Franco-ontarien	70.9% (±6.5)
Franco-acadien	79.4% (± 6.9)

## 3. DISCUSSION

Le TMB est composé de quatre listes de 35 monosyllabes, similaires au plan du degré de la familiarité des mots et de leur intelligibilité dans le bruit. Ce matériel vocal peut être utilisé avec les enfants dont le vocabulaire est jugé équivalent ou supérieur à celui des enfants de cinq ans, de même qu'auprès des adultes. Tel qu'attendu, les résultats de la présente étude montrent un effet d'âge qui devra être considéré dans l'interprétation clinique des résultats. Par contre, bien qu'une différence statistiquement significative soit notée entre certains groupes d'âge à l'étude, la différence entre le pourcentage moyen pour chacun des groupes ne dépasse que rarement 2.86%, qui est l'équivalent du pourcentage accordé à chaque mot du TMB dans le calcul du taux de réussite. Par exemple, le taux de réussite moyen est de 63.5%, 65.3 % et 65.9 % chez les enfants de sept ans, huit ans et neuf ans respectivement. Au plan clinique, cette différence dans les performances n'est pas significative puisqu'elle ne dépasse pas le pourcentage accordé à un seul mot (2.86%). L'ensemble des analyses suggère que les données des groupes d'enfants de sept, huit et neuf ans pourront être amalgamées lors de

l'établissement des normes finales, de même que ces des enfants de 10, 11 et 12 ans.

Ces résultats vont de pair avec ceux de Wilson et coll. (2010) dans l'établissement des données normatives pour le test *Words-in-Noise (WIN)* (Wilson, 2003; Wilson et McArdle, 2007), qui consiste aussi à une tâche de répétition de monosyllabes présentées dans un bruit de verbiage. Le but de cette épreuve, destinée aux auditeurs anglophones, consiste essentiellement à déterminer le rapport s/b auquel la personne peut identifier les mots correctement 50% du temps. Pour le test *WIN*, on note que le plus grand changement au plan de la reconnaissance de mots dans le bruit survient entre l'âge de six et sept ans (Wilson et coll., 2010). Les performances demeurent stables entre neuf et 12 ans, et les adolescents âgés de 17 ans obtiennent des performances légèrement meilleures que celles mesurées auprès des enfants de 12 ans.

La différence de 20% entre les performances des enfants de 6 ans et celles des adultes de la même région (Moncton, NB) s'apparente à celle obtenue avec la version franco-canadienne du HINT (Vaillancourt et coll., 2008). Selon les résultats de cette étude, une différence équivalente à 2.3 dB était notée entre les performances des enfants de 6 ans et celle des adultes. En utilisant la règle du 1 dB = 10% proposée par Picard et Bradley (2001), la différence de 20% entre les enfants de 6 ans et les adultes au TMB concorde bien avec les résultats obtenus avec le HINT franco-canadien pour enfants.

À la connaissance des auteurs, aucune étude n'a examiné l'effet des différences linguistiques régionales sur les performances à des épreuves cliniques employées en audiologie. Les études publiées portent plutôt sur l'effet des variantes linguistiques culturelles entre différentes nationalités et ce, essentiellement pour des épreuves en anglais. Des études révèlent que pour une épreuve d'écoute impliquant des stimuli verbaux, les performances sont différentes selon le pays d'origine des participants, alors que d'autres études ne montrent pas cette différence (Cameron, Barker et Newall, 2003; Dawes, 2011; Keith, Katbamna, Tawfik et Smolak, 1987; Marriage, King, Briggs et Lutman, 2001; Sockalingam et coll., 2004). Selon une étude de Clopper et Bradlow (2008), la reconnaissance de la parole d'un interlocuteur dont l'accent est familier serait plus précise que celle émise par un interlocuteur dont l'accent n'est pas familier lorsqu'en présence de bruit de fond. Bien qu'il soit impossible de généraliser le résultat de la présente étude en raison de la petite taille de l'échantillon, les données recueillies montrent une différence significative (8.46%) entre les performances mesurées auprès des adultes francophones de la région de Moncton (NB) et ceux de la région d'Ottawa (ON) pour les performances mesurées au TMB. Il est possible que cette différence soit liée aux limites méthodologiques, par exemple, la collecte de données a été effectuée par une personne et un équipement différents aux deux sites. D'autres données devront être cueillies pour vérifier cette tendance car les résultats d'un individu ne présentant pas de

difficulté d'écoute de la parole dans le bruit et ayant un dialecte régional différent de celui de l'épreuve de mesure utilisée pourraient être interprétés comme s'il avait un déficit réel. La qualité des traitements et recommandations serait, dans un tel cas, compromise, tout comme l'interprétation de résultats de recherche visant à documenter des différences ou des effets de traitement sur la perception de la parole dans le bruit.

#### 4. CONCLUSION

Le développement du TMB contribue à réduire la pénurie d'outils cliniques standardisés permettant d'évaluer les capacités d'écoute de la parole dans le bruit dans cette langue. En plus d'un effet d'âge, les résultats de cette étude suggèrent que le dialecte d'une population a un impact sur la performance à ce type d'épreuve. D'autres mesures doivent être effectuées avant son application clinique, soit compléter la collecte de données normatives et vérifier l'effet du dialecte sur les performances.

#### REMERCIEMENTS

Des remerciements sont adressés à Véronique Vaillancourt pour ses commentaires judicieux lors de la révision du manuscrit.

#### RÉFÉRENCES

- Ahissar, M., Lubin, Y., Putter-Katz, H. et Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nature Neuroscience*, 9, 1558-1564.
- American Speech-Language-Hearing Association (ASHA). (2005). *(Central) Auditory Processing Disorders*. [Technical Report]. Récupéré du site de l'association: <http://www.asha.org/members/deskrefjournals/deskref/default>
- Bamiou, D. E., Musiek, F. E. et Luxon, L. M. (2001). Aetiology and clinical presentations of auditory processing disorders: a review. *Archives of Diseases in Childhood*, 85, 361-365.
- Bilger, R.C., Nuetzel, J.M., Rabinowitz, W.M. et Rzeckowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research*, 27, 32-38.
- Boudreau, A. (2012). (Communication personnelle). La définition du «chiac». Université de Moncton, Moncton, Nouveau-Brunswick.
- Brady, S., Shankweiler, D. et Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology*, 35, 345-367.
- Cameron, S., Barker, R. et Newall, P. (2003). The effect of linguistic background on the Macquarie Pediatric Speech Intelligibility Test. *The Australian and New Zealand Journal of Audiology*, 25, 95-98.
- Canadian Interorganizational Steering Group for Speech-Language Pathology and Audiology (CISG). (2012). *Canadian guidelines on auditory processing disorder in*



