

CANADIAN JOURNAL OF SPEECH-LANGUAGE PATHOLOGY AND AUDIOLOGY

CJSLPA • RCOA

REVUE CANADIENNE D'ORTHOPHONIE ET D'AUDIOLOGIE

Fall ► Autumn, 2011

Volume 35, No. 3



FROM THE EDITOR

Tim Bressmann

Speech audiometry with non-native English speakers:

The use of digits and Cantonese words as stimuli

Stefka H. Marinova-Todd, Carrie K. Siu, Lorient M. Jenstad

Exploration of the utility of a brief swallow screening protocol with comparison to concurrent videofluoroscopy

Catrina M. Steele, Sonja M. Molfenter, Gemma L. Bailey, Rebecca Cliffe Polacco, Ashley A. Waito, Dana C. B. H. Zoratto, Tom Chau

Développement de la version québécoise francophone du Children's Communication Checklist – 2 (CCC-2).

Traduction, adaptation et équivalence conceptuelle

Marie Vézina, Catherine Samson-Morasse, Julie Gauthier-Desgagné, Marion Fossard, Audette Sylvestre

Performance of French-speaking Quebec Adults on the Boston Naming Test

Patricia M. Roberts, Nathalie Doucet

Book Reviews:

Science of Successful Supervision and Mentorship

Ian Roth

Interventions for Speech Sound Disorders in Children

Bojana Radovanovic

Purpose and Scope

The Canadian Association of Speech-Language Pathologists and Audiologists (CASLPA) is the only national body that supports and represents the professional needs of speech-language pathologists, audiologists and supportive personnel inclusively within one organization. Through this support, CASLPA champions the needs of people with communications disorders. The association was founded in 1964 and incorporated under federal charter in 1975. CASLPA's periodical publications program began in 1973.

The purpose of the *Canadian Journal of Speech-Language Pathology and Audiology* (CJSLPA) is to disseminate contemporary knowledge pertaining to normal human communication and related disorders of communication that influence speech, language, and hearing processes. The scope of the Journal is broadly defined so as to provide the most inclusive venue for work in human communication and its disorders. CJSLPA publishes both applied and basic research, reports of clinical and laboratory inquiry, as well as educational articles related to normal and disordered speech, language, and hearing in all age groups. Classes of manuscripts suitable for publication consideration in CJSLPA include tutorials; traditional research or review articles; clinical, field, and brief reports; research notes; and letters to the editor (see Information to Contributors). CJSLPA seeks to publish articles that reflect the broad range of interests in speech-language pathology and audiology, speech sciences, hearing science, and that of related professions. The Journal also publishes book reviews, as well as independent reviews of commercially available clinical materials and resources.

The *Canadian Journal of Speech-Language Pathology and Audiology* is supported by a grant in Aid to Scholarly Journals, provided by the Canadian Social Sciences and Humanities Research Council (grant # 651-2008-0062), for the period January 2009 to December 2011.

CASLPA Vision and Mission

Vision

The Canadian Association of Speech-Language Pathologists and Audiologists ...the national voice and recognized resource for speech-language pathology and audiology.

Mission

The Canadian Association of Speech-Language Pathologists and Audiologists ...supporting and empowering our members to maximize the communication and hearing potential of the people of Canada.

Indexing

CJSLPA is indexed by:

- CINAHL – Cumulative Index to Nursing and Allied Health Literature
- Elsevier Bibliographic Databases (SCOPUS)
- ERIC Clearinghouse on Disabilities and Gifted Education
- ProQuest – CSA Linguistics and Language Behavior Abstracts (LLBA)
- PsycInfo
- Thomson Gale (Academic Onefile)
- EBSCO Publishing Inc. (CINHAL Plus with full text)
- Directory of Open Access Journals

Online Archive

CJSLPA is now an open-access publication. For full-text articles and archives, visit: www.cjslpa.ca

Advertising

All inquiries concerning the placement of advertisements in CJSLPA should be directed to pubs@caslpa.ca. Acceptance of an ad does not in any way constitute CASLPA's endorsement of the product/service or company. CASLPA reserves the right to reject any ad if the advertisement, organization, product, or service is not compatible with CASLPA's mission or vision. CASLPA does not accept responsibility for the accuracy of statements by advertisers.

2010 CJSLPA Reviewers

Michael Kieft, Christiane Séguin, Natasha Trudeau, Elena Nicoladis, Joanne Deluzio, Robert Harrison, Kai Uus, André Lafargue, Frank Russo, Andrea Macleod, Audette Sylvestre, Elizabeth Crais, Claude Vincent, Francois Bergeron, Anita McGinty, Lori Skibbe, Michel Habib, Jean-Pierre Gagné, Marianne Pouplier, Sharynne McLeod, Carol To, Lydia So, (Barbara) May Bernhardt, Caroline Bowen, Nicole Mueller, Glenda Mason, Peter Flipsen Jr., Melanie Campbell, Denise Hayward, Lori Swanson, Laurent Motron, Carole Sénéchal, Margaret Cheesman, Cory Portnuff, Phaedra Royle, Renée Béland, Glen Nowell, Francois Bergeron, Lendra Friesen

Copyright

© 2011, CASLPA

Copyright of the Canadian Journal of Speech-Language Pathology and Audiology is held by the Canadian Association of Speech-Language Pathologists and Audiologists (CASLPA). Appropriate credit must be given (CASLPA, publication name, article title, volume number, issue number and page number[s]) but not in any way that suggests CASLPA endorses you or your use of the work. You may not use this work for commercial purposes. You may not alter, transform, or build upon this work. Contact pubs@caslpa.ca

Vol. 35, No. 3
Fall 2011

Editor

Tim Bressmann, Ph.D.
University of Toronto

Managing Editor/Layout

Olga Novoa

Director of Communications

Angie D'Aoust

Associate Editors

Andrea MacLeod, Ph.D.
Université Laval
(Language, English submissions)

Vincent Gracco, Ph.D.
McGill University
(Speech, English submissions)

Navid Shahnaz
University of British Columbia
(Audiology, English submissions)

Joël Macoir, Ph.D.
Université Laval
(Speech & Language,
French submissions)

Benoît Jutras, Ph.D.
Université de Montréal
(Audiology, French submissions)

Assistant Editors

Candace Myers, M.Sc.
CancerCare Manitoba
(Material & Resource Reviews)

Glen Nowell, M.Sc.
Hamilton Health Sciences
(Book Reviews)

Review of translation

Benoît Jutras, Ph.D.
Université de Montréal

Translation

Genevieve Charbonneau

ISSN 1913-200X

CJSLPA is published quarterly by the Canadian Association of Speech-Language Pathologists and Audiologists (CASLPA). Publications Agreement Number: # 40036109. Return undeliverable Canadian addresses to: CASLPA, 1 Nicholas Street, Suite 1000, Ottawa, Ontario K1N 7B7. Address changes should be sent to CASLPA by e-mail to pubs@caslpa.ca or to the above-mentioned address.

Objet et Portée

L'Association canadienne des orthophonistes et audiologistes (ACOA) est l'association professionnelle nationale reconnue des orthophonistes et des audiologistes du Canada. L'Association a été fondée en 1964 et incorporée en vertu de la charte fédérale en 1975. L'Association s'engage à favoriser la meilleure qualité de services aux personnes atteintes de troubles de la communication et à leurs familles. Dans ce but, l'Association entend, entre autres, contribuer au corpus de connaissances dans le domaine des communications humaines et des troubles qui s'y rapportent. L'Association a mis sur pied son programme de publications en 1973.

L'objet de la *Revue canadienne d'orthophonie et d'audiologie* (RCOA) est de diffuser des connaissances relatives à la communication humaine et aux troubles de la communication qui influencent la parole, le langage et l'audition. La portée de la Revue est plutôt générale de manière à offrir un véhicule des plus compréhensifs pour la recherche effectuée sur la communication humaine et les troubles qui s'y rapportent. La RCOA publie à la fois les ouvrages de recherche appliquée et fondamentale, les comptes rendus de recherche clinique et en laboratoire, ainsi que des articles éducatifs portant sur la parole, le langage et l'audition normaux ou désordonnés pour tous les groupes d'âge. Les catégories de manuscrits susceptibles d'être publiés dans la RCOA comprennent les tutoriels, les articles de recherche conventionnelle ou de synthèse, les comptes rendus cliniques, pratiques et sommaires, les notes de recherche, et les courriers des lecteurs (voir Renseignements à l'intention des collaborateurs). La RCOA cherche à publier des articles qui reflètent une vaste gamme d'intérêts en orthophonie et en audiologie, en sciences de la parole, en science de l'audition et en diverses professions connexes. La Revue publie également des critiques de livres ainsi que des critiques indépendantes de matériel et de ressources cliniques offerts commercialement.

La Revue canadienne d'orthophonie et d'audiologie est appuyée par une subvention d'Aide aux revues savantes accordée par le Conseil de recherches en sciences humaines du Canada (subvention no. 651-2008-0062), pour la période de janvier 2009 à décembre 2011.

ACOA : VISION ET MISSION

Vision

L'Association canadienne des orthophonistes et audiologistes : porte-parole nationale et ressource reconnue dans le domaine de l'orthophonie et de l'audiologie.

Mission

L'Association canadienne des orthophonistes et audiologistes appuie et habilite ses membres en vue de maximiser le potentiel en communication et en audition de la population canadienne.

Inscription au Répertoire

RCOA est répertoriée dans:

- CINAHL – Cumulative Index to Nursing and Allied Health Literature
- Elsevier Bibliographic Databases (SCOPUS)
- ERIC Clearinghouse on Disabilities and Gifted Education
- ProQuest – CSA Linguistics and Language Behavior Abstracts (LLBA)
- PsycInfo
- Thomson Gale (Academic Onefile)
- EBSCO Publishing Inc. (CINHAL Plus with full text)
- Directory of Open Access Journals

Archive en-ligne

Les articles et les archives de la RCOA sont maintenant disponibles au public à www.cjslpa.ca

Publicité

Toutes les demandes visant à faire paraître de la publicité dans la RCOA doivent être adressées au Bureau national. L'acceptation d'une annonce publicitaire ne signifie absolument pas que l'ACOA fait la promotion du produit, du service ou de la compagnie. L'ACOA se réserve le droit de rejeter une annonce si le message, l'organisation, le produit ou le service n'est pas compatible avec la mission, la vision ou les valeurs de l'ACOA. L'ACOA n'assume pas la responsabilité de l'exactitude des déclarations des annonceurs.

Réviseurs de la RCOA 2010

Michael Kieffe, Christiane Séguin, Natasha Trudeau, Elena Nicoladis, Joanne Deluzio, Robert Harrison, Kai Uus, André Lafargue, Frank Russo, Andrea Macleod, Audette Sylvestre, Elizabeth Crais, Claude Vincent, Francois Bergeron, Anita McGinty, Lori Skibbe, Michel Habib, Jean-Pierre Gagné, Marianne Pouplier, Sharynne McLeod, Carol To, Lydia So, (Barbara) May Bernhardt, Caroline Bowen, Nicole Mueller, Glenda Mason, Peter Flipsen Jr., Melanie Campbell, Denise Hayward, Lori Swanson, Laurent Motron, Carole Sénéchal, Margaret Cheesman, Cory Portnuff, Phaedra Royle, Renée Béland, Glen Nowell, Francois Bergeron, Lendra Friesen

Droit d'auteur

© 2011, ACOA

C'est l'Association canadienne des orthophonistes et audiologistes (ACOA) qui détient le droit d'auteur de la Revue canadienne d'orthophonie et d'audiologie. Il faut mentionner la source (ACOA, nom de la publication, titre de l'article, numéro du volume, numéro de parution et nombre de pages), mais sans laisser entendre que l'ACOA vous approuve ou approuve l'utilisation que vous faites du texte. Il est interdit d'utiliser le document à des fins commerciales. Il est interdit de modifier, transformer ou développer le texte. Contactez pubs@caslpa.ca

Vol. 35, N° 3
Automne 2011

REVUE CANADIENNE D'ORTHOPHONIE ET D'AUDIOLOGIE

Rédacteur en chef

Tim Bressmann, Ph.D.
University of Toronto

**Directrice de la rédaction /
mise en page**
Olga Novoa

Directrice des communications
Angie D'Aoust

Rédacteurs en chef adjoints

Andrea MacLeod, Ph.D.
Université Laval
(Langage, soumissions en anglais)

Vincent Gracco, Ph.D.
McGill University
(Parole, soumissions en anglais)

Navid Shahnaz
University of British Columbia
(Audiologie, soumissions en anglais)

Joël Macoir, Ph.D.
Université Laval
(Parole et langage, soumissions
en français)

Benoît Jutras, Ph.D.
Université de Montréal
(Audiologie, soumissions en français)

Rédacteurs adjoints

Candace Myers, M.Sc.
CancerCare Manitoba
(Évaluation des ressources)

Glen Nowell, M.Sc.
Hamilton Health Sciences
(Évaluation des ouvrages écrits)

Révision de la traduction

Benoît Jutras, Ph.D.
Université de Montréal

Traduction

Genevieve Charbonneau

ISSN 1913-200X

La RCOA est publiée quatre fois l'an par l'Association canadienne des orthophonistes et audiologistes (ACOA). Numéro de publication : #40036109. Faire parvenir tous les envois avec adresses canadiennes non reçus au 1, rue Nicholas, bureau 1000, Ottawa (Ontario) K1N 7B7. Faire parvenir tout changement à l'ACOA au courriel pubs@caslpa.ca ou à l'adresse indiquée ci-dessus.

TABLE OF CONTENTS

From the Editor218

Tim Bressmann

Article 1.....220

**Speech audiometry with non-native English speakers:
The use of digits and Cantonese words as stimuli**

Associate editor Elizabeth Fitzpatrick

Stefka H. Marinova-Todd, Carrie K. Siu, Lorianne M. Jenstad

Article 2.....228

**Exploration of the utility of a brief swallow screening
protocol with comparison to concurrent videofluoroscopy**

Catriona M. Steele, Sonja M. Molfenter, Gemma L. Bailey, Rebecca
Cliffe Polacco, Ashley A. Waito, Dana C. B. H. Zoratto, Tom Chau

Article 3244

**Développement de la version québécoise francophone
du Children's Communication Checklist – 2 (CCC-2).
Traduction, adaptation et équivalence conceptuelle**

Guest editor Phaedra Royle

Marie Vézina, Catherine Samson-Morasse, Julie Gauthier-Desgagné,
Marion Fossard, Audette Sylvestre

Article 4.....254

**Performance of French- speaking Quebec Adults
on the Boston Naming Test**

Associate editor Jeff Small

Patricia M. Roberts, Nathalie Doucet

Book Reviews268

Science of Successful Supervision and Mentorship

Ian Roth

Interventions for Speech Sound Disorders in Children

Bojana Radovanovic

TABLE DES MATIÈRES

Mot de rédacteur en chef218

Tim Bressmann

Article 1.....220

**Audiométrie vocale chez des personnes dont la langue
maternelle n'est pas l'anglais : l'utilisation de chiffres et
de mots en cantonais comme stimuli**

Rédactrice en chef adjointe Elizabeth Fitzpatrick

Stefka H. Marinova-Todd, Carrie K. Siu, Lorianne M. Jenstad

Article 2228

**Exploration de l'utilité d'un bref protocole de dépistage
des troubles de déglutition en comparaison avec une
vidéofluoroscopie simultanée**

Catriona M. Steele, Sonja M. Molfenter, Gemma L. Bailey, Rebecca
Cliffe Polacco, Ashley A. Waito, Dana C. B. H. Zoratto, Tom Chau

Article 3244

**Development of a Quebec French version of the
Children's Communication Checklist – 2 (CCC-2).
Translation, adaptation and conceptual equivalence**

Rédactrice invitée Phaedra Royle

Marie Vézina, Catherine Samson-Morasse, Julie Gauthier-Desgagné,
Marion Fossard, Audette Sylvestre

Article 4.....254

**Résultats du Boston Naming Test chez des adultes
francophones de Québec**

Rédacteur en chef adjoint Jeff Small

Patricia M. Roberts, Nathalie Doucet

Évaluation des ouvrages écrits268

Science of Successful Supervision and Mentorship

Ian Roth

Interventions for Speech Sound Disorders in Children

Bojana Radovanovic

From the Editor

FALL ISSUE



DR. ELIZABETH FITZPATRICK WILL BE THE NEW EDITOR OF THE CANADIAN JOURNAL OF SPEECH-LANGUAGE PATHOLOGY AND AUDIOLOGY

The *Canadian Association of Speech-Language Pathologists and Audiologists* has completed the search for the incoming editor and Dr. Elizabeth Fitzpatrick has accepted the position. Elizabeth is an Associate Professor at the École des sciences de la réadaptation at the University of Ottawa. She already knows the CJSPLA very well through her work as associate editor for audiology from 2008-2010. Elizabeth is both a frequent reviewer as well as a regular contributor, and she has been enthusiastic and steadfast in her support of the journal. On behalf of the whole team of the CJSPLA, we would like to congratulate Elizabeth on her appointment.

2010 Editor's award

This year, the 2010 Editor's Award was presented at the awards banquet at the *Canadian Association of Speech-Language Pathologists and Audiologists's* (CASPLA) 2011 annual conference in Montreal, Quebec. The winner of the 2010 Editor's Award was the paper 'A Pilot Study to Evaluate a New Early Screening Instrument for Speech and Language Delays,' by Janis Carscadden, Pamela Corsiatto, Lita Ericson, Robin Illchuk, Carrie Esopenko, Erin Sterner, Gregory D. Wells and Scott Douglas Oddie (CJSPLA, 34, 87-95). On behalf of the *Canadian Association of Speech-Language Pathologists and Audiologists*, I congratulate the authors for receiving this award!

The two other papers nominated for the award deserve an honourable mention. Louise Duchesne, Ann Sutton, François Bergeron and Natacha Trudeau published a study entitled 'Le développement lexical précoce des enfants porteurs d'un implant cochléaire' (CJSPLA, 34, 132-145). Pamela Millett and Neil Purcell were nominated for their contribution 'Effect of sound field amplification on grade one reading outcomes' (CJSPLA, 34, 17-24).

Current issue

There are four papers in the current issue of the CJSPLA. In the first paper, 'Speech audiometry with non-native English speakers: The use of digits and Cantonese words as stimuli', Stefka H. Marinova-Todd, Carrie K. Siu and Lorianne M. Jenstad examine the validity of English speech audiometry with non-native English speakers.

The second paper was written by Catriona M. Steele, Sonja M. Molfenter, Gemma L. Bailey, Rebecca Cliffe Polacco, Ashley A. Waito, Dana C. B. H. Zoratto and Tom Chau. It is entitled 'Exploration of the utility of a brief swallow screening protocol with comparison to concurrent videofluoroscopy.' This article compares the results of a brief swallowing screening protocol and videofluoroscopy of the same swallows in 40 adults.

Marie Vézina, Catherine Samson-Morasse, Julie Gauthier-Desgagné, Marion Fossard, and Audette Sylvestre contributed an article entitled 'Développement de la version québécoise francophone du Children's Communication Checklist - 2 (CCC-2). Traduction, adaptation et équivalence conceptuelle.' In this study, the translation and validation of an assessment tool for francophone children's pragmatic abilities is described.

The fourth and final paper is entitled 'Performance of French-speaking Quebec adults on the Boston Naming Test' and was authored by Patricia M. Roberts and Nathalie Doucet. This study examined the stimuli of the Boston Naming Test for name agreement in Quebec French with 45 older adults.

The current issue of the CJSPLA includes two book reviews. Ian Roth reviews 'Science of Successful Supervision and Mentorship', by Linda Carozza, and Bojana Radovanovic reviews 'Interventions for Speech Sound Disorders in Children' by A. Lynn Williams, Sharynne McLeod and Rebecca J. McCauley.

Tim Bressmann, Ph.D., Associate Professor
Department of Speech-Language Pathology, University of Toronto
Rehabilitation Sciences Building
160-500 University Avenue, Toronto, ON M5G 1V7, Canada
Tel. +1-416-978-7088, Fax -1596
tim.bressmann@utoronto.ca, www.slp.utoronto.ca
Associate Professor, Faculty of Dentistry, University of Toronto
Associate Scientist, Toronto Rehabilitation Institute

Mot de rédacteur en chef

NUMÉRO D'AUTOMNE



DR ELIZABETH FITZPATRICK SERA LA NOUVELLE RÉDACTRICE EN CHEF DE LA REVUE CANADIENNE D'ORTHOPHONIE ET D'AUDIOLOGIE.

L'Association canadienne des orthophonistes et audiologistes a terminé sa recherche du futur rédacteur en chef, car madame Elizabeth Fitzpatrick a accepté le poste. Elizabeth est professeure adjointe à l'École des sciences de la réadaptation de l'Université d'Ottawa. Elle connaît déjà très bien la RCOA, grâce à son travail de rédactrice en chef adjointe pour la section audiologie de 2008 à 2010. Elizabeth travaille sur une base régulière à titre de réviseure et de collaboratrice et a toujours fait preuve d'enthousiasme et de constance dans son appui à la revue. Au nom de toute l'équipe de la RCOA, nous aimerions féliciter Elizabeth pour sa nomination.

Prix du rédacteur en chef 2010

Cette année, le prix du rédacteur en chef 2010 a été présenté lors du banquet de remise des prix, qui a eu lieu au Congrès annuel 2011 de l'Association canadienne des orthophonistes et audiologistes (ACOA) à Montréal, Québec. L'article gagnant du prix du rédacteur en chef 2010 est « Une étude pilote pour évaluer un nouvel instrument de dépistage précoce des retards de la parole et du langage », de Janis Carscadden, Pamela Corsiatto, Lita Ericson, Robin Illchuk, Carrie Esopenko, Erin Sterner, Gregory D. Wells et Scott Douglas Oddie (RCOA, 34, 87-95). Au nom de l'Association canadienne des orthophonistes et audiologistes, je tiens à féliciter les auteurs pour ce prix!

Les deux autres articles qui étaient en nomination pour ce prix méritent une mention honorable. Louise Duchesne, Ann Sutton, François Bergeron et Natacha Trudeau ont publié une étude intitulée « Le développement lexical précoce des enfants porteurs d'un implant cochléaire » (RCOA, 34, 132-145). Pamela Millett et Neil Purcell étaient pour leur part en nomination pour leur article « Effet de l'amplification en champ libre sur les performances de lecture des élèves de première année » (RCOA, 34, 17-24).

Présent numéro

Il y a quatre articles dans le présent numéro de la RCOA. Le premier article, « Audiométrie vocale chez des personnes dont la langue maternelle n'est pas l'anglais : l'utilisation de chiffres et de mots en cantonais comme stimuli », a été rédigé par Stefka H. Marinova-Todd, Carrie K. Siu et Lorienne M. Jenstad. Ils examinent la validité d'une audiométrie vocale en anglais chez des personnes dont la langue maternelle n'est pas l'anglais.

Le deuxième article a été écrit par Catriona M. Steele, Sonja M. Molfenter, Gemma L. Bailey, Rebecca Cliffe Polacco, Ashley A. Waito, Dana C. B. H. Zoratto et Tom Chau. Il s'intitule « Exploration de l'utilité d'un bref protocole de dépistage des troubles de déglutition en comparaison avec une vidéofluoroscopie simultanée ». Cet article compare les résultats d'un bref protocole de dépistage des troubles de déglutition et d'une vidéofluoroscopie des mêmes déglutitions chez 40 adultes.

Marie Vézina, Catherine Samson-Morasse, Julie Gauthier-Desgagné, Marion Fossard, et Audette Sylvestre ont écrit l'article « Développement de la version québécoise francophone du Children's Communication Checklist – 2 (CCC-2). Traduction, adaptation et équivalence conceptuelle ». Cette étude décrit la traduction et la validation d'un outil pour évaluer la pragmatique chez les jeunes francophones du Québec.

Le quatrième et dernier article s'intitule « Résultats du Boston Naming Test chez des adultes francophones de Québec » et a été écrit par Patricia M. Roberts et Nathalie Doucet. Cette étude a examiné les stimuli du Test de dénomination de Boston, auprès de 45 adultes, pour évaluer la dénomination des noms en français québécois.

Vous trouverez également dans le présent numéro de la RCOA deux comptes rendus de livre. Ian Roth analyse pour nous « Science of Successful Supervision and Mentorship », de Linda Carozza, et Bojana Radovanovic analyse « Interventions for Speech Sound Disorders in Children », de A. Lynn Williams, Sharynne McLeod et Rebecca J. McCauley.

Tim Bressmann, Ph.D., professeur agrégé
Département d'orthophonie, Université de Toronto
Édifice des Sciences de la réadaptation
160-500 Avenue University, Toronto, ON M5G 1V7, Canada
Tel. +1-416-978-7088, Fax -1596

tim.bressmann@utoronto.ca, www.slp.utoronto.ca

Professeur agrégé, Faculté de médecine dentaire, Université de Toronto
Scientifique agrégé, Institut de réadaptation de Toronto

- **Speech audiometry with non-native English speakers: The use of digits and Cantonese words as stimuli**
- **Audiométrie vocale chez des personnes dont la langue maternelle n'est pas l'anglais : l'utilisation de chiffres et de mots en cantonais comme stimuli**

Stefka H. Marinova-Todd
 Carrie K. Siu
 Lorienne M. Jenstad

KEY WORDS

SRT

WRS

DIGITS MULTILINGUAL

SPEECH AUDIOMETRY

NON-NATIVE ENGLISH

CANTONESE

Stefka H. Marinova-Todd,
 Ed.D.
 School of Audiology
 and Speech Sciences
 University of British
 Columbia
 2177 Wesbrook Mall
 Vancouver, B.C.
 Canada

Carrie K. Siu,
 M.Sc. Aud(C)
 3740 Albert St #1104
 Burnaby, B.C. V5C 5Y7
 Canada

Lorienne M. Jenstad,
 Ph.D.
 School of Audiology
 and Speech Sciences
 University of British
 Columbia
 2177 Wesbrook Mall
 Vancouver, B.C.
 Canada

Abstract

This pilot study investigated validity of English speech audiometry with non-native English speakers. Two widely used procedures in speech audiometry, the Speech Reception Threshold and the Word Recognition Score, were administered to 45 adults with English as their second language, 30 of whom were Cantonese native speakers. The effects of test stimuli (English words, English digits and Cantonese words) on the test performance were analyzed. English digit pair stimuli were found to be more accurate predictors of pure-tone average than English word stimuli for all participants, while Cantonese words elicited the lowest speech audiometric thresholds from the Cantonese-speaking participants. In terms of word recognition scores, the effect of noise was largest when testing was done in the second language. The subjects with hearing impairment were not disadvantaged when tested in their first language, Cantonese, but they had significantly lower scores when tested in their second language, English. The results from this study are of theoretical importance. In order to determine their clinical significance, more research with larger sample sizes is necessary. We conclude that clinicians should use caution when interpreting the results from a speech test when assessing non-native English-speaking clients.

Abrégé

Cette étude pilote a examiné la validité d'une audiométrie en anglais chez des personnes dont la langue maternelle n'est pas l'anglais. Un groupe de quarante-cinq adultes, pour qui l'anglais est la langue seconde et dont trente parlent le cantonais comme langue première, a passé deux tests très répandus en audiométrie vocale : le seuil d'intelligibilité et le pourcentage de reconnaissance des mots. Les effets des stimuli (mots anglais, chiffres anglais et mots cantonais) sur les résultats du test ont été analysés. Les chiffres en anglais étaient des indicateurs plus précis de la moyenne des sons purs que les mots en anglais pour tous les participants, alors que les mots en cantonais ont obtenu les seuils d'audiométrie vocale les plus bas pour les participants parlant le cantonais. En ce qui concerne le pourcentage de reconnaissance des mots, les répercussions du bruit étaient plus grandes lorsque le test était effectué dans la langue seconde. Les sujets souffrant de troubles auditifs n'étaient pas désavantagés lorsque le test était effectué dans leur langue première, le cantonais, mais ils ont obtenu des résultats significativement plus bas lorsque testés dans leur langue seconde, l'anglais. Les résultats de cette étude ont une importance théorique. Afin de déterminer leur signification clinique, il est nécessaire d'effectuer davantage de recherches sur des échantillons plus grands. Nous avons conclu que les cliniciens devraient faire preuve de prudence lorsqu'ils interprètent les résultats d'un test oral effectué chez des patients dont la langue maternelle n'est pas l'anglais.

A survey conducted by the Canadian Association of Speech-Language Pathologists and Audiologists in 2003 showed that among its 423 registered audiologists, 93% spoke English as their first language and the rest spoke French as their first language. However, the 2006 Census conducted by Statistics Canada revealed 19.7% of Canadians spoke neither English nor French as their first language. This percentage is much higher in some of the major cities. For example, 41.7% of the citizens in the Greater Vancouver metropolitan area of British Columbia report neither English or French as a first language. Eighteen percent of the population in this metropolitan area reported that Chinese was their mother tongue (Statistics Canada, 2006). The underrepresentation of language minority members in the Audiology profession is similarly seen in the United States, where only 7% of certified Audiologists identified themselves as belonging to racial minority groups (American Speech, Language and Hearing Association, 2009). These statistics demonstrate the need for a linguistically and culturally sensitive approach towards conducting speech audiometry with non-native English speakers. When hearing tests using English speech stimuli are administered to people for whom English is not their first language, knowledge and proficiency of English, as well as hearing sensitivity, may contribute significantly to test performance.

SPEECH AUDIOMETRY WITH NON-NATIVE ENGLISH SPEAKERS

Beverley-Ducker (2003) reported an increasing need to prepare for, and respond to, racial, ethnic and linguistic diversity of current and future caseloads as well as a need to develop culturally and linguistically sensitive assessment tools. She also pointed out a need to conduct research on the reliability of speech audiometric test results with non-native English speakers. Although most clinicians are aware of the need for more linguistically sensitive assessment tools, they still use English speech tests for non-English-speaking clients because of their availability, longevity, research support, and most importantly, compatibility with their own language (Ramkissoon & Khan, 2003). In some cases, audiologists who use alternative speech audiometry tests use subsets of the English standardized word lists. However, research showed that using a smaller list than the standardized list resulted in better Speech Reception Thresholds (SRTs) due to familiarization with test stimuli, thus sacrificing test validity (Ramkissoon, Proctor, Lansing, & Bilger, 2002). Another alternative is to administer speech audiometry in the client's native language. Speech tests have been developed and standardized for *monolingual* speakers

of languages other than English (e.g., *Arabic*: Ashoor & Prochazka, 1982; *Canadian French*: Vaillancourt, Laroche, Mayer, Basque, Nali, Eriks-Brophy, Soli, & Giguère, 2005; *Cantonese*: Wong & Soli, 2005; *Danish*: Wagener, Josvassen, & Ardenkjoer, 2003; *Mandarin*: Nissen, Harris, & Slade 2007; *Russian*: Harris, Nissen, Pola, McPherson, Tavartkiladze, & Eggett 2007; *Spanish*: Ferrer, 1960; Zubic, Irizarry, Rosen, Feudo, Kelly, & Strome, 1983; and *Swedish*: Hällgren, Larsby, & Arlinger, 2006.), but it is not clear whether these tests are appropriate for bilingual speakers. In the field of bilingualism and second language acquisition, it has been established that bilingual individuals perform differently from monolinguals on a variety of tasks, including both language-based and non-language-based tasks (Michael & Gollan, 2005). It is unknown how bilinguals perform on speech tests that have been developed and standardized on monolingual populations alone.

