



HAL
open science

Une méthode d'évaluation de la compréhension orale par choix d'image : application à de la parole dégradée par simulation de la presbyacousie

Cynthia Magnen, Julien Tardieu, Lionel Fontan, Pascal Gaillard, Nathalie
Spanghero-Gaillard

► To cite this version:

Cynthia Magnen, Julien Tardieu, Lionel Fontan, Pascal Gaillard, Nathalie Spanghero-Gaillard. Une méthode d'évaluation de la compréhension orale par choix d'image : application à de la parole dégradée par simulation de la presbyacousie. 31ème édition des Journées d'études sur la parole (JEP 2016), Jun 2016, Paris, France. hal-01572462

HAL Id: hal-01572462

<https://hal.science/hal-01572462>

Submitted on 8 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Une méthode d'évaluation de la compréhension orale par choix d'image : application à de la parole dégradée par simulation de la presbyacousie

Cynthia Magnen¹ Julien Tardieu¹ Lionel Fontan² Pascal Gaillard³ Nathalie Spanghero-Gaillard⁴

(1) MSHS-T USR3414, Université de Toulouse et CNRS, Toulouse, France

(2) ARCHEAN Technologies, Montauban, France

(3) CLLE UMR5263, Université de Toulouse et CNRS, Toulouse, France

(4) LERASS-CPST EA827, Université de Toulouse, Toulouse, France

cynthia.magnen@univ-tlse2.fr, julien.tardieu@univ-tlse2.fr,
lfontan@archean.fr, pascal.gaillard@univ-tlse2.fr,
nathalie.spanghero@univ-tlse2.fr

RESUME

Nous présentons une méthode permettant d'évaluer la compréhension de la parole dégradée par simulation des effets de la presbyacousie, dans le calme et dans le bruit. Cette méthode intègre des phrases signifiantes et implique pour l'auditeur de sélectionner, parmi un ensemble de quatre images, celle qui correspond à l'énoncé qu'il entend. Le test présente de nombreux avantages méthodologiques comme l'immédiateté du score et le fait qu'il ne nécessite pas de faire répéter la phrase entendue. Les résultats obtenus montrent un effet significatif de la dégradation et du bruit du fond. La cohérence de ces effets avec les études précédentes sur la presbyacousie permet de valider cette méthode. Par ailleurs, la nature exacte du score mesuré dans ce test est discutée en le comparant avec le score d'intelligibilité obtenu par répétition d'items dans une précédente étude.

ABSTRACT

A method for assessing listening comprehension using image selection : application to speech degraded by presbycusis simulation.

This paper presents a method for assessing the comprehension of speech degraded by presbycusis simulation in both quiet and babble noise. This method uses meaningful sentences and asks the participant to select, among four images, the one that best corresponds to what he/she hears. This test has many methodological advantages such as : immediate score and the participant is not asked to repeat the sentence. The results show a significant effect of both the degradation and the background noise. These effects are coherent with previous studies on presbycusis which provides a validation of our method. In addition, the exact nature of the score measured in this test is discussed by comparing it to the intelligibility score measured with repeated sentences in a previous study.

MOTS-CLES : Presbyacousie, intelligibilité, compréhension, perception, parole dans le bruit.

KEYWORDS: Presbycusis, intelligibility, comprehension, perception, speech in noise.

1 Introduction

On pense souvent, et assez légitimement, qu'une bonne audition est garante de la compréhension orale d'une situation de communication. De sorte que si l'on veut étudier les processus sous-jacents à la compréhension, il semble évident de décrire en premier lieu les mécanismes liés à l'audition. Dans l'étude des surdités, par exemple, afin de mesurer les pertes auditives des patients et le gain apporté par l'appareillage auditif, les audioprothésistes évaluent l'intelligibilité de la parole en recourant à des tests de perception de mots ou de phrases. Or, plusieurs travaux ont contribué à montrer que la perception auditive ne peut être vue comme une simple activité d'identification. Les études sur les phénomènes de restauration phonémique (Sivonen *et al.*, 2006) et d'illusions sémantiques (Nieuwland, Van Berkum, 2005) ont notamment traduit l'existence d'attentes perceptives chez l'auditeur en montrant que des processus « descendants » (*top-down*, du sens vers la forme) influencent les traitements dits « de plus bas niveau ». C'est-à-dire qu'une information de nature auditive est identifiable par un auditeur si elle intégrée dans la dynamique d'une situation et donc, déjà saisie dans un projet d'utilisation.