- children and adults. Récupéré du site de l'association: [http://www.cshhpb.org/docs/canadian\\_guidelines\\_on\\_auditory\\_processing\\_disorder\\_in\\_children\\_and\\_adults\\_english\\_final\\_2012.pdf](http://www.cshhpb.org/docs/canadian_guidelines_on_auditory_processing_disorder_in_children_and_adults_english_final_2012.pdf)
- Chevalier, G. et Rodrigue, L. (2009). *Les mots pour parler des maux : Lexique différentiel des termes acadiens liés à la santé*. Consortium national de la formation en santé: Moncton.
- Clopper, C. G. et Bradlow, A. R. (2008). Perception of Dialect Variation in Noise: Intelligibility and Classification. *Language and Speech*, 51, 175-198.
- Cole, R.A., et Jakimik, J. (1980). A model of speech perception. Dans R. Cole (dir.), *Perception and production of fluent speech* (p. 133-163). Hillsdale, NJ: Erlbaum.
- Dawes, P. (2011). The SCAN-A in testing for auditory processing disorder in a sample of British adults. *International Journal of Audiology*, 50, 107-111.
- DeBow, A. et Green, W. B. (2000). A Survey of Canadian Audiological Practices: Pure Tone and Speech Audiometry. *Journal of Speech-Language Pathology and Audiology*, 24, 153-161.
- Elliott, L. L., Connors, S., Kille, E., Levin, S., Ball, K. et Katz, D. (1979). Children's understanding of monosyllabic nouns in quiet and in noise. *Journal of the Acoustical Society of America*, 66, 12-21.
- Epstein, A., Giolas, T. G. et Owens, E. (1968). Familiarity and intelligibility of monosyllabic word lists. *Journal of Speech, Language, and Hearing Research*, 11, 435-438.
- Fallon, M., Trehub, S. E. et Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *Journal of the Acoustical Society of America*, 108, 3023-3029.
- Johnson, C. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language and Hearing Research*, 43, 144-156.
- Johnston, K., John, A. B., Kreisman, N. V., Hall, J. W. et Crandell, C. C. (2009). Multiple benefits of personal FM system use by children with auditory processing disorder (APD). *International Journal of Audiology*, 48, 371-383.
- Keith, R. W., Katbamna, B., Tawfik, S. et Smolak, L. H. (1987). The effect of linguistic background on staggered spondaic word and dichotic consonant vowel scores. *British Journal of Audiology*, 21, 21-26.
- Krishnamurti, S. (2007). Monaural low-redundancy speech tests. Dans F. E. Musiek et G. D. Chermak (dir.), *Handbook of (Central) Auditory Processing Disorders. Auditory Neuroscience and Diagnosis. Volume 1*. (p. 193-206). San Diego: Plural Publishing.
- Labov, W. et Ash, S. (1997). Understanding Birmingham. Dans C. Bernstein, T. Nunnally, et R. Sabino (éditeurs), *Language variety in South Revisited* (p. 508-573). Tuscaloosa: University Press.
- Lagacé, J. (2010). Développement du Test de Mots dans le Bruit : Mesure de l'équivalence des listes et données préliminaires sur l'effet d'âge. *Acoustique canadienne*, 38, 19-30.
- Leclerc, J. (2012). *L'aménagement linguistique dans le monde: Ontario*. Récupéré le 3 février 2013 du site <http://www.tlfg.ulaval.ca/axl/amnord/ontario.htm>.
- Marriage, J., King, J., Briggs, J. et Lutman, M. E. (2001). The reliability of the SCAN test: results from a primary school population in the UK. *British Journal of Audiology*, 35, 199-208.
- Mason, H.M. (1946). Understandability of speech in noise as affected by region of origin of speaker and listener. *Speech Monographs*, 13, 54-68.
- Muchnik, C., Roth, D. A.-E., Othman-Jebara, R., Putter-Katz, H., Shabtai, E. L. et Hildesheimer, M.. (2004). Reduced Medial Olivocochlear Bundle System Function in Children with Auditory Processing Disorders. *Audiology and Neurootology*, 9, 107-114.
- Nittrouer, S. et Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and adults. *Journal of the Acoustical Society of America*, 87, 2705-2715
- Ordre des orthophonistes et audiologistes du Québec (OOAQ). (2007). *Révision des pratiques courantes en audiologie entourant le trouble de traitement auditif (TTA) chez l'enfant*. Rapport du comité ad hoc. Montréal: Canada.
- Perrin, F. et Grimault, N. (2005). *Fonds sonores*. Laboratoire Unités Mixtes de Recherche, Centre National de la Recherche Scientifique 5020, Lyon, France.
- Picard, M. et Bradley, J. S. (2001). Revisiting speech interference in classrooms. *Audiology*, 40, 221-244.
- Pichora-Fuller, K. M., Schneider, B. et Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America*, 97, 593-608.
- Sockalingam, R., Morrison, R., Cole, D., James, S., Morin, S. et Potter, S. (2004). Comparing SCAN-A scores between speakers of standard Australian English and American English: A preliminary study. *The Australian and New Zealand Journal of Audiology*, 26, 110-115.
- Stollman, M. H. P., van Velzen, E. C. W., Simkens, H. M. F., Snik, A. F. M. et van den Broek, P. (2003). Assessment of auditory processing in 6-year-old language-impaired children. *International Journal of Audiology*, 42, 303-310.
- Tabri, D., Smith Abou Chacra, K. et Pring, T. (2010). Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language and Communication Disorders*, 46, 1-12.
- Vaillancourt, V., Laroche, C., Giguère, C. et Soli, S.D. (2008). Establishment of age-specific normative data for the Canadian French version of the Hearing in Noise Test for Children. *Ear & Hearing*, 29, 453-466.
- Wilson, R. H. (2003). Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *Journal of the American Academy of Audiology*, 14, 453-470.