Research findings suggest that when the clinician does not speak the client's first language, it may be more valid to use digit pairs in SRT measures than spondees. Ramkissoon and colleagues (2002) compared the SRT of native English speakers and non-native speakers using digit pairs versus standardized spondees as stimuli. They assessed the accuracy of the two stimuli by comparing the SRT obtained with the pure tone average (PTA), and found that compared with compound words, digit pairs more accurately measured the hearing threshold for speech of non-native English speakers.

THE CURRENT STUDY

The present pilot study aimed to expand upon previous research on issues relating to speech audiometry and non-native English speakers by exploring the most appropriate auditory stimuli for accurate measurement of SRT and Word Recognition Scores (WRSs) for this group. In the first part of the study, we aimed to replicate Ramkissoon et al.'s (2002) study, which had the goal of determining whether digit pairs or spondees were more accurate indicators of hearing thresholds for non-native English speakers. In addition, the present study aimed to determine which speech audiometry test stimuli (i.e., English spondees, English digits or Cantonese spondees) led to the most accurate measure of hearing sensitivity in a Cantonese-speaking group. The second part investigated Cantonese speakers' word recognition performance in quiet and noise using Cantonese versus English stimuli.

The present study had the following main research question: Is there a test stimulus effect on performance in speech audiometry for non-native English speakers? Additional specific questions addressed were:

1. Does SRT differ depending on stimuli used

(English spondees vs. digit pairs vs. first language spondees)?

2. Do WRS in quiet and in noise differ depending on language of stimuli used (first language or second language)?

METHOD

Participants

Participants were recruited by word-of-mouth, by email, and by advertisements posted at local university campuses and at various libraries and community centers across a major metropolitan city in Canada. A total of 45 non-native English speakers participated (see Table 1). Among them, 30 spoke Cantonese as their native language. The native languages of the remaining 15 subjects were Tagalog, Japanese, German, Bulgarian, Punjabi, Mandarin, and French. The subjects' ages ranged from 19 to 69 years, with a mean age of 48 years for the Cantonese-speaking group and 37 years for the group who spoke other languages. The Cantonese speakers were evenly divided in terms of gender (15 males; 15 females), as were the speakers of other languages (7 males; 8 females). Sixteen out of the 45 participants (2 in the non-Cantonese group and 14 in the Cantonese group) had a hearing loss: Eight subjects had hearing loss in the high frequencies (2000-4000 Hz), 1 subject had hearing loss in the low frequencies (250-1500Hz), and 7 subjects had hearing loss in both high and low frequencies. Normal hearing was considered pure tone thresholds of 25 dB HL or better from 250 to 4000 Hz in both ears.

Table 1

Background characteristics for all subjects (N=45).

	Cantonese Mean (SD)	Non-Cantonese Mean (SD)
N	30	15
PTA (dB HL)	12.75 (15.92)	6.22 (10.75)
Length of Residence in L2 Country (yrs)	12.53 (4.31)	11.87 (13.05)
Age of first exposure to L2 (yrs)	8.5 (2.98)	11.33 (8.30)
Years of L2 instruction	10.87 (4.69)	9.33 (4.64)
Age at time of testing	48.17 (15.24)	37 (12.60)

Stimuli

Stimuli for the English SRT were the 18 spondees of the American Speech and Hearing Association half list A (American Speech and Hearing Association, 1988), adapted from the CID W-1 word list. Stimuli for the digit SRT test were compiled in the same way as the study conducted by Ramkissoon et al. (2002). Two individual numbers from "1" to "9", excluding "7", were paired. The number "7" was excluded because it has two syllables

and, when paired with another digit, would result in digit pairs with more than two syllables. We obtained 56 digit pairs with no item containing a repeated number. To match the number of stimuli for the CID-spondees, 18 pairs were randomly selected from these 56 pairs for use in the study.

Stimuli for the WRS were words from the NU-6 List 3A. This list, among the six lists developed by Tillman and Carhart (1966), consisted of 50 phonemically balanced monosyllabic words. For this study, two 25-word lists - the first and last 25 words of NU-6 List 3A - were randomly used during testing. Half-lists (25 words), instead of full lists of 50 words, were used because a survey conducted in 2000 indicated that this was the list of choice by most Canadian audiologists (DeBow & Green, 2000). Therefore, half-lists were used to increase the efficiency and clinical applicability of this study.

Stimuli for the Cantonese SRT and WRS comprised lists of Cantonese spondees and monosyllabic words. At the time of the present study, standardized Cantonese word lists were not available. The stimuli in the present study were developed by Cantonese-speaking clinical audiologists and were discussed with three native Cantonese speakers who had academic training in linguistics. Eventually, these lists were deemed to adequately represent all Cantonese speech sounds in a phonemically-balanced manner. The lists consisted of common words that would be familiar to adult participants who are native speakers of Cantonese.

Procedure

Bilateral pure tone air conduction thresholds at 500, 1000, 2000, and 4000 Hz and SRTs were measured using standard clinical procedures (American Speech and Hearing Association, 1978, 1979, 1988). Pure-tone averages were calculated from thresholds at 500, 1000, and 2000 Hz, using the Fletcher 2-frequency average (Fletcher, 1950) where appropriate (i.e., where adjacent thresholds differed by 20 dB or more). For testing word

recognition in quiet, subjects were presented with 25 words from NU-6 List 3A to both ears simultaneously at a supra-threshold level, i.e. the higher of 45 dB HL or the Most Comfortable Level. Participants were instructed to repeat aloud the words heard and the percentage of correctly repeated words was calculated.

Word recognition in noise testing was performed binaurally via two standard CD players, with the recorded NU-6 words routed through one CD player and the Auditec cafeteria noise (Auditec of St. Louis, St. Louis, MO) routed through the second player. The twenty-five NU-6 List 3A words along with background cafeteria noise were presented diotically via supra-aural earphones (TDH 50-P, Telephonics, Farmingdale, NY). The two half-lists were randomized depending on whether they were tested in quiet or noise for each listener. The words were presented at the same intensity as when testing in quiet, while the noise was presented at an intensity of 5 dB lower than the presentation level of the words (signal-to-noise ratio of +5 dB).

For the Cantonese-speaking participants, in addition to word recognition in English, word recognition in Cantonese was tested using the same procedure as in English. A 25-word list was presented at the same intensity as described above in quiet and in noise. The noise used was the same track of cafeteria noise as used in English testing, presented at a signal-to-noise level of +5 dB. Two recorded Cantonese word lists were used and their order of presentation was randomized among participants. Half of the subjects heard the stimuli in English first, and the other half heard the stimuli in Cantonese first. All Cantonese stimuli were spoken by a Cantonese-English bilingual female speaker and were digitally recorded with computer software and equated for peak intensity.

All testing was performed using a diagnostic audiometer (Grason-Stadler, GSI 61, Eden Prairie, MN) in a sound-treated booth where the ambient noise met the ANSI standard for audiometric testing (ANSI, 1979). At the end of testing, each participant was briefed on the test results.

Results

In a first step of the analysis, we confirmed for each of the stimulus sets that measures from both ears of listeners were highly correlated and not significantly different. Therefore, the average of scores from the left and right ears were used as the measurement. The effects of hearing loss on SRT and WRS were also analyzed.

SRT Results

A mixed-design ANOVA was used to investigate any differences in SRT performance between the Cantonese-speaking group and the non-Cantonese-speaking group,

and between the hearing-impaired and normal-hearing groups. The data were analyzed using a 2 (language) x 2 (hearing status) x 2 (stimulus type) repeated-measures ANOVA, where language (Cantonese and non-Cantonese speakers) and hearing status (normal hearing and hearing impaired) were the between-subject factors, and the stimulus type (CID spondees vs. digits) was the within-subject factor. There was a significant main effect of hearing status ($F(1, 41) = 54.89, p < 0.0001$), where hearing-impaired individuals had higher SRTs than normal-hearing individuals, a significant main effect of language ($F(1, 41) = 7.16, p = 0.01$), a significant main effect of stimulus type ($F(1, 41) = 40.55, p < 0.0001$), such that subjects had a significantly lower mean SRT when the stimuli were digit pairs than when they were English spondees, and a significant interaction between hearing status, language and hearing test ($F(1, 41) = 7.21, p = 0.01$). Subsequent analysis revealed that in the normal-hearing group there was only a significant main effect of stimulus type ($F(1, 27) = 11.49, p = 0.002$) indicating that both Cantonese and non-Cantonese-speaking subjects had lower (better) thresholds for digits than spondees; however, the main effect of language group ($F(1, 27) = 0.01, p = 0.93$) and the interaction between language group and stimulus type ($F(1, 27) = 1.36, p = 0.25$) were not significant. Similarly, in the hearing-impaired group there was a significant main effect of stimulus type ($F(1, 14) = 17.83, p = 0.0009$) but no significant main effect of language ($F(1, 14) = 3.36, p = 0.09$). However, the interaction between stimulus type and language was significant ($F(1, 14) = 5.09, p = 0.03$) and revealed that the difference between the digits and spondee scores was greater among the non-Cantonese speakers. Due to the very small sample size of the non-Cantonese group, caution needs to be applied when interpreting the results of the hearing-impaired non-Cantonese group. Results are displayed in Figure 1.

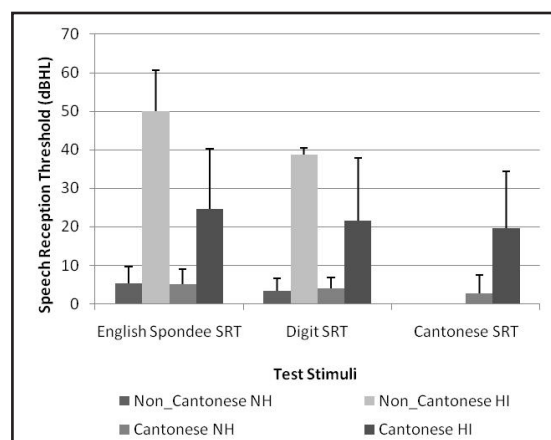


Figure 1: Comparison of the mean SRT between the normal hearing group and the hearing-impaired group of non-Cantonese ($N = 15$) and Cantonese-speaking subjects ($N=30$) (NH = normal hearing; HI = hearing impaired).

To determine accuracy of the SRT measure, SRT values were compared with pure tone averages. Paired *t*-tests revealed a significant difference between PTA (*Mean* = 10.57, *SD* = 14.67) and CID-SRT (*Mean* = 13.22, *SD* = 16.77), $t(89) = -3.87, p = 0.0002$. No significant difference was found between PTA and Digit-SRT (*Mean* = 10.89, *SD* = 15.39), $t(89) = -0.58, p = 0.56$. Correlational analyses revealed a high correlation between both Digit-SRT and PTA ($r = 0.94, p < 0.0001$) and between CID-SRT and PTA ($r = 0.92, p < 0.0001$). Therefore, digit pairs were found to be accurate in predicting the hearing threshold for speech for non-native speakers and provide a better alternative to the CID W-1 spondees. Moreover, for the group on average, the difference between PTA and SRT was less than 6 dB for both stimulus types, which is considered to be a good agreement (Brandy, 2002).

COMPARISONS BETWEEN ENGLISH AND CANTONESE TEST MATERIALS

The largest group of subjects was from Cantonese background, and we examined their performance on the English and Cantonese SRT. The data were analyzed using a 2 (hearing status) \times 3 (stimulus type) repeated-measures ANOVA, where hearing status (normal hearing and hearing impaired) was the between-subjects factor, and the stimulus type (English spondees vs. digits vs. Cantonese spondees) was the within-subject factor. The descriptive statistics on the English and Cantonese tests are presented in Table 2. The univariate tests of repeated measures revealed a significant main effect of stimulus type ($F(2,56) = 21.29, p < 0.0001$), a significant main effect of hearing status ($F(1, 28) = 20.60, p < 0.0001$), and a significant interaction between hearing status and stimulus type ($F(2,56) = 3.51, p = 0.04$). Subsequent analyses revealed that it was only in the hearing-impaired group that there was a significant main effect of stimulus type ($F(2, 13) = 21.60, p < 0.0001$), and an examination of Figure 1 indicated that the hearing-impaired subjects had significantly lower Cantonese SRT scores than digits ($p = 0.009$), and their Cantonese and digit SRT scores were significantly lower than their English SRT scores ($p = 0.0001$ and $p = 0.003$, respectively).

When compared with the mean pure-tone average (*Mean* = 12.75, *SD* = 15.92), a significant difference was found between PTA and English SRT (*Mean* = 14.17, *SD* = 17.03), $t(59) = -2.14, p = 0.0369$, and between PTA and Chinese SRT (*Mean* = 10.67; *SD* = 16.53), ($t(59) = 3.16, p = 0.0025$), while no significant difference was found between PTA and Digit SRT (*Mean* = 12.25; *SD* = 16.45), $t(59) = 0.75, p = 0.4566$. Correlation analyses revealed strong positive relationships between PTA and English spondee SRT ($r = 0.953, p < 0.0001$), PTA and Digit SRT ($r = 0.950, p < 0.0001$), and PTA and Cantonese SRT ($r = 0.951, p < 0.0001$). Therefore, for the Cantonese-speaking participants, digit pairs resulted in more accurate measures of hearing sensitivity than either English or Cantonese stimuli. For the group on average, the difference between PTA and SRT was less than 6 dB for both stimulus types, which is considered to be a good agreement (Brandy, 2002).

PERFORMANCE ON WRS

WRS was conducted binaurally on all participants but only the Cantonese speakers were tested with the Cantonese stimuli. Because a main goal of the study was to compare performance for stimuli in the native language to performance for stimuli in the non-native language, only results from the 30 Cantonese-speaking participants are presented here. The descriptive statistics on word recognition are presented in Table 2, with the data in rationalized arcsine units (RAUs; Studebaker, 1985). The data were analyzed using a 2 (hearing status) \times 2 (noise) \times 2 (stimulus language) repeated-measures ANOVA, where hearing status (normal hearing and hearing impaired) was the between-subject factor, and noise (noise vs. quiet) and stimulus language (English vs. Cantonese words) were the within-subject factors. There was a significant main effect of hearing status ($F(1, 28) = 7.61, p = 0.01$), a significant main effect of noise ($F(1, 28) = 57.65, p < 0.0001$), and a significant main effect of stimulus language ($F(1, 28) = 73.24, p < 0.0001$). While the interaction among all three factors was not significant ($F(1, 28) = 0.78, p = 0.38$), the interactions between stimulus language and hearing status

Table 2

Word recognition scores (WRS) in RAUs for different conditions for the Cantonese-speaking subjects only (N = 30).

	English Quiet Mean (SD)	English Noise Mean (SD)	Cantonese Quiet Mean (SD)	Cantonese Noise Mean (SD)
Normal Hearing (N = 16)	95.00 (5.16)	88.00 (8.39)	99.25 (1.61)	96.00 (4.38)
Hearing Impaired (N = 14)	85.79 (13.37)	73.43 (16.05)	96.86 (6.11)	93.43 (7.12)

($p(1, 28) = 9.29, p = 0.005$), and stimulus language and noise ($F(1, 28) = 6.02, p = 0.02$) were both statistically significant. The two interaction effects are displayed on Figure 2 and Figure 3 respectively.

Upon closer examination of Figure 2 and the simple main effects it became apparent that the word recognition scores in English were significantly lower than the word recognition scores in Cantonese for both the normal-hearing ($F(1, 15) = 23.14, p < 0.0001$) and hearing-impaired groups ($F(1, 13) = 47.02, p < 0.0001$). On the other hand, while the English scores of the hearing-impaired group are significantly lower than those of the normal-hearing group ($F(1, 28) = 10.19, p = 0.003$), there was no significant difference between the groups when tested in Cantonese ($F(1, 28) = 1.82, p = 0.19$). In other words, the effect of hearing impairment was greater when tested in the second language than in the native language.

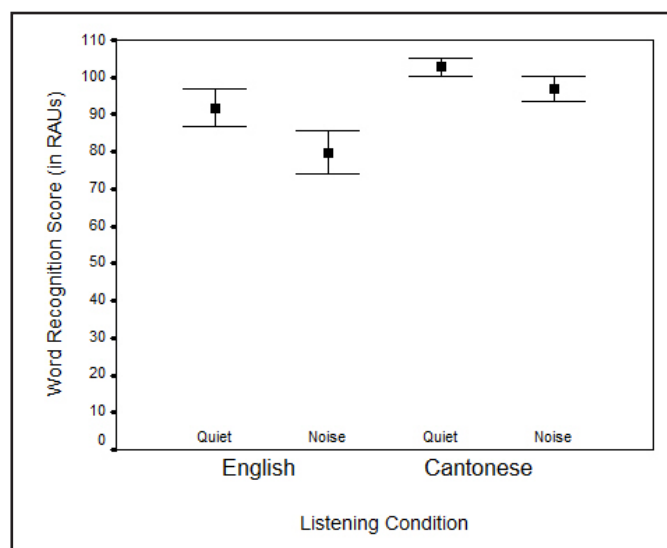


Figure 2: Interaction between the Noise and Language variables, collapsed across hearing status for the Cantonese-speaking subjects ($N = 30$).

Upon closer examination of Figure 3 and the simple main effects, it was found that the word recognition scores were lower in English than in Cantonese when tested both in quiet ($F(1, 29) = 28.96, p < 0.0001$) and in noise ($F(1, 29) = 45.35, p < 0.0001$). Moreover, subjects' WRSs were lower when tested in noise than when tested in quiet in both English ($F(1, 29) = 34.61, p < 0.0001$) and Cantonese ($F(1, 29) = 16.32, p < 0.0001$). However, the effect of noise was greater when the subjects were tested in their second language, English.

Overall, for this subgroup of Cantonese speakers, the mean supra-threshold word recognition was significantly better when stimuli were presented in Cantonese than when they were presented in English. This finding applied to testing in quiet and in noise, and to both normal-hearing and hearing-impaired listeners.

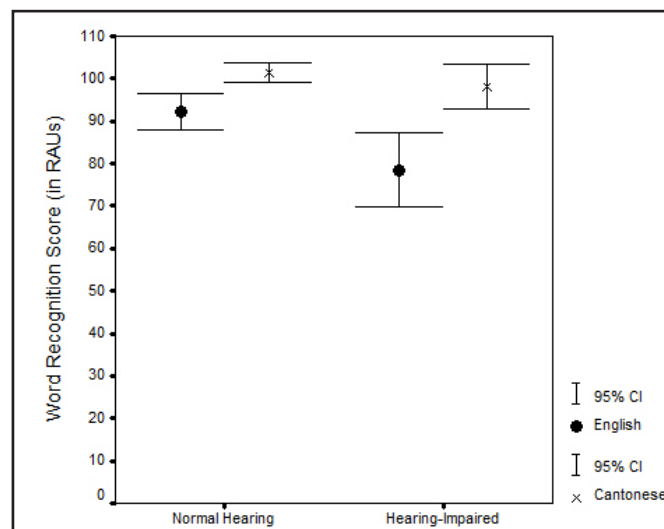


Figure 3: Interaction between Language and Hearing status, collapsed across noise for the Cantonese-speaking subjects ($N = 30$).

DISCUSSION

The present pilot study confirmed Ramkissoon and colleagues' (2002) finding that compared with CID spondee, digit pairs elicited more accurate SRTs that closely approximated the PTA. Similarly, we conclude that compared to CID spondee, digit pairs should facilitate more accurate SRT testing for non-native speakers of English. In addition, results from the Cantonese participants suggested that PTA might not accurately reflect hearing sensitivity for speech in all languages. The standard calculation for PTA was derived from the acoustic spectrum of English speech sounds, in which the majority of sounds lies within 500-2000 Hz. The lower threshold obtained from using Cantonese spondee compared to the PTA suggests that the frequencies 500-2000 Hz may not be an accurate representation of Cantonese speech sounds. The frequency-importance function for Cantonese sentences shows that low-frequency information is more important for sentence recognition in Cantonese than in English (Wong, Ho, Chua, & Soli, 2007). The authors of that study suggested that the increased low-frequency importance was in part due to the tonal nature of the Cantonese language. It can be speculated that similar differences in frequency importance might also apply to the spondaic materials used in the current study.

A significant effect due to stimulus language was also found in the WRS results. Both in quiet and in noise, performance of the Cantonese-speaking participants for supra-threshold word recognition was significantly better when stimuli were in their first language than in English. Moreover, the effect of noise was greater when the stimuli were presented in the second language than in the native language of the participants. This effect is consistent with previous research, in which a decrease in word recognition

performance in noise in a non-native language has been documented (von Hapsburg & Pena, 2002). Most of the past studies compared non-native speakers' performance with that of native speakers. Despite effort in matching individual participants' characteristics between the native and non-native groups, inter-subject variability, nevertheless, could not be entirely eliminated. The present study compared performance between a native and a non-native language within subjects, thereby eliminating the effect of inter-subject variability on the results.

We also found that for both normal-hearing and hearing-impaired groups of Cantonese listeners, noise had little effect on WRS when stimuli were in their native language. In an effort to keep experimental conditions consistent, the same cafeteria noise was used in the measurement of both English and Cantonese WRS. Due to the fact that the background speech in the cafeteria noise was in English, a release of masking effect might have resulted during Cantonese WRS measurement because the signal and noise were in different languages. As quantified by the frequency-importance function, the important acoustic cues within Cantonese speech span a frequency range different from English speech, possibly enabling a release of masking when the competing background noise was in English while the signal was in Cantonese. Generalizing our results beyond the lab, it could be hypothesized that Cantonese speakers in English-speaking societies may be less affected by background English speech, and that Cantonese hearing-impaired people may experience less social impairment than their English-speaking counterparts.

IMPLICATIONS

The major findings of the present study are 1) for non-native speakers of English, digit pairs as SRT stimuli more accurately measured hearing threshold for English speech than CID W-1 spondee; 2) for the Cantonese participants, digit pairs also more accurately measured hearing threshold for English speech than either English or Cantonese spondee, however, Cantonese spondee elicited a better hearing threshold than English stimuli; 3) for the Cantonese participants, performance for monosyllabic word recognition was significantly better when stimuli were in Cantonese than when they were in English; this effect existed across testing conditions (quiet and noise) and despite the presence of hearing loss; and 4) for the Cantonese participants, the presence of hearing loss affected word recognition in English significantly more than word recognition in Cantonese.

Of importance for clinicians, the present study found the language of test stimuli to significantly affect test accuracy in speech audiometry for non-native English

speakers. Although the average discrepancy between SRT and PTA was within 6 dB for both stimulus types, which is considered "good" agreement (Brandy, 2002), the individual data show that for the CID spondee, 29 of 90 ears were not in good agreement (i.e., more than 6 dB discrepancy), and for the digit spondee, only 17 of 90 ears were not in good agreement. The clinical importance of this needs to be determined via larger and more varied sample sizes. As the consistency between SRT and PTA is an important indicator of pseudohypacusis, and testing in the non-native language may result in discrepancy between SRT and PTA, pseudohypacusis may be misdiagnosed when insufficient knowledge of the native language may be the true cause of SRT-PTA discrepancy. Therefore, clinicians serving multicultural clients should be aware that it may not be only hearing sensitivity that they are measuring, but also their clients' language ability.

The interesting supplemental finding that hearing-impaired listeners performed differently from normal hearing listeners only in their non-native language (English) has significant clinical implications. As most clients at an Audiology clinic have a certain degree of hearing loss, extra caution should be taken in applying and interpreting English WRS for non-native speakers of English because a hearing loss increases the confounding effect of language on their performance in English WRS. This evidence strongly supports the use of speech recognition tests in the client's native language. Many versions of the Hearing In Noise Test are being developed and validated in languages other than English (e.g., Cantonese; Wong and Soli 2005). The clinician should be aware of the main language groups in their community and, if appropriate test materials have been developed for that language, ensure that these speech tests are available in their clinic. However, it is important that these tests are administered by native speakers of the language whenever possible because the use of such tests may increase the possibility of "auditor errors" (e.g., Nelson & Chaiklin, 1970), which could presumably increase when administered by non-native speakers of the language.

FUTURE DIRECTIONS

As this was a pilot study, we included a small group of hearing-impaired subjects to assure applicability and generalization of results to clinical populations. Due to the small sample size, however, all hearing-impaired participants were categorized into one group for comparison to the normal hearing group. In future studies, the use of a larger number of hearing-impaired participants is recommended in order to define the

relationship between hearing loss and the effects noted in the current study. We chose to include a clinician-developed WRS test in Cantonese that was the only option available to us at the time. In future studies, researchers should aim to develop standardize speech tests in Cantonese with known psychometric properties. Moreover, it would be informative to include an English-speaking comparison group as a control for the English speech tests, as well as to provide a context for the interpretation of the results.

REFERENCES

- American National Standards Institute (1979). *American National Standard Specification for Audiometers*. New York: American National Standards Institute, Inc.
- American Speech, Language and Hearing Association (2009). *Summary Counts by Ethnicity and Race*. Retrieved from <http://www.asha.org/uploadedFiles/2009MemberCounts.pdf>
- American Speech and Hearing Association (1988). *Determining the Threshold Level for Speech*. [Guidelines] Retrieved from <http://www.asha.org/policy>
- American Speech and Hearing Association (1979). Guidelines for determining the threshold level for speech. *ASHA*, 21, 353–356.
- American Speech and Hearing Association (1978). Guidelines for manual pure tone speech audiometry. *ASHA*, 20, 297–301.
- Ashoor, A.A., & Prochazka Jr., T. (1982). Saudi Arabic speech audiometry. *Audiology*, 21, 493–508.
- Beverly-Ducker, K. (2003). Multicultural issues in Audiology. *ASHA Division 9 Newsletter*, 13, 12–15. doi:10.1044/hhdc13.1.12
- Brandy, W. (2002). Speech audiometry. In J. Katz (Ed.), *Handbook of Clinical Audiology*, 5th ed (pp. 96–110). Baltimore, MD: Lippincott Williams and Wilkins.
- DeBow, A., & 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.
- Ferrer, O. (1960). Speech audiometry: A discrimination test for Spanish language. *Laryngoscope*, 70, 1541–1551.
- Fletcher, H. (1950). A method of calculating hearing loss for speech from an audiogram. *Acta Oto-laryngologica Supplementum*, 90, 26–37.
- Hällgren, M., Larsby, B., & Arlinger, S. (2006). A Swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition. *International Journal of Audiology*, 45, 227–237.
- Harris, R.W., Nissen, S.L., Pola, M.G., McPherson, D.L., Tavartkiladze, G.A., & Eggett, D.L. (2007). Psychometrically equivalent Russian speech audiometry materials by male and female talkers. *International Journal of Audiology*, 46, 47–66.
- Michael, E., & Gollan, T.H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J.F. Kroll & A.M.B. de Groot (Eds.), *The handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). New York: Oxford University Press.
- Nelson D.A., & Chaiklin J.B. (1970). Writedown versus talkback scoring and scoring bias in speech discrimination testing. *Journal of Speech and Hearing Research*, 13, 645–654.
- Nissen, S.L., Harris R.W., & Slade, K.B. (2007). Development of speech reception threshold materials for speakers of Taiwan Mandarin. *International Journal of Audiology*, 46, 449–458.
- Ramkissoon, I., & Khan, F. (2003). Serving multilingual clients with hearing loss: How linguistic diversity affects audiologic management. *ASHA Leader*, 8 (3), 1, 10–11, 27. Retrieved from <http://www.asha.org/Publications/leader/2003/030218/030218a.htm>
- Ramkissoon, I., Proctor, A., Lansing, C.R., & Bilger, R.C. (2002). Digit speech recognition thresholds for non-native speakers of English. *American Journal of Audiology*, 11, 23–28.
- Statistics Canada (2006). Population by Mother Tongue: 2006 Census. Retrieved from <http://www40.statcan.gc.ca/l01/cst01/demo12g-eng.htm>
- Studebaker, G. A. (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 455–462.
- Vaillancourt, V., Laroche, C., Mayer, C., Basque, C., Nali, M., Eriks-Brophy, A., Soli, S.D., & Giguère, C. (2005). Adaptation of the HINT (hearing in noise test) for adult Canadian Francophone populations. *International Journal of Audiology*, 44, 358–69.
- Von Hapsburg, D., & Pena, E.D. (2002). Understanding bilingualism and its impact on speech audiometry. *Journal of Speech, Language and Hearing Research*, 45, 202–213.
- Wagener, K., Josvassen, J.L., & Ardenkjær, R. (2003). Design, optimization and evaluation of a Danish sentence test in noise. *International Journal of Audiology*, 42, 10–17.
- Wong, L.L., Ho, A.H., Chua, E.W., & Soli, S.D. (2007). Development of the Cantonese speech intelligibility index. *Journal of the Acoustical Society of America*, 121, 2350–2361.
- Wong, L.L., & Soli, S.D. (2005). Development of the Cantonese Hearing In Noise Test (CHINT). *Ear and Hearing*, 26, 276–289.
- Zubick, H.H., Irizarry, L.M., Rosen, L., Feudo, P., Kelly, J.H., & Strome, M. (1983). Development of speech-audiometric materials for native Spanish-speaking adults. *Audiology*, 22, 88–102.

AUTHOR'S NOTE

Correspondence should be sent to Stefka H. Marinova-Todd, Ed.D., School of Audiology and Speech Sciences, University of British Columbia, 2177 Wesbrook Mall Vancouver, B.C., V6T 1Z3, Canada. E-mail: stefka@audiospeech.ubc.ca ▶

Received date : February 01, 2010

Accepted date : August 10, 2010

KEY WORDS

SWALLOWING

DYSPHAGIA

DEGLUTITION

SCREENING

VIDEOFLUOROSCOPY

ASPIRATION

- **Exploration of the utility of a brief swallow screening protocol with comparison to concurrent videofluoroscopy**
- **Exploration de l'utilité d'un bref protocole de dépistage des troubles de déglutition en comparaison avec une vidéofluoroscopie simultanée**

Catriona M. Steele
 Sonja M. Molfenter
 Gemma L. Bailey
 Rebecca Cliffe Polacco
 Ashley A. Waito
 Dana C. B. H. Zoratto
 Tom Chau

Abstract

This study involved a direct blinded comparison between the results of a brief, standardized swallowing screening protocol and videofluoroscopy of exactly the same swallows. Forty adults participated. Each participant completed a brief swallow screening protocol involving tongue lateralization, voluntary cough, a voice task, and 2 swallowing tasks (3 swallows of 5cc thin liquid barium suspension and a cup-drinking task). We collected time-linked radiographic data and a high-definition movie of the participant's face, head and neck. The movie data were rated by 7 blinded clinicians (nurses and speech-language pathologists) for evidence of clinical signs associated with dysphagia. The videofluoroscopy data were rated by a separate panel of blinded speech-language pathologists for evidence of penetration-aspiration and post-swallow pharyngeal residues. Predictive statistics were calculated for the movie rating results, compared to the videofluoroscopy results. The results showed that none of the screening questions met our criteria for adequate predictive power: sensitivity, specificity and negative predictive values > 0.6, a false negative rate < 0.2 and a positive likelihood ratio > 1.0. We conclude that swallow screening decisions based on a series of 3-4 thin liquid swallows do not have good clinical utility for detecting dysphagia or penetration-aspiration. We discuss a number of issues in swallow screening research that may have contributed to the difference in these results compared to other studies.