1.1 La mesure de l'intelligibilité de la parole en audiométrie vocale

Dans le domaine de la réhabilitation auditive, les capacités de discrimination phonémique et de compréhension du langage de patients atteints de troubles auditifs sont mesurées grâce à un examen d'audiométrie vocale. Cet examen est complémentaire de l'audiométrie tonale qui détecte les seuils absolus de perception de sons purs. En plus de s'intéresser aux répercussions sociales de la perte auditive, l'audiométrie vocale permet de différencier entre les distorsions d'origine endocochléaire et les atteintes centrales de l'audition. Le diagnostic ainsi établi permet de choisir une technique de réhabilitation adaptée (Bouccara *et al.* 2005). En seconde intention, le praticien juge l'efficacité des moyens de compensation mis en place tels que la prothèse ou l'implant cochléaire par un nouvel examen (Bonfils, Avan, 2005).

Les tests utilisés dans cette optique permettent d'établir un score d'intelligibilité en cotant en pourcentage le taux de reconnaissance du matériel verbal présenté au sein d'une liste (mots, phrases, logatomes). Selon les pratiques, les listes sont soit enregistrées, soit énoncées par l'expérimentateur. Au cours de ces tests, on demande aux patients de répéter des listes de stimuli sonores émis à différentes intensités par voie aérienne ou par voie osseuse. Ainsi, sont décrits : 1) le seuil d'intelligibilité, intensité à partir de laquelle 50 % des mots sont compris, 2) le maximum d'intelligibilité, pourcentage maximum de mots compris et 3) le pourcentage de discrimination, pourcentage de stimuli sonores compris à 35 dB au dessus du seuil d'intelligibilité. L'audiométrie vocale peut aussi se réaliser avec l'adjonction de bruit perturbant afin de confirmer les scores obtenus pour les tests en cabine (les scores étant souvent inférieurs dans le bruit).

1.2 La diversité du matériel linguistique

Le matériel sonore utilisé est très variable. Il peut s'agir de listes de sons complexes le plus souvent signifiants et possiblement non signifiants. Les plus utilisées en France sont les listes de mots dissyllabiques de Fournier (1951) et les listes de mots de 3 phonèmes de Lafon (1964). En proposant la répétition erronée du mot comme unité d'erreur, les listes de Fournier ont pour objectif de mesurer l'intelligibilité. Les listes de Lafon mesurent quant à elles l'identification phonémique par un comptage du nombre de phonèmes perçus ou erronés. De nombreuses critiques

ont été formulées sur l'élaboration de ces listes. Parmi elles, nous pouvons évoquer la désuétude du vocabulaire, le français ayant beaucoup évolué en une cinquantaine d'années, la fréquence d'occurrence des mots employés s'en trouve modifiée (Garnier *et al.*, 1997). L'ordonnement interne des items a également été soulevé puisque la présentation successive de différents termes peut induire des effets d'amorçage sémantique (Estienne, Piérart, 2006). Enfin, la qualité de l'équilibrage phonétique dans les listes de Fournier (Lafon, 1964), comme l'inégalité de longueur des stimuli pour les listes de Lafon et le manque de reproductibilité des tests (Garnier *et al.*, 1997) sont autant de limites évoquées dans l'utilisation clinique de ce matériel.

Jusqu'à il y a peu, le contexte phrastique était moins utilisé que le mot pour évaluer l'intelligibilité de la parole. Paradoxalement, les effets de la suppléance mentale étaient considérés comme un inconvénient dans l'évaluation des capacités auditives réelles. D'autant que plus l'unité traitée est grande plus la suppléance mentale est influente dans la réalisation de la tâche (Garnier *et al.* 1997). Aujourd'hui, l'utilisation de la parole dans ce contexte gagne en popularité, car elle est plus représentative des situations de communication quotidiennes (indices contextuels et bruit de fond) que les listes de mots isolés diffusées en milieu calme. Les tests de répétition de phrases semblent donc plus valides du point de vue de la prédictibilité (Schoepflin, 2015). Le plus populaire est le test HINT (pour *Hearing in Noise Test* - Nilsson *et al.*, 1994) qui évalue la reconnaissance de 250 phrases (25 listes de 10 phrases) variant de 3 à 7 mots; le score est calculé à partir du nombre total de mots correctement identifiés dans les phrases.