- Wilson, R. H., Farmer, N. M., Gandhi, A., Shelburne, E. et Weaver, J. (2010). Normative data for the Words-in-Noise Test for 6- to 12-year-old children. *Journal of Speech, Language, and Hearing Research*, 43, 1111-1121.
- Wilson, R. H. et McArdle, R. (2007). Intra- and inter-session test, retest reliability of the Words-in-Noise (WIN) Test. *Journal of the American Academy of Audiology*, 18, 813-825.
- Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F.-X., et Lorenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proceedings of the National Academy of Sciences, USA*, 102, 14110-14115.



www.odeon.dk

# Discover new heights in acoustics design



Odeon Room Acoustics Software

**Canadian Acoustical Association**  
**Minutes of the Board of Directors Meeting**  
Montreal, Quebec  
June 4, 2013

Present: Christian Giguère (chair), Jérémie Voix, Hugues Néliste, Stan Dosso, Dalila Giusti, Chantal Laroche, Roberto Racca, Kathy Pichora-Fuller, Frank Russo, Sean Pecknold, Bryan Gick

The meeting was called to order at 7:15 p.m. Minutes of the previous Board of Directors' meeting on October 9<sup>th</sup> 2012 were approved as published in the December 2012 issue of *Canadian Acoustics*. (*Moved by R. Racca, seconded by J. Voix, carried*).

**President's Report**

Christian Giguère thanked everyone involved in the several initiatives that have been launched in the past year, such as the online membership database and dues payment system, the online publishing of *Canadian Acoustics*, the establishment of two new CAA Awards, and the revamped website. In addition to providing better services to our members, these initiatives will facilitate the management of the Association and increase its visibility. After serving for several years as President, Christian indicated that it is time for a change-over and he indicated he will not seek re-election at the AGM.

*(Approval of report moved by S. Pecknold, Seconded by F. Russo, carried)*

**Secretary's Report**

Chantal Laroche reported that the new system put in place by Jeremie Voix is officially in place since December 2012. CAA now accepts online payments. Very few members are still paying by cheque or by fax. There were a few bugs to deal with, but as of June 4<sup>th</sup>, membership and renewals are up by 25 from last year (see Table). There are 445 active members. With respect to routine CAA communications:

- Annual filing with Corporations Canada was submitted
- Renewal to I-INCE has been paid by Dalila Giusti for 2013
- Access Copyright: \$94.10 received in October 2012

Category	Paid 2011	Change from 2012
Member	290	16
Emeritus	1	0
Student	78	4
Sustaining subscriber	51	4
Indirect subscribers		
- Canada	9	0
- USA	6	0
- International	5	0
Direct subscribers	5	1
<b>Total</b>	445	+25

Secretarial operating costs from September 1<sup>st</sup> 2012 to May, 31<sup>st</sup> 2013 totaled \$458.89. The secretary is requesting a budget of \$500 to reimburse postal box fees and to cover expenses for the rest of the fiscal year. (*Moved by S. Pecknold, Seconded by D. Gick, carried*)

The Treasurer, Dalila Giusti, has arranged for the Association to subscribe to Directors' liability insurance, and the Board members are now covered against eventual fraud.

For indirect USA and international subscribers, it was proposed to charge an extra amount to cover the mailing cost of the journal. The amount will be determined at a later date. These members will be

informed of this change during the membership renewal period.

A 6-month free promotional offer was announced during ICA 2013. It was proposed that new enrolled members received a free paper copy of the September issue via mail. (not seconded). It was then proposed that people who signed for a free offer would receive an online copy of Canadian Acoustics for 6 months (Moved C. Laroche, Seconded K.P.-Fuller, carried).

*(Approval of report moved by D. Giusti, seconded by R. Racca, carried)*

### **Treasurer's Report**

The Treasurer, Dalila Giusti, submitted a report including a preliminary financial statement for the fiscal year. CAA finances are in reasonable standing. The 2012 Conference (Banff) made a substantial profit. Overall, Association revenues well exceeded expenses and the costs for student awards.

Dalila will transfer \$40,000 from the Operating Account to the Capital Account to cover Student Awards. \$9,450 has been distributed in awards at the 2012 Conference.

Travel funds for students were discussed. It was proposed to cover 50% of traveling cost to a maximum of \$250 per student (pending submission of an estimate budget at the date of abstract submission) and a maximum cap amount of \$7,500 for the next two years, funds permitted. (Moved by R. Racca, Seconded S. Dosso, carried). Kathy has offered to review the generic form used for travel funds request.

The CAA fiscal year end has been moved to December 30.

*(Approval of report moved by B. Gick, seconded by S. Pecknold, carried)*

### **Editor's Report**

Frank Russo mentioned that the Editorial Board counts 15 members who have been renewed/replaced (no vacancies) and

committed for 2 years from January, 2013. An international Editorial team is considered in the near future. An Advisory Board has been suggested by Kathy to increase the visibility of the Journal.

The reviewing time is approximately 8 weeks. There are 6 papers in press for the June issue (coordinated by Josée Lagacé) and 5 papers are currently under review. 6 more papers are in the online system. The Journal was available online just in time for the 40<sup>th</sup> anniversary (Phase 1). The new website and logo are also ready (Phase 2) and the cost was \$1600 as budgeted. The initiation of new website and logo were approved through emails. Sean has created a mailbox ([clerk@caa-aca.ca](mailto:clerk@caa-aca.ca)) in order to track down the decisions made between face-to-face meetings.

It is hoped that there will be an agreement with EBSCO by the end of June. It was proposed that once we have an agreement in place, two Executive Board members will sign the agreement (Moved by F. Russo, Seconded by B. Gick, carried).

The September issue of Canadian Acoustics will feature contributions from new scholars. For the purposes of this issue, a new scholar is defined as someone who is within 10 years of the terminal degree at the time of submission (this pertains to the corresponding author only). Authors who have not previously published articles in the journal are especially welcome to submit. July 1 is the submission deadline.

According to Frank, we should consider submitting an application within the next 2 years to become a journal with a listed impact factor (IF). Many authors refer to IF before considering submission.

### **CAA Conferences – Past, Present & Future**

2012 (Banff): Stan Dosso reported that there were approximately 100 papers with 136 attendees and a profit of \$21,299.54.

2013 (Montreal, ICA 2013): Mike Stinson reported *that they are* approximately 2500 participants (600 students) at the meeting. This is one of the largest meetings in Acoustics ever held.

2014 (Winnipeg): Karen Turner is in charge of the local organization. She is collecting information about hotel venues. She will also need to finalize her committee. A few names were suggested, Christian Giguère will do the follow up. Jeremie Voix mentioned that a separate conference system, similar to the system now hosting the journal, could be used for registration and handling of abstracts.

2015: Halifax has been suggested. To be discussed with local people.