Abrégé

Cette étude portait sur la comparaison à l'aveugle des résultats d'un bref protocole normalisé de dépistage des troubles de déglutition et d'une vidéofluoroscopie des mêmes déglutitions. Quarante adultes y ont participé. Chaque participant a effectué un bref protocole de dépistage des troubles de déglutition, c'est-à-dire une latéralisation de la langue, une toux volontaire, un exercice de voix et 2 exercices de déglutition (trois gorgées barytées de 5cc et un test du verre d'eau). Nous avons recueilli les données radiographiques en ordre chronologique et une vidéo en haute définition du visage, de la tête et du cou des participants. Les données de la vidéo ont été notées à l'aveugle par sept cliniciens (infirmières et orthophonistes) afin de trouver des signes cliniques liés à la dysphagie. Les données de la vidéofluoroscopie ont été notées à l'aveugle par un panel séparé d'orthophonistes afin de trouver des signes de pénétration et d'aspiration de résidus ainsi que des résidus pharyngés post déglutition. Des statistiques de prédiction ont été calculées pour le classement des résultats de la vidéo afin de les comparer à ceux de la vidéofluoroscopie. Les résultats ont démontré qu'aucune des questions du dépistage ne répondait à nos critères de prévisibilité : les valeurs prédictives de la sensibilité, la spécificité et la négativité > 0,6, un taux de faux négatif < 0,2 et un ratio positif de possibilité > 1,0. Nous avons conclu que les résultats d'un dépistage des troubles de déglutition basés sur une série de 3-4 gorgées de liquide clair ne sont pas utiles cliniquement pour détecter la dysphagie ou les troubles de pénétration-aspiration. Nous discutons d'un certain nombre de questions liées à la recherche en dépistage de trouble de déglutition qui ont peut-être contribué à faire la différence entre ces résultats et ceux des autres études.

Catriona M. Steele, Ph.D.,
 CCC-S-LP, BRS-S,
 ASHA Fellow^{1,2,3,4}
 Sonja M. Molfenter,
 M.H.Sc.¹
 Gemma L. Bailey,
 M.H.Sc.¹
 Rebecca Cliffe Polacco,
 M.H.Sc.¹
 Ashley A. Waito, B.A.¹
 Dana C. B.H. Zoratto,
 M.H.Sc.^{1,3}
 Tom Chau, Ph.D.^{3,4}

1. Toronto Rehabilitation Institute, Toronto, ON, Canada
2. Department of Speech-Language Pathology, University of Toronto, Toronto, ON, Canada
3. Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada
4. Bloorview Research Institute, Holland Bloorview Kids Rehab, Toronto, ON, Canada

The early identification of dysphagia and aspiration risk through swallow screening has been recognized as best practice in many guidelines, particularly those applying to the management of stroke (e.g., Canadian Stroke Network, 2005; Joint Commission for the Accreditation of Healthcare Organizations, 2004; Scottish Intercollegiate Guidelines Network, 2004). According to the World Health Organization definition, “Screening tests sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not intended to be diagnostic” (CCI Conference on preventive aspects of chronic disease, 1951). A screening does not provide sufficient information to support management decisions for those who fail the test by showing evidence of the clinical sign in question; rather, “Persons with positive or suspicious findings must be referred... for diagnosis and necessary treatment” (CCI Conference on preventive aspects of chronic disease, 1951).

Swallow screenings are supposed to be simple tests that can be administered by a variety of trained healthcare professionals. In the ideal world, a dysphagia screening should provide a quick and accurate indication of a patient’s risk of aspiration, and the likelihood that they

have dysphagia. Existing guidelines fail to clearly define the content and procedures required for valid and reliable swallow screening. For example, the guideline published by the Joint Commission for the Accreditation of Healthcare Organizations (2004), which was removed from their accreditation standards in 2009, left room for a screening protocol to range from very quick and simple tests (Suiter & Leder, 2008) right up to full clinical bedside swallow examinations (Logemann, Veis & Colangelo, 1999).

One response to the mandate to provide swallow screening has been for speech-language pathologists to design protocols for use within their local facilities. A search on the internet, using www.google.ca and the search term “dysphagia screening tool” leads to more than 29,000 results and numerous examples of such tools (e.g., Grey-Bruce Health Network Dysphagia Screening Tool, 2008; Iredell Dysphagia Screen, 2006; Lothian Dysphagia Screening Test, 2005; Oklahoma Dysphagia Screening Tool, n.d.; St. George Dysphagia Screening Tool, n.d.). Researchers have also responded to this mandate by designing and testing different screening protocols. Many of these protocols overlap in their core elements, as shown in Table 1.

Table 1
Comparison of protocol components in previously published swallow screening tools.

Test	Diagnostic Criteria	Readiness/Alertness Criteria	Secretion Management/Drooling	Respiratory Rate	Facial Muscle or Oral Motor Testing	Sensory Testing	Baseline Voice Quality Appraisal	Baseline Cough Appraisal	Water Swallows	Post-Swallow Cough or Voice Observation	Other
Standardized Swallow Assessment (Perry, 2001)	Stroke	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Speech criteria
Massey Bedside Swallowing Screen (Massey & Jedlicka, 2002)	Stroke	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Speech criteria
VAMC Nursing Admission Dysphagia Screening Tool (Bravata et al., 2009)	Acute Ischemic Stroke	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Patient or family report of dysphagia; speech criteria
Royal Brisbane and Women’s Hospital Dysphagia Screening Tool (Cichero, Heaton & Bassett, 2009)	Stroke	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Patient or family report of dysphagia; speech criteria
TOR-BSST© (Martino et al., 2009a)	Stroke	Implied	No	No	Yes	No	Yes	No	Yes	Yes	

Volume-Viscosity Screening Test (Clave et al., 2008)	Risk for dysphagia	Implied	Yes	Yes	Yes	No	Yes	No	(Yes)	Yes	Pulse Oximetry ($\geq 3\%$ drop); begins with nectar at controlled volumes and proceeds based on tolerance
Daniels Swallow Screen (Daniels et al., 1998)	Stroke	Implied	No	No	No	Yes	Yes	Yes	Yes	Yes	Speech criteria
MGH-SST (Cohen, 2008)	Neuroscience admissions	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
3-oz Water Swallow Test (Suiter & Leder, 2008)	No	Implied	No	No	No	No	No	No	Yes	Yes	Failure to complete drinking of 3-oz without stopping
Gugging Swallow Screen (Trapl et al., 2007)	Stroke	Yes	Yes	No	No	No	Yes	Yes	(Yes)	Yes	Begins with saliva swallow followed by semisolid before proceeding to liquid
ASSIST - Acute Swallow Screen in Stroke and TIA (2009)	Stroke	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Speech criteria

Almost all reported swallow screening protocols involve water swallowing, but the precise volume and method of administering water differs across tests. Some tests involve between 3 and 10 repeated sips of small volumes of water (Cohen, 2008, 2009; Martino et al., 2009a). In some cases, initial sips of water are followed by cup drinking (Cohen, 2008; Cichero, Heaton & Bassett, 2009). Other protocols start with small, controlled volumes and increase volume gradually (Clave et al., 2008). Some tests involve the rapid and continuous drinking of a large volume of water (Suiter & Leder, 2008) while others begin with thickened water and progress to thin liquids only when thicker items are tolerated without evidence of difficulty (Clave et al., 2008). Regardless of the specific procedures for administering water, all of these tests look for similar signs of difficulty on water swallowing tasks: a) difficulty completing the task; b) coughing; c) change in voice quality post swallow (specifically vocal wetness); or d) respiratory difficulty.

The predictive power of a screening test is most easily understood by using a two by two contingency table, plotting the test result against the presence or absence of the target problems, which, in the case of swallowing, are aspiration or dysphagia (Sackett, Straus, Richardson,

Rosenberg & Haynes, 2000). This allows the calculation of indices such as sensitivity (the proportion of people with the underlying problem who have a positive test result, i.e., they fail the test), specificity (the proportion of those who do NOT have the underlying problem who have a negative test result, i.e. they pass the test), and negative predictive value (the proportion of those who pass the test who do not have the underlying problem). In order for a test to have clinical utility, it should score well on all 3 of these indices. It is also desirable for a test to have a low false-negative rate, so that true cases of the target underlying problem are not missed. A full discussion of these measures can be found in McCullough et al. (2005). Likelihood ratios (which are not susceptible to sample prevalence bias when this does not match prevalence in the broader population) are another important index to include when exploring the utility of a screening measure (Schoenfeld, 2009). Likelihood ratios compare the proportions of patients with and without the disease who have been given the diagnostic test and divide the true-positive rate by the false-positive rate (i.e., sensitivity/1-specificity). Thus, the likelihood ratio represents the probability that a given diagnostic test result would be expected in a patient who has the underlying target disorder (Scherokman, 1997).

Research on swallow screening has attempted to validate specific approaches and demonstrate the predictive power of screening tests (as a whole) and of particular component items, for detecting penetration-aspiration and dysphagia. Several systematic reviews concur that no single specific screening approach or procedural element has adequate predictive power for detecting dysphagia or aspiration (Bours, Speyer, Lemmens, Limburg & de Wit, 2009; Martino, Pron & Diamant, 2000; Perry & Love, 2001). Three kinds of screening validation studies can be found in the literature: a) comparisons against clinical bedside examinations; b) comparisons against flexible endoscopic examinations of swallowing; and c) comparisons against videofluoroscopy (VF). When critically reading the literature on swallow screening performance, clinicians need to ask several questions:

1. Was there any bias or skew in the sample in which the screening test was studied?
2. Are the items in the protocol (individually and combined) logical, reasonable and valid measures for detecting the underlying problem?
3. How much variation was encountered in screening (or validation procedure) results across raters?
4. Were there any differences in the number and types of swallows that were compared from the screening to the validation procedure, and how might this affect the conclusions?
5. What was the delay between the screening and the validation procedure, and is this a potential concern?
6. Were the individuals who judged the screening blinded to the validation procedure (true status) results, and *vice versa*?

Within the literature, some swallow screening studies have been performed exclusively in stroke patients while others have been performed in heterogeneous populations. Similarly, some studies have recruited all incoming stroke patients while others have used convenience samples of referred individuals. Sample-related considerations impact the denominators in sensitivity and specificity calculations, which represent the true occurrence of an abnormal or normal swallowing status. When all other elements are kept equal, studies of people with heterogeneous etiologies (compared to etiologically focused samples), or of all incoming patients within a focused target group (compared to referred samples), are prone to recording lower estimates of the true incidence of an underlying problem, which is likely to increase sensitivity and lower specificity.

Conversely, in studies of more focused etiological groups, or of referred samples, the estimated incidence of the underlying problem is likely to be higher than that seen in the broader population, leading to lower sensitivity and increased specificity.

Methodological issues can also impact the reported sensitivity and specificity of a screening tool. For example, the type and number of items included in a screening protocol can influence the power of the test for detecting an underlying problem. Martino, Streiner, Maki & Diamant (2009b) showed that sensitivity for detecting dysphagia improved as the number of water swallows in a screening protocol increased. The chances of a binary verdict that the “problem exists” being correct increases with the number of opportunities that the patient has to demonstrate that problem. Similarly, a recent endoscopic study found that patients who aspirate silently on smaller volumes of liquid are more likely to demonstrate an overt sign of aspiration if a greater volume or number of swallows is obtained (Leder, Suiter and Green, 2010).

Table 2 summarizes the predictive power, validation methods, sample characteristics and use of blinding from a selection of swallow screening test validation studies reported in the literature. With the exception of the Veterans Affairs Medical Center Nursing Admission Dysphagia Screening Tool (Bravata et al., 2009), reported sensitivities are generally quite high across screening studies. On the other hand, specificities are generally poor, even when blinding has been used. This suggests a general trend towards over-identifying dysphagia through swallow screening. While it may be argued that it is preferable to over-identify, rather than under-identify, a health condition that is associated with negative outcomes and health care costs, it may also be argued that over-identification involves unnecessary health-care expenditures, as well as negative quality of life consequences when interventions like diet texture restriction are unnecessarily or overzealously applied.

OBJECTIVES

The purpose of our study was to conduct a direct blinded comparison between the results of a brief, standardized swallowing screening protocol and VF of exactly the same swallows, thereby removing the contributions of time lag and test circumstances to differences in test results. Our hypothesis was that clinically observed signs of swallowing difficulty (i.e. failure of specific questions on the swallow screening observation form) would be associated with the occurrence of physiologically abnormal pharyngeal phase swallowing on the videofluoroscopy, and that this relationship would demonstrate good clinical utility

Table 2**Comparison of the methodology and results of previous swallow screening tool validation studies.**

<u>Test</u>	<u>Validation</u>	<u>Population</u>	<u>Sensitivity</u> <u>%</u>	<u>Specificity</u> <u>%</u>	<u>Negative</u> <u>Predictive</u> <u>Value</u> <u>%</u>	<u>Positive</u> <u>Likelihood</u> <u>Ratio</u>	<u>Blinding?</u>
Standardized Swallow Assessment (Perry, 2001)	Chart-documented evidence of dysphagia	Stroke	97	90	Not reported	9.70	Not reported
Massey Bedside Swallowing Screen (Massey & Jedlicka, 2002)	Chart-documented evidence of dysphagia	Stroke	100	100	Not reported	N/A	Not reported
VAMC Nursing Admission Dysphagia Screening Tool (Bravata et al., 2009)	S-LP evaluation of swallowing	Stroke	29	84	68	1.81	Not reported
Royal Brisbane and Women's Hospital Dysphagia Screening Tool (Cichero, Heaton & Bassett, 2009)	S-LP Clinical Swallow Examination and Chart Review	Stroke	95	97	98	31.6	No
TOR-BSST© (Martino et al., 2009a)	VFSS confirmation of dysphagia using P-A Scale and MASA dysphagia subscore	Acute stroke	96	64	93	2.60	Yes
Volume-Viscosity Screening Test (Clave et al., 2008)	VFSS confirmation of aspiration and other abnormal swallowing parameters	Heterogeneous	100	29	Not reported	1.40	Yes
Daniels Swallow Screen (Daniels et al., 1998)	VFSS confirmed aspiration	Acute stroke	92	66	Not reported	4.46	Yes
MGH-SST (Cohen, 2008)	FEES confirmation of dysphagia and/or penetration-aspiration	Neuroscience admissions	89	61	87	2.28	Yes
3-oz Water Swallow Test (Suiter & Leder, 2008)	FEES immediately beforehand	Heterogeneous	96	46	98	1.80	No
Gugging Swallow Screen (Trapl et al., 2007)	FEES measures of aspiration (P-A scale ≥ 5)	Stroke	100	50-69	100	3.23	Yes

Abbreviations: FEES = S-LP = Speech-Language Pathologist; VFSS = Videofluoroscopic Swallowing Study; MASA = Mann Assessment of Swallowing Ability; VAMC = Veterans Affairs Medical Center; MGH-SST = Massachusetts General Hospital Swallow Screening Tool; Flexible Endoscopic Examination of Swallowing; P-A = Penetration-Aspiration.

of the swallow screening process. We were interested in evaluating the strength of specific clinical signs for detecting underlying problems, and in comparing the clinical utility of screening judgments made by nurses (RNs) with those made by speech-language pathologists (S-LPs). We expected that these would not differ significantly.

METHODS

Participants

Data were collected from a gender-balanced convenience sample of 40 consenting adults (mean age: 67 years), referred for VF at one of two hospitals. Ethical considerations regarding radiation risk led us to use this convenience sample, rather than recruiting individuals from the broader population. Etiologies were mixed, and included inpatients as well as outpatients referred by community physicians for initial investigation of swallowing complaints. Our questions were not specific to any particular diagnostic group. Therefore, medical diagnostic information was not captured, other than to confirm the absence of a history of tracheostomy, head and neck cancer, and any surgery to the head and neck other than routine tonsillectomy or adenoidectomy. The study received human subjects approval from the institutional review boards of the participating hospitals.

Screening Protocol

We selected 5 tasks for inclusion in a brief swallow screening protocol, based on a review of the swallow screening literature (see Table 1). The protocol most closely resembled part 2 of the Massachusetts General Hospital Swallow Screening Test (MGH-SST; Cohen, 2008, 2009), and included the following steps:

- a. a tongue lateralization task;
- b. a baseline voluntary cough task;
- c. a baseline phonation task;
- d. a sequence of three single-sip swallow tasks (5 cc per sip) with thin liquid, with each sip followed by a repetition of the phonation task;
- e. a cup-drinking task with thin liquid, again followed by a repetition of the phonation task.

A readiness-for-testing component, such as that detailed in part 1 of the MGH-SST, was not specifically included. Given the research requirement that all participants be able to consent to the study, it can reasonably be assumed that all of our participants would have received a passing score on such a component. The selected phonation task was the utterance “ha-ha-ha-ha”, as used in the MGH-SST protocol. We excluded the evaluation of pharyngeal sensation based on evidence that this component does not add useful information to

screening outcomes (Martino et al., 2009a). Due to the fact that the collection of concurrent VF data involved radiation exposure, the cup-drinking task required the patient to “take several continuous sips” from a cup. This was not a complete 3 oz liquid swallowing challenge, as used in some screening protocols (Cohen, 2008, 2009; Suiter and Leder, 2008; Cichero, Heaton and Bassett, 2009), but is consistent with usual videofluoroscopy procedures (Martin-Harris et al., 2008; Logemann, 1993).

Data Collection Procedures

The experiment was set up in the videofluoroscopy suite, using a high-definition camcorder, positioned to capture an image of the patient’s face, head and neck, and a high quality stage microphone to capture sound. A dual-axis accelerometer (Analog Devices, ADXL 322) was also placed on the patient’s neck in midline over the cricoid cartilage. Time-linked data collection from all channels was controlled by LabVIEW software (National Instruments, Toronto, Canada,). The VF was captured through the hospital’s fluoroscopy equipment at 30 frames per second. Figure 1 shows an example of time-linked movie and VF images for a participant during a thin liquid swallow. The acoustic (microphone) and accelerometry data have been reported elsewhere (Waito, Bailey, Molfenter, Zoratto & Steele, 2010; Zoratto, Chau & Steele, 2010) and were not included in the specific analyses for this manuscript.



Figure 1. Time-linked screen captures from the movie and videofluoroscopic recording channels, showing a participant swallowing thin liquid barium.

The data collection protocol proceeded as follows:

1. A movie recording was taken while the participant performed the baseline tongue lateralization, volitional cough, and phonation tasks.
2. Time-linked movie and VF recordings were collected during the swallowing of three 5-cc volumes of thin liquid barium suspension (40% w/v), with each swallow followed by repetition of the phonation task. These boluses were administered by teaspoon and a command swallow paradigm was used.

3. A thin liquid barium suspension cup-drinking task (i.e. 3-5 consecutive sips of unrestricted volume) was recorded in the time-linked movie and fluoroscopy channels, again followed by a phonation sample.

The videofluoroscopy then continued for clinical investigative purposes, but these additional swallows were not part of this research study.

Data Processing and Rating

The steps involved in data processing and rating for the screening movies are illustrated in Figure 2. The movies were organized in clips capturing the entire swallow screening sequence, and randomized for rating. Title screens were added to introduce each task in the screening protocol (e.g. “tongue lateralization”, “single sip #1”). An algorithm written in MATLAB 2010a (Mathworks, Natick, MA, USA) was used to remove low amplitude background noise from the audio channel of all the movie recordings by filtering out data below 10% of the maximal amplitude found in each signal. This removed contamination from any background conversation and enabled the rater to focus clearly on the voice samples produced by the participant when rating.

The screening movies were rated by 4 S-LPs, in their first year of practice, and 3 registered nurses from an acute care hospital. Using new graduate S-LPs conferred the advantage that none of the raters had any previous knowledge of the participating patients. The acute-care hospital from which the RN judges were drawn was not using a specific swallow screening protocol at the time of the study. We intentionally did not provide extensive didactic training in the identification of the clinical signs of interest because we wanted to determine the utility of the swallow screening tool in the hands of health-care professionals who should, by virtue of their professional knowledge and skills, be able to focus on the specific questions raised, without additional training. Training in the scoring of each sign was conducted with 3 cases not included in the experimental dataset. This allowed us to clarify any questions that arose, and to make sure that all raters were attuned to the signs of interest.

After watching the entire screening movie for a particular participant, the rater was asked to record forced-choice judgments (normal, abnormal, uncertain) regarding the different clinical signs of interest (tongue lateralization; baseline voluntary cough; baseline voice quality; post-swallow spontaneous coughing or change in voice-quality; and overall pass-fail, with fail indicating a need for further swallowing assessment). We asked the raters to give a single rating for each question, across the entire movie sequence for each participant, mirroring the conventional expectation that swallow screenings

lead to a single pass-fail result. We took the conservative position of classifying scores of “uncertain” as reflecting the presence of an abnormal clinical sign (i.e., fail) for the purposes of further analysis. This was motivated by the assumption that any question of abnormality should be followed up in the context of usual swallow screening in the clinical setting. The percent agreement across all raters within each professional group and across all 7 raters was calculated for each question for each movie. In cases where there was no clear majority consensus in the screening result for a particular question within a professional group, repeated rating by two of the original raters was conducted.

In addition to providing forced choice answers on the five primary screening questions, raters were also asked to report observations of “any other signs of difficulty” and were given a comments box in which they could elaborate on these observations.

The VF ratings were treated in a similar manner, with the important distinction that the recordings were spliced into clips, randomized and rated at the level of the individual swallowing task (single 5cc sip; cup drinking task). This decision was motivated by the desire to collect an accurate gold-standard answer regarding the presence or absence of dysphagia and aspiration, which would not be biased by knowledge regarding previous swallows

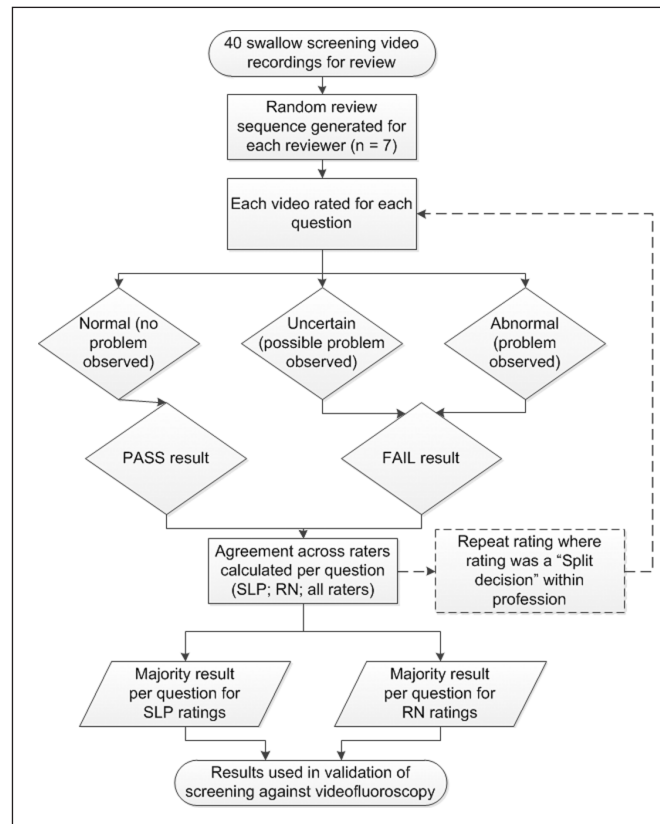


Figure 2. Flow-chart illustrating the data processing and rating steps for determining swallow screening results based on the movies of the swallow screening sequence.

during rating. For 3 participants, video quality concerns (such as shoulder interference or a premature turn-off of the fluoroscopy stream) resulted in the availability of only 1 single sip recording. In a further 9 cases, only 2 single sip clips were of adequate quality to permit rating. Consequently, the VF data set included a total of 106 X 5cc single sip swallows. Cup drinking sequences were available for 36 participants.

The VF rating process is illustrated in Figure 3. A panel of 7 S-LPs (separate from those involved in the movie ratings) completed a 6-hour training session in the rating procedures prior to the study. The individual swallow clips were rated using the 8-point Penetration-Aspiration Scale (Rosenbek, Robbins, Roecker, Coyle & Wood, 1996) and Eisenhuber's ordinal scales capturing residues in the valleculae and pyriform sinuses, with each residue location scored separately (Eisenhuber et al., 2005). The rating assignment (a total of 1275 feature ratings) was randomized across raters, and was organized in batches (10 swallow clips per batch), for which a rater was required to rate a single feature (aspiration or residue). The final data set included 3 ratings per feature for each swallow. These ratings were reviewed for concordance across raters. For 6 clips, consensus was not established in the initial ratings of penetration-aspiration, while 3 and 5 clips, respectively, showed a lack of consensus for vallecular and pyriform sinus residue ratings. These clips were entered into a new rating set, which was reviewed in a live consensus session attended by two of the original raters and a research assistant facilitator. Each clip was played 3 times and the raters recorded scores independently. Scores were then declared. In the case that these new scores differed, the clip was reviewed carefully and discussed until an agreed score was reached.

Once the VF ratings were complete, binary penetration-aspiration and cumulative residue disposition scores were derived for each participant across all 4 swallowing tasks in the screening sequence. In this way, the resolution of the binary screening and VF verdicts was equalized. Any occurrence of penetration-aspiration scale scores ≥ 3 (i.e. material entering the supraglottic space without subsequent ejection) resulted in a disposition classification of "penetration-aspiration present". A cumulative residue score (i.e. vallecular residue score plus pyriform sinus residue score) ≥ 2 (representing either 2 or more occurrences of mild residue or at least one occurrence of moderate residue) resulted in a disposition classification of "residue present". A classification of "dysphagia present" was assigned in the case of either "penetration-aspiration present" and/or "residue present". As such, the threshold for classifying a patient as truly having dysphagia was low, with the exception that a single episode of mild

residue was considered insufficient to classify a person as having dysphagia.

ANALYSIS

We calculated the predictive value of each abnormal screening test result (from the screening movie ratings) for dysphagia and penetration-aspiration on VFSS, looking separately at the results for the RN and S-LP raters. Two-by-two contingency tables were prepared for each of the five movie rating results, compared to the videofluoroscopic dispositions of aspiration-present and dysphagia-present. Data were analyzed at the level of the individual patient (across the series of thin liquid swallows and non-swallowing tasks collected during the screening protocol), rather than on a swallow-by-swallow basis. The criterion of considering a particular result to have adequate predictive utility was defined *a priori* as the combination of sensitivity, specificity and negative predictive value scores > 0.6 , a false negative rate < 0.2 and a positive likelihood ratio > 1.0 .

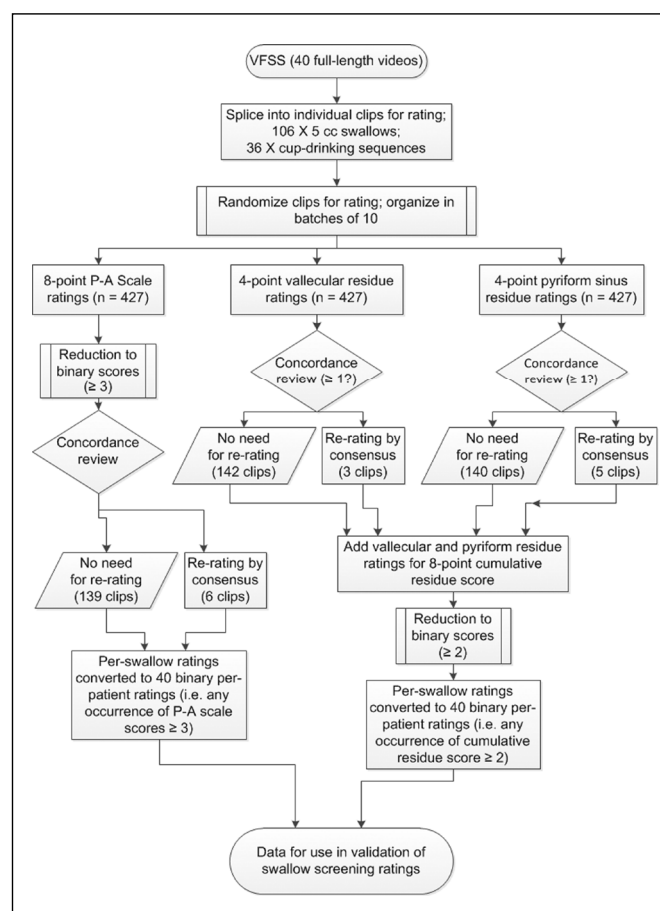


Figure 3. Flow-chart illustrating the data processing and rating steps for determining true aspiration and dysphagia status based on blinded review of the videofluoroscopies.

RESULTS

Descriptive statistics for the frequency (percent occurrence) of abnormal findings in the movie ratings are shown by rater group (S-LPs, RNs, combined) in Table 3. In Table 4, the intra-class correlations for inter-rater agreement in the binary disposition scores for each question are shown by rater group. Table 4 also shows the

agreement between professions for the final disposition scores assigned to each participant, considering all raters within each professional group. This agreement was generally strong, as indexed by Cohen's Kappa scores of $\kappa > 0.6$. The kappa score for ratings of baseline voice quality fell in the moderate agreement range ($\kappa = 0.4$ to 0.6 ; Capozzoli, McSweeney & Sinha, 1999).

Table 3

The frequencies of unanimous and partial-agreement in swallow screening ratings by speech-language pathologist and nurse judges during the rating of abnormal clinical signs from movies of patients performing a brief swallow screening protocol.

<u>Screening Question</u>	<u>Split decision (50% agreement)</u>	<u>S-LPs</u>		<u>RNs</u>		<u>Pooled across all 7 raters</u>			
		<u>75% agreement</u>	<u>100% agreement</u>	<u>67% agreement</u>	<u>100% agreement</u>	<u>57% agreement</u>	<u>71% agreement</u>	<u>86% agreement</u>	<u>100% agreement</u>
Tongue lateralization	6%	22%	72%	33%	67%	13%	10%	30%	48%
Baseline voluntary cough	15%	30%	55%	48%	53%	25%	13%	33%	30%
Baseline voice quality	20%	29%	51%	58%	42%	25%	20%	30%	25%
Post swallow cough, throat clear or voice change	3%	22%	75%	25%	75%	10%	13%	33%	45%
Overall pass-fail result	13%	20%	67%	45%	55%	20%	25%	25%	30%

Table 4

Intra-class correlation and Cohen's Kappa statistics for inter-rater agreement (within, across and between professional groups) during the rating of abnormal clinical signs from movies of patients performing a brief swallow screening protocol.