1.3 Objectif de l'étude

L'objectif de l'étude est de présenter une nouvelle méthode permettant d'évaluer la compréhension de la parole dégradée par simulation des effets de la presbycusie, dans le calme et dans le bruit. Cette méthode intègre des phrases signifiantes et propose aux participants d'apparier ces phrases à des images qui les illustrent. Afin de valider notre méthode, nous comparerons les scores de compréhension par choix d'images aux scores obtenus dans une étude précédente avec une tâche de répétition de phrases (*cf.* Fontan *et al.*, 2014). L'étude montrait notamment une baisse significative du score de reconnaissance des phrases en fonction de la dégradation et du bruit de fond.

2 Le test de compréhension de la parole

2.1 Choix d'un test d'appariement phrase/image

Afin d'évaluer la compréhension de la parole dégradée, nous avons choisi d'élaborer un test d'appariement énoncé oral / image. Ce type de test est déjà utilisé dans une visée diagnostique pour évaluer la capacité de patients aphasiques à traiter des énoncés dont la complexité syntaxique va croissant (*cf.* Paradis, 1989), ainsi qu'en didactique des langues pour mesurer le niveau de compréhension d'un apprenant de langue seconde (Spanghero-Gaillard, 2008). La tâche proposée implique pour l'auditeur de sélectionner, parmi un ensemble de quatre images, celle qui correspond à l'énoncé qu'il entend. Elle présente l'avantage de proposer un contexte commun au locuteur et à l'auditeur représenté par un ensemble visuel de référents possibles (le jeu d'images). Le locuteur y fait référence dans son énoncé, référence à laquelle doit répondre l'auditeur par une tâche de sélection. En adaptant ce test à la compréhension de la parole dégradée, nous ne testons

plus la capacité à répéter des mots ou des phrases. Nous nous intéressons à la compréhension d'un énoncé dégradé dans une situation de communication s'approchant davantage de la vie réelle, c'est-à-dire où le contexte n'est pas figé mais sujet à interprétation. Ce cadre est avantageux mais il n'est pas évident à mettre en place car il exige de prédéfinir un ensemble de réponses limité. Les critères retenus pour constituer le matériel expérimental sont présentés ci-après.

2.2 Élaboration du matériel linguistique

L'élaboration des variantes de l'énoncé n'est pas une tâche triviale. Il s'agit de définir des ambiguïtés reflétant la perception de l'énoncé dégradé par simulation des effets de la presbycusie. Pour résoudre cette difficulté, nous avons utilisé les données d'un test précédent (cf. Test 2 dans Fontan *et al.*, 2014). Le test consistait à faire répéter des phrases issues de la version française du test HINT (Vaillancourt *et al.*, 2005) à une trentaine d'auditeurs francophones (âgés entre 18 et 30 ans). Les phrases étaient dégradées selon les mêmes conditions que pour cette présente étude (cf. section 3.3). Les résultats constituent un corpus d'énoncés soit conformes d'un point de vue phonétique et sémantique, soit déviants par rapport à l'énoncé de base. Au sein de ce corpus, nous avons sélectionné les 33 phrases qui pouvaient être facilement illustrées (ex.: « L'oiseau s'envole du nid »). Puis, pour chacune des phrases sélectionnées, nous avons relevé les trois énoncés déviants les plus fréquemment cités (ex.: « L'oiseau s'envole de nuit », « Les oiseaux s'envolent du nid », « Les oiseaux sont près du nid »). En fonction du niveau de dégradation entendu, les énoncés produits pouvaient donner lieu à des scènes irréelles plutôt rigolotes et incongrues (ex. pour « Les dragons crachent du feu »: « Le dragon crache du feu », « Les dragons crachent des pneus », « Le chaton crache du feu »). En tout, le corpus audio est constitué de 132 phrases assertives (3 variantes de 33 énoncés, pour 12 phrases d'entraînement et 120 phrases de test).