### **Awards**

Hugues Nélisse presented a report summarizing decisions from all individual prize coordinators. Eligible applications were available for all awards and winners have been selected, except for the Student Prize in Architectural Acoustics (no candidacy) and the Bell Student Prize in Speech Communication and Speech (1 candidacy but not eligible). Winners will be announced at the joint ICA/ASA/CAA Awards Ceremony on June 5<sup>th</sup> and in the September issue of *Canadian Acoustics*. The Directors' award will be selected later as the process has not been launched yet. Two new awards have been created (Bregman Prize in Psychological Acoustics and Northwood Prize in Architectural and Room Acoustics). A motion was put forward that Aeracoustics would sponsor a new award after John Bradley's name and that an ad-hoc committee would work on the exact name and description of the award as well as the financial terms.

*(Moved by KP Fuller. Seconded by B. Gick, carried)*

### **Acoustical Standards Committee**

Christian Giguère presented a report on behalf of Tim Kelsall, the Chair of the Acoustical Standards Committee. The Committee produced the "CAA Guide 101 to Acoustical Standards" and asked that the BoD endorses it and approves its posting on the CAA website. A long conversation followed. The Board

members asked that the Committee clarifies its expectations and explain the kind of endorsement that is sought from the CAA BoD, especially as it may expose the Board to potential liability issues. *(Moved by D. Giusti, Seconded by S. Pecknold, Carried)*. Christian said he will talk to Tim Kelsall about this as well as the need to formalize the terms of reference for the Committee now that documents are being produced.

### **CAA Website**

Jérémie Voix reported on the Journal Website and Membership Database migration effort that is still within budget (\$6,500 budgeted, \$5000 expended) and ahead of or within schedule. The journal online migration was completed by December 24<sup>th</sup> 2012 ahead of schedule (French was to be unrolled in June 2013). The Automated membership management was debugged in time for automated renewal emails to be sent on February 1<sup>st</sup>, 2013. The Online editorial system is live since January and used since February. All journal back-issues were digitized and put online, this is 161 issues, 2149 articles, and more than 15,000 pages that have been put online just before Christmas 2012! This huge effort was made possible thanks to the relentless scanning, coding and scripting work of Dr. Eugen Popovici (more than 100 hours of PC crunching time and 350 hours of hard work); the personal hard-copies archives kindly lent (in chronological issue order) by Dr. John Bradley, Dr David Quirt, Dr. Ramani Ramakrishnan, Dr. Christian Giguère and Dr. Chantal Laroche; the mental support and encouraging words of Dr. Bryan Gick and Dr. Frank Russo; the financial support of the Canadian Acoustical Association enabled by its Board of Directors; the frantic & friendly supervision of Dr. Jérémie Voix. Finally the journal is now indexed by Google Scholar and Open Access Initiative and now gets excellent coverage from search engines and publishing databases.

The remaining item is to finalize the journal online publishing Latex-XML transformation scripts with Mr popovici. Jeremie Voix requested a little extra budget of \$600 for this. *(Approved.)*

Jeremie Voix also mentioned the need to move the

JCAA website to the CAA legacy servers with Sean Peckold's help.

Sean Peckold reported that some Sustaining Subscribers complained that their logos were sometime missing or out of date. It is suggested that the secretary could ask the sustaining members each time he/she receive an automated subscription notification and post that into the "note" field of the subscription info.

### **Other business**

CAA Logo: Dalila Giusti explains the history of the blue logo. Frank Russo proposed to adopt the new logo. It is decided to go for another round of edits.

---

ICA Special: Jérémie reported on the ICA initiative (upload of all individual complete PDF of past issues, printing of the posters for the CAA table, printing of the pamphlet in every congress bag, work at the booth) and informed the Board that he called for some help and hired Ms Rebecca Reich on his remaining \$1500 budget. More than 50 ICA congress members subscribed to the 6-month free online subscription.

### **Adjournment**

Meeting adjourned at 00:15a.m. (*Moved by J. Voix, seconded by [R. Racca](#), carried*)

## **Canadian Acoustical Association**

### **Minutes of Annual General Meeting**

Montreal, Quebec  
June 5, 2013

#### **Call to Order**

President Christian Giguère called the meeting to order at 6:50 p.m. with 18 members in attendance, and presented the Agenda for acceptance (*Moved by A. Behar, seconded by S. Dosso, carried.*)

Minutes of the previous Annual General Meeting held on October 11, 2012 in Banff were approved as printed in the December 2012 issue of *Canadian Acoustics*. (*Moved by J. Voix, seconded by S. Abel, carried*)

#### **President's Report**

Christian Giguère briefly summarized his report to the Board meeting on June 4<sup>th</sup>. Christian Giguère mentioned several initiatives that have been launched in the past year, such as the online membership database and dues payment system, the online publishing of *Canadian Acoustics*, the establishment of two new CAA Awards, and the revamped website. After serving for several years as President, Christian indicated that it is time for a change-over.

#### **Secretary's Report**

Chantal Laroche provided an overview of membership and operational activities.

- The total of 445 paid renewals and new memberships is up from last year, presumably due to reminders sent by email to members.
- Jeremie Voix has put in place a new system that allows CAA members to pay online.

*(Acceptance of Secretary's report moved by N. Ellaham, seconded by S. Abel, carried.)*

#### **Treasurer's Report**

On behalf of Dalila Giusti, Christian Giguère presented an overview of her written report to the Board on CAA finances. CAA is in good

financial standing, with \$231,527 in TD Canada Trust fund and \$122 123 in operating fund. The Banff Conference has generated \$21,299 in net revenue. The fiscal year now ends on December 30<sup>th</sup>.

The Treasurer's report was accepted. (*Moved by A. Behar, seconded by S. Dosso, carried*)

#### **Editor's Report**

Frank Russo mentioned that the Editorial Board counts 15 members. An international Editorial team will be considered in the near future. An Advisory Board is now in place with Jérémie Voix, Brian Gick and Frank Russo.

The reviewing time is approximately 8 weeks. There are 17 papers in queue for publication. The Journal is now available online, and the new website and logo are also ready.

A 6 month free promotional offer was announced during ICA 2013. New subscribers will freely receive an online copy of *Canadian Acoustics* for 6 months.

It is hoped that there will be an author agreement signed with EBSCO by the end of June. The September issue of *Canadian Acoustics* will feature contributions from new scholars. July 1 is the submission deadline.

It is foreseen that an application will be submitted within the next 2 years to become a journal with a listed impact factor (IF).

All journal back-issues were digitized and put online (161 issues, 2149 articles, and more than 15,000 pages). This huge effort was made possible thanks to the relentless work of Dr. Eugen Popovici and the personal hardcopy archives kindly lent by Dr. John Bradley, Dr. David Quirt, Dr. Ramani Ramakrishnan, Dr. Christian Giguere and Dr. Chantal Laroche. Ms Rebecca Reich was also thanked for her contribution to the online system and her presence at the CAA table during ICA 2013.