Screening Question	Intra-class Correlation (S-LPs)			Intra-class Correlation (RNs)			Intra-class Correlation (Combined)			Cohen's Kappa for Inter-profession Agreement on Binary Dispositions
	-	95% Confidence Interval		-	95% Confidence Interval		-	95% Confidence Interval		
	Mean	Lower Boundary	Upper Boundary	Mean	Lower Boundary	Upper Boundary	Mean	Lower Boundary	Upper Boundary	
Tongue lateralization	0.84	0.73	0.91	0.53	0.21	0.74	0.84	0.75	0.91	0.66
Baseline voluntary cough	0.79	0.66	0.89	0.74	0.57	0.86	0.88	0.81	0.93	0.79
Baseline voice quality	0.78	0.64	0.88	0.45	0.07	0.69	0.81	0.70	0.89	0.43
Post swallow cough, throat clear or voice change	0.90	0.83	0.94	0.86	0.76	0.92	0.91	0.86	0.95	0.65
Overall pass-fail result	0.73	0.55	0.85	0.69	0.48	0.83	0.84	0.74	0.91	0.64

In total, 14 participants were determined to show penetration-aspiration scores of 3 or higher on at least one of the swallowing tasks in the screening protocol. Dysphagia was judged to be present in 24 participants. The results of the cross-tabulation of the screening movie and VF disposition scores are shown by profession in Table 5. None of the test results in this study met the *a priori* criterion that we had specified for adequate predictive power (the combination of sensitivity, specificity and negative predictive value scores > 0.6 , a false negative rate < 0.2 and a positive likelihood ratio > 1.0).

The diagnostic utility of the “other signs of difficulty” question was not specifically tested in a cross-tabulation, due to the fact that the question was open ended.

However, the items reported under this parameter were interesting: 13 patients were described to exhibit delayed swallow or prolonged transit times; 14 were described to need multiple swallows per bolus; 7 were described to exhibit bolus control problems with associated anterior spill or drooling. On a percentage basis, there were no obvious differences in the frequencies of specific types of comments (delay; multiple swallows, bolus control) between the S-LP and RN raters. Other comments, describing 1-3 patients per condition, included observations regarding breathing, behaviors (grimacing, impulsivity, obvious effort to swallow) and visible facial paresis. Figure 1 shows one such example, in which facial asymmetry was noticeable in the screening movie.

Table 5

The predictive utility of abnormal clinical signs for detecting aspiration and dysphagia, when observed by speech-language pathologists and nurses for patients performing a brief swallow screening protocol, and validated against concurrent videofluoroscopy.

Validation	Rater Group	Measure	Sensitivity	Specificity	Negative Predictive Value	Positive Predictive Value	False Positive Rate	False Negative Rate	Positive Likelihood Ratio
Penetration-Aspiration on VFSS	S-LP	Abnormal Baseline Tongue Lateralization	14%	72%	60%	22%	78%	40%	0.51
		Abnormal Baseline Voluntary Cough	36%	60%	63%	33%	67%	38%	0.89
		Abnormal Baseline Voice Quality	57%	56%	70%	42%	58%	30%	1.30
		Post Swallow Cough, Throat Clear or Voice Change	21%	52%	54%	20%	80%	46%	0.45
		Overall Screen Result of Fail	64%	16%	44%	30%	70%	56%	0.77
	RN	Abnormal Baseline Tongue Lateralization	7%	80%	61%	17%	83%	39%	0.36
		Abnormal Baseline Voluntary Cough	36%	80%	69%	50%	50%	31%	1.79
		Abnormal Baseline Voice Quality	50%	64%	70%	44%	56%	30%	1.39
		Post Swallow Cough, Throat Clear or Voice Change	50%	52%	65%	37%	63%	35%	1.04
		Overall Screen Result of Fail	57%	44%	65%	36%	64%	35%	1.02
Dysphagia on VFSS	S-LP	Abnormal Baseline Tongue Lateralization	21%	73%	37%	56%	44%	63%	0.78
		Abnormal Baseline Voluntary Cough	42%	67%	42%	67%	33%	58%	1.25
		Abnormal Baseline Voice Quality	58%	67%	50%	74%	26%	50%	1.75
		Post Swallow Cough, Throat Clear or Voice Change	38%	60%	38%	60%	40%	63%	0.94
		Overall Screen Result of Fail	71%	13%	22%	57%	43%	78%	0.82
	RN	Abnormal Baseline Tongue Lateralization	13%	80%	36%	50%	50%	64%	0.63
		Abnormal Baseline Voluntary Cough	25%	73%	38%	60%	40%	62%	0.94
		Abnormal Baseline Voice Quality	54%	80%	52%	81%	19%	48%	2.71
		Post Swallow Cough, Throat Clear or Voice Change	46%	47%	35%	58%	42%	65%	0.86
		Overall Screen Result of Fail	58%	47%	41%	64%	36%	59%	1.09

DISCUSSION

This study provides greater detail regarding the predictive power of individual clinical signs for detecting aspiration and dysphagia using a brief swallow screening protocol. Our results identified differences in the power of particular indicators, when judged by RNs compared to S-LPs. Overall, predictive power measures were poor compared to prior studies of longer screening protocols in which a separate (indirect) instrumental gold standard test has been used for validation (see Table 2). These results were not what we had expected and prompted us to consider a variety of explanations. The most obvious explanation is the fact that our study involved a direct comparison of clinical judgments and blinded VF ratings for exactly the same swallows. This is, to our knowledge, the first study to report such a direct comparison. In this section, we review the details of our study, its limitations, and several other possible explanations for differences between our study results and those of previous studies in the swallowing literature (see Table 2).

ITEM ANALYSIS

Of the five questions that were asked during the swallow screening movie rating, the observation of abnormal tongue lateralization had the lowest sensitivity, both for aspiration and for dysphagia detection. This clinical sign was not one that judges had difficulty agreeing on, either within or across professions, although the intra-class correlation amongst RN judges was only 0.53. When considered in isolation, the identification of abnormal tongue lateralization led to excessively high false positive rates for the identification of aspiration and chance-performance for identifying dysphagia. On this basis, we would argue that difficulty in lateralizing the tongue has questionable validity as a clinical sign of dysphagia and should not be interpreted as an indication of aspiration.

Abnormalities in the ability to produce a voluntary cough at baseline showed poor sensitivity for detecting aspiration or dysphagia according to the ratings of both professional groups, although intra-class correlations for inter-rater agreement were fairly strong. Nurse judgments of this sign showed good specificity (80%) and low false-positives (31%), leading to a positive likelihood ratio of 1.79 for aspiration. The speech-language pathology ratings of this sign showed modest specificity (67%), a low false-positive rate (33%) and a positive likelihood ratio of 1.25 for dysphagia. These findings suggest that this parameter may provide useful swallow screening data when combined with other indicators.

The validity of abnormal voice quality, either at baseline or post-swallow, has received extensive

discussion elsewhere in the literature (Groves-Wright, Boyce & Kelchner, 2010; Warms & Richards, 2000; Waito et al., 2010). In this study, S-LP judges showed better agreement (ICCs = 0.78) than RNs (0.45) for perceptual judgments of baseline voice quality. This is probably not surprising, given the focus on perceptual voice assessment in speech-language pathology training programs. When all raters were pooled as a single rating group, consensus on this indicator was poor. Nonetheless, specificities for aspiration were modest for both rater groups with positive likelihood ratios of 1.3-1.39. Interestingly, this clinical sign had the strongest performance, in terms of likelihood ratios, for detecting dysphagia. Specificities for this sign were reasonably strong, consistent with previous findings in the literature (Waito et al., 2010). When RNs determined that baseline voice quality was abnormal, the patient was 2.71 times more likely to have dysphagia, as per the blinded videofluoroscopy ratings. These results support the continued consideration of baseline voice quality in swallow screenings, noting, however, that perceptual judgments of this feature may suffer from high inter-rater variability.

The observation of a post-swallow cough, throat clear, or change in voice quality was agreed upon most easily by judges. However, sensitivity and specificity of this measure for aspiration were poor, and false positives were high, particularly in the S-LP ratings. This sign yielded essentially equivocal findings regarding the possibility of underlying dysphagia. These results suggest that observations of coughing and voice quality post-swallow do not lead to accurate impressions regarding the presence of aspiration or dysphagia.

Finally, we asked judges to give an overall impression rating regarding the pass-fail status of a patient on the entire swallow screening protocol. This was not a score that was derived based on the results of the other screening questions, but an overall rating indicating the judge's opinion about whether the patient would require a referral for more detailed swallowing assessment, if they were seen in a typical emergency-room screening context. Agreement on this item was reasonably good, although RNs displayed greater variation than S-LPs. This parameter showed poor specificity for aspiration and dysphagia based on the S-LP ratings, and equivocal results for the RNs. The false negative rates on this question are somewhat alarming, particularly with respect to identifying dysphagia. Interestingly, according to this overall impression question, patients were more likely to be incorrectly classified as not having dysphagia by a speech-language pathologist than by a nurse!

DIFFERENCES IN SCREENING PROTOCOLS

As mentioned previously, the specific tasks in our brief swallow screening protocol were chosen based on a review of the swallow screening literature, and most closely resemble part 2 of the protocol used at the Massachusetts General Hospital (MGH-SST; Cohen, 2008, 2009). Our study did not consider the relationship between the thin liquid swallows examined in the swallow screening sequence and swallows of both liquid and other stimuli that may have occurred over the course of the entire videofluoroscopy protocol. We acknowledge that the choice to perform our screening protocol in a time-linked manner in the videofluoroscopy suite imposed a constraint on the number of swallows elicited in the screening test. From a mathematical perspective, the reduced number of swallows used in our study to determine true aspiration and dysphagia status would be most likely to lead to lower sensitivity statistics, in comparison to studies in which all swallows from a subsequent instrumental exam have been used. The observed result seems to confirm this expectation.

SAMPLE CONSIDERATIONS

Previous studies in which blinded validation of swallow screening protocols has been reported have included validation samples ranging in size from 50 to 100 participants (Clave et al., 2008; Cohen, 2008, 2009; Daniels et al., 1998; Martino et al., 2009a; Trapl et al., 2007). Our sample size, with 40 participants, was slightly smaller, but comparable to the validation sample sizes of 50-70 in the studies by Martino et al. (2009a), Daniels et al. (1998), and Trapl et al. (2007), with the important qualification that we studied a heterogeneous sample rather than the stroke population evaluated in all three of these prior studies.

With the exception of likelihood ratio measures, it is probably not appropriate to compare predictive power statistics for a protocol tested in a specific group to one tested in a heterogeneous sample because the prevalence of the underlying disorder in the sampled population may differ (Schoenfeld, 2009). Although swallow screening protocols have been most rigorously tested in etiologically specific groups (most commonly stroke) in the literature (e.g. Daniels et al., 2008; Martino et al., 2009a; Trapl et al., 2007), it is fairly common practice to use swallow screening protocols in broader patient populations (e.g., Suiter and Leder, 2008). It was in this general context that we wanted to explore our questions regarding swallow screening utility. In retrospect, it is unfortunate that we did not collect additional diagnostic and etiological information regarding our participants, which might have permitted a subgroup analysis of those with stroke,

brain injury or neurologic disease compared to those with unknown or non-neurogenic etiologies. However, to do such a subgroup analysis with adequate statistical power would have required a much larger sample size. Heterogeneity in the etiology of our participants may have led to increased sensitivities and lower specificities than studies of more etiologically-specific samples. As a rule, sensitivity results in our study were considerably lower than those reported in studies limited to the stroke population (Martino et al., 2009a; Perry, 2001; Massey & Jedlicka, 2002; Cichero, Heaton & Bassett, 2009; Daniels, et al., 2008; Trapl, et al., 2007). On the other hand, our sample comprised individuals referred for the investigation of suspected swallowing disorders. This meets one of the key criteria outlined by Sackett and Haynes for validity of a test, namely "independent, blind comparison of test results with a reference standard among a consecutive series of patients suspected (but not known) to have the target disorder" (Sackett & Haynes, 2002). To use a referred sample like this in a test validation study risks inflating the estimate of the true occurrence of the target underlying disorder, leading to reduced sensitivity and heightened specificity compared to studies of all-comers. Both of the prior studies in which swallow screening has been evaluated in heterogeneous populations appear to have been done in referred samples (Clave et al., 2008; Suiter & Leder, 2008). The sensitivities (and sample sizes) in our study were lower than reported in these studies, but specificities were comparable.

DATA COLLECTION PROCEDURES

Although our method of collecting time-linked data permitted a direct comparison across screening and videofluoroscopic data, the use of movies rather than live observation may well have altered the information that was considered by raters for the screening procedure. Screening procedures are supposed to be simple and transparent. In order to minimize the possibility of variable screening decisions across clinicians, such decisions should be made based on clinical signs that are directly evaluated and explicitly considered, rather than other contextual information which may be apparent. These requirements lay behind our method of posing specific and direct questions during the movie rating portion of our study. Nonetheless, qualitative review of the comments provided by raters, particularly regarding other signs of concern noted during the review suggests that both S-LP and nursing raters were attuned to contextual information beyond that which was queried in the screening protocol questions. This finding suggests that the predictive power reported in previous studies where the decisions have not been derived through direct, itemized questions may involve an unknown contribution of context.

RATING PROCEDURES AND DEFINITIONS

In reviewing the previous literature on screening, we found it quite difficult to ascertain the operational definitions used to define the target abnormalities of interest. Definitions for penetration-aspiration and dysphagia are lacking in many previous studies, as are the methodological details necessary to support replication. It is common to rate a wide number of physiological parameters on videofluoroscopy, ranging from lip and soft-palate closure to aspiration (Martin-Harris et al., 2008). In this study, we decided that the target underlying disorder of dysphagia would be defined as an abnormality resulting in a functional consequence (i.e., penetration-aspiration or pharyngeal residues). Our methods of classifying true disposition from the videofluoroscopy were intentionally set with low thresholds for failure (any single occurrence of a penetration-aspiration score ≥ 3 and/ or any single occurrence of a cumulative residue score ≥ 2). It is possible that our definition for dysphagia was more focused and less inclusive than that in other studies, where the criterion for a disposition of dysphagia has been reported as “any abnormality on videofluoroscopy” (Daniels et al., 1998; Martino et al., 2009). If so, one would expect higher sensitivities and lower specificities to be found. On the other hand, our threshold for classifying the screening results as abnormal was intentionally set to err on the side of identifying a problem in the event that at least 50% of the raters within a professional group queried the presence of an abnormality. This low-threshold definition also runs the risk of inflated sensitivity, reduced specificity and increasing false-negatives in validation against a gold-standard method like videofluoroscopy. We found both lower sensitivities and lower specificities with higher false negatives compared to previous studies where the classification threshold definitions are not transparently reported. This suggests the possibility that *both* our clinical screening and our VF classification thresholds may have been set too broadly, and that narrower definitions might need to be established.

POSSIBLE LIMITATIONS OF DIRECT COMPARISON

As noted in our introduction, prior studies of swallow screening performance have compared the single pass-fail result from a screen to the subsequent occurrence of problems, anywhere during the course of an instrumental examination. In our study, we took the approach of limiting our focus only to the limited number of thin liquid swallows in the screening protocol. It is possible that this direct focus led us to miss the power of the screening protocol to predict problems on a more comprehensive

instrumental examination. In such a case, the result would have been increased sensitivity and reduced specificity. The fact that we saw reductions in both of these metrics could be interpreted as an argument against this possibility. It is our opinion that swallowing performance on a thin liquid screening challenge is not logically likely to predict true performance on different tasks, and that to draw such connections is to overstate the power of the original screening test. It should be remembered that the purpose of swallow screening tests is to obtain an accurate initial impression of swallowing performance, which should determine the appropriateness of sending a patient for further assessment. Water or thin liquids are arguably the most likely stimuli to elicit aspiration, and are therefore a good choice for items that are most likely to identify patients who require additional assessment. The emphasis on developing screening protocols that are also good predictors of dysphagia is interesting to consider in this respect. Our experience with this study has caused us to reflect that the five questions included in our study were thematically more oriented to detecting possible aspiration than to other aspects of dysphagia like residue. Indeed, we did not ask patients whether they had a sensation of residue, nor did we ask judges to comment specifically on any observations that might suggest residue (like multiple swallows per bolus). Based on these reflections, we would argue that it is theoretically improbable that the observation of tongue mobility, cough, and voice quality before and after swallows of water would be good predictors of post-swallow residue, either with water or with other stimuli.

CONCLUSIONS

It is difficult to compare results across different screening validation studies in the literature because methods of determining the disposition of dysphagia or aspiration have not been clearly described, sample inclusion criteria differ, and blinding has not always been used. In this study, we performed a direct (blinded) comparison of screening judgments to concurrent VF ratings, removing concerns that the screening and validation tests might capture variable physiological performance in a patient due to circumstantial differences and the passage of time. We had expected to see reasonably good correspondence between the observation of abnormal clinical signs and the true occurrence of aspiration and dysphagia. Our findings, therefore, came as a surprise and have caused us to reflect critically on the goals of swallow screening as a process. Based on our study, we have to conclude that swallow screening decisions based on a series of 3-4 thin liquid swallows do not have good clinical utility for detecting penetration-aspiration or dysphagia with thin liquid stimuli. In

particular, judgments regarding tongue lateralization and post-swallow cough or voice quality changes were not found to have predictive value.

Our results suggest that swallow screening results are imperfect, and suffer both from over-identification and under-identification of the underlying target disorders. In particular, the screening protocol tested in this study yielded a high number of false negative decisions, suggesting that patients with underlying dysphagia or penetration-aspiration would not have been detected by clinicians applying the screening. This finding is of concern, given the objective of swallow screening programs to accurately identify patients at risk for dysphagia and its consequences. Of the five focused questions that were asked during the rating of the screening movies, only the observations of abnormal baseline voice quality and voluntary cough showed acceptable predictive performance for identifying aspiration and dysphagia.

The findings are challenging to consider in terms of clinical practice recommendations. We suggest that clinicians need to recognize that swallow screenings are a very preliminary first step test towards identifying swallowing problems, and that they need to be treated as such. In this respect, we come back to the original expectations regarding screening put forward by the World Health Organization: "Screening tests sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not intended to be diagnostic" (*CCI Conference on preventive aspects of chronic disease*, 1951). However, we believe that the current results also suggest that swallow screening tests may need to involve a greater number of swallows, or be repeated on more than one occasion over the first few days of a patient's admission to hospital if the goal is not to miss patients who may have dysphagia. Furthermore, our results point to the importance of following up with clinical and instrumental swallowing assessments in a timely manner in patients who are identified through swallow screening protocols. This will provide greater detail to inform patient management and should help to limit the over-zealous use of diet restrictions in patients whose swallow screening results are false positives. Certainly, our results reinforce the importance of recognizing that swallow screenings do not provide sufficient information for patient management and cannot be used to replace more detailed swallowing assessment in patients at risk for dysphagia.

REFERENCES

- ASSIST: Acute Swallow Screen in Stroke and TIA. (2009). Retrieved June 24, 2010 from http://www.ircst.health.nsw.gov.au/_data/assets/pdf_file/0006/95685/Can_i_ASSIST_you_with_that.pdf
- Bours, G. J., Speyer, R., Lemmens, J., Limburg, M. & de Wit, R. (2009). Bedside screening tests vs. videofluoroscopy or fibreoptic endoscopic evaluation of swallowing to detect dysphagia in patients with neurological disorders: systematic review. *Journal of Advanced Nursing*, 65(3), 477-493.
- Bravata, D. M., Daggett, V. S., Woodward-Hagg, H., Damush, T., Plue, L., Russell, S. et al. (2009). Comparison of two approaches to screen for dysphagia among acute ischemic stroke patients: Nursing admission screening tool versus National Institutes of Health Stroke Scale. *Journal of Rehabilitation Research and Development*, 46(9), 1127-1134.
- Canadian Stroke Network (2005). *SCORE (Stroke Canada Optimization of Rehabilitation through Evidence). Post-Stroke Evidence-Based Recommendations: Screening for risk of pressure ulcers, falls, dysphagia, cognitive disorders and depression*. Retrieved June 24, 2010 from http://www.trilliumhealthcentre.org/west_GTA_stroke_network/professional/SCORE.pdf
- Capozzoli, M., McSweeney, L., and Sinha, D. (1999). Beyond Kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics*, 27(1), 3 - 23.
- CCI Conference on preventive aspects of chronic disease. (1951). Retrieved June 24, 2010. from http://whqlibdoc.who.int/php/WHO_PHP_34.pdf
- Cichero, J. A., Heaton, S. & Bassett, L. (2009). Triaging dysphagia: nurse screening for dysphagia in an acute hospital. *Journal of Clinical Nursing*, 18(11), 1649-1659.
- Clave, P., Arreola, V., Romea, M., Medina, L., Palomera, E. & Serra-Prat, M. (2008). Accuracy of the volume-viscosity swallow test for clinical screening of oropharyngeal dysphagia and aspiration. *Clinical Nutrition*, 27(6), 806-815.
- Cohen, A. K. (2008). MGH- Swallow Screening Tool (MGH-SST): Validation and implementation in acute neuro patients. Retrieved June 24, 2010, from [http://www.docstoc.com/docs/21656737/MGH--Swallow-Screening-Tool-\(MGH-SST\)-Validation-and](http://www.docstoc.com/docs/21656737/MGH--Swallow-Screening-Tool-(MGH-SST)-Validation-and)
- Cohen, A. K. (2009). Creating a swallow screening program at Mass General Hospital: A model for development and implementation. *Perspectives on Swallowing and Swallowing Disorders*, 18(4), 123-128.
- Daniels, S.K., Brailey, K., Priestly, D.H., Herrington, L.R., Weisberg, L.A. & Foundas, A.L. (1998). Aspiration in patients with acute stroke. *Arch Phys. Med Rehabil*, 79(1), 14-19.
- Eisenhuber, E., Schima, W., Stadler, A., Schober, E., Schibany, N., Denk, D. M. et al. (2005). Evaluation of swallowing disorders with videofluoroscopy in Austria: a survey. *European Journal of Radiology*, 53(1), 120-124.
- Grey-Bruce Health Network Dysphagia Screening Tool (2008). Retrieved June 24, 2010, from <http://www.gbhn.ca/ebc/documents/DYSPHAGIASCREENINGTOOLJuly2008.pdf>
- Groves-Wright, K. J., Boyce, S. & Kelchner, L. (2010). Perception of wet vocal quality in identifying penetration/aspiration during swallowing. *Journal of Speech, Language and Hearing Research*, 53(3), 620-632.
- Iredell Dysphagia Screen (2006). Retrieved June 24, 2010, from http://www.enscn.org/Hospitals%20Plans%20of%20Care%20Templates/Iredell_dysphagia.screen.09.pdf
- Joint Commission on the Accreditation of Healthcare Organizations (2004). *Stroke Performance Measurement Implementation Guide*. Retrieved June 24, 2010 from http://www.jointcommission.org/CertificationPrograms/PrimaryStrokeCenters/implementation_guide.htm
- Leder, S. B., Suiter, D. M. & Green, B. G. (2010). Silent Aspiration Risk is Volume-dependent. *Dysphagia*, online first. doi: 10.1007/s00455-010-9312-2
- Logemann, J. A. (1993). *Manual for the videofluoroscopic study of swallowing*. 2nd edition. Austin, TX: Pro-Ed.
- Logemann, J.A., Veis, S. & Colangelo, L. (1999). A screening procedure for oropharyngeal dysphagia. *Dysphagia* 14, 44-51.
- Lothian Dysphagia Screening Test (2005). Retrieved June 24, 2010, from http://www.lothianstrokemc.scot.nhs.uk/publications/documents/dysphagiascreeningtestv4a_270705.pdf
- Martin-Harris, B., Brodsky, M. B., Michel, Y., Castell, D. O., Schleicher, M., Sandidge, J. et al. (2008). MBS measurement tool for swallow impairment--MBSImp: establishing a standard. *Dysphagia*, 23(4), 392-405.
- Martino, R., Pron, G. & Diamant, N.E. (2000). Screening for oropharyngeal dysphagia in stroke: Insufficient evidence for guidelines. *Dysphagia* 15(1), 19-30.
- Martino, R., Silver, F., Teasell, R., Bayley, M., Nicholson, G., Streiner, D. L. et al. (2009a). The Toronto Bedside Swallowing Screening Test (TOR-BSST): development and validation of a dysphagia screening tool for patients with stroke. *Stroke*, 40(2), 555-561.
- Martino, R., Streiner, D. L., Maki, E., Diamant, N. (2009b). A sensitivity

analysis to determine whether ten teaspoons of water are really necessary. *Dysphagia*, 24(4),473.

Massey, R. & Jedlicka, D. (2002). The Massey bedside swallowing screening. *Journal of Neuroscience Nursing*, 34(5), 252-260.

McCullough, G. H., Rosenbek, J. C., Wertz, R. T., McCoy, S., Mann, G. & McCullough, K. (2005). Utility of clinical swallowing examination measures for detecting aspiration post-stroke. *Journal of Speech-Language Hearing Research*, 48(6), 1280-1293.

Oklahoma Dysphagia Screening Tool (n.d.). Retrieved June 24, 2010, from http://www.integris-health.com/NR/rdonlyres/22E4D1AA-E0D6-4F78-9AC5-C49ED00993A4/0/Dysphagia_Screening.pdf

Perry, L. (2001). Screening swallowing function of patients with acute stroke. Part II: Detailed evaluation of the tool used by nurses. *Journal of Clinical Nursing*, 10, 474-481.

Perry, L. & Love, C. P. (2001). Screening for dysphagia and aspiration in acute stroke: a systematic review. *Dysphagia* 16(1), 7-18.

Rosenbek, J.C., Robbins, J.A., Roecker, E.B., Coyle, J.L. & Wood, J.L. (1996). A penetration-aspiration scale. *Dysphagia* 11(2), 93-98.

Sackett, D.L. & Haynes, R.B. (2002). The architecture of diagnostic research. *British Medical Journal*, 324(7336), 539-541.

Sackett, D.L., Straus, S.E., Richardson, W.S., Rosenberg, W. & Haynes, R.B. (2000). *Evidence-Based Medicine: How to practice and teach EBM*. Toronto, ON: Churchill Livingstone

Scherokman, B. (1997). Selecting and interpreting diagnostic tests. *The Permanente Journal*, 1(2), 1-7.

Schoenfeld, P. (2009). Is this diagnostic test useful? *Am J Gastroenterol*, 104(9), 2139-2142.

Scottish Intercollegiate Guidelines Network (2004). *Management of patients with stroke: Identification and management of dysphagia*. Retrieved June 24, 2010, from <http://www.sign.ac.uk/pdf/sign78.pdf>

St. George Dysphagia Screening Tool (n.d.). Retrieved June 24, 2010, from <http://www.stgeorgedgp.asn.au/division-programs/uploadedFiles/1236656727857-1959.pdf>

Suiter, D. M. & Leder, S. B. (2008). Clinical utility of the 3-ounce water swallow test. *Dysphagia*, 23(3), 244-250.

Trapl, M., Enderle, P., Nowotny, M., Teuschl, Y., Matz, K., Dachenhausen, A. et al. (2007). Dysphagia bedside screening for acute-stroke patients: the Gugging Swallowing Screen. *Stroke*, 38(11), 2948-2952.

Waito, A., Bailey, G. L., Molfenter, S. M., Zoratto, D. C. & Steele, C. M. (2010). Voice-quality Abnormalities as a Sign of Dysphagia: Validation against Acoustic and Videofluoroscopic Data. *Dysphagia*, online first. doi: 10.1007/s00455-010-9282-4

Warms, T. & Richards, J. (2000). "Wet Voice" as a predictor of penetration and aspiration in oropharyngeal dysphagia. *Dysphagia*, 15(2), 84-88.

Zoratto, D. C., Chau, T. & Steele, C. M. (2010). Hyolaryngeal excursion as the physiological source of swallowing accelerometry signals. *Physiological Measures*, 31(6), 843-855.

collection and processing.

AUTHOR'S NOTE

Correspondence should be sent to Catriona M. Steele, Ph.D., Senior Scientist, Toronto Rehabilitation Institute, 550 University Avenue, 12th floor (research) Toronto, ON., M5G 2A2, Canada. E-mail: steele.catriona@torontorehab.on.ca ▶

Received date : August 25, 2010

Accepted date : February 14, 2011

ACKNOWLEDGMENTS

This study was funded by a grant to the first author from the Ontario Centres of Excellence, Ontario Ministry of Research and Training, with industry support from Panacis Medical Inc. The first author also acknowledges support from a Canadian Institutes of Health Research New Investigator Award (Grant #69521). The authors acknowledge the support of Toronto Rehabilitation Institute which receives funding under the Provincial Rehabilitation Research Program from the Ministry of Health and Long-term Care in Ontario. The views expressed are those of the authors and do not reflect those of the ministry.

The authors would like to acknowledge the assistance of Janice Bennett, Nicole Buesselberg, Joon Lee, Mitsuko Takeuchi, Ka Lun Tam and Annemarie Weeda with data



KEY WORDS

TROUBLE DU LANGAGE

TROUBLE PRAGMATIQUE

PRAGMATIQUE

CHILDREN COMMUNICATION
CHECKLIST (CCC)

TRADUCTION

ADAPTATION

ÉQUIVALENCE
LINGUISTIQUE

ÉQUIVALENCE CULTURELLE

ÉQUIVALENCE
CONCEPTUELLE

VALIDITÉ DE CONTENU

VALIDITÉ DE FORME

VALIDITÉ DE CONSTRUIT

Marie Vézina, M.Sc.¹Catherine Samson-
Morasse, M.Sc.²Julie Gauthier-
Desgagné, M.Sc.³Marion Fossard, Ph.D.^{4,5}Audette Sylvestre, Ph.D.^{4,6}

1. Centre de santé et de services sociaux – Institut universitaire gériatrique de Sherbrooke, Sherbrooke, QC Canada
2. Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, QC Canada
3. Centre Hospitalier Universitaire de Québec, Québec, QC Canada
4. Département de réadaptation, Programme de maîtrise en orthophonie, Université Laval, Québec, QC Canada
5. Centre de Recherche Université Laval – Robert Giffard, Beauport, Québec, QC Canada
6. Centre Interdisciplinaire de Recherche en Réadaptation et Intégration Sociale, Québec, QC Canada

► **Développement de la version québécoise francophone du Children's Communication Checklist – 2 (CCC-2). Traduction, adaptation et équivalence conceptuelle.**

► **Development of a Quebec French version of the Children's Communication Checklist – 2 (CCC-2). Translation, adaptation and conceptual equivalence.**

Marie Vézina

Catherine Samson-Morasse

Julie Gauthier-Desgagné

Marion Fossard

Audette Sylvestre

Abrégé

Une évaluation orthophonique complète doit tenir compte de chacune des composantes du langage, soit les aspects de forme (morphologie, syntaxe, phonologie), de contenu (sémantique et lexique) et d'utilisation (pragmatique). Les aspects de forme et de contenu sont relativement bien couverts par les différents outils d'évaluation disponibles pour la population franco-québécoise, contrairement aux habiletés d'utilisation. Pourtant, l'évaluation de la composante pragmatique du langage peut s'avérer déterminante pour l'établissement d'un diagnostic et le choix des interventions chez un jeune, en orthophonie comme en pédopsychiatrie. En effet, l'évaluation des habiletés pragmatiques peut contribuer au diagnostic différentiel entre les troubles du langage (TL) et d'autres troubles développementaux, notamment les troubles envahissants du développement (TED). La présente étude a été réalisée en réaction au manque d'outils validés pour évaluer la pragmatique chez les jeunes franco-québécois. Il s'agit d'une étude de validation en franco-québécois du *Children Communication Checklist* (CCC-2, Bishop, 2006), un outil réputé d'évaluation de la pragmatique dont les qualités métrologiques ont été démontrées. Dans un premier temps, le CCC-2 a été traduit et adapté en français québécois suivant une méthodologie rigoureuse, conforme aux recommandations de la Commission internationale pour l'adaptation des tests. La méthodologie employée a permis la réalisation d'une traduction-adaptation du CCC-2 en franco-québécois qui est équivalente linguistiquement et culturellement à la version source de l'instrument. Par la suite, diverses analyses qualitatives ont été réalisées afin de s'assurer que les paramètres mesurés par l'outil soient équivalents dans les populations américaine (population source) et franco-québécoise (population cible). À cette fin, trois ouvrages réputés au Québec libellant les caractéristiques du phénomène étudié ont été utilisés comme cadre référentiel, afin de s'assurer que les items du test concordaient avec la conceptualisation québécoise du construit. Ces analyses ont démontré l'équivalence conceptuelle des versions américaine et franco-québécoise du CCC-2. Cette étude a donc permis la création de la version franco-québécoise du CCC-2 et la démonstration de l'équivalence linguistique, culturelle et conceptuelle par rapport à la version source.