2.3 Représentations imagées des énoncés : conception, réalisation et accord

Les phases de conception et de réalisation des représentations imagées des 132 énoncés ont été menées en étroite collaboration avec deux graphistes professionnels (Jérémy Villy¹ et Olivier Subra²). Des croquis réalisés par nos soins leur ont été soumis en première intention afin de préciser le cahier des charges. D'un point de vue de la réalisation, il était important que le participant puisse interpréter le sens du dessin sans qu'il n'y ait d'ambiguïté sur la signification graphique. Les illustrateurs ont donc choisi de travailler en ligne claire, avec des aplats de couleur en reprenant les codes historiques de la bande dessinée, popularisés notamment avec Tintin, Lucky Luke ou Spirou. Un pré-test a été mené auprès de cinq personnes francophones natives âgées entre 20 et 30 ans. Ce pré-test a consisté à appairer chacune des phrases écrites aux 4 possibilités visuelles. Chaque phrase était présentée en même temps que les 4 images correspondantes sur un écran informatique via une interface dédiée. Une fois l'image choisie, les participants évaluaient leur réponse sur une échelle indiciaire de confiance allant de 1 à 7. Lorsque l'indice de confiance était en deçà de 5, le participant justifiait son choix par un commentaire libre écrit. Le matériel visuel a été modifié en fonction des résultats de ce test et grâce au recueil des commentaires afin qu'aucune ambiguïté ne subsiste quant à l'appariement phrase/image.

¹ <http://jeremyvilly.com/>

² <http://www.admarginem.fr/>

3 Méthode

3.1 Participants

Trente participants ont été sélectionnés et rémunérés pour cette expérience. Le profil des participants répondait aux critères d'inclusion suivants : étudiants francophones natifs, âgés de 18 à 30 ans inclus, sans problème de vue non corrigé par des lentilles ou des lunettes. Le niveau d'audition de chaque participant a été vérifié par un audiogramme tonal (critère d'inclusion : moyenne des pertes entre 2kHz et 8kHz < 15 dB).

3.2 Stimuli

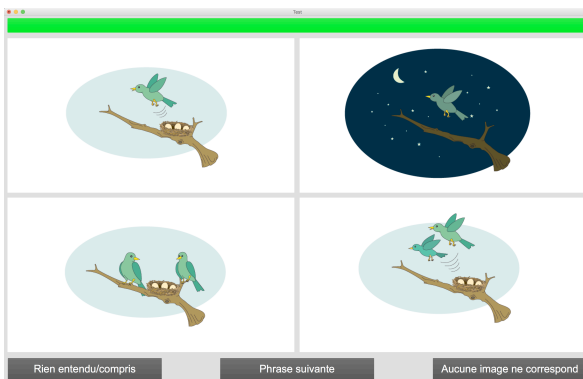


FIGURE 1. Exemple de fenêtre présentant quatre images ainsi que les boutons de réponses pour le test de compréhension de phrases.

Audio. L'ensemble des 132 phrases (*cf.* section 2.2 pour la description du matériel linguistique) a été enregistré auprès de deux locuteurs adultes (un homme et une femme de 46 et 47 ans) et d'une enfant (12 ans). L'enregistrement a eu lieu dans la cabine audiométrique PETRA (<http://petra.univ-tlse2.fr>), avec un microphone omnidirectionnel Sennheiser MD46, une console de mixage TASCAM DM-3200 et un ordinateur MacPro équipé du logiciel Reaper. Les phrases d'une durée moyenne de 1273 ms (écart type = 186 ms) ont ensuite été égalisées en sonie par un panel d'auditeurs puis mixées avec un bruit de fond vocal de type « brouhaha » et un rapport signal sur bruit de 5 dBA. Les phrases avec et sans bruit de fond ont enfin été dégradées par simulation de la presbycusie suivant une variation en dix niveaux. Le premier niveau correspond au signal non dégradé et les 9 niveaux suivant vont d'un âge théorique de 60 ans à 110 ans. La procédure de traitement de signal est détaillée dans Fontan *et al.* (2014).

Visuels. Les images décrites dans la section 2.3 sont présentées par planches de quatre images de 820x440 pixels sur un écran Apple Cinema Display de 20 pouces avec une résolution de 1680x1050 pixels. La fenêtre présentant les quatre images et les boutons de réponse occupe tout l'écran et est illustrée en Figure 1.