#### **Award Coordinator's Report**

Hugues Nélisse presented a report summarizing decisions from all individual prize

coordinators. Eligible applications were available for all awards and winners have been selected, except for the Student Prize in Architectural Acoustics (no candidacy) and the Bell Student Prize in Speech Communication and Speech (1 candidacy but not eligible). Winners have been announced at the ICA Awards Ceremony on June 5<sup>th</sup> and will appear in the September issue of *Canadian Acoustics*. Two new awards have been created (Bregman Prize in Psychological Acoustics and Northwood Prize in Architectural and Room Acoustics).

### **Past and Future Meetings**

2012 (Banff): Stan Dosso reported that they were approximately 100 papers with 136 attendees and a profit of \$21,299.54.

2013 (Montreal, ICA 2013): Mike Stinson reported *that they are* approximately 2500 participants (600 students) at the meeting. This is one of the largest meeting in Acoustics ever held.

2014 (Winnipeg):

Christian Giguère mentioned that Karen Turner is in charge of the local organization. She is collected information about hotel venues. She will also need to finalize her committee. A few names were suggested, Christian Giguère will do the follow up. He also mentioned that the online system includes a separate conference system that could be used for registration and handling of abstracts.

2015: Halifax has been suggested. To be discussed with local people.

### **Acoustical Standards Committee**

Christian Giguère gave some background. The CAA Standards Committee is proposing adoption of the CAA Guide 101 to Acoustic Standards. The BoD wish to formalize the terms of reference for the Committee and clear liability issues before going ahead with the guide's approval.

### **Website**

Sean Pecknold presented the new website. Christian Giguère thanked Sean and his team for the good work accomplished.

### **Nominations and Election**

CAA corporate bylaws require that each year we elect the Executive and Directors. The Past President, Stan Dosso, presented nominations and managed the election process.

For the election process, Stan read the name of the nominee(s) and asked if there were other nominees from the floor.

- Jeremie Voix served as a Director for the maximum term of 6 years. Alberto Behar is nominated to replace him as new Director. The other Directors (namely Roberto Racca, Kathy Pichora-Fuller, Sean Pecknold, Clair Wakefield, Karen Turner, Bryan Gick and Hugues Nelisse) indicated their willingness to stand.
- Chantal Laroche and Dalila Giusti indicated their willingness to stay as Executive Secretary and Treasurer, respectively.
- Christian Giguère is stepping out as President and becomes Past President. Stan Dosso is stepping out after 10 years of service on the Executive. Frank Russo was nominated as the new President. Jérémie Voix was nominated as the new Editor-in-Chief.

There were no other nominations from the floor, so these nominees were declared elected by acclamation. Stan Dosso and Christian Giguère were thanked for their service. The new Board of Directors will be in effect on October 15<sup>th</sup> 2013.

### **Adjournment**

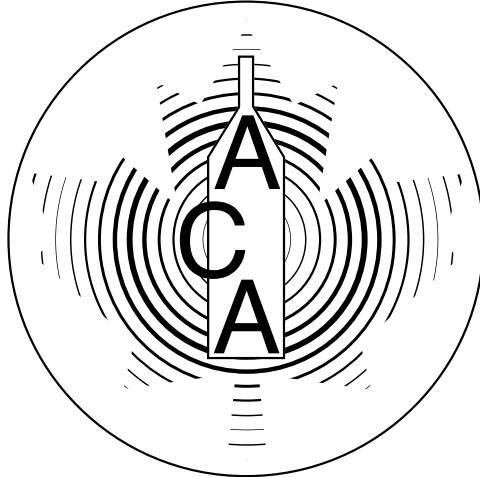
Meeting adjourned at 8 p.m. (*Moved by J. Voix, seconded by P. Van Delden, carried*)



Canadian Acoustical Association  
Association Canadienne d'Acoustique

---

**PRIZE ANNOUNCEMENT • ANNONCE DE PRIX**



**Prize**

EDGAR AND MILLICENT SHAW POSTDOCTORAL PRIZE IN ACOUSTICS  
ALEXANDER G. BELL GRADUATE STUDENT PRIZE IN SPEECH COMMUNICATION AND HEARING  
ECKEL GRADUATE STUDENT PRIZE IN NOISE CONTROL  
FESSENDEN GRADUATE STUDENT PRIZE IN UNDERWATER ACOUSTICS  
RAYMOND HETU UNDERGRADUATE STUDENT PRIZE IN ACOUSTICS  
GRADUATE STUDENT PRIZE IN ARCHITECTURAL AND ROOM ACOUSTICS  
BREGMAN GRADUATE STUDENT PRIZE IN PSYCHOLOGICAL ACOUSTICS

**Prix**

PRIX POST-DOCTORAL EDGAR ET MILLICENT SHAW EN ACOUSTIQUE  
PRIX ETUDIANT ALEXANDER G. BELL EN COMMUNICATION ORALE ET AUDITION (2<sup>E</sup> OU 3<sup>E</sup> CYCLE)  
PRIX ETUDIANT ECKEL EN CONTROLE DU BRUIT (2<sup>E</sup> OU 3<sup>E</sup> CYCLE)  
PRIX ETUDIANT FESSENDEN EN ACOUSTIQUE SOUS-MARINE (2<sup>E</sup> OU 3<sup>E</sup> CYCLE)  
PRIX ETUDIANT RAYMOND HETU EN ACOUSTIQUE (1<sup>ER</sup> CYCLE)  
PRIX ETUDIANT EN ACOUSTIQUE ARCHITECTURALE ET ACOUSTIQUE DES SALLES (2<sup>E</sup> OU 3<sup>E</sup> CYCLE)  
PRIX ETUDIANT BREGMAN EN PSYCHOACOUSTIQUE (2<sup>E</sup> OU 3<sup>E</sup> CYCLE)

**Deadline for Applications:  
April 30<sup>th</sup> 2014**

**Date limite de soumission des demandes:  
30 Avril 2014**

Consult CAA website for more information  
Consultez le site Internet de l'ACA pour de plus amples renseignements  
(<http://www.caa-aca.ca>)



### Application for Membership

CAA membership is open to all individuals who have an interest in acoustics. Annual dues total \$90.00 for individual members and \$40.00 for student members. This includes a subscription to *Canadian Acoustics*, the journal of the Association, which is published 4 times/year, and voting privileges at the Annual General Meeting.