Abstract

A comprehensive speech and language assessment should cover every components of language, including form (morphology, syntax, and phonology), content (semantics and lexicon), and use (pragmatics). While the former two components of language are relatively well covered by the different evaluation tools available for Quebec french population, the latter is lacking. Yet, appraisal of the pragmatic aspects of language can prove decisive in making a diagnosis and choosing appropriate interventions for a child, both for speech and language assessment, and for child psychiatry. Appraisal of the pragmatic aspects of language can indeed contribute to making a differential diagnosis between language disorders and other developmental disorders, such as pervasive developmental disorders (PDD). This research was conducted in response to the lack of validated tools to evaluate the pragmatic aspects of language of Quebec french children. It provides a translation in Quebec French and a validation study of the Children Communication Checklist (CCC-2, Bishop, 2006), a well-recognized evaluation tool of the pragmatic aspects of language. First, we translated and adapted the CCC-2 tool in Quebec French using a rigorous methodology in accordance with the International Test Commission. The chosen methodology ensured that the resulting Quebec French version of the CCC-2 is linguistically and culturally equivalent to the original version. Second, we carried out various qualitative analyses to test if the parameters measured by the tool are the same in the American population (source population) and the Quebec population (target population). To do this, we used three prominent tools in Quebec which describe the characteristics of the subject under study as a framework of reference in order to ensure that the test items are in accordance with the conceptualization of the construct in Quebec. These analyses confirmed the conceptual equivalence of the American and Quebec French versions of the CCC-2.

INTRODUCTION

Une évaluation complète des habiletés langagières d'un enfant doit permettre l'analyse de chacune des dimensions de la fonction cognitive complexe qu'est le langage, soit celles relatives aux aspects de forme (morphologie, syntaxe, phonologie), de contenu (sémantique et lexique) et d'utilisation (pragmatique). La pragmatique fait référence à l'utilisation fonctionnelle du langage dans un contexte de communication réelle (Owens, 2009). Elle implique l'expression d'intentions de communication via le recours à des formes langagières adaptées à l'interlocuteur et au contexte. Les enfants présentant des troubles du langage (TL) ont fréquemment des déficits pragmatiques. Or, les déficits pragmatiques ne sont pas exclusifs aux troubles du langage. Ils peuvent également s'observer dans diverses problématiques comme l'autisme, le syndrome d'Asperger, les troubles envahissants du développement (TED) non spécifiés, le syndrome de William ou le trouble de déficit d'attention avec ou sans hyperactivité (van Balkom et Verhoeven, 2004). Les interrelations entre la pragmatique et les autres composantes du langage diffèrent selon ces divers troubles donnant lieu à des profils de difficultés de communication bien distincts. L'évaluation de la pragmatique et la mise en lumière des interrelations entre cette composante et les autres paramètres du langage peut donc grandement contribuer au diagnostic différentiel entre divers troubles du développement et, par ricochet, au choix des interventions.

Les enfants présentant un TL de même que ceux présentant un TED font face à des difficultés de communication importantes. Chez les enfants présentant un TL, les difficultés de langage sont généralement primaires, c'est-à-dire que les autres sphères de développement sont relativement bien développées (DSM-IV-TR, 2003). Les enfants présentant un TED, quant à eux, se caractérisent non seulement par leur atteinte au plan de la communication, mais aussi par une altération des interactions sociales, des comportements stéréotypés et/ou des intérêts restreints, et souvent par des perturbations au plan sensoriel (DSM-IV-TR, 2003). Malgré tout, le diagnostic différentiel entre les TL et les TED peut être très ardu, surtout dans la petite enfance. L'évaluation orthophonique des différentes composantes langagières, particulièrement de la pragmatique, peut alors contribuer fortement au diagnostic différentiel.

Chez les enfants présentant un TL, les déficits pragmatiques sont souvent secondaires aux altérations des autres composantes langagières (van Balkom et Verhoeven, 2004). En effet, il apparaît logique qu'un enfant qui présente des difficultés de compréhension et de formulation de phrases (déficits au plan de la

sémantique, de la morphologie et de la syntaxe) éprouve de la difficulté à bien participer à l'interaction en contexte de conversation (impact sur la pragmatique). De plus, les enfants avec un TL ont tendance à compenser leurs difficultés langagières par d'autres modes de communication comme les expressions faciales, les gestes, le regard (DSM-IV-TR, 2003). Au contraire, chez les enfants présentant un TED, les déficits pragmatiques sont souvent prépondérants par rapport aux atteintes des autres sphères du langage et non pas consécutifs. L'évaluation orthophonique des différentes composantes langagières, particulièrement de la pragmatique, peut alors contribuer à l'identification de l'étiologie de l'altération de la communication et contribuer grandement au diagnostic différentiel entre le TL et le TED. Une évaluation complète en pédopsychiatrie est cependant nécessaire pour poser un diagnostic de TED.

Compte tenu de l'importance de la pragmatique pour le diagnostic différentiel des portraits cliniques et pour l'intervention, il est de première importance d'utiliser des mesures d'évaluation valides et fiables de cette composante du langage. Cette condition est essentielle à une approche clinique correspondant à une démarche scientifique rigoureuse. Or, très peu d'instruments permettent l'évaluation de la pragmatique et, à ce jour, seule l'édition la plus récente du Children's Communication Checklist (CCC : Bishop, 1998), le CCC-2 (Bishop, 2006), présente des qualités métrologiques satisfaisantes (Adams, 2002). Cet outil, validé auprès des populations américaine et britannique (Bishop, 2006), permet d'identifier des comportements pragmatiques subtils, difficiles à provoquer ou à observer à l'intérieur d'un échantillon de langage spontané (Adams, 2002; Bishop, 1998; Gauthier-Desagné, Samson-Morasse et Vézina, 2008).

Le CCC a initialement été développé par Bishop (1998). Tout en permettant d'évaluer les différentes dimensions du langage, l'accent principal de cet outil d'évaluation portait alors sur la pragmatique et visait à identifier, parmi les jeunes présentant un trouble persistant du langage, ceux présentant des déficits pragmatiques prédominants comparativement aux altérations dans les autres sphères du langage. Tel qu'expliqué précédemment, un tel profil communicatif peut être indicateur d'un TED. La version 2 de l'instrument (CCC-2) permet non seulement d'identifier les jeunes présentant des déficits pragmatiques prédominants, mais aussi de dépister ceux présentant des difficultés langagières significatives par rapport à la population générale.

Le CCC-2 est un questionnaire auto-administré soit par les parents, soit par l'enseignant de l'enfant. Il comporte 70 items décrivant des comportements de communications distribués en dix sous-échelles :

(A) phonologie et fluidité, (B) morphosyntaxe, (C) sémantique, (D) cohérence, (E) initiation de la conversation, (F) langage stéréotypé, (G) utilisation du contexte, (H) communication non verbale, (I) relations sociales, (J) intérêts. Pour chaque item, le répondant doit se prononcer sur la fréquence d'occurrence du comportement décrit en utilisant une échelle en 3 points qui permet de préciser l'absence du comportement, son caractère occasionnel ou constant (Bishop, 2006). Les résultats aux différentes sous-échelles sont d'abord pondérés de manière à ce que les différents paramètres du langage puissent être comparés entre eux. Deux scores composites sont ensuite calculés à partir des scores pondérés. Le « score composite de communication générale » (score CCG) est dérivé à partir de la somme des scores pondérés aux échelles A à H. Le score CCG peut être utilisé dans une perspective de dépistage pour identifier les jeunes présentant des difficultés communicationnelles significatives. Le second score dérivé, l'« index différentiel d'interaction sociale » (score IDIS), correspond à la différence entre la somme des sous-échelles E, H, I, J et la somme des sous-échelles A, B, C, D. Il constitue un indice pouvant contribuer au diagnostic différentiel entre les TL et les TED. Ce score permet d'identifier les jeunes pour lesquels une évaluation plus approfondie des troubles du spectre autistique est jugée pertinente (Bishop, 2006). Enfin, une analyse qualitative du patron global de résultats, prenant en compte les scores aux différentes sous-échelles et les scores composites, permet de compléter l'évaluation.

Les propriétés métrologiques du CCC-2 ont été vérifiées auprès de jeunes présentant un développement typique et auprès de diverses populations cliniques, notamment des enfants présentant un TL, un trouble pragmatique ou un TED (Bishop, 2006 ; Norbury, Nash, Baird et Bishop, 2004). L'adaptation américaine de l'instrument, parue en 2006, a été normée auprès d'un vaste échantillon composé de 950 jeunes américains âgés de 4 :0 à 16:11 ans issus de la population générale (100 jeunes échantillonnés pour chaque tranche d'âge entre 4:0 et 13:11 ans; 150 jeunes échantillonnés pour l'intervalle d'âge 14:0-16:11). La composition de l'échantillon a été stratifiée selon l'origine ethnique, la région géographique et le niveau de scolarité des parents de manière à optimiser la représentativité de l'échantillon par rapport à la population américaine générale.

Il est bien démontré qu'un outil d'évaluation du langage développé pour une langue et une culture données ne peut être utilisé auprès d'une population différente sans traduction et adaptation préalables (Geisinger, 2003). Puisque les concepts à la base des définitions des troubles du langage peuvent varier d'une population et

d'une langue à l'autre, il est impératif d'établir si le construit mesuré par l'outil à la même signification dans les deux populations et s'il est influencé ou non par la langue et la culture dans lesquelles il est mesuré. L'établissement de cette comparaison constitue la mesure de l'équivalence conceptuelle. Elle permet d'optimiser la validité des conclusions tirées à partir de l'adaptation de l'instrument pour une nouvelle population et d'éviter de négliger des aspects du phénomène étudié qui seraient spécifiques à la culture ou la langue cible.

OBJECTIFS DE L'ÉTUDE

Considérant l'importance de l'évaluation de la pragmatique lors des démarches de diagnostic différentiel entre diverses pathologies du développement, l'absence d'un outil d'évaluation valide et fiable en franco-québécois constitue une lacune majeure en orthophonie, spécialement lors d'interventions relevant du domaine de la pédopsychiatrie. C'est dans le but de combler cette lacune que la présente étude a été réalisée. Elle constitue la première étape du processus de normalisation du CCC-2 en français québécois. L'objectif premier de cette étude est de traduire et adapter le CCC-2 en se basant sur une méthodologie assurant l'équivalence linguistique et culturelle entre la version source et la version franco-québécoise de l'instrument. Le deuxième objectif vise à valider l'équivalence conceptuelle du construit mesuré par le CCC-2 entre la population américaine et la population franco-québécoise.

Objectif 1 : Traduction-adaptation du CCC-2, équivalence linguistico-culturelle

Méthodologie

La traduction et l'adaptation linguistico-culturelle du CCC-2 en franco-québécois sont basées sur les lignes directrices proposées par la Commission Internationale d'adaptation des tests (Geisinger, 2003). La procédure retenue implique non seulement la traduction directe des consignes et items qui permet l'équivalence sémantique, mais elle tient également compte des particularités linguistico-culturelles du franco-québécois en ce qui a trait aux interprétations culturelles, à la familiarité des termes et à leur fréquence dans la langue d'usage (Peña, 2007). Ainsi, cette procédure de traduction vise à obtenir l'équivalence tant linguistique que culturelle.

La procédure retenue pour obtenir cette équivalence comprend quatre étapes combinant la traduction directe, la rétro-traduction, l'approche de révision par comité et les prétests. Lors de ce processus de validation, six personnes ont été mises à contribution soit à titre de traductrices (deux étudiantes franco-québécoise de 2^e cycle en orthophonie, expertes du domaine, connaisseuses des

cultures source et cible, l'une présentant un bilinguisme simultané anglais français dès la naissance, l'autre un bilinguisme successif après l'entrée à l'école acquis lors de cours d'anglais langue seconde et de stages d'immersion), traductrices faisant partie du comité de révision (deux autres étudiantes franco-québécoise de 2^e cycle en orthophonie, expertes du domaine, connaisseuses des cultures source et cible et présentant un bilinguisme successif acquis progressivement lors des études et de voyages à l'étranger) et de réviseurs (une professionnelle possédant des études de niveau universitaire, non spécialiste du domaine et monolingue ainsi que la mère d'un garçon âgé de 9:7 ans présentant un TL, ne connaissant ni l'instrument ni le domaine d'expertise).

Étape 1. Une traduction de la version américaine du CCC-2 vers le français a été réalisée par une première traductrice (version 1).

Étape 2. Une rétro-traduction de la version francophone vers l'anglais a ensuite été effectuée par une autre traductrice travaillant à l'aveugle de la version source du CCC-2 en anglais (version 2). Les deux versions anglaises, soit la version originale anglophone et la version rétro-traduite, ont été comparées par la seconde traductrice sur la base de l'adéquation sémantique et terminologique (version 3). Les différences résultant d'une altération au plan sémantique ont été adaptées dans la version en franco-québécois (version 4).

Étape 3. Par la suite, deux autres traductrices ont ajouté leur contribution au processus afin de former le comité de révision. Les traductrices de ce comité ont procédé à l'étape de révision de la traduction en franco-québécois en comparant la version originale anglophone du CCC-2 à la version 4 francophone. Des modifications ont été proposées à la suite de cette comparaison puis débattues en comité pour parvenir à un consensus (versions 5 et 6).

Étape 4. La première étape du prétest a été réalisée auprès d'une professionnelle possédant des études de niveau universitaire non spécialiste du domaine et monolingue. Ce prétest a été réalisé dans le but de détecter toute ambiguïté de sens pouvant être liée à la traduction, à l'usage de termes trop spécifiques ou à un niveau de langage inadéquat (version 7). Le prétest s'est poursuivi en soumettant la version 7 à un utilisateur potentiel de l'instrument, soit la mère d'un garçon âgé de 9:7 ans présentant un TL. Dans ce cas également, la personne ne connaissait ni l'instrument ni le domaine d'expertise. Les modifications découlant des prétests ont été apportées à la version franco-québécoise selon les commentaires des deux participantes, conduisant à l'obtention de la version finale traduite et adaptée linguistiquement et culturellement à la population franco-québécoise : le

CCC-2 version franco-québécoise.

Résultats

Tel que prévu par le processus de traduction, la seule traduction directe du CCC-2 vers le franco-québécois (étape 1) n'a pas été suffisante pour assurer l'équivalence linguistico-culturelle. En effet, des améliorations ont dû être apportées à travers les étapes du processus de traduction/adaptation tant au plan d'équivalence sémantique (équivalence de sens entre les deux langues) que l'équivalence au plan culturel (termes spécifiques au franco-québécois, familiarité des termes utilisés, formulation de phrases).

Comme le CCC-2 comporte des items évaluant le développement des aspects formels du langage (phonologie, morphologie et syntaxe), ceux-ci ont dû non seulement être traduits en franco-québécois, mais ont également fait l'objet d'une analyse supplémentaire visant à assurer que les termes choisis en franco-québécois référaient à la même étape de développement ou indice de difficultés qu'en anglais américain. Certains items ont pu être traduits directement. Par exemple l'item 2 en anglais « pronounces crocodile as “cockodile” » a été traduit directement par « l'enfant prononce “cocodile” pour crocodile ». Il s'agit en effet d'un mot trisyllabique tant en français qu'en anglais qui a le même niveau de complexité au plan phonologique. De plus, il s'agit d'un mot fréquent pour les enfants tant en anglais qu'en franco-québécois. Cet item directement traduit était donc jugé équivalent au plan linguistique et culturel. D'autres items ont dû faire l'objet d'une adaptation pour assurer l'équivalence au plan linguistico-culturel. Par exemple, l'item 36 « leaves off past tense – ed endings on words (e.g. says “John kick the ball” instead of “John kicked the ball”) » a été adapté au franco-québécois par « n'utilise pas le temps passé des verbes (ex : l'enfant dit « hier, je fais du ski » au lieu de dire « hier, j'ai fait du ski »). Le morphème -ed qui apparaît tôt dans le développement normal de l'anglais comme marqueur du passé (environ 30 à 36 mois, Paul, 2007) n'existe pas en franco-québécois. Pour que l'item évalué traduise le construit dans la version franco-québécoise, il a été adapté par le passé composé, marqueur morphosyntaxique du franco-québécois qui apparaît aux alentours du même âge dans le développement normal (Sylvestre, 2010).

L'étape de comparaison de la version source de l'instrument à la version rétro-traduite (étape 2) a relevé d'autres différences. Les différences relatives à l'utilisation de synonymes ou de formes équivalentes ont été ignorées. Certaines divergences étaient liées à des usages spécifiques à la culture et langue franco-québécoise. Par exemple, les prénoms utilisés dans certains items (ex : John, Sally, Johnny, soit des prénoms fréquents aux États-Unis)

ont été remplacés par Jean, un prénom très fréquent en franco-québécois. Cette adaptation permet de respecter le caractère de fréquence d'usage dans la langue et la culture cible.

Par la suite, après la comparaison entre la version franco-québécoise du CCC-2 et la version source de l'instrument (étape 3), d'autres améliorations ont été apportées. Des modifications ont été proposées au plan sémantique (ex : *langue maternelle* au lieu de *langue première*), de la formulation (*au cours des trois derniers mois* au lieu de *pour au moins trois mois*) et de formes linguistiques relevant d'une traduction directe (*énoncé préféré* au lieu de *énoncé favori*). Ces modifications ont par la suite été débattues en comité pour parvenir à un consensus (versions 5 et 6).

Lors des étapes de prétest (étape 4), les commentaires apportés par les réviseurs ont permis de retravailler certaines formulations ambiguës pour les personnes non familières avec le domaine d'expertise (version 7). Par exemple, l'item 65 « smiles appropriately when talking to people » avait initialement été traduit « utilise un sourire adéquat lorsqu'il parle avec les gens ». Cette formulation a été modifiée par « sourit de manière adéquate lorsqu'il parle avec les gens » suite au commentaire d'une réviseure qui comprenait difficilement comment le sourire pouvait être « utilisé ». Cette formulation est souvent employée par les orthophonistes qui caractérisent le sourire comme un moyen de communication non verbal que l'enfant « utilise ». Les néophytes et donc potentiels répondants à ce questionnaire n'entrevoient pas le sourire comme pouvant être « utilisé ». Cette formulation a donc été adaptée. Notons cependant qu'aucun élément majeur du questionnaire traduit et adapté ne portait à confusion. Ces modifications ont conduit à la version finale du processus de traduction-adaptation, soit le CCC-2 *version franco-québécoise*.

Discussion

Les spécialistes du domaine de la psychométrie affirment qu'une évaluation menant à un diagnostic valide est directement liée à l'utilisation d'outils et de procédures culturellement et linguistiquement validés. Le CCC-2 est un instrument d'évaluation orthophonique innovateur qui a été soumis à une procédure de traduction/adaptation pour étendre son utilisation à la population franco-québécoise. Une méthodologie rigoureuse en plusieurs étapes a été utilisée lors de ce processus de validation pour optimiser l'équivalence linguistico-culturelle de la version traduite et, par ricochet, augmenter sa validité interne. En effet, la traduction-adaptation de l'outil a été effectuée selon des normes strictes reconnues internationalement, suivant les « lignes directrices de la Commission internationale des tests ».

De plus, les traductrices ont été choisies de manière à refléter les recommandations des spécialistes, soit être experts du domaine de connaissances, être bilingues et connaître à la fois la culture source et la culture cible (Geisinger, 2003; Massourbe, 2002; van de Vijver & Leung, 1997, dans Geisinger, 2003). Notons cependant que trois traductrices sur quatre présentaient un bilinguisme successif plutôt que simultané. Leur niveau de compétence en langue anglaise était inférieur à leurs compétences en français. Certaines subtilités linguistiques auraient donc pu leur échapper. Cependant, les étapes ultérieures du processus de traduction-adaptation ont certainement permis de corriger ces lacunes, le cas échéant. Par ailleurs, bien qu'elles soient connaisseuses des cultures source et cibles, les traductrices étaient toutes d'origine franco-québécoise et non américaine. Des subtilités inhérentes à la culture américaine propre auraient donc pu leur échapper. Cela est toutefois peu probable, puisque toutes les traductrices ont été fortement exposées à la culture américaine relative aux construits mesurés par le CCC-2, particulièrement lors de leurs études universitaires en orthophonie. En effet, le choix des ouvrages de références et les enseignements universitaires au Québec sont fortement influencés par les pratiques et les recherches américaines.

Par ailleurs, il aurait été pertinent de soumettre la version traduite du CCC-2 à davantage de réviseurs lors des étapes de pré-test. Une seule mère d'un enfant présentant un TL a complété le questionnaire à cette étape de validation, alors qu'il s'agit pourtant de la clientèle cible à qui l'outil est destiné. L'opinion de réviseurs issus de différentes classes socio-économiques ou ayant différents niveaux de scolarité aurait peut-être permis de mettre en lumière d'autres ambiguïtés dans le questionnaire. Cette lacune pourrait nuire à la validité externe lors des procédures ultérieures de validation de l'outil.

Néanmoins, vu la méthodologie rigoureuse employée et les améliorations apportées au questionnaire tout au long du processus, le critère d'équivalence linguistico-culturel est rencontré. La méthodologie utilisée pour la traduction va au-delà de la traduction directe et inclut les éléments d'équivalence culturelle, en tenant compte de l'interprétation culturelle, de la familiarité et de la fréquence des formulations linguistiques. La version franco-québécoise du CCC-2 est donc adaptée linguistiquement et culturellement pour permettre une interprétation équivalente des items du test. Ainsi, l'extension de son utilisation à d'autres populations francophones hors Québec devrait passer par un même processus d'adaptation linguistico-culturel. Bien qu'il s'agisse de la même langue, les termes utilisés peuvent être non familiers pour une autre population francophone,

tout comme les exemples retenus peuvent ne pas refléter la réalité culturelle d'une autre population. Ainsi, le CCC-2 version franco-québécoise est considéré comme une adaptation valide, soit équivalente linguistiquement et culturellement à la version source de l'instrument pour la population franco-québécoise.

Objectif 2 : Équivalence conceptuelle

Méthodologie

Le second objectif de recherche consiste à valider l'équivalence conceptuelle, c'est-à-dire à vérifier si le construit mesuré par le CCC-2 version américaine peut être généralisé à la population franco-québécoise. En d'autres termes, il s'agit de vérifier si le construit sous-jacent à chacun des items du CCC-2 est compris de la même manière par la population franco-québécoise et s'il entraîne des réponses similaires au questionnaire, comparativement à la population américaine (Peña, 2007). Il est largement reconnu que la conceptualisation du langage et de ses pathologies diffère à travers les groupes culturels et linguistiques. Inévitablement, les outils utilisés dans l'évaluation orthophonique d'un enfant doivent représenter et respecter les normes langagières et communicationnelles de la culture et de la langue cibles, et ce, afin de rendre compte du profil langagier de l'enfant et d'avoir le potentiel de déceler la présence d'un déficit. Une interprétation culturellement différente des concepts d'un test peut affecter la façon dont les individus d'un groupe culturel répondent aux questions (Peña, 2007). L'équivalence conceptuelle a donc été évaluée par le biais de diverses analyses qualitatives décrites ci-dessous.

Le CCC-2 est un instrument qui évalue à la fois les aspects d'utilisation (pragmatique), de forme (morphologie, syntaxe, phonologie) et de contenu (sémantique et lexique) du langage. Ces différents paramètres ont donc tous été considérés pour juger de l'équivalence conceptuelle. Tous les items du questionnaire CCC-2 ont été passés en revue par trois étudiantes franco-québécoises de niveau maîtrise en orthophonie, ayant donc une connaissance approfondie du concept franco-québécois du développement du langage et de la communication ainsi que de ses indices de difficultés. L'analyse qualitative a porté à la fois sur le choix des items, leur formulation et les exemples donnés pour illustrer chacune des forces ou difficultés langagières, en comparaison avec le cadre de référence franco-québécois. Le but de cette analyse était de s'assurer que chaque item du CCC-2 respectait la conceptualisation de la norme du développement du langage et de la communication ainsi que ses altérations pour la population franco-québécoise.

Pour la conceptualisation franco-québécoise de l'utilisation du langage (sous-échelles D à H), le cadre référentiel utilisé pour l'analyse de l'équivalence

conceptuelle se compose de trois outils franco-québécois reconnus dans la pratique orthophonique québécoise et fréquemment utilisés pour évaluer la pragmatique. Ces trois outils libellent des stades de développement du langage et de la communication ainsi que des indices de difficultés langagières en franco-québécois. Il s'agit de *S.O.S. com : Situations d'Observation Structurées de la Communication* (Groupe de travail des orthophonistes des commissions scolaires de la Montérégie, 2002)¹, des *Guide et outils cliniques : Trouble primaire du langage/dysphasie* (Ordre des orthophonistes et des audiologistes du Québec, 2004)² et de la grille *Aperçu de la chronologie du développement de la communication chez les enfants* (Sylvestre, 2010)³. En ce qui concerne les aspects de forme et de contenu du CCC-2 (sous-échelles A, B, C), l'analyse a été réalisée à partir d'un seul outil comme cadre de référence franco-québécois, soit la grille *Aperçu de la chronologie du développement de la communication chez les enfants* (Sylvestre, 2010). Cet outil franco-québécois décrit de manière détaillée le développement des aspects de forme et de contenu. Cette grille critériée est fondée sur les connaissances issues de 32 ouvrages de références franco-québécois, américains ou européens. Elle est régulièrement mise à jour en fonction des nouvelles publications scientifiques. Elle inclut notamment les résultats des travaux récents de Sutton (2007) et Trudeau (2010) portant sur le développement du langage en franco-québécois. De plus, les balises développementales présentées dans cette grille de développement ont été validées via la consultation d'orthophonistes franco-québécoises ayant une vaste expérience clinique auprès des jeunes enfants. Cet outil est d'ailleurs fortement utilisé par les orthophonistes du Québec francophone. À noter que la grille *Aperçu de la chronologie du développement de la communication chez les enfants* (Sylvestre, 2010) est un outil qui décrit le développement typique du langage et non les déviations. Elle a tout de même été retenue pour la présente étude comme base de comparaison pour les items 51 à 70 du questionnaire qui traitent des forces langagières, soit des indices de développement normal des jeunes. Elle a également été utilisée pour les items 1 à 50 portant sur les difficultés langagières, en référence aux stades de développement typique présentés qui sont considérés déviants s'ils perdurent dans le temps.

Sur la base de ces outils utilisés comme cadre référentiel, les indices de difficultés ou de forces langagières décrites dans le CCC-2 ont été passés en revue afin d'identifier s'ils se trouvaient également cités dans un ou plusieurs des outils franco-québécois et respectaient donc la conceptualisation franco-québécoise. Ensuite, une réflexion sur les pratiques orthophoniques américaines et franco-québécoises a permis de juger globalement de la proximité culturelle de ces deux populations en regard

du construit mesuré. Les constats issus de cette réflexion seront présentés dans la discussion.

Résultats

Conceptualisation de la pragmatique

La comparaison entre les items relatifs à l'utilisation du langage du CCC-2 version américaine et le contenu des outils franco-québécois retenus comme cadre de référence conceptuel permet de confirmer le recouvrement complet entre la conceptualisation des deux populations. Tel qu'illustré succinctement dans le tableau 1, les 35 items du CCC-2 version américaine qui font référence à des indices d'altération ou de forces de la pragmatique (sous-échelles D à H) sont également cités dans l'un ou plusieurs des outils franco-québécois de manière analogue.

Conceptualisation des sphères de forme et de contenu

Les résultats de l'analyse comparative des items du CCC-2 faisant référence aux sphères de forme et de contenu présentés dans le tableau 2 permettent également de constater un recouvrement quasi complet de la conceptualisation de ces sphères entre la culture anglo-américaine et franco-québécoise. En effet, la quasi-totalité des indices de difficultés du développement de la forme et du contenu en langue anglaise (anglais américain) sont également retrouvés dans le développement franco-québécois. La seule difficulté qui n'est pas relevée en franco-québécois concerne un processus phonologique appliqué sur un phonème que ne possède pas le franco-québécois, le /θ/. Pour respecter le construit mesuré,

Tableau 1.

Comparaison des contenus d'items référant à l'utilisation du langage entre le CCC-2 version américaine et trois outils d'évaluation franco-québécois* (extrait)

Contenus d'items du CCC-2 version américaine référant au discours et à la pragmatique	Habilité/déficit du discours et de la pragmatique selon les outils franco-québécois	Guide et outils cliniques : Trouble primaire du langage	Aperçu de la chronologie du développement de la communication chez les enfants	S.O.S. com
D : Cohérence du discours				
Utilise des termes tels que <i>il</i> ou <i>ça</i> sans clarifier de quoi il/elle parle (ex : en racontant un film, l'enfant dit « il était vraiment génial » sans mentionner à quoi le <i>il</i> fait référence).	Évalue mal les informations pertinentes à donner à l'interlocuteur pour être compris. Ajuste mal son discours aux différents éléments du contexte (situation, interlocuteur)	X	X	
F : Langage stéréotypé				
Répète ce que les autres viennent de dire (ex : si on demande à l'enfant « qu'est-ce que tu as mangé? » il/elle répond par « qu'est-ce que tu as mangé? »).	Utilise la répétition / écholalie		X	X
H : Communication non verbale				
Utilise adéquatement les gestes pour se faire comprendre.	(Force) Hoche la tête pour accompagner le langage (oui / non), hoche la tête pour remplacer le langage	X		X

* Les trois outils d'évaluation franco-québécois utilisés sont 1) S.O.S. com : Situations d'Observation Structurées de la Communication (Groupe de travail des orthophonistes des commissions scolaires de la Montérégie, 2002), 2) Guide et outils cliniques : Trouble primaire du langage/dysphasie (Ordre des orthophonistes et des audiologistes du Québec, 2004) et 3) Chronologie du développement de la communication chez les enfants (Sylvestre, 2010)

Tableau 2.

Comparaison des items référant à la forme et au contenu du CCC-2 version américaine avec le contenu de l'outil *la Chronologie du développement de la communication chez les enfants* (Sylvestre, 2006) (extrait)

Items du CCC-2	Phénomène langagier	Indicateurs d'un développement langagier atypique en franco-québécois	Adaptations nécessaires lors de la traduction
A – Phonologie et fluence			
Item 38	Erreurs phonologiques sur les mots longs	Oui	L'exemple utilisé en anglais pour un mot long (4 syllabes) n'a pas été repris puisqu'il ne s'agit pas d'un mot long en français. Il a été remplacé par un mot comportant le même nombre de syllabes et la même structure phonologique en français.
B – Syntaxe			
Item 27	Omission de constituants syntaxiques	Oui	Pour le premier exemple « me got ball » pour « I've got the ball », un exemple avec le verbe aimer a été utilisé car l'omission de constituants syntaxiques ne pouvait être reproduite avec le verbe avoir (moi ai balle).
C- Sémantique			
Item 32	Difficultés d'accès lexical	Oui	Les mots en anglais ont été substitués par des mots similaires au plan phonologique en franco-québécois.

cet item a été remplacé par un processus phonologique qui s'actualise dans les mêmes âges (antériorisation des palatales).