3.3 Procédure

Le test de compréhension a été effectué dans la cabine audiométrique PETRA (cf. section 3.3). Les participants étaient assis à un mètre des haut-parleurs (Tannoy Precision 6D), avec un niveau de diffusion des phrases non dégradées (condition 0) de 60 dBA et un niveau de bruit de fond vocal pour les phrases non dégradées à 55 dBA. L'interface a été créée avec le logiciel Max/MSP permettant la présentation concomitante des stimuli visuels et audio (voir Figure 1). Après un entraînement sur 12 phrases, les participants entendaient les 120 phrases de test. Ils avaient pour consigne de sélectionner l'image correspondant à la phrase entendue. Ils étaient également averti que les énoncés qu'ils entendraient pouvaient être totalement incongrus. Une barre latérale verte s'allumait en haut de l'écran pour indiquer la diffusion de la phrase. Dès lors que la barre latérale était active, ils devaient répondre le plus rapidement possible, avec la possibilité de répondre avant la fin de la diffusion de la phrase (les temps de réaction sont mesurés à partir du début de chaque phrase). Un bouton « Phrase suivante » permettait de passer à une nouvelle phrase. Si les participants jugeaient que le signal de parole n'était pas compréhensible ou audible, ils pouvaient sélectionner le bouton « rien entendu/compris ». Si aucune image ne correspondait à ce qu'ils avaient perçu, les participants pouvaient sélectionner le bouton « Aucune image ne correspond ». L'enregistrement des temps de réaction et des réponses choisies par les participants a été effectué via le logiciel de passation dédié. Le plan expérimental comporte 2 facteurs intra sujet : la dégradation (10 niveaux) et la condition de bruit de fond (2 niveaux).

4 Résultats

4.1 Scores de compréhension

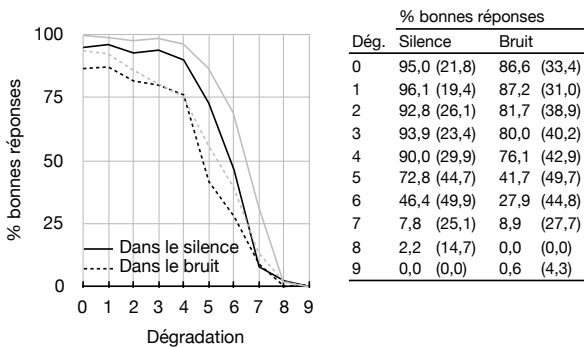


FIGURE 2. Pourcentages de bonnes réponses en fonction du niveau de dégradation, dans le silence et dans le bruit, écarts types entre parenthèses. En grisé, sont présentés les résultats du test d'intelligibilité par répétition de phrases issus de Fontan *et al.* (2014).

Le score est de 100 % si l'appariement énoncé/image est correctement réalisé ou 0 % sinon (même si l'image et le stimulus audio présentent des éléments communs). Les résultats sont présentés sur la Figure 2 et montrent que plus la dégradation du signal de parole augmente, plus les scores de compréhension diminuent, dans le silence comme dans le bruit (différence moyenne = 10,8 %). Les données ont été analysées en utilisant un modèle linéaire mixte généralisé et ont révélé un effet significatif de la dégradation ($\chi^2(9) = 2416,7; p < 0,001$) et du bruit ($\chi^2(1) = 95,3; p < 0,001$). La Figure 2 présente en grisé les scores obtenus dans (Fontan *et al.* 2014) avec une tâche

de répétition de phrases issues du test HINT, la comparaison entre ces deux résultats est discutée en section 5.

4.2 Type de réponse et temps de réaction

L'analyse du type de réponses permet d'observer le nombre de fois que les participants ont sélectionné un des trois types de réponses possibles: choix d'images, bouton « Aucune image ne correspond » et bouton « Rien entendu/compris ». Les résultats montrent que les participants n'ont quasiment jamais utilisé le bouton « Aucune image ne correspond » (1 fois dans le silence et 6 fois dans le bruit sur une totalité de 3600 réponses). La Figure 3 (Gauche) présente le nombre de fois (en %) que les participants ont utilisé le bouton « Rien entendu/compris » en fonction de la dégradation et du bruit de fond. Ce pourcentage dépasse les 50% à partir des dégradations 6 et 7, limite au delà de laquelle les participants n'arrivent plus à choisir une des images proposées.