### Subscriptions to *Canadian Acoustics* or Sustaining Subscriptions

Subscriptions to *Canadian Acoustics* are available to companies and institutions at a cost of \$90.00 per year. Many organizations choose to become benefactors of the CAA by contributing as Sustaining Subscribers, paying \$400.00 per year (no voting privileges at AGM). The list of Sustaining Subscribers is published in each issue of *Canadian Acoustics* and on the CAA website.

*Please note that online payments will be accepted at <http://jcaa.caa-aca.ca>*

#### Address for subscription / membership correspondence:

Name / Organization \_\_\_\_\_  
Address \_\_\_\_\_  
City/Province \_\_\_\_\_ Postal Code \_\_\_\_\_ Country \_\_\_\_\_  
Phone \_\_\_\_\_ Fax \_\_\_\_\_ E-mail \_\_\_\_\_

#### Address for mailing *Canadian Acoustics*, if different from above:

Name / Organization \_\_\_\_\_  
Address \_\_\_\_\_  
City/Province \_\_\_\_\_ Postal Code \_\_\_\_\_ Country \_\_\_\_\_

#### Areas of Interest: (Please mark 3 maximum)

- |  |   |   |
|--|---|---|
| 1. Architectural Acoustics               | 5. Psychological / Physiological Acoustic | 9. Underwater Acoustics                   |
| 2. Engineering Acoustics / Noise Control | 6. Shock and Vibration                    | 10. Signal Processing / Numerical Methods |
| 3. Physical Acoustics / Ultrasound       | 7. Hearing Sciences                       | 11. Other                                 |
| 4. Musical Acoustics / Electro-acoustics | 8. Speech Sciences                        |   |

For student membership, please also provide:

\_\_\_\_\_  
(University)

\_\_\_\_\_  
(Faculty Member)

\_\_\_\_\_  
(Signature of Faculty Member)

\_\_\_\_\_  
(Date)

I have enclosed the indicated payment for:

- CAA Membership \$ 90.00  
 CAA Student Membership \$ 40.00

Corporate Subscriptions (4 issues/yr)

- \$90 including mailing in Canada  
 \$98 including mailing to USA,  
 \$105 including International mailing

- Sustaining Subscription \$400.00  
(4 issues/yr)

**Please note that the preferred method of payment is by credit card, online at <http://jcaa.caa-aca.ca>**

**For individuals or organisations wishing to pay by cheque, please download the application form at <http://www.caa-aca.ca> and mail to:**

**Executive Secretary, The Canadian Acoustical Association, PO Box 74068, Ottawa, Ontario, K1M 2H9, Canada**



### Formulaire d'adhésion

L'adhésion à l'ACA est ouverte à tous ceux qui s'intéressent à l'acoustique. La cotisation annuelle est de 90.00\$ pour les membres individuels, et de 40.00\$ pour les membres étudiants. Tous les membres reçoivent *l'Acoustique Canadienne*, la revue de l'association publiée quatre fois par année, et ont droit de vote à l'assemblée générale annuelle.

### Abonnement à la revue *Acoustique Canadienne* et abonnement de soutien

Les abonnements à la revue *Acoustique Canadienne* sont disponibles pour les corporations et institutions au coût annuel de 90.00\$. Plusieurs organisations choisissent de devenir bienfaiteurs de l'ACA en souscrivant à un abonnement de soutien de 400.00\$ par année (sans droit de vote à l'AGA). La liste des abonnés de soutien est publiée dans chaque numéro de la revue *Acoustique Canadienne* et sur le site internet de l'ACA.

*Noter que les paiements en ligne sont effectués à <http://jcaa.caa-aca.ca>*

#### *Pour correspondance administrative et financière:*

Nom / Organisation \_\_\_\_\_  
Adresse \_\_\_\_\_  
Ville/Province \_\_\_\_\_ Code postal \_\_\_\_\_ Pays \_\_\_\_\_  
Téléphone \_\_\_\_\_ Téléc. \_\_\_\_\_ Courriel \_\_\_\_\_

#### *Adresse postale pour la revue Acoustique canadienne*

Nom / Organisation \_\_\_\_\_  
Adresse \_\_\_\_\_  
Ville/Province \_\_\_\_\_ Code postal \_\_\_\_\_ Pays \_\_\_\_\_

#### **Cocher vos champs d'intérêt:** (maximum 3)

- |   |                               |  |
|---|-------------------------------|--|
| 1. Acoustique architecturale                | 5. Physio / Psycho-acoustique | 9. Acoustique sous-marine                        |
| 2. Génie acoustique / Contrôle du bruit     | 6. Chocs et vibrations        | 10. Traitement des signaux / Méthodes numériques |
| 3. Acoustique physique / Ultrasons          | 7. Audition                   | 11. Autre  |
| 4. Acoustique musicale / Électro-acoustique | 8. Parole                     |  |

Prière de remplir pour les étudiants et étudiantes:

\_\_\_\_\_ (Université) (Nom d'un membre du corps professoral) \_\_\_\_\_ (Signature du membre du corps professoral) \_\_\_\_\_ (Date)

Cocher la case appropriée:

- Membre individuel 90.00 \$  
 Membre étudiant(e) 40.00 \$

Abonnement corporatif (4 numéros/année)

- 90 \$ à l'intérieur du Canada  
 98 \$ vers les États-Unis  
 105 \$ tout autre envoi international

- Abonnement de soutien 400.00 \$  
(4 numéros/année)

**Noter que la méthode de paiement privilégiée est par carte de crédit, en ligne à <http://jcaa.caa-aca.ca>**

**Les individus et organisations préférant payer par chèque doivent télécharger le formulaire d'adhésion disponible à <http://www.caa-aca.ca> et le poster à l'adresse suivante:**

**Secrétaire exécutif, Association canadienne d'acoustique,  
CP 74068, Ottawa, Ontario, K1M 2H9, Canada**

## INSTRUCTIONS TO AUTHORS FOR THE PREPARATION OF MANUSCRIPTS

**Submissions:** The original manuscript and two copies should be sent to the Editor-in-Chief. The manuscript can also be submitted electronically.

**General Presentation:** Papers should be submitted in camera-ready format. Paper size 8.5" x 11". If you have access to a word processor, copy as closely as possible the format of the articles in Canadian Acoustics 39(1) 2011. All text in Times-Roman 10 pt font, with single (12 pt) spacing. Main body of text in two columns separated by 0.25". One line space between paragraphs.