DISCUSSION

Le CCC-2 version américaine évalue les sphères d'utilisation (pragmatique), de forme (morphologie, syntaxe, phonologie) et de contenu (sémantique, lexique) du langage. Les items sélectionnés pour opérationnaliser ces composantes sont considérés être, selon la culture américaine, des indices d'atypies discriminants ou de forces dans le développement du langage. Le développement de ces composantes est propre à la langue d'usage, de sorte que des différences inhérentes à la langue peuvent modifier le construit mesuré par l'outil. Une analyse de ces items s'aurait donc essentielle pour déterminer s'ils pouvaient s'appliquer ou non au développement du langage en franco-québécois et pour proposer des adaptations, si jugé nécessaire.

En ce qui concerne la pragmatique, l'analyse qualitative des items du CCC-2 permet de constater

de grands recouvrements entre la conceptualisation de cette sphère du langage telle qu'opérationnalisée dans le questionnaire et le cadre de référence franco-québécois. En effet, tous les indicateurs de développement atypique ou harmonieux de la pragmatique présents dans le CCC-2 sont répertoriés dans l'un ou l'autre des ouvrages de références franco-québécois retenus comme cadre de référence. Il en est de même pour les items du CCC-2 relatifs aux sphères de forme et de contenu du langage. Ces items s'appliquent donc également à la population franco-québécoise et conservent ainsi leur signification clinique. Ces analyses qualitatives renforcent la validité de contenu du questionnaire et constituent un indicateur additionnel d'équivalence conceptuelle. Elles sont basées, non pas sur des normes, mais sur des outils et ouvrages de références dont les critères ont été établis en fonction d'observations cliniques. Il est regrettable qu'il y ait peu de normes sur le développement du langage franco-québécois pour valider nos conclusions.

Par ailleurs, mêmes si les indicateurs du CCC-2 évaluant les sphères de forme et de contenu se retrouvent

également dans les ouvrages de références franco-québécois utilisés, cela ne signifie pas pour autant qu'il s'agisse d'indicateurs des troubles du langage (TL / SLI) tout aussi discriminants en anglais américain qu'en franco-québécois. Des études récentes mettent effectivement en évidence des différences ainsi que des similarités dans les marqueurs développementaux des TL (SLI) en français et en anglais (Royle & Thordardottir, 2008; Paradis, Crago, & Rice, 2003). Ces marqueurs sont des indices précis de forme et du contenu du langage qui caractérisent fortement le développement du langage chez les enfants présentant un TL. Il est donc important de noter que la présence d'items concernant les sphères formelles du langage dans le CCC-2 ne vise pas à identifier les enfants présentant un TL (Bishop, 2006). Ces items permettent plutôt d'identifier les enfants qui pourraient présenter des difficultés de langage significatives et qui devraient être évalués en orthophonie. D'ailleurs, ces items réfèrent à des marqueurs d'altération évidents en langue anglaise américaine, ils ne permettent donc pas une analyse en profondeur. Pour ces raisons, la méthode de comparaison qualitative retenue pour la présente étude a été jugée suffisante. Il est certain qu'une étude plus approfondie des marqueurs de déficits de forme et de contenu retenus dans la version anglophone de l'outil, en comparaison avec ceux des TL du français québécois, aurait permis une plus grande validité de la traduction de l'outil, particulièrement pour les sous-échelles A, B et C.

CONCLUSION

La méthodologie rigoureuse employée, conforme aux lignes directrices de la Commission internationale pour l'adaptation des tests, a permis la réalisation d'une traduction-adaptation du CCC-2 en franco-québécois qui est équivalente linguistiquement et culturellement à la version source de l'instrument. Par ailleurs, les analyses réalisées ont permis de démontrer l'équivalence conceptuelle des versions américaine et franco-québécoise du CCC-2. Les résultats concluants à l'analyse qualitative de l'équivalence des construits relatifs aux différentes sphères langagières indiquent une forte équivalence conceptuelle entre les deux versions. La proximité culturelle entre les populations franco-québécoise et américaine renforce également l'équivalence conceptuelle. Ces résultats permettent de statuer que le construit mesuré par le CCC-2 version américaine sous toutes les facettes étudiées est équivalent pour les populations américaine et franco-québécoise et conserve sa signification, avec les adaptations réalisées. La démonstration de l'équivalence conceptuelle renforce la validité interne du CCC-2 version franco-québécoise, puisque les conclusions tirées à partir de cet outil ne sont pas biaisées par des dimensions inhérentes à la langue et/ou la culture. Ces indices de

validité étant désormais bien documentés, le CCC-2 version franco-québécoise pourra être soumis aux étapes ultérieures de la validation, soit la normalisation, puis la démonstration de l'équivalence métrique et de l'utilité clinique auprès de la population franco-québécoise.

RÉFÉRENCES

- Adams, C. (2002). Practitioner Review: The assessment of language pragmatics. *Journal of Child Psychology and Psychiatry*, 43(8), 973-987.
- American Psychiatric Association – DSM-IV-TR. *Manuel diagnostique et statistique des Troubles mentaux*, 4^e édition, Texte Révisé (Washington DC, 2000). Traduction française par J.-D. GUELFY et al., Masson, Paris, 2003, 1120 pages.
- Bishop, D. V. M. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry*, 39(6), 879-891.
- Bishop, D. V. M. (2006). *Children's Communication Checklist – 2nd Edition*, U. S. Edition. San Antonio, TX : PsychCorp.
- Gauthier-Desgagné, J., Samson-Morasse, C., Vézina, M. (2008). *Étude de validation en franco-québécois du Children's Communication Checklist-2 (CCC-2)*. Essai de maîtrise inédit, Université Laval, Québec, Canada.
- Geisinger, K.F. (2003). Testing and assessment in cross-cultural psychology. *Handbook of Psychology: Assessment Psychology*, (10), 95-117. Published Online: (2003, Apr 15). doi: 10.1002/0471264385.wei1005.
- Groupe de travail des orthophonistes des commissions scolaires de la Montérégie (2002). *SOS.com : Situations d'Observations Structurées de la communication*. Saint-Hyacinthe, QC : ÉRS-DI-TED.
- Massourbe, C., Lang, F., Jaeger, B., Jullien, M., & Pellet, J. (2002). La traduction des questionnaires et des tests : techniques et problèmes / Translating Questionnaires and Tests: Techniques and Problems. *Canadian Journal of Psychiatry*, 47(1), 63-69.
- Norbury, C. F., Nash, M., Baird, G. & Bishop, D. V. M. (2004). Using a parental checklist to identify diagnostic groups in children with communication impairment : a validation of the Children's Communication Checklist-2. *International Journal of Language and Communication Disorders*, 39(3), 345-364.
- Ordre des orthophonistes et des audiologistes du Québec. (2004). *Guide et outils cliniques : Troubles primaires du langage/Dysphasie*. Montréal, Québec.
- Owens, R. E. (2005). *Language development: An introduction*. Boston: Pearson Education.
- Paradis et al. French-English Bilingual Children with SLI: Who do they compare with their monolingual peers? *Journal of Speech, Language and Hearing Research*, 46(2), 113-127.
- Paul, R. (Eds.) (2007). *Language Disorders from Infancy through Adolescence, Assessment and Intervention*, 3rd edition. St-Louis, Missouri: Mosby Elsevier.
- Peña, E.D. (2007). Lost in translation: Methodological consideration in cross-cultural research. *Child Development*, 78(4), 1255-1264.
- Royle, P. & Thordardottir, E. T. (2008). Elicitation of the *passé composé* in French preschoolers with and without specific language impairment. *Applied Psycholinguistics*, 29(3), 341-365.
- Sutton, A., Trudeau, N., Poupart, A., Grenon, V., Thordardottir, E. & Lessard, N. (2007). *Développement phonologique chez les enfants francophones âgés de 18 à 54 mois*. Congrès de l'Association Canadienne des Orthophonistes et des Audiologistes, Moncton, NB.
- Sylvestre, A. (2010). *Chronologie du développement de la communication chez les enfants*. Document inédit, Université Laval, Québec, Canada .
- Sylvestre, A., Desmarais, C., & Gobeil, S. (2007). Notes de cours ORT-66275 : *Troubles développementaux du langage*. Document inédit, Université Laval, Québec, Canada.
- Trudeau, N. (2010). *Émergence des temps de verbes entre 16 et 30 mois*. Présentation au 78^e Congrès de l'Acfas, Montréal, mai
- van Balkom, H., & Verhoeven, L. (2004). Pragmatic disability in children with specific language impairments. Dans L. Verhoeven, & H. van Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp. 283-289). Mahwah, NJ: Lawrence Erlbaum Associates.
- van de Vijver, F., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1(2), 89-99.

REMERCIEMENTS

Nous tenons à remercier tout spécialement Isabelle Morasse et Geneviève Samson, pour leur contribution aux étapes de prétest de l'adaptation-traduction de l'outil.


NOTES MARGINALES

¹S.O.S. com : Situations d'Observation Structurées de la Communication (Groupe de travail des orthophonistes des commissions scolaires de la Montérégie, 2002), est un guide d'observation et d'évaluation de la communication chez l'enfant présentant un TED. Il a été élaboré par les orthophonistes du groupe d'intérêt en autisme de la région de la Montérégie et fait état d'indices de développement pour la pragmatique et la régie de l'échange pour une clientèle scolaire de tout âge.

²Guide et outils cliniques : Trouble primaire du langage/dysphasie (Ordre des orthophonistes et audiologistes du Québec, 2004) est un guide québécois d'évaluation de la dysphasie. Il fait également état d'indices de déficits pragmatiques et discursifs aux plans réceptif et expressif.

³Aperçu de la chronologie du développement de la communication chez les enfants (Sylvestre, 2010) est une grille détaillant la chronologie du développement de l'ensemble des sphères de la communication, notamment de la pragmatique et du discours.

NOTE DES AUTEURS

Prière d'adresser toute correspondance à : Audette Sylvestre, Ph.D., Programme de maîtrise en orthophonie, Département de réadaptation, Université Laval, 1050, Avenue de la Médecine, bureau 4412, Québec (Québec) G1V 0A6, Canada. Courriel: audette.sylvestre@rea.ulaval.ca 

Date soumis : le 21 mai 2010

Date accepté : le 2 février 2011

- **Performance of French-speaking Quebec adults on the Boston Naming Test**
- **Résultats du Boston Naming Test chez des adultes francophones de Québec**

Patricia M. Roberts
Nathalie Doucet

KEY WORDS

BOSTON NAMING TEST

FRENCH

NAMING

LOW EDUCATION

CULTURE

SENSITIVITY

NAME AGREEMENT

Abstract

The Boston Naming Test (BNT) is an American test, widely used in Canada, in both English and French. This study examined the 60 stimuli of the BNT for name agreement in Quebec French as a first step towards developing scoring rules. Sensitivity of the test to age and education differences was also tested. Forty-five Quebec francophones (ages 50 to 79) with a mean of 8.6 years of education and little knowledge of English were asked to name all 60 pictures in French. The responses were scored after consulting 3 different dictionaries. There are 13 pictures with synonyms in French, and several additional items where it is unclear what names should be accepted. Nearly half the items are French-English cognates. Despite these problems with the stimuli, a 2 by 2 ANOVA (age by education) found significant effects ($p < .008$) for age and for education, indicating that the test, even with no cultural adaptation, may be sufficiently sensitive to detect different levels of naming ability. Future studies will focus on replacing some of the problematic pictures to achieve better name agreement and reduce the number of cognates.

Abrégé

Le Boston Naming Test est un test américain largement utilisé au Canada anglais et français. Cette étude a vérifié les 60 stimuli du test en français québécois, servant ainsi de première étape dans l'élaboration de règles de notation. La sensibilité du test aux variables âge et éducation a aussi été testée. Quarante-cinq francophones du Québec (âgés de 50 à 79 ans) ayant une moyenne de 8,6 années d'éducation et une faible connaissance de l'anglais ont dû nommer 60 images en français. Les réponses ont été cotées après la consultation de trois dictionnaires différents. Il y a 13 images qui ont des synonymes en français et plusieurs autres images où il est difficile de déterminer quels noms devraient être acceptés. Près de la moitié des images sont des mots jumeaux en français et en anglais. En dépit de ces problèmes avec les stimuli, une analyse de variance à deux variables (âge par éducation) a trouvé un effet significatif pour l'âge et l'éducation ($p < .008$), indiquant que le test, même sans adaptation culturelle, pourrait être suffisamment sensible pour identifier différents niveaux de capacité de dénomination. Les études à venir seront axées sur le remplacement de certaines images problématiques, afin d'atteindre une meilleure concordance des noms et de réduire le nombre de mots jumeaux.

Patricia M. Roberts,
Faculty of Health Sciences,
University of Ottawa,
Ottawa, ON, Canada

Nathalie Doucet, Faculty
of Health Sciences,
University of Ottawa,
Ottawa, ON, Canada

The most widely used test of confrontation naming ability may be the Boston Naming Test (Kaplan, Goodglass & Weintraub, 2001). Published in the United States, and not designed as an international instrument, it has nonetheless enjoyed widespread use in North America, England, and Australia (Katz et al., 2000). There are studies of the test, with or without changes to the stimuli, in at least nine different languages. Garcia and Desrochers (1997) found that the Boston Naming Test (BNT) is often used to assess French-speaking patients in Canada. This is surprising, given that there are no published norms or even French scoring guidelines for the test. As this paper will show, scoring the BNT in French is far from simple.

THREE CHARACTERISTICS OF AN IDEAL TEST OF NAMING

Visually unambiguous: Ideally, a confrontation naming test is made up of pictures that are visually unambiguous (Snodgrass & Vanderwart, 1980). Clinical experience shows that some of the BNT drawings are frequently misidentified. The standardized cues to be given when a visual misidentification is made (e.g., saying *snake* for *pretzel*) or when the wrong part of the picture is named (e.g., *man* instead of *stilts*) compensate for this, to some degree. Rossion and Pourtois (2004) recommend that future editions of the BNT use coloured drawings as this would reduce the number of misidentifications.

Name agreement: Ideally, all pictures should have only one possible correct name (Lecours & Lhermitte, 1979; Lezak, 1995; 2004). Cuetos, Ellis, and Alvarez (1999) and Marquez et al., (2007) propose using 85% name agreement as a criterion for stimuli. A high level of name agreement for test stimuli is important for at least six reasons. Having a single possible correct answer:

1. makes scoring easy and quick, which is important in clinical settings;
2. enhances inter-rater reliability. If the patient changes clinicians over the course of their recovery, each new clinician can administer the test, confident that their scoring is neither stricter nor looser than that of their colleagues. Similarly, across published studies, scores can be compared without idiosyncratic scoring rules acting as a confounding factor;
3. makes it possible to give phonemic cues. If there are two or three possible names for a given picture, it is impossible to know which one the person being tested is trying to find (Bender, Martin Garcia, & Barr, 2010; Roberts & Bois, 1998; 1999). Giving a cue that does not match the word being

searched for, or the word on the tip of the tongue of the person being tested could lead him/ her to mistakenly conclude that the planned answer was wrong;

4. reduces response time and allows response times to be compared across items and experimental groups. Mitchell (1989) found that response time was slower for items with more than one possible name;
5. allows analysis based on the properties of that single answer. If there are two possible names for a picture (e.g., *harmonica* or *mouth organ*), it is impossible to analyse results based on factors such as familiarity, frequency and word length. All of these factors influence response time and accuracy of naming (e.g., Rochefort & Williams, 1962; Dent, Johnson, & Humphreys, 2008). Word frequency affects naming success in many aphasic adults (Nickels & Howard, 1995);
6. allows analysis based on likely age of acquisition. Hodgson and Ellis (1998) have found that age of acquisition influences naming speed and accuracy. These effects persist in aphasic speakers (e.g., Cuetos, Aguado, Izura, & Ellis, 2002; Nickels & Howard, 1995). For a recent review see Johnson and Barry's recent work (2006).

Sensitivity: A test should accurately discriminate between different levels of ability and allow for testing a range of ages or levels without encountering floor or ceiling effects. It has been shown that the BNT does this quite well in studies of English-speaking North Americans, when a range of ages and education levels and adequate power levels are used (Neils, et al., 1995; Randolph, Lansing, Ivnik, Cullem, & Hermann, 1999; Welch, Doineau, Johnson, & King, 1996; Zec, Burkett, Markwell, & Larsen, 2007a). Some studies with low power and/or fairly narrow ranges of age or education show little or no correlation between BNT scores and education or age (e.g., Nicholas, Brookshire, MacLennan, Schumacher, & Porrazzo, 1989). Neurologically intact adults younger than 70 years with at least a high school education tend to score near the maximum (54 to 57 out of 60), in studies in the USA and Canada (Randolph et al., 1999; Tombaugh & Hubley, 1997; Zec et al., 2007a, 2007b) but this is not the target group for this test.

In a growing body of studies in different languages, the BNT has accurately discriminated between groups with different levels of education (e.g., Allegri, et al., 1997; Patricacou, Psallida, Pring, & Dipper, 2007; Rami et al., 2008) and those with and without language impairments

(e.g., Cheung, Cheung, & Chan, 2004; Marquez de la Plata et al., 2008; Tsolaki, Tsantali, Lekka, Kiosseoglu, & Kazis, 2003).

TESTS ACROSS LANGUAGES AND CULTURES

When a naming test is used to assess speakers of languages other than the one it was written in, two additional characteristics to consider are cultural appropriateness and cognates.

Cultural appropriateness

Many authors have pointed out the need to consider cultural factors when tests “travel” from one country or language to another (e.g., Ardila, 2007; Ardila, Rosselli,

& Puente, 1994; Garcia & Desrochers, 1997). If pretzels are not a common snack food in New Zealand or Greece, and if testing shows a large discrepancy between the percent correct in different cultural groups/countries, then this item should be either replaced, or retained but moved to a better place in the sequence from easy to difficult items. Some thought and testing are needed to ensure that either the replacement item is of roughly the same difficulty as the original item it is replacing and that multiple replacements balance each other to maintain the range of difficulty of the original test samples.

It is important not to overextend the notion of cultural appropriateness as applied to tests (see also Roberts, 2001, 2008). Lapointe-Goupil, Everett, Rousseau, Giguère, Laplante, and Keller (2004), in the only published study of

Table 1

Examples of studies with different versions of the BNT in different languages and countries

Items in bold font were found to be problematic in the present study

Language/country	Version	Reference
English – Canada	English – original	Tombaugh & Hubley, 1997
English – New Zealand	English – original but recommend replacing pretzel and beaver	Barker-Collo, 2007
French – Canada	Original but noted culturally inappropriate items	Lapointe-Goupil et al., 2004
French – Switzerland	Found low sensitivity to education for the items: abacus, beaver, cactus, dart, hanger, igloo, pelican, seahorse, stethoscope , scroll , unicorn, latch , protractor, tripod, volcano, whistle, yoke . Replaced these in final, shortened version of the test for use in French	Thuillard-Colombo & Assal, 1992
Swedish	Original	Tallberg, 2005
Greek – Greece	Replaced 4 items: pretzel , door knocker , stethoscope , scroll with, respectively, a kind of cake, mailbox, blood pressure instrument, ancient Greek column	Patricacou, 2007
Spanish – Argentina, Spain	Original Spanish version replaces 15/60 items: whistle; saw; toothbrush; mushroom; hanger; wheelchair; pretzel ; seahorse; wreath; escalator ; door knocker ; asparagus; latch ; scroll ; trellis. However, the more recent Neuronorma project (Peña – Casanova et al., 2009) kept all original items except pretzel (Spain)	The original Spanish version is used in Argentina (Allegri et al., 1997) and Spain (Rami et al., 2007); a newer version was created by Peña – Casanova et al. (2009)
Cantonese – Hong Kong	Selected 30 items from original test	Cheung, Cheung, & Chan, 2004

the BNT with Quebec French-speaking adults, state that sphinx and sea horse should be replaced as these items are “not very representative of Quebec context” (p. 204: “peu représentatifs du contexte québécois”). By the same logic, *camel*, *rhinoceros*, and *pyramid* would also have to be replaced. These things are not part of American or English Canadian culture either, but it is reasonable to expect educated American and Canadian adults to know what they are. Similarly, Thuillard-Colombo and Assal (1992) and Lapointe-Goupil et al. (2004) both object to noose because it has negative associations. One assumes this would be true in all cultures. However, this is not a valid, culturally-based reason to delete it from the test.

Aside from the published Spanish version of the test (Garcia-Albea, Sanchez-Bernardos, & del Viso-Pabon, 1986), researchers and clinicians generally use the published English version, asking patients to respond in their own language. Some studies in European countries have recommended deleting or replacing some items as culturally inappropriate. Other studies (see Table 1) have used the 60 item test as is and found significant effects for age and education (e.g., Marien, Mampaey, Vervaeke, Saerens, & De Deyn, 1998; Tallberg, 2005). The original 60 pictures were used, but regionalisms and synonyms were scored as correct for 15 pictures in Italian (Riva, Nichelli & Devoti, 2000) and 37 in Dutch (Marien et al., 1998).

Some authors have chosen to make quite massive changes. For example, Kim and Na (1999) replaced 49 of the 60 pictures with no published data to justify the replacements. In more measured approaches, the first step has been to field test the 60 pictures to explore name agreement and whether the item is familiar to people in that specific country. For Greek adults, Patricacou and colleagues (2007) asked 10 well-educated adults to name “culturally specific” items. Four items which 8 of the 10 volunteers failed to name correctly were replaced. The name agreement for these (newly drawn) pictures was then checked, informally, with 10 new adult participants. In selecting replacement items, the authors tried to stay within the same general category. This approach contrasts with that of Thuillard-Colombo and Assal (1992), assessing French-speaking Swiss adults. The authors created an 85 item experimental version of the test but then, prior to testing, removed nine items they felt were culturally inappropriate. These included the items *wheelchair* and *noose*, both seen as too negative based on the authors’ judgement, without any testing.

The starting premise of the current study is that before deciding that a particular item is problematic, the original test should be given, and the results scrutinized for name agreement, difficulty level, and sensitivity. Intuition is not a sufficient basis for modifying a test.

Cognates

Ideally, a naming test used to assess bilingual speakers should have a low number of cognate words in the pairs of languages tested. Cognates are words with the same meaning and very similar phonology and/or spelling in two languages. A test with as few cognates as possible is desirable because:

1. cognate words are processed differently in bilingual speakers (Costa, Santeseban & Caño, 2005; de Groot & Nas, 1991);
2. cognates are easier to name than non-cognates (Gollan, Fennema-Notestine, Montoya & Jernigen, 2007; Lemhofer, Dijkstra, & Michel, 2004; Roberts & Deslauriers, 1999);
3. in clinical practice, it is sometimes difficult to tell if a patient is pronouncing a word in the target language or not, if the two words are very similar in the two languages;
4. if the words are cognates, it is possible to “make up” the equivalent word in the other language and end up with a correct answer.

The goals of the present, exploratory study were to assess the name agreement in French for the 60 BNT items and to determine if the unmodified BNT is sensitive to the effects of age and education in a sample of elderly speakers of Quebec French.

METHOD

Participants

Forty-five adults (28 women and 17 men) aged between 50 and 79 years old volunteered to participate in the study. They were from three towns in Quebec (Drummondville, Trois-Rivières and Lévis) where 70% to 75% of citizens report being only able to speak French (Statistics Canada, 2006). Most participants responded to notices and announcements made in church halls and community centres. Some were located via friends and acquaintances of the second author. Overall, education levels were low (mandatory education laws were not in effect during their youth). For the entire group, the mean number of years of education was 8.6 years (4 to 22 years of education for the younger group under 65 years and 3 to 14 years of education for the group aged 65 and over). Ten participants in the younger age group and 13 in the older group had less than 9 years of education. The mean education levels for each age group are listed in Table 2. The most common occupations were housewife (12), textile or clothing factory worker (8), construction worker or

Table 2**Age and education levels of participants**

		Age			
Education		Below 65 years		65 or more	
		Mean (SD)	Range	Mean (SD)	Range
		All		All	
	< 9 years	58.2 (3.8) <i>n</i> = 10	51 to 63	71.3 (5.01) <i>n</i> = 13	65 to 78
	≥ 9 years	56.5 (4.35) <i>n</i> = 13	51 to 64 <i>n</i> = 9	73.0 (3.84)	66 to 79
		Mean Age		Mean Age	
		65.6 (7.98) <i>n</i> = 23		63.2 (9.26) <i>n</i> = 22	

carpenter (4) and mechanic (3). Seven participants were office workers (2 teachers, 1 secretary, 1 receptionist, 2 pastoral care workers, 1 government office worker). The other occupations (one of each) were butcher, clerk/housewife, seamstress/housewife, day labourer, farmer, janitor, paper factory worker, landlord, mailman, jewelry craftsman, and steelworker.

In analyses, participants were divided into two groups for level of education (< 9 years versus 9 or more years of education), and two age groups (50–69 versus 70–79 years old) as shown in Table 2. The mean number of years of education was significantly different for the two education groups at 6.00 versus 11.2 [$t(27.55)$, corrected for unequal variances] = 7.07, $p < .001$]. The two groups based on age also significantly differed in mean age: 57.2 versus 72.0 [$t(43) = -11.44$, $p < .001$]. The younger group mean number of years of education was 9.7 (range 4 to 22, SD 4.1) while the older group mean years of education was 7.3 (range 3 to 14, SD 2.5). The difference in education levels between the younger and older groups was not significant [$t(43) = 1.04$, $p = 0.306$].

Language background was documented using a relatively short questionnaire administered in French, by a native speaker of Quebec French. All participants stated that French was their native language and the language used at least 95% of the time in their daily lives (work, home, TV/computer, social life). The only other language knowledge reported was English. None of the participants reported learning any English before age 10. Of the 13/45 who stated that they “spoke another language”, 11 were in the group with higher level of education. Of these, 11/13 claimed to speak, understand, read, and write French “much better” than English (the options were: same, a little better, much better). The other two

reported speaking French (1) or understanding French (1) “a little better than” English. None of the 13 reported using English more than 5% of the time at work or at home. Only 3 reported ever used English more than 5% of the time. One woman reported speaking English to a friend for one year when she was 11 or 12 years old. Another woman spent one year in the USA when she was in her early twenties, but reports never using English at home or at work. Thus, even those reporting some knowledge of English rarely used it.

All participants stated they had no communication problem, memory disorders, head injury or stroke and no neurologic or psychiatric troubles. None reported a history of alcohol or substance abuse.

Each volunteer was seen individually in his or her home or in a local community centre. All 60 pictures were presented in order. If no answer was given after 20 seconds or if an answer was a clear misidentification of the picture (e.g., *snake* for *pretzel*), the researcher gave a semantic cue. No phonemic cues were given. The total correct was answers with no cue plus those following the semantic cue.

Scoring

It was necessary to develop a list of acceptable answers to score the test. In English, all items have (ostensibly) only one correct, one-word name, if the person is told that all names are single words (to avoid composites like park bench, tennis racquet, Christmas wreath; however, see Roberts, Garcia, Desrochers, & Hernandez, 2002 and Zec, Markwell, Burkett, & Larsen, 2005, for regional synonyms for some items in English). This is not the case in French. For some items, the only possible answer is more than one word (e.g., *brosse à dents*). For others, there is more than one correct, widely used name: Seahorse

can be either *hippocampe* or *cheval de mer*. Snail can be *colimaçon* or *escargot*. These items were scored as correct if either name was given.

Some items have names widely used in some regions that are not correct, standard French. These regionalisms were accepted if at least one Canadian dictionary of French listed them as correct (Bélisle, 1979; Boulanger, 1992; Proteau, 1991). The expected names, synonyms in standard French, and regional variants are presented in Table 3.

If the participant gave more than one answer (it could be a _____ or a _____), they were asked to pick one and if they could not choose, the second answer was the one

we scored. If a participant named the object and then said something like “But I don’t know”, we scored the name. The correct response plus an adjective was scored as correct, since participants were not told to give only one word answers, as noted above (e.g., *grosse maison* for *maison*). Errors in pronunciation were accepted if only one phoneme was incorrect and the resulting response was not another real word. For example, *captus* for *cactus* was scored as correct. All scoring was checked by both authors and any disagreements resolved through discussion and careful review of the 3 dictionaries of Canadian French chosen as the basis for scoring decisions.

Table 3

Translations and alternate responses for English and French
All names listed in this table were counted as correct responses

English (alternate or synonym)	French (synonym)	Non-standard, regional
1. bed	lit	
2. tree	arbre	
3. pencil	crayon	
4. house	maison	
5. whistle	sifflet	
6. scissors (pair of scissors)	ciseaux (paire de ciseaux)	
7. comb	peigne	
8. flower	fleur	
9. saw (hand saw)	scie	égoïne
10. toothbrush	brosse à dents	
11. helicopter	hélicoptère	
12. broom	balai	
13. octopus	pieuvre	
14. mushroom (toadstool)	champignon	
15. hanger (clothes or coat hanger)	cintre (support)	
16. wheelchair	chaise roulante (fauteuil roulant)	
17. camel	chameau	
18. mask	masque	
19. pretzel	bretzel (pretzel)	
20. bench	banc	
21. racket	raquette	
22. snail	escargot (calimaçon)	
23. volcano	volcan	
24. sea horse	hippocampe (cheval de mer)	
25. dart	fléchette (dard)	
26. canoe	canot (canoë)	
27. globe	globe terrestre	mappemonde, boule terrestre
28. wreath	couronne	
29. beaver	castor	
30. harmonica (mouth organ)	harmonica	musique à bouche, ruine babines

31. rhinoceros (rhino)	rhinocéros	
32. acorn	gland	
33. igloo	igloo	
34. stilts	échasses	
35. dominoes	dominoes	
36. cactus	catus	
37. escalator	escalier roulant	
38. harp	harpe	
39. hammock	hamac	
40. knocker (door knocker)	marteau (heurtoir)	
41. pelican	pélican	
42. stethoscope	stéthoscope	
43. pyramid	pyramide	
44. muzzle	muselière	
45. unicorn	licorne (unicorne)	
46. funnel	entonnoir	
47. accordion (squeeze box)	accordéon	
48. noose	corde de potence (nœud coulant)	
49. asparagus (asparagus spears)	asperge	
50. compass	compas	
51. latch	loquet	barrure, clenche
52. tripod	trépied	
53. scroll	manuscrit	parchemin
54. tongs	pincettes (pinces)	
55. sphinx	sphynx	
56. yoke	joug	
57. trellis	treillis	tuteur
58. palette	palette	planche à/ pour peindre planche à peinture; planche à toile
59. protractor	rapporteur d'angles	
60. abacus	boulier (boulier compteur)	

Bold indicates items that are very difficult to score in French. See text for alternate names.

Responses with adjective added (e.g., park bench, tennis racquet, Halloween mask) were accepted as correct and are not listed in this table.