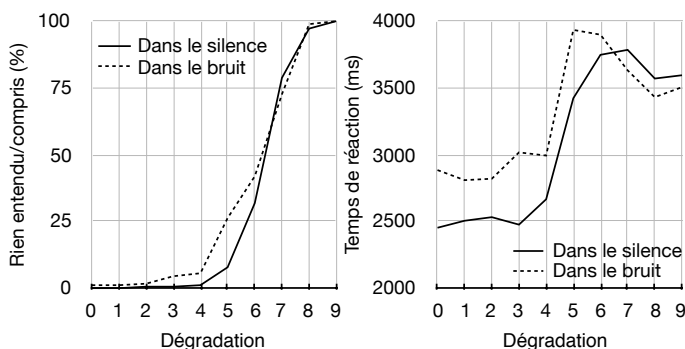


FIGURE 3. Gauche : nombre d'appui sur le bouton « Rien entendu/compris » (en %) en fonction de la dégradation. Droite : Temps de réaction (ms) en fonction de la dégradation. Toutes les réponses de participants sont prises en comptes.

Dans l'analyse des temps de réaction, toutes les réponses des participants sont prises en compte (Figure 3 - Droite) afin d'observer l'évolution de la difficulté de la tâche dans son ensemble. Les données ont été analysées à l'aide d'un modèle linéaire mixte et montrent un effet significatif de la dégradation ($F(9; 3504) = 53,4; p < 0,001$) et du bruit de fond ($F(1; 3504) = 141,7; p < 0,001$). Les données dans le silence présentent une forte augmentation du temps de réaction à partir de la dégradation 4 puis une baisse à partir de la dégradation 7. Le bruit de fond est aussi responsable d'une forte augmentation des temps de réaction avec une augmentation moyenne de 217 ms.

5 Discussion

Le test de compréhension proposé dans ce papier diffère des tâches habituellement proposées dans les tests d'intelligibilité en audiométrie vocale car il s'agit pour le participant de sélectionner une image qui correspond à l'énoncé entendu. Les résultats obtenus avec de la parole dégradée par simulation de la presbyacousie montrent un effet significatif de la dégradation et du bruit de fond. Ces effets sont cohérents avec les études précédentes sur les effets de la presbyacousie sur l'audition (Moore, 2007, Fontan *et al.* 2014), c'est-à-dire une baisse des scores en fonction de la gravité de la presbyacousie et de la présence de bruit de fond vocal. Par ailleurs, la comparaison

des scores de compréhension obtenus dans ce test avec les scores d'intelligibilité obtenus par répétition de phrases dans Fontan et al. (2014) nous permet de discuter la nature exacte du score mesuré. D'une part, les scores de compréhension sont plus faibles que les scores d'intelligibilité, avec des différences jusqu'à 24% à la dégradation 7 dans le silence (*cf.* figure 2). Ainsi, les participants sont meilleurs en répétition de phrases qu'en choix d'images. D'autre part, les temps de réaction observés dans les deux études présentent une augmentation puis une baisse entre les dégradations 4 et 7 mais avec une augmentation totale de 1119 ms dans le test de compréhension contre 321 ms dans le test d'intelligibilité.

Les différences entre les deux tâches traduisent la complexité de la tâche de choix d'images. Pour y répondre, le participant doit accomplir trois objectifs en parallèle et de façon dynamique: interpréter un message acoustique plus ou moins dégradé, interpréter des scènes visuelles et enfin apparier l'énoncé compris à l'image correspondante. Il semblerait que deux stratégies soient mises en place par les participants pour répondre à ces objectifs. Lorsque la dégradation est faible (en dessous de 4), les participants entendent le message et l'interprètent facilement. L'appariement entre l'énoncé et l'image correspondante est donc très rapide car le participant sélectionne simplement les éléments visuels correspondant à ce qu'il a entendu. À partir d'une dégradation plus importante (à partir de 4), la tâche devient plus difficile car le participant cherche des indices lui permettant de restaurer les éléments phonétiques manquants dans le message acoustique. Ces indices sont disponibles parmi les quatre scènes imagées qui présentent une grande diversité d'éléments visuels. Cette complexité est accrue par le fait que les scènes visuelles sont très proches sémantiquement. Des mesures oculométriques permettraient d'étudier plus en détail ces deux stratégies en observant les zones et durées de focalisation ou la dilatation pupillaire par exemple.