**Margins:** Top - 0.75"; bottom - 0.75" minimum; sides - 0.75".

**Title:** Bold, Times New Roman 14 pt with 14 pt spacing, upper case, centered.

**Authors/addresses:** Names and full mailing addresses, 10 pt with single (12 pt) spacing, upper and lower case, centered. Names in bold text.

**Abstracts:** English and French versions. Headings, 12 pt bold, upper case, centered. Indent text 0.5" on both sides.

**Headings:** Headings to be in 12 pt bold, Times-Roman font. Number at the left margin and indent text 0.5". Main headings, numbered as 1, 2, 3, ... to be in upper case. Sub-headings numbered as 1.1, 1.2, 1.3, ... in upper and lower case. Sub-sub-headings not numbered, in upper and lower case, underlined.

**Equations:** Minimize. Place in text if short. Numbered.

**Figures/Tables:** Keep small. Insert in text at top or bottom of page. Name as "Figure 1, 2, ..." Caption in 9 pt with single (12 pt) spacing. Leave 0.5" between text.

**Line Widths:** Line widths in technical drawings, figures and tables should be a minimum of 0.5 pt.

**Photographs:** Submit original glossy, black and white photograph.

**Scans:** Should be between 225 dpi and 300 dpi. Scan: Line art as bitmap tiffs; Black and white as grayscale tiffs and colour as CMYK tiffs;

**References:** Cite in text and list at end in any consistent format, 9 pt with single (12 pt) spacing.

**Page numbers:** In light pencil at the bottom of each page. For electronic submissions, do not number pages.

**Reprints:** Can be ordered at time of acceptance of paper.

## DIRECTIVES A L'INTENTION DES AUTEURS PREPARATION DES MANUSCRITS

**Soumissions:** Le manuscrit original ainsi que deux copies doivent être soumis au rédacteur-en-chef. Le manuscrit peut être aussi acheminé par voie électronique.

**Présentation générale:** Le manuscrit doit être soumis avec mise en page en format de publication. Dimension des pages, 8,5" x 11". Si vous avez accès à un système de traitement de texte, dans la mesure du possible, suivre le format des articles dans l'Acoustique canadienne 39(1) 2011. Tout le texte doit être en caractères Times-Roman, 10 pt et à simple (12 pt) interligne. Le texte principal doit être en deux colonnes séparées d'un espace de 0.25". Les paragraphes sont séparés d'un espace d'une ligne.

**Marges:** Haut - 0.75"; bas - minimum 0.75"; côtés, - 0.75".

**Titre du manuscrit:** Caractères gras, Times New Roman 14 pt, avec espace interligne de 14 pt, lettres majuscules, texte centré.

**Auteurs/adresses:** Noms et adresses postales. Lettres majuscules et minuscules, 10 pt à simple (12 pt) interligne, texte centré. Les noms doivent être en caractères gras.

**Sommaire:** En versions anglaise et française. Titre en 12 pt, lettres majuscules, caractères gras, texte centré. Paragraphe 0.5" en alinéa de la marge, des 2 cotés.

**Titres des sections:** Tous en caractères gras, 12 pt, Times-Roman. Premiers titres: numéroter 1, 2, 3, ..., en lettres majuscules; sous-titres: numéroter 1.1, 1.2, 1.3, ..., en lettres majuscules et minuscules; sous-sous-titres: ne pas numéroter, en lettres majuscules et minuscules et soulignés.

**Équations:** Minimiser le nombre et les numéroter. Insérer directement dans le texte les équations très courtes.

**Figures/Tableaux:** De petites tailles. Les insérer dans le texte au haut ou au bas de la page. Les nommer "Figure 1, 2, 3, ..." Légende en 9 pt à simple (12 pt) interligne. Laisser un espace de 0.5" entre le texte.

**Largeur des traits:** La largeur des traits sur les schémas techniques doivent être au minimum de 0.5 pt pour permettre une bonne reproduction.

**Photographies:** Soumettre la photographie originale sur papier glacé, noir et blanc.

**Figures numérisées:** Doivent être au minimum de 225 dpi et au maximum de 300 dpi. Les schémas doivent être en format bitmap tif. Les photos noir et blanc doivent en format tif sur une échelle de tons de gris et toutes les photos couleurs doivent être en format CMYK tif.

**Références:** Les citer dans le texte et en faire la liste à la fin du document, en format uniforme, 9 pt à simple (12 pt) interligne.

**Pagination:** Au crayon pâle, au bas de chaque page. Ne pas paginer si le manuscrit est envoyé par voie électronique.

**Tirés-à-part:** Ils peuvent être commandés au moment de l'acceptation du manuscrit.

# The Canadian Acoustical Association l'Association Canadienne d'Acoustique



## PRESIDENT PRÉSIDENT

**Christian Giguère**  
Université d'Ottawa  
Ottawa, Ontario  
K1H 8M5  
(613) 562-5800 x4649  
cgiguere@uottawa.ca

## PAST PRESIDENT PRÉSIDENT SORTANT

**Stan Dosso**  
University of Victoria  
Victoria, British Columbia  
V8W 3P6  
(250) 472-4341  
sdosso@uvic.ca

## EXECUTIVE SECRETARY SECRÉTAIRE EXÉCUTIF

**Chantal Laroche**  
P. O. Box 74068  
Ottawa, Ontario  
K1M 2H9  
(613) 562-5800 # 3066  
claroche@uottawa.ca

---

## TREASURER TRÉSORIER

**Dalila Giusti**  
Jade Acoustics  
411 Confederation Parkway, Unit 19  
Concord, Ontario  
L4K 0A8  
(905) 660-2444  
dalila@jadeacoustics.com

## EDITOR-IN-CHIEF RÉDACTEUR EN CHEF

**Frank A. Russo**  
Dept. of Psychology  
Ryerson University  
350 Victoria Street  
Toronto, Ontario  
M5B 2K3  
(416) 979-5000, x. 2647  
russo@ryerson.ca

## WEBMASTERS

**Sean Pecknold (Association)**  
sean.pecknold@drdc-rddc.gc.ca  
<http://www.caa-aca.ca>

**Jérémie Voix (Journal)**  
voix@caa-aca.ca  
<http://jcaa.caa-aca.ca>

---

## DIRECTORS DIRECTEURS

**Bill Gastmeier**  
HGC Engineering  
(905) 826-4940  
bgastmeier@hgcengineering.com

**Sean Pecknold**  
DRDC Atlantic  
(902) 426-3100  
sean.pecknold@drdc-rddc.gc.ca

**Karen Turner**  
Protec Hearing  
(204) 771-9330  
protec@escape.ca

**Bryan Gick**  
University of British Columbia  
(604) 822-4817  
gick@mail.ubc.ca

**Kathy Pichora-Fuller**  
University of Toronto at Mississauga  
(905) 828-3865  
k.pichora.fuller@utornot.ca

**Jérémie Voix**  
École de technologie supérieure,  
Université de Québec  
(514) 396-8437  
voix@caa-aca.ca

**Hugues Nélisse**  
IRSST  
(514) 288-1551 x221  
Hugues.nelisse@irsst.qc.ca

**Roberto Racca**  
JASCO  
(250) 483-3300  
rob@jasco.com

# SUSTAINING SUBSCRIBERS / ABONNES DE SOUTIEN

The Canadian Acoustical Association gratefully acknowledges the financial assistance of the Sustaining Subscribers listed below. Their annual donations (of \$350.00 or more) enable the journal to be distributed to all at a reasonable cost.