RESULTS

Name Agreement

In the English BNT, disregarding items often given two-word names (tennis racquet, coat hanger, Halloween mask...), there are only 3 items with more than one correct name in Canadian English: mushroom/ toadstool (drawing is ambiguous), harmonica/ mouth organ, and accordion/ squeeze box. Thirteen items were found to have two possible names in standard French, as listed in most dictionaries. These are *saw*, *harmonica*, *hanger*, *wheelchair*, *pretzel*, *snail*, *seahorse*, *dart*, *canoe*, *globe*, *unicorn*, *tongs*, and *abacus*. Table 4 shows the percent correct for each item, including synonyms.

In addition, there were 6 items where scoring was very difficult, even after consulting dictionaries (*escalator*, *noose*, *latch*, *scroll*, *yoke*, and *palette*). This was partly due to a lack of consensus among native speakers on the names and partly due to the nature of French. Many object names are composite nouns, with *à* or *de* linking words that make up a single name: toothbrush – brosse à dents (brush for teeth) and cheval de mer – seahorse. This means that when people say “corde à _____” for *noose*, it is difficult to set a dividing line between “appropriate, multi-word name” and circumlocution or description given because the person does not know the name of this picture. Few participants gave either of the dictionary labels for this item (*noeud coulant* or *corde de potence*). *Palette* is also

difficult to score for the same reason. *Planche* has many meanings in French, so the name of this particular object cannot be a single word. It is unclear at what point the many variations become descriptive, circumlocution type errors: *planche à peinture*, *planche de peintre*, *planche pour peindre*, *planche pour (la/les) peinture(s)*.

Scoring is made more difficult by the tendency of Canadian French to accept the English word, in some instances, or the English word, given a French twist. Thus *escalator* pronounced as if it were a French word is seen by many people as wrong (based on an informal check with native French speakers enrolled in speech-language pathology) yet it was found in one of the three chosen dictionaries and, therefore, scored as correct. It is unsatisfactory that the choice of dictionary could lead to different scoring decisions on some responses. The need to distinguish between commonly used and lexically correct labels requires that a clinician make scoring judgements that are not required in English. Cumulatively, these could change an individual's score by 6 or more out of 60. These scoring dilemmas reduce the inter-rater reliability to unacceptable levels, in the absence of any published BNT scoring guidelines in French.

In Table 4, errors in one phoneme are included in the total correct for the relevant word. These errors were (number of participants who made the error): *hélico_tère* 1; *bratzel* 1; *fretzel* 1; *ninocéros* 3; *rhonocéros* 1; *captus* 9; *tactus* 1; *_ téthoscope* 2; *mususlière* 1; *entonnoi_2*. Of the 22 instances of these slips of the tongue, only 4 were made by participants over 70 years old. All 9 *captus/cactus* errors were by adults under age 70.

Table 4

Percent of participants correctly naming each picture (due to rounding, totals and subtotals do not always add up) % of answers for each possible correct name shown in parentheses

1. bed	100
2. tree	98
3. pencil	98
4. house	89
5. whistle	98
6. scissors	100
7. comb	100
8. flower	91
9. saw - (scie 56) (égoïne 44)	100
10. toothbrush	98
11. helicopter	87
12. broom	91
13. octopus	87

14. mushroom	98
15. hanger (cintre 49) (support 51)	100
16. wheelchair (chaise roulante 71) (fauteuil roulant 18)	89
17. camel	80
18. mask	87
19. pretzel (pretzel 49) (bretzel 7)	56
20. bench	100
21. raquet	93
22. snail (escargot 24) (calimaçon 36)	60
23. volcano	89
24. sea horse (hippocampe 16) (cheval de/s mer/s 31)	47
25. dart (fléchette 13) (dard 51)	64
26. canoe (canot 73) (canoë 7) (chaloupe 7)	87
27. globe (globe terrestre 76) (mappemonde 11) (boule terrestre 4)	91
28. wreath	98
29. beaver	84
30. harmonica (harmonica 16) (musique à bouche 78)	94
31. rhinoceros	75
32. acorn	78
33. igloo	71
34. stilts	40
35. dominoes	62
36. cactus	67
37. escalator (escalier roulant 58) (escalier mobile 7) (escalier électrique 4) (marches roulantes 4) (escalator 4)	78
38. harp	76
39. hammock	89
40. knocker (marteau 2) (heurtoir 4)	6

41. pelican	31
42. stethoscope	34
43. pyramid	76
44. muzzle	55
45. unicorn (licorne)	42
46. funnel	85
47. accordion	100
48. noose	
(corde de potence 2)	
(noeud coulant 13)	
(câble avec noeud coulant 2)	28
(corde de pendu 5)	
(câble à pendu 2)	
(câble de pendaison 4)	
49. asparagus	62
50. compass	69
51. latch	
(loquet 2)	
(barrure 58)	73
(clenche 9)	
(verrou 4)	
52. tripod	62
53. scroll	
(parchemin 36)	38
(manuscrit 2)	

54. tongs	71
(pincettes 4)	
55. sphinx	29
56. yoke	
(joug 22)	
(attelage 13)	
(licou 4)	45
(collier 4)	
(harnais 2)	
57. trellis	
(treillis 18)	29
(tuteur 7)	
58. palette	
(palette 22)	
(planche à peintre 7)	
(planche pour peintre 2)	34
(planche à peinture 2)	
(planche à toile 2)	
59. protractor	18
60. abacus	
(boulier 47)	
(boulier compteur 9)	58
(abaque 2)	

Table 5**Mean number of correct names**

	Years of education					
	≤ 9 years		9 years or more		Total	
	Standard	Dialect	Standard	Dialect	Standard	Dialect
<u>50-64 years old</u>						
Mean	38.2	43.9	43.54	48.8	41.2	46.6
S.D.	(5.75)	(6.05)	(4.03)	(3.92)	(5.45)	(5.42)
Range	29 to 49	31 to 52	36 to 51	40 to 56	29 to 51	31 to 56
n	10	10	13	13	23	23
<u>64-79 years old</u>						
Mean	33.3	37.6	39.11	44.2	35.7	40.3
S.D.	(9.05)	(8.13)	(4.43)	(6.28)	(7.92)	(8.02)
Range	18 to 54	22 to 50	31 to 44	33 to 52	18 to 54	22 to 54
n	13	13	9	9	22	22
<u>All groups (n=45)</u>						
Mean	35.43	40.35	41.7	46.9	38.5	43.56
S.D.	(8.02)	(7.8)	(4.7)	(5.4)	(7.2)	(7.5)
Range	18 to 54	22 to 52	31 to 51	33 to 56	18 to 54	22 to 56

Note:

Standard: accepted standard French translation, with or without adjective; Dialect: also accepted regionalisms and colloquialisms

The four people with the highest education levels (22, 16, 15, and 14 years of education) obtained relatively high scores (49, 49, 48 and 42, respectively). These were the only participants with post-secondary education. However, the three highest scores on the test (56, 52 and 52) were obtained by individuals with 12, 10 and 9 years of education, respectively.

SENSITIVITY TO AGE AND EDUCATION

The number of correct answers varied with age (higher scores for the younger groups) and with education (higher scores for those with more than 9 years of schooling). As Table 5 shows, this was true for both methods of scoring. Accepting regionalisms and other variants raised the mean group scores by 4 to 7 points.

STATISTICAL ANALYSIS.

After checking for homogeneity of variance (Levene tests for each proposed comparison, all $p > .1$), two analyses of variance were done (2 x 2 : education x age), one using standard French and the other with regionalisms accepted (see Table 6). Regardless of the scoring method (accepting only standard French responses or also accepting regionalisms) there were significant effects for education and for age, both with large effect sizes (Cohen's d). There were no significant age by education interactions in either ANOVA.

DISCUSSION

Name Agreement on Stimuli

The BNT stimuli, as a set, failed to meet minimum levels of name agreement in French. The 85% standard discussed in the introduction was not met for 17 of the pictures. If the test were developed in French, no one

would select so many items with more than one name. Multiple possible names make it difficult to score the test in a reliable way. Table 3 is not a definitive list; there are too many items where it remains unclear which answers should be considered correct. Depending on which dictionary of Canadian French is used to guide decisions about alternative names for pictures, the scoring rules would change.

Interpretation of test results may be particularly problematic for French-English bilingual speakers. Depending on the level of resemblance needed to be considered cognates, the BNT has between 23 and 30 French-English cognates. In some cases, one possible French name is a cognate and the other is not (e.g., *dart* could be *dard* or *fléchette*). The large number of cognate words helps those with some knowledge of English to score higher on the test in French by guessing. Roberts and Deslauriers (1999) showed that this strategy (dubbed the Bilingual Chipmunk Strategy) is used by both aphasic and non-aphasic, bilingual adults. The issue of cognates has been raised elsewhere (Roberts et al., 2002) and will be explored in future studies where participants have higher levels of proficiency and greater use of both languages. Given the limited English knowledge and patterns of use by the participants in the current study, the cognate status of possible picture names likely had little to no impact on scores.

Table 6

ANOVA (2 x 2) for Age and Education

Accepting only Standard French responses					
	df	F	p	Partial eta ²	Power
Age	1	5.98	.019	.127	.66
Education	1	8.54	.006	.172	.81
Age by Education	1	0.015	.904	.000	.052
Error	41	(39.82)			
Accepting Regionalisms					
Age	1	8.127	.007	.165	.80
Education	1	9.123	.004	.182	.84
Age by Education	1	0.209	.650	.005	.073
Error	41	(39.56)			

Note: Values enclosed in parentheses represent Mean square errors.

Replacing 4 to 6 items on the test would simplify the scoring. This approach would be similar to that followed for the Greek and Australian adaptations (Cruise, Worrall, & Hickson, 2000; Patricacou, Psallida, Pring, & Dipper, 2007), while being somewhat less radical than some of the Spanish adaptations. Candidates for replacement, based on poor name agreement and scoring ambiguities are *tongs*, *latch*, *yoke*, *palette*, *noose*, and *escalator*. The fact that two of these are French-English cognates (*palette* and, for some, *escalator*) provides a further reason for replacing them. *Door knocker* was correctly named by 97% of English speaking Canadians, with at least 12 years of schooling (Tombaugh & Hubley, 1997) but by only 6% of the present sample. Any revised version of the test will need to consider extreme differences like this in balancing the overall difficulty level (and hence the sensitivity) of the test. A study of a larger and more representative sample of Canadian French speakers (on the 60 BNT items and on possible replacement items) will be necessary before making decisions about replacements. Some additional items might be replaced to reduce the number of French-English cognate items in the test.

Age and Education Effects

Despite these scoring and stimulus problems, the mean score on the BNT was significantly affected by both education and age, no matter how the test was scored. Thus, our results are consistent with those of other studies (see Table 7) and they indicate a degree of validity for the test, in French. Furthermore, the mean scores are similar to those of English-speaking Americans with similar, low levels of education. There were phonemic errors on 7 different words which we scored as correct. Scoring these as errors would slightly lower the overall group mean for the younger group but it would remain similar to that of adults with little education in other studies. In their study of American English speakers, Neils et al. (1995) said “according to our results for non-institutionalized subjects, the cut-off for the sixth to ninth grade education group should be 35; and the cut-off for the two higher educational groups should be 40” (p. 1148). By this standard, the unmodified test administered in French correctly identified 21/22 participants with 9 or more years education as scoring within normal limits. Of the 21 participants with 6 to 8 years of education, only

Table 7
Comparison with other studies (the sub-group most similar to the present study)

	Age range	n	Education (years)	Mean BNT Score/60	S.D.
French-Quebec					
Roberts & Doucet (standard) (dialect)	50-79	45	8.6	38.5	7.2
				43.6	7.5
Lapointe-Goupil 2004	n/a (mean 69.9)	17	8.8	44.4	6.6
Mariën et al. 1998 (Dutch)	55-91	92	<10	49.6	6.4
Patricacou 2007 (Greek)	20 to >71	26	0 to 6	33.9	8.5
American-low education					
Neils et al. (1995)(USA-Eng)	65-74	12	6 to 9	47.6	6.1
	75-84	19	6 to 9	42.8	10.9
	65-74	xx	10 to 12	53.0	6.6
Nicholas et al., 1989 (USA-Eng)	40-78		13.9	54.5	
Welch & al., 1996	80+	20	<12	40.7	11.3
	75-79	17	<12	36.6	10.9

Note: Age and education are in years; n/a means not available; standard = correct French ; Dialect = regionalisms also accepted.

2 scored below this American cut-off score. Both had 6 years education and their scores were of 31 and 22 /60. Since there was no screening for cognitive ability, it is unclear whether these low scores (especially the 22/60) are due to the BNT stimuli or to neurocognitive problems.

The low level of education and the fact that only 3 of the 45 completed university make it hard to generalise results to the broader population of Francophones in Quebec. With the introduction of mandatory education laws in the 1960s (requiring that all children remain in school until age 16), groups composed of people born since the mid-1950s will have considerably more education than those in this study. Therefore, it will be important to replicate this study with a larger and more representative sample of adults, in Quebec and in other regions where French is spoken (especially regions in Ontario and New Brunswick) before drawing final conclusions about the validity of the BNT in French.

Education effects on BNT scores have been found in various languages in most studies with adequate power (e.g., Marien et al., 1998; Neils et al., 1995; Patricacou et al., 2007; Thuillard-Colombo & Assal, 1992; Welch et al., 1996). As in other studies (e.g., Randolph, Lansing, Ivnik, Cullum, & Hermann, 1999; Tombaugh & Hubley, 1997; Welch et al., 1996), education influenced scores slightly more than age. Given that only 3 of the 45 participants completed university, and most did not complete high school, the range of education levels was not ideal. With a broader and more balanced distribution, there would likely be a stronger education effect. The fact that a significant education effect was found, despite the small *n* and a skewed distribution in education levels, testifies to the potential sensitivity and validity of the BNT in French.

Nevertheless, much of the variance in scores remains unaccounted for. Education and age together only explain 34% of the variability. Future studies will be necessary to identify further sources, including level of bilingualism, reading habits, and the frequency of occurrence of each possible name in French compared to English.

Despite the problems with the stimuli, it is both encouraging and puzzling that the unchanged BNT was sensitive to differences in education and age. Significant effects were found with both lexically strict scoring rules and with broader rules that accepted regional and colloquial names. Lapointe-Goupil et al. (2004) dismissed the BNT as “culturally biased” because items such as *sphinx* are not part of Quebec culture. The authors proposed a new, 36 item test using pictures from the Snodgrass and Vanderwart set (Bellefleur, 2001). However, their own results show that the French BNT (with no details provided as to how it was scored, and with small groups) was sensitive to differences in performance

between a control group and a group of adults with clinical depression. These results should encourage rather than discourage the use of the BNT in Quebec.

There are a number of advantages to using the same test across different languages, with adaptations where necessary to ensure roughly equal difficulty and sensitivity. Although other naming tests have been developed in French, including the Masters thesis of Bellefleur (see Lapointe-Goupil et al., 2004) and a European test (Metz-Lutz, Kremin, Deloche, Hannequin, Ferrand, Perrier, et al., 1991), these tests are not easy to obtain (especially now that the Snodgrass and Vanderwart pictures are copyrighted and must be purchased). The age of acquisition for some names and the name itself may not always be the same for the different varieties of French spoken in Europe and in Canada. Using the BNT (modified or not) facilitates comparisons across languages, especially for studies of bilingualism and bilingual language acquisition and loss.

Before recommending or discarding the BNT for use with French-speaking Canadians, more study is needed.

ACKNOWLEDGMENTS

The authors thank the churches and community centres that allowed us to recruit participants and M. Guilbault, M. Tremblay, L. Roy and J-P. Allard for facilitating the recruitment process.

REFERENCES

- Acevedo, A., & Loewenstein, D. A. (2007). Performance on the Boston Naming Test in English-Spanish bilingual older adults: Some considerations. *Journal of the International Neuropsychological Society*, 13(2), 212-214.
- Allegri, R. F., Mangone, C.A., Villavicencio, A. F., Rymberg, S., Taragano, F. E., & Baumann, D. (1997). Spanish Boston Naming Test norms. *The Clinical Neuropsychologist*, 11, 416-420.
- Ardila, A. (2007). Toward the development of a cross-linguistic naming test. *Archives of Clinical Neuropsychology*, 22, 297-307.
- Ardila, A., Rosselli, M., & Puente, A. E. (1994). *Neuropsychological evaluation of the Spanish speaker*. New York: Plenum.
- Barker-Collo, S. (2007). Boston Naming Test performance of older New Zealand adults. *Aphasiology*, 21, 1171-1180. doi:10.1080/02687030600821600
- Bellefleur, S. (2001). Développement d'un outil de dénomination adapté à la population francophone nord-américaine : normes préliminaires. Unpublished manuscript. Laval University.
- Bélisle, L.-A. (1979). *Dictionnaire Nord-Américain de la langue Française*. Montréal: Beauchemin.
- Bender, H. A., Martin Garcia, A., Barr, W. B. (2010). An interdisciplinary approach to neuropsychological test construction: Perspectives from translation studies. *Journal of the International Neuropsychological Society*, 16, 227 - 232. doi:10.1017/S1355617709991378227
- Boulanger, J.-C. (1992). *Dictionnaire québécois d'aujourd'hui : langue française, histoire et géographie*. St-Laurent, QC: Dicorobert.
- Cheung, R.W., Cheung, M. C., & Chan, A. S. (2004). Confrontation naming in Chinese patients with left, right, or bilateral damage. *Journal of the International Neuropsychological Society*, 10, 46-53. doi:10.1017/S1355617704101069
- Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language*, 94, 94-103. doi:10.1016/j.bandl.2004.12.002
- Cruise, M. N., Worrall, L.E., & Hickson, L. M. H. (2000). Boston Naming Test results for healthy older Australians : A longitudinal and cross-sectional study. *Aphasiology*, 12, 143-155.

- Cuetos, F., Aguado, G., Izura, C., & Ellis, A. W. (2002). Aphasic naming in Spanish: Predictors and errors. *Brain and Language*, 82, 344-365.
- Cuetos, F., Ellis, A., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, 31, 650-658.
- de Groot, A. M. B., & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, 30, 90-123.
- Dent, K., Johnston, R. A., & Humphreys, G. W. (2008). Age of acquisition and word frequency effects in picture naming: A dual-task investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(2), 282-301.
- Garcia, L. J. & Desrochers, A. (1997). L'évaluation des troubles du langage et de la parole chez l'adulte francophone. *Journal of Speech Language Pathology and Audiology*, 21, 271-293.
- Garcia-Albea, J. E., Sanchez-Bernardos, M. L., & del Viso-Pabon, S. (1986). Test de Boston para el diagnóstico de la afasia: Adaptación española [Boston Naming Test for aphasia diagnosis: Spanish version]. *de la afasia y de trastornos relacionados* [Assessment of aphasia and related disorders] (2nd ed., pp. 129-198). Madrid, Spain: Editorial Medica Panamericana.
- Gollan, T. J., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. J. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society*, 13, 197-208.
- Hodgson, C. & Ellis, A. W. (1998). Last in, first to go: age of acquisition and naming in the elderly. *Brain and Language*, 64, 146-163.
- Johnson, R. A. & Barry, C. (2006). Age of acquisition and visual processing. *Visual Cognition*, 13, 789-845. doi:10.1016/S0093-934X(03)00262-1
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.
- Katz, R., Hallowell, B., Code, C., Armonstrong, E., Roberts, P., Pound, C., & Katz, L. (2000). Multinational comparison of aphasia management practices. *International Journal of Language and Communication Disorders*, 33, 303-314.
- Kim, H. & Na, D. L. (1999). Normative data on the Korean version of the Boston Naming Test. *Journal of Clinical and Experimental Neuropsychology*, 21, 127-133.
- Lapointe-Goupil, R., Everett, J., Rousseau, F., Giguère, V., Laplante, L., & Keller, E. (2004). Étude comparative entre le Boston Naming Test et le Test de Dénomination Bellefleur. *Revue québécoise de psychologie*, 25(2), 203-212.
- Lecours, A. R. & Lhermitte, F. (1979). *Laphasie*. Montréal: Les Presses de l'Université de Montréal.
- Lemhöfer, K., Dijkstra, T., & Michel, M.-C. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, 19, 585-611.
- Lezak, M., D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment*, 4th edition. New York: Oxford University Press.
- Lezak, M. (1995). *Neuropsychological assessment*. 3rd edition. New York: Oxford University Press.
- Mariën, P., Mampaey, E., Vervaeke, A., Saerens, J., & De Deyn, P. P. (1998). Normative data for the Boston Naming test in native Dutch-speaking Belgian elderly. *Brain and Language*, 65, 447-467. doi:10.1006/brln.1998.2000
- Marquez de la Plata, M., Vicioso, B., Hynan, L., Evans, H. M., Diaz-Arrastia, R., Lacruz, L., & Cullum, C. M. (2007). Development of the Texas Spanish Naming Test: a test for Spanish speakers. *The Clinical Neuropsychologist*, 22, 288-304. doi: 10.1080/13854040701250470
- Metz-Lutz, M. N., Kremin, H., Deloche, G., Hannequin, D., Ferrand, I., Perrier, D., Quint, S., Dordain, M., Bunel, G., Cardebat, D., Larroque, C., Lota, A. M., Pichard, B. & Blavier, A. (1991). Standardisation d'un test de dénomination orale: contrôle des effets de l'âge, du sexe et du niveau de scolarité chez les sujets adultes normaux. *Revue de Neuropsychologie*, 1(1), 73-95.
- Mitchell, D. B. (1989). How many memory systems? Evidence from aging. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 31-49.
- Neils, J., Baris, J. M., Dell'aira, A. L., Nordloh, S. J., Weiler, E., & Weisger, B. (1995). Effects of age, education and living environment on Boston Naming test performance. *Journal of Speech and Hearing Research*, 38, 1143-1149.
- Nicholas, L. E., Brookshire, R. H., MacLennan, D. L., Schumacher, J. G., & Porrazzo, S. A. (1989). Revised administration and scoring procedures for the Boston Naming Test and norms for non-brain-damaged adults. *Aphasiology*, 3, 569-580. doi: 10.1080/02687038908249023
- Nickels, L., & Howard, D. (1995). *Aphasic naming- what matters?* *Neuropsychologia*, 33, 1281-1303. doi:10.1016/0028-3932(95)00102-9
- Patricacou, A., Psallida, E., Pring, T., & Dipper, L. (2007). The Boston Naming Test in Greek: Normative data and the effects of age and education on naming. *Aphasiology*, 21, 1157-1170. doi: 10.1080/02687030600670643
- Peña-Casanova et al. (2009). Spanish multicenter normative studies (NEURONORMA project): Norms for Boston Naming Test and Token Test. *Archives of Clinical Neuropsychology*, 24, 343-354. doi:10.1093/arclin/acp039
- Ponton, M. O., & Ardila, A. (1999). The future of neuropsychology with Hispanic populations in the United States. *Archives of Clinical Neuropsychology*, 14, 565-580.
- Proteau, L. (1991). *Le français populaire au Québec et au Canada*. Boucherville, Qc: Publications Proteau.
- Rami, L., Serradell, M., Bosch, B., Caprile, C., Sekler, A., Villar, A., Canal, R., & Molinuevo, J. L. (2008). Normative data for the Boston Naming Test and the Pyramids and Palm Trees Test in the elderly Spanish population. *Journal of Clinical and Experimental Neuropsychology*, 30, 1-6. doi: 10.1080/13803390701743954
- Randolph, C., Lansing, A. E., Ivnik, R. J., Cullum, C. M., & Hermann, B. P. (1999). Determinants of confrontation naming performance. *Archives of Clinical Neuropsychology*, 14, 489-496. doi:10.1016/S0887-6177(98)00023-7
- Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and confrontation naming in children. *Brain and Language*, 71, 267-284. doi:10.1006/brln.1999.2166
- Roberts, P. M. (2001). Aphasia assessment and treatment in bilingual and multicultural populations. In R. Chahey (Ed.), *Language intervention strategies in adult aphasia*. 4th edition (pp. 208-232). Baltimore: Lippincott, Williams & Wilkins.
- Roberts, P. M. (2008). Issues in assessment and treatment for bilingual and culturally diverse patients. In R. Chahey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 245-275). Baltimore: Lippincott Williams & Wilkins.
- Roberts, P. M. & Bois, M. (1998, June). Picture-name agreement for French-English bilingual adults. Poster presented at TENNET XII (Theoretical and Experimental Neuropsychology - Neuropsychologie expérimentale et théorique). Montreal.
- Roberts, P. M., & Bois, M. (1999). Picture name agreement for French-English bilingual adults. *Brain and Cognition*, 40, 238-241.
- Roberts, P. M. & Deslauriers, L. (1999). Picture naming of cognate and non-cognate nouns in bilingual aphasia. *Journal of Communication Disorders*, 32, 1-14. doi:10.1016/S0021-9924(98)00026-4
- Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology*, 16, 635-645. doi:10.1080/02687030244000220
- Rocheffort, G. & Williams, M. (1962). Studies in the development and breakdown of the use of names. *Journal of Neurology, Neurosurgery and Psychiatry*, 25, 222-233.
- Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic level object recognition. *Perception*, 33, 217-236. doi:10.1068/p5117
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 6, 174-215.
- Statistics Canada 2006 Census: <http://www12.statcan.gc.ca/census-recensement/2006/dp-pd/prof/92-591/index.cfm?Lang=E>
- Tallberg, I. M. (2005). The Boston Naming Test in Swedish: Normative data. *Brain and Language*, 94, 19-31. doi:10.1016/j.bandl.2004.11.004
- Thomas, J. C., Fozard, J. L. & Waugh, N. C. (1977). Age-related differences in naming latency. *American Journal of Psychology*, 90, 499-509.
- Thuillard-Colombo, F. & Assal, G. (1992). Adaptation française du test de dénomination de Boston versions abrégées. *Revue Européenne de Psychologie Appliquée*, 42(1), 67-71.
- Tombaugh, T. N. & Hubley, A. M. (1997). The 60 item Boston Naming Test: norms for cognitively intact adults aged 25 to 88 years. *Journal of Clinical and Experimental Neuropsychology*, 19, 922-932. doi: 10.1080/01688639708403773
- Tsolaki, M., Tsantali, E., Lekka, S., Kiosseoglou, S., & Kazis, S. (2003). Can the Boston Naming Test be used as clinical tool for differential diagnosis in dementia. *Brain and Language*, 87, 185-186. doi:10.1016/S0093-934X(03)00262-1
- Welch, L. W., Doineau, D., Johnson, S. & King, D. (1996). Educational and gender normative data for the Boston Naming Test in a group of older adults. *Brain and Language*, 53, 260-266.
- Zec, R. F., Burkett, N. R., Markwell, S. J., Larsen, D. L. (2007a). A cross-sectional study of the effects of age, education, and gender on the Boston Naming Test. *The Clinical Neuropsychologist*, 21, 587-616. doi:

10.1080/13854040701220028

Zec, R. F., Burkett, N.R., Markwell, S.J., & Larsen, D.L. (2007b). Normative data stratified for age, education, and gender on the Boston Naming Test. *The Clinical Neuropsychologist*, 21, 617-637. doi: 10.1080/13854040701339356

Zec, R.F., Markwell, S.J., Burkett, N.R., & Larsen, D.L. (2005). A longitudinal study of confrontation naming in the "normal" elderly. *Journal of the International Neuropsychological Society*, 11, 716-726. doi: 10.1017/S1355617705050897

AUTHOR'S NOTE

Correspondence (in French or English) should be sent to Patricia M. Roberts, Audiology and Speech-Language Pathology, 451 Smyth Road, Room 3071, University of Ottawa, Ottawa, ON., K1H 8M5, Canada. E-mail: proberts@uottawa.ca ►

Received date: March 26, 2010

Accepted date: April 20, 2011

Book Reviews Évaluation des livres

Title: Science of Successful Supervision and Mentorship

Author: Linda Carozza, Ph.D.

Date: 2010

Publisher: Plural Publishing Inc.

Cost: \$75

Reviewer: Ian Roth, MHSc
Toronto Western Hospital



SCIENCE OF SUCCESSFUL SUPERVISION AND MENTORSHIP

Student supervision is a responsibility bestowed on senior clinicians with little or no specific training or preparation. Linda Carozza's *Science of Successful Supervision and Mentorship* is a thorough, practical, and easy-to-digest book that, like a good mentor, not only provides guidance but also nudges the reader toward self-reflection and discovery. The author, who is a Speech-Language Pathologist and Assistant Professor in the Department of Communication Sciences at St. John's University in Queens, NY, states that clinical expertise should not be seen as sufficient when supervising a student or new practitioner. Rather, she believes that formal training in management and conflict resolution, as well as gaining an understanding of different conceptual models of supervision, are mandatory steps en route to developing supervisory skills. She acknowledges that this is a tall order – especially considering that supervisors are doing “double duty” in that they are responsible for the progress of both the client and the student. However, the book takes a pragmatic approach to supervision and mentorship, and the author presents reasonable and realistic approaches to create the “win-win” that she cites as the gold-standard in supervisory relationships. Despite the name of the book, Carozza regards supervision as much an art as a science.

The book is organized in three sections of text followed by 10 appendices. Carozza first identifies common issues to do with supervision and mentorship, such as distance/virtual supervision, and she presents multiple approaches and perspectives. This section lays the theoretical groundwork on which supervisors can build their clinical

practica. She then discusses how mentors and supervisors can develop formal expertise in the area of supervision. In the third section, she provides various practical examples of how to create supportive mentorships. Finally, Carozza dedicates approximately 100 pages (almost a third of the book) to useful appendices, including several performance evaluations that supervisors and supervisees can use to assess each other in the context of a clinical internship.

The premise of this book is that being a good supervisor or mentor begins very similarly to how one might begin a clinical relationship. Carozza argues that good supervision, like good clinical practice, should be grounded in evidence. However, since relevant evidence is not always available for clinical supervisory situations, she suggests beginning with data-driven observations of the student and then moving into more formal interactions using surveys and questionnaires, many examples of which are included in the book. Carozza cites and explains the relevant research that has been compiled by her own group, by other S-LPs, and by other health professions as well as from the corporate world. While she acknowledges that it is not always possible for the average S-LP to heed all her advice because of time constraints, most of what is discussed in the book should be feasible to even the busiest clinician. For example, she has taught her students to be “clinical peers,” offering each other feedback in clinical situations and reducing the need for the supervisor's presence in many situations. Carozza presents different appropriate paths that a supervisor can take in a supervisory relationship based on different issues and angles that may have to be considered.

Carozza stresses the importance of self-reflection. As a supervisor, encouraging a student to be aware of her or his own conversational styles, having “metapragmatic awareness” and using “reflective practice” is invaluable in new clinicians' learning. Reflection allows learners to become aware of their communicative and clinical skills, and therefore places them in a position to improve independently. On the other hand, the author gives the example of a non-learning response, such as “presumption,” which might involve a student not seeking clarification in what would otherwise be a learning opportunity. Rather than discredit students who exhibit this type of response, Carozza uses this example to help the reader see different perspectives on this topic: For example, in some cultures it is seen as impolite to advertise a lack of understanding, and not clarifying may be less about ignorance and more about how that particular student was socialized.

Multiculturalism is an issue that the author revisits frequently in the book because of its growing relevance in even the smallest communities in North America.