6 Conclusions et perspectives

Le test de compréhension proposé dans cette étude permet d'obtenir des scores cohérents avec les résultats obtenus par des tests plus classiques (répétition de mots ou de phrases) dans le domaine de l'audiométrie. Il se présente comme un test plus facile à mettre en oeuvre (pas besoin de faire répéter, pas besoin de transcrire les réponses, adapté à un public plus vaste) tout en se rapprochant d'une situation de communication réelle par l'utilisation de phrases et d'un contexte visuel. La complexité de la tâche (faire correspondre ce qui est entendu avec les indices trouvés dans les images) illustre une partie des difficultés rencontrées par les patients atteints de presbycusie dans leur vie quotidienne. Les scores plus faibles en compréhension que dans les tests d'intelligibilité plus classiques permettent d'interroger le lien entre intelligibilité et compréhension. Dans notre cas, l'intelligibilité surestime la compréhension c'est-à-dire qu'au-delà d'une certaine dégradation, les scores d'intelligibilité ne permettent pas de savoir si le message a bien été compris. Ce résultat est à explorer davantage au regard des résultats contradictoires obtenus dans des études précédentes interrogeant également ce lien (*cf.* Fontan *et al.*, 2015).

Remerciements

Cette étude a été réalisée dans le cadre d'un projet financé par la Région Midi-Pyrénées.

Références

BONFILS P. et AVAN P. (2005). Evaluation du système auditif. In Dulguerov P. & Remacle M. *Précis d'audiophonologie et de déglutition, Tome 1 : l'oreille et les voies de l'audition* (pp.149-163). Marseille : Solal.

BOUCCARA D., AVAN P., MOSNIER I., BOZORG GRAYELI A., FERRARY E. et STERKERS O. (2005). Réhabilitation auditive, *Medecine Sciences* 21(2), 190-197.

ESTIENNE F. et PIÉRART B. (2006). *Les bilans de langage et de voix. Fondements théoriques et pratiques*. Paris : Masson.

FONTAN L., MAGNEN C., TARDIEU J., et GAILLARD P. (2014). Simulation des effets de la presbycousie sur l'intelligibilité et la compréhension de la parole dans le silence et dans le bruit. *Actes des Journées d'étude sur la parole (30e édition, Le Mans)*, 694-702.

FONTAN L., TARDIEU J., GAILLARD P. et RUIZ R. (2015). Relationship Between Speech Intelligibility and Speech Comprehension in Babble Noise. *Journal of Speech Language and Hearing Research* 58, 977-986.

FOURNIER J.-E. (1951). *Audiométrie vocale : les épreuves d'intelligibilité et leurs applications au diagnostic, à l'expertise et à la correction prothétique des surdités*. Maloine.

GARNIER T., VESSON J.F., AUDRY J.C. et AZEMA B. (1997). Épreuves vocales, applications. In *Précis d'audioprothèse: l'appareillage de l'adulte. Tome 1, Le bilan d'orientation prothétique* (pp. 203-247). Paris : Éditions du Collège National d'Audioprothèse (1ère éd. 1997).

LAFON J.C. (1964). *Le test phonétique et la mesure de l'audition*. Eindhoven : Éditions Centrex.

MOORE B. C. J. (2007). *Cochlear hearing loss : Physiological, psychological and technical Issues*. Wiley.

NILSSON M., SOLI S. et SULLIVAN J.A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America* 95(2), 1085-1099.

NIEUWLAND M. S. et VAN BERKUM J. J. A. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research* 24(3), 691-701.

PARADIS M. (1989). *Bilingual Aphasia test (French version)*. Mahwah : Lawrence Erlbaum.

SCHOEPFLIN J. (2015). Back to basics : Speech Audiometry, <http://www.audiologyonline.com/>.

SIVONEN P., MAESS B., LATTNER S. et FRIEDERICI A. D. (2006). Phonemic restoration in a sentence context: Evidence from early and late ERP effects. *Brain Research* 1121(1), 177-189.

SPANGHERO-GAILLARD N. (2008). Comment l'être humain comprend? Quelques éléments de réflexion à partir de l'apprentissage d'une langue étrangère, Habilitation à Diriger des Recherches, Spécialité Sciences du Langage, Université Toulouse - Le Mirail, Toulouse.

VAILLANCOURT V., LAROCHE C., MAYER C., BASQUE C., NALI M., ERIKS-BROPHY A., SOLI S. D. et GIGUERE C. (2005). Adaptation of the HINT (Hearing in Noise Test) for adult Canadian Francophone populations. *International Journal of Audiology* (44), 358-369.