



HAL
open science

Caractérisation psychophysique du bruit interne dans le traitement auditif des modulations temporelles : effets du vieillissement et de lésions cochléaires

Sarah Attia

► To cite this version:

Sarah Attia. Caractérisation psychophysique du bruit interne dans le traitement auditif des modulations temporelles : effets du vieillissement et de lésions cochléaires. Neurosciences. Université Paris sciences et lettres, 2021. Français. NNT : 2021UPSLE054 . tel-04003404

HAL Id: tel-04003404

<https://theses.hal.science/tel-04003404>

Submitted on 24 Feb 2023

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.

THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à l'École Normale Supérieure

**Caractérisation psychophysique du bruit interne
dans le traitement auditif des modulations temporelles :
effets du vieillissement et de lésions cochléaires**

Soutenue par

Sarah ATTIA

Le 6 décembre 2021

Ecole doctorale n° 540

**Ecole doctorale Lettres, Arts,
Sciences humaines et
sociales**

Spécialité

**Sciences cognitives,
neurosciences, psychologie**

Composition du jury :

Catherine, SEMAL Professeur, Institut National Polytechnique de Bordeaux	<i>Président</i>
Brian, MOORE Professeur, Université de Cambridge	<i>Rapporteur</i>
Diane, LAZARD Praticien Hospitalier, Institut de l'Audition	<i>Rapporteur</i>
Stephan, EWERT Professeur associé, Université d'Oldenbourg	<i>Examineur</i>
Jan, WOUTERS Professeur, Université de Louvain	<i>Examineur</i>
Christian, LORENZI Professeur, École normale supérieure	<i>Directeur de thèse</i>

Table des acronymes

AM : Amplitude Modulation

AMFR : Amplitude Modulation Following Response

ASSR : Auditory Steady State Response

dB HL : dB Hearing Level

dB SL : dB Sensation Level

dB SPL : dB Sound Pressure Level

EEG : Electroencephalography

EFR : Envelope Following Response

ENV : Envelope

EPSM : Envelope Power Spectrum Model

ERB : Equivalent Rectangular Bandwidth

FM : Frequency Modulation

FMFR : Frequency Modulation Following Response

Hlo : Old Hearing Impaired

IPD : Interaural Phase Differences

IPM-FR : Interaural Phase Modulation Following Response

ISI : Interspike Intervals

NHy : Young Normal Hearing

NHo : Old Normal Hearing

nAFC: n-Alternative Forced Choice

PEA : Potentiels Evoqués Auditifs

PEMO : Perception Model

RSB : Rapport Signal sur Bruit

SRT : Speech Reception Threshold

TDS : Théorie de la Détection du Signal

TFS : Temporal Fine Structure

TMTF : Temporal Modulation Transfer Function

Cette thèse de doctorat a été réalisée au sein de :

L'équipe Audition

Laboratoire des systèmes Perceptifs,
Département d'études cognitives,
École normale supérieure,
Paris Sciences et Lettres Research University,
UMR CNRS 8248,
29 rue d'Ulm,
75005 Paris, France

Sous la supervision du Professeur Christian Lorenzi.

Ce travail de recherche doctorale a été soutenu financièrement par le *Collège National des Audioprothésistes (CNA)*, par la société *Entendre SAS* ainsi que la bourse ANR-17-EURE-0017.

Résumé

Les informations de modulation d'amplitude (AM) et de modulation de fréquence (FM) lentes jouent un rôle crucial dans l'analyse des scènes auditives et la communication. Le but de ce programme de recherche doctorale est de caractériser le rôle de la variabilité neuronale (le « bruit interne ») dans le traitement auditif d'AM et de FM chez le sujet humain jeune et âgé normo-entendant et le sujet âgé malentendant à partir de mesures comportementales et de simulations informatiques. Les mesures comportementales sont obtenues avec le paradigme dit de « double-passe ». La modélisation informatique utilise soit un banc de filtres cochléaires suivi d'un banc de filtres de modulation, d'un bruit Gaussien additif et d'un mécanisme de décision par appariement de gabarit, soit un modèle mathématique dérivé de la théorie de la détection du signal (TDS). Cette approche combinant mesures psychophysiques et modélisation permet de préciser les mécanismes du traitement de l'AM et de la FM chez le sujet entendant et les effets du vieillissement et de lésions cochléaires sur le bruit interne dans le domaine de l'AM et de la FM. Trois études sont menées au cours de ce programme de recherche: i) une première étude portant sur un groupe de sujets normo-entendants jeunes ; ii) une seconde comparant deux groupes de sujets normo-entendants jeunes et âgés ; iii) une troisième comparant deux groupes de sujets normo-entendants jeunes et malentendants âgées. Pour chaque étude, des mesures sont réalisées à l'aide de la méthode des stimuli constants et d'une procédure en choix forcé afin d'estimer la capacité des sujets à détecter une AM et une FM sinusoïdales appliquées à un son pur de 500 Hz. La première étude utilise une cadence d'AM et de FM de 2 et 20 Hz, alors que les deux autres études n'utilisent qu'une cadence de 2 Hz. Un bruit de modulation de variance donnée est appliqué à tous les stimuli. Les signaux cibles d'AM et de FM sont présentés à trois niveaux de performance. Tous les stimuli utilisés au cours de la première session (la première « passe ») sont stockés et présentés à nouveau dans une deuxième session (la deuxième « passe ») après quelques minutes. Le pourcentage d'accord entre les réponses des sujets est calculé pour les essais correspondants dans lesquels les signaux sont identiques entre la première et de la deuxième passe. Le bruit interne des sujets est estimé à partir de ces données empiriques soit en ajustant les réponses du modèle informatique de traitement de modulation par modification de la variance du bruit Gaussien ajouté à la sortie des filtres de modulation (première étude), soit à l'aide d'un algorithme basé sur la TDS (deuxième et troisième études). Les données empiriques et de modélisation de la première étude portant sur les sujets normo-entendants jeunes indiquent qu'AM et FM - quelle que soit leur cadence - sont détectées par des mécanismes sensoriels distincts possédant leur propre source de bruit interne. Le modèle du banc de filtres de modulation permet de rendre compte des données de détection d'AM lente et rapide, confirmant sa capacité à offrir une approche unifiée de la perception d'informations d'enveloppe temporelle. Les résultats de la deuxième étude indiquent qu'avant l'âge de soixante ans, le bruit interne n'augmente pas dans des tâches de détection d'AM et de FM. Enfin, les résultats de la dernière étude indiquent qu'une perte neurosensorielle n'altère pas le bruit interne dans la détection d'AM et de FM. La variabilité neuronale associée au codage et à la détection de modulations temporelles lentes et rapides ne semble donc pas augmenter avec l'âge et à la suite de lésions cochléaires. Le paradigme de double-passe - malgré des limites pratiques évidentes liées à la durée des tests - apporte une nouvelle mesure des capacités supralimaires susceptible de compléter le diagnostic des pertes neurosensorielles.

Mots clés : modulation d'amplitude, modulation de fréquence, bruit interne, vieillissement, perte auditive.

Abstract

Slow amplitude modulation (AM) and frequency modulation (FM) information plays a crucial role in auditory scene analysis and communication. The aim of this doctoral research programme is to characterise the role of neural variability ("internal noise") in the auditory processing of AM and FM in young and elderly normal-hearing and hearing-impaired humans using behavioural measurements and computer simulations. The behavioural measurements are obtained using the so-called "double-pass" paradigm. Computational modelling uses either a cochlear filter bank followed by a modulation filter bank, additive Gaussian noise and a template matching decision mechanism, or a mathematical model derived from signal detection theory (SDT). This approach, combining psychophysical measurements and modelling, makes it possible to specify the mechanisms of AM and FM processing in normal-hearing subjects and the effects of ageing and cochlear lesions on internal noise in the AM and FM domain. Three studies are being conducted in this research programme: i) a first study involving a group of young normal-hearing subjects; ii) a second comparing two groups of young and elderly normal-hearing subjects; iii) a third comparing two groups of young and elderly normal-hearing subjects. For each study, measurements are made using the constant stimulus method and a forced choice procedure to estimate the ability of subjects to detect sinusoidal AM and FM applied to a pure tone of 500 Hz. The first study used an AM and FM rate of 2 and 20 Hz, while the other two studies used a rate of only 2 Hz. A modulation noise is applied to all stimuli. The AM and FM target signals are presented at three performance levels. All stimuli used in the first session (the first "pass") are stored and presented again in a second session (the second "pass") after a few minutes. The percentage of agreement between subjects' responses is calculated for the corresponding trials in which the signals are identical between the first and second passes. The subjects' internal noise is estimated from these empirical data either by adjusting the responses of the computational model by modifying the variance of the Gaussian noise added to the output of the modulation filters (first study) or by using a SDT-based algorithm (second and third studies). The empirical and modelling data from the first study of young normal-hearing subjects indicate that AM and FM - regardless of their rate - are detected by distinct sensory mechanisms with their own internal noise source. The modulation filterbank model is able to account for both slow and fast AM detection data, confirming its ability to provide a unified approach to the perception of temporal envelope information. The results of the second study indicate that before the age of sixty, internal noise does not increase in AM and FM detection tasks. Finally, the results of the last study indicate that sensorineural hearing loss does not alter internal noise in AM and FM detection. The neural variability associated with the encoding and detection of slow and fast temporal modulations therefore does not appear to increase with age and following cochlear damage. The double-pass paradigm - despite obvious practical limitations related to test duration - provides a new measure of suprathreshold abilities that may complement the diagnosis of sensorineural hearing loss.

Mots clés : amplitude modulation, frequency modulation, internal noise, ageing, hearing loss.

Table des matières

Table des acronymes	1
Résumé.....	3
Abstract	4
Table des matières.....	5

Chapitre I : Introduction

I.1. Nature et rôle des indices d'AM et de FM dans la perception auditive	8
I.1.a. Préambule.....	8
I.1.b. Caractérisation et rôle des indices d'AM et de FM dans la perception des sons naturels.....	9
I.1.c. Caractérisation et rôle des indices d'AM et de FM dans la perception des signaux de parole	13
I.1.d. Caractérisation et rôle des indices d'AM et de FM dans la perception des sons musicaux.....	23
I.1.e. Caractérisation des indices de modulation temporelle en fonction de leur représentation dans le système auditif périphérique.....	29
I.2. Encodage et traitement des indices d'AM et de FM par le système auditif.....	35
I.2.a. Préambule.....	35
I.2.b. Traitement de l'AM par le système auditif	37
I.2.b.i. Traitement par le sujet normo-entendant adulte jeune	37
I.2.b.ii. Etudes développementales	42
I.2.c. Traitement de la FM par le système auditif.....	44
I.2.c.i. Traitement par le sujet normo-entendant adulte jeune	44
I.2.c.ii. Etudes développementales.....	52
I.2.d. Des modèles de traitement de modulation	53
I.3. Effets du vieillissement et d'une perte auditive neurosensorielle.....	58
I.3.a. Préambule.....	58
I.3.b. Effets du vieillissement sur le traitement des indices d'AM et de FM	60

I.3.b.i. Effets du vieillissement sur la sensibilité à l'AM	60
I.3.b.ii. Effets du vieillissement sur la sensibilité à la FM	68
I.3.c. Effets de la perte auditive neurosensorielle sur le traitement des indices d'AM et de FM	77
I.3.c.i. Effets de la perte auditive neurosensorielle sur la détection d'AM.....	78
I.3.c.ii. Effets de la perte auditive neurosensorielle sur la détection de FM.....	82
I.3.d. Origines probables des troubles de codage temporel.....	86
I.3.d.i. Réduction du nombre de synapses ou de neurones.....	86
I.3.d.ii. Altération du mécanisme de verrouillage en phase	89
I.3.d.iii. Perte d'inhibition centrale	93
I.4. Caractérisation psychophysique et neurophysiologique du bruit interne	93
I.4.a. Préambule.....	93
I.4.b. Bruit interne.....	95
I.4.b.i. Définitions.....	95
I.4.b.ii. Effets neuronaux et comportementaux	98
I.4.c. Méthodes d'estimation du bruit interne.....	103
I.4.c.i. Méthode du bruit équivalent	103
I.4.c.ii. Masquage de piédestal.....	106
I.4.c.iii. Méthode de double-passe	108
I.4.c.iv. Estimation à partir de la pente de la fonction psychométrique.....	112
I.4.c.v. Méthode de classification de limite.....	113
I.4.c.vi. Méthode d'estimation basée sur le gel des fluctuations externes.....	115
I.4.c.vii. Ajustement de modèles	116
I.4.d. Effets du bruit interne en perception auditive et visuelle	119
I.4.d.i. Trajectoires développementales.....	119
I.5. Objectifs et plan de thèse	142
I.5.a. Problématique	142

Chapitre II :

Application du paradigme de double-passe à la détection d'AM et de FM chez des sujets normo-entendants jeunes

II.1. Double-pass consistency for amplitude- and frequency- modulation detection in normal-hearing listeners	145
--	-----

II.1.a. Résumé.....	145
II.1.b. Article.....	146

Chapitre III:

**Application du paradigme de double-passe à la détection d’AM et de FM
chez des sujets normo-entendants jeunes et âgés**

III.1. No effect of age on sensitivity and double-pass consistency for amplitude-and frequency-modulation detection.....	179
III.1.a. Résumé.....	179
III.1.b. Article.....	180

Chapitre IV:

**Application du paradigme de double-passe à la détection d’AM et de FM
chez des sujets normo-entendants jeunes et malentendants âgés**

IV.1. No effect of sensorineural hearing loss on double-pass consistency for amplitude- and frequency- modulation detection.....	208
IV.1.a. Résumé	208
IV.1.b. Article.....	209

Chapitre V:

Discussion et conclusion

V.1. Précisions sur le paradigme expérimental.....	232
V.1.a. Choix du paradigme expérimental.....	232
V.1.b. Choix des paramètres expérimentaux	233
V.1.b.i. Cadence de modulation	233
V.1.b.ii. Caractéristiques du bruit externe.....	234
V.1.b.iii. Niveaux de sensibilité (performance) ciblés.....	236
V.1.c. Choix théoriques	236
V.2. Principales conclusions de la première étude	237
V.3. Principales conclusions de la deuxième étude	239
V.4. Principales conclusions de la troisième étude.....	240
V.5. Discussion.....	241
V.6. Limites	245
V.7. Applications cliniques.....	247
Références.....	249

Chapitre I : Introduction

I.1. Nature et rôle des indices d'AM et de FM dans la perception auditive

I.1.a. Préambule

Le système auditif humain décompose les signaux acoustiques complexes large-bande tels que la parole, la musique ou les sons de l'environnement grâce à un banc de filtres passe-bande situés sur la membrane basilaire dans la cochlée (Moore, 2012). Ces filtres sont étroits en basse fréquence audio et s'élargissent en haute fréquence. De fait, leur résolution spectrale est meilleure en basse fréquence et leur résolution temporelle est meilleure en haute fréquence. Cette organisation spécifique résulte probablement d'une forme d'adaptation des mécanismes cochléaires aux statistiques des signaux naturels tels que les vocalisations animales et les sons dits « environnementaux » d'origine géophysique (Lewicki, 2002). Toutefois, cette adaptation s'étend probablement au-delà de la simple dimension audio-fréquentielle.

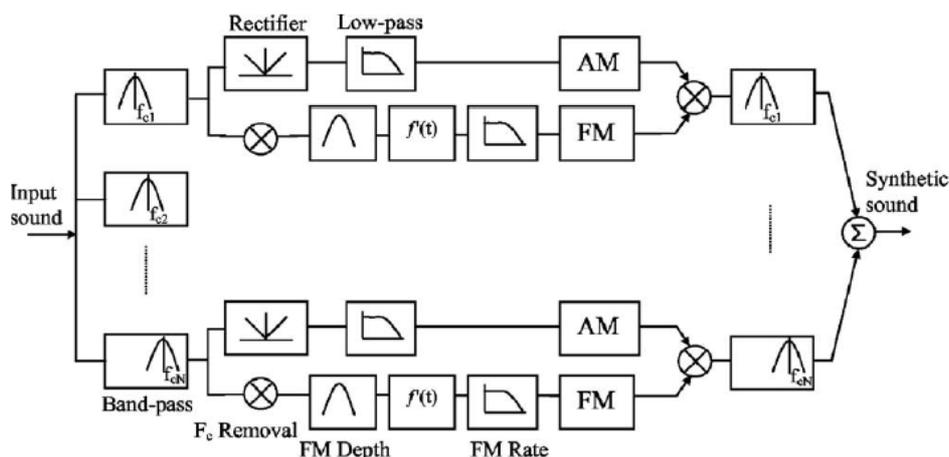


Figure 1. Vocodeur décomposant les sons en bande étroite grâce à des filtres d'analyse passe-bande. Les signaux filtrés sont ensuite rectifiés et filtrés en passe-bas afin d'extraire la composante d'AM. En parallèle, la composante de FM des sons est extraite et filtrée. Le signal vocodé final peut être reconstruit en combinant les composantes d'AM et de FM, puis en additionnant ces signaux de bande après re-filtrage par des filtres identiques aux filtres d'analyse. D'après Stickney et al. (2007).

Les études manipulant les signaux acoustiques à l'aide de vocodeurs (*e.g.*, Dudley et al., 1939 ; Van Tasell et al., 1987; Zeng et al., 2005) décrivent chaque signal à bande étroite issu de la décomposition fréquentielle d'un signal complexe tel que la parole à l'aide d'un banc de filtres accordés comme véhiculant des informations (des indices acoustiques) dynamiques de modulation d'amplitude (AM, l'« enveloppe temporelle » acoustique du signal) et de modulation de fréquence (FM, la « structure temporelle fine » acoustique du signal). La figure 1 illustre le schéma de traitement d'un vocodeur. Conformément à cette vue dichotomique, chaque signal à bande étroite issu du filtrage en un point donné de la membrane basilaire dans la cochlée humaine peut être considéré comme un signal dynamique modulé à la fois en amplitude et en fréquence (Moore, 2014). Il est alors tout à fait concevable que le système auditif des primates (et ainsi de l'humain) ait évolué de manière à élaborer des stratégies sensorielles efficaces d'extraction et de traitement des indices d'AM et de FM qui constituent l'information pertinente sur laquelle la reconnaissance de sources est fondée (Rodriguez et al., 2010 ; Koumura et al., 2019) . Les recherches portant sur la contribution de ces modulations temporelles à la détection, discrimination et identification des signaux acoustiques complexes sont nombreuses. Depuis plus d'un demi-siècle, l'attention de la communauté scientifique s'est portée sur l'importance de ces indices temporels vis-à-vis la perception de la parole puis plus récemment vis-à-vis la perception de la musique, des sons de l'environnement et des vocalisations des animaux non-humains.

Ce chapitre a pour but de détailler le rôle et la contribution des indices d'AM et de FM dans la perception auditive. Il s'articule en trois parties portant sur la caractérisation et rôle des indices d'AM et de FM dans la perception : i) des sons naturels ; ii) des sons de parole ; en enfin iii) des sons produits par des instruments musicaux.

I.1.b. Caractérisation et rôle des indices d'AM et de FM dans la perception des sons naturels

Les « **paysages sonores** » (**soundscapes**) **naturels** sont composés de sons simples ou complexes produits par des sources acoustiques de nature **biologique** (*e.g.*, vocalisations animales humaines et non humaines) ou **non biologique** (*e.g.*, bruit d'une source d'eau, bruit de la pluie, bruissements du vent) spécifiques au biotope considéré. Les paysages sonores urbains se distinguent fortement des paysages sonores naturels car ils sont fortement imprégnés par l'action de l'homme (*e.g.*, par des signaux produits par des machines ou des véhicules, etc.) (Pijanowski et al., 2011 ; Sueur & Farina, 2015 ; Farina & Gage, 2017). Un paysage sonore - qu'il soit naturel ou urbain – est *structuré* car il correspond à un ensemble restreint de combinaisons de sons spécifiques dont il est possible de quantifier le degré de redondance statistique (Attneave, 1954 ; Field, 1987 ; Singh & Teunissen, 2003). Cette approche théorique repose sur l'hypothèse fondamentale formulée par Barlow en 1953, à savoir que « les systèmes sensoriels sont adaptés à l'encodage des signaux naturels » (pour revue, voir Garcia-Lazaro et al., 2006 ; Theunissen & Elie, 2014). Cette hypothèse a conduit de

nombreux chercheurs à étudier plus particulièrement la structure statistique des sons naturels afin de caractériser et quantifier les composantes les plus pertinentes (Voss and Clarke, 1975 ; Attias & Schreiner, 1998 ; Nelken et al., 1999 ; Lewicki, 2002 ; Escabí et al., 2003 ; Singh and Theunissen, 2003 ; Lesica & Grothe, 2008) utilisées pour élaborer des stratégies de traitement optimal de l'information sensorielle.

Cette structure statistique peut être mise en évidence par le calcul du **spectre de puissance de modulation** (modulation power spectrum, MPS) permettant de visualiser la puissance de modulation (code couleur) en fonction de la fréquence de modulation temporelle (en abscisse) et de la fréquence de modulation spectrale (en ordonnée) comme le montre la figure 2 ci-dessous extraite de Woolley et al. (2005). Ce spectre de puissance de modulation révèle que les vocalisations animales (humaines et non humaines) se distinguent nettement des sons environnementaux (pluie, vent, feu, cours d'eau) de par la présence de modulation d'amplitude lentes (<50 Hz) et de modulations spectrales élevées reflétant la structure harmonique des vocalisations étudiées dans cet exemple (Singh & Theunissen, 2003 ; Elliott & Theunissen, 2009 ; Theunissen & Elie, 2014).

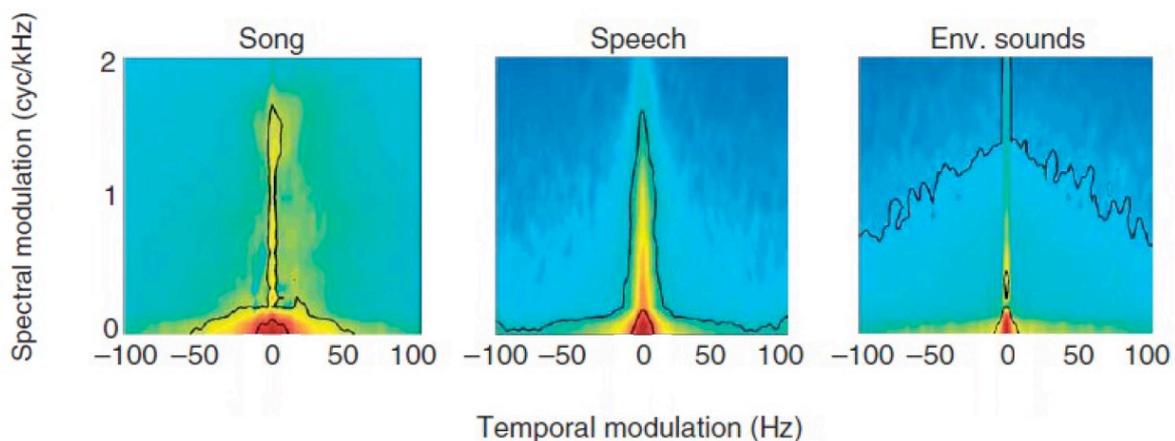


Figure 2. Spectres de puissance de modulation (MPS) de trois signaux différents (chant d'oiseau, parole, sons environnementaux – feu, pluie, cours d'eau, vent dans les feuillages). La puissance de modulation des signaux biologiques (chant d'oiseau et parole) est maximale pour les cadences de modulation temporelle lentes et les cadences de modulation spectrale élevées (en raison de la riche structure harmonique de ces signaux). Extrait de Woolley et al. (2005).

Des études psychoacoustiques manipulant des **signaux environnementaux « urbains »** (bruits de voiture ou d'avion, bruits de machines, etc.) à l'aide de vocodeurs d'AM (*i.e.*, des algorithmes permettant de dégrader les indices de FM tout en préservant les indices d'AM dans une série de bandes audio-fréquentielles ; voir section I.1.c plus bas) révèlent que la capacité des auditeurs humains à identifier les sons environnementaux repose à la fois sur la perception d'indices d'AM et de FM (Gygi et al., 2004 ; Shafiro, 2008 ; voir figure 3 ci-dessous). Cette hypothèse a également été validée par les études psychoacoustiques démontrant que

les auditeurs humains sont capables de discriminer des « **textures auditives** » (sons environnementaux produits par la superposition de plusieurs évènements acoustiques similaires) sur la base de statistiques moyennées dans le temps calculées à partir des fluctuations d'enveloppe temporelle en sortie des filtres cochléaires et/ou de filtres de modulation (McDermott & Simoncelli, 2011 ; McWalter & Dau, 2017 ; McWalter & McDermott, 2018).

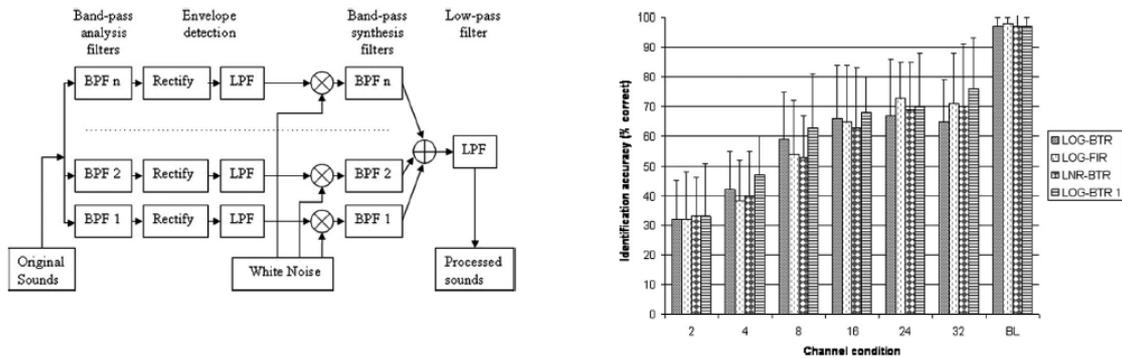


Figure 3. Schéma de gauche : Vocodeur d'AM utilisé pour estimer les performances d'identification de sons environnementaux chez des sujets humains normo-entendants. Le vocodeur décompose les sons en bande étroite grâce à des filtres d'analyse. Les signaux filtrés sont ensuite rectifiés et filtrés en passe-bas afin d'extraire leur enveloppe temporelle (la composante d'AM). La porteuse d'origine (structure temporelle fine ou FM) des sons environnementaux est remplacée par un bruit à bande étroite. Les bruits modulés par les enveloppes temporelles sont additionnés, filtrés puis présentés pour identification aux sujets. Graphique de droite : performances moyennes d'identification sur le groupe de sujets, en fonction de la résolution fréquentielle du vocodeur (2-32 bandes) et dans une condition à bande limitée (BL). La performance augmente avec la résolution fréquentielle mais ne dépasse pas 76% de réponses correctes, indiquant que l'identification est possible sur la seule base des indices d'AM mais que les indices de FM jouent également un rôle. Extrait de Shafiro (2008).

Dans l'ensemble ces résultats suggèrent que le système auditif humain exploite les indices de modulation temporelle véhiculés par les scènes acoustiques naturelles et que leurs changements devraient être utilisés pour évaluer les caractéristiques biologiques et géophysiques d'un habitat donné et les caractéristiques des vocalisations animales non humaines.

Plus récemment, Thoret et al. (2020) caractérisent et étudient les variations des indices d'AM et de FM véhiculés par les paysages sonores de quatre habitats différents (clairière, maquis, prairie, rivière) d'une même écorégion de Californie et ce à quatre moments de la journée et sur une année entière. Des **spectres d'AM et de FM** sont calculés par bande de fréquence audio en tenant compte des effets du filtrage cochléaire humain. Ces spectres d'AM et de FM sont calculés différemment des « spectres de puissance de modulation (MPS) » cités plus haut. Ils permettent de visualiser la puissance de modulation (code couleur) en fonction de la fréquence de modulation temporelle (en abscisse) et de la bande fréquentielle audio (en ordonnée). L'analyse de ces spectres, tels que représentés en figure 4 pour l'AM, indique que les indices de modulation temporelle (AM comme FM) fournissent des marqueurs acoustiques saillants des paysages sonores naturels et de leurs variations temporelles au gré de la journée

et de la saison. Ces indices de modulation temporelle présentent des cadences comprises entre quelques Hertz et quelques dizaines de Hertz, et apparaissent généralement dans les régions fréquentielles audio moyennes et hautes. Ces mêmes analyses révèlent le rôle important joué par la composante biologique (dite «biophonique » ; *i.e.*, les vocalisations animales) de ces paysages sonores dans les distinctions observées entre habitat, moment de la journée et saison. Plus précisément, cette étude indique que : i) les indices biophoniques correspondent à des indices de modulation temporelle lents (< 10 Hz) et/ou rapides (30-100 Hz), avec un pic autour de 2 Hz pour les oiseaux et autour de 50 Hz pour les insectes dans une gamme de fréquence audio moyenne voire élevée (> 1-3 kHz) ; ii) les variations diurnes et saisonnières sont associées à des changements significatifs des indices d'AM et de FM ; iii) les indices géophoniques (vent, pluie, cours d'eau) correspondent à des indices de modulation relativement lents (< 10 -50 Hz) dans la gamme de fréquence audio plutôt basse (<1-2 kHz). Enfin, pour chaque habitat, la puissance des bandes de modulation lentes (10-30 Hz) et élevées (30-100 Hz) est plus élevée au printemps, reflétant des variations de biophonie entre les saisons.

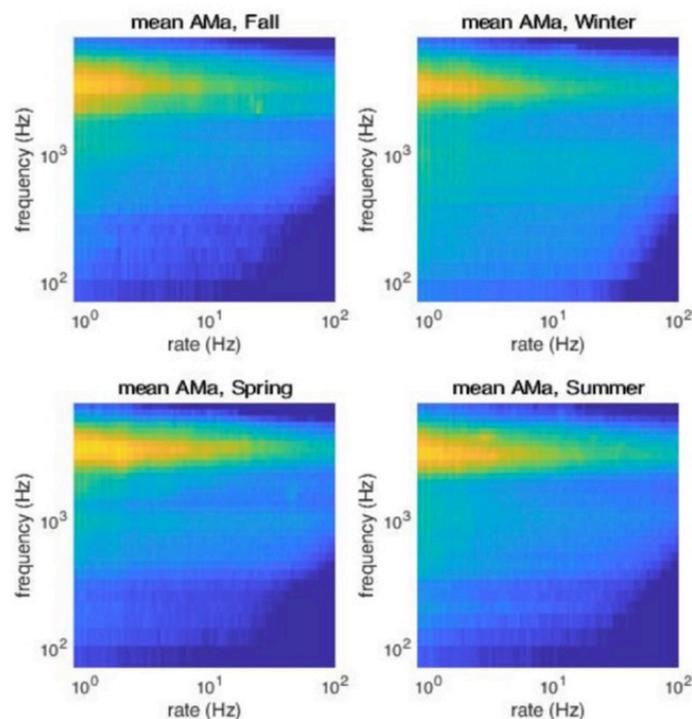


Figure 4. Spectres de modulation d'amplitude calculés à partir de bases d'enregistrements sonores réalisés dans quatre habitats du parc national des Séquoia en Californie, à quatre moments de la journée et sur une année entière. Chaque spectre d'AM est calculé pour une bande de fréquence audio, correspondant à la sortie d'un filtre cochléaire (dont la fréquence centrale est présentée en ordonnée). Le code de couleur indique le niveau de puissance de modulation en fonction de la cadence de modulation (en abscisse). Ces spectres sont moyennés sur l'ensemble des quatre habitats et des quatre moments de la journée. Chaque graphique montre le spectre d'AM pour chacune des quatre saisons (automne, hiver, printemps, été). Extrait de Thoret et al. (2020).

Thoret et al. (2020) calculent les spectres d'AM et de FM de leur base de paysages sonores et classifient ces spectres par un algorithme d'apprentissage machine (support vector machine, SVM). En phase test, les performances de classification sont largement au-dessus du niveau de réponse au hasard, indiquant que les paysages sonores naturels véhiculent des informations de modulation temporelle potentiellement utiles pour l'identification du biotope, du moment de la journée et de la saison par un auditeur humain.

I.1.c. Caractérisation et rôle des indices d'AM et de FM dans la perception des signaux de parole

Ces résultats démontrent donc un partitionnement spécifique des spectres d'AM et de FM en fonction du paysage sonore. Ces spectres s'avèrent très différents de ceux obtenus à partir d'enregistrements de sons de parole (Houtgast & Steeneken, 1973 ; Steeneken & Houtgast, 1980; Varnet et al., 2017). La figure 5 illustre les spectres d'AM calculés par Steeneken & Houtgast (1980) sur un échantillon de 60 secondes de parole d'une langue donnée et ce dans deux bandes de fréquence audio distinctes (centrées respectivement sur 250 et 2000 Hz). Les spectres d'AM révèlent que la puissance de modulation est concentrée en dessous de 5-10 Hz environ pour la parole.

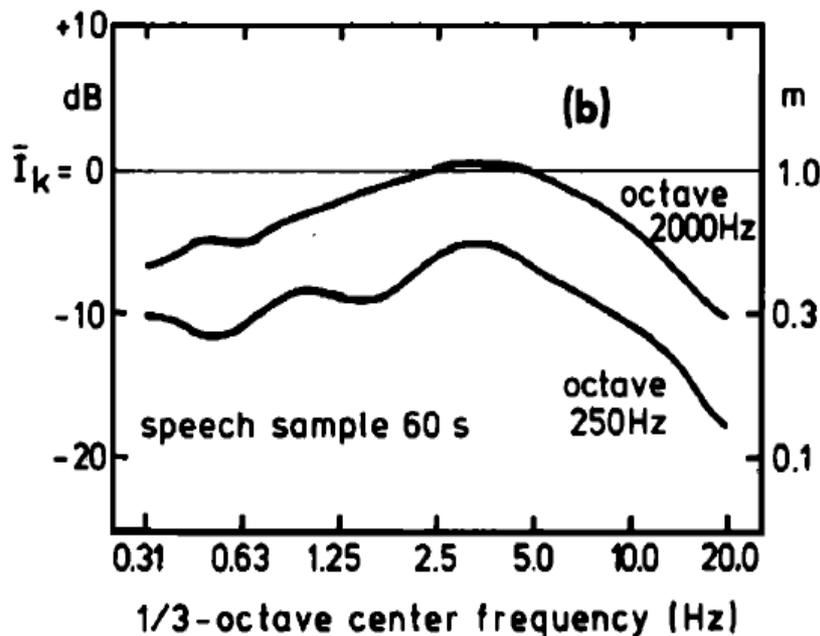


Figure 5. Spectres de modulation d'amplitude calculés à partir d'un échantillon de 60 secondes de parole dans deux bandes de fréquence audio. Le spectre d'AM montre l'intensité moyenne de l'enveloppe (à gauche) ou l'index de modulation m (à droite) en fonction de la cadence de modulation (plus précisément, la fréquence centrale de filtres de modulation de largeur 1/3 d'octave). Extrait de Steeneken & Houtgast (1980).

L'étude de Varnet et al. (2017) étend notablement les résultats de l'étude « pionnière » d'Houtgast et Steeneken (1973) en réalisant les analyses de spectres d'AM comme de FM sur un corpus de phrases enregistrées dans dix langues différentes (français, anglais, espagnol, turque, zoulou, japonais, etc.) par plusieurs locuteurs.

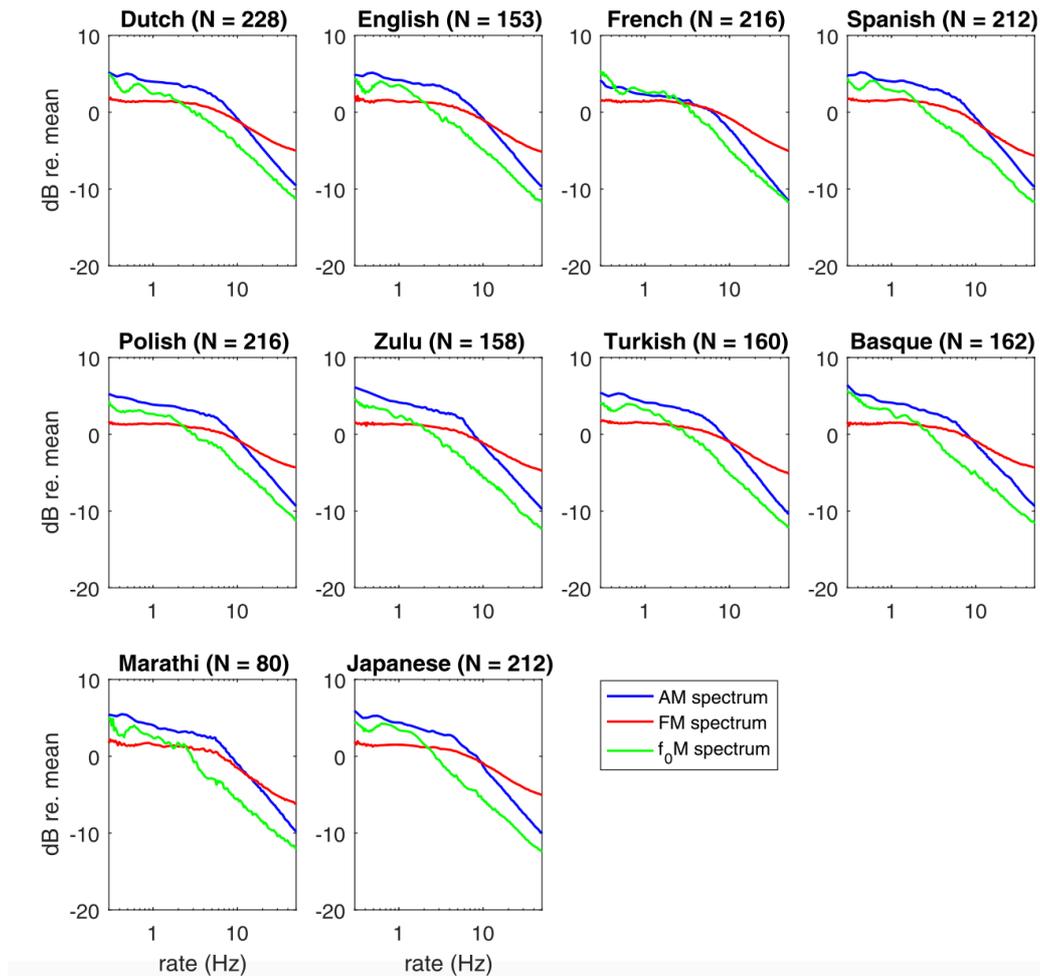


Figure 6. Spectres de modulation d'amplitude (AM) et de fréquence (FM) pour 10 langues. Les spectres ont été moyennés sur l'ensemble des phrases et locuteurs. Les lignes bleues représentent les spectres de modulation d'amplitude; les lignes rouges représentent les spectres de FM (modulations de la fréquence instantanée au sein de chaque bande de fréquence); les lignes vertes représentent les spectres de modulation de la fréquence fondamentale (F₀) du signal de parole. Dans chaque graphique, la puissance de modulation est représentée en fonction de la cadence de modulation. Extrait de Varnet et al. (2017).

Cette étude révèle que les spectres d'AM et de FM des sons de parole sont très similaires entre les langues, la puissance de modulation étant généralement maximale en dessous de 5-10 Hz (voir la figure 6). La forme générale de type « passe-bas » de ces spectres reflète le fait

que les sons de parole - quelle que soit la structure rythmique et syntaxique de la langue - véhiculent principalement des modulations temporelles d'amplitude et de fréquence *lentes*, une propriété commune à tous les sons naturels (Voss & Clarke, 1975, Singh & Theunissen, 2003). Dans le cas de la parole, les composantes de FM lentes reflètent les modulations temporelles de la fréquence de vibration des cordes vocales et les composantes d'AM lentes, les modulations temporelles des propriétés géométriques du tractus vocal (*e.g.*, Story, 2014 ; voir les simulations d'un modèle de production de parole représentées en figure 7). La similarité de ces spectres de modulation à travers les langues résulte probablement des *contraintes fondamentales imposées par les mécanismes articulatoires et perceptifs communs à toutes les langues* (Arai & Greenberg, 1997).

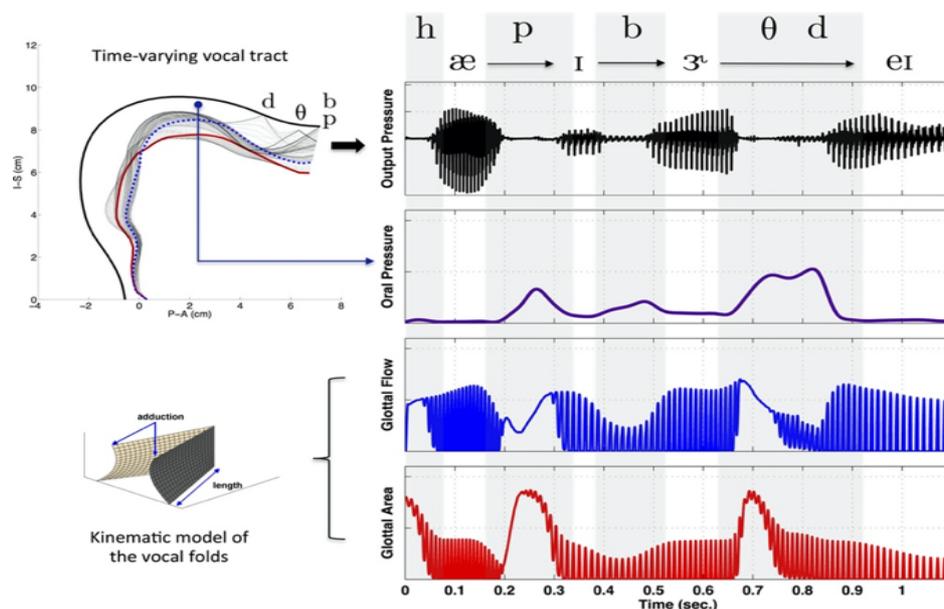


Figure 7. Simulation de la phrase "Happy birthday" à l'aide du modèle de production de parole de Story (2014). Les variations dynamiques des caractéristiques géométriques du tractus vocal (à gauche) engendrent des changements dans l'enveloppe temporelle du signal acoustique (la composante d'AM ; à droite : deuxième graphique à partir du haut). Extrait de Story (2014).