L'Association Canadienne d'Acoustique tient à témoigner sa reconnaissance à l'égard de ses Abonnés de Soutien en publiant ci-dessous leur nom et leur adresse. En amortissant les coûts de publication et de distribution, les dons annuels (de \$350.00 et plus) rendent le journal accessible à tous nos membres.

## **ACI Acoustical Consultants Inc.**

Mr. Steven Bilawchuk - (780) 414-6373  
stevenb@aciacoustical.com - Edmonton, AB

## **Acoustec Inc.**

Dr. J.G. Migneron - (418) 834-1414  
courrier@acoustec.qc.ca - St-Nicolas, QC

## **ARMTEC**

Ron Galloway - 905-521-0999  
armtec.com

## **Dalimar Instruments Inc.**

Mr. Daniel Larose - (514) 424-0033  
daniel@dalimar.ca - Vaudreuil-Dorion, QC

## **Engineers for Change, Inc.**

Richard Joseph Peppin - 301-984-3375

## **Hatch Associates Ltd.**

Mr. Tim Kelsall - (905) 403-3932  
tkelsall@hatch.ca - Mississauga, ON

## **Jade Acoustics Inc.**

Ms. Dalila Giusti - (905) 660-2444  
dalila@jadeacoustics.com - Concord, ON

## **MJM Conseillers en Acoustique Inc.**

M. Michel Morin - (514) 737-9811  
mmorin@mjmqc.ca - Montréal, QC

## **Pinchin Environmental Ltd.**

(905) 363-0678; FAX: (905) 363-0681  
info@pinchin.com - Mississauga, ON

## **SNC/Lavalin Environment Inc.**

M. Jean-Luc Allard - (514) 651-6710  
jeanluc.allard@snclavalin.com - Longueuil, QC

## **Soundtrap Inc.**

Roger Foulds - 705-357-1067  
soundtrap.ca

## **True Grit Consulting Ltd.**

Ms Ina Chomyshyn - 807-626-5640  
tgcl.ca

## **Vibro-Acoustics**

Mr. Tim Charlton - (800) 565-8401  
tcharlton@vibro-acoustics.com

## **Xscala Sound & Vibration**

Jim Ulicki - 403-274-7577  
xscala.com

## **ACO Pacific Inc.**

Mr. Noland Lewis - (650) 595-8588  
acopac@acopacific.com - Belmont, CA

## **Aercoustics Engineering Ltd**

Mr. John O'Keefe - (416) 249-3361  
aercoustics@aercoustics.com - Rexdale, ON

## **Bruel & Kjaer North America Inc.**

Mr. Andrew Khoury - (514) 695-8225  
andrew.khoury@bksv.com - Pointe-Claire, QC

## **DuraSystems Barriers Inc.**

Fred Woo - 905-660-4455  
durasystems.com

## **FFA Consultants in Acoustics and Noise Control**

Robyn Sharek

## **HGC Engineering Ltd.**

Mr. Bill Gastmeier - (905) 826-4044  
info@hgcengineering.com - Mississauga, ON

## **Kinetics Noise Control Inc.**

Mehrzad Salkhordeh - 905-670-4922  
kineticsnoise.com

## **Novel Dynamics Test Inc.**

Mr. Andy Metelka - (519) 853-4495  
metelka@aztec-net.com - Acton, ON

## **Pyrok Inc.**

(914) 777-7770; FAX: (914) 777-7103  
info@pyrokinc.com - Mamaroneck, NY

## **Soft dB Inc.**

M. André L'Espérance - (418) 686-0993  
contact@softdb.com - Sillery, QC

## **Stantec Consulting Ltd**

Zohreh Razavi - 604-606-8472  
stantec.com

## **Valcoustics Canada Ltd.**

Dr. Al Lightstone - (905) 764-5223  
solutions@valcoustics.com - Richmond Hill, ON

## **Wakefield Acoustics Ltd.**

Mr. Clair Wakefield - (250) 370-9302  
nonoise@shaw.ca - Victoria, BC

## **Acoustic Solutions LTD**

Isidor Gliener  
acousticsolutions.com

## **AMEC**

Mr. Frank Babic - 905.568.2929  
amec.com

## **Conestoga-Rovers & Associates**

Tim C. Wiens - 519-884-0510 xx2352  
craworld.com

## **Eckel Industries of Canada Ltd.**

Mr. Bruce Allan - (613) 543-2967  
eckel@eckel.ca - Morrisburg, ON

## **H.L. Blachford Ltd.**

Mr. Dalton Prince - (905) 823-3200  
amsales@blachford.ca - Mississauga, ON

## **Integral DX Engineering Ltd.**

Mr. Greg Clunis - (613) 761-1565  
greg@integraldxengineering.ca - Ottawa, ON

## **Mc SQUARED System Design Group**

Mr. Wade McGregor - (604) 986-8181  
info@mcsquared.com - North Vancouver, BC

## **Peutz & Associés**

M. Marc Asselineau +33 1 45230500  
m.asselineau@peutz.fr - Paris, FRANCE

## **Scantek Inc.**

Mr. Richard J. Peppin, (410)-290-7726  
peppinr@scantekinc.com - Columbia, MD

## **Sound & Vibration Solutions Canada, Inc.**

Mr. Andy Metelka - 519-853-4495

## **State of the Art Acoustik Inc.**

Dr. C. Fortier - (613) 745-2003,  
sota@sota.ca - Ottawa, ON

## **Vendatech Inc.**

Behrou Ghazizadeh - 416-787-8797  
vendatech.com

## **West Caldwell Calibration Labs**

Mr. Stanley Christopher - (905) 595-1107  
info@wcccl.com - Brampton, ON