She suggests that all mentors should receive training in cultural diversity, and she discusses the different issues one might be exposed to when supervisor and student are from a similar cultural background versus the issues that must be considered when they are from different cultural backgrounds. Carozza examines multiculturalism from various perspectives. She describes the results of a survey that she conducted of S-LP supervisors from different multicultural backgrounds. She also includes a case study and introduces the reader to ASHA's S.T.E.P. (Student To Empowered Professional) Mentorship Program, which allows students – particularly those from racial and ethnic minorities – to find a suitable mentor. This comprehensive treatment of the issue provides context and inspires divergent thinking around a subject that might not get as much attention as it should.

Interestingly, the terms “supervisor” and “mentor” are not used synonymously in this book. The author provides an explanation of how a supervisor (someone who “...seeks to encourage and enhance the development of a practitioner within... an organization”) differs from a mentor (someone who “...seeks to develop the mentee in a much broader life and professional construct, regardless of the particular workplace setting”). A mentor is meant to model behaviour rather than actually teach. However, this distinction is largely academic. For readers who are interested in honing their supervisory skills, the exacting distinction between a supervisor and a mentor is less important.

Carozza recommends for both supervisors and mentors that they be attentive to the clinical learning process rather than be too outcome-driven. It is the process that leads the learner to become an independent decision maker adept at handling new clinical situations. She suggests that a student's self-sufficiency is increased if the supervisor adopts a consultative role, as opposed to that of a teacher or counsellor.

In many ways, the book was very reassuring to read. Thankfully, it was not filled with many time-consuming extra steps that a clinician in a supervisory position could not realistically add to her or his repertoire. Nor did the book suggest that there are finite or fixed ways of providing quality leadership. Rather, the book is characterized by an emphasis on process, organization and vision that, if employed by a seasoned clinician, would allow a learner to discover and refine his or her own professional and clinical skills. The book is broad enough to be appropriate for speech-language pathologists and other health care practitioners, regardless of the population with whom they work. Because it is so inclusive and theoretical, the text does not offer ready-made solutions to some of the practical obstacles to supervision, such as how to recruit

more mentors, how to maximize safety to clients, etc. However, the book exposes the reader to the many paths he or she can take to further explore an issue that is relevant in a specific situation. It is this kind of process-oriented guidance that shows that Linda Carozza is practising what she preaches. ►

Book Reviews Évaluation des livres

Title: Interventions for Speech Sound Disorders in Children

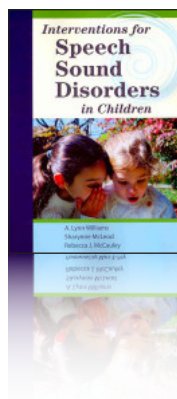
Authors: A. Lynn Williams Ph.D.,
Sharynne McLeod Ph.D.,
Rebecca J. McCauley Ph.D.

Date: 2010

Publisher: Paul H. Brookes
Publishing Co.

Cost: \$79.95

Reviewer: Bojana Radovanovic
University of Toronto, Department of
Speech-Language Pathology



INTERVENTIONS FOR SPEECH SOUND DISORDERS IN CHILDREN

Part of the *Communication and Language Intervention Series*, this edited book is a collection of 25 chapters written by authors from around the world. The editors, all experts in the field of childhood speech disorders, assembled this book with two purposes in mind: “to thoroughly describe, analyze and generally expose readers to a broad spectrum of intervention approaches designed for all types of SSD (speech sound disorders) and what is known about their efficacy; and to provide clinicians with enough information to enable them to select an appropriate intervention for a specific child with whom they are working.”

The editors have divided the book into three major sections. The first, ‘Direct Speech Production Intervention,’ contains seven interventions that directly target a child’s speech production, from reducing homonymy and increasing stimulability to boosting generalizability and establishing consistency. Chapters include ‘Minimal Pair Intervention,’ where Elise Baker outlines the well-known method of pairing words produced as homonyms by the child in an effort to create contrast, and ‘Multiple Oppositions Intervention’ by A. Lynn Williams who describes this variation of the minimal pair contrastive approach, used for children who collapse several target sounds into a single error. Elise Baker and A. Lynn Williams then weigh in together on

‘Complexity Approaches to Intervention,’ introducing the reader to a technique that uses complex targets to expand a child’s knowledge of the phonological system. ‘Core Vocabulary Intervention’ (a chapter written by Barbara Dodd, Alison Holm, Sharon Crosbie and Beth McIntosh) explains the approach used to establish consistency for a child who exhibits inconsistent production of the same lexical item. Raúl F. Prezas and Barbara Williams Hodson report on ‘The Cycles Phonological Remediation Approach,’ which identifies deficiencies in phonological patterns and uses cycles of presentation to increase intelligibility. ‘The Nuffield Centre Dyspraxia Programme,’ by Pam Williams and Hilary Stephens, outlines the British intervention approach based on motor learning theory. The final technique ‘Stimulability Intervention’ is described by A. Lynn Williams and the late Adele W. Miccio as a technique that enhances the restricted phonetic inventories of children through the pairing of consonants with alliterative characters and with hand or body movements.

‘Speech Interventions in Broader Contexts,’ the second section of the book, covers 12 interventions which target speech production with a focus on broader aspects of communication, such as language and literacy. Goals of these interventions include increasing phonological awareness, establishing prosodic structures and expanding a child’s vocabulary. In the ‘Psycholinguistic Intervention’ chapter, Joy Stackhouse and Michelle Pascoe write about the psycholinguistics framework and its use in selecting and interpreting assessment data, and in planning and evaluating interventions. Anne Hesketh then describes ‘Metaphonological Intervention: Phonological Awareness Therapy,’ which uses a child’s phonological awareness to support and drive speech change. The ‘Computer-Based Interventions’ chapter (by Yvonne Wren, Sue Roulstone and A. Lynn Williams) describes how software can increase access to therapy materials and decrease preparation time for busy clinicians. The next chapter, ‘Speech Perception Interventions,’ written by Susan Rvachew and Françoise Brosseau-Lapr , covers interventions aimed at helping the child develop a detailed internal model of the acoustic-phonetic characteristics of the phonemes that are being misarticulated. B. May Bernhardt, Karen D. Bopp, Bonnie Daudlin, Susan M. Edwards and Susan E. Wastie discuss ‘Nonlinear Phonological Intervention’ developed from phonological theories that describe the hierarchical representation of phonological form from the prosodic phrase to the individual feature. ‘Dynamic Systems and Whole Language Intervention’ is a chapter by Paul R. Hoffmann and Janet A. Norris and details interventions within the context of interactive storybook reading. ‘Morphosyntax Intervention,’ as explained by Ann A.

Tyler and Allison M. Haskill, involves targeting speech indirectly through language structures such as grammatical morphemes. The next chapter, Stephen M. Camarata's 'Naturalistic Intervention for Speech Intelligibility and Speech Accuracy,' presents an intervention that focuses on increasing overall intelligibility rather than individual phoneme accuracy. Caroline Bowen contributed 'Parents and Children Together (PACT) Intervention' where intelligibility is targeted through family centred, broad-based approaches. 'Enhanced Milieu Teaching with Phonological Emphasis for Children with Cleft Lip and Palate' is described by Nancy J. Scherer and Ann P. Kaiser as a naturalistic paradigm that both facilitates vocabulary and speech sound production. 'PROMPT: A Tactually Grounded Model' is one of the final interventions in this section, and stands for Prompts for Restructuring Oral Muscular Phonetic Targets. It is explained by Deborah A. Hayden, Jennifer Eigen, Anne Walker and Lisa Olsen as a sensory-motor and cognitive-linguistic model that helps the therapist develop assessment and intervention frameworks. Concluding this section, Nicole Watts Pappas reports on 'Family-Friendly Intervention' and presents a model of modifying any speech intervention for use in a family setting.

The third and final section is entitled 'Interventions for Achieving Speech Movements' and outlines 4 different interventions, which can be used to assist children in developing the articulatory movements. Chapters in this section include 'Visual Feedback Therapy with Electropalatography,' in which Fiona E. Gibbon and Sara E. Wood focus on electropalatography and the tongue-palate contact information it provides as an instrumental feedback device in speech therapy. Chapter 22 (by B. May Bernhardt, Joseph P. Stemberger and Penelope Bacsfalvi) introduces 'Vowel Intervention,' which outlines the use of nonlinear phonological theory and articulatory visual feedback to improve vowel production. Megan M. Hodge describes 'Developmental Dysarthria Interventions' that focus on increasing control of speech muscles as a way of increasing intelligibility in this population. The final intervention discussed, 'Nonspeech Oral Motor Intervention' by Heather M. Clark explores several oral motor exercises (OMEs) that target the functions thought to underlie speech sound production. The volume concludes with a 'Future Directions' chapter where the book editors discuss conclusions and research implications for children with speech sound disorders.

Each of the three sections also includes a table which compares the interventions across 10 factors, including client age, primary population, intervention agents, key interventional components, broad goals, basis of

target selection, level of focus, session type, technology/materials requires and the key codes of the International Classification of Functioning, Disability and Health: Children and Youth Version. Standardized headings across chapters facilitate comparisons between interventions. The subheadings cover topics such as theoretical and empirical basis for the intervention, practical requirements, assessments, special considerations and case studies. The book also includes a supplementary DVD, in which the chapter authors introduce the technique and demonstrated the interventions. The DVD clips are well made and add interest. They further help readers appreciate the differences between the different approaches.

The editors state that their book addresses students of Speech-Language Pathology, clinical practitioners who work with children with speech sound disorders, professors who teach in this field, and parents of children. Though well suited to the first three groups mentioned, this book would be too challenging and detailed a read for most parents. It is a very extensive volume, and parents would probably benefit more from a book specific to the disorder of their child. As well, the language used throughout this book would be too scientific and technical for the average parent.

In summary, this book simultaneously presents well-organized and detailed information about each intervention while providing the reader with the ability to easily and quickly compare interventions on a number of aspects. Overall, this is a great resource, bringing together theory, evidence, step-by-step explanations, multimedia supplementation and more. It would be especially useful for clinicians and researchers working in the field of speech sound disorders. ►

INFORMATION FOR CONTRIBUTORS

The Canadian Journal of Speech-Language Pathology and Audiology (CJSLPA) welcomes submissions of scholarly manuscripts related to human communication and its disorders broadly defined. This includes submissions relating to normal and disordered processes of speech, language, and hearing. Manuscripts that have not been published previously are invited in English and French. Manuscripts may be tutorial, theoretical, integrative, practical, pedagogic, or empirical. All manuscripts will be evaluated on the basis of the timeliness, importance, and applicability of the submission to the interests of speech-language pathology and audiology as professions, and to communication sciences and disorders as a discipline. Consequently, all manuscripts are assessed in relation to the potential impact of the work on improving our understanding of human communication and its disorders. All categories of manuscripts submitted will undergo peer-review to determine the suitability of the submission for publication in CJSLPA. The Journal has established multiple categories of manuscript submission that will permit the broadest opportunity for dissemination of information related to human communication and its disorders. The categories for manuscript submission include:

Tutorials: Review articles, treatises, or position papers that address a specific topic within either a theoretical or clinical framework.

Articles: Traditional manuscripts addressing applied or basic experimental research on issues related to speech, language, and/or hearing with human participants or animals.

Clinical Reports: Reports of new clinical procedures, protocols, or methods with specific focus on direct application to identification, assessment and/or treatment concerns in speech, language, and/or hearing.

Brief Reports: Similar to research notes, brief communications concerning preliminary findings, either clinical or experimental (applied or basic), that may lead to additional and more comprehensive study in the future. These reports are typically based on small “*n*” or pilot studies and must address disordered participant populations.

Research Notes: Brief communications that focus on experimental work conducted in laboratory settings. These reports will typically address methodological concerns and/or modifications of existing tools or instruments with either normal or disordered populations.

Field Reports: Reports that outline the provision of services that are conducted in unique, atypical, or nonstandard settings; manuscripts in this category may include screening, assessment, and/or treatment reports.

Letters to the Editor: A forum for presentation of scholarly/clinical differences of opinion concerning work previously published in the Journal. Letters to the Editor may influence our thinking about design considerations, methodological confounds, data analysis, and/or data interpretation, etc. As with other categories of submissions, this communication forum is contingent upon peer-review. However, in contrast to other categories of submission, rebuttal from the author(s) will be solicited upon acceptance of a letter to the editor.

SUBMISSION OF MANUSCRIPTS

Contributors should use the electronic CJSLPA manuscript submission system at <http://cjslpa.coverpage.ca> to submit articles. If you are unable to use the electronic system, please send a file containing the manuscript, including all tables, figures or illustrations, and references in MS Word or WordPerfect format via e-mail to the Editor at: tim.bressmann@utoronto.ca.

Along with copies of the manuscript, a cover letter indicating that the manuscript is being submitted for publication consideration should be included. The cover letter must explicitly state that the manuscript is original work, that it has not been published previously, and that it is not currently under review elsewhere. Manuscripts are received and peer-reviewed contingent upon this understanding.

The author(s) must also provide appropriate confirmation that work conducted with humans or animals has received ethical review and approval. Failure to provide information on ethical approval will delay the review process. Finally, the cover letter should also indicate the category of submission (i.e., tutorial, clinical report, etc.). If the editorial staff determines

that the manuscript should be considered within another category, the contact author will be notified.

All submissions should conform to the publication guidelines of the Publication Manual of the American Psychological Association (APA), 6th Edition. A confirmation of receipt for all manuscripts will be provided to the contact author prior to distribution for peer review. CJSLPA seeks to conduct the review process and respond to authors regarding the outcome of the review within 90 days of receipt. If a manuscript is judged as suitable for publication in CJSLPA, authors will have 30 days to make necessary revisions prior to a secondary review.

The author is responsible for all statements made in his or her manuscript, including changes made by the editorial and/or production staff. Upon final acceptance of a manuscript and immediately prior to publication, the contact author will be permitted to review galley proofs and verify its content to the publication office within 72 hours of receipt of galley proofs.

ORGANIZATION OF THE MANUSCRIPT

All copies should be typed, double-spaced, with a standard typeface (12 point, noncompressed font) on high quality 8 ½ X 11 paper. All margins should be at least one (1) inch. An electronic copy of the manuscript should be submitted directly to the editor. Author identification for the review process is optional; if blind-review is desired, the documents should be prepared accordingly (cover page and acknowledgments blinded). Responsibility for removing all potential identifying information rests solely with the author(s). All submissions should conform to the publication guidelines of the most current edition of the Publication Manual of the American Psychological Association (APA). The APA manual is available from most university and commercial bookstores. Generally, the following sections should be submitted in the order specified.

Title Page: This page should include the full title of the manuscript, the full names of the author(s) with academic degrees, each author's affiliation, and a complete mailing address for the contact author. An electronic mail address also is recommended.

Abstract: On a separate sheet of paper, a brief yet informative abstract that does not exceed one page is required. The abstract should include the purpose of the work along with pertinent information relative to the specific manuscript category for which it was submitted.

Key Words: Following the abstract and on the same page, the author(s) should supply a list of key words for indexing purposes.

Tables: Each table included in the manuscript must be typewritten double-spaced and placed at the end of the document. Tables should be numbered consecutively beginning with Table 1. Each table must have a descriptive caption. Tables should serve to expand the information provided in the text of the manuscript, not to duplicate information.

Potential Conflicts of Interest and Dual Commitment

As part of the submission process, the author(s) must explicitly identify if any potential conflict of interest or dual commitment exists relative to the manuscript and its author(s). Such disclosure is requested so as to inform CJSIPA that the author or authors have the potential to benefit from publication of the manuscript. Such benefits may be either direct or indirect and may involve financial and/or other nonfinancial benefit(s) to the author(s). Disclosure of potential conflicts of interest or dual commitment may be provided to editorial consultants if it is believed that such a conflict of interest or dual commitment may have had the potential to influence the information provided in the submission or compromise the design, conduct, data collection or analysis, and/or interpretation of the data obtained and reported in the manuscript submitted for review. If the manuscript is accepted for publication, editorial acknowledgement of such potential conflict of interest or dual commitment may occur within the publication.

Illustrations: All illustrations to be included as part of the manuscript must also be submitted in their original file format separate from the manuscript. High resolution (at least 300 dpi) files in any of the following formats must be submitted for each graphic and image: JPEG, TIFF, AI, PSD, GIF, EPS or PDF. For other types of computerized illustrations, it is recommended that CJSIPA production staff be consulted prior to preparation and submission of the manuscript and associated figures/illustrations.

Legends for Illustrations: Legends for all figures and illustrations should be typewritten (double-spaced) on a separate page with numbers corresponding to the order in which figures/illustrations appear in the manuscript.

Page Numbering and Running Head: The text of the manuscript should be prepared with each page numbered, including tables, figures/illustrations, references, and appendices. A short (30 characters or less) descriptive running title should appear at the top right hand margin of each page of the manuscript.

Acknowledgments: Acknowledgments should be typewritten (double-spaced) on a separate page. Appropriate acknowledgment for any type of sponsorship, donations, grants, technical assistance, and to professional colleagues who contributed to the work, but are not listed as authors, should be noted.

References: References are to be listed consecutively in alphabetical order, then chronologically for each author. Authors should consult the most current edition of the APA publication manual for methods of citing varied sources of information. Journal names and appropriate volume number should be spelled out and italicized. All literature, tests and assessment tools, and standards (ANSI and ISO) must be listed in the references. All references should be double-spaced.

Participants in Research Humans and Animals

Each manuscript submitted to CJSIPA for peer-review that is based on work conducted with humans or animals must acknowledge appropriate ethical approval. In instances where humans or animals have been used for research, a statement indicating that the research was approved by an institutional review board or other appropriate ethical evaluation body or agency must clearly appear along with the name and affiliation of the research ethics and the ethical approval number. The review process will not begin until this information is formally provided to the Editor.

Similar to research involving human participants, CJSIPA requires that work conducted with animals state that such work has met with ethical evaluation and approval. This includes identification of the name and affiliation of the research ethics evaluation body or agency and the ethical approval number. A statement that all research animals were used and cared for in an established and ethically approved manner is also required. The review process will not begin until this information is formally provided to the Editor.

RENSEIGNEMENTS À L'INTENTION DES COLLABORATEURS

La Revue canadienne d'orthophonie et d'audiologie (RCOA) est heureuse de se voir soumettre des manuscrits de recherche portant sur la communication humaine et sur les troubles qui s'y rapportent, dans leur sens large. Cela comprend les manuscrits portant sur les processus normaux et désordonnés de la parole, du langage et de l'audition. Nous recherchons des manuscrits qui n'ont jamais été publiés, en français ou en anglais. Les manuscrits peuvent être tutoriels, théoriques, synthétiques, pratiques, pédagogiques ou empiriques. Tous les manuscrits seront évalués en fonction de leur signification, de leur opportunité et de leur applicabilité aux intérêts de l'orthophonie et de l'audiologie comme professions, et aux sciences et aux troubles de la communication en tant que disciplines. Par conséquent, tous les manuscrits sont évalués en fonction de leur incidence possible sur l'amélioration de notre compréhension de la communication humaine et des troubles qui s'y rapportent. Peu importe la catégorie, tous les manuscrits présentés seront soumis à une révision par des collègues afin de déterminer s'ils peuvent être publiés dans la RCOA. La Revue a établi plusieurs catégories de manuscrits afin de permettre la meilleure diffusion possible de l'information portant sur la communication humaine et les troubles s'y rapportant. Les catégories de manuscrits comprennent :

Tutoriels : Rapports de synthèse, traités ou exposés de position portant sur un sujet particulier dans un cadre théorique ou clinique.

Articles : Manuscrits conventionnels traitant de recherche appliquée ou expérimentale de base sur les questions se rapportant à la parole, au langage ou à l'audition et faisant intervenir des participants humains ou animaux.

Comptes rendus cliniques : Comptes rendus de nouvelles procédures ou méthodes ou de nouveaux protocoles cliniques

portant particulièrement sur une application directe par rapport aux questions d'identification, d'évaluation et de traitement relativement à la parole, au langage et à l'audition.

Comptes rendus sommaires : Semblables aux notes de recherche, brèves communications portant sur des conclusions préliminaires, soit cliniques soit expérimentales (appliquées ou fondamentales), pouvant mener à une étude plus poussée dans l'avenir. Ces comptes rendus se fondent typiquement sur des études à petit « n » ou pilotes et doivent traiter de populations désordonnées.

Notes de recherche : Brèves communications traitant spécifiquement de travaux expérimentaux menés en laboratoire. Ces comptes rendus portent typiquement sur des questions de méthodologie ou des modifications apportées à des outils existants utilisés auprès de populations normales ou désordonnées.

Comptes rendus d'expérience : Comptes rendus décrivant sommairement la prestation de services offerts en situations uniques, atypiques ou particulières; les manuscrits de cette catégorie peuvent comprendre des comptes rendus de dépistage, d'évaluation ou de traitement.

Courrier des lecteurs : Forum de présentation de divergences de vues scientifiques ou cliniques concernant des ouvrages déjà publiés dans la Revue. Le courrier des lecteurs peut avoir un effet sur notre façon de penser par rapport aux facteurs de conception, aux confusions méthodologiques, à l'analyse ou l'interprétation des données, etc. Comme c'est le cas pour d'autres catégories de présentation, ce forum de communication est soumis à une révision par des collègues. Cependant, contrairement aux autres catégories, on recherchera la réaction des auteurs sur acceptation d'une lettre.

PRÉSENTATION DE MANUSCRITS

Pour soumettre un article, les auteurs doivent utiliser le système de soumission électronique de l'ACOA à l'adresse <http://cjslpa.coverpage.ca>. Si vous ne pouvez pas utiliser le système électronique, veuillez envoyer par courriel un fichier Word ou WordPerfect contenant le manuscrit, y compris tous les tableaux, les figures ou illustrations et la bibliographie. Adressez le courriel au rédacteur en chef à l'adresse tim.bressmann@utoronto.ca.

On doit joindre aux exemplaires du manuscrit une lettre d'envoi qui indiquera que le manuscrit est présenté en vue de sa publication. La lettre d'envoi doit préciser que le manuscrit est une œuvre originale, qu'il n'a pas déjà été publié et qu'il ne fait pas actuellement l'objet d'un autre examen en vue d'être publié. Les manuscrits sont reçus et examinés sur acceptation de ces conditions. L'auteur (les auteurs) doit (doivent) aussi fournir une attestation en bonne et due forme que toute recherche impliquant des êtres humains ou des animaux a fait l'objet de l'agrément d'un comité de révision déontologique. L'absence d'un tel agrément retardera le processus de révision. Enfin, la lettre d'envoi doit également préciser la catégorie de la présentation (i.e. tutoriel, rapport clinique, etc.). Si l'équipe d'examen juge que le manuscrit devrait passer sous une autre catégorie, l'auteur-contact en sera avisé.

Toutes les présentations doivent se conformer aux lignes de conduite présentées dans la publication *Manual of the American Psychological Association* (APA), 6^e Édition. Un accusé de réception de chaque manuscrit sera envoyé à l'auteur-contact avant la distribution des exemplaires en vue de la révision. La RCOA cherche à effectuer cette révision et à informer les auteurs des résultats de cette révision dans les 90 jours de la réception. Lorsqu'on juge que le manuscrit convient à la RCOA, on donnera 30 jours aux auteurs pour effectuer les changements nécessaires avant l'examen secondaire.

L'auteur est responsable de toutes les affirmations formulées dans son manuscrit, y compris toutes les modifications effectuées par les rédacteurs et réviseurs. Sur acceptation définitive du manuscrit et immédiatement avant sa publication, on donnera l'occasion à l'auteur-contact de revoir les épreuves et il devra signifier la vérification du contenu dans les 72 heures suivant réception de ces épreuves.

ORGANISATION DU MANUSCRIT

Tous les textes doivent être écrits à double interligne, en caractère standard (police de caractères 12 points, non comprimée) et sur papier 8 ½" X 11" de qualité. Toutes les marges doivent être d'au moins un (1) pouce. Un fichier électronique du manuscrit doit être présenté directement au rédacteur en chef. L'identification de l'auteur est facultative pour le processus d'examen : si l'auteur souhaite ne pas être identifié à ce stade, il devra préparer un fichier électronique dont la page couverture et les remerciements seront voilés. Seuls les auteurs sont responsables de retirer toute information identificatrice éventuelle. Tous les manuscrits doivent être rédigés en conformité aux lignes de conduite les plus récentes de l'APA. Ce manuel est disponible dans la plupart des librairies universitaires et commerciaux. En général, les sections qui suivent doivent être présentées dans l'ordre chronologique précisé.

Page titre : Cette page doit contenir le titre complet du manuscrit, les noms complets des auteurs, y compris les diplômes et affiliations, l'adresse complète de l'auteur-contact et l'adresse de courriel de l'auteur contact.

Abrégé : Sur une page distincte, produire un abrégé bref mais informateur ne dépassant pas une page. L'abrégé doit indiquer l'objet du travail ainsi que toute information pertinente portant sur la catégorie du manuscrit.

Mots clés : Immédiatement suivant l'abrégé et sur la même page, les auteurs doivent présenter une liste de mots clés aux fins de constitution d'un index.

Tableaux : Tous les tableaux compris dans un même manuscrit doivent être écrits à double interligne sur une page distincte. Les tableaux doivent être numérotés consécutivement, en commençant par le Tableau 1. Chaque tableau doit être accompagné d'une légende et doit servir à compléter les renseignements fournis dans le texte du manuscrit plutôt qu'à reprendre l'information contenue dans le texte ou dans les tableaux.

Conflits d'intérêts possibles et engagement double

Dans le processus de présentation, les auteurs doivent déclarer clairement l'existence de tout conflit d'intérêts possibles ou engagement double relativement au manuscrit et de ses auteurs. Cette déclaration est nécessaire afin d'informer la RCOA que l'auteur ou les auteurs peuvent tirer avantage de la publication du manuscrit. Ces avantages pour les auteurs, directs ou indirects, peuvent être de nature financière ou non financière. La déclaration de conflit d'intérêts possibles ou d'engagement double peut être transmise à des conseillers en matière de publication lorsqu'on estime qu'un tel conflit d'intérêts ou engagement double aurait pu influencer l'information fournie dans la présentation ou compromettre la conception, la conduite, la collecte ou l'analyse des données, ou l'interprétation des données recueillies et présentées dans le manuscrit soumis à l'examen. Si le manuscrit est accepté en vue de sa publication, la rédaction se réserve le droit de reconnaître l'existence possible d'un tel conflit d'intérêts ou engagement double.

Illustrations : Toutes les illustrations faisant partie du manuscrit doivent être annexer avec chaque exemplaire du manuscrit. Chaque manuscrit doit être accompagné d'un fichier électronique pour chaque image et graphique en format JPEG, TIFF, AI, PSD, GIF, EPS ou PDF, compression minimale 300 ppp. Pour les autres types d'illustrations informatisées, il est recommandé de consulter le personnel de production de la RCOA avant la préparation et la présentation du manuscrit et des figures et illustrations s'y rattachant.

Légendes des illustrations : Les légendes accompagnant chaque figure et illustration doivent être écrits à double interligne sur une page distincte et identifiées à l'aide d'un numéro qui correspond à la séquence de parution des figures et illustrations dans le manuscrit.

Numérotation des pages et titre courant : Chaque page du manuscrit doit être numérotée, y compris les tableaux, figures, illustrations, références et, le cas échéant, les annexes. Un bref (30 caractères ou moins) titre courant descriptif doit apparaître dans la marge supérieure droite de chaque page du manuscrit.

Remerciements : Les remerciements doivent être écrits à double interligne sur une page distincte. L'auteur doit reconnaître toute forme de parrainage, don, bourse ou d'aide technique, ainsi que tout collègue professionnel qui ont contribué à l'ouvrage mais qui n'est pas cité à titre d'auteur.

Références : Les références sont énumérées les unes après les autres, en ordre alphabétique, suivi de l'ordre chronologique sous le nom de chaque auteur. Les auteurs doivent consulter le manuel de l'APA le plus récent pour obtenir la façon exacte de rédiger une citation. Les noms de revues scientifiques et autres doivent être rédigés au long et imprimés en italiques. Tous les ouvrages, outils d'essais et d'évaluation ainsi que les normes (ANSI et ISO) doivent figurer dans la liste de références. Les références doivent être écrits à double interligne.

Participants à la recherche – êtres humains et animaux

Chaque manuscrit présenté à la RCOA en vue d'un examen par des pairs et qui se fonde sur une recherche effectuée avec la participation d'être humains ou d'animaux doit faire état d'un agrément déontologique approprié. Dans les cas où des êtres humains ou des animaux ont servi à des fins de recherche, on doit joindre une attestation indiquant que la recherche a été approuvée par un comité d'examen reconnu ou par tout autre organisme d'évaluation déontologique, comportant le nom et l'affiliation de l'éthique de recherche ainsi que le numéro de l'approbation. Le processus d'examen ne sera pas amorcé avant que cette information ne soit formellement fournie au rédacteur en chef.

Tout comme pour la recherche effectuée avec la participation d'être humains, la RCOA exige que toute recherche effectuée avec des animaux soit accompagnée d'une attestation à l'effet que cette recherche a été évaluée et approuvée par les autorités déontologiques compétentes. Cela comporte le nom et l'affiliation de l'organisme d'évaluation de l'éthique en recherche ainsi que le numéro de l'approbation correspondante. On exige également une attestation à l'effet que tous les animaux de recherche ont été utilisés et soignés d'une manière reconnue et éthique. Le processus d'examen ne sera pas amorcé avant que cette information ne soit formellement fournie au rédacteur en chef. ►



Newfoundland Labrador

CASLPA 2012

to be held in St. John's, Newfoundland and Labrador, promises to be an invaluable professional development opportunity. With a fantastic list of speakers and events to choose from, this conference will be a memorable one.

St. John's, one of the oldest cities in North America, is located on the most easterly tip of the continent, quickly helping one realize that land, sea and sky truly do embrace like old friends. Its picturesque landscape, winding coastlines, and rugged terrain offer incredible scenery for nature lovers. From bird-watching, to boat tours, to the possibility of spotting 10,000 year-old glacial giants in the Atlantic Ocean, there will never be a disappointing moment.

While embracing fellow colleagues and catching up with old friends between conference speakers, enjoy some valuable personal time shopping along the quaint streets of downtown St. John's or walking through beautiful Bowering Park.

Feeling hungry? St. John's offers exciting and diverse dining experiences. Your taste buds will surely be tantalized as you try local delights including cod tongues, touts, and fish-and-brewis. On your way, take a stroll along the two blocks of bars and pubs which make up George Street; the centre of the city's entertainment district with a unique reputation that has spread far and wide!



One social event you can take advantage of is a *Rally in the Alley*, a wonderful opportunity to experience a wide range of establishments on George Street. Start the evening with supper with the crowd, followed by a tour of many venues along the infamous street!

To find out more about what is happening in St. John's, and within our beautiful province, please visit the following websites:

www.newfoundlandlabrador.com

www.stjohnskiosk.com



ST. JOHN'S, NEWFOUNDLAND
MAY 9-12, 2012
DELTA ST. JOHN'S

Join Us!

Planning is in full swing as we prepare for an unforgettable conference; so don't forget to visit the **CASLPA website** for frequent updates on confirmed speakers, program details and registration information!





CASLPA-ACOA

© 2011, CASLPA