A la suite des travaux pionniers de Flanagan (Flanagan, 1980), une longue série d'études empiriques a examiné le rôle des indices d'AM et de FM dans la perception - et le plus souvent, l'*intelligibilité* - de la parole (*e.g.*, Drullman, 1995 ; Shannon et al., 1995 ; Smith et al., 2002 ; Zeng et al., 2004, 2005 ; Gilbert & Lorenzi, 2006 ; Hopkins et al., 2010 ; Kates, 2011 ; Heinz & Swaminathan, 2009). Dans ces nombreuses études, la composante d'AM est souvent dénommée « enveloppe temporelle (ENV) » et la composante de FM, « structure temporelle fine (TFS pour temporal fine structure) ». Ces concepts d'enveloppe temporelle et de structure temporelle fine renvoient ici à des indices physiques (acoustiques), extraits de signaux de parole à bande étroite ou large.

Pour évaluer la contribution de ces deux indices sur l'intelligibilité de la parole, une méthode généralement utilisée est celle des **vocodeurs** (*cf.* figures 1 et 3) dans laquelle les signaux de

paroles sont filtrés en un nombre fini de bandes fréquentielles appelées « bandes d'analyse ». Dans sa version la plus communément utilisée, le nombre et la bande passante de ces filtres d'analyse peuvent être systématiquement variés afin de manipuler la résolution spectrale du vocodeur (1, 2, 4, ... n bandes). Notons que la largeur des bandes d'analyse ne devrait pas être inférieure à celle des filtres cochléaires (le nombre de bande ne devrait donc pas dépasser environ 32 bandes entre 80 et 8000 Hz, ce qui permet d'obtenir des bandes dont la largeur est précisément égale à une bande passante rectangulaire équivalente, ERB (Equivalent Rectangular Bandwidth)) [la bande passante de chaque filtre auditif (ou « cochléaire ») est typiquement comprise entre 11 et 17 % de la fréquence centrale (Moore, 1996) et peut être approximée à l'aide de l'échelle ERB]. Puis ces signaux de bande sont démodulés de façon à extraire l'enveloppe ou la structure fine du signal d'origine au sein de chaque bande d'analyse. Ce traitement repose sur l'hypothèse que ces indices sont des composantes indépendantes du signal à bande étroite. Différentes stratégies d'extraction d'enveloppe temporelle au sein de chaque canal ont été proposées : i) la première consiste à effectuer un redressement mono-alternance puis un filtrage passe-bas avec une fréquence de coupure inférieure à la fréquence critique de la bande d'analyse (la moitié de la bande passante du filtre d'analyse). Cette méthode permet d'extraire une enveloppe temporelle dont la forme (la résolution temporelle) dépend de la fréquence de coupure et de la pente du filtre passe-bas (Moore, 2014) ; ii) la deuxième méthode consiste à utiliser le signal analytique et la transformée de Hilbert ; dans ce cas, l'enveloppe temporelle n'est pas limitée. La structure temporelle fine peut être extraite en divisant le signal de bande par l'enveloppe temporelle extraite à partir de ces opérations, ou en estimant la fréquence instantanée à partir de la transformée de Hilbert. Les vocodeurs permettent ainsi de manipuler ces deux composantes d'enveloppe et de structure fine du signal de bande, en remplaçant l'une des deux composantes originelles par des composantes extraites d'autres signaux (dans le cas d'un vocodeur à bruit (noise vocoder) ou à ton (tone vocoder), ou dans le cas d'un « chimériseur » (chimaerizer)), ou en filtrant d'une manière ou d'une autre ces mêmes composantes de façon à éliminer sélectivement certains traits (features) de modulation dans le signal vocodé. Les procédures communément utilisées pour supprimer entièrement la structure temporelle fine d'origine sont appliquées par deux types de vocodeur : les vocodeurs utilisant comme porteuses (*i.e.*, structure temporelle fine) au sein de chaque bande i) un bruit (des bandes de bruit Gaussien sont modulées par l'enveloppe de la parole dans chaque bande) ou ii) un son pur (l'enveloppe module une simple sinusoïde dont la fréquence est égale à la fréquence centrale de la bande d'analyse) (Kates, 2011). Enfin, à l'issue de ces processus d'extraction et de manipulation des composantes d'enveloppe et de structure fine, les signaux de bande sont reconstruits en appliquant l'enveloppe temporelle à la structure fine, en les refiltrant par les filtres d'analyse, puis en les ajoutant pour composer le signal vocodé final. Les signaux de parole vocodés sont alors présentés aux oreilles des sujets humains pour identification.

Une première façon d'étudier le rôle des indices acoustiques d'AM dans la perception de la parole consiste à évaluer la *fidélité de transmission* de ces indices via un canal donné : une

salle, un auditorium, une prothèse, ou encore ... un simple vocodeur. Cette évaluation peut être réalisée à l'aide de la **fonction de transfert de modulation (MTF)** du canal, qui estime la réduction de la profondeur de modulation d'un signal modulé donné produite dans chaque série de bande de fréquence audio (Steeneken & Houtgast, 1980 ; Houtgast & Steeneken, 1985). Dans une étude princeps menée par (Drullman, 1995), les effets nuisibles produits par un bruit masquant ou par différents types de distorsions temporelles (*e.g.*, compression rapide, etc.) sur l'intelligibilité de parole vocodée sont ainsi interprétés à la lumière de leurs effets spécifiques sur la MTF du vocodeur. Ces derniers sont illustrés en figure 8 ci-dessous.

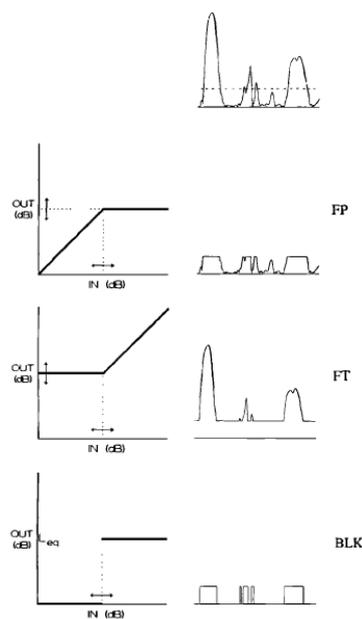


Figure 8. Distorsions d'enveloppe temporelle utilisées afin d'évaluer le rôle des indices d'enveloppe dans l'intelligibilité de phrases. Le graphique du haut représente l'enveloppe temporelle de la phrase utilisée à titre d'exemple. Les trois graphiques du bas illustrent les effets de distorsions éliminant les pics (FP) et les creux (FT) d'enveloppe, ou remplaçant des segments entiers de l'enveloppe d'origine par des modulations carrées plus ou moins longues (BLK). Extrait de Drullman (1995).

Les résultats obtenus par Drullman (1995) démontrent que l'intelligibilité de la parole est fortement réduite lorsque que l'enveloppe temporelle du signal est distordue, les pics d'enveloppe jouant un rôle plus important que les creux. Dans une autre étude princeps menée par (Drullman et al., 1994), l'intelligibilité de phrases présentées dans un bruit stationnaire masquant est évaluée en filtrant en passe-bas ou en passe-haut les indices d'enveloppe temporelle au sein de chaque bande de fréquence à l'aide d'un vocodeur préservant la structure temporelle fine. Les résultats expérimentaux montrent que l'élimination d'indices d'enveloppe temporelle n'affecte que très peu les seuils d'identification (speech reception threshold, SRT) tant que les modulations temporelles comprises entre 4 et

16 Hz sont préservées. Ainsi, les indices d'AM jouant un rôle clé dans la compréhension orale sont relativement lents (4-16 Hz), et correspondent aux modulations d'amplitude associées au débit syllabique dans la parole continue.

Une autre manière d'étudier le rôle de l'AM et de la FM dans la perception de la parole consiste à dégrader l'information spectrale du signal de parole à des degrés variables grâce à des vocodeurs d'AM. Dans une expérience réalisée au sein de l'étude de Drullman (1995) citée plus haut, l'importance relative des indices temporels d'enveloppe et de structure fine du signal de parole au sein de chaque bande de fréquence est évaluée à l'aide d'un vocodeur à bruit (noise vocodeur) remplaçant la structure temporelle fine d'origine des phrases par des bruits Gaussiens. Les mesures d'intelligibilité de ces signaux vocodés indiquent que les signaux vocodés restent entièrement intelligibles en l'absence d'information de structure temporelle fine, confirmant le rôle clé joué par l'enveloppe temporelle.

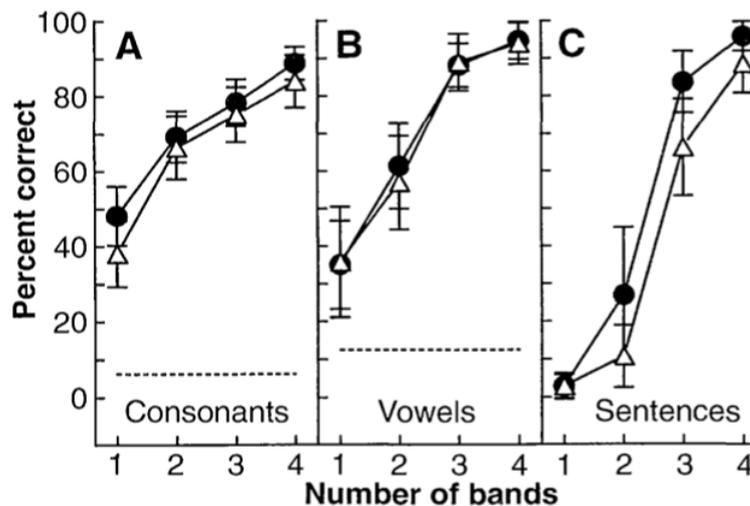


Figure 9. Scores d'identification de consonnes (A), de voyelles (B), de phrases (C) moyennés sur huit auditeurs normo-entendants en fonction du nombre de bandes spectrales. Les cercles pleins indiquent les résultats obtenus pour des enveloppes filtrées en passe-bas avec des fréquences de coupure de 50, 160 et 500 Hz (les résultats ne diffèrent pas significativement entre conditions de fréquence de coupure au-delà de 16 Hz). Les triangles ouverts indiquent les résultats obtenus pour des enveloppes filtrées en passe-bas avec une fréquence de coupure de 16 Hz. Extrait de Shannon et al. (1995).

Ce résultat est confirmé par une étude similaire menée par Shannon et al. (1995) avec un autre type de vocodeur à bruit. Dans cette étude canonique, le signal acoustique est divisé en différentes bandes de fréquence et l'enveloppe temporelle est extraite au sein de chaque canal fréquentiel par redressement mono-alternance suivi d'un filtrage passe-bas (fréquences de coupure : 16, 50, 160 et 500 Hz). Un bruit est ensuite modulé par l'enveloppe du signal dans la bande d'analyse du signal d'origine et toutes les bandes sont ensuite additionnées. La tâche consiste à mesurer l'intelligibilité de consonnes (16 consonnes médiales a/C/a), de voyelles (8 voyelles de type h/V/d) et de phrases simples dans chaque condition (fréquence de coupure utilisée pour le filtrage passe-bas d'enveloppe), lorsque ces signaux sont présentés dans le silence. Comme le montre la figure 9, la performance d'identification de la parole sur

l'ensemble des trois mesures augmente avec le nombre de bandes spectrales, ou en d'autres termes, avec la résolution fréquentielle des signaux. Une performance de reconnaissance des phrases quasi parfaite (90 à 100 % de performance d'intelligibilité) est atteinte pour un nombre limité à 4. Ainsi, dans des conditions d'écoutes favorables (dans le silence), une intelligibilité parfaite peut être obtenue sur la seule base d'indices d'AM extraits au sein d'un nombre limité de bandes de fréquence. Comme déjà observé par Drullman (1995), les indices de FM (*i.e.*, de structure temporelle fine) ne semblent pas jouer de rôle dans la perception dans le silence.

Smith et al. (2002) utilisent de la parole chimérique obtenue en combinant par vocodage l'enveloppe et la structure temporelle fine de deux sons distincts, ce afin d'étudier la poids perceptif (*i.e.*, la contribution relative de l'enveloppe temporelle physique (AM) et de la structure temporelle fine physique (FM)) dans la perception de la parole dans différentes situations d'écoute. Plus précisément, les **chimères auditives** sont synthétisées à partir de deux signaux acoustiques différents (*e.g.*, deux phrases distinctes) de même durée, filtrés en passe-bande afin de décomposer chaque son en un certain nombre de bandes de fréquences comprises entre 80 et 8.8 kHz. En sortie de chaque filtre d'analyse, une transformée de Hilbert est appliquée afin d'extraire l'enveloppe temporelle de l'un et la structure temporelle fine de l'autre. Le son chimérique est obtenu en multipliant l'enveloppe temporelle du premier son par la structure temporelle fine du deuxième son au sein de chaque canal fréquentiel. Enfin, les signaux de bande sont additionnés sur l'ensemble des bandes. Des chimères constituées de parole (où seule l'enveloppe ou la structure fine est conservée) et de bruit de même composition spectrale permettent d'étudier le rôle respectif de chaque indice tandis que des chimères constituées de deux signaux de parole (présentant un seul indice d'AM ou de FM) permettent d'étudier la contribution simultanée des deux indices dans un nombre de bandes de fréquence donné.

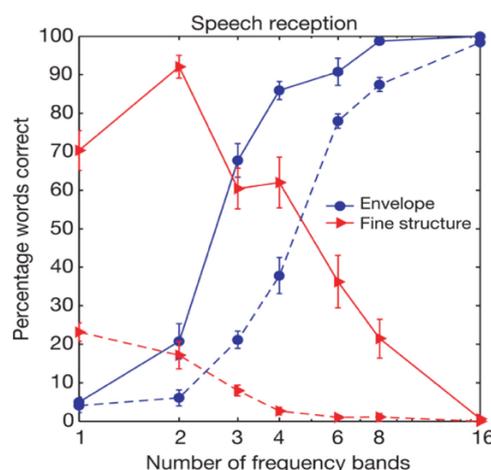


Figure 10. Scores d'intelligibilité de phrases (nombre de mots identifiés correctement) « chimériques » basés soit sur i) les informations d'enveloppe temporelle physique (AM, traits continus bleus) ou sur ii) les informations de structure temporelle fine physique (FM, traits continus rouges). Les traits continus correspondent aux conditions où l'information est véhiculée uniquement

par l'enveloppe ou la structure fine (l'autre composante étant remplacée par du bruit). Les traits en pointillés correspondent aux conditions où les informations d'enveloppe et de structure fine sont en conflit car extraites de phrases différentes. Extrait de Smith et al. (2002).

Les résultats représentés en figure 10 révèlent que, lorsque seule l'information d'enveloppe temporelle est transmise par une chimère de parole (trait continu bleu), les scores d'intelligibilité s'améliorent à mesure que le nombre de bandes fréquentielles augmente et atteint 85 % de réponses correctes à partir de 4 bandes fréquentielles. L'intelligibilité est alors quasiment parfaite. *A contrario*, lorsque seule l'information de structure temporelle fine est transmise par une chimère de parole (trait continu rouge), les sujets sont capables d'identifier correctement les signaux de parole avec seulement une ou deux bandes de fréquences ; au-delà, la performance d'identification chute drastiquement et à partir de 8 bandes de fréquence, les sujets répondent au hasard : *ils identifient donc principalement la parole sur la base des indices d'enveloppe temporelle lorsque la résolution fréquentielle du vocodeur approche la résolution fréquentielle humaine*. Dans une seconde tâche, les chimères auditives obtenues à partir de phrases distinctes sont présentées. Les résultats indiquent clairement que les auditeurs comprennent la phrase essentiellement sur la base des indices d'enveloppe temporelle physique (trait pointillé bleu) lorsque le nombre de bande de fréquence augmente et approche la résolution fréquentielle du système auditif humain. L'intelligibilité basée sur les indices de structure temporelle - quant à elle - est quasi nulle dans ces conditions (trait pointillé rouge). Les résultats de cette expérience de « conflit perceptif » corroborent et complètent donc ceux obtenus par Drullman (1995) et Shannon et al. (1995) avec des vocodeurs à bruit. En effet, ils démontrent que la compréhension orale de phrases est principalement liée à l'utilisation des indices d'enveloppe temporelle physique (la composante d'AM) et non ceux de structure temporelle fine (la composante de FM) par les auditeurs humains.

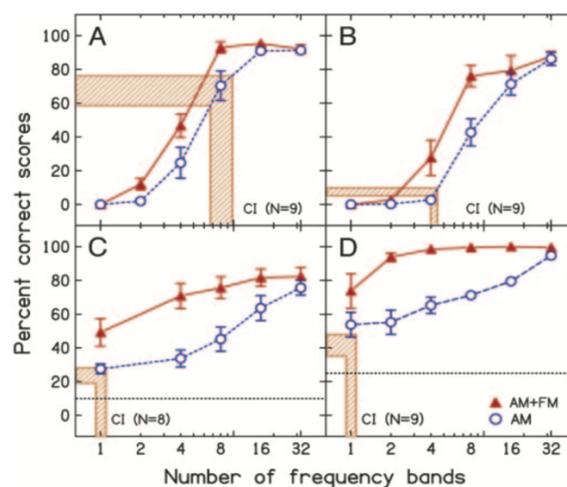


Figure 11. Scores d'intelligibilité (exprimés en pourcentage correct) de parole traitée (AM seule ou AM+FM) en fonction du nombre de bande de fréquence d'analyse chez des auditeurs normo-entendants (courbes). Les données sont comparées à celles de sujets malentendants portant un implant cochléaire (zone hachurée). Les triangles rouges correspondent à la condition « AM+FM » tandis que les cercles bleus ouverts caractérisent la condition « AM seule ». La figure 8A présente

l'intelligibilité de phrases en anglais dans le silence. La figure 8B présente l'intelligibilité de phrases en anglais en présence d'une voix concurrente (rapport signal sur bruit de 5 dB). La figure 8C présente les scores de reconnaissance de locuteur (homme, femme, garçon, fille). La figure 8D présente les scores de reconnaissance de mots en Mandarin. Extrait de Zeng et al. (2005).

Zeng et al. (2005) étendent plus avant l'étude de la contribution relative de ces deux types d'indices à l'identification de la parole en utilisant un autre type de vocodeur et en testant les auditeurs dans d'autres conditions d'écoute (identification de phrases en anglais dans le silence, identification de phrases en anglais en présence d'une voix concurrente, reconnaissance de locuteur (homme, femme, garçon, fille), identification de mots dans une langue à ton). Ici, le rôle des indices d'AM seuls et des indices d'AM et de FM combinés est évalué en fonction du nombre de bandes du vocodeur, variant de 1 à 34. Cette fois et contrairement aux études précédentes, les résultats représentés en figure 11 indiquent que les performances des auditeurs sont significativement plus faibles dans la condition d'AM seule lorsque le nombre de bande d'analyse est inférieur ou égal à 16, et ceux avant tout dans les tâches d'identification de parole masquée (dans le bruit), d'identification de mots prononcés avec une langue à ton, ou encore dans la tâche de reconnaissance de locuteur. Lorsque le nombre de bande atteint la résolution fréquentielle du système auditif humain (une trentaine de bandes entre 80 et 8000 Hz), la performance des auditeurs est au plafond dans toutes les conditions testées et il n'est plus possible d'évaluer l'effet du vocodage.

Ainsi, ces expériences suggèrent que i) les indices d'enveloppe temporelle physique (AM) jouent un rôle essentiel dans l'identification de la parole des langues non tonales comme l'anglais lorsque les signaux sont présentés dans le silence ; ii) les indices de structure temporelle fine physique (FM) pourraient malgré tout jouer un rôle important dans la perception de la parole lorsque les informations mélodiques et de timbre sont nécessaires fournissant des indices de ségrégation auditive (e.g., identification de parole en présence de parole concurrente), des informations indexicales (e.g., reconnaissance de locuteur) ou encore des indices phonétiques (e.g., identification d'une langue à ton).

Dans une expérience complémentaire, Zeng et al. (2005) estiment les seuils d'intelligibilité de phrases (SRT) masquées par une voix concurrente et vocodées de façon à ne conserver que i) les indices d'enveloppe temporelle physique (AM) au sein de 34 bandes de fréquence ou ii) une combinaison d'indices d'AM et de structure temporelle fine physique (FM) au sein du même nombre de bandes. La figure 12 (graphique de gauche) indique que les stimuli vocodés en (AM) et (AM+FM) présentent des spectres de puissance à long terme identiques à ceux des signaux originaux. Les mesures d'intelligibilité (figure 12, graphique de droite) démontrent que l'ajout d'informations de FM au sein des 34 bandes de fréquence conduit à une amélioration significative et substantielle des capacités d'identification de phrases masquées : l'apport de la FM conduit à un gain de 3 à 5 dB SRT d'intelligibilité. Dans l'ensemble, ces données de reconnaissance de la parole par vocodeur suggèrent donc qu'AM et FM contribuent à l'intelligibilité de la parole mais de manière complémentaire. La contribution de

l'AM s'observe essentiellement lorsque la parole est présentée dans le silence, et la contribution de la FM n'est observée que lorsque le signal de parole est dégradé (par masquage par un autre signal de parole, par exemple) ou lorsque la tâche requiert d'utiliser des informations mélodiques (dans le cas d'une tâche d'identification de locuteur, par exemple). Enfin, l'avantage de la FM ne semble pas lié à des indices purement spectraux à long terme.

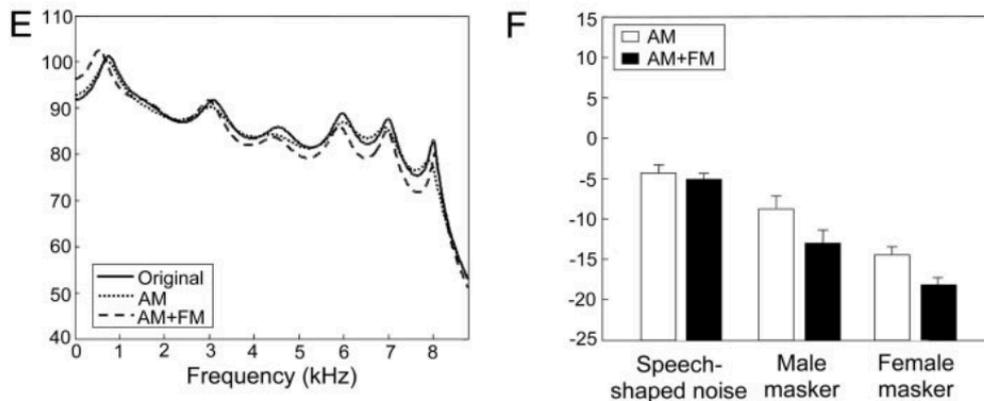


Figure 12. Spectre de puissance à long terme (E) de phrases non traitée (ligne continue) et des phrases vocodées (en AM ou en AM+FM, lignes en pointillés ou interrompues) sur 34 bandes d'analyse. Seuils de reconnaissance de la parole (F) exprimés en dB SRT et mesurés dans trois conditions de masquage (masque=bruit dont le spectre à long terme est identique à celui de la parole cible, locuteur concurrent masculin ou locuteur concurrent féminin). Extraites de Zeng et al. (2005).

De nombreuses études de nature similaire suggèrent que l'information de structure temporelle fine physique améliore significativement l'intelligibilité de la parole lorsque cette dernière est masquée par un bruit modulé lentement (e.g., 8 Hz) en amplitude. Ce phénomène est généralement dénommé « démasquage de modulation » ou « **speech masking release** ». Dans une telle situation d'écoute, les auditeurs semblent tirer avantage d'une amélioration locale (à court terme) du rapport signal sur bruit (RSB), leur permettant ainsi d'« entrevoir » (glimpse) des informations utiles sur le signal de parole cible. Ce mécanisme perceptif impliquant potentiellement des traitements de bas (sensoriel) et de haut (cognitif) niveau prend parfois le nom d'« écoute dans les vallées du bruit » (**listening in the valleys/dips**) (pour une revue, voir Füllgrabe et al., 2006 ; Léger et al., 2012). Les sujets normo-entendants présentent en effet des SRTs plus bas (meilleurs) ou une meilleure performance d'identification de syllabes et de mots dans un bruit modulé en amplitude que dans un bruit stationnaire (Duquesnoy, 1983 ; Festen & Plomp, 1990 ; Peters et al., 1998). Hopkins et Moore (2009) ont repris ce paradigme expérimental et montrés que, pour des auditeurs normo-entendants jeunes, les SRTs mesurés sur des phrases traitées par un vocodeur à bruit s'améliorent plus en présence d'un bruit masquant modulé en amplitude plutôt que dans le cas d'un bruit masquant stationnaire lorsque l'information de structure temporelle fine physique (FM) est transmise dans un certain nombre de bandes de fréquence. Des résultats similaires sont obtenus pour des syllabes isolées (Gnansia et al., 2008, 2009). Cela suggère que les sujets normo-entendants jeunes utilisent un changement d'information de FM dans les

vallées du bruit de fond modulé pour identifier la présence d'un signal de parole et coder ses propriétés.

En résumé, les résultats de ces différentes études basées sur des vocodeurs indiquent que, si les informations de phase instantanée dérivant de la FM ne sont pas nécessaires dans des tâches d'écoute simples (dans le silence, par exemple), elles se révèlent essentielles dans des tâches plus complexes comme dans le cas de la reconnaissance de la parole en présence d'une voix concurrente ou d'un bruit masquant présentant des fluctuations d'amplitude relativement lentes.

I.1.d. Caractérisation et rôle des indices d'AM et de FM dans la perception des sons musicaux

Tout comme les signaux de parole, les sons produits par des instruments de musique présentent une structure spectrale et temporelle complexe. Ce sont des signaux non stationnaires et leurs caractéristiques temporelles résultent de la variation temporelle de résonateurs (*e.g.*, la flûte) ou de cordes (*e.g.*, la guitare). De manière générale, le signal peut être modélisé comme une somme de sinusoides dont les amplitudes instantanées $A_i(t)$ (équation 1) et les fréquences instantanées $f_i(t)$ (équation 2) varient dans le temps avec la phase instantanée $\varphi_i(t)$ du signal $s(t)$ tel que :

$$s(t) = \sum_{i=1}^M A_i(t) (\sin(\varphi_i(t))) \quad (1)$$

$$f_i(t) = x = \frac{1}{2\pi} \frac{d\varphi_i(t)}{dt} \quad (2)$$

Cette approche suppose donc que tout son musical présentant une bande passante spécifique est modélisable par une combinaison d'indices d'AM et de FM. Ces indices peuvent être extraits à l'aide d'algorithmes de classification des signaux tels que ceux développés pour la classification la parole (Sekhar & Sreenivas, 2004). Les avancées récentes dans le domaine du traitement du signal audio ont conduit à des analyses plus sophistiquées de différents aspects de la régularité de la musique tels que la **rythme**, le **timbre** ou la **mélodie**, trois attributs caractéristiques des sons musicaux (Müller, 2015 ; Meredith, 2016). Ces attributs sont discutés ci-dessous.

Le **rythme** est un attribut fondamental de la musique et du langage (Liberman & Prince, 1977). Le rythme musical est structuré en plusieurs unités telles que les accents, le tempo ou différentes durées de notes (noires, croches, etc.) (Patel, 2010). Daikoku et Goswami (2020), associent cette organisation à une hiérarchie rythmique spécifique d'indices acoustiques d'AM correspondant aux temps de montée de l'enveloppe d'amplitude (voir figure 13 ci-dessous). En effet, les patterns rythmiques sont corrélés aux modulations d'amplitude inférieures à 40 Hz liées à l'attaque du son (« **onset** ») ou à l'accent d'une note. Cette

hiérarchie résulterait de l'encodage réalisé par un réseau de neurones du cortex auditif synchronisant les décharges neuronales sur la phase des enveloppes temporelles lentes inférieures à 40 Hz (plus spécifiquement entre 0.2 et 32 Hz) véhiculant l'intensité du signal (Shamma, 2001 ; Joris et al., 2004 ; Ding et al. 2017). Ce mécanisme cortical contribue à l'identification d'unités musicales comme les notes, le battement ou les mesures au sein d'un morceau de musique de la même manière que les fibres auditives nerveuses émettent des décharges neuronales synchronisées sur le rythme de la parole (Poehpel, 2003 ; Giraud & Poehpel, 2012; Obleser & Kayser, 2019). Ainsi, les modulations temporelles lentes des sons musicaux permettraient de percevoir le début ou la fin (*attaque/chute*) de séries de notes se succédant rapidement et conduisant à la perception du tempo, du mouvement ou de l'ensemble d'un groupe musical (Gordon, 1987). L'étude de Daikoku et Goswami (2020) est la première à suggérer l'émergence d'unités musicales à partir du traitement auditif des modulations temporelles. D'autres résultats suggèrent que les mécanismes perceptifs basés sur cette hiérarchie d'indices d'AM sont impliqués à la fois dans le traitement du langage et de la musique et que ces mécanismes perceptifs universels peuvent aussi expliquer les régularités entre les cultures musicales (McPherson et al., 2020).

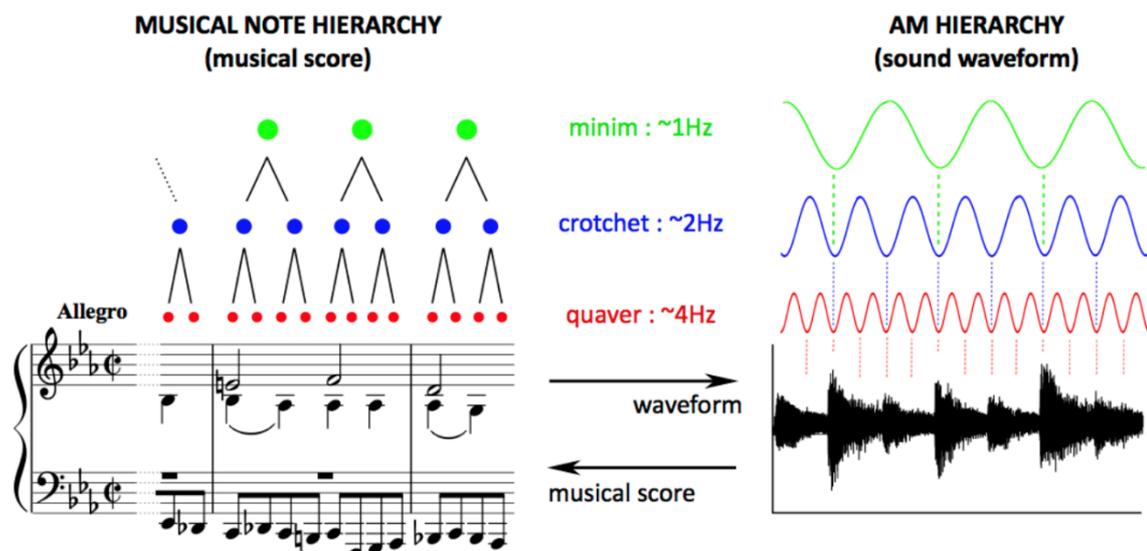


Figure 13. Structure hiérarchique du rythme dans la musique. Correspondance entre la représentation par partition musicale et la forme d'onde acoustique correspondante d'une partie des 33 variations d'une valse d'Anton Diabelli, Op. 120 de Ludwig Van Beethoven. Le rythme musical peut être organisé de manière hiérarchique. Extrait de Daikoku & Goswami (2020).

Le **timbre** est l'attribut perceptif permettant de faire la distinction entre deux sons ayant la même hauteur, la même sonie et la même durée (ANSI, 1973). L'enveloppe temporelle physique constitue un des attributs essentiels des signaux de musique pouvant varier d'un instrument à l'autre (ou même au sein d'un instrument) en fonction de la hauteur, de la sonorité et de l'expression d'une note de musique. Néanmoins, même si cette dernière joue un rôle essentiel dans la reconnaissance de la parole dans le silence, elle ne fournit pas suffisamment d'information à elle seule pour la perception de la hauteur ou l'appréciation de

la musique (Gfeller & Lansing, 1991 ; Gfeller et al., 2002, McDermott, 2004 ; Kong et al., 2004 ; Galvin et al., 2008). Néanmoins, une étude récente démontre que **les informations d'AM lentes et saillantes (cf. spectres de modulation d'amplitude en figure 14 ci-dessous, présentant un pic autour de 2 Hz) contribuent à la perception du style musical ou du timbre et à l'identification du type d'instrument** (Ding et al., 2017). La perception du timbre d'un instrument est également importante pour la ségrégation et la répartition des sources sonores dans un ensemble musical, ce qui permet d'entendre différentes compositions mélodiques individuellement ou en combinaison (Galvin et al., 2008). C'est pourquoi la reconnaissance d'instrument sur la base du timbre a été amplement étudiée, notamment à partir de signaux acoustiques modifiés par différentes techniques de traitement du signal audio (Berger, 1964 ; Saldanha & Corso, 1964). Les travaux récents d'Hajda (2007) démontrent que pour des sons purs brefs émis par des instruments produisant des sons impulsifs, l'intégrité de la structure temporelle du signal joue un rôle crucial dans l'identification de l'instrument de musique. Cependant, pour des sons purs longs, la partie soutenue est suffisante pour permettre la reconnaissance en l'absence d'indices acoustiques temporels. Ceci a également été étudié chez les personnes malentendantes équipées d'un implant cochléaire (*e.g.*, Gfeller & Lansing, 1991 ; Gfeller et al., 2002 ; McDermott & Looi, 2004 ; Nimmons et al., 2008). Dans ces études, les notes de musique ou les sons joués par différents instruments (*e.g.* violon, flûte, etc.) sont présentés à des auditeurs normo-entendants et implantés cochléaires dont la tâche consiste à identifier un instrument de musique spécifique. Les auditeurs implantés cochléaires, qui font généralement état de grandes difficultés de perception et d'appréciation musicale, sont généralement moins performants que les auditeurs normo-entendants jeunes, et présentent des scores d'identification atteignant 40 à 60 % de réponses correctes suivant le nombre d'alternatives possibles comparé à 90-100 % d'identification correcte chez les auditeurs normo-entendants (Gfeller et al., 2002; McDermott & Looi, 2004). Les mauvais scores des auditeurs implantés cochléaires sont probablement liés au fait que la résolution temporelle et spectrale d'un implant cochléaire est limitée (*e.g.*, Smith et al., 2002). Cependant, Pressnitzer et al. (2005) constatent que les auditeurs implantés cochléaires sont aussi sensibles que les auditeurs normo-entendants aux changements d'attaque et de centre de gravité spectrale qui constituent deux composantes du timbre des instruments (McAdams et al., 1995), suggérant que les auditeurs implantés cochléaires pourraient être entraînés à identifier correctement les sources sonores d'instruments. Enfin, plus récemment, l'étude de Kong et al. (2011) révèle que la perception du timbre chez les auditeurs normo-entendants implique des dimensions multiples incluant l'enveloppe temporelle, la forme spectrale et les détails spectraux fins (« irrégularité spectrale ») (McAdams et al., 1995 ; Kong et al., 2011 ; Caclin et al., 2005). Les auditeurs peuvent utiliser ces indices spectraux de manière indépendante ou en combinaison (Samson et al., 1997). Kong et al. (2012) démontrent également que les auditeurs implantés cochléaires perçoivent le timbre sur la base de deux dimensions seulement : l'enveloppe temporelle et la forme spectrale.

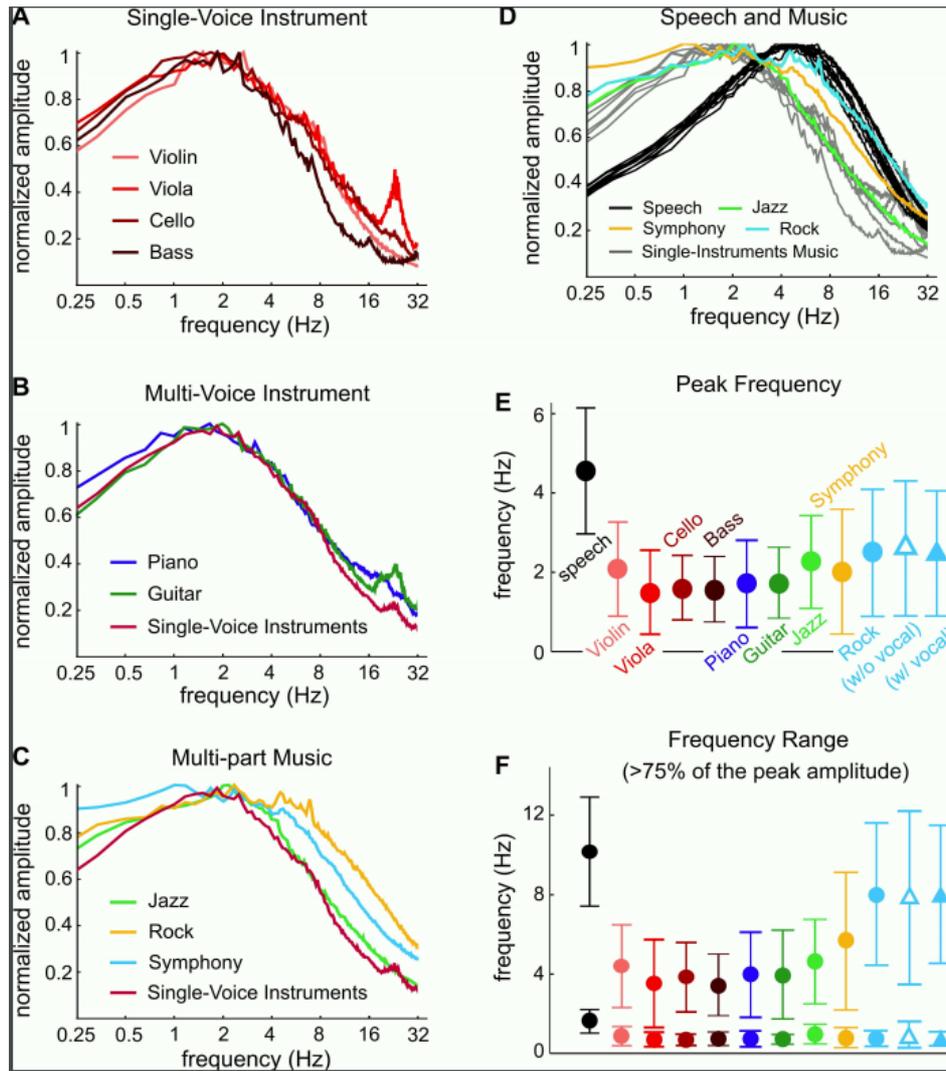


Figure 14. (A) Spectres de modulation d'amplitude d'un thème de musique classique joué par quatre instruments différents (violon, alto, violoncelle, basse) caractérisés par un pic entre 1 et 2 Hz (inférieur à celui de la parole, qui se trouve à 4-5 Hz). (B) Spectres de modulation d'amplitude de signaux musicaux joués par un piano ou une guitare. Les spectres ne diffèrent pas du spectre moyen d'instruments uniques. (C) Spectres de modulation d'amplitude de trois types de musique : quel que soit le style musical, les spectres sont semblables jusqu'à 4 Hz, bien que la musique symphonique se démarque avec plus d'énergie dans les fréquences de modulation basses (inférieures à 1.5 Hz). Au-delà de 4 Hz, les spectres de modulation de la musique rock et symphonique contiennent plus d'énergie que pour un seul instrument ou pour la musique jazz. La musique rock présente un pic de modulation plus large autour de 2-3 Hz avec beaucoup plus d'énergie de modulation que pour la parole dans les basses fréquences de modulation et moins d'énergie entre 2 et 16 Hz. (D) Spectres de la parole vs. spectres de modulation de différents styles musicaux. Les spectres de modulation sont clairement distinguables : le pic de modulation pour la musique est inférieur à 2 Hz tandis que celui de la parole est supérieur à 4 Hz. (E) : le pic de modulation de la parole est nettement supérieur à celui de la musique et de tous les styles musicaux analysés ici. Extrait de Ding et al. (2017).

La **mélodie** est un attribut résultant d'une succession de notes ayant chacune une fréquence caractéristique (Zhang et al., 2020). Elle est perçue sur la base **d'indices rythmiques** et de **hauteur** (pitch), constituée de motifs de hauteur pouvant être reproduits à travers différents niveaux (« transposés » en termes musicaux) (Dowling & Barlett, 1981 ; Andrews & Dowling, 1991). Dowling (1978) suggère que le contour mélodique et l'échelle musicale sont deux composantes essentielles dans la mémorisation des patterns de hauteur des mélodies. L'importance de ces deux types d'information semble varier avec la familiarité des signaux musicaux (Galvin et al., 2009). Le contour mélodique, c'est-à-dire le schéma des montées et descentes de la hauteur d'un son est une caractéristique importante de la mélodie. Les auditeurs normo-entendants jeunes (adultes) s'appuient à la fois sur le contour et sur la structure rythmique pour reconnaître une mélodie (Andrews & Dowling, 1991). De plus, contrairement à la parole, la reconnaissance de la mélodie nécessite considérablement plus d'indices de structure temporelle fine et d'indices spectraux, et les informations mélodiques ne sont donc pas bien retransmises par le processeur vocal d'un implant cochléaire. Grâce à leur chimériser, Smith et al. (2002) étudient le poids perceptif des indices d'enveloppe temporelle et de la structure temporelle fine physiques dans la reconnaissance de mélodies. Des chimères auditives basées sur deux mélodies différentes sont synthétisées et les signaux chimériques sont synthétisés pour mettre en conflit les indices d'enveloppe temporelle et la structure temporelle fine. Les résultats de cette expérience montrent que **les auditeurs identifient la mélodie sur la base des informations de structure temporelle fine physique (indices de FM)** entre 1 et 32 bandes fréquentielles. Au-delà (48-64 bandes spectrales), ils identifient principalement la mélodie sur la base de l'enveloppe temporelle physique. Le point d'intersection où l'enveloppe commence à dominer sur la structure fine se produit lorsque la bande passante des filtres d'analyse devient plus étroite que la bande passante des filtres cochléaires chez l'humain. Dans ce cas, l'information sur la fréquence de chaque composante peut être extraite à partir de la distribution spectrale des signaux.

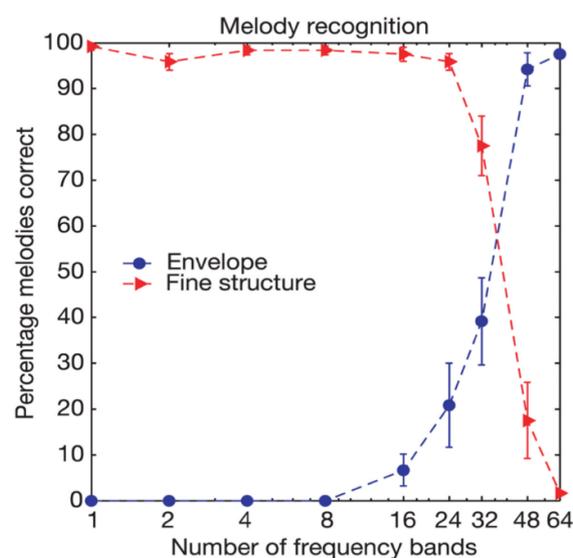


Figure 15. Résultats de la tâche de reconnaissance de mélodies basée sur l'enveloppe ou la structure temporelle fine de chimères auditives. Les chimères de mélodie-mélodie contiennent des mélodies conflictuelles dans l'enveloppe et la structure fine. Extrait de Smith et al. (2002).

Dans une autre étude basée sur un vocodeur, Zeng et al. (2005) comparent la capacité d'auditeurs normo-entendants jeunes à utiliser les indices d'AM et de FM dans la reconnaissance de mélodies construites avec et sans indices rythmiques. Les résultats de cette étude sont présentés en figure 16. Ils indiquent que les performances d'identification de mélodies vocodées sur la seule base des indices d'AM sont quasi parfaites lorsque l'indice rythmique est fourni aux sujets, et ce indépendamment du nombre de bandes. Les indices de FM ne semblent réaliser aucune contribution mais les performances sont au plafond en condition « AM seule ». Cependant, lorsque l'indice rythmique est supprimé, les indices de FM réalisent une contribution substantielle à la performance d'identification de mélodies. Trente-deux bandes spectrales sont nécessaires pour atteindre une reconnaissance parfaite de la mélodie en condition « AM seule » tandis que seulement 4 bandes sont requises dans la condition combinant indices d'AM et de FM. Entre 4 et 16 bandes, la condition AM+FM est associée à une performance significativement meilleure que la condition AM.

Prises ensemble, ces études montrent que l'information de structure temporelle fine physique (indices de FM) est nécessaire pour la reconnaissance de la musique en l'absence d'indices rythmiques. Ces indices véhiculent des informations nécessaires à la perception de la mélodie. Les informations d'enveloppe temporelle physique (AM) - quant à elles - sont également nécessaires pour la reconnaissance de la musique : elles jouent un rôle essentiel dans la perception du rythme et du timbre.

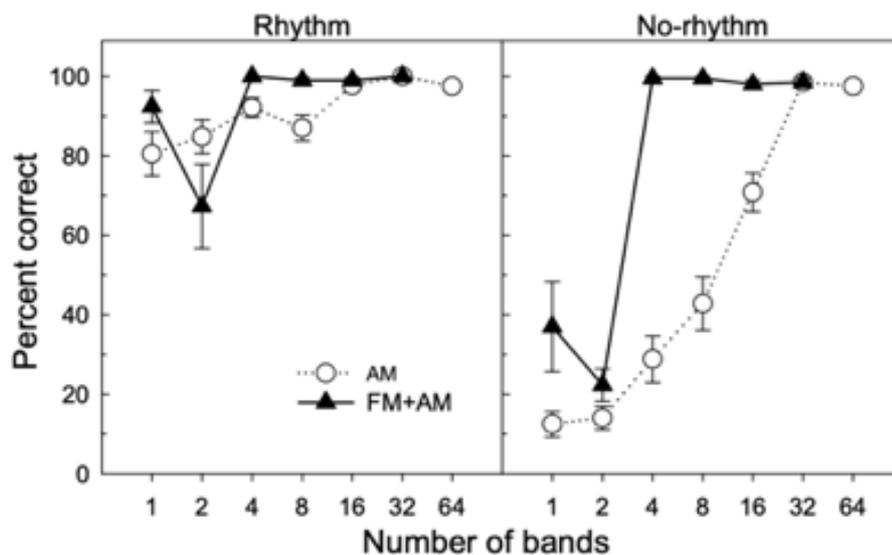


Figure 16. Performance de reconnaissance de mélodies en fonction du nombre de bandes fréquentielles du vocodeur en présence (à gauche) et en l'absence (à droite) d'indices rythmiques pour des signaux vocodés ne contenant que les indices d'AM (cercles ouverts) ou une combinaison d'indices d'AM et de FM (triangles fermés). Extrait de Zeng et al. (2005).

I.1.e. Caractérisation des indices de modulation temporelle en fonction de leur représentation dans le système auditif périphérique

La dichotomie AM/FM permet de décrire les composantes temporelles d'un signal acoustique tel qu'un paysage sonore naturel, un signal de parole ou le son d'un instrument de musique. Cette description purement physique peut être réalisée à différents niveaux de résolution spectrale (par exemple, en faisant varier le nombre et la largeur des bandes d'analyse fréquentielle d'un vocodeur ; cf. figure 1). Cette dichotomie peut être affinée en distinguant plusieurs types de composantes de modulation temporelle (l'enveloppe temporelle, ENV ; la structure temporelle fine, TFS) *en fonction de leur représentation dans le système auditif périphérique animal* (humain ou non-humain) (Heinz & Swaminathan, 2009 ; Shamma & Lorenzi, 2013 ; Moore, 2014). Cette nouvelle codification des indices de modulation temporelle véhiculés ou évoqués par un signal acoustique peut s'avérer utile lorsqu'il s'agit d'établir une distinction claire entre la description physique (c'est-à-dire acoustique) et la description biologique (ou perceptive) de ces indices d'ENV et de TFS.

L'enveloppe temporelle (**ENV**) et la structure temporelle fine (**TFS**) changent effectivement en fonction de l'étape de traitement considérée au sein du système auditif périphérique (Moore, 2014) :

- Avant le filtrage des sons larges bandes par la membrane basilaire dans la cochlée, ces indices temporels véhiculent un contenu purement physique (p) et sont alors dénommés **ENV_p et TFS_p** . L' ENV_p et la TFS_p correspondent donc à la composante d'**AM** et de **FM** du signal de bande (étroite ou large). Pour un signal de parole, l'enveloppe temporelle « physique » (ENV_p) correspond aux fluctuations d'amplitude lentes du signal dont les cadences sont essentiellement comprises entre environ 2 et 50 Hz (Rosen, 1992). Cette dernière peut être décrite comme véhiculant des indices acoustiques de durée, temps de montée et de descente du signal de parole, ou encore de tempo. Les modulations d'amplitude les plus lentes (inférieures à 16 Hz) sont déterminées par le débit syllabique, à savoir le nombre de syllabes par seconde (Greenberg et al., 2003 ; Goswami & Leong, 2013). La structure temporelle fine « physique » (TFS_p) - quant à elle - correspond aux fluctuations plus rapides, supérieures à 50 Hz. Elles correspondent à la périodicité du signal. Entre 50 et 500 Hz, elles véhiculent l'information relative au fondamental laryngé F_0 (la cadence de vibration des cordes vocales).
- A l'issue de la décomposition élaborée par la membrane basilaire (BM) dans la cochlée, ces indices sont dénommés **ENV_{BM} et TFS_{BM}** . Ils sont illustrés en figure 17 ci-dessous. ENV_{BM} et TFS_{BM} reflètent les effets de filtrage fréquentiel réalisés par la cochlée. En effet, les filtres cochléaires s'élargissant en haute fréquence audio, l' ENV_{BM} estimée pour une fréquence caractéristique de 4803 Hz présente des fluctuations plus rapides que celle estimée pour une fréquence caractéristique de 369 Hz.

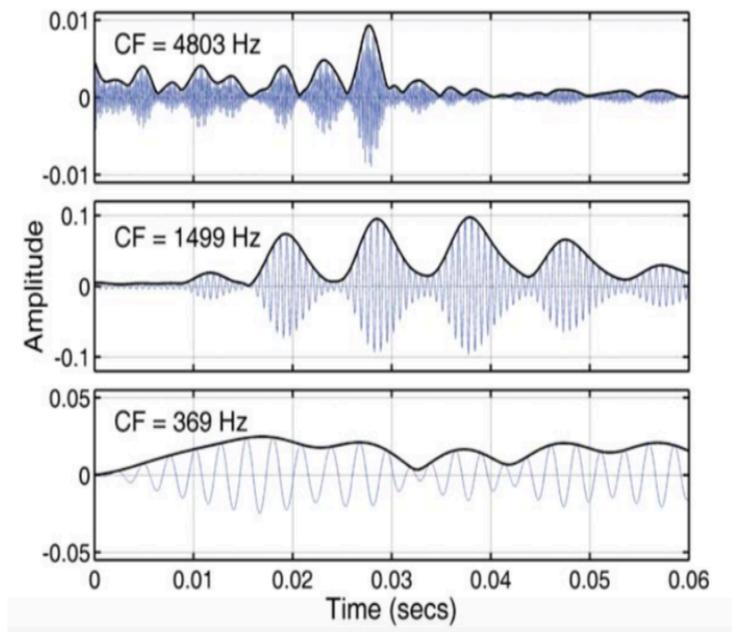


Figure 17. Sortie des filtres cochléaires simulés en réponse à un signal de parole /en/ du mot « sens » pour trois fréquences caractéristiques (369, 1499 et 4803 Hz). Dans chaque graphique, les traits épais et fins représentent respectivement l'enveloppe temporelle (ENV_{BM}) et la structure temporelle fine (TFS_{BM}) en un point donné de la membrane basilaire (BM) dans la cochlée. Extrait de Moore (2014).

- Après transduction par les cellules ciliées internes et codage neuronal (N) par les fibres afférentes de type I du nerf auditif, ces indices sont enfin représentés par des trains de potentiels d'action et dénommés ENV_N et TFS_N . L'enveloppe temporelle neurale, ENV_N , est fortement affectée par des traitements non-linéaires ayant lieu lors de l'étape de transduction et de transmission synaptique entre la cellule ciliée interne et les fibres du nerf auditif (e.g., l'adaptation nerveuse à court terme, rehaussant l'attaque du signal ; Joris et al., 2004 ; Scheidt et al., 2010). La structure temporelle fine neurale, TFS_N , est quant à elle fortement dépendante des caractéristiques de verrouillage en phase (phase locking) des décharges des fibres du nerf auditif sur la phase instantanée du signal transmis par la membrane basilaire. Elle n'est probablement encodée que jusqu'à environ 1-3 kHz chez l'humain (Verschooten et al., 2019). Sur le plan perceptif, L'enveloppe temporelle neurale, ENV_N , évoque des variations de sonie et de rythmicité dans les signaux de parole qui jouent un rôle clé dans l'identification de la parole (Drullman et al., 1994 ; Drullman, 1995 ; Shannon et al., 1995 ; Elliott & Theunissen, 2009). Sur le plan perceptif, la structure temporelle fine neurale, TFS_N , évoque des informations liées à la mélodie, au voisement et au timbre de la voix (Rosen, 1992).

La figure 18 schématise ces différents niveaux de représentation des indices de modulation temporelle en fonction de l'étape de traitement dans le système auditif périphérique.

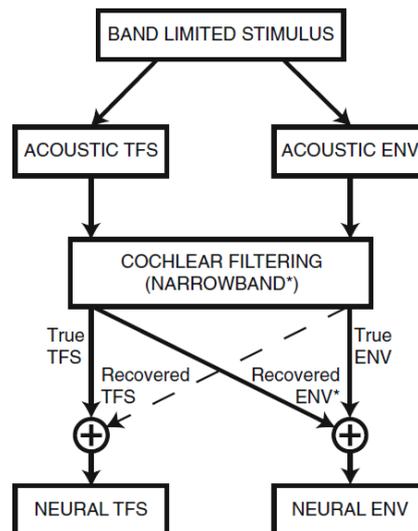


Figure 18. Représentation schématique des différents indices d'enveloppe temporelle (ENV) et de structure temporelle fine (TFS) véhiculés par un signal à bande étroite traité par le système auditif périphérique. Enveloppe et TFS acoustiques correspondent respectivement aux composantes d'AM et de FM du signal de bande. On note que la TFS acoustique (FM ou TFS_p) peut se reconvertir en TFS neurale (TFS_N) et en enveloppe neurale (ENV_N) dans les réponses des fibres du nerf auditif. Extrait de Heinz et Swaminathan (2009).

Cette représentation a le mérite de souligner l'absence d'indépendance entre l'enveloppe temporelle et la structure temporelle fine du signal (Gilbert & Lorenzi, 2006 ; Shamma & Lorenzi, 2013). En effet, les indices physiques de TFS (TFS_p) peuvent se reconvertir sous forme d'indices d' ENV_{BM} en sortie des filtres cochléaires (Gilbert & Lorenzi, 2006), ou encore dans les réponses des neurones du tronc cérébral, ENV_N (Shamma & Lorenzi, 2013).

Conformément à cette vue, les travaux de modélisation de Ghitza (2001) démontrent que les indices d'enveloppe temporelle physique (ENV_p) des signaux de parole dégradés peuvent être **reconstruits** en sortie des filtres cochléaires à partir des indices de structure temporelle fine physique (TFS_p) car les composantes d'enveloppe et de fréquence instantanée du signal ne sont pas indépendantes. Cette relation est décrite par deux théorèmes fondamentaux de traitement du signal : i) dans tout signal à bande limitée de type $s_i(t) = a_i(t)\cos\phi_i(t)$, l'enveloppe $a_i(t)$ et la phase $\phi_i(t)$ du signal sont reliées (Voelcker, 1966) ; ii) si $\phi_i(t)$ est un signal à bande limitée et si $\cos\phi_i(t)$ est l'entrée d'un filtre passe-bande alors la sortie du filtre a une enveloppe qui est reliée à la phase $\phi_i(t)$ (Rice, 1973). Ghitza (2001) prédit également que la quantité d'enveloppe reconstruite diminue à mesure que le nombre de canaux d'analyse augmente et que leur largeur diminue. Par conséquent, l'intelligibilité de la parole mesurée avec des signaux dont l'enveloppe temporelle physique (ENV_p) est dégradée et dont la structure fine temporelle physique (TFS_p) est préservée devrait être influencée par ces indices d'enveloppe reconstruits, et ce surtout dans le cas où un nombre restreint de bandes d'analyse est utilisé. Ce mécanisme déjà décrit dans des études psychoacoustiques portant sur des signaux non linguistiques (Zwicker, 1952 ; Saberi & Hafter, 1995) prend souvent le nom de « **conversion de la FM en AM** » (en d'autres termes, les variations de fréquence

instantanée peuvent se convertir en variations dynamiques d'amplitude (*i.e.* en signal d'AM) en sortie des filtres auditifs lorsque l'excursion de fréquence de la FM est plus large que la bande passante des filtres auditifs).

Les travaux de modélisation de Ghitza (2001) permettent d'expliquer pourquoi Smith et al. (2002) observent que l'intelligibilité d'un signal de parole « chimérisé » de façon à ne contenir que des indices de TFS_p est excellente lorsque le processus de vocodage est réalisé avec seulement quelques bandes de fréquence (voir figure 10 ci-dessus ; traits continus rouge). Dans ce cas, les bandes de fréquence sont beaucoup plus larges que celles des filtres cochléaires humains (> 1 ERB). En effet, des scores d'identification correcte de 90 % sont obtenus avec une et deux bandes spectrales larges et ces scores se détériorent à mesure que la résolution fréquentielle du vocodeur augmente jusqu'à atteindre le niveau du hasard à partir de 8 bandes plus étroites. Ces résultats s'accordent avec les résultats d'études plus anciennes démontrant que l'opération d'écrêtage réduisant la transmission de l'enveloppe temporelle physique (ENV_p) ne dégrade que légèrement l'intelligibilité de la parole (Licklider, 1946 ; Licklider & Pollack, 1948).

Dans une étude théorique, Zeng et al. (2004) suggèrent que le phénomène de reconstruction d'enveloppe dépend du rapport entre la bande passante d'analyse (la bande passante des filtres utilisés pour vocoder les signaux de parole) et la bande passante des filtres cochléaires. Gilbert et Lorenzi (2006) testent cette hypothèse en évaluant la capacité d'auditeurs normo-entendants jeunes à identifier des signaux de parole (des syllabes isolées) sur la base d'indices d'enveloppe reconstruits à la sortie des filtres cochléaires. Les signaux de parole originaux sont vocodés de deux manières : la première consiste à préserver uniquement les indices de structure fine physique TFS_p (l'enveloppe temporelle étant susceptible de se reconstruire par la suite en sortie des filtres cochléaires) tandis que la seconde consiste à synthétiser des signaux dont l'enveloppe temporelle a été reconstruite par simulation du filtrage réalisé par des filtres cochléaires (un banc de filtres gammatone est utilisé pour cela). Les premiers résultats montrent une forte corrélation entre l'enveloppe physique d'origine contenue dans le signal de parole non traité (ENV_p) et l'enveloppe reconstruite en sortie des filtres cochléaires (ENV_{BM}) lorsqu'une seule bande d'analyse extrêmement large est utilisée. Comme prédit par Ghitza (2001), les indices d'enveloppe sont reconstruits à partir de l'information de structure fine physique en sortie des filtres cochléaires lorsque le traitement repose sur une seule bande de fréquence. L'augmentation du nombre de bande d'analyse (>8) a un effet délétère sur la reconstruction d'enveloppe. Ainsi, les résultats des tests d'intelligibilité démontrent que ces syllabes isolées dont l'enveloppe est reconstruite naturellement en sortie des filtres auditifs - c'est-à-dire les signaux de parole vocodés ne contenant que de la structure fine physique) (cercles pleins figure 19 ci-dessous) sont très intelligibles (performance comprise entre 90 et 100%) quel que soit le nombre de bandes d'analyse (1 à 16). Par contre, les syllabes dont l'enveloppe est obtenue par simulation du processus de reconstruction cochléaire (cercles ouverts, figure 10) ne sont intelligibles que

lorsque le nombre de canaux utilisés pour l'extraction de la TFS_p est faible (en accord avec Ghitza, 2001). Lorsque le nombre de canaux est supérieur à 8, la capacité d'identification des syllabes est médiocre. **Cette étude indique que les indices d'enveloppe reconstruits (ENV_{BM}) à partir de la structure fine physique (TFS_p) permettent d'identifier des syllabes sans signification lorsque la bande d'analyse est plus large que celle des filtres cochléaires humains. Leur contribution à l'identification des syllabes s'effondre lorsque la bande d'analyse s'approche de celle des filtres cochléaires. Toutefois, les informations de structure fine physique (TFS_p) semblent véhiculer des informations phonétiques utilisables par les auditeurs et ce indépendamment du processus de reconstruction.**

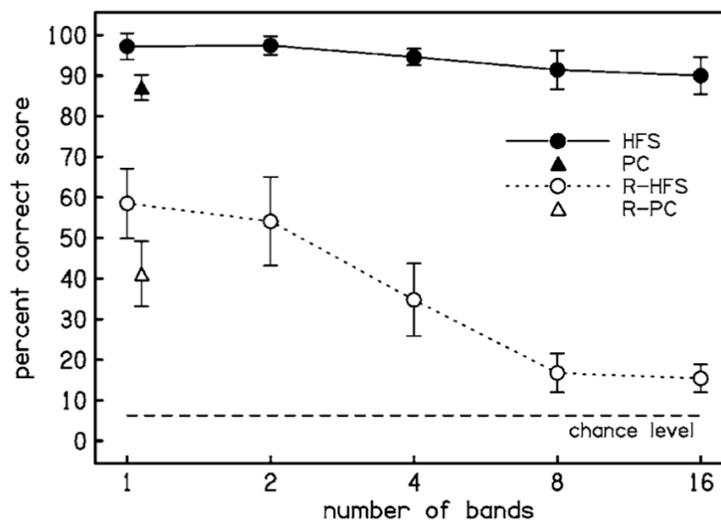


Figure 19. Scores d'identification obtenus pour des syllabes isolées vocodées de façon à ne préserver que l'information de structure fine physique TFS_p (cercles pleins) au sein d'un nombre variable de bandes de fréquence (nombre de bandes en abscisse) et pour des syllabes vocodées de façon à ne véhiculer que des informations d'enveloppe temporelle « reconstruites » à la sortie d'un modèle de banc de filtres cochléaires (banc de filtres gammatone), ENV_{BM} (cercles ouverts). L'intelligibilité des syllabes est élevée (>50%) lorsque le signal de TFS_p est vocodé à partir de 1 ou 2 bandes de fréquence larges. Elle chute au niveau du hasard (trait interrompu) au-delà de 8 bandes de fréquence. La résolution fréquentielle humaine est approchée à partir de 32 bandes de fréquence. Extrait de Gilbert et Lorenzi (2006).

Les indices d'enveloppe reconstruits à partir de la TFS physique (TFS_p) sont souvent interprétés comme un artefact produit par les vocodeurs (Zeng et al., 2004 ; Gilbert & Lorenzi, 2006). Il est toutefois possible d'envisager que la capacité à extraire des informations sur l'ENV physique à partir de la TFS physique représente une forme de traitement biologique « robuste » du signal (Heinz & Swaminathan, 2009, Swaminathan & Heinz, 2012 ; Shamma & Lorenzi, 2013). En effet, ce mécanisme de reconstruction peut s'avérer très utile lorsque l'enveloppe temporelle physique est dégradée par le canal de transmission (Licklider, 1946; Licklider & Pollack, 1948; Won et al., 2012, 2014). Bruit masquant, écrêtage, compression d'amplitude, réverbération sont autant de distorsions de l'enveloppe physique produites par des canaux de transmission variés (habitat naturel, salle, prothèse auditive, téléphone, etc.) dont les effets délétères sur l'intelligibilité sont susceptibles d'être compensés par ce processus biologique de reconstruction, dont l'origine est cochléaire et/ou neurale.

Il s'ensuit logiquement qu'une **perte de sélectivité fréquentielle cochléaire - telle qu'elle apparaît généralement dans le cas de lésions cochléaires associées à une perte auditive neurosensorielle - affecte l'efficacité de ce processus de reconstruction**. Cette hypothèse a été testée par Lorenzi et al. (2012). Dans cette étude, les capacités d'identification de syllabes sans signification sont mesurées chez des sujets normo-entendants et malentendants présentant une perte neurosensorielle modérée. Les syllabes sont traitées par vocodeur de façon à ne conserver que l'information de FM (TFS_p) au sein de trois bandes de fréquence larges (>1 ERB). Les résultats (représentés en figure 20 ci-dessous) indiquent que les sujets malentendants présentent une performance d'identification significativement inférieure à celle des sujets normo-entendants, conformément à l'hypothèse selon laquelle le processus de reconstruction est moins efficace chez ces derniers car la largeur de leurs filtres cochléaires s'approche de celle des filtres d'analyse du vocodeur.

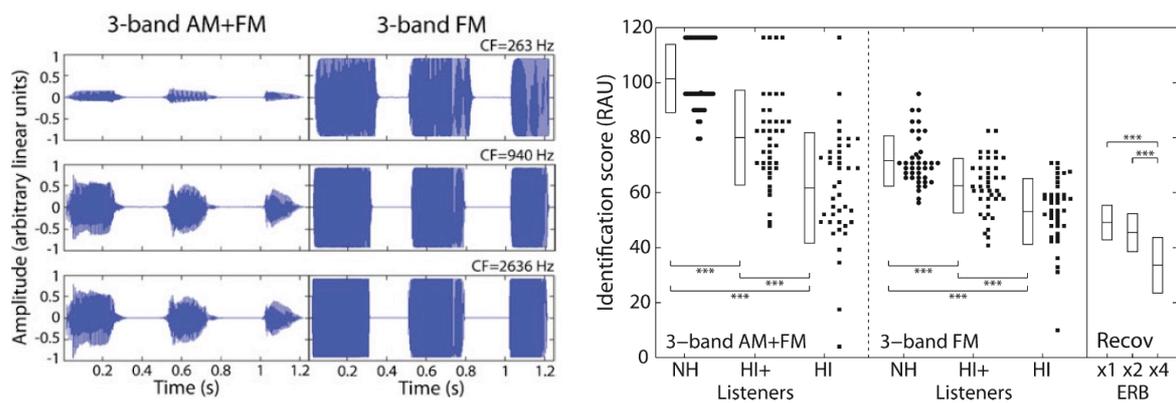


Figure 20. Graphique de gauche : signaux de parole (e.g., syllabes /ababa/ici) utilisés dans les expériences d'identification. Ces signaux sont décomposés en 3 bandes de fréquence larges (dont la fréquence centrale, CF, est égale à 263, 940 et 2636 Hz). Les signaux de bande sont soit intacts (condition 3-band AM+FM) soit vocodés de façon à ne conserver que l'information de TFS_p (FM) au sein des 3 bandes de fréquence (condition 3-band FM). Graphique de droite : scores d'identification obtenus pour des syllabes intactes (3-band AM+FM) et vocodées (3-band FM) mesuré chez des sujets normo-entendants (NH) et malentendant (HI) à deux niveaux de présentation (celui utilisé pour les NH (niveau HI) et un niveau plus élevé (niveau HI+) de façon à compenser la perte d'audibilité des malentendants). Les performances des sujets malentendants sont significativement inférieures à celles des sujets normo-entendants, car la reconstruction d'enveloppe est moins efficace chez ces derniers. Extrait de Lorenzi et al. (2012).

En résumé, les signaux acoustiques - qu'il s'agisse de sons environnementaux, de parole ou de musique - peuvent être décrits comme une somme de composantes fréquentielles « porteuses » modulées plus ou moins lentement en amplitude (AM) et en fréquence (FM). Les caractéristiques de ces modulations temporelles diffèrent fortement entre catégories de signaux. Ces différences peuvent être mises en évidence par l'estimation des spectres de modulation (spectres d'AM, de FM, MPS) des signaux. Le système auditif humain analyse ces modulations physiques à la sortie de chaque filtre cochléaire, puis de structures neurales plus centrales pour conduire à la perception (détection, discrimination, identification) du signal. Ces indices physiques d'AM et de FM sont ainsi recodés sous la forme d'indices d'« enveloppe temporelle » (ENV, variation du taux de décharge moyen) et de « structure temporelle fine » (TFS, verrouillage en phase des décharges sur la phase instantanée du signal à bande étroite) dans les trains d'impulsions nerveuses des neurones auditifs. Ces indices auditifs d'ENV et de TFS neurales jouent des rôles importants et parfois complémentaires dans la perception auditive des sons environnementaux, de parole ou de musique. Il est toutefois crucial de ne pas assimiler les indices physiques d'AM et FM aux indices auditifs d'ENV et de TFS neurale car il n'y a pas de correspondance bi-univoque entre AM/FM et ENV/TFS. En effet, la FM peut être à la fois convertie sous forme d'ENV et de TFS neurales dans le système auditif périphérique. Des lésions du système auditif périphérique sont susceptibles de réduire ce processus de conversion de la FM sous forme d'ENV, et de participer aux troubles de reconnaissance systématiquement observés chez les personnes malentendantes.

I.2. Encodage et traitement des indices d'AM et de FM par le système auditif

I.2.a. Préambule

Les indices d'AM et de FM sont reçus, encodés et traités par le système auditif, conduisant à la détection, discrimination et l'identification de signaux acoustiques complexes tels que les sons environnementaux, les sons de parole, ou les sons produits par les instruments de musique. Différentes méthodes psychoacoustiques et neurophysiologiques ont été employées pour caractériser la manière dont le système auditif animal humain ou non humain démodule ces sons complexes. En psychoacoustique, celles-ci consistent à mesurer les capacités de détection ou de discrimination d'indices d'AM et de FM de quatre manières: i) en faisant varier un **paramètre expérimental** dans la tâche (*e.g.*, la cadence, la profondeur ou la phase de modulation); ii) en utilisant des techniques de **masquage de modulation ou d'interférence** entre les indices d'AM et de FM (masquage simultané ou postérieur, au sein ou entre bandes fréquentielles audio); iii) en comparant les capacités de détection, discrimination ou identification d'indices d'AM et de FM entre **différentes populations d'auditeurs** (sujets normo-entendants jeunes, sujets normo-entendants âgés et sujets

malentendants jeunes ou âgés) ; ou encore iv) en caractérisant le **développement** des capacités de détection, discrimination ou identification auditive depuis la naissance jusqu'à l'âge adulte. En neurophysiologie, différentes mesures neuronales sont utilisées pour identifier la manière dont ces indices sont traités. Elles ne seront que brièvement décrites dans cette thèse. Ces mesures s'appuient essentiellement sur l'estimation du *taux de décharges moyen* (le nombre de potentiels d'action évoqués par différents cycles de modulation) et le *taux de synchronisation* des trains de potentiels d'action sur l'enveloppe temporelle ou la structure temporelle fine. De nombreuses autres mesures peuvent en être dérivées.

Néanmoins, de nombreux débats perdurent quant à la manière dont les indices d'AM et de FM sont encodés dans les réponses du nerf auditif humain en termes d'indices neuronaux d'enveloppe et de structure fine comme mentionné à la fin du chapitre précédent. En effet, ces études démontrent qu'il n'existe pas de correspondance exacte entre les indices d'AM (ENV_p) et de FM (TFS_p) véhiculés par le signal *acoustique* et l'enveloppe temporelle (ENV_N) et la structure temporelle fine (TFS_N) *neurales* des réponses du nerf auditif ou du tronc cérébral (Heinz & Swaminathan, 2009 ; Shamma & Lorenzi, 2013).

Aujourd'hui, les mécanismes sensoriels de codage sensoriel et de traitement central de l'AM sont suffisamment bien compris pour déboucher sur des **modèles informatiques (dits « computationnels »)** capables de prédire les capacités de détection, de discrimination et d'identification de sujets humains normo-entendants avec une précision remarquable (Dau, 1996 ; 1997a ; 1997b ; Ewert & Dau, 2000 ; 2004 ; Biberger & Ewert, 2016 ; Wallaert et al., 2017, 2018). La nature exacte des mécanismes de codage sensoriel et de traitement central de la FM, quant à elle, fait encore débat. Toutefois, de nombreuses données issues d'études psychophysiques, neurophysiologiques et de simulation militent en faveur de l'idée selon laquelle les mécanismes responsables de l'encodage de l'AM et de la FM se recoupent partiellement, et ce tout au moins pour les indices de FM qualifiés de « rapides » (>5-10 Hz ; Zwicker, 1952 ; Saberi & Hafter, 1995 ; Moore & Sek, 1996 ; Shamma & Lorenzi, 2013 ; Wallaert et al., 2018 ; Whiteford et al., 2020). Dans ce cas, les indices neuronaux correspondraient principalement à des indices d'enveloppe temporelle (ENV_N). D'autres études suggèrent que le codage sensoriel de la FM ferait appel à un deuxième mécanisme de nature purement temporelle (Siebert, 1970 ; Goldstein & Sruлович, 1977 ; Moore & Sek, 1996 ; Shamma & Lorenzi, 2013 ; Wallaert et al., 2018), et reposerait sur des indices neuronaux de structure temporelle fine (TFS_N).

L'importance de la dimension temporelle des signaux acoustiques se manifeste très tôt pendant le développement typique au travers de la spécialisation anatomique et fonctionnelle du système auditif (pour une revue, voir Joris et al., 2004). Une étude psychophysique récente démontre que la capacité d'extraction de ces indices serait fonctionnelle dès l'âge de six mois de vie (Cabrera et al., 2015). Ce phénomène pourrait ainsi contribuer à l'apprentissage

précoce du langage et fournir des informations utiles permettant de comprendre comment le système auditif d'un sujet adulte a acquis cette capacité remarquable de traitement des signaux de communication.

Ce deuxième chapitre a pour objectif de récapituler les principales théories actuelles décrivant les mécanismes sensoriels responsables du traitement de l'AM (dans une première partie) et de la FM (dans une seconde partie). Les résultats d'études psychoacoustiques menées chez l'adulte ou l'enfant (études dites « développementales ») et d'études de modélisation computationnelle sont présentés et discutés ci-dessous.

I.2.b. Traitement de l'AM par le système auditif

I.2.b.i. Traitement par le sujet normo-entendant adulte jeune

Un des moyens d'évaluer le codage sensoriel de l'AM par le système auditif humain consiste à estimer la capacité d'auditeurs normo-entendants à **détecter** les modifications d'amplitude d'un son au cours du temps. Ceci est réalisé en mesurant les seuils et les performances de détection de signaux modulés en amplitude. Le signal d'AM utilisé est généralement synthétisé à partir d'un son pur (une onde sinusoïdale) modulé par un autre son pur dont l'équation est la suivante : $s(t) = [1 + m \sin(2\pi f_m t)] \sin(2\pi f_c t)$ (1)

Cette équation représente la forme de l'onde acoustique $s(t)$ d'un son pur de fréquence porteuse f_c , dont l'amplitude est modulée par une fonction de modulation (la modulante) de cadence plus lente (f_m) à une profondeur de modulation m comprise entre 0 et 1. Lorsque f_c est largement supérieur à f_m le premier terme $[1 + m \sin(2\pi f_m t)]$ constitue l'enveloppe temporelle d'amplitude du signal. Plusieurs types de porteuses peuvent être utilisées : i) un son pur [comme c'est le cas en (1)] ; ii) un train d'impulsions brèves (clics) ; ou encore iii) un bruit à bande étroite ou large. Un son pur présente l'avantage de n'introduire aucune fluctuation susceptible d'interférer avec la perception de l'AM cible, mais les composantes spectrales latérales (sidebands) engendrées par la modulation (à $f_c + f_m$ et $f_c - f_m$ dans le cas d'une AM sinusoïdale de cadence f_m) sont susceptibles d'être détectées par l'observateur lorsque la cadence de modulation est élevée et lorsque les composantes spectrales sont résolues par les filtres cochléaires (Kohlrausch et al., 2000). En effet, la modulation d'amplitude introduit des composantes spectrales (dites « composantes latérales ») de part et d'autre de la fréquence porteuse f_c . Plus la cadence de modulation est élevée, plus l'écart entre f_c et ces composantes latérales s'accroît (et ces dernières sont susceptibles d'être détectées par un filtre cochléaire accordé sur une fréquence distincte de f_c). Dans ce cas, la modulation n'est plus perçue en tant que telle (*i.e.*, de manière temporelle, c'est-à-dire dynamique), comme grâce à un indice stationnaire de nature spectrale (et donc tonale). Un bruit « porteuse » peut être préféré à un son pur, car il masque potentiellement ces indices spectraux – tout au moins lorsque la cadence de modulation est lente. Toutefois, un bruit à

bande étroite présente des « fluctuations aléatoires intrinsèques » dont le spectre de modulation (de forme triangulaire) est déterminé par la largeur de bande du signal (Lawson & Uhlenbeck, 1950). Ce type de porteuse s'avère utile pour sonder les mécanismes de perception d'AM car ces fluctuations aléatoires sont susceptibles de masquer les modulations cibles à détecter (Dau et al., 1997a ; 1997b). Enfin, un bruit large bande peut être utilisé afin d'empêcher l'observateur de détecter les composantes spectrales individuelles du spectre du stimulus (Viemeister, 1979) car le spectre à long terme d'un bruit blanc modulé en amplitude est équivalent à celui d'un bruit blanc non modulé.

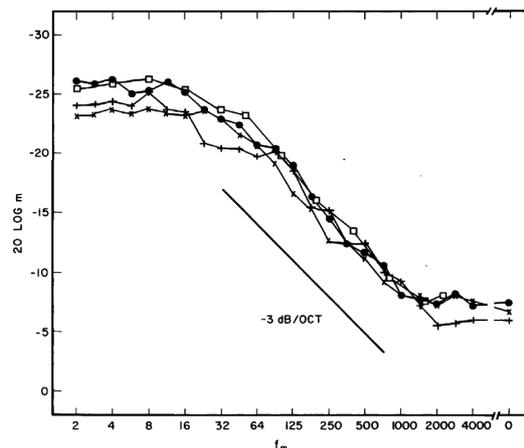


Figure 21. Fonction de transfert de modulation temporelle (TMTF) mesurée sur des sujets normo-entendants adultes avec un bruit large bande comme porteuse. La TMTF présente les seuils ($20 \log m$, en ordonnée) en fonction de la cadence de modulation f_m (en abscisse). Les « bons » seuils sont représentés en haut (e.g., -25 dB autour de 4 Hz), et les « mauvais » seuils en bas (e.g., -6 dB autour de 1000 Hz). Le plus mauvais seuil mesurable est de 0 dB ($20 \log m$ avec $m=1$ ou 100%). Extrait de Viemeister (1979).

La fonction décrivant la relation entre le seuil de détection d'AM (m ou $20 \log m$ au seuil) et la fréquence de modulation f_m prend le nom de **fonction de transfert de modulation temporelle (TMTF)** (Viemeister, 1979). Cette fonction, lorsqu'elle est mesurée avec un bruit large bande comme porteuse (Viemeister, 1979), est essentiellement une fonction de type passe-bas avec une fréquence de coupure autour de 50 Hz et une pente de 3 dB par octave (*i.e.*, les seuils de détection d'AM sont meilleurs pour des cadences de modulation lentes (e.g., 10 Hz) que rapides (e.g. 100 Hz). La limite supérieure de détection de la modulation est de 1-2 kHz. Cette fonction est représentée en figure 21. Cette fonction, lorsqu'elle est mesurée avec un son pur comme porteuse (Kohlrausch et al., 2000), présente deux régimes distincts : elle est essentiellement de type passe-bas avec une fréquence de coupure autour de 100-150 Hz, puis les seuils s'améliorent à des cadences de modulation plus élevées. Plus la fréquence porteuse augmente, plus la cadence de modulation à partir de laquelle cette amélioration de sensibilité est observée augmente. Cette fonction est représentée en figure 22.

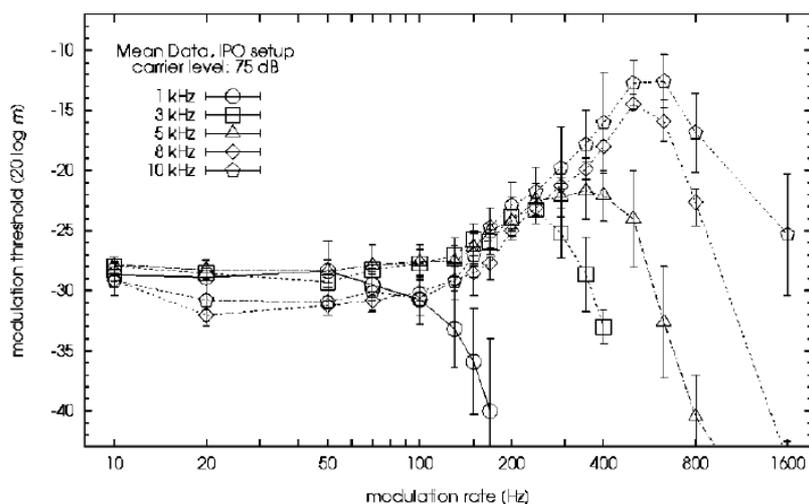


Figure 22. Fonction de transfert de modulation temporelle (TMTF) mesurée sur des sujets normo-entendants adultes avec un son pur de fréquence égale à 1, 3, 5, 8 ou 10 kHz comme porteuse. Les seuils sont relativement constants (et égaux à -30 dB) jusqu'à environ 100 Hz et se dégradent au-delà, puis ils s'améliorent fortement. La fréquence à laquelle le seuil s'améliore augmente avec la fréquence porteuse. La dégradation de sensibilité observée autour de 100 Hz révèle le fait que l'acuité temporelle est limitée. L'amélioration de sensibilité observée au-delà de 100 Hz résulte de la résolution des composantes spectrales latérales par les filtres cochléaires. Extrait de Kohlrausch et al. (2000).

L'amélioration de sensibilité observée au-delà de 100-150 Hz résulte de la résolution des composantes spectrales latérales (induites par la modulation) par les filtres cochléaires. Par ailleurs, plus la fréquence porteuse augmente, plus la largeur des filtres cochléaires accordés sur cette porteuse augmente (la bande passante des filtres cochléaires étant proportionnelle à la fréquence centrale des filtres), et la résolution des composantes spectrales latérales (induites par la modulation) par les filtres cochléaires diminue. La fréquence de coupure (f_{cutoff}) à partir de laquelle les seuils de détection d'AM se dégradent est donc deux fois plus basse avec un bruit large bande qu'avec un son pur. Ceci s'explique par les effets de masquage de modulation produits par les fluctuations aléatoires intrinsèques véhiculées par le bruit porteuse (Dau et al., 1997a ; 1997b). Les seuils de détection d'AM se dégradent toutefois au-delà de 100-150 Hz en l'absence de telles fluctuations masquantes, suggérant que l'acuité temporelle est limitée, la constante de temps associée à cette limitation étant d'environ 1 ms ($\tau=1/2\pi f_{cutoff}$).

Ces données de sensibilité à l'AM corroborent l'idée selon laquelle la détection d'AM est soumise à deux processus sensoriels rétro-cochléaires successifs opérant sur l'enveloppe temporelle neurale des signaux (ENV_N). L'un des processus correspond une **décomposition de l'enveloppe neurale des signaux à bande étroite issu du filtrage cochléaire grâce à un banc de filtres accordés dans le domaine de l'enveloppe temporelle** (Bacon & Grantham, 1989 ; Houtgast, 1989 ; Dau et al., 1997a ; 1997b ; Lorenzi et al., 1997 ; Ewert & Dau, 2000 ; Millman et al., 2002). Ces filtres sont relativement larges (leur facteur de qualité Q est de l'ordre de 1 ; Ewert & Dau, 2000 ; Lorenzi et al., 2001) et la plus haute fréquence centrale de

ces filtres est probablement de l'ordre de 120 Hz (Moore et al., 2009). Ces filtres - dont la forme est déduite d'expériences de masquage de modulation - sont représentés en figure 23. L'autre processus, indépendant du premier, correspond à un **filtrage de type passe-bas de l'enveloppe neurale** avec une fréquence de coupure de l'ordre de 100-150 Hz (Kohlrausch et al., 2000) (ou en d'autres termes, un processus d'intégration temporelle dont la constante de temps est d'environ 1 ms). Cette modélisation du traitement d'AM dépasse les premiers modèles de détection de modulation à base de « pattern d'excitation » cochléaires (Zwicker 1952) - et est compatible avec les données électrophysiologiques obtenues chez l'animal humain ou non-humain dans les structures et circuits du tronc cérébral et du cortex auditif (Joris et al., 2004).

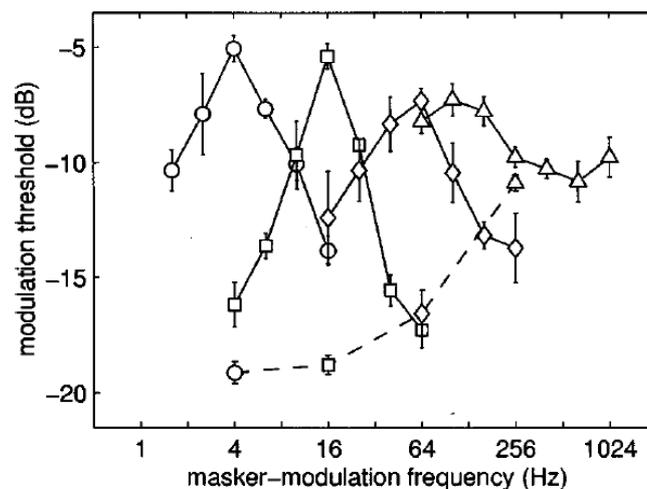


Figure 23. Seuils de détection d'AM masqués pour des cadences de modulation cibles de 4 Hz (cercles), 16 Hz (carrés), 64 Hz (diamants) et 256 Hz (triangles). La profondeur de la modulation cible au seuil est représentée en fonction de la cadence de l'AM masqué. Extrait d'Ewert et Dau (2000).

La sensibilité à la modulation n'est pas uniquement déterminée par les caractéristiques de la porteuse et la cadence de modulation. Elle n'est pas non plus déterminée uniquement par des processus purement sensoriels. A titre d'exemple, les seuils de détection de modulation s'améliorent (*i.e.*, diminuent) lorsque le nombre de cycles de modulation augmente. Ce phénomène prend le nom d'« **intégration temporelle** (Viemeister, 1979 ; Sheft & Yost, 1990) et s'observe pour l'AM comme pour la FM (Hartmann & Klein, 1980). L'intégration temporelle est interprétée comme résultant des propriétés du **mécanisme central (cognitif) de prise de décision** dans la tâche de détection. Il s'agirait soit d'un mécanisme de combinaison d'observations (looks) multiples (Viemeister & Wakefield, 1991), soit d'un mécanisme d'appariement de gabarit (Dau et al., 1997b), à savoir un mécanisme de nature corrélationnelle, dont l'efficacité augmente avec le nombre d'échantillons qui lui sont proposés. Dans le cas d'un mécanisme d'appariement de gabarit, le sujet prend sa décision en comparant la sortie de ses filtres de modulation en réponse à un signal donné à une représentation interne stockée en mémoire à long terme (un gabarit ; template). Les travaux de Wallaert et al. (2016, 2017, 2018) indiquent que l'intégration temporelle pour l'AM est plus

importante pour une cadence de modulation de 20 Hz que pour une cadence de 2 Hz. Ceci est interprété comme reflétant des limitations cognitives imposées par la *mémoire auditive à court terme*, car pour un même nombre de cycles, la durée du signal d'AM de cadence lente est 10 fois supérieure à celle du signal d'AM de cadence rapide. La constante de temps de ce registre de stockage des indices d' ENV_N en **mémoire à court terme** est estimée à environ 1 s (Ardoint et al., 2008).

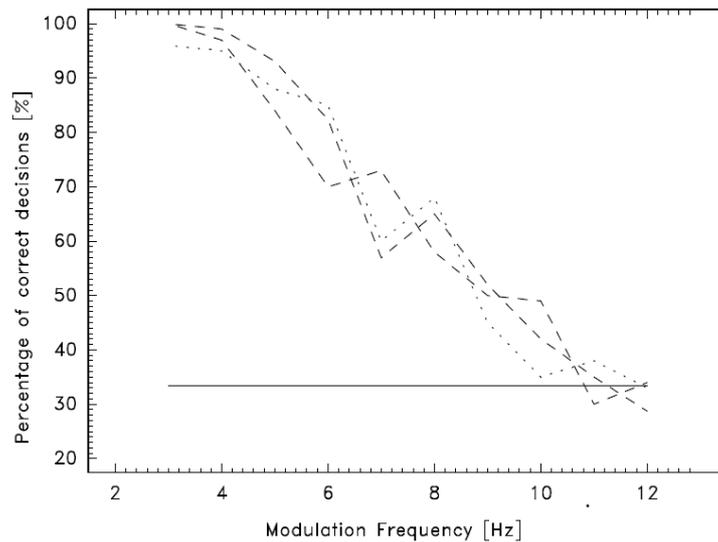


Figure 24. Performance de discrimination d'un changement de phase de départ d'une AM sinusoïdale (en abscisse) en fonction de la cadence de l'AM (en abscisse). La porteuse est un son pur de 5 kHz, la durée de l'AM est de 500 ms et la profondeur d'AM est de 100%. Le paradigme est de type 3IAFC. L'intervalle cible contient une AM sinusoïdale dont la phase de départ est tirée aléatoirement entre 0 et 365° d'essai en essai. Les AM présentées dans les deux intervalles de comparaison présentent une phase différent de l'AM cible de 180°. Extrait de Dau (1996).

Enfin, un aspect important du traitement de l'AM est rarement discuté et pris en compte, à savoir le fait que la capacité à suivre une AM dépend de sa cadence. En effet, les expériences psychophysiques portant sur la sensibilité à la **phase de départ** d'une AM sinusoïdale révèlent que la capacité à discriminer des différences de phase se dégrade abruptement au-delà de 5 Hz et est abolie à partir de 10 Hz (Dau, 1996). Cette limite s'accorde bien avec le changement de percept associé aux AM lentes vs. rapides. En dessous de cette cadence, l'AM est perçue comme une fluctuation de sonie. Au-dessus de cette cadence, l'AM évoque une sensation de rugosité puis de hauteur. Ceci est illustré dans la figure 24 ci-dessus. Dans une version du modèle computationnel de traitement auditif utilisant un banc de filtres de modulation, ceci se traduit par le fait que les basses cadences de modulation (< 10 Hz) sont transformées sans aucune perte d'information sur la phase de modulation alors que pour les cadences de modulation supérieures à 10 Hz, seule l'enveloppe temporelle des signaux en sortie des filtres de modulation (*i.e.*, la « venelope ») est analysée plus en détails (Dau et al., 1997).

I.2.b.ii. Etudes développementales

Les études comportementales tentant de caractériser la maturation des capacités de détection d'AM chez l'enfant sont rares. Les résultats de ces études indiquent que les seuils de détection d'AM tendent à diminuer (s'améliorer) avec l'âge ; toutefois, certains résultats semblent contradictoires : si certaines études indiquent que les seuils de détection d'AM s'améliorent davantage à certaines cadences (lentes ou rapides), d'autres rapportent une diminution des seuils sur l'ensemble des cadences testées (Hall & Grose, 1994; Lorenzi et al., 2000 ; Banai et al., 2011 ; Dawes & Bishop, 2008; Peter et al., 2014 ; Buss et al., 2019 ; Cabrera et al., 2019 ; Walker et al., 2019).

Comme indiqué en figure 25 ci-dessous, Hall et Grose (1994) observent que les enfants de 4-5 ans et 6-7 ans présentent des seuils de détection d'AM significativement plus élevés (plus mauvais) que les adultes et ce pour toutes les cadences de modulation testées comprises entre 5 et 200 Hz. Par contre, les seuils de détection d'AM ne diffèrent pas significativement entre les enfants de 9-10 ans et les adultes. Point important, les TMTF des différents groupes se superposent parfaitement, indiquant que la forme passe-bas des TMTF – et donc l'acuité temporelle auditive – a atteint son niveau adulte dès 4-5 ans. Des mesures de TMTF obtenues sur le nourrisson corroborent cette conclusion en suggérant que l'acuité temporelle pourrait être mature bien plus tôt, à savoir dès les premiers mois de vie (Walker et al., 2019).

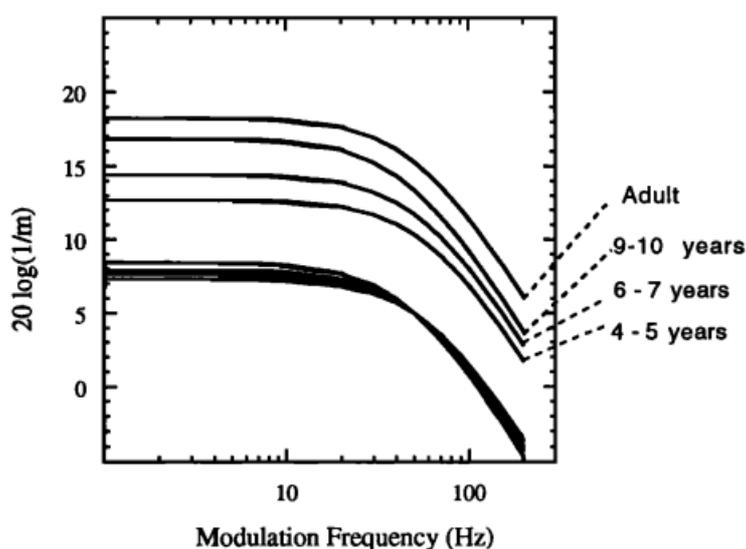


Figure 25. Fonctions de transfert de modulation temporelle (TMTF) résultant d'ajustements (moindres carrés) aux données empiriques moyennes obtenues sur 4 groupes d'âge. Ces fonctions sont superposées dans la partie basse du graphique afin de démontrer la similitude des formes des fonctions. Extrait de Hall et Grose (1994).

Les résultats de Peter et al. (2014) représentés en figure 26 ci-dessous diffèrent de ceux obtenus par Grose et Hall (1994) ou Walker et al., (2019). Cette étude porte sur trois groupes

d'auditeurs (le premier âgé de 8 à 10 ans, le second âgé de 12 à 17 ans et le troisième âgé de 18 à 32 ans). Les mesures révèlent des effets développementaux limités à la cadence la plus lente de 4 Hz.

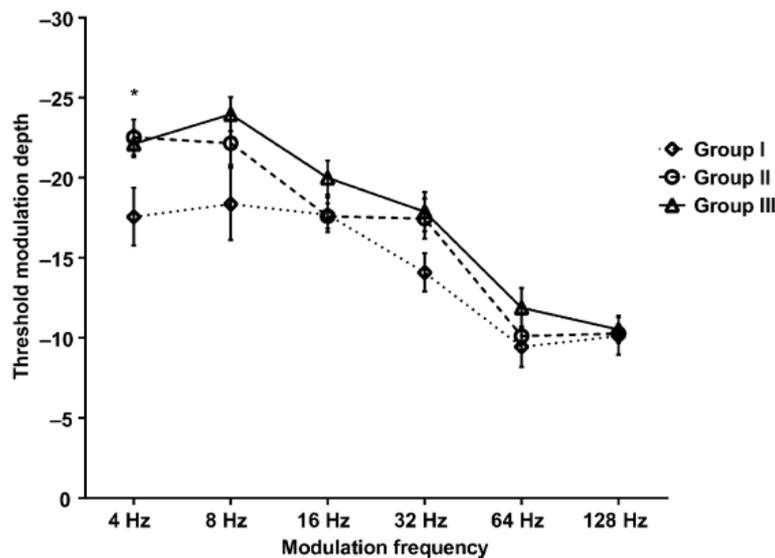


Figure 26. Fonctions de transfert de modulation temporelle (TMTF) de trois groupes d'auditeurs. Le groupe I est constitué de 10 enfants âgés de 8 à 10 ans ; le groupe II est constitué de 10 enfants âgés de 12 à 17 ans et le groupe III est constitué de 10 adultes âgés de 18 à 32 ans. Les cercles, triangles et losanges indiquent les seuils de détection moyens pour chaque cadence de modulation. L'astérisque indique la cadence de modulation pour laquelle les groupes diffèrent. Extrait de Peter et al. (2014).

En effet, comme le montre la figure 26, les seuils de détection d'AM à 4 Hz sont significativement plus élevés (moins bons) chez les enfants les plus jeunes (âgés entre 8 et 11 ans) que chez les adolescents (12-18 ans) et les adultes. Ceci suggère que leur capacité à détecter une AM *lente* n'est pas arrivée à maturation avant l'âge de 12-18 ans. Peter et al. (2014) n'observent pas d'effet développemental pour les autres cadences de modulation testées (8, 16, 32, 64 et 128 Hz).

Contrairement à ces derniers résultats, d'autres études indiquent que les seuils de détection d'AM mesurés pour une cadence *rapide* s'améliorent plus lentement avec l'âge que pour une AM lente. L'étude de Moore et al. (2011) dont les résultats sont représentés en figure 27 démontre ainsi que les seuils de détection d'une AM rapide de 125 Hz s'améliorent plus lentement avec l'âge qu'une AM lente de 8 Hz.

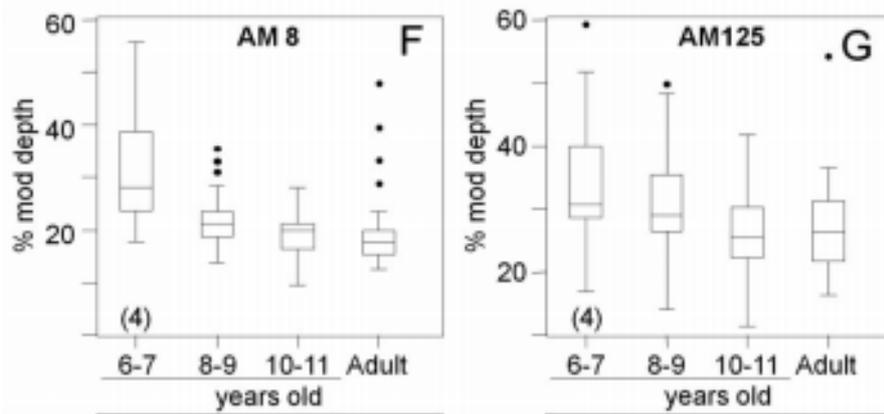


Figure 27. Boîtes à moustaches comparant la médiane et la variabilité des seuils de détection d'AM entre différents groupes d'âge. Les points extrêmes correspondent aux observations extrêmes ou aberrantes. Extrait de Moore et al. (2011).

Ces résultats sont conformes à ceux obtenus par Buss et al. (2019) dans une étude mesurant les seuils de détection d'AM à trois cadences de modulation différentes (16, 64 et 256 Hz) chez des enfants âgés entre 5 et 11 ans et des adultes. Les résultats de cette étude indiquent que la détection d'AM s'améliore avec l'âge aux cadences rapides (64 et 256 Hz) mais pas à la cadence lente (16 Hz).

Contrairement aux conclusions de l'étude de Hall et Grose (1994) et Walker et al. (2019), ces dernières études suggèrent que les trajectoires développementales diffèrent selon la cadence d'AM (ou tout au moins entre cadences lentes et rapides). Ces dissociations sont compatibles avec l'hypothèse d'une sélectivité dans le traitement d'AM (Houtgast, 1989 ; Bacon & Grantham, 1989 ; Dau et al., 1997). Certains résultats peuvent également être interprétés comme révélant une amélioration de l'acuité temporelle pendant les premières années de vie. Enfin, l'amélioration globale de la sensibilité à l'AM observée par certains auteurs comme Hall et Grose (1994) ou Cabrera et al. (2019) entre environ 5 ans et l'âge adulte peut être interprétée comme reflétant une amélioration de l'« **efficacité de traitement (processing efficiency)** » de l'AM, à savoir la *capacité à utiliser optimalement les indices sensoriel d'AM (ENV_N) dans une tâche de détection*. Cette optimalité peut être réduite par des facteurs stochastiques (le bruit interne, qui sera défini plus loin dans cette thèse), et des facteurs déterministes liés aux limites dans les capacités de mémorisation, de sélection attentionnelle, ou encore de prise de décision (Cabrera et al., 2019).

I.2.c. Traitement de la FM par le système auditif

I.2.c.i. Traitement par le sujet normo-entendant adulte jeune

Si les mécanismes de traitement de l'AM font aujourd'hui l'objet d'un consensus au sein de la communauté des sciences de l'audition, cela n'est pas encore le cas pour le traitement auditif de la FM, pour lequel deux mécanismes sensoriels sont encore opposés et débattus:

i) Zwicker (1952) et Maiwald (1967a, b) suggèrent qu'**un seul et même mécanisme** de nature spatiale (tonotopique) est impliqué dans le traitement de l'AM et de la FM. Ce mécanisme est basé sur le concept de pattern d'excitation – à savoir le niveau d'excitation neuronale moyen (*i.e.*, la sortie intégrée dans le temps) des filtres cochléaires en réponse à un son donné (Moore & Glasberg, 1983 ; 1987). Zwicker et Maiwald supposent que les changements dynamiques des sons en amplitude ou en fréquence sont détectés à partir des variations du niveau d'excitation neuronale produit au sein du système auditif périphérique sur la pente basse fréquence (pour la FM) ou haute fréquence (pour l'AM) du pattern d'excitation évoqué par le signal modulé. En d'autres termes, l'atténuation différentielle des filtres cochléaires convertit les variations de FM en variation d'AM (et donc d'excitation sensorielle) à la sortie des filtres cochléaires. Les filtres accordés en dessous de cette fréquence porteuse sont optimaux pour la détection de FM car les variations d'excitation neuronale – encodées par des changements de taux moyen de décharge dans les neurones auditifs - sont maximales à cet endroit. Notons que le modèle du banc de filtres de modulation développé plus récemment par Dau et al. (1997) inclut implicitement un tel mécanisme spatial (*i.e.*, tonotopique), mais en lui conférant un aspect dynamique, prenant en compte la nature précise des indices temporels. Ce mécanisme est en quelque sorte « spatio-temporel ». En effet, dans le modèle du banc de filtres de modulation, tout son est initialement décomposé par un banc de filtres cochléaires, puis démodulé par un banc de filtres de modulation et un filtre passe-bas. Toute variation d'enveloppe temporelle neurale (ENV_N) engendrée par le signal de FM en sortie des filtres cochléaires puis des filtres de modulation sera alors susceptible d'être détectée par le mécanisme décisionnel, qu'il s'agisse d'un appariement de gabarit utilisant l'information de phase d'enveloppe ou d'un simple changement de puissance de modulation.

ii) *A contrario*, d'autres études (*e.g.*, Feth, 1972 ; Coninx, 1978 ; Demany & Semal, 1986 ; Moore & Sek, 1996) supposent que l'AM et la FM d'un son donné sont détectées par **deux mécanismes indépendants** : un mécanisme basé sur un code purement *temporel* utilisant les informations neurales représentées par verrouillage en phase des décharges des neurones auditifs périphériques ou du tronc cérébral sur la structure temporelle fine des sons (TFS_N , ces informations étant supposées être encodées uniquement pour des fréquences porteuses basses (<1-4 kHz) et des cadences de modulation lentes (<5 Hz)) et un mécanisme *spatial* (tonotopique) basé sur les indices de pattern d'excitation dans les autres cas. Les données électrophysiologiques unitaires obtenues par Paraouty et al. (2016) dans le noyau cochléaire antéroventral du cochon d'inde suggèrent que le mécanisme temporel pourrait être opérationnel pour des cadences de modulation plus élevées que celles postulées par ces auteurs. En tout cas, l'incapacité à utiliser des informations de TFS_N évoquées par une FM rapide (> 5 Hz) et portée par un son pur de fréquence inférieure à 1-4 kHz – incapacité mise sur le compte d'une forme d'« inertie » (*sluggishness*) dans les traitements – résulterait donc de limitations centrales, intervenant au-delà du tronc cérébral bas.

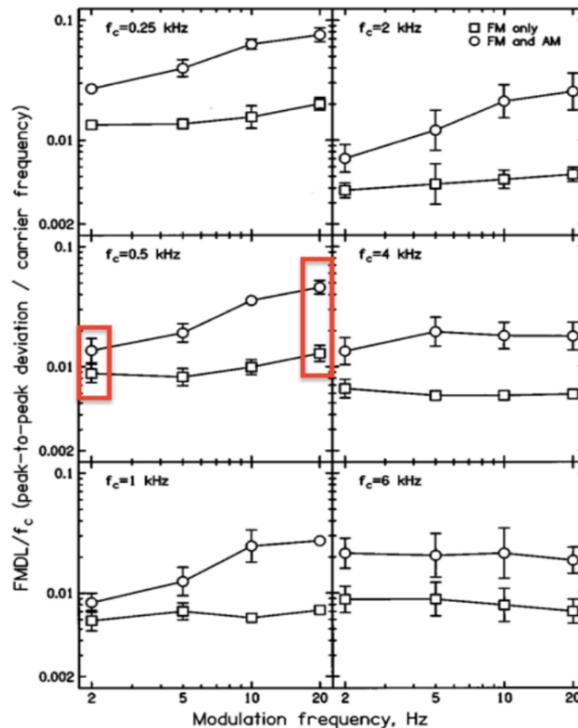


Figure 28. Seuils de détection de FM avec (cercle) et sans (carré) AM interférente représentés en fonction de la fréquence de modulation. L'axe des ordonnées indique les seuils de détection de FM exprimés comme la différence crête à crête divisée par la fréquence centrale. L'axe des abscisses indique la fréquence de modulation exprimée en Hertz. Chaque graphe illustre les résultats pour une fréquence porteuse spécifique (valeur f_c indiquée en haut de chaque graphe). Les deux encadrés rouges indiquent deux conditions expérimentales testées dans notre recherche doctorale. Dans la première condition testée (encadré gauche), l'ajout d'une AM de phase aléatoire (employée afin de perturber le codage spatial dans la tâche de détection de FM) n'affecte pas significativement le seuil de détection de FM à une cadence de modulation lente (2 Hz) et une fréquence porteuse basse (500 Hz). Par contre, le seuil de détection de FM se dégrade à mesure que la cadence de modulation augmente (les deux courbes se séparent lorsque la cadence de modulation augmente, la séparation étant maximale pour une cadence de 20 Hz (encadré de droite)). Ceci est conforme avec l'hypothèse d'un double mécanisme : lorsque la fréquence porteuse est basse (<1-4 kHz), la détection de FM dépendrait principalement d'un mécanisme de codage temporel pour des fréquences de modulation lentes (<5 Hz) et d'un mécanisme spatial pour des fréquences de modulation rapides (>5 Hz). Extrait de Moore et Sek (1996).

Dans une tâche à deux alternatives en choix forcé, Moore et Sek (1996) mesurent les seuils de détection d'une FM sinusoïdale en présence d'une AM sinusoïdale interférente (de même cadence que la FM, mais de phase aléatoire d'essai en essai) ou en l'absence d'AM interférente au sein des deux intervalles. L'AM interférente est introduite afin de limiter la possibilité d'utiliser des indices spatiaux de pattern d'excitation. Différentes fréquences porteuses (0.25-6 kHz) et fréquences de modulation (2, 5, 10 et 20 Hz) sont utilisées afin d'évaluer le rôle du mécanisme de verrouillage en phase sachant que la limite de ce mécanisme est estimée autour de 1-4 kHz chez l'humain (Verschooten et al., 2019). La figure 28 ci-dessus présente les seuils de détection de FM en fonction de la cadence de modulation pour les deux conditions expérimentales testées (FM seule (carrés) ou FM en présence d'AM (cercles)). D'une manière générale, les seuils de détection de FM seule (en l'absence d'AM)

changent très peu en fonction de la cadence de modulation. Ces seuils augmentent plus nettement avec la cadence de modulation lorsque la porteuse est égale à 2 kHz et diminuent lorsque la porteuse est égale à 6 kHz. Ces observations sont en accord avec les résultats de Sek et Moore (1995) et corroborent l'hypothèse d'un mécanisme de verrouillage en phase sur la structure temporelle fine des sons pour des porteuses basses et des cadences lentes. L'ajout de l'AM interférente dégrade généralement les seuils de détection de FM. Dans les conditions où la porteuse est basse (comprise entre 0.25 kHz et 2 kHz), la dégradation est minimale lorsque la cadence de modulation de FM est lente (*e.g.*, 2 Hz) et augmente avec la cadence de modulation pour culminer à 20 Hz, la plus rapide cadence de FM testée. Dans les conditions où la porteuse est élevée (4-6 kHz), la dégradation produite par l'AM est relativement constante quelle que soit la cadence de FM. Ces résultats s'accordent avec l'hypothèse d'un double mécanisme temporel et spatial (code tonotopique) pour la détection de FM. Le mécanisme temporel dominerait dans une seule condition, à savoir lorsque la fréquence porteuse est basse (car le mécanisme de verrouillage en phase n'est opérant qu'en dessous de 1-4 kHz) et lorsque la cadence de modulation est lente (car le mécanisme de lecture des indices de verrouillage en phase est supposé souffrir d'une certaine « inertie » (*sluggishness*)). Le mécanisme spatial (tonotopique) - quant à lui - dominerait dans deux conditions : i) lorsque la fréquence porteuse est élevée (car le mécanisme de verrouillage en phase n'est plus opérant) et ii) lorsque la fréquence porteuse est basse et la cadence de modulation est rapide (rendant le mécanisme de lecture des indices de verrouillage en phase inopérant).

Paraouty et al. (2016) et King et al. (2019) étendent ces travaux princeps et précisent les effets d'interférence entre AM et FM en combinant mesures psychophysiques et modélisation informatique basée sur différentes versions d'un modèle auditif utilisant un banc de filtre de modulation. Dans l'étude menée par Paraouty et al. (2016), les seuils de détection sont mesurés pour trois catégories d'auditeurs (sujets normo-entendants jeunes, sujets normo-entendants âgés de 46 à 67 ans, sujets malentendants âgés de 46 à 67 ans). King et al. (2019) limitent leur investigation à des sujets normo-entendants jeunes. Les effets perceptifs d'interférence rapportés par Moore et Sek (1996) entre AM et FM sont observés dans chaque étude: la détection de FM est effectivement altérée en présence d'AM ; réciproquement, la détection d'AM est altérée par la présence de FM. Toutefois, l'ensemble de ces résultats n'est pas correctement simulé à l'aide de modèles computationnels utilisant uniquement des indices d'enveloppe temporelle neurale (ENV_N), à savoir des indices dynamiques de pattern d'excitation. Plus précisément, les modèles prédisent des effets d'interférence bien plus élevés que ceux observés sur les sujets entendants et malentendants. Ceci suggère qu'un autre mécanisme (peut-être temporel) intervient dans le codage de FM, réduisant ainsi les effets d'interférence.

D'une manière générale,

- Lorsque la porteuse est inférieure à 4 kHz et lorsque la cadence de modulation est très basse (<5-10 Hz), la FM serait essentiellement détectée grâce à un mécanisme temporel utilisant des indices neuronaux de verrouillage en phase (TFS_N).
- Lorsque la porteuse est élevée (>1-4 kHz) ou encore lorsque la cadence de modulation est rapide (>5-10 Hz) et ce quelle que soit la fréquence de la porteuse, la FM serait détectée grâce à un autre mécanisme de nature spatio-temporelle utilisant les indices neuronaux d'enveloppe temporelle (ENV_N) obtenus par reconversion FM-AM en sortie du filtre cochléaire.
- Cette dépendance spécifique des mécanismes temporels et spatio-temporels vis-à-vis de la fréquence porteuse et de la cadence de modulation expliquent les effets de masquage produits par une AM interférente sur la détection d'une FM cible et vice et versa.

Toutefois, d'autres résultats ne corroborent pas l'hypothèse d'un double code pour le traitement de la FM. Notamment, certaines expérimentations animales indiquent que les patterns de décharge unitaire de neurones du cortex auditif primaire sont similaires en réponse à une AM et à une FM (e.g., Eggermont, 1994 ; Liang et al., 2002). Ceci suggère qu'il n'existe pas d'unité neuronale corticale spécialisée dans le traitement d'un type spécifique de modulation temporelle. Par ailleurs, des expériences de neuroimagerie utilisant la technique d'imagerie par résonance magnétique (IRM) fonctionnelle indiquent que les mêmes régions du cortex auditif primaire répondent à ces deux types de modulation chez l'homme (Hart et al., 2003).

Une étude corrélationnelle menée sur une cohorte importante de sujets remet également en cause l'hypothèse des deux codes pour la perception de la FM. Whiteford et Oxenham (2015) utilisent ici une approche « différentielle » basée sur l'étude des différences individuelles dans un groupe de 100 sujets normo-entendants jeunes afin de tester cette hypothèse. Les seuils de détection de FM individuels mesurés pour une cadence lente (1 Hz) et une cadence rapide (20 Hz) avec une porteuse de 500 Hz sont comparés à i) des mesures de seuils de détection d'AM réalisées dans les mêmes conditions et d'autres mesures binaurales ou monaurales susceptibles de refléter : ii) la précision du codage temporel basé sur le verrouillage en phase (seuils de discrimination de différence interaurale de temps (ITD) ou iii) la précision du codage spatial par la mesure de sélectivité fréquentielle (réalisée grâce à une mesure de masquage proactif). Les résultats de cette étude ne s'accordent pas avec l'idée d'un double-code pour la détection de FM. En effet, toutes les mesures de seuils de détection d'AM et de FM s'avèrent être corrélées les unes avec les autres. Par ailleurs, les corrélations observées entre les différentes mesures ne sont pas plus importantes pour celles censées refléter les capacités de verrouillage de phase (par ex. la détection de FM lente et la sensibilité aux ITD), ou pour celles censées refléter la précision du codage spatial (la détection FM rapide et les patterns de

masquage). Whiteford et al. (2017) étendent cette première investigation à une population de 85 sujets dont l'âge est compris entre 20 et 79 ans. A nouveau, les seuils de détection de FM et d'AM sont corrélés les uns aux autres. L'âge est plus fortement corrélé à la sensibilité à la FM qu'à la sensibilité à l'AM, mais l'effet de l'âge s'avère similaire pour la détection de FM lente et rapide. Nous reviendrons plus loin sur les effets du vieillissement.

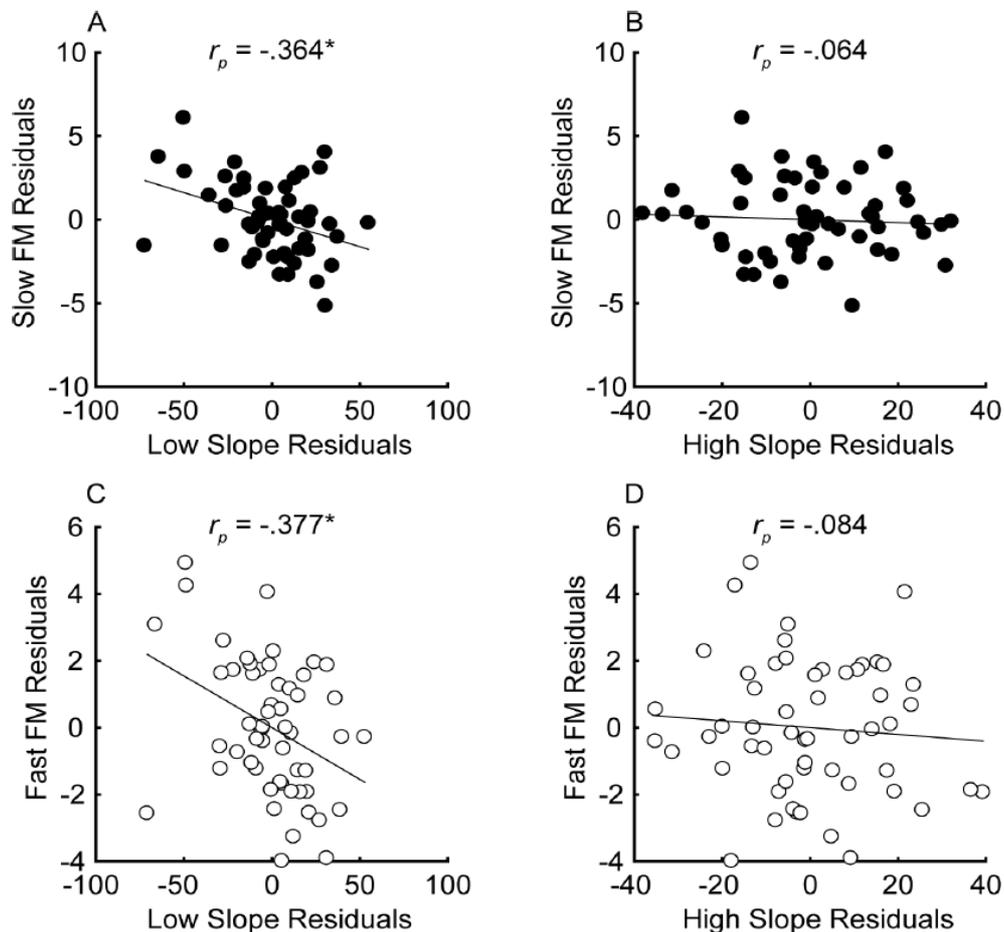


Figure 29. Corrélations partielles entre la pente du pattern de masquage (en abscisse) et le seuil de détection de FM (ordonnée) pour une FM lente (A et B) et rapide (C et D), après prise en compte de la part de variance due à l'audibilité et à l'âge. Extrait de Whiteford et al. (2020).

Enfin, dans une troisième étude menée sur une cohorte de 56 sujets dont 47 présentent une perte auditive neurosensorielle, Whiteford et al. (2020) montrent que les seuils de détection de FM lente et rapide sont corrélés de manière forte (et égale) aux pentes « basse fréquence » des patterns de masquage (cf. figure 20 ci-dessus) révélant ainsi une relation entre sélectivité fréquentielle et sensibilité à la FM lente et rapide. Les analyses de régression linéaire multiple montrent que l'effet combiné de l'audibilité, de l'âge, de la sensibilité à l'AM et des pentes des patterns de masquage rend compte de plus de 50 % de la variance totale des seuils de détection de FM lente et rapide. Ces résultats suggèrent donc que *l'efficacité du codage de la FM lente dépend de la sélectivité du filtrage cochléaire, comme prédit par un mécanisme spatio-temporel unique utilisant les indices d'ENV_N issus de la conversion de la FM en AM à la sortie des filtres cochléaires.*

Dans une expérience complémentaire, Whiteford et al. (2020) mesurent la capacité de 20 sujets normo-entendants jeunes à percevoir des différences de cohérence (un décalage de phase) entre deux AM sinusoïdales de même cadence appliquées respectivement à deux sons purs de fréquences audio distinctes et faisant office de porteuse. Cette expérience est réalisée afin de simuler le codage purement spatial de la FM sur des sujets normo-entendants. La différence de phase à percevoir entre ces deux AM est constante et égale à 180° . Les fréquences porteuses sont séparées de $2/3$ ou $4/3$ d'octaves. Une bande de bruit est présentée entre les deux porteuses afin d'éliminer tout indice acoustique de fréquence instantanée (TFS_P) produit par ces deux sinusoïdes faisant office de porteuse lorsqu'elles sont présentées simultanément. Les auteurs observent que les performances de discrimination des participants sont moins bonnes pour des fréquences porteuses élevées et pour des cadences d'AM rapides, deux aspects de la détection de FM (limites du verrouillage en phase et « inertie (sluggishness) ») généralement attribués aux limites dans les capacités de codage des indices de TFS_N évoqués par une FM. Ces résultats corroborent donc l'hypothèse selon laquelle une FM lente appliquée à une porteuse basse fréquence serait détectée à partir des indices d' ENV_N résultant de la conversion de la FM en AM à la sortie des filtres cochléaires plutôt que par des indices de TFS_N .

Ces derniers résultats s'opposent à ceux obtenus par Moore et Sek (2019) et illustrés en figure 30 ci-dessous. Cette étude démontre que : i) la capacité de discrimination entre une AM et une FM de même cadence appliquées à une porteuse basse fréquence de 1 kHz et présentées au même niveau de détectabilité (d' ; e.g., $d'=1$) est meilleure pour une cadence lente (2 Hz) que pour une cadence rapide (10 Hz) ; ii) cet effet ne peut s'expliquer en terme de perte de sensibilité à une différence de phase d'AM évoquée par la FM sur les pentes basse et haute fréquence du pattern d'excitation lorsque l'AM est rapide car la capacité à discriminer un changement de phase d'AM de cette nature est au niveau du hasard lorsque l'AM est présentée aux profondeurs de modulation permettant aux sujets de discriminer aisément AM et FM. Comme pour l'expérience de Whiteford et al. (2020), les auteurs mesurent la performance de discrimination d'un décalage de phase de 180° entre deux AM quasi-trapézoïdales de même cadence appliquées respectivement à deux porteuses sinusoïdales de 762 et 1296 Hz, avec une bande de bruit entre les deux porteuses afin de masquer tout indice intra-canal. Cette performance est mesurée pour des cadences d'AM comprises entre 2 et 20 Hz. Les performances de discrimination de phase d'AM sont relativement constantes pour les cadences d'AM de 2 à 10 Hz, mais se dégradent à 20 Hz. La performance de discrimination est proche du niveau de réponse au hasard ($d' < 0.5$) pour les profondeurs d'AM proches (et juste supérieures) au seuil de détection d'AM. Pour ces valeurs de profondeurs, la discrimination entre une AM et une FM est bien supérieure au hasard. En d'autres termes, les participants normo-entendants jeunes peuvent distinguer l'AM de la FM à une cadence de 2 Hz pour des profondeurs de modulation où la discrimination de phase d'AM est proche du hasard. Ce résultat suggère à nouveau que l'hypothèse d'un mécanisme spatial basé sur la conversion de la FM en AM ne peut expliquer la discrimination AM-FM à une cadence de 2 Hz.

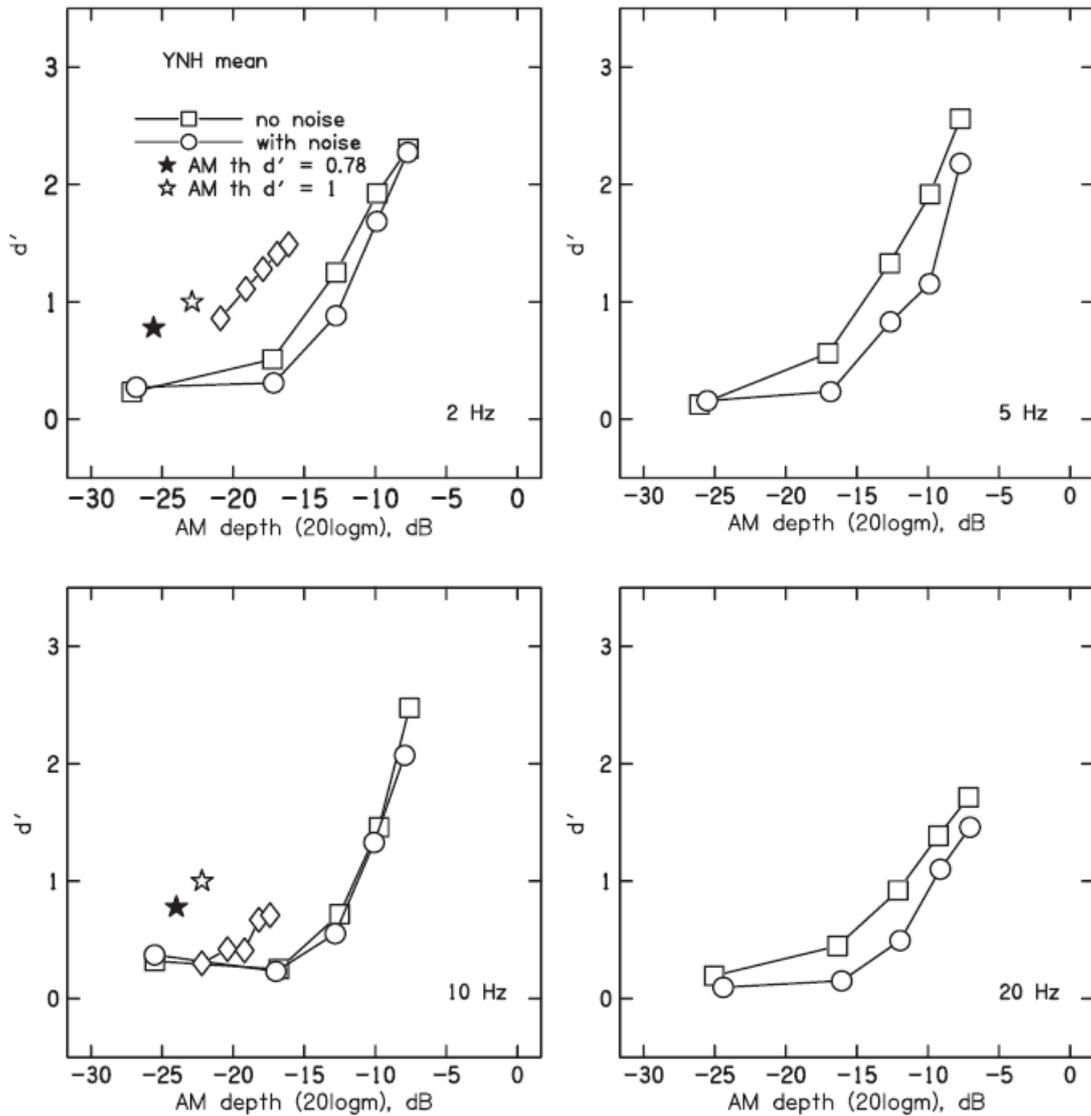


Figure 30. Fonctions psychométriques pour la discrimination d'un décalage de phase d'AM entre deux porteuses. Chaque figure correspond aux fonctions psychométriques mesurées à une cadence d'AM différente (de 2 à 20 Hz). Dans chaque figure, les courbes (cercles et carrés ouverts) représentent les fonctions psychométriques moyennes obtenues sur l'ensemble des sujets normo-entendants jeunes de l'étude. Ces fonctions présentent la sensibilité (d' , en ordonnée) en fonction de la profondeur d'AM (20logm, en abscisse). Les cercles correspondent aux données obtenues avec un bruit masquant inséré entre les deux porteuses ; les carrés correspondent aux données obtenues sans bruit masquant. Les étoiles fermées et ouvertes indiquent les valeurs de profondeur (20logm) correspondant au seuil de détection de l'AM (étoile fermée : $d'=0.78$; étoile ouverte : $d'=1$). Les données de discrimination entre AM et FM à 2 et 10 Hz sont reproduites dans deux figures (courbes avec losanges ouverts). Extrait de Moore et Sek (2019).

I.2.c.ii. Etudes développementales

Quelques études comportementales complètent celles décrites plus haut en évaluant la maturation de la sensibilité à la FM chez l'enfant. Dans une étude menée sur une cohorte de 98 enfants âgés entre 6 et 10 ans et 18 adultes, Dawes et Bishop (2008) mesurent le développement de la capacité à détecter une FM sinusoïdale à l'aide d'une procédure adaptative en choix forcé. La porteuse est de 500 Hz pour la FM de cadence égale à 2 ou 40 Hz et de 1000 Hz pour la FM de cadence égale à 240 Hz. Les résultats de cette étude présentés en figure 31 ci-dessous indiquent que les trajectoires développementales diffèrent entre la FM la plus rapide et les FM les plus lentes. En effet, le seuil de détection d'une FM de 240 Hz atteint le niveau adulte dès 7 ans, alors qu'il n'atteint ce niveau qu'à partir de 9 ans à 2 et 40 Hz. Les enfants sont donc plus précoces dans le traitement des indices de FM rapides. Toutefois, aucune différence n'apparaît dans les trajectoires développementales observées à 2 et 40 Hz. Ces données confirment celles d'une autre étude menée par Sutcliffe et Bishop (2005).

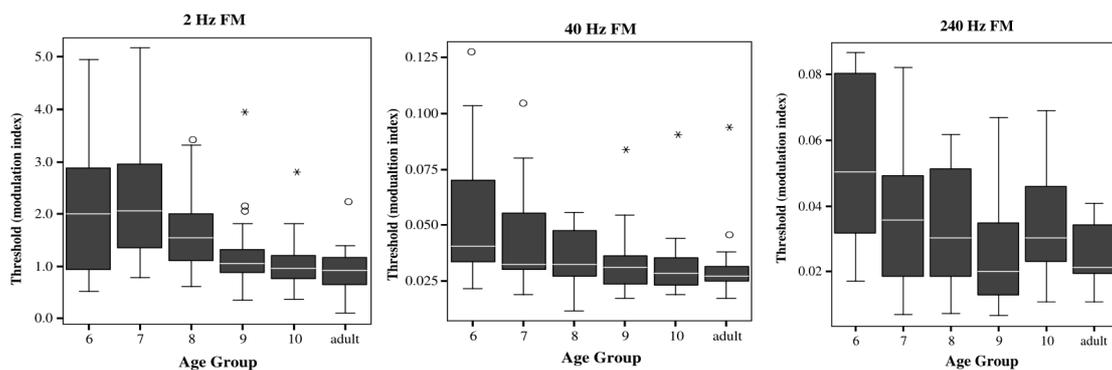


Figure 31. Seuils de détection de FM représentés par des boîtes à moustaches (en ordonnée) en fonction de l'âge (en abscisse). La médiane est indiquée par une ligne blanche horizontale. Les bords des cases correspondent aux quartiles supérieur et inférieur. Extrait de Dawes et Bishop (2008).

Une autre étude menée par Banai et al. (2011) évalue deux aspects de la performance de détection d'AM et de FM dans une tâche adaptative chez des enfants et des adultes: la « performance moyenne » (*i.e.*, le seuil) et une mesure dénommée « performance consistency » (la variabilité intra-individuelle estimée par l'écart-type calculé sur les dernières inversions de l'escalier psychophysique). Notons que cette deuxième mesure est liée à la variance du *bruit interne* pour la tâche en question. Nous reviendrons plus loin sur cette notion. Les seuils de détection de FM sont évalués pour trois cadences de modulation (2, 20 et 240 Hz) et une porteuse de 1000 Hz à l'aide d'une procédure adaptative à 3 alternatives en choix forcé. Les trajectoires développementales obtenues sont comparées pour la détection d'AM et de FM chez 12 jeunes enfants (8-10 ans) et 11 pré-adolescents (11-12 ans). Les résultats présentés en figure 32 indiquent que la détection de FM atteint le niveau adulte estimé par ces deux mesures dès l'âge de 8 ans, tandis que la détection d'AM n'arrive à maturité qu'à partir de 11-12 ans et continue à s'améliorer sans pour autant changer en

termes de variabilité (écart-type) et donc de niveau de bruit interne. De plus, les données individuelles des enfants indiquent que le degré d’immaturité des seuils de détection moyens et de « performance consistency » ne sont pas corrélés entre les tâches de détection d’AM et de FM. L’ensemble de ces résultats suggère que *les trajectoires développementales pour la détection d’AM et de FM sont partiellement distinctes*. Par conséquent, **différents mécanismes semblent responsables de la maturation de la capacité à détecter AM et FM**.

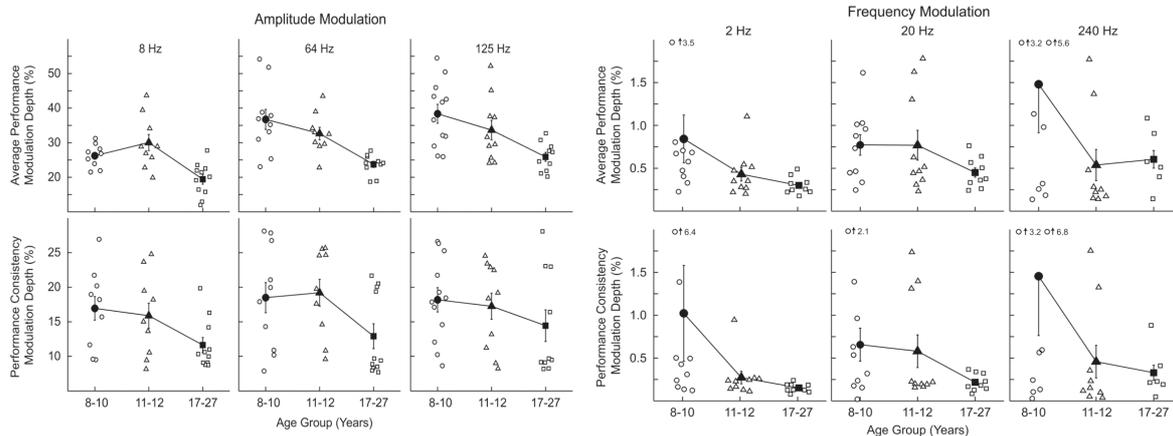


Figure 32. Performance de détection d’AM (figures de gauche) et de FM (figures de droite) pour différents groupes d’âge. La rangée du haut représente la performance moyenne (le seuil). La ligne du bas représente la variabilité individuelle (ou « performance consistency »). Chaque colonne représente une cadence de modulation (première colonne : 2 Hz, deuxième colonne: 20 Hz, troisième colonne: 240 Hz). Le groupes d’âge (8-10 ans ; 11-12 ans ; 17-27 ans) est représenté en abscisse. La performance moyenne exprimée en termes de profondeur de modulation (rangée du haut) ou par la cohérence de la performance (rangée du bas) est représentée en ordonnée. Extrait de Banai et al. (2011).

I.2.d. Des modèles de traitement de modulation

Pris ensemble, ces résultats psychophysiques obtenus chez l’adulte normo-entendant jeune ou âgé et chez l’enfant pointent vers deux mécanismes distincts, l’un de nature purement temporelle, l’autre de nature spatio-temporelle, tous deux étant impliqués dans le codage et la perception des indices d’AM et de FM. Ces mécanismes représenteraient AM et FM en termes d’indices neuronaux de TFS_N et d’ ENV_N au sein du système auditif périphérique et du tronc cérébral bas. Cette conception implique deux architectures ou voies de traitement distinctes.

1-Celle impliquée dans la représentation de l’AM et de la FM rapide ainsi que de la FM lente portée par des composantes fréquentielles haute fréquence (>1-4 kHz) fait aujourd’hui l’objet d’un consensus. Cette architecture de traitement met l’accent sur les traitements spatio-temporels et est dénommée « **modèle à banc de filtres de modulation** » (Dau et al., 1997) car elle encode ces informations d’AM et de FM sous la forme d’indices d’ ENV_N . Dans cette architecture, l’information de fréquence instantanée (TFS) est éliminée au moment de l’étape

de démodulation. Cette architecture peut se décliner de deux manières selon qu'elle conserve ou non l'information de phase d'enveloppe.

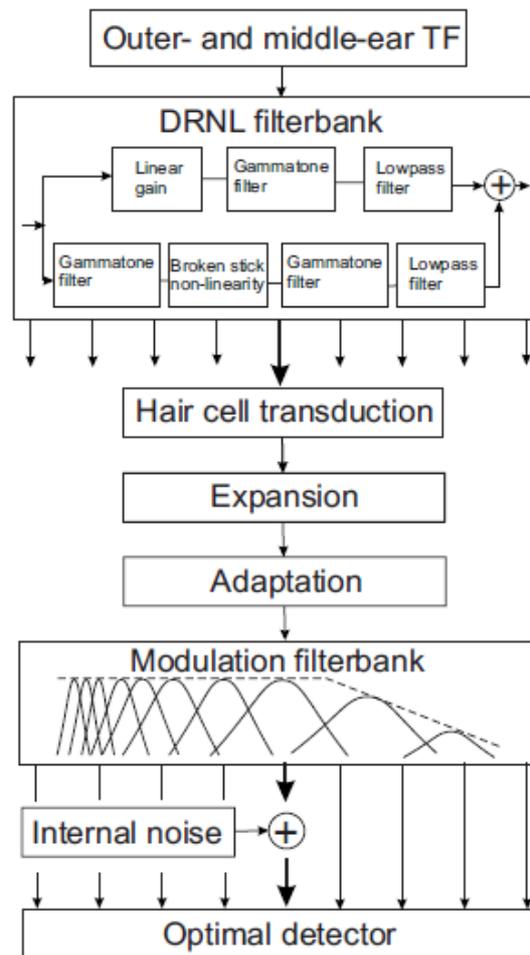


Figure 33. Modèle à banc de filtres de modulation de type « PEMO ». Le traitement du signal acoustique est réalisé en plusieurs étapes : filtrage par l'oreille externe et moyenne, suivi d'une décomposition dans le domaine audio-fréquentiel via un banc de filtres non linéaire. La sortie de ces filtres est compressée en amplitude, redressée, puis expansée et transformée par un mécanisme d'adaptation rehaussant les éléments transitoires. L'enveloppe temporelle en sortie de chaque canal fréquentiel audio est ensuite extraite via un filtrage passe-bas et une batterie de filtres de modulation (un banc de filtre par canal audio-fréquentiel). Les enveloppes temporelles ainsi extraites sont dégradées par un bruit interne additif et soumises à un mécanisme de décision de nature corrélacionnelle (« optimal detector ») opérant par appariement de gabarit (le gabarit est une représentation du signal à détecter stockée en mémoire à long terme). Extrait de Jepsen et al. (2008).

La version de ce modèle conservant l'information de phase d'enveloppe en sortie des filtres de modulation (et ce pour les filtres accordés sur des cadences inférieures à 5-10 Hz) utilise un mécanisme de décision basé sur un appariement de gabarit. Elle prend parfois le nom de PEMO (pour « perception model » ; Dau et al., 1997 ; Jepsen et al., 2008). Le schéma en figure 33 illustre les grandes étapes du modèle PEMO. La version de ce modèle éliminant l'information de phase d'enveloppe n'utilise que les indices de puissance moyenne d'enveloppe temporelle calculés en sortie des filtres de modulation comme mécanisme de

décision. Elle perd donc une partie des informations dynamiques du signal. Elle prend le nom d'EPSM (pour « envelope power spectrum model » ; Ewert & Dau, 2000).

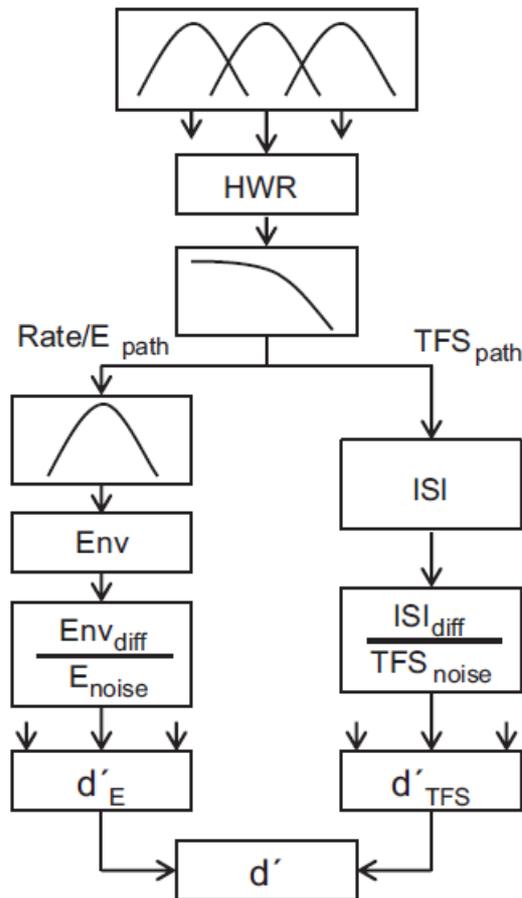


Figure 34. Modèle à double voie. Le traitement du signal acoustique est réalisé en plusieurs étapes. Les signaux (AM ou FM, ou encore signaux mélangeant AM et FM) sont présentés à un banc de filtres cochléaires. La sortie des filtres cochléaires est redressée et filtrée en passe-bas en-dessous de 1 kHz afin de respecter les limites du verrouillage en phase des fibres du nerf auditif. La voie de gauche (rate/ENV) traite les signaux résultants à la manière du modèle à banc de filtres de modulation décrit en figure 31. La voie de droite (TFS) extrait les indices de fréquence instantanée des mêmes signaux. Chaque voie possède sa propre source de bruit interne (E_{noise} ; TFS_{noise}). Un indice de sensibilité est calculé à la sortie de chaque voie. Ces indices sont combinés afin de prendre la décision finale par comparaison à un seuil (un critère). Extrait d'Ewert et al. (2020).

2-L'architecture de traitement impliquée dans la représentation de la FM lente portée par des composantes fréquentielles basse fréquence (<1-4 kHz) n'a fait l'objet que d'une seule forme de modélisation computationnelle (Ewert et al., 2020). Le mécanisme représentant AM et FM sous la forme d'indices de TFS_N extrait les variations de fréquence instantanée à la sortie des filtres cochléaires par estimation des intervalles inter-spikes (ISI, intervalles de temps entre passages à zéro successifs). Cette architecture de traitement est dénommée « modèle à double voie » (Ewert et al., 2020) car elle encode ces informations d'AM et de FM sous la forme d'indices d' ENV_N et de TFS_N grâce à des mécanismes (« voies ») distincts, puis elle combine ces indices afin de prendre une décision finale. Le schéma en figure 34 illustre les

grandes étapes du modèle « à double voie ». La voie extrayant AM et FM sous forme d'indices de TFS_N à partir de l'estimation des ISI correspond à la voie de droite (« TFS path »). La voie extrayant AM et FM sous forme d'indices d' ENV_N (« ENV path ») correspond à une implémentation du modèle EPSM décrit plus haut. Notons que chaque voie possède sa propre source de bruit interne additif. Nous reviendrons plus loin sur ce point.

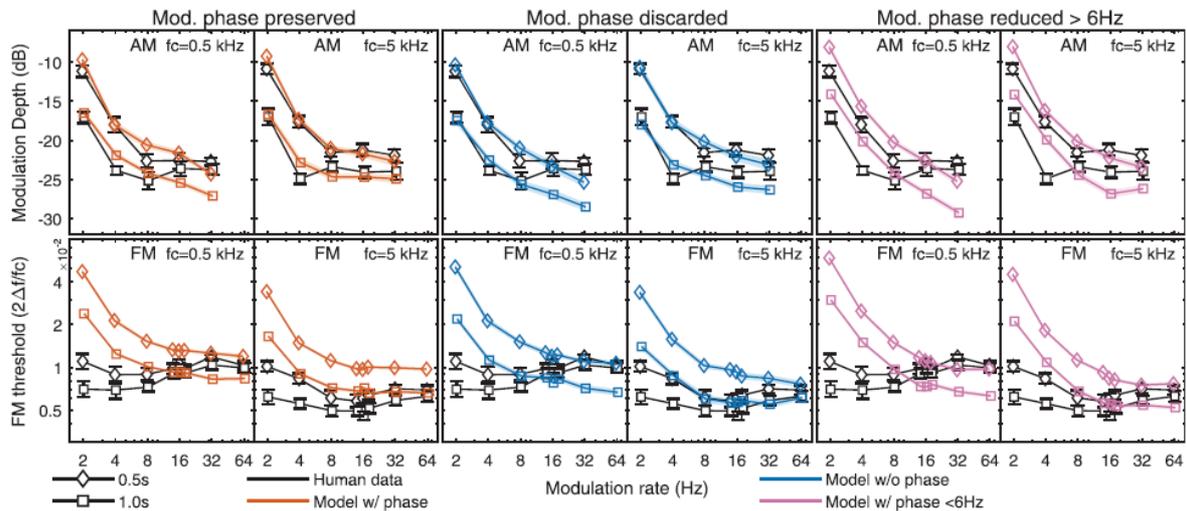


Figure 35. Seuils de détection d'AM (en haut) et de FM (en bas) en fonction de la cadence de modulation, pour deux fréquences porteuses ($f_c=0.5$ et 5 kHz) et deux durées de présentation des stimuli (0.5 et 1 sec). Les traits noirs connectant des symboles ouverts noirs correspondent aux données empiriques obtenues sur des sujets normo-entendants jeunes. Les courbes de couleur correspondent aux données simulées par différentes implémentations du modèle à banc de filtres de modulation, PEMO. Les simulations présentées dans les figures de gauche sont obtenues lorsque l'information de phase d'enveloppe est préservée quelle que soit la cadence de modulation. Celles présentées dans les figures centrales sont obtenues lorsque l'information de phase d'enveloppe est entièrement éliminée. Celles présentées dans les figures de droite sont obtenues lorsque l'information de phase n'est préservée que pour les filtres de modulation accordés en dessous de 6 Hz. Extrait de King et al. (2019).

La revue de littérature réalisée dans le cadre de ce deuxième chapitre de thèse indique toutefois que plusieurs résultats expérimentaux obtenus chez l'adulte et chez l'enfant ne sont pas entièrement compatibles avec l'hypothèse d'un mécanisme purement temporel et du modèle à double voie. En effet, certaines données électrophysiologiques ou de neuroimagerie, ainsi que des données psychophysiques issues entre autres d'études corrélacionnelles militent en faveur d'un seul mécanisme de nature spatio-temporelle, tel qu'implémenté par le modèle à banc de filtres de modulation (PEMO ou EPSM). **Des études supplémentaires semblent nécessaires afin de clarifier ce débat.** Notons toutefois que si plusieurs implémentations du modèle à banc de filtres de modulation (PEMO, EPSM) permettent de reproduire les capacités de détection d'une AM quelle que soit la cadence et d'une FM rapide, ces mêmes modèles sont en échec lorsqu'il s'agit de reproduire les capacités de détection d'une FM, et ce pour une porteuse basse fréquence (500 Hz) ou haute fréquence (5000 Hz) (Paraouty et al., 2016 ; Wallaert et al., 2018 ; King et al. 2019). Or, de tels modèles

utilisent toute l'information d'enveloppe (ENV_N) disponible à la sortie des filtres cochléaires et des canaux de modulation. Les résultats de ces études de simulation ne sont pas compatibles avec l'hypothèse d'un mécanisme spatio-temporel pour la détection de FM lente. Les données de détection d'AM et de FM obtenues sur des sujets normo-entendants jeunes et les données prédites par plusieurs implémentations du modèle PEMO (King et al., 2019) sont représentées ci-dessous en figure 35. Ces implémentations diffèrent quant au degré de préservation de la phase d'enveloppe à la sortie des canaux de modulation. Clairement, le modèle PEMO prédit les capacités de détection d'une FM rapide mais sous-estime les capacités de détection de FM lente quelle que soit la porteuse.

En résumé, les informations d'AM et de FM sont potentiellement encodées par deux mécanismes distincts. Quelle que soit la porteuse et la cadence de modulation, l'AM serait essentiellement détectée grâce à un mécanisme spatio-temporel unique utilisant des indices neuronaux d'enveloppe temporelle (ENV_N). Ce mécanisme fait appel à une décomposition d'enveloppe grâce à un banc de filtres accordés dans le domaine des modulations, un bruit interne additif, un registre de stockage en mémoire à court terme d'environ une seconde et un mécanisme de décision par appariement de gabarit ou un mécanisme de décision basé sur des indices de puissance de modulation à long terme. Lorsque la porteuse est inférieure à 4 kHz et lorsque la cadence de modulation est très basse (<5-10 Hz), la FM serait essentiellement détectée grâce à un mécanisme purement temporel utilisant des indices neuronaux de verrouillage en phase sur la structure fine du signal à bande étroite (TFS_N). Toutefois, lorsque la porteuse est élevée (>1-4 kHz) ou encore lorsque la cadence de modulation est rapide (>5-10 Hz) et ce quelle que soit la fréquence de la porteuse, la FM serait détectée grâce au mécanisme spatio-temporel utilisant les indices neuronaux d'enveloppe temporelle (ENV_N) obtenus par reconversion FM-AM en sortie des filtres cochléaires. Cette conception à « double voie » du traitement d'AM et de FM est encore débattue et contestée au vu de données physiologiques et psychophysiques. Cette conception « à double voie » prédit que des lésions du système auditif produites par le vieillissement ou des pertes cochléaires sont susceptibles de produire des déficits de détection d'AM et de FM et plus particulièrement des dissociations entre traitement d'AM et de FM. Ces déficits pourraient participer aux troubles de reconnaissance systématiquement observés chez les personnes presbycousiques.

I.3. Effets du vieillissement et d'une perte auditive neurosensorielle

I.3.a. Préambule

Les adultes âgés manifestent souvent des difficultés de compréhension de la parole, et plus particulièrement dans des conditions d'écoute difficiles (e.g., Plomp & Mimpen, 1979 ; CHABA, 1998 ; Dubno et al., 2008). Généralement, ce déficit d'intelligibilité est associé à une perte auditive neurosensorielle diagnostiquée sur la base d'un audiogramme tonal (Van Rooij & Plomp, 1990). En audiologie clinique, cet outil de diagnostic consiste à mesurer le seuil de détection absolu en réponse à des sons purs, fréquence par fréquence. Ce test repose pour partie sur le mécanisme actif dans la cochlée et sa capacité à amplifier les sons faibles. Or, certains auditeurs se plaignent de problèmes de compréhension de la parole en situation bruyante tout en présentant des seuils audiométriques normaux (Grose et al., 2019 ; Saunders et al., 1989 ; King & Stephens, 1992). D'autres personnes avec une perte auditive en haute fréquence présentent des déficits d'intelligibilité - modestes ou massifs - en basse fréquence audio, où leurs seuils audiométriques sont normaux (Léger et al., 2012a,b,c ; 2014). Ces éléments suggèrent fortement qu'un simple audiogramme tonal n'est pas suffisant pour détecter une perte auditive neurosensorielle ; des **troubles auditifs dits « supraliminaires »** (au-dessus du seuil absolu) correspondant à des distorsions spectrales, temporelles, ou spectro-temporelles contribuent aux difficultés auditives des personnes malentendantes, et nécessitent des tâches auditives plus complexes pour être détectées (Grose et al., 2009 ; Plack et al., 2014).

L'importance des indices de modulation temporelle (AM et FM) dans la reconnaissance de la parole, des sons musicaux ou des sons environnementaux laisse donc présager que les difficultés auditives, et plus spécifiquement les troubles d'intelligibilité de la parole sont liés à un déficit d'encodage et de représentation des informations d'enveloppe temporelle (ENV_N) et de structure temporelle fine (TFS_N) à un niveau périphérique ou central (Pichora-Fuller & Souza, 2003). Les aspects du traitement auditif temporel des auditeurs âgés (tout en contrôlant la contribution de la perte cochléaire généralement associée au vieillissement) sont donc évalués à l'aide de différentes méthodes psychophysiques : i) mesures de **détection d'intervalles de silence** (Strouse et al., 1998 ; Phillips et al., 2000 ; Snell et al., 2002 ; Helfer & Vargo, 2009), ii) mesures de **détection d'AM et de FM** (He et al., 2007 ; 2008), iii) mesures de **masquage proactif** (Gifford et al., 2007) ; iv) mesures des capacités binaurales de discrimination de **différences interaurales de phase et de temps** (Divenyi & Haupt, 1997 ;

Grose & Mamo, 2010 ; Hopkins & Moore, 2011), mesures des capacités de **discrimination d'intervalles** ou de **durée** (Gordon-Salant & Fitzgibbons, 1993, 1999 ; Fitzgibbons & Gordon-salant, 1995 ; Strouse et al., 1998) et de **discrimination du rythme** et de **l'ordre temporel** (Trainor & Trehub, 1989 ; Humes & Christopherson, 1991 ; Gordon-Salant & Fitzgibbons, 1999 ; Fitzgibbons et al., 2006) et/ou de **mesures objectives** basées sur des techniques électrophysiologiques (Ross et al., 2007 ; Grose & Mamo, 2010 ; 2012).

Force est de constater que l'ensemble des résultats de mesures de traitement temporel associées à la perception de la parole ne sont pas entièrement cohérents. Ceci est probablement lié au fait que les études utilisent des méthodologies différentes et que les facteurs cognitifs ne sont pas toujours pris en compte dans l'analyse des résultats (CHABA, 1998). Ceci a conduit à l'hypothèse que l'effet de l'âge sur la relation entre le *traitement temporel* et la *perception de la parole* serait plus apparent dans des situations complexes du point de vue des stimuli et des tâches auditives (Pichora-Fuller & Souza, 2003). De nombreuses études confirment que l'acuité temporelle se détériore avec l'âge (Snell & Frisina, 2000 ; Gordon-Salant & Fitzgibbons, 1999). Il est à noter qu'un déficit d'acuité temporelle peut également être observé dans certaines pathologies telles que la *neuropathie auditive* (Zeng et al., 1999 ; 2012) ou la *sclérose en plaque* (Rappaport et al., 1994) et qu'un tel déficit d'acuité temporelle peut être la cause de troubles auditifs (Schneider, 1997 ; Schneider et al., 2002 ; Snell et al., 2002). Il est toutefois difficile de dissocier l'effet du vieillissement naturel de celui causé par une perte auditive neurosensorielle (à savoir une perte d'origine sensorielle ou neurale ; Moore, 2007). En effet, comme nous le verrons plus loin, de nombreuses études démontrent qu'une perte auditive neurosensorielle a des effets délétères et complexes sur la sensibilité à l'AM et la FM : alors que la perte auditive neurosensorielle dégrade la capacité à détecter la FM (*e.g.*, Lacher-Fougere & Demany, 1998 ; Moore & Skrodzka, 2002 ; Buss et al., 2004 ; Strelcyk & Dau, 2009 ; Ernst & Moore, 2012 ; Johannesen et al., 2016 ; Paraouty et al., 2016 ; Wallaert et al., 2018 ; Whiteford et al., 2020), elle préserve voire améliore la capacité à détecter l'AM, selon que les seuils de détection d'AM sont mesurés à des niveaux sonores équivalents (dB SPL) (*e.g.*, Bacon & Gleitman, 1992 ; Moore & Glasberg, 2001 ; Sek et al. 2015 ; Schlittenlacher & Moore, 2016) ou à des niveaux de sensation équivalents (dB SL) (Bacon & Gleitman, 1992 ; Moore et al., 1992 ; Moore et al., 1996 ; Füllgrabe et al., 2003 ; Wallaert et al., 2017).

Ce troisième chapitre a pour objectif de récapituler les principales études psychophysiques et quelques études neurophysiologiques importantes traitant des effets du vieillissement et/ou de la perte auditive neurosensorielle sur le codage et la perception de l'AM et de la FM.

I.3.b. Effets du vieillissement sur le traitement des indices d'AM et de FM

Différentes méthodes ont été utilisées pour dissocier les effets de l'âge et de la perte auditive (diagnostiquée sur la base de l'audiogramme tonal) afin de caractériser l'effet de l'âge *per se* sur le traitement des indices d'AM et de FM (Pichora-Fuller & Souza, 2003) : i) comparaison de deux groupes d'auditeurs (auditeurs normo-entendants jeunes vs. auditeurs normo-entendants âgés) ; ii) comparaison de deux groupes d'auditeurs dont les seuils auditifs moyens ne diffèrent que de quelques décibels HL sur l'ensemble des fréquences testées ou bien iii) évaluation des différences inter-individuelles à partir des corrélations entre des mesures audiométriques tonales et psychophysiques.

I.3.b.i. Effets du vieillissement sur la sensibilité à l'AM

Takahashi et Bacon (1992) mesurent les capacités de détection d'une AM sinusoïdale d'un groupe de dix auditeurs normo-entendants jeunes (d'âge moyen égal à 26 ans) et de trois groupes de dix participants âgés (de moyenne d'âge respectivement égale à 54.3 ans ; 64.8 ans ; 72.2 ans) présentant des seuils audiométriques tonaux normaux ou une perte auditive neurosensorielle légère à modérée en haute fréquence. Les fonctions de transfert de modulation temporelle (TMTF) sont représentées en figure 36. Elles ont été obtenues à l'aide d'une procédure adaptative à deux alternatives en choix forcé consistant à détecter le signal d'AM cible appliqué à un bruit large-bande. Notons qu'un bruit introduit des fluctuations aléatoires intrinsèques susceptibles de masquer la modulation cible (Dau et al., 1997). Les cadences de modulation sont comprises entre 2 et 1024 Hz et les signaux sont présentés à 35 dB SPL. Les résultats indiquent que les TMTF des auditeurs âgés présentent une forme similaire mais les seuils de détection sont moins bons à 2 Hz qu'à 4 et 8 Hz, conférant aux TMTF l'apparence d'une fonction passe-bande plutôt que passe-bas. Les sujets âgés présentent également des seuils de détection d'AM légèrement plus mauvais aux cadences de modulation très rapides. Néanmoins, l'analyse statistique ne révèle pas de différences significatives de sensibilité à l'AM entre les différents groupes d'âge mais une corrélation significative est obtenue entre la sensibilité à l'AM lente et l'âge lorsque l'audibilité (la contribution des seuils audiométriques tonaux) est contrôlée. Cet effet de l'âge sur la sensibilité à l'AM a déjà été suggéré par Bacon et Viemeister (1985).

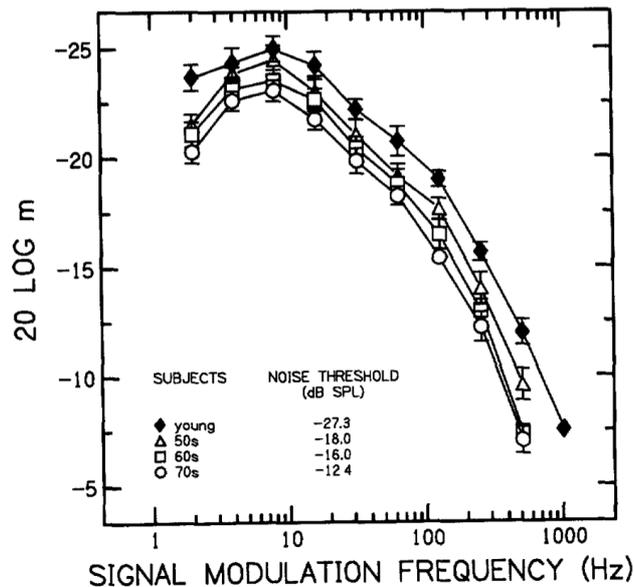


Figure 36. Fonctions de transfert de modulation temporelle (TMTF) moyennes pour chaque groupe d'âge. Les seuils de détection d'AM sont représentés (en ordonnée) en fonction de la cadence de modulation (en abscisse). Les groupes de sujets âgés sont représentés par des symboles ouverts et le groupe de sujets jeunes est représenté par des symboles fermés. Le seuil absolu moyen pour un bruit large bande est indiqué pour chaque groupe en bas de la figure. Extrait de Takahashi et Bacon (1992).

L'absence d'effet clair de l'âge est cohérente avec les études portant sur la compréhension de la parole dans le bruit chez des sujets âgés (Gelfand et al., 1986 ; Humes & Roberts, 1990 ; Humes et al., 1991). Toutefois, il convient de noter que la tranche d'âge testée ne dépasse pas 77 ans. L'étude de Humes et Christopherson (1991) révèle des déficits de compréhension de la parole sur un groupe d'âge compris entre 65 et 75 ans. **Ces données comportementales suggèrent que les effets du vieillissement sur la détection d'AM n'apparaissent clairement qu'à partir d'un certain âge. L'effet semble plus marqué aux basses cadences d'AM mais les résultats suggèrent que le vieillissement pourrait également affecter le traitement des AM rapides.**

L'enregistrement et l'analyse des **potentiels auditifs évoqués (PEA)** en réponse à une AM sinusoïdale (EFR : Envelope Following Response ; AMFR : AM Following Response, ASSR : Auditory Steady State Response) chez des auditeurs normo-entendants jeunes et âgés offre une approche alternative aux mesures psychophysiques (Gransier et al., 2021). Il est possible d'utiliser ces mesures car Rees et al. (1986) et Roß et al. (2000) démontrent que les TMTF mesurées à l'aide de cette technique objective présentent des traits similaires à ceux décrits par Viemeister (1979) sur des TMTF « comportementales ». Par ailleurs, Picton et al. (1987) soulignent une correspondance entre les seuils de détection d'AM et les AMFR obtenus sur un même participant, suggérant que les potentiels évoqués auditifs en réponse à l'AM peuvent être utilisés comme une mesure objective fiable permettant d'estimer les capacités de détection d'AM. La discrimination d'intensité et la détection d'AM étant des processus étroitement liés (Wojtczak & Viemeister, 1999), des différences dans les potentiels auditifs évoqués en réponse à l'AM entre deux groupes de participants sont susceptibles de refléter

des différences de discrimination d'intensité. Plus récemment, Gransier et al. (2021) mesurent des TMTF à l'aide d'ASSR chez l'humain sur la gamme 0.5-100 Hz et démontrent une plus grande variabilité inter-individuelle dans les réponses aux cadences les plus lentes, inférieures à 20 Hz, et dont l'origine est corticale.

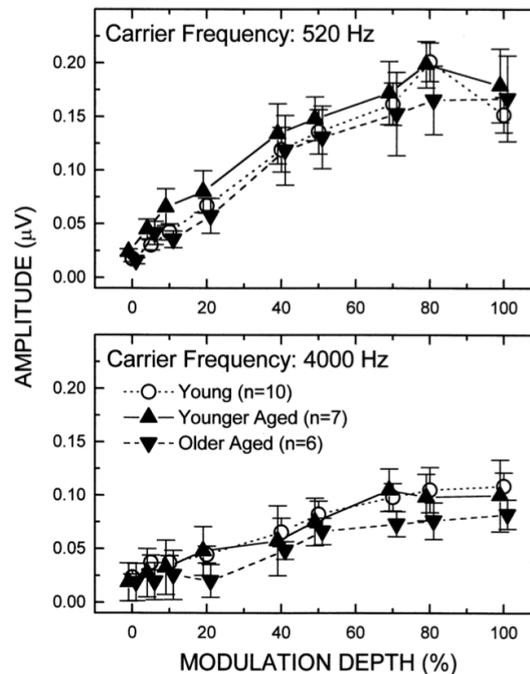


Figure 37. Amplitude moyenne des potentiels évoqués auditifs en réponse à une AM sinusoïdale pour une porteuse de 520 Hz (figure du haut) et de 4000 Hz (figure du bas) chez des sujets jeunes (cercles ouverts) et âgés (triangles) en fonction de la profondeur de la modulation (en %). Extrait de Boettcher et al. (2001).

Sur la base de ces études, Boettcher et al. (2001) comparent les potentiels évoqués de dix sujets normo-entendant jeunes (22-29 ans), de sept sujets normo-entendants âgés (60-65 ans) et de six sujets normo-entendants plus âgés (66-72 ans). Leurs seuils audiométriques sont tous compris entre 0 et 25 dB HL entre 250 Hz et 8000 Hz. Les signaux sont des sinusoïdes de porteuse égale à 520 et 4000 Hz, modulées en amplitude à 40 Hz à des profondeurs de modulation comprises entre 0 et 100 %. Ces sons sont présentés de manière monaurale à un niveau de 65 dB SPL. **Les résultats électrophysiologiques n'indiquent pas d'effet significatif du vieillissement sur l'amplitude ou la phase des réponses évoquées par l'AM lorsque la porteuse est de 520 Hz ou de 4000 Hz.** En effet, les résultats obtenus sur les deux groupes (*cf.* figure 37) sont quasi identiques. Néanmoins, l'effectif restreint à 6-7 participants âgés est peut-être insuffisant pour faire apparaître des différences dans les potentiels évoqués étant donné le degré de variabilité dans l'amplitude des réponses.

Les résultats obtenus par Boettcher et al. (2001) sont similaires à ceux précédemment rapportés par Stapells et al. (1984) qui n'obtiennent pas de différences significatives dans l'amplitude des AMFR de sujets âgés entre 21 à 61 ans. Ces résultats sont également similaires à ceux obtenus par Johnson et al. (1988) qui ne démontrent pas d'effet du vieillissement sur

les AMFRS mesurées en réponse à AM de 40 Hz appliquée à une porteuse de 4000 Hz chez des sujets normo-entendant jeunes (36-40 ans ; n=5) et plus âgés (65-77 ans ; n=5).

Purcell et al. (2004) évaluent l'acuité temporelle (la capacité à discriminer des changements rapides dans l'enveloppe d'un son) sur deux groupes d'auditeurs : quatorze sujets normo-entendant jeunes (18-43 ans) et treize sujets normo-entendant âgés (60-78 ans) en combinant les mesures comportementales de détection d'AM et de détection d'intervalles de silence et des mesures objectives de PEA (EFR). Dans cette étude, une procédure adaptative à deux alternatives en choix forcé estime la cadence de modulation maximale (variable d'intérêt) à partir de laquelle les participants sont capables de détecter une AM présentant une profondeur de modulation de 25 % appliquée à un bruit blanc. Une technique de balayage des cadences de modulation est utilisée et la tâche est réalisée sur un nombre d'essais restreint afin de maintenir une durée de test raisonnable. Les résultats indiquent une **décroissance significative de l'acuité temporelle auditive avec l'âge** : tandis que la cadence de modulation maximale est de 567 Hz pour les participants normo-entendants jeunes, elle passe à 264 Hz chez les participants normo-entendants âgés. Un effet significatif de l'âge est observé lorsque les trois mesures respectives sont corrélées deux à deux. En combinant ces trois méthodes, Purcell et al. (2004) montrent que **le vieillissement dégrade l'acuité temporelle. Néanmoins, dans cette étude, les effets de l'âge sur l'acuité temporelle pourraient être liés à la présence d'une perte auditive modérée en haute fréquence audio chez les participants âgés.**

Leigh-Paffenroth et Fowler (2006) utilisent également des mesures électrophysiologiques pour quantifier la capacité du système auditif à *verrouiller ses réponses en phase* sur l'enveloppe du stimulus (Rickards & Clark, 1982). Ces réponses neuronales sont mesurées sur des auditeurs normo-entendants jeunes (âge moyen : 29.6 ans ; n=11) et plus âgés (âge moyen : 69,9 ans ; n=11). Les stimuli sont présentés à un niveau supraliminaire de 65 dB SPL et constituent des sons purs de 500 Hz et 2000 Hz modulés en amplitude à des cadences de modulation de 20, 40 et 90 Hz. Les profondeurs de modulation sont comprises entre 20 % (où le verrouillage en phase devrait être difficile) et 100 % (où le verrouillage en phase devrait être important) (Kuwada et al., 1986). La figure 38 (figure de gauche) indique que lorsque la porteuse est basse (500 Hz), les auditeurs les plus jeunes (barres grises) présentent un nombre plus élevé de réponses neuronales verrouillées en phase que les auditeurs plus âgés (barres blanches) tandis qu'à 2000 Hz, les auditeurs jeunes et âgés présentent autant de réponses neuronales à 20 qu'à 90 Hz et plus de réponses à 40 Hz comparé aux auditeurs normo-entendants jeunes (barres grises).

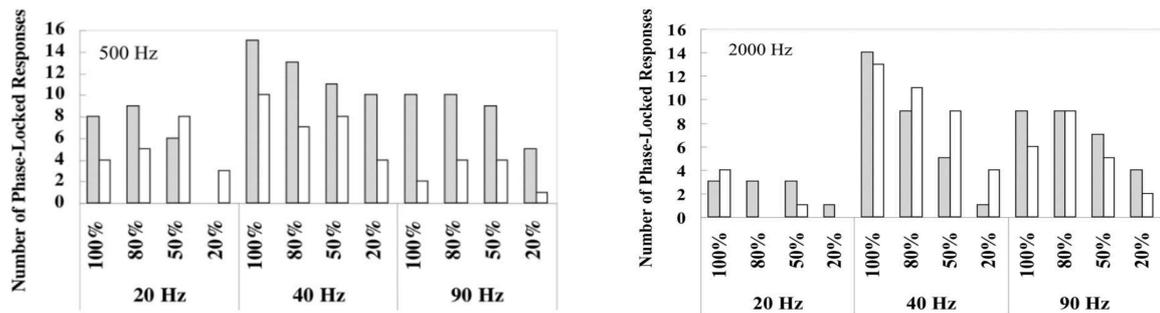


Figure 39. Nombre de réponses verrouillées en phase (en ordonnée) en fonction de la cadence de modulation pour différentes profondeurs de modulation (en abscisse) et pour une porteuse de 500 Hz (figure de gauche) ou de 2000 Hz (figure de droite). Le verrouillage en phase est déterminé à l'aide d'un algorithme de détection automatique basé sur les propriétés statistiques du bruit de fond (de l'EEG) et sur les propriétés statistiques de l'ASSR. Les barres grises représentent les données des sujets normo-entendants jeunes ; les barres blanches représentent les données des sujets normo-entendants plus âgés. Le nombre maximal de réponses est de 16. Extrait de Leigh-Paffenroth et Fowler (2006).

Ainsi, les résultats de cette étude suggèrent que chez l'humain, le codage neural des indices d'ENV_N dépend de la cadence, de la porteuse et de l'âge :

- Premièrement, les réponses évoquées par les AM les plus lentes sont générées dans certaines régions du cortex auditif, alors que celles évoquées par les AM les plus rapides sont générées dans des régions du thalamus. Les réponses provenant du tronc cérébral (évoquées par les cadences rapides) diminuent avec l'âge (Roß et al., 2002).
- Deuxièmement, si le codage temporel des modulations portées par un son pur de haute fréquence n'est pas affecté avec l'âge, celui-ci est réduit dans la région basse fréquence audio. Les auteurs expliquent cette différence en termes de réduction d'innervation neuronale dans cette région fréquentielle (Leigh-Paffenroth & Fowler, 2006).

He et al. (2008) évaluent les effets de l'âge sur la détection d'AM en comparant les fonctions de transfert de modulation temporelle de huit sujets jeunes (âge moyen de 21 ans) et de huit sujets plus âgés (âge moyen de 71 ans) présentant des seuils absolus inférieurs à 20 dB HL (ANSI, 1996). Les seuils de détection d'AM sont mesurés à l'aide d'une méthode dite « *de maximum de vraisemblance* » décrite par Green (1993) dans une tâche « oui/non » à un intervalle. Dans cette tâche, la profondeur de modulation est modifiée de manière adaptative pour converger vers le seuil. Les TMTF (voir figure 40 ci-dessous) sont mesurées à l'aide d'un son pur (la porteuse) de fréquence égale à 500 ou 4000 Hz modulé à des cadences de modulation comprises entre 5 et 250 Hz. Tous les sujets sont initialement entraînés à une cadence de 5 Hz avant d'être testés aux autres cadences. Les résultats indiquent que **les seuils de détection d'AM sont significativement plus élevés (moins bons) chez les auditeurs plus âgés aux cadences de modulation rapides.**

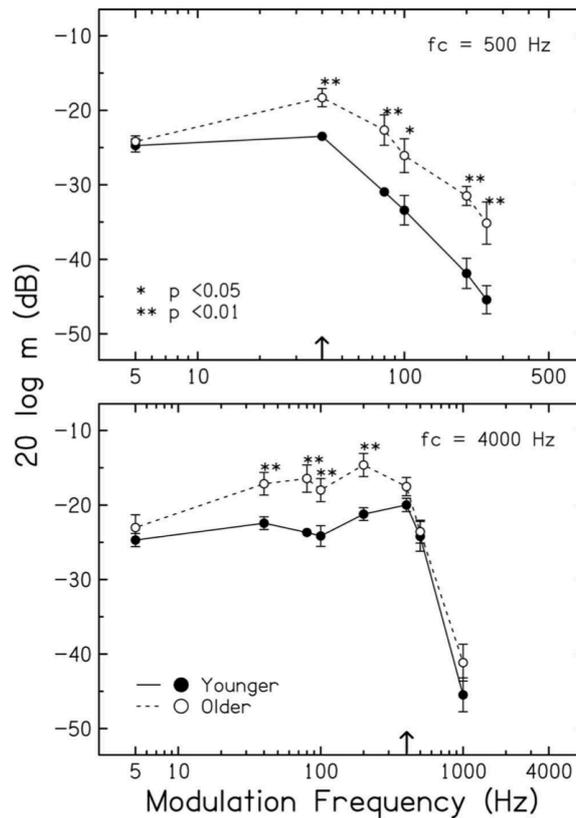


Figure 40. Fonctions de transfert de modulation temporelle pour une porteuse de 500 Hz (figure du haut) et de 4000 Hz (figure du bas) mesurées sur des deux groupes d'auditeurs (les jeunes étant représentés par les cercles pleins et les âgés étant représentés par les cercles ouverts). Extrait de He et al. (2008).

De plus, la forme des courbes diffère entre les deux groupes d'âge. Comme prédit, pour les cadences de modulation inférieures à la fréquence de transition (où la détection d'AM est basée principalement sur des indices purement temporels), les différences liées à l'âge augmentent avec la cadence de modulation avec une même tendance dans les fréquences audios basses et élevées. Contrairement aux sujets normo-entendants jeunes, **les seuils de détection d'AM des sujets plus âgés augmentent progressivement à mesure que la cadence de modulation se rapproche de la fréquence de transition suggérant un déclin de la résolution temporelle**. Ainsi, cette étude démontre des effets significatifs liés à l'âge bien que les fréquences de transition soient similaires chez les auditeurs normo-entendants jeunes et âgés (car la sélectivité fréquentielle n'est pas affectée par le vieillissement).

Schoof et Rosen (2014) évaluent le déclin des capacités auditives et cognitives associées au vieillissement à l'aide de tâches de détection d'AM et de FM ainsi que des tâches mesurant la mémoire de travail, l'attention et la rapidité de traitement auditif. Dix-neuf participants normo-entendants jeunes (âgés entre 19 et 29 ans ; âge moyen de 24 ans) et dix-neuf participants normo-entendants plus âgés (60-72 ans ; âge moyen : 64 ans) sont testés dans une tâche à trois alternatives en choix forcé dans laquelle l'AM cible est appliquée à une porteuse constituée par un bruit large bande (1-4 kHz). Les TMTF sont mesurées pour des cadences de modulation égales à 10, 20, 40, 80 et 160 Hz et les signaux sont présentés à un

niveau de 70 dB SPL. Avant de démarrer la tâche, les participants sont familiarisés dans quelques conditions. Contrairement à He et al. (2008), **les résultats de cette étude ne montrent pas d'effet significatif de l'âge sur les seuils de détection d'AM**. Ceci pourrait être lié au fait que les effets du vieillissement n'apparaissent que pour des cadences de modulation élevées, supérieures à 200 Hz (Purcell et al., 2004 ; Grose et al., 2009). Ceci pourrait également résulter du fait que le bruit faisant office de porteuse véhicule des fluctuations aléatoires intrinsèques susceptibles de masquer l'AM cible, et ce avant tout aux hautes cadences de modulation où les filtres de modulation sont plus larges. Un tel phénomène de masquage n'a pas lieu dans l'étude de He et al. (2008). Néanmoins, l'étude de Schoof et Rosen (2014) révèle que le vieillissement serait associé à un déclin cognitif, plus spécifiquement dans la mémoire de travail et la rapidité de traitement.

Füllgrabe et al. (2015) mesurent les TMTF de neuf sujets jeunes (âgés entre 18 et 27 ans ; âge moyen 23 ans) et vingt et un sujets plus âgés (60-79 ans ; âge moyen : 67 ans) à l'aide d'une procédure à 3 alternatives en choix forcé à un niveau de 30 dB SL. Les auditeurs (appariés sur la base de leur audiogramme et de leur niveau éducatif) sélectionnent l'intervalle (parmi les trois) contenant l'AM cible. La porteuse est un son pur de 4000 Hz et les cadences sont égales à 5, 30, 90, et 180 Hz. Les participants sont entraînés sur deux mesures répétées avant de démarrer la tâche, et ce pour cadence de modulation. Les résultats sont présentés en figure 41.

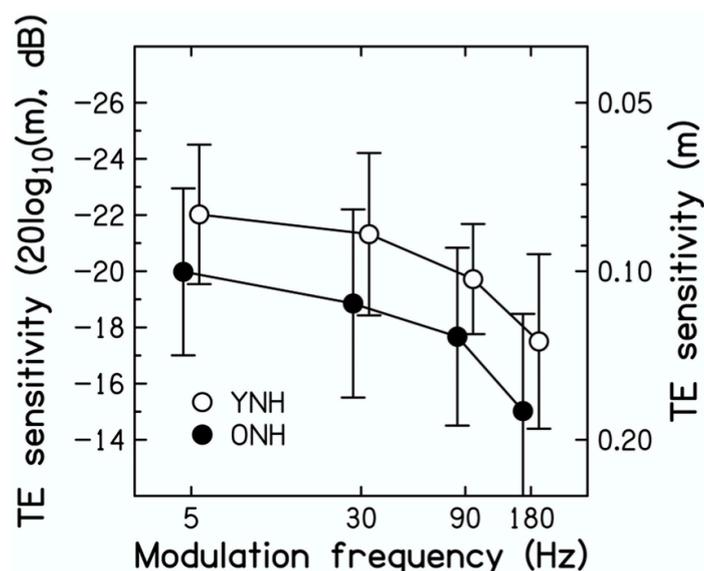


Figure 41. Fonctions de transfert de modulation temporelle (moyennées sur l'ensemble du groupe) des sujets normo-entendants jeunes (cercles ouverts) et plus âgés (cercles fermés). Les seuils de détection d'AM sont exprimés en dB (20log₁₀m) (axe des ordonnées de gauche) ou sur une échelle linéaire de profondeur de modulation (m ; axe des ordonnées de droite). L'axe des abscisses indique la cadence de modulation. Extrait de Füllgrabe et al. (2015).

Les seuils de détection d'AM des sujets âgés sont significativement plus élevés (dégradés) de 2 à 2.5 dB sur l'ensemble des cadences de modulation testées (entre 5 et 180 Hz). Dans cette étude et contrairement à d'autres études comme celle de Leigh-Paffenroth et Fowler (2006) citée plus haut, *l'effet du vieillissement est indépendant de la cadence de modulation.* Les résultats de cette étude suggèrent donc que **sous l'effet du vieillissement, les seuils de détection d'AM se dégradent globalement (pour toutes les cadences) en raison d'une diminution de l'« efficacité de traitement » (une moindre capacité à utiliser optimalement les indices d'ENV_N disponibles) et non à cause d'une diminution de la résolution temporelle.**

Wallaert et al. (2016) mesurent les seuils de détection d'AM à un niveau de 40 dB SL sur des auditeurs normo-entendants jeunes (22-28 ans) et des auditeurs normo-entendants plus âgés (44-66 ans) en utilisant une porteuse de 500 Hz et deux cadences de modulation (2 et 20 Hz). Le nombre de cycles de modulation varie entre 2 et 9 afin d'évaluer l'intégration temporelle de l'AM chez chaque groupe de sujet. En accord avec les résultats de He et al. (2008) et Füllgrabe et al. (2015), Wallaert et al. (2016) montrent que l'effet du facteur groupe sur les seuils de détection d'AM est significatif et ce *aux deux cadences testées.* Toutefois, cet **effet de l'âge - bien que significatif - est modeste (l'élévation des seuils étant de l'ordre de 2 dB).**

Paraouty et Lorenzi (2017) changent d'approche afin d'étudier l'effet du vieillissement et utilisent une méthode corrélative. Ici, les auteurs évaluent les effets du vieillissement sur une cohorte de soixante-dix sujets âgés entre 20 et 70 ans (âge moyen : 42 ans). Les seuils de détection d'AM sont mesurés à l'aide d'une procédure adaptative à deux alternatives en choix forcé estimant le seuil correspondant à 70,7 % de réponses correctes (Levitt, 1971). La cadence d'AM est lente (5 Hz) et la porteuse est un son pur de 500 Hz. Les résultats montrent que **les seuils de détection d'AM ne sont pas significativement corrélés avec l'âge,** en accord avec les résultats de l'étude de Takahashi et Bacon (1992) et de Schoof et Rosen (2014) et plus récemment celle de Whiteford et al. (2017). Les auteurs suggèrent que les différences observées dans les seuils de détection d'AM pourraient être liées à des facteurs procéduraux dont l'influence peut être réduite par l'entraînement (comme dans l'étude de He et al., 2008) ou en testant un grand nombre de conditions (comme dans le cas de l'étude de Wallaert et al., 2016).

Quoiqu'il en soit, ces différentes études suggèrent que les effets du vieillissement sur la sensibilité à l'AM, s'ils existent, sont de *faible amplitude.* Ceci est corroboré par les résultats d'une étude menée sur une cohorte importante de sujet (n=102) par Carcagno et Plack (2021). Ces derniers estiment qu'à des niveaux de présentation élevés, **les seuils de détection d'AM mesurés avec une porteuse de 2 kHz à 25, 50 et 100 Hz augmentent avec l'âge avec un taux de 0,1-1,5 dB par décennie d'âge.**

En résumé, les résultats de ces nombreuses études psychophysiques et électrophysiologiques sont parfois divergents. Certaines d'entre elles n'indiquent pas d'effet de l'âge sur la sensibilité à l'AM tandis que d'autres indiquent que le vieillissement affecte la sensibilité à l'AM à toutes les cadences (lentes et rapides) ou uniquement lorsque la cadence de modulation est rapide. Certaines études concluent que la résolution temporelle en soi n'est pas affectée par l'âge ; des facteurs non-sensoriels associés au vieillissement seraient responsables de ces déficits de sensibilité à l'AM, à savoir des facteurs cognitifs ou procéduraux susceptibles de réduire l'« efficacité de traitement » de l'ENV_N, à savoir la capacité à utiliser de manière optimale l'information sensorielle disponible à la sortie des filtres de modulation centraux. Un bruit interne plus important, des stratégies d'écoute sous-optimales, des déficits mnésiques ou attentionnels pourraient être responsables de cette réduction d'efficacité de traitement de l'AM. Quoiqu'il en soit, l'effet du vieillissement sur la sensibilité à l'AM, lorsqu'il est observé, est généralement modeste.

I.3.b.ii. Effets du vieillissement sur la sensibilité à la FM

Plusieurs études électrophysiologiques et psychophysiques indiquent que le codage des indices de TFS_N monauraux ou binauraux se dégrade sous l'effet du vieillissement et ce indépendamment de la perte auditive (Ross et al., 2007 ; Clinard et al., 2010 ; Grose & Mamo, 2010 ; Hopkins & Moore, 2011 ; Moore et al., 2012 ; Füllgrabe, 2013 ; King et al., 2014 ; Füllgrabe et al., 2015).

Boettcher et al. (2001) évaluent l'effet de l'âge sur la détection de FM à l'aide d'une mesure objective de PEA obtenue en réponse à un signal de FM sinusoïdale (mesure de « FMFR » ou « *frequency-modulation following response* »). Cette mesure est supposée refléter la fidélité du codage des indices de TFS_N monauraux et pourrait constituer un corrélât neuronal de la détection auditive de FM (Picton et al., 1987 ; John et al., 2001). Ici, l'amplitude des FMFR est mesurée en réponse à une FM portée par un son pur de fréquence égale à 500 Hz. Dans une première expérience, cette porteuse est modulée à une cadence de 3,91 ; 7,81 ; 15,63 ; 31,25 et 37,91 Hz tandis que la profondeur de modulation est maintenue à 80 %. Dans une deuxième expérience, la cadence de modulation est fixée à 38 Hz et la profondeur de modulation varie entre 0 et 80 %. Dans les deux expériences, les signaux sont présentés à un niveau de 80 dB SPL.

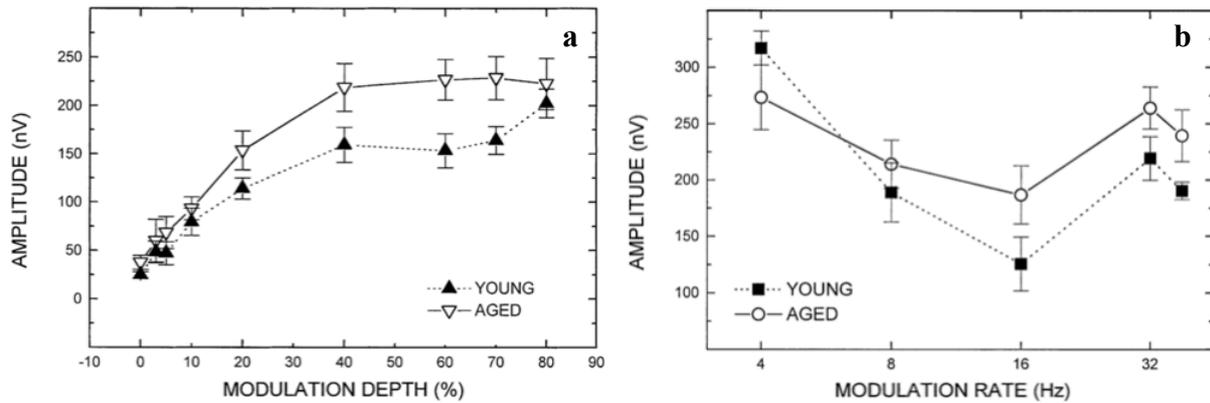


Figure 42. Amplitude moyenne des FMFR tracées en fonction de la cadence de modulation (à droite) et de la profondeur de modulation (à gauche). Dans la figure de droite, la profondeur de modulation est fixée à 80 %. Dans la figure de gauche, la cadence de modulation est fixée à 38 Hz. Extrait de Boettcher et al. (2001).

Les résultats de ces expériences sont présentés en figure 42. **Bien que la variabilité individuelle soit élevée, l'amplitude de la FMFR est significativement plus grande chez les personnes âgées sur l'ensemble des valeurs de profondeur de modulation testées** (figure 42a) et pour plusieurs valeurs de cadences de modulation : 8, 16, 32 et 38 Hz (figure 42b). Ces résultats révèlent des différences entre sujets jeunes et âgés, mais ils sont contre-intuitifs : si le codage des indices temporels est dégradé (distordu) par le vieillissement, on s'attendrait plutôt à une baisse qu'à une hausse de l'amplitude des FMFR chez les sujets âgés. Ces résultats neurophysiologiques n'apportent donc pas de réponse claire quant à l'effet de l'âge mais ils sont néanmoins cohérents avec ceux obtenus par Picton et al. (1987) qui obtiennent des amplitudes de FMFR plus grandes pour les réponses de latence moyenne et tardive de sujets âgés dans certaines conditions expérimentales (pour des cadences de modulation égales à 4,9 et 39,1 Hz et pour des profondeurs de modulation comprises entre 10 et 70 %).

He et al. (2007) mesurent les seuils de détection de FM sur huit sujets normo-entendants jeunes et huit sujets normo-entendants plus âgés à l'aide de deux méthodes psychophysiques différentes. L'une de ces deux procédures est une tâche à stimuli constant utilisant un paradigme oui/non. La porteuse est un son pur de 500 ou 4000 Hz, modulé à une cadence de 5 Hz à l'aide d'une FM de forme sinusoïdale ou quasi-sinusoïdale. Avant de démarrer la tâche, les participants sont entraînés à la détection de FM. Les résultats obtenus à partir de la procédure à stimuli constants indiquent que **les seuils de détection de FM se dégradent sous l'effet du vieillissement et ce avant tout en basse fréquence** plutôt qu'en haute fréquence (cf. fonctions psychométriques présentées en figure 43).

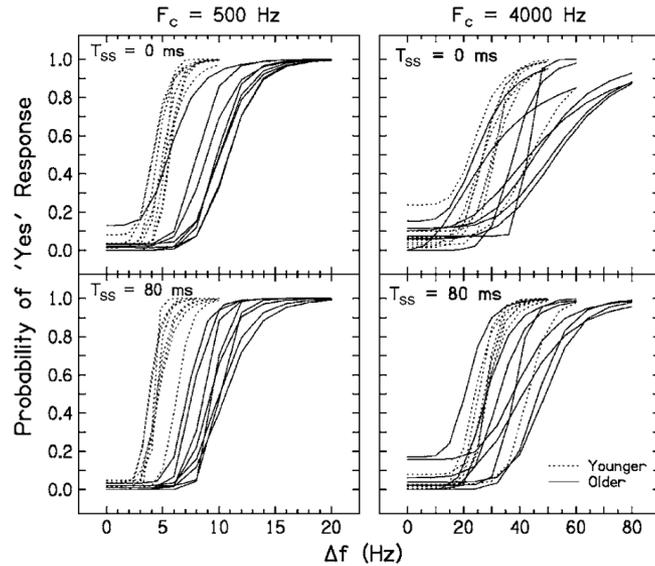


Figure 43. Fonctions psychométriques individuelles de huit sujets normo-entendants jeunes (courbes en pointillés) et de huit sujets plus âgés (courbes continues) mesurées avec une tâche à stimuli constants de type oui-non. Les résultats obtenus pour une porteuse de 500 Hz et de 4000 Hz sont présentés respectivement dans les colonnes de gauche et de droite. Les lignes du haut et du bas représentent respectivement les résultats obtenus pour des signaux de durée égale à 0 ms et à 80 ms. Les fonctions psychométriques des sujets âgés sont décalées vers la droite par rapport aux sujets jeunes. Extrait de He et al. (2007).

Les auteurs concluent que l'effet du vieillissement s'explique par un déclin du mécanisme temporel de verrouillage en phase des décharges neuronales sur les indices de TFS_N impliqué dans la détection de FM.

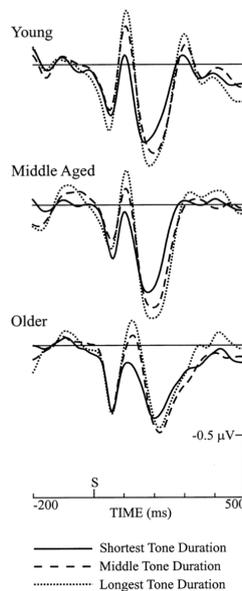


Figure 44. Potentiels auditifs moyennés sur l'ensemble des participants de chaque groupe pour le stimulus de durée la plus courte, intermédiaire et longue chez des adultes normo-entendants jeunes ($n=10$; âge moyen : 27 ans ; [21-31 ans]), d'âge intermédiaire ($n=10$; âge moyen : 47 ans ; [37-58 ans]) et plus âgés ($n=9$; âge moyen : 69 ans ; [60-78 ans]). La négativité est tracée vers le haut. La lettre "S" dans l'échelle indique le début du stimulus. L'onde N1 représente le premier pic négatif, l'onde P2 représente le second pic positif, l'onde N2 représente le second pic négatif. Extrait de Ostroff et al. (2003).

La majorité des études visant à caractériser le traitement de la FM par les auditeurs normo-entendants jeunes et âgés utilisent des tâches psychophysiques monaurales (cf. Boettcher et al., 2001 ; He et al., 2007). Cependant, certaines mesures de capacités auditives binaurales dépendent - du moins en partie - de la fidélité du codage des indices la structure temporelle fine (TFS_N): c'est notamment le cas dans les estimations du « niveau de masquage binaural » (BMLD pour binaural masking level difference) (Strelcyk & Dau, 2009) et dans la discrimination des différences interaurales de phase (IPD) (Yost, 1975 ; Lacher-Fougere & Demany, 2005). De plus, il paraît de plus en plus évident que **des troubles supraliminaire de nature temporelle apparaissent relativement tôt dans le processus du vieillissement**, à savoir à un âge dit « intermédiaire » (middle-age) c'est-à-dire à partir de 45-55 ans (Grose et al., 2006; Helfer & Vargo, 2009). Ces différences s'expliquent par des modifications anatomiques et physiologiques apparaissant très tôt au sein du système auditif. Des mesures psychophysiques de détection ou de discrimination d'intervalle comme dans les travaux de Artieda et al. (1992) et plus récemment de Rammasayer et Classen (1997) soulignent le fait que la capacité à détecter des événements sonores brefs est associée à un certain niveau d'activité de dopamine élevé dans les noyaux gris centraux. Or, des études d'imagerie cérébrale suggèrent que dans certaines régions cérébrales, les récepteurs à la dopamine diminuent durant toute la vie de l'adulte (Kaasinen et al., 2000) et ce, de manière évidente à partir de 36 ans (Mozley et al., 1999).

Des mesures électrophysiologiques révèlent un déficit de codage temporel chez des auditeurs normo-entendants d'âge intermédiaire et des anomalies dans les réponses de latence longue (N1-P2) à des sons purs de différentes durées. Ces ondes, présentées en figure 44 montrent que **le traitement de la durée des sons est modifié sous l'effet du vieillissement** : les courbes (traits pleins et pointillés) présentent dans l'ensemble une amplitude inférieure à celle des sujets normo-entendants jeunes ; par ailleurs, les latences entre les ondes N1 et P2 sont significativement différentes entre les sujets normo-entendants jeunes et d'âge intermédiaire. Néanmoins, le fait que les adultes jeunes et d'âge intermédiaire présentent des ondes N1 et P2 semblables suggère que **le déclin lié à l'âge dans le traitement temporel des sons démarre après la cinquième décennie de la vie**.

A la suite de ces travaux électrophysiologiques, Grose et Mamo (2010) évaluent le traitement auditif de la TFS_P des sons par des sujets d'âge intermédiaire à l'aide d'une tâche de discrimination de différence de phase interaurale (IPD). Trois groupes de participants sont testés dans cette étude : treize sujets normo-entendants jeunes (âge moyen : 22 ans ; étendue : [18-27 ans]), dix sujets d'âge intermédiaire (âge moyen : 47 ans ; étendue : [40-55 ans]) et dix sujets plus âgés (âge moyen : 68 ans ; étendue : [63-75 ans]). Tous les participants présentent des seuils audiométriques inférieurs à 20 dB HL entre 250 et 2000 Hz. Les signaux sont des sons purs modulés en amplitude avec une AM sinusoïdale de cadence égale à 5 Hz, dont la profondeur de modulation est fixée à 100 % et dont la durée est de 800 ms (soit quatre périodes de modulation). En raison de l'AM, les signaux présentent donc quatre moments où

l'amplitude atteint zéro. Les différences interaurales de phase sont introduites dans ces passages à zéro afin d'éviter de produire des artefacts. Les sons sont présentés de manière binaurale à un niveau de 65 dB SPL. Une procédure adaptative à trois alternatives en choix forcé convergeant vers 79 % de réponses correctes est utilisée pour mesurer les seuils de détection d'IPD. Cette expérience consiste donc à utiliser des phases interaurales fixes (0 ou 2π radians) et à faire varier la fréquence de la porteuse d'une AM sinusoïdale. Le seuil mesuré ici correspond la plus haute fréquence audio de la porteuse pour laquelle le signal présentant une inversion de phase interaurale dans les passages à zéro de l'AM peut être discriminé de sa version diotique (sans inversion de phase interaurale). Les résultats de cette expérience sont donc supposés être reliés à la limite (ou fréquence de coupure) du phénomène de verrouillage en phase sur les indices de TFS_N . Ces résultats présentés en figure 45 ci-dessous démontrent que **la capacité de discrimination de différences interaurales de phase diminue avec l'âge et que ce déficit est évident chez des auditeurs d'âge intermédiaire**. Cela suggère que **la dégradation du traitement des indices de TFS_N démarre tôt (dans le système auditif dit « pré-sénescence »)**. Ces résultats sont comparés à ceux de Ross et al. (2007) qui enregistrent des PEA dans des conditions similaires : la gamme de fréquence « seuil » est plus étendue pour les auditeurs âgés que pour les auditeurs normo-entendants jeunes ; par ailleurs, cette fréquence seuil diminue en fonction de l'âge. Les résultats de Grose et Mamo (2010) confirment ceux de Ross et al. (2007) : la dégradation de l'audition binaurale liée au vieillissement est significative à partir d'un âge intermédiaire.

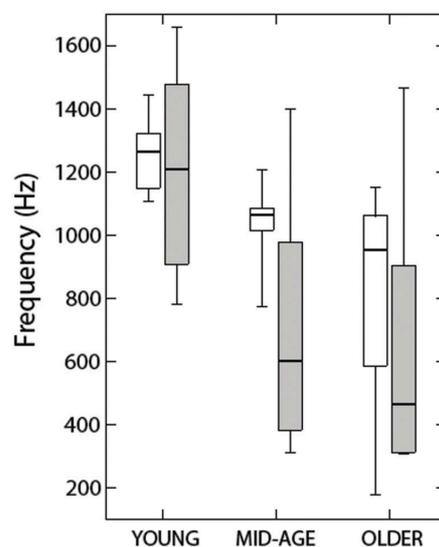


Figure 45. Limite supérieure de fréquence pour la discrimination des IPD 0- et π -radians pour chaque groupe d'âge. Les rectangles ouverts correspondent aux données de cette étude ; les rectangles ombragés sont ceux de Ross et al. (2007). Chaque rectangle englobe le 25e à 75e percentile ; la barre horizontale épaisse au sein de chaque rectangle indique la médiane. Les bouchons sur les lignes verticales indiquent les 10e et 90e percentiles. Extrait de Grose et Mamo (2010).

Dans la continuité de cette étude, Grose et Mamo (2012) évaluent les capacités de détection de FM autour de 500 Hz en conditions monaurale, diotique et dichotique (en utilisant un mode de stimulation binaural où la fonction de modulation est en opposition de phase entre les

deux oreilles). Trente-six participants répartis en trois groupes distincts de 12 participants sont testés : un groupe de sujets normo-entendants jeunes âgés entre 19 et 29 ans ; un groupe de sujets normo-entendants d'âge intermédiaire (entre 43 et 57 ans) et un groupe de sujets plus âgés (65-77 ans). Les signaux de FM sont produits à partir d'une porteuse (dont la fréquence audio varie aléatoirement entre 460 et 540 Hz) modulée à une cadence de 2 Hz et présentés à un niveau sonore de 65 dB SPL. Cinq conditions de présentation sont testées : 1) monaurale avec présentation des sons sur l'oreille gauche ; 2) monaurale avec présentation des sons sur l'oreille droite ; 3) binaurale avec présentation de deux sons identiques aux deux oreilles ; 4) dichotique avec présentation d'un son pur à gauche et une FM à droite ; 5) dichotique avec une FM présentée sur l'oreille droite et gauche, la phase de la fonction de modulation étant inversée entre les deux oreilles. Une tâche adaptative à trois alternatives en choix forcé est utilisée afin d'estimer le point de la fonction psychométrique correspondant à 79,4 % de réponses correctes. Les résultats indiquent que **les seuils de détection de FM des sujets normo-entendants jeunes sont significativement meilleurs (plus bas) que ceux des sujets normo-entendants d'âge intermédiaire et plus âgés**. Dans l'ensemble, la détection dichotique de FM est supérieure à la détection monaurale quel que soit l'âge des participants. Les données obtenues à partir des trois conditions binaurales sont présentées en figure 46. Les analyses statistiques indiquent que les trois groupes d'âge diffèrent tous significativement et que les participants d'âge intermédiaire et plus âgés présentent des seuils de détection de FM significativement plus élevés (moins bons) que les participants normo-entendants jeunes. Les analyses indiquent également une interaction significative entre le groupe et la condition testée, révélant des différences dans le bénéfice tiré de la transition d'un mode « dichotique » à un mode « dichotique » entre les groupes d'âge. En effet, comme indiqué en figure 46, le groupe de sujets plus âgés montre non seulement de moins bonnes performances dans l'ensemble des conditions mais également une moindre capacité à bénéficier des indices de TFS dichotiques.

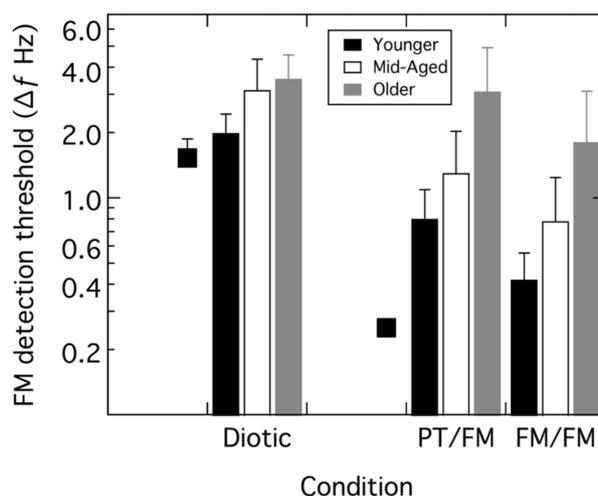


Figure 46. Seuils de détection de FM dans trois conditions binaurales : diotique, dichotique (son pur (PT) vs. FM), dichotique (FM vs. FM). L'ombrage des barres indique le groupe d'âge (cf. encadré central). Les barres d'erreurs indiquent l'écart-type. Les carrés pleins montrent les données de Witton et al. (2000). Extrait de Grose et Mamo (2012).

Dans l'ensemble, les résultats de cette étude indiquent que **les mesures des capacités de détection de FM dichotique semblent plus sensibles au vieillissement que les mesures des capacités de détection de FM monaurales**. Toutefois, il est important de constater que ces mesures dichotiques sont fortement affectées par toute asymétrie dans les seuils absolus entre les deux oreilles, ce qui pourrait limiter l'usage clinique d'une telle méthode.

Plutôt que de mesurer les PEA en réponse à deux AM (oreille droite et gauche) soumises à un seul déphasage de phase interaurale, une technique dite « IPM-FR » ou « Interaural-phase modulation following response » mesure les PEA en réponse à des déphasages périodiques intégrés dans la TFS_p d'une AM sinusoïdale. McAlpine et al. (2016) démontrent que cette technique est plus efficace pour évaluer l'encodage neuronal des différences interaurales de phase car elle produit un plus grand nombre de réponses neuronales sur une période de temps plus courte. La réponse est ensuite évaluée par analyse spectrale dans la bande de fréquence correspondant à la vitesse de changement de phase. A l'aide de cette technique, Koerner et al. (2000) observent une dégradation du traitement temporel sous l'effet du vieillissement : **les réponses neuronales observées des sujets normo-entendants âgés sont plus faibles que celles des sujets normo-entendants jeunes**.

Dans une autre étude psychophysique combinant de nombreuses mesures sensorielles et cognitives, Schoof et Rosen (2014) mesurent les seuils de détection monauraux d'une FM sinusoïdale de cadence lente (2 Hz) portée par un son pur de 1000 Hz et présentée à un niveau de 70 dB SL. Les mesures sont réalisées sur deux groupes de dix-neuf participants jeunes (19-29 ans) et dix-neuf participants âgés (60-72 ans). **Contrairement aux études précédentes, les auteurs n'observent pas de différences significatives entre les seuils de détection de FM des deux groupes**. Cette étude remet donc en cause l'idée que le vieillissement affecterait le traitement temporel auditif et appelle à réplication et extensions. L'implication de facteurs cognitifs dans les effets d'âge est soulevée par cette étude (les sujets âgés présentant un déclin de la mémoire de travail et de la rapidité de traitement). Par ailleurs, la détection de FM pouvant faire appel à des indices d' ENV_w obtenus par conversion de la FM en AM en sortie des filtres cochléaires, il est possible que les sujets âgés changent de stratégie à mesure que leurs capacités de codage temporel se dégradent. Ces aspects sont explorés dans plusieurs études conduites par Wallaert et collaborateurs.

Wallaert et al. (2016) évaluent les effets du vieillissement sur la détection de FM et comparent ces mesures à celles obtenues pour la détection d'AM. Les auteurs mesurent des seuils de détection de FM en présence ou en absence d'AM interférente : ici, l'AM non informative est utilisée afin de limiter la capacité des auditeurs à utiliser les indices de pattern d'excitation issus de la conversion de la FM en AM à la sortie des filtres cochléaires. Les mesures sont réalisées sur quinze sujets normo-entendants jeunes (âge moyen : 24 ans ; étendue : 22-28 ans) et quatorze sujets normo-entendants âgés (âge moyen : 57 ans ; étendue : 47-66 ans). FM (et AM) sont appliquées à une porteuse de 500 Hz et présentées à 40 dB SL. Deux cadences de modulation de 2 et 20 Hz sont testées. Par ailleurs, le nombre de cycles de modulation (N) varie systématiquement entre 2 et 9. L'effet de N permet d'étudier l'intégration temporelle pour la détection de modulation, un phénomène a priori contraint par la nature du mécanisme

décisionnel et des capacités de stockage en mémoire échoïque (Wallaert et al., 2017, 2018). Une tâche détection à deux alternatives en choix forcé estime le seuil correspondant à 70.7 % de réponses correctes sur la fonction psychométrique (Levitt, 1971). Les résultats sont représentés en figure 47.

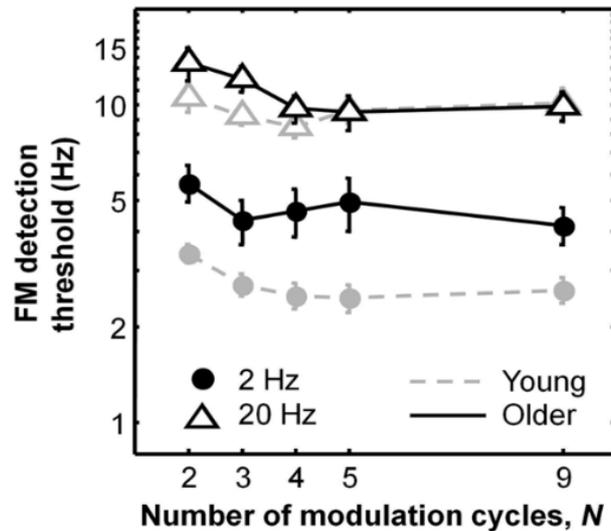


Figure 47. Seuils de détection moyens de FM d'auditeurs normo-entendants jeunes (traits en pointillés) et plus âgés (traits continus) en fonction du nombre de cycles de modulation, N (en abscisses), et de la cadence de modulation (cercles fermés : 2 Hz ; triangles ouverts : 20 Hz). Les seuils sont exprimés en Hz sur une échelle logarithmique. Extrait de Wallaert et al. (2016).

Pour la détection de FM, l'effet du groupe est significatif : dans la plupart des conditions, les seuils de détection de FM sont plus élevés (moins bons) chez les auditeurs âgés (sauf pour $n=5$ ou $n=9$ et pour une cadence de 20 Hz). Les seuils sont significativement plus élevés quelle que soit la cadence de modulation, l'effet délétère du vieillissement étant toutefois plus important à 2 Hz qu'à 20 Hz (d'un facteur 1.7 pour les auditeurs normo-entendants âgés à 2 Hz et d'un facteur 1.4 à 20 Hz). En comparaison, l'effet délétère de l'âge est similaire à 2 et 20 Hz en détection d'AM (Wallaert et al., 2017). Ces résultats sont interprétés comme reflétant un effet du vieillissement sur le mécanisme temporel permettant d'utiliser efficacement les indices de TFS_N dans la détection de FM lente. Pour les deux groupes, le seuil de détection de FM s'améliore à mesure que le nombre de cycles N augmente et ce phénomène d'« intégration temporelle » ne diffère pas significativement entre les groupes. Ce dernier résultat suggère que les mécanismes de prise de décision supposés être responsables du phénomène d'intégration temporelle (appariement de gabarit ou combinaison d'observations multiples) ne sont pas affectés par le vieillissement. Il en va de même pour les mécanismes de stockage des informations sensorielles en mémoire échoïque, impliqués dans la détection des signaux les plus longs (à 2 Hz lorsque $n=9$). Ainsi, dans l'ensemble, cette étude indique que **le vieillissement dégrade sélectivement la sensibilité aux indices de TFS_N utilisés pour la détection de FM lente; toutefois, ce déclin ne semble pas lié à un déclin des traitements centraux.**

Enfin, dans une étude corrélationnelle, Whiteford et al. (2017) mesurent les seuils de détection d'AM et de FM à des cadences lente (2 Hz) et rapide (20 Hz) et ce avec une porteuse de 500 Hz. Deux modes de présentation sont testés : un mode monaural et un mode binaural. Quarante-vingt-cinq adultes âgés entre 20 et 80 ans (dont quinze participants pour chaque décennie entre 20 et 69 ans et dix participants par décennie entre 70 et 79 ans) sont testés. Après avoir contrôlé la sensibilité à l'AM à la même cadence dans la condition diotique, les seuils de détection de FM lente sont corrélés avec l'âge. Ceci n'est pas le cas pour les seuils de détection d'AM mesurés en mode diotique, quelle que soit la cadence. Ces résultats indiquent que **les seuils de détection de FM sont plus affectés par l'effet de l'âge que les seuils de détection d'AM** (cf. figure 48 ci-dessous). L'absence de corrélation significative entre les seuils de détection de FM lente (mesurés en mode dichotique) et l'âge après avoir contrôlé la sensibilité à l'AM lente (mesurée en mode dichotique) pourrait signifier que les performances dans les tâches binaurales sont limitées par des facteurs centraux, opérant durant ou après l'intégration binaurale de l'information temporelle. Ainsi, les résultats de précédentes études révélant des effets du vieillissement sur le traitement de la FM lente et rapide en condition binaurale (Grose & Mamo, 2010) seraient à mettre sur le compte d'une variabilité centrale dans le codage binaural des indices de TFS_N .

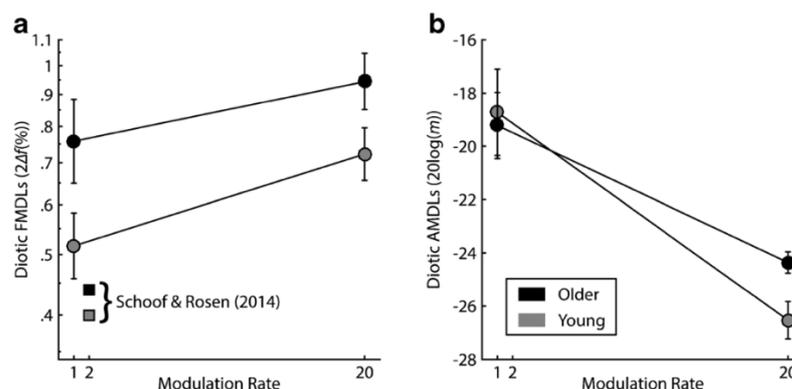


Figure 48. Seuils de détection moyen de FM (figure a) et d'AM (figure b) mesurés en condition diotique pour des sujets normo-entendants jeunes (cercles gris) et normo-entendants plus âgés (cercles noirs) en fonction de la cadence de modulation exprimée en Hertz. Les carrés représentent les résultats obtenus par Schoof et Rosen (2014). Extrait de Whiteford et al. (2017).

En résumé, ces études sur la détection de FM présentent quelques incohérences. Toutefois, pris ensemble, ces travaux indiquent que le vieillissement n'affecte pas la détection d'AM et de FM de la même manière. La sensibilité à la FM est généralement dégradée sous l'effet du vieillissement et ce déficit semble débiter relativement tôt dans la vie. La dégradation de la sensibilité à la FM apparaît plutôt aux cadences lentes et lorsque la porteuse est basse, en accord avec l'idée que la détection de FM lente repose sur le traitement d'indices de TFS_N mais d'autres hypothèses sont envisageables. Le déclin des capacités de détection de FM lente observé chez les sujets âgés est difficilement explicable en termes de facteurs centraux (non-sensoriels) associés au vieillissement. *A contrario*, la dégradation de la sensibilité à l'AM chez les sujets âgés - souvent observée à toutes les cadences - pourrait refléter un déclin général d'efficacité de traitement des indices d' ENV_N . Ces études portant sur le vieillissement pointent donc vers des mécanismes ou voies distinctes pour le traitement de l'AM et de la FM.

I.3.c. Effets de la perte auditive neurosensorielle sur le traitement des indices d'AM et de FM

Une vaste littérature épidémiologique révèle que la sensibilité auditive humaine se dégrade progressivement avec l'âge en démarrant plus rapidement dans les hautes que dans les basses fréquences (Gordon-Salant, 2005). Par exemple, les hommes dans leur neuvième décennie présentent une différence de 40 dB HL entre basses et hautes fréquences (*i.e.*, 250 Hz (30 dB HL) vs. 8000 Hz (70 dB HL) (Gordon-Salant, 2005). Une autre étude évalue les seuils auditifs tonaux d'un millier d'auditeurs adultes âgés entre 20 et 90 ans n'ayant aucun antécédent de traumatisme sonore sur une période de 30 ans chez l'homme et de 17 ans chez la femme. La figure 49 montre les changements longitudinaux des seuils audiométriques durant toute la durée de vie à cinq fréquences audio. Les données indiquent que **le déclin de la sensibilité auditive - et donc, de l'audibilité (la capacité à entendre les sons faibles) - s'accélère au-delà de 20 à 30 ans chez les hommes et au-delà de 50 ans chez la femme.**

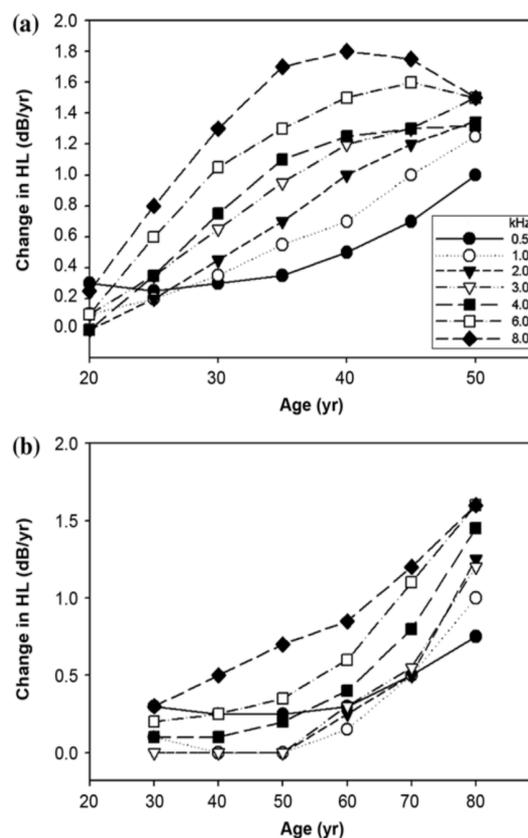


Figure 49. Changements longitudinaux moyens sur 10 ans des seuils audiométriques (dB HL) de sons purs en fonction de la fréquence pour des hommes (figure a) et des femmes (figure b), d'après l'étude longitudinale de Baltimore sur le vieillissement. Extrait de Gordon-Salant (2005).

Cette perte auditive neurosensorielle associée au vieillissement prend le nom de presbycusis. Toutefois, la perte auditive neurosensorielle peut survenir de bien différentes manières (prise de substances ototoxiques, traumatismes sonores, vieillissement, pathologies congénitales ou encore infections) (Henry & Heinz, 2013). D'un point de vue physiologique, ceci se traduit par :

- Une atteinte des *cellules ciliées externes* (le mécanisme actif cochléaire) *et internes* (le mécanisme de transduction neuro-électrique) (Moore, 2014).
- Une atteinte des *neurones du ganglion spiral* constituant le canal unique de relai d'information entre la cochlée et le tronc cérébral (Kujawa & Liberman, 2006 ; 2009).
- Une atteinte de la *strie vasculaire* ayant pour rôle de fournir de l'énergie pour la transduction des cellules ciliées internes et l'amplification du mouvement de la membrane basilaire par les cellules ciliées externes (Schmiedt et al., 2002).

Ces événements sont susceptibles de conduire à une dégradation de la représentation des indices d' ENV_N et de TFS_N dans la périphérie et dans les centres auditifs, à savoir des distorsions supraliminaire de nature temporelle qui en retour pourraient avoir des conséquences perceptives mesurables à l'aide de tâches comportementales.

I.3.c.i. Effets de la perte auditive neurosensorielle sur la détection d'AM

Si certaines études démontrent que la sensibilité à l'AM n'est pas affectée par la perte auditive neurosensorielle et qu'elle peut parfois être meilleure que celle observée chez des auditeurs normo-entendants (Bacon & Gleitman, 1992 ; Moore et al., 1992 ; Moore & Glasberg, 2001 ; Wallaert et al., 2016), d'autres études indiquent que la perte auditive neurosensorielle dégrade la sensibilité monaurale et binaurale de l'AM (King et al., 2014). Ces différences sont probablement à mettre sur le compte de disparités méthodologiques, et plus particulièrement de disparités dans : i) les paramètres expérimentaux utilisés (*type de porteuse* : son pur, bruit à bande étroite ou bruit à bande large; *niveau sonore* utilisé pour comparer les groupes) et ii) le statut auditif des participants (*perte auditive unilatérale ou bilatérale, perte fluctuante (maladie de Ménière) ou stable, perte plate ou en pente*).

Les premières études mesurant les TMTF chez les sujets malentendants avec un bruit large bande comme porteuse indiquent que les auditeurs malentendants sont généralement moins sensibles aux cadences de modulation élevées que les auditeurs normo-entendants (Formby, 1982 ; Lamore et al., 1984 ; Bacon & Viemeister, 1985). Cependant, ce déficit apparent d'acuité temporelle pourrait résulter du fait que les hautes fréquences audio (pour lesquelles les filtres cochléaires - à savoir, ceux susceptibles de transmettre les modulations les plus rapides - sont les plus larges) sont généralement inaudibles pour les auditeurs malentendants (Bacon & Viemeister, 1985). Un meilleur contrôle de la « bande passante audible » (listening

bandwidth) est donc nécessaire. Les études de Bacon et Gleitman (1992) et Moore et al. (1992) discutées ci-dessous prennent en compte ce facteur.

Bacon et Gleitman (1992) mesurent les TMTF de cinq sujets normo-entendants et huit sujets présentant une perte auditive neurosensorielle *plate*. La perte auditive plate garantit un niveau de sensation (SL) relativement constant sur l'ensemble des fréquences audiométriques. Ainsi, les participants peuvent écouter les signaux dans une bande passante assez large. Les capacités de détection d'AM sont mesurées à l'aide d'une tâche adaptative à deux alternatives en choix forcé estimant le seuil correspondant à 70,7 % de réponses correctes sur la fonction psychométrique (Levitt, 1971). La porteuse est un bruit large-bande modulé sinusoïdalement en amplitude dans un des deux intervalles, dans l'autre elle n'est pas modulée. L'utilisation d'une porteuse large bande limite l'usage d'indices spectraux. Le niveau de la porteuse varie entre 10 et 50 dB SPL, et pour un niveau de porteuse donné, la cadence de modulation varie entre 2 et 1024 Hz. Les résultats indiquent que lorsque les niveaux estimés en dB SPL sont identiques entre les deux groupes de participants testés, les TMTF sont semblables pour les deux groupes d'auditeurs testés tandis que ***lorsque les niveaux de sensation (dB SL) sont égalisés entre les deux groupes (par exemple, à 20 dB SL), les seuils de détection d'AM des sujets malentendants sont meilleurs que ceux des sujets normo-entendants et ce quelle que soit la cadence de modulation.***

Moore et al. (1992) contrôlent explicitement la « bande passante d'écoute » de leurs auditeurs en mesurant les TMTF sur des sujets présentant une perte neurosensorielle avec un bruit centré sur 2 kHz et dont la largeur de bande est d'une octave. Sur la gamme de fréquences couverte par le bruit faisant office de porteuse, les sujets normo-entendants et malentendants présentent des seuils audiométriques constants. Ainsi, les tâches sont égalisées en termes d'audibilité. De plus, afin de s'assurer que les sujets n'utilisent pas d'information dans des zones fréquentielles situées en dehors de la bande passante nominale du bruit, la porteuse modulée est présentée dans un bruit de fond large bande non modulé. ***Les résultats indiquent qu'à niveaux SPL et SL identiques, les seuils de détection d'AM sont similaires entre les participants normo-entendants jeunes et les participants malentendants présentant une perte unilatérale.***

Moore et Glasberg (2001) comparent les TMTF de quatre sujets normo-entendants jeunes âgés entre 23 et 54 ans et de trois auditeurs malentendants âgés entre 70 et 84 ans. Les TMTF sont estimées avec une porteuse sinusoïdale grâce une tâche à trois alternatives en choix forcé estimant le seuil correspondant à 79,4 % de réponses correctes sur la fonction psychométrique. Quatre fréquences sont utilisées pour la porteuse sinusoïdale : 1000, 2000, ou 5000 Hz. Les cadences de modulation sont comprises entre 10 et 640 Hz. Les participants normo-entendants jeunes sont testés à un niveau de 30 et 80 dB SPL tandis que les participants malentendants sont testés uniquement à 80 dB SPL. Les résultats indiquent que : ***i) les sujets malentendants présentent de meilleurs seuils de détection d'AM que les sujets***

normo-entendants jeunes quand les stimuli sont présentés à un même niveau SL. Ces résultats s'accordent avec les études de Jerger et al. (1959), Buus et al. (1982) et de Moore, 1995 qui indiquent qu'à niveau de sensation identique, la résolution d'amplitude est meilleure chez les sujets malentendants ; ii) **la fréquence de coupure des TMTF est similaire entre les deux groupes** suggérant que la résolution temporelle n'est pas affectée par la perte auditive ; iii) **une absence d'amélioration (diminution) des seuils de détection d'AM aux cadences de modulation rapides chez les auditeurs malentendants**, résultant de la perte de sélectivité fréquentielle cochléaire et de la dégradation de la résolution des bandes spectrales latérales. L'amélioration paradoxale de sensibilité à l'AM observée chez les auditeurs malentendants est interprétée comme résultant d'une magnification perceptive des indices d' ENV_N évoqués par l'AM. Cette magnification perceptive résulterait de la perte de la compression instantanée d'amplitude faisant suite à la perte du mécanisme actif cochléaire.

Füllgrabe et al. (2003) évaluent les seuils de détection d'AM sur neuf sujets normo-entendants jeunes (d'âge moyen égale à 22 ans) et sept sujets malentendants (d'âge moyen égale à 66 ans) présentant une perte de 40 à 45 dB HL à 2 kHz. Les auteurs utilisent une porteuse sinusoïdale de 2 kHz. La procédure employée est une tâche adaptative à deux alternatives en choix forcé estimant le seuil correspondant à 70,7 % de réponses correctes. Les cadences de modulation sont comprises entre 4 et 87 Hz. **Les résultats indiquent que les participants malentendants sont en moyenne meilleurs que les sujets normo-entendants jeunes pour les deux cadences les plus lentes (4 et 16 Hz) lorsqu'ils sont testés à un niveau de sensation similaire (en dB SL).**

Wallaert et al. (2017, 2018) évaluent les effets de la perte auditive neurosensorielle sur la sensibilité et l'intégration temporelle de l'AM en mesurant les seuils de détection d'AM à un niveau de 40 dB SL chez des auditeurs normo-entendants jeunes et âgés et des auditeurs âgés présentant une perte auditive neurosensorielle légère à modérée. Les auteurs utilisent une porteuse sinusoïdale de 500 Hz et des cadences de modulation de 2 et 20 Hz. Le nombre de cycles de modulation, N, varie entre 2 et 9. Une tâche adaptative à deux alternatives en choix forcé estime le seuil correspondant à 70,7 % de réponses correctes (Levitt, 1971). Les résultats présentés en figure 50 indiquent que : (i) l'âge (en soi) dégrade la sensibilité à l'AM mais n'affecte pas l'intégration temporelle des indices d'AM ; (ii) **la perte auditive neurosensorielle améliore la sensibilité à l'AM (gain de 8 dB) ainsi que l'intégration temporelle (de 3 à 4 dB).** Des simulations du modèle à banc de filtres de modulation utilisant une stratégie de décision basée sur un appariement de gabarit sont comparées aux données empiriques. Ce modèle limite la détection d'AM grâce à trois bruits internes en sortie des filtres de modulation (un bruit additif, un bruit multiplicatif, et un bruit additif mnésique utilisant une constante de temps d'environ 1 s). Les simulations obtenues en éliminant la compression instantanée du modèle et en élargissant les filtres cochléaires par un facteur 2 suggèrent que **l'amélioration de la sensibilité et de l'intégration temporelle de l'AM chez les auditeurs malentendants résulte principalement de la diminution voire de la perte de la compression d'amplitude**

produite par le mécanisme actif cochléaire. Les résultats suggèrent également que la perte auditive neurosensorielle est associée à une augmentation substantielle de bruit interne **additif**. Nous reviendrons plus loin sur ce résultat important des simulations informatiques réalisées par Wallaert et al. (2017, 2018).

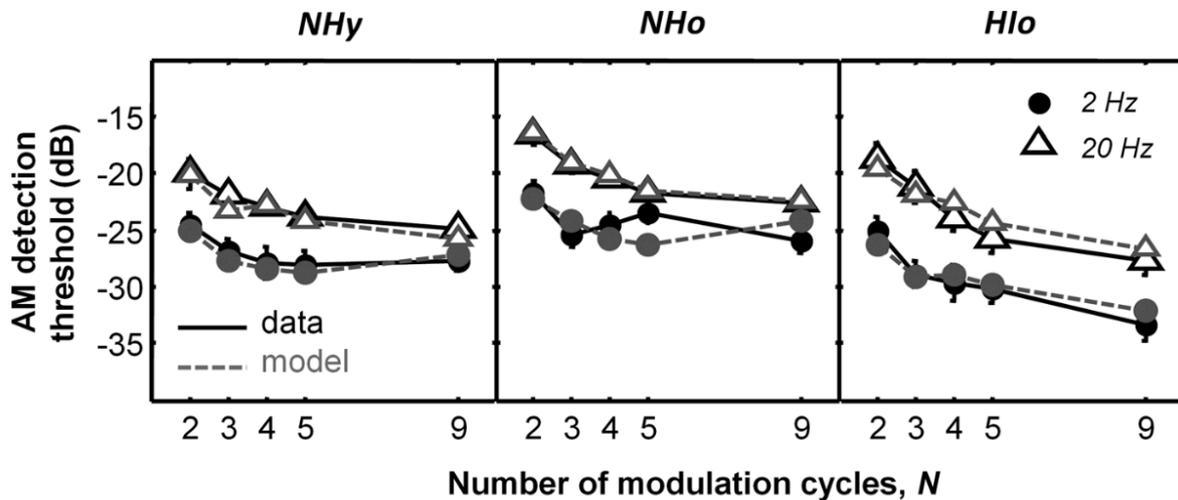


Figure 50. Seuils de détection d'AM moyens (symboles noirs et lignes continues) pour le groupe de participants normo-entendants jeunes (à gauche), normo-entendants âgés (au centre) et malentendants âgés (à droite). Les seuils de détection d'AM simulés par le modèle du banc de filtres de modulation (symboles gris et lignes pointillées) sont présentés pour chaque groupe d'auditeur. Les cercles indiquent les seuils de détection mesurés pour une cadence lente (2 Hz) tandis que les triangles indiquent les seuils de détection mesurés pour une cadence rapide (20 Hz). Extrait de Wallaert et al. (2017).

L'étude de Paraouty et al. (2016) complète ces investigations sur la détection de modulation en mesurant les *effets d'interférence entre AM et FM* ainsi que les effets du vieillissement et de la perte auditive sur cet effet d'interférence. Ici, les capacités de détection d'AM sont mesurées (en présence et en l'absence d'une FM de même cadence que l'AM cible) à l'aide d'une tâche adaptative à deux alternatives en choix forcé estimant le seuil correspondant à 70,7 % de réponses correctes pour une porteuse sinusoïdale de 500 Hz et une cadence lente de 5 Hz sur trois groupes d'auditeurs : dix sujets normo-entendants jeunes (19-40 ans); dix sujets normo-entendants plus âgés (46-67 ans) et dix sujets malentendants âgés entre 46 et 64 ans présentant une perte auditive bilatérale légère à modérée. Tous les sujets sont testés à un niveau de 40 dB SL. **Les résultats indiquent que l'effet d'interférence de la FM sur la détection d'AM est comparable pour les trois groupes de participants testés, suggérant ainsi que ni l'âge ni la perte auditive n'affectent les mécanismes responsables de l'interférence perceptive entre AM et FM. Enfin, les seuils de détection d'AM des participants malentendants sont similaires à ceux des participants normo-entendants jeunes et plus âgés.** Les auteurs suggèrent que l'absence d'amélioration de la sensibilité à l'AM observée chez les sujets malentendants est liée au fait que les participants de cette étude présentent une perte auditive relativement légère. L'effet d'interférence, quant à lui, est expliqué en

termes de conversion de la FM en AM à la sortie des filtres cochléaires. Une modélisation informatique corrobore cette interprétation.

Dans l'ensemble, les résultats indiquent que la sensibilité à l'AM est soit préservée, soit améliorée par une perte auditive neurosensorielle. La bande passante audible et le niveau de présentation des stimuli jouent un rôle important dans la comparaison des performances entre auditeurs normo-entendants et malentendants. L'amélioration de la sensibilité à l'AM est généralement expliquée en termes de recrutement de sonie, à savoir en raison d'une magnification perceptive des indices d' ENV_N résultant de la perte du mécanisme actif cochléaire. Des données de simulation suggèrent toutefois qu'une augmentation du bruit interne est nécessaire afin de compenser les effets bénéfiques du recrutement de sonie. Prises ensemble, ces données révèlent que si des lésions cochléaires ne modifient pas voire rehaussent la sensibilité à l'AM, cela ne signifie pas pour autant que la représentation auditive interne des indices d' ENV_N évoqués par l'AM est normale dans le cas d'une perte neurosensorielle : ces indices d' ENV_N sont susceptibles d'être distordus sous l'effet de modifications périphériques et centrales.

I.3.c.ii. Effets de la perte auditive neurosensorielle sur la détection de FM

De nombreuses études suggèrent que le traitement des indices monauraux et binauraux de TFS_N décline sous l'effet de la perte auditive et du vieillissement (Moore & Skrodzka, 2002 ; Lacher-fougère & Demany, 2005 ; Grose & Mamo, 2010 ; Hopkins & Moore, 2011 ; King et al., 2014). Ceci pourrait expliquer pourquoi la détection monaurale de FM se dégrade généralement sous l'effet du vieillissement (He et al., 2007 ; Grose & Mamo, 2012) ou de la perte auditive (Lacher-Fougère & Demany, 1998 ; Moore & Skrodzka, 2002 ; Buss et al., 2004 ; Lorenzi et al., 2006 ; Strelcyk & Dau, 2009).

Zurek et Formby (1981) mesurent les seuils de détection de FM à l'aide d'une tâche adaptative à deux alternatives en choix forcé sur huit auditeurs présentant une perte auditive neurosensorielle. Les auteurs utilisent une cadence de modulation de 3 Hz et des fréquences porteuses comprises entre 125 et 4000 Hz. Les sujets sont testés à un niveau de 25 dB SL. Les résultats indiquent que : **les seuils de détection de FM se dégradent à partir d'un certain niveau de perte auditive et ce de manière plus importante dans les basses fréquences audio.** Ce résultat est en accord avec les études indiquant une détérioration générale de la discrimination fréquentielle chez des sujets présentant une perte auditive modérée. D'après les auteurs, le déficit en basse fréquence serait lié à un trouble d'encodage des indices de TFS_N provenant : i) d'une atteinte plus importante du mécanisme temporel (verrouillage en phase sur la TFS_{BM}) supposé agir dans les basses fréquences que du mécanisme spatial/tonotopique ; ou ii) d'un dommage des cellules ciliées accordées sur les basses fréquences invisibles à

l'audiogramme et conduisant à ce que les basses fréquences soient détectées au niveau de la queue (tail) de la courbe d'accord des filtres cochléaires.

En accord avec les travaux de Zurek et Formby (1981), Moore et Glasberg (1986) constatent que **la perte auditive neurosensorielle a un effet délétère sur la détection de FM lente** (pour une cadence de 4 Hz). Cet effet délétère **augmente à mesure que la fréquence porteuse diminue** (de 2000 Hz à 500 Hz). Lacher-Fougère et Demany (1998) mesurent les seuils de détection de FM à l'aide d'une tâche adaptative sur cinq auditeurs normo-entendants jeunes et de sept auditeurs malentendants âgés (présentant une perte de 30 à 75 dB à 500 Hz). Les auteurs utilisent une porteuse sinusoïdale de 500 Hz et une cadence de modulation de 2 Hz (pour laquelle le mécanisme de verrouillage en phase est supposé être dominant) ou 10 Hz (pour laquelle le mécanisme de codage spatial est supposé être dominant). Étant donné que la sélectivité fréquentielle est généralement dégradée chez les auditeurs malentendants, *les seuils de détection de FM devraient être anormalement élevés lorsque le mécanisme de détection le plus efficace impliqué dans la détection est le mécanisme spatial* (donc pour une cadence de modulation lente de 10 Hz). Les résultats confirment que **la perte auditive dégrade la détection de FM quand elle excède 50 dB HL à la fréquence testée (500 Hz)**. Toutefois, les seuils de détection ne sont pas affectés par la cadence de modulation. Cela suggère que la perte auditive affecte les mécanismes spatiaux et temporels de manière équivalente.

Moore et Skrozdka (2002) réalisent une étude complémentaire portant sur la détection de FM *masquée* chez des auditeurs malentendants. Les seuils de détection de FM sont mesurés sur trois sujets normo-entendants jeunes et quatre sujets malentendants âgés présentant une perte neurosensorielle modérée en utilisant une tâche à deux alternatives en choix forcé. La porteuse est de 0,25, 0,5, 1, 2,4 ou 6 kHz et cette dernière est modulée à des cadences de 2, 5, 10 ou 20 Hz. Les auditeurs malentendants sont testés à 85 dB SPL tandis que les sujets normo-entendants jeunes sont testés à 70 dB SPL. Les seuils de détection de FM sont mesurés en présence d'AM (de profondeur de modulation fixée à 0,33) ou en l'absence d'AM. Comme indiqué plus haut, cette dernière est utilisée afin de limiter la capacité des auditeurs à utiliser des indices de reconversion de FM en AM. Les résultats présentés en figure 51 indiquent que **les seuils de détection de FM sont significativement moins bons chez les participants malentendants que chez les participants normo-entendants**. D'autre part, ces résultats indiquent que **l'effet interférant de l'AM est plus important chez les auditeurs malentendants** (et ce quelle que soit la fréquence porteuse et la cadence de modulation). L'effet d'interférence est approximativement constant chez les auditeurs malentendants. Un profil différent est obtenu chez les auditeurs normo-entendants : dans ce cas, l'effet délétère de l'AM sur les seuils de détection d'une FM est observé dans toutes les conditions à l'exception du cas où une FM lente est appliquée à une porteuse basse fréquence, conformément à l'hypothèse d'un mécanisme purement temporel (Moore & Sek, 1996). Chez les auditeurs malentendants, la dégradation globale des seuils de détection de FM par l'ajout

d'une AM dans l'ensemble des conditions suggère que ces derniers n'utilisent pas efficacement l'information temporelle de TFS_N . Les auteurs proposent donc que *les seuils de détection de FM des auditeurs malentendants sont essentiellement basés sur les indices de pattern d'excitation estimés à partir de filtres cochléaires élargis sous l'effet de la perte neurosensorielle* (reconversion d'indices de FM en AM) et que ces indices sont - de fait - fortement dégradés par l'ajout de l'AM interférente. L'étude de modélisation de Paraouty et al. (2016) portant sur l'interférence perceptive entre AM et FM confirme cette interprétation.

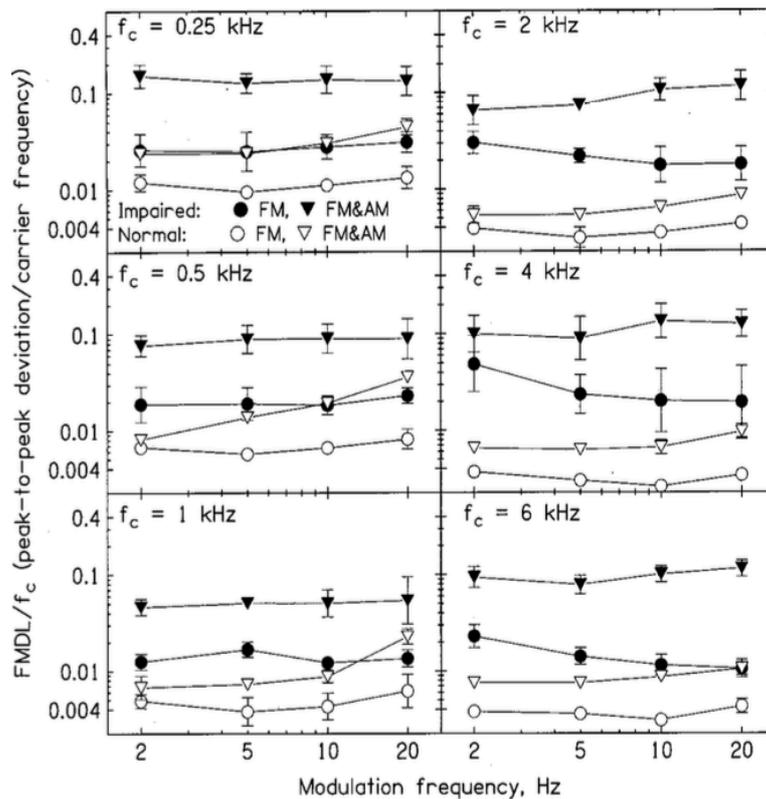


Figure 51. Seuils de détection de FM moyens en fonction de la cadence de modulation. Chaque figure montre les résultats obtenus pour une fréquence porteuse spécifique. Les cercles indiquent les seuils de détection de FM seule ; les triangles indiquent les seuils de détection de FM mesurés en présence d'AM. Les symboles ouverts indiquent les seuils de détection des auditeurs normo-entendants jeunes. Les symboles fermés indiquent les seuils de détection des auditeurs malentendants âgés. Les barres d'erreur indiquent l'écart-type. Extrait de Moore et Skrodzka (2002).

Quelques études complètent ce panorama. Buss et al. (2004) testent quatorze auditeurs normo-entendants jeunes et douze auditeurs présentant une perte auditive neurosensorielle dans une tâche de détection de FM lente (2 Hz) appliquée à une porteuse de 500 et 1000 Hz. Conformément aux études précédentes, les résultats indiquent que les seuils de détection de FM sont dégradés de manière significative aux deux fréquences porteuses chez les auditeurs malentendants, l'effet de la fréquence porteuse étant similaire dans les deux groupes. Par la suite, Kortlang et al. (2016) mesurent les seuils de détection d'une FM lente et aléatoire (dont la cadence de modulation varie entre 1 et 4 Hz) dans le silence ou dans un bruit masquant (bruit Gaussien à bande étroite filtré entre 1 et 4 Hz) en présence ou en l'absence d'une AM aléatoire (utilisée afin de limiter la détectabilité des indices d'AM provenant de la

reconversion FM-AM). Les seuils sont mesurés sur six auditeurs normo-entendants jeunes (25-30 ans), six auditeurs normo-entendants plus âgés (65-77 ans) et six auditeurs malentendants âgés (60-82 ans). Une tâche adaptative à trois alternatives en choix forcé estime le seuil correspondant à 70,7 % de réponses correctes. Les signaux sont présentés à un niveau de 65 dB HL pour les sujets normo-entendants jeunes et de 80 dB HL pour les sujets malentendants. Les résultats de détection de FM individuels et moyens sont présentés en fonction des conditions expérimentales en figure 52. Les seuils de détection de FM se dégradent en présence d'une AM et à mesure que le niveau du bruit masquant augmente. Dans l'ensemble, **les auditeurs malentendants et normo-entendants âgés présentent des seuils significativement dégradés par rapport aux auditeurs normo-entendants jeunes et ce quelle que soit la condition expérimentale testée.**

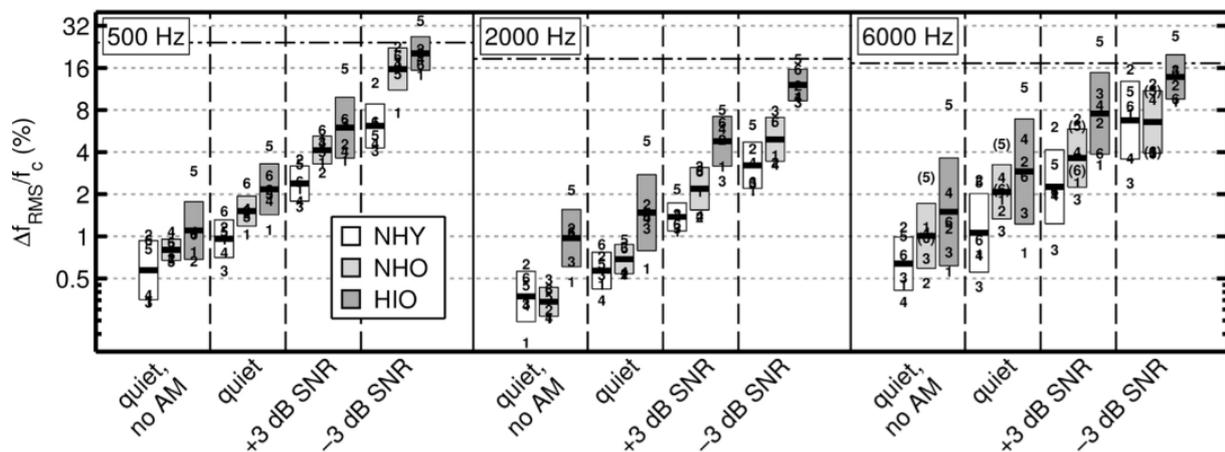


Figure 52. Seuils de détection de FM aléatoire en fonction de la condition expérimentale testée (moyennes et écart-types géométriques). Chaque figure présente les données obtenues pour une fréquence porteuse donnée (500, 2000, 6000 Hz). NHy : sujets normo-entendants jeunes ; NHO : sujets normo-entendants âgés ; HIO : sujets malentendants âgés. Extrait de Kortlang et al. (2016).

En résumé, ces études s'accordent toutes sur le fait que les seuils de détection de FM se dégradent significativement - et parfois substantiellement - sous l'effet d'une perte auditive neurosensorielle. Les résultats suggèrent que les mécanismes de traitement purement temporels ou spatio-temporels de la FM sont affectés de manière équivalente. La dégradation de ces mécanismes résulte probablement : i) de l'élargissement des filtres cochléaires causé par la perte du mécanisme actif cochléaire et/ou ii) d'une réduction de l'efficacité du mécanisme d'encodage et/ou de lecture des indices de TFS_N . Plusieurs études (e.g., Buss et al., 2004b ; Lopez-Poveda et al., 2017) notent que la sensibilité à une FM lente portée par un son pur basse fréquence est un bon prédicteur de l'intelligibilité de la parole filtrée ou masquée chez les auditeurs malentendants, confirmant l'importance des indices de FM dans la reconnaissance de la parole.

I.3.d. Origines probables des troubles de codage temporel

Notre revue de littérature met en lumière des troubles auditifs temporels associés au vieillissement et à la perte neurosensorielle. Comme discuté plus haut, ces troubles sont susceptibles de refléter - pour partie - des troubles de codage sensoriel des indices d' ENV_N et de TFS_N à un niveau relativement périphérique (nerf auditif et tronc cérébral bas). Toutefois, force est de constater que d'autres hypothèses sont envisageables (trouble cognitif, perte d'efficacité de traitement, ...). Plusieurs pathologies peuvent conduire à des modifications du codage sensoriel des indices d' ENV_N et de TFS_N , et ce parfois même lorsque les seuils audiométriques sont normaux ou quasi normaux. Ces hypothèses sont présentées ci-dessous.

I.3.d.i. Réduction du nombre de synapses ou de neurones

Pendant longtemps, les études sur la perte auditive neurosensorielle considèrent que les cellules ciliées externes sont les éléments les plus vulnérables (dans le cas où une élévation des seuils à l'audiogramme clinique est observée) et dans la majorité des cas, les fibres auditives nerveuses sont supposées dégénérer bien longtemps après (Lieberman et al., 2017). Très récemment, des études animales et humaines révèlent que dans le cas d'une perte auditive neurosensorielle (acquise suite à une exposition au bruit, au vieillissement ou à la prise de substances ototoxiques), les connexions synaptiques périphériques sont les éléments les plus vulnérables et les synapses du nerf auditif peuvent être détruites même dans le cas où les cellules ciliées survivent (Stamatakis et al., 2006 ; Sergeyenko et al., 2013 ; Lieberman, 2017). Chez l'homme, une étude de l'os temporal post-mortem de 20 sujets normo-entendants âgés entre 0 et 89 ans révèle une **perte massive de plus de 60% d'axones périphériques** chez un certain nombre de sujets âgés de plus de 50 ans (Wu et al., 2019). Les données de cette étude sont présentées en figure 53 et indiquent que la perte de fibres auditives nerveuses liée à l'âge est presque trois fois plus importante que celle des cellules ciliées internes (perte rarement supérieure à 15 % quel que soit l'âge). Il y a donc une augmentation régulière du degré de désafférentation des cellules ciliées survivantes sous l'effet du vieillissement. En d'autres termes, un grand nombre de neurones auditifs de l'oreille interne âgée sont déconnectés de leurs cellules ciliées internes. Or, quand une synapse est perdue, sa fibre auditive associée devient silencieuse (chaque fibre étant connectée à une cellule ciliée interne) et **l'amplitude de la réponse neuronale diminue de manière proportionnelle à la quantité de synapses perdues** (Sergeyenko et al., 2013). Il est souvent envisagé que les fibres auditives affectées par cette synaptopathie cochléaire soient *des fibres à seuil haut et à taux de décharge spontané bas* : Furman et al. (2013) démontrent que la

synaptopathie cochléaire induite par sur-exposition au bruit chez des rongeurs âgés est sélective pour ce type de fibres auditives qui représentent 40 % de la population totale des fibres auditives nerveuses. Schmiedt et al. (1996) notent également une réduction de l'activité de cette catégorie de fibres auditives chez des chats âgés et celles-ci seraient également les plus vulnérables à une forte exposition au bruit chez l'homme (Liberman, 1978).

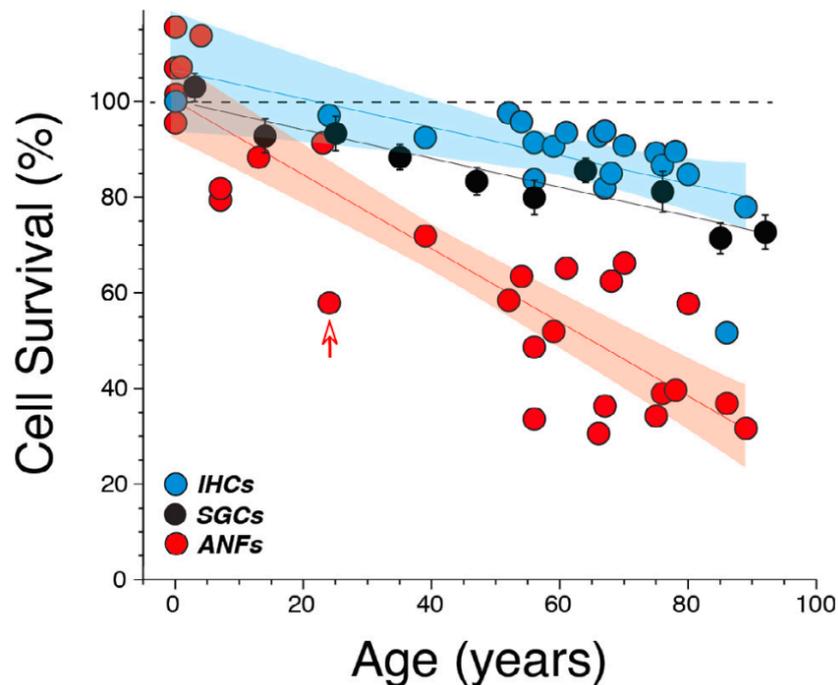


Figure 53. Pourcentage de cellules survivantes en fonction de l'âge (étude post-mortem chez l'humain). Les cercles bleus indiquent les cellules ciliées internes. Les cercles noirs indiquent les cellules du ganglion spiral. Les cercles rouges indiquent les axones périphériques des fibres auditives nerveuses. Cette figure montre que la perte d'axones périphériques liée au vieillissement est plus prononcée que celles des cellules du ganglion spiral et des cellules ciliées internes. Chaque point représente la survie cellulaire dans chaque cas, moyenné sur tous les points d'échantillonnage aux fréquences audiométriques comprises entre 0,25 et 8 kHz et normalisés en pourcentage de survie. La flèche rouge indique le seul cas exposé à une ototoxine connue (cisplatine). Les limites à 95 % de confiance sont indiquées pour les régressions des données de cellules ciliées internes et des axones des fibres auditives nerveuses. Extrait de Wu et al. (2019).

Or ces fibres auditives périphériques présentent des gammes dynamiques plus grandes que celles des fibres à taux de décharge spontané haut, une meilleure aptitude à préserver l'information temporelle et plus spécifiquement d'AM que les autres catégories de fibres et sont aussi plus robustes au bruit (Schmiedt et al., 1996 ; Bharadwaj et al., 2014). Tout ceci conduit donc à penser que la perte d'unités neurales pourrait dégrader l'extraction et la représentation du signal dans les voies auditives centrales, contribuant de manière significative au déclin régulier des capacités auditives (Wu et al., 2019). **Le vieillissement, est donc associé à une désafférentation neuronale consécutive à la « synaptopathie cochléaire » aussi dénommée « perte auditive cachée »** ; l'ensemble de ces éléments pourrait contribuer à des difficultés de compréhension de la parole dans le bruit apparaissant

à un âge intermédiaire (Grose et al., 2006) et plus avancé tout au long de la vie (Snell & Frisina, 2000). Ceci pourrait également contribuer à des troubles de localisation spatiale nécessitant l'utilisation d'indices binauraux (Noble et al., 1994) malgré des seuils auditifs normaux, la perte de synapses ne provoquant pas d'élévation de seuils électrophysiologiques ou audiométriques à moins qu'elle ne soit complète à 80-90 % (Lobarinas et al., 2016).

D'après Bharadwaj et al. (2014), **la perte de fibres auditives nerveuses à seuil haut et à taux de décharge spontané bas est susceptible d'affecter plus particulièrement le codage des indices d'ENV_N** véhiculés par les sons complexes comme la parole. Dans une cochlée saine, la convergence de multiples entrées stochastiques produites par ces fibres auditives périphériques conduit à l'amélioration de la précision temporelle dans le pattern de décharges neurales de plusieurs types de cellules dans les noyaux cochléaires (Joris et al., 1994 ; Oertel et al., 2000). Ainsi, il est probable qu'une réduction du nombre de fibres auditives dégrade la précision avec laquelle les indices d'ENV_N (mais également ceux de TFS_N) sont transmis aux structures auditives centrales (Lopez-Poveda & Barrios, 2013). Cette prédiction est vérifiée par une étude de simulation informatique menée par Goodman et al. (2018). Les auteurs simulent les réponses du noyau cochléaire antéroventral à des signaux d'AM à l'aide d'un modèle de neurone formel de type « leaky integrate and fire » recevant un nombre variable d'entrées périphériques. Comme indiqué dans la figure 54 ci-dessous, la réduction du nombre d'entrées - simulant ainsi l'effet d'une synaptopathie cochléaire - altère drastiquement la TMTF simulée de ces neurones de type « chopper » (sustained et transient) du noyau cochléaire. **Les effets d'une synaptopathie cochléaire sur le codage temporel pourraient donc se manifester à un niveau central (ici, le tronc cérébral bas).**

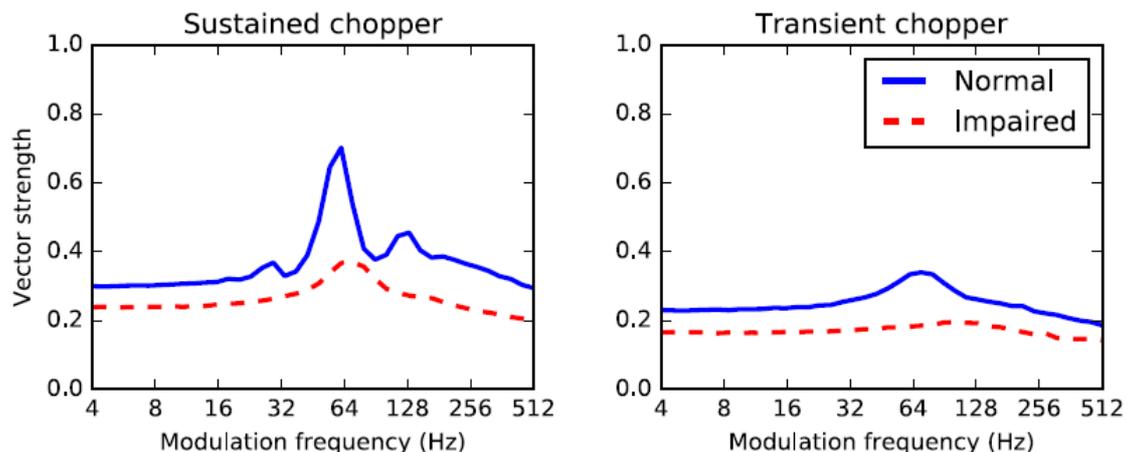


Figure 54. TMTF simulée pour un modèle mathématique de neurone du noyau cochléaire antéroventral. Chaque figure représente l'indice de synchronisation en fonction de la cadence de modulation d'un son pur modulé sinusoidalement en amplitude. La figure de gauche représente les résultats de simulation d'un modèle de neurone de type « Sustained Chopper ». La figure de droite représente les résultats de simulation d'un modèle de neurone de type « Transient Chopper ». La courbe continue bleue montre les résultats de simulation lorsque le nombre d'entrées (i.e., le nombre de fibres auditives périphériques convergeant sur le modèle de neurone Chopper) est égal à 50. La courbe en pointillés rouges montre les résultats de simulation lorsque le nombre d'entrées (i.e., le nombre de fibres auditives périphériques) est réduite à 10. Extrait de Goodman et al. (2018).

Parallèlement à cela, des études neurophysiologiques réalisées sur des animaux non humains suggèrent qu'à un niveau périphérique (*i.e.*, dans les fibres afférentes de type I), une surexposition au bruit n'affecte pas le mécanisme de verrouillage en phase sur les indices de TFS_N dans le silence tandis qu'il améliore - paradoxalement - le codage des indices d' ENV_N (Miller et al., 1997 ; Kale & Heinz, 2010 ; 2012 ; Henry et al., 2014). Ce point est détaillé plus bas et plusieurs explications sont apportées à cela. *A contrario*, le codage des indices de TFS_N par les fibres auditives périphériques serait réduit dans des environnements bruyants (Henry & Heinz, 2013) et ce essentiellement en raison de l'élargissement des filtres cochléaires et de la dégradation du rapport signal-sur-bruit à la sortie des filtres, et altéré pour le traitement de sons complexes comme, par exemple, des voyelles (Miller et al., 1997). Toutefois, plusieurs études dont celle réalisée par Harrison et Evans (1979) ne démontrent pas d'altération du mécanisme de verrouillage en phase des fibres périphériques sur les indices de TFS_N à la suite d'une atteinte des cellules ciliées externes produite par la prise de substances ototoxiques.

I.3.d.ii. Altération du mécanisme de verrouillage en phase

A un niveau périphérique (*i.e.*, au niveau des fibres auditives afférentes de type I), le mécanisme de verrouillage en phase neuronal sur les indices d' ENV_N et de TFS_N est altéré à des degrés divers par un trauma acoustique ou par une substance ototoxique. Plusieurs études importantes menées par Heinz et collaborateurs fournissent une caractérisation détaillée de ces altérations et des explications précises des causes de ces troubles du codage temporel.

Kale et Heinz (2010 ; 2012) évaluent les effets d'une perte auditive neurosensorielle induite par un traumatisme sonore irréversible provoqué volontairement par les expérimentateurs sur le codage temporel périphérique. Les auteurs utilisent ici des sons modulés et des bandes de bruits véhiculant des indices d' ENV_N et de TFS_N et enregistrent des trains d'impulsion nerveuses dans les fibres auditives périphériques de rongeurs (chinchillas). La figure 55 ci-dessous compare le codage temporel de l' ENV_N entre une population de fibres auditives de rongeurs sains et une population de fibres auditives de rongeurs lésés (Kale & Heinz, 2010). Trois métriques différentes sont utilisées afin de quantifier le verrouillage en phase sur l' ENV_N en réponse à des sons purs modulés en amplitude ou des formants. Chaque point indique la valeur maximale de chaque métrique calculée. ***Ces représentations n'indiquent pas de dégradation du codage de l' ENV_N par les fibres auditives périphériques lésées (lignes rouges) mais a contrario un rehaussement du codage de l' ENV_N*** (valeurs minimales et maximales plus élevées pour les trois métriques) et ce dans la région fréquentielle du trauma acoustique (entre 1 et 4 kHz). Ce résultat électrophysiologique est conforme avec l'hypothèse formulée par Moore (2014) selon laquelle la perte du mécanisme actif et de la compression d'amplitude instantanée produite par la contraction des cellules ciliées externes peut conduire à une amplification de la représentation interne de la profondeur de modulation du signal AM. Toutefois, d'autres explications sont à considérer, comme par exemple une augmentation de

la réponse transitoire des fibres nerveuses résultant d'une atteinte des cellules ciliées internes (Scheidt et al., 2010).

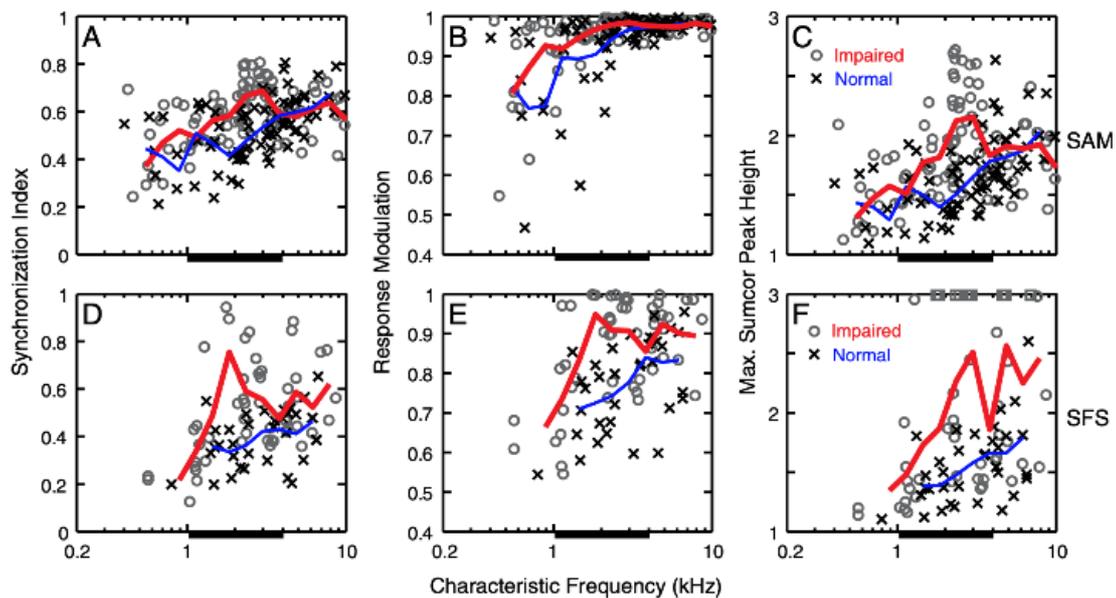


Figure 55. Quantification du codage temporel des indices d'ENV_N en fonction de la fréquence caractéristique de la fibre nerveuse périphérique à l'aide de trois métriques distinctes (une métrique par colonne : Index de synchronisation ; Response modulation ; Max Sumcorr Peak Height). Les données sont représentées pour des fibres auditives périphériques saines (croix) et lésées (cercles) en réponse à des sons purs modulés en amplitude (figures A, B, C) et en réponse à des formants (figures D, E, F). Les lignes pleines bleues (fibres saines) et rouges (fibres lésées) représentent les moyennes pondérées sur l'ensemble de la population neuronale. La barre horizontale indique la région (1-4 kHz) pour laquelle les seuils des fibres sont significativement élevés. Extrait de Kale et Heinz (2010).

Des patterns similaires sont obtenus entre les fibres auditives périphériques saines et lésées (voir figure 56A) lorsque l'amplitude du verrouillage en phase neural est estimée pour les indices de TFS_N (Kale & Heinz, 2010). Ainsi, le mécanisme de verrouillage en phase sur la structure temporelle fine n'est pas affecté pour les fibres auditives individuelles soumises à un traumatisme sonore permanent.

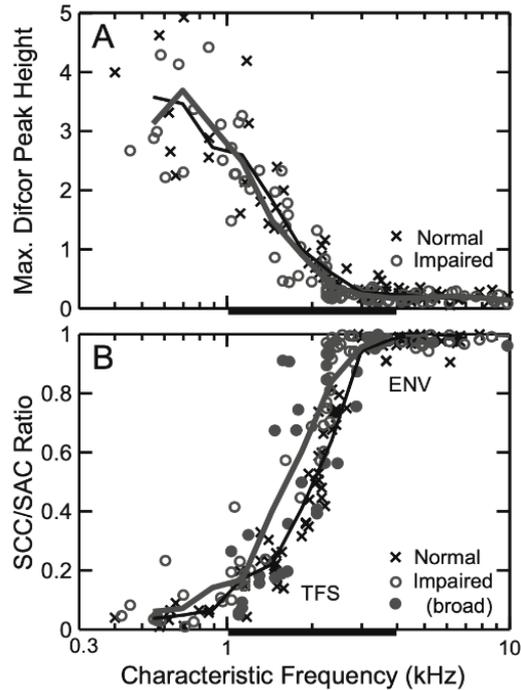


Figure 56. Amplitude du verrouillage en phase des réponses neuronales sur la TFS_N ou l' ENV_N d'un son pour des fibres auditives périphériques saines (croix) ou lésées (cercles ouverts). La figure A montre l'amplitude du verrouillage en phase sur la TFS_N (Max Diffcorr Peak Height) en réponse à des sons purs modulés en amplitude en fonction de la fréquence caractéristique. La figure B quantifie le verrouillage en phase sur la TFS_N et l' ENV_N à l'aide d'une métrique distincte (SCC/SAC Ratio) pour laquelle une valeur proche de 0 indique que le verrouillage en phase sur la TFS_N est prédominant tandis qu'une valeur proche de 1 indique que le verrouillage en phase sur l' ENV_N est prédominant. Les cercles pleins représentent des fibres auditives périphériques dont la courbe d'accord est élargie par rapport à la normale (et dont la sélectivité fréquentielle est dégradée). La barre horizontale indique la région (1-4 kHz) pour laquelle les seuils des fibres sont significativement élevés. Extrait de Kale et Heinz (2010).

Cependant, bien que cette capacité de verrouillage en phase sur la TFS_N ne soit pas dégradée, le rehaussement du codage de l' ENV_N rapporté plus haut en figure 55 *déplace la transition* entre le verrouillage sur la TFS_N et le verrouillage sur l' ENV_N vers des fréquences caractéristiques plus basses au sein des fibres auditives lésées (figure 56B). Ces données révèlent donc **un déficit relatif du codage périphérique de la TFS_N par rapport à l' ENV_N .**

Si le déficit temporel observé par Kale et Heinz (2010) – bien que significatif – soit de taille modeste, les résultats obtenus par Henry et al. (2016) suggèrent la possibilité d'un déficit beaucoup plus substantiel. Chez l'animal sain (*i.e.*, normo-entendant), le verrouillage en phase des réponses d'une fibre auditive périphérique sur les indices de TFS_N et d' ENV_N évoqués par un son (par exemple, une bande de bruit) concorde avec la fréquence caractéristique de la fibre (*cf.* figure 57, colonnes de gauche montrant les réponses de fibres accordées respectivement en basse et haute fréquence). Le codage est donc tonotopique. Lorsque l'animal est soumis à un traumatisme acoustique permanent, le mécanisme actif se dégrade et la pointe de la courbe d'accord (tuning curve) des fibres auditives périphériques de la région fréquentielle lésée tend à disparaître (*cf.* figure 57, graphiques du haut : colonnes de gauche

versus colonnes de droite). Par conséquent, la queue (basse fréquence) de la courbe d'accord de la fibre joue un rôle accru dans la réponse aux sons.

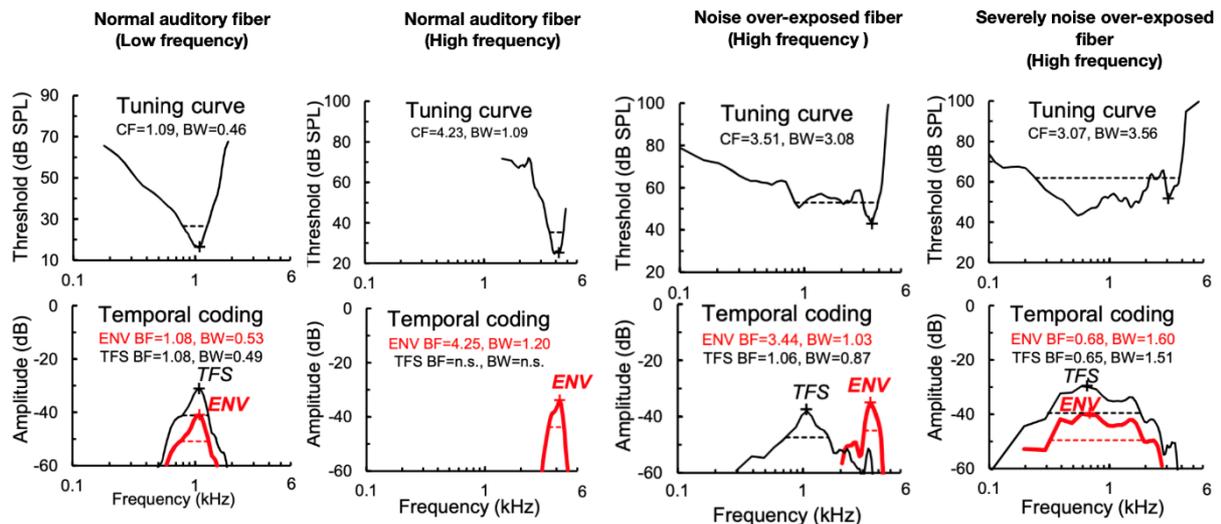


Figure 57. Figures de gauche : courbes d'accord (en haut) de fibres auditives périphériques saines et codage d'ENV_N et TFS_N (en bas) en réponse à un bruit large bande centré sur 1 kHz (figure a) et 4 kHz (figure b). Figures de droite : courbes d'accord (en haut) de fibres auditives périphériques lésées et codage d'ENV_N et TFS_N (en bas) en réponse à un bruit large-bande centré sur 4 kHz. Extrait de Henry et al. (2016).

Le codage tonotopique des fibres auditives périphériques lésées est donc profondément distordu et ceci introduit une discordance dans le codage des indices de TFS_N et d'ENV_N réalisé par chaque fibre. Ainsi, dans les fibres basales dont la fréquence caractéristique est supérieure à 1.5 kHz, la fréquence dominante maximale des réponses verrouillées sur les indices de TFS_N passe à 2.4 kHz dans les fibres exposées au bruit, alors qu'elle n'est que de 4.5 kHz dans les fibres saines. Par comparaison, les réponses aux indices d'ENV_N correspondent toujours à la fréquence caractéristique dans les fibres exposées au bruit. De fait, cette fibre auditive périphérique dont la position tonotopique n'a pas changé encode dorénavant les indices de TFS_N à une position spatiale différente (décalée vers les basses fréquences). Les conséquences perceptives d'une telle distorsion du codage temporel périphérique sont très certainement importantes. A l'aide d'une tâche psychophysique de transposition fréquentielle, Bernstein et Oxenham (2003) démontrent qu'une telle dissociation entre codage de TFS_N et d'ENV_N altère fortement la perception de la hauteur tonale et fondamentale des sons chez des auditeurs normo-entendants.

I.3.d.iii. Perte d'inhibition centrale

Les mécanismes de décodage des informations temporelles extraites à un niveau périphérique devraient dépendre - au moins partiellement - des relations inhibitrices entre les neurones de fréquences caractéristiques différentes (Shamma, 1985). Différentes études menées sur l'animal suggèrent que la quantité de neurotransmetteurs inhibiteurs (glycine et GABA) ayant une action dans les voies auditives centrales diminue avec l'âge et une perte auditive neurosensorielle dans le noyau cochléaire et le colliculus inférieur (Zhao & Evans, 1990). Or, ces neurotransmetteurs régulent les propriétés des réponses spectro-temporelles des neurones situés dans les noyaux cochléaires (Caspary et al., 1993 ; Zhao & Evans, 1990), dans le complexe olivaire supérieur (Moore & Caspary, 1982) et au niveau du colliculus inférieur (Pollak & Park, 1993). Si l'équilibre relatif entre excitation et inhibition neuronale est altéré au sein des noyaux auditifs du système auditif âgé ou présentant des dommages cochléaires, les propriétés de réponses neuronales devraient être fortement affectées. A minima, le décodage des indices de TFS_N extraits par un mécanisme de verrouillage en phase est donc probablement perturbé (Shamma & Lorenzi, 2013).

En résumé, les altérations observées ou susceptibles d'être observées dans le codage périphérique et central (nerf auditif, tronc cérébral) des indices de TFS_N et d' ENV_N à la suite d'une synaptopathie cochléaire, d'une perte auditive neurosensorielle et du vieillissement devraient fortement affecter les capacités de détection et de discrimination d'AM et de FM chez l'humain.

I.4. Caractérisation psychophysique et neurophysiologique du bruit interne

I.4.a. Préambule

La **variabilité** est une caractéristique intrinsèque du comportement humain. Elle se manifeste à la fois dans les activités perceptives et dans tout type d'action impliquant le système moteur, et ce même lorsque les conditions externes sont maintenues aussi constantes que possible (Faisal et al., 2008). Elle ne peut être ignorée dans les études de psychophysique sensorielle. Il n'est donc pas étonnant qu'elle prenne la forme d'un des principes de base de la théorie de la détection du signal (TDS) (Green & Swets, 1966).

L'ensemble des perturbations aléatoires d'un système sensoriel supposées être à l'origine de cette variabilité sont regroupées sous le terme générique de **bruit interne**, un concept particulièrement bien illustré par la théorie de l'information (Shannon & Weaver, 1949) sur

laquelle la TDS se fonde. Dans cette théorie décrivant la transmission de l'information à travers un canal, le bruit du canal d'un système s'ajoute au signal d'origine pendant la transmission. Cette théorie part du principe que le bruit de l'environnement externe déforme le signal d'entrée et qu'il représente une nuisance indésirable interférant avec la transmission du signal d'origine. Par conséquent, le degré de divergence entre les signaux d'entrée et de sortie d'un système fournit une estimation de la quantité de bruit présent dans un canal. Ce bruit peut être décrit comme un ensemble de composantes **déterministes** (mesurable par la performance moyenne) ou **stochastiques** (mesurable par la dispersion des valeurs autour de la moyenne) au sein du canal. Le rapport entre la performance moyenne et le degré de variabilité d'un système sensoriel permet d'accéder à la magnitude du signal par rapport à la quantité de bruit présent dans un signal. Quelle que soit son origine, les approches basées sur le paradigme du traitement de l'information et la TDS supposent que le bruit interne contribue à minimiser l'opération de certaines structures sensorielles, aboutissant à une discrimination *sous-optimale* du signal d'entrée (Green & Swets, 1966 ; Deutsch & Newell, 2004).

Au vu de la théorie de l'information et de la TDS, détection et discrimination sensorielles sont réalisées en présence d'une activité neuronale aléatoire continue (*activité neuronale spontanée*) ; par conséquent, le signal neuronal doit être d'une magnitude suffisamment importante pour émerger et être pris en compte lors de la prise de décision de l'observateur (Pardhan et al., 1996). Il s'ensuit que les niveaux relatifs de **bruit externe** (ensemble des incertitudes liées à l'environnement de test) et de **bruit interne** (au sein du système sensoriel ou du système nerveux central) contrôlent la réponse de l'observateur réel à une entrée sensorielle. Par exemple, lorsque le bruit externe est très faible par rapport au bruit interne, le niveau de bruit interne contraint la détection ou la discrimination du signal. A l'inverse, lorsque le niveau de bruit externe est élevé, celui-ci peut « noyer » l'effet du bruit interne. Ainsi, les représentations internes des stimuli physiques sont imparfaites et des stimuli identiques peuvent conduire à des perceptions différentes (Green, 1964 ; Green & Swets, 1966 ; Gold et al., 1999).

Depuis une cinquantaine d'années, différentes méthodes psychophysiques ont été développées afin d'évaluer la **magnitude** et la **nature (additive vs. multiplicative)** du bruit interne : ces méthodes incluent notamment la méthode du bruit équivalent (Nagaraja, 1964), le paradigme de double-passe (Green, 1964) , les observations multiples du signal dans le bruit (Swets et al., 1959). Il est également possible de réaliser une estimation directe de la variabilité à partir de la fonction psychométrique (Jesteadt et al., 2003) ou par ajustement de modèles perceptifs aux données empiriques (Oxenham & Buss, 2000 ; Wallaert et al., 2017 ; 2018 ; Ewert et al., 2020). La **localisation** du bruit interne sur le trajet des voies auditives est difficile à effectuer mais elle peut être explorée en combinant des mesures neurophysiologiques et des mesures psychophysiques (Faisal et al., 2008). Ainsi, les études s'intéressant aux propriétés stochastiques des réponses des fibres du nerf auditif suggèrent

qu'une proportion significative de bruit interne serait d'origine périphérique (*e.g.*, Carney, 1993 ; Colburn, 1973 ; 1977 ; Javel & Viemeister, 2000 ; Siebert, 1968 ; Heinz et al., 2001).

Ces techniques ont été majoritairement utilisées afin d'évaluer l'**impact** du bruit interne sur des tâches auditives et visuelles et ce afin d'étudier/évaluer: i) les trajectoires développementales (Buss et al., 2009 ; Banai et al., 2011 ; Falkenberg et al., 2014); ii) l'apprentissage perceptif (Jones et al., 2013); iii) les effets du vieillissement sur la perception (Pardhan, 2004) ; iv) des pathologies bien spécifiques telles que la dyslexie (Hornickel & Kraus, 2013 ; Ziegler et al., 2009), l'amblyopie (« œil paresseux ») (Levi et al., 2007), la dégénérescence maculaire (McAnany et al., 2013) et l'autisme (Milne, 2011 ; Dinstein et al., 2012).

Par ailleurs, le bruit interne peut être vu comme une **variable d'ajustement** dans les modèles computationnels visant à prédire quantitativement le comportement perceptif et notamment dans les modèles auditifs dédiés à la perception d'AM et de FM. Ainsi, la comparaison des prédictions avec les résultats empiriques en détection et discrimination d'AM suggère que le bruit interne pour le traitement des indices d'enveloppe temporelle (ENV_p) est non seulement additif (sa variance étant indépendante du stimulus) mais également multiplicatif (sa variance étant proportionnelle à la magnitude du stimulus le long de la dimension étudiée) (Ewert & Dau, 2004 ; Wallaert et al., 2017).

Ce chapitre a pour objectif de définir le concept du bruit interne ainsi que ses méthodes d'estimation. L'impact du bruit interne sur le développement, le vieillissement ou dans certaines pathologies sera discuté au travers de plusieurs études menées en modalité auditive et visuelle.

I.4.b. Bruit interne

I.4.b.i. Définitions

Comme tout dispositif physique, le cerveau humain opère en présence de variabilité. Néanmoins, les calculs sont effectués avec une rapidité et une précision remarquables, pouvant atteindre les limites physiques fixées par les capteurs sensoriels (Hecht et al., 1942). Le cerveau humain semble donc avoir développé des stratégies robustes vis-à-vis des multiples sources de variabilité biologique (London et al., 2010). D'un point de vue opérationnel, la variabilité s'observe à partir de l'enregistrement des réponses neuronales et/ou comportementales d'un organisme à des présentations répétées de signaux identiques. Les enregistrements électrophysiologiques varient d'essai en essai (Tomko & Crapper, 1974 ; Tolhurst et al., 1983 ; Shadlen & Newsome, 1998). Il en va de même sur le plan comportemental, cette variabilité représentant alors le facteur limitant la possibilité de prédire la réponse d'un sujet d'essai en essai par l'expérimentateur (Green, 1964).

L'interprétation de cette variabilité est difficile car elle peut provenir de deux sources distinctes :

i) la première dépend des **propriétés déterministes** du système étudié. L'aspect déterministe des systèmes sensoriels et des modèles computationnels de perception résulte du fait que la perception ou la sortie des modèles est entièrement contrainte par les valeurs des paramètres du système et des conditions initiales et ce à n'importe quelle échelle (moléculaire, neuronale ou comportementale). Le système déterministe - par définition - ne comprend aucun élément aléatoire en soi. Dans ce cas, la variabilité apparente des réponses du système sensoriel ou du modèle résulte d'un comportement instable pendant l'état initial du système ou bien de la méconnaissance de cet état initial. Ceci peut conduire à une amplification exponentielle de petites perturbations avec le temps, un phénomène couramment dénommé « *effet papillon* » (d'après les travaux de Lorenz qui illustre l'effet de cette incertitude sur les conditions initiales sur les mouvements atmosphériques ; Lorenz, 1963 ; 1969).

ii) la seconde dépend des **propriétés stochastiques** du système étudié. La stochasticité représente l'ensemble des processus pouvant être matérialisés par une distribution de probabilités ou un pattern pouvant être analysé statistiquement mais ne pouvant pas être prédit correctement. Cette forme de variabilité survient dans des phénomènes naturellement et véritablement stochastiques et décrit l'ensemble des perturbations aléatoires indépendantes du signal d'entrée (Faisal et al., 2008). Elle provient d'une variabilité d'essai en essai à l'échelle moléculaire ou neuronale et constitue donc une forme de **bruit intrinsèque** (car il contribue à la variabilité sans apporter aucune information et ne peut pas être éliminé) (London et al., 2010).

Le **bruit interne** représente une quantité inférée de la *variabilité biologique*, sa présence étant déduite de la manifestation de ses effets (Green, 1964). Il ne peut être observé directement par les physiologistes : en effet, sachant que le traitement sensoriel n'est pas complètement compris, il est difficile de dissocier le « bruit interne » du signal. Ainsi, le bruit interne est synonyme de la présence d'effets aléatoires qui ne sont pas totalement compris tant au niveau physiologique que psychophysique. Il représente donc l'incertitude dans une réponse interne à une entrée sensorielle, qui, contrairement au **bruit externe**, est généré par des sources intrinsèques de l'observateur (Green, 1964 ; Green & Swets, 1966 ; Jones et al., 2013).

Ce bruit interne intervient à toutes les étapes de traitement et prend donc différentes formes (voir figure 58 ci-dessous). Dès la première étape où l'énergie du stimulus acoustique est transformée en signal mécanique par le mouvement des cellules ciliées, une variabilité peut survenir au travers du mouvement brownien des cellules ciliées (Denk et al., 1989). A ce stade, le processus de transduction sous-jacent amplifie le signal audio d'entrée et le convertit en impulsions électriques (Faisal et al., 2008). Un bruit sensoriel interne déjà présent ou engendré durant le processus d'amplification (bruit transducteur) (Lillywithe et al., 1979)

augmente la variabilité d'essai en essai. Par conséquent, les niveaux de bruit sensoriel fixent le seuil de perception pour les étapes ultérieures de traitement de l'information car si le signal est plus faible que le bruit, celui-ci ne peut pas être distingué après amplification. Ceci est compatible avec le théorème de « l'inégalité de traitement des données » qui indique qu'une étape de traitement donnée (même en l'absence de bruit) ne peut pas extraire plus d'information qu'aux étapes précédentes. Par conséquent, pour réduire le bruit, les organismes fournissent en retour une énergie métabolique et structurale élevée dès le premier étage de traitement (étage sensoriel). Par la suite, des sources de bruit interne peuvent résulter de l'encodage et de la transmission neuronale dans le système auditif périphérique (Javel & Viemeister, 2000) ou central (e.g., Vogels et al., 1989).

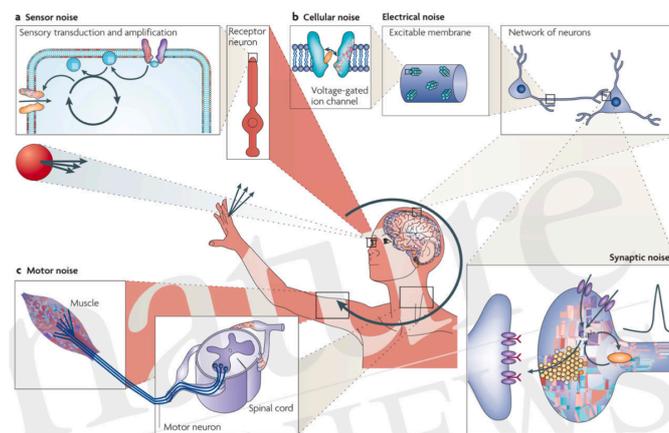


Figure 58. Boucle sensori-motrice impliquée dans une tâche comportementale (attraper une balle), indiquant les différentes sources de bruit interne à toutes les étapes du système nerveux. a) Les sources de bruit sensoriel affectent la transduction des signaux. b) Les sources de bruit cellulaire affectent les canaux ioniques des membranes excitables, la transmission synaptique et les interactions du réseau. c) Les sources de bruit moteur affectent les motoneurones et les muscles. Dans cette tâche comportementale, le système nerveux doit donc agir en présence de bruit dans la détection du signal, le traitement de l'information sensorielle et la réalisation du mouvement. Extrait de Faisal et al. (2008).

Le bruit interne se distingue donc des limitations systématiques pesant sur la performance de l'observateur (Berg, 2004 ; Berg & Green, 1990). Dans des tâches sensorielles, ces facteurs systématiques décrivent principalement la manière dont l'auditeur est capable d'intégrer l'information au sein ou entre les canaux sensoriels (Jones et al., 2013), ou de prendre en compte la totalité de l'information sensorielle disponible. Le concept d'**efficacité d'échantillonnage** (sampling efficiency) recouvre ces différentes notions. Le terme d'« efficacité d'échantillonnage » est introduit initialement par Burgess et al. (1981) (voir également Burgess et Barlow, 1983) mais ce même concept est parfois dénommé « efficacité de calcul » par Pelli (1981) ou encore « efficacité centrale » par Barlow (1977). Une efficacité d'échantillonnage réduite résulterait donc d'une intégration inadéquate (sous-optimale) des informations sensorielles disponibles (le sujet n'attribuant pas les bons poids perceptifs aux informations clés), ou encore la perte d'information sensorielle résultant de capacités

mnésiques ou attentionnelles réduites (e.g., certaines informations sont systématiquement oubliées ou ignorées).

I.4.b.ii. Effets neuronaux et comportementaux

Dans certains cas, ces sources de bruit semblent avoir des **effets bénéfiques** sur le comportement de réseaux neuronaux impliqués dans le traitement de l'information sensorielle (Faisal et al., 2008). L'une des améliorations constatées correspond au processus de **résonance stochastique** découvert dans les neurones visuels du chat. Cette dite résonance stochastique engendre une *amélioration de la détection et de la transmission de signaux faibles en présence d'un bruit synaptique*, dans des systèmes nécessitant l'atteinte d'un certain seuil (Faisal et al., 2008 ; Destexhe et al., 2001 ; Shu et al., 2003). Ceci se traduit par une facilitation de l'initiation des potentiels d'action et une robustesse dans la formation des réseaux neuronaux en présence de bruit, contribuant ainsi à l'adaptation aux changements dynamiques dans l'environnement.

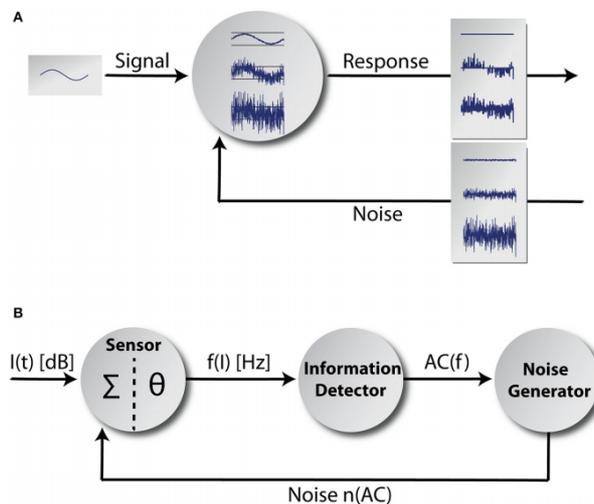


Figure 59. Modèle de résonance stochastique proposé par Krauss et al. (2016). (A) : principe de la résonance stochastique. Un signal infra-liminaire est présenté à l'entrée d'un détecteur et la réponse du détecteur est indiquée pour trois niveaux différents d'intensité de ce bruit interne. Si la quantité est trop faible (signal du haut), le signal ne peut pas être transmis et si la quantité de bruit est trop importante (signal du bas), le signal disparaît dans le bruit. En revanche, pour un niveau de bruit optimal (signal du milieu), la transmission du signal est la meilleure. (B) : modèle développé par Krauss et al. (2016) composé de trois unités fonctionnelles : (i) capteur supposé être localisé dans la cochlée entre une cellule ciliée interne (CCI) et une fibre afférente; (ii) unité recevant l'information provenant du capteur et qui calcule la fonction d'autocorrélation des variations du taux de décharge des fibres afférentes provenant d'un capteur, reflétant son contenu informationnel ; (iii) générateur de bruit blanc contrôlé par le détecteur d'information, retournant le bruit au capteur via des connexions retour (boucle de rétroaction). Extrait de Krauss et al. (2016).

Dans le système auditif, le processus de résonance stochastique semble participer d'un phénomène de régulation biologique faisant suite à un dommage cochléaire. Krauss et al. (2016) suggèrent que ce processus permet de réguler l'entrée sensorielle à la suite d'un traumatisme sonore ayant conduit à l'apparition d'acouphène: les sons d'intensité infraliminaire au sein d'une synapse endommagée peuvent en effet être transmis

efficacement dans les voies auditives et donc détectés grâce à l'action d'un bruit interne régulant l'entrée sensorielle (voir figure 59a présentant le principe de résonance stochastique) (Benzi et al., 1981 ; Collins et al., 1996 ; Levin & Miller, 1996). L'existence d'un **bruit interne additif** pourrait ainsi maximiser la transmission de l'information (Wiesenfeld & Moss, 1995). En effet, dans les systèmes de détection auto-adaptatifs de signaux basés sur la résonance stochastique, le niveau de bruit optimal est ajusté en continu grâce à une boucle de rétro-contrôle afin de maintenir l'efficacité de transmission de l'information même dans le cas où les propriétés du signal d'entrée changent (Krauss et al., 2016). Dans ce contexte de résonance stochastique, le contenu de l'information transmise est quantifié en comparant l'entrée et la sortie sensorielles (Shannon, 1948). Moss et al. (2004) démontrent que cette « information mutuelle » - lorsqu'elle est représentée en fonction de la magnitude du bruit interne - atteint un maximum dans une gamme de fréquences spécifique indiquant le *niveau de bruit optimal* additionné au signal d'entrée pour atteindre une transmission d'information optimale par résonance stochastique (voir figure 59a ci-dessus).

Les simulations obtenues à l'aide du modèle de Krauss et al. (2016) sont illustrées en figure 60. Ces simulations indiquent qu'à la suite d'une réduction chronique de l'entrée sensorielle au niveau du système auditif périphérique (suite à une lésion cochléaire), la résonance stochastique permet d'améliorer la transmission de l'information sensorielle en optimisant la fonction d'autocorrélation des fibres auditives périphériques. Ceci se produirait à mesure que la perte auditive augmente, conduisant à une amélioration des seuils auditifs (gain de +4 dB pour une perte auditive de 20 dB ; + 6 dB pour une perte de 40 dB). Pour compenser l'élévation des seuils auditifs à la suite d'une lésion cochléaire, le bruit interne est additionné à la boucle de résonance stochastique sur l'entrée sensorielle et doit être augmenté en permanence pour compenser la perte de fibres, constituant donc probablement la raison de l'hyperactivité neuronale généralement associée à l'acouphène. ***Au vu de cela, l'acouphène pourrait donc apparaître comme un mécanisme compensatoire visant à restaurer les seuils auditifs à la suite d'une perte auditive grâce au bruit interne. Dans ce cas particulier, l'augmentation du bruit interne apparaît donc bénéfique.***

La régulation du bruit interne devrait agir au niveau des projections efférentes du complexe olivaire supérieur sur les cellules ciliées internes ou encore au niveau du noyau cochléaire dorsal (Krauss et al., 2016). Par conséquent, l'action du bruit interne varie suivant la région fréquentielle affectée par le traumatisme sonore car elle implique différentes catégories de fibres auditives périphériques comme l'illustre la figure 60 ci-dessous. A l'aide du modèle de résonance stochastique proposé par Krauss et al. (2016), on peut émettre l'hypothèse qu'à la suite d'un traumatisme sonore, la résonance stochastique devrait être plus efficace dans les basses fréquences (inférieures à 3-4 kHz) que dans les hautes fréquences. En effet, en basse fréquence, les fibres auditives présentent majoritairement un taux de décharge spontanée élevé et des seuils bas. En raison de ce seuil bas, une faible quantité de bruit interne devrait être nécessaire pour conduire à une résonance stochastique efficace. Au contraire, les fibres

auditives qui sont les premières atteintes dans la partie basale de la cochlée codant les hautes fréquences (Zimmermann et al., 1995) présentent majoritairement un taux de décharge spontanée bas et des seuils élevés. Ces dernières devraient donc avoir besoin d'une plus grande quantité de bruit interne pour maximiser le processus de résonance stochastique. Ceci, pourrait expliquer en partie pourquoi les acouphènes apparaissent généralement dans les hautes fréquences : le système tente d'améliorer la transmission des sons par résonance stochastique, induisant une élévation globale des seuils auditifs chez les sujets acouphéniques dans les hautes fréquences seulement. En effet, **la forte quantité de bruit interne pourrait être perçue comme un acouphène qui en retour pourrait réduire le bénéfice apporté par la résonance stochastique sur les seuils absolus** (Krauss et al., 2016).

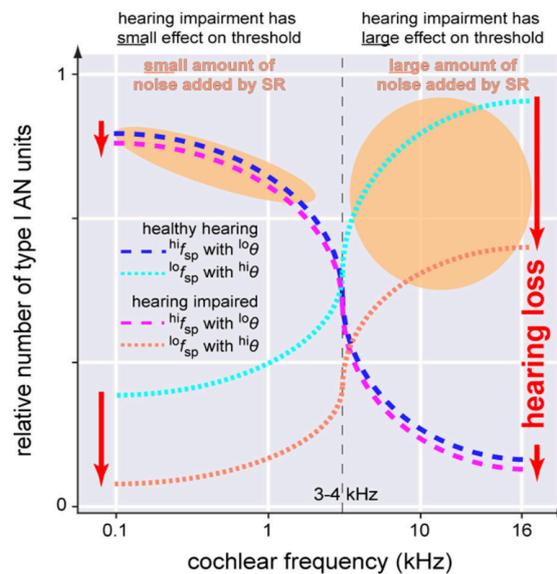


Figure 60. Asymétrie anatomique des distributions de seuil des fibres auditives de type I innervant une cellule ciliée interne. Dans une cochlée saine, les neurones présentant des taux de décharge spontanée élevés et des seuils bas (pointillés en bleu foncé) sont majoritaires en dessous de 3-4 kHz alors que les neurones présentant des taux de décharge spontanée bas et des seuils élevés sont observés dans toute la cochlée et majoritairement en haute fréquence (pointillés en bleu clair). A la suite d'un traumatisme sonore, ces neurones à taux de décharge spontanée bas et à seuils élevés sont affectés en premier et le plus sévèrement (pointillés en rouge) tandis que la première catégorie de fibres auditives (à taux de décharge spontanée bas et à seuils élevés) est en grande partie épargnée (pointillés magentas). La résonance stochastique augmente l'intensité du bruit pour compenser la perte auditive (zones oranges en basse et haute fréquence). Il est intéressant de noter que l'intensité du bruit en hautes fréquences doit être beaucoup plus élevée, ce qui peut engendrer des acouphènes dans cette région fréquentielle. Extrait de Krauss et al. (2016).

Cette étude ouvre de nouvelles perspectives sur la compréhension des acouphènes et pourrait déboucher sur des approches thérapeutiques. Dans ce contexte, il est possible d'envisager à la fois l'ajout d'un bruit externe pour réduire le processus de résonance stochastique et donc le bruit interne ou encore des stratégies de suppression de ce processus dans les hautes fréquences où les acouphènes sont généralement perçues (Krauss et al., 2016).

Toutefois, bien que ce bruit interne puisse avoir des effets bénéfiques sur le traitement de l'information auditive, il est plus souvent identifié comme étant lié à des **déficits auditifs en environnement bruyant** comme le démontrent récemment Resnik et Polley (2021). Dans cette étude combinant mesures comportementales et enregistrements corticaux chez le petit animal (rongeur), les déficits auditifs dans le bruit sont caractérisés à la suite d'une dégénérescence neuronale modérée des fibres afférentes de type I (CND pour cochlear neural degeneration) sur l'ensemble des fréquences induite pharmacologiquement par injection d'ouabaïne. Les résultats présentés en figure 61(H) indiquent que la dégénérescence neuronale ne provoque pas d'élévation des seuils (absolus) de détection de sons purs dans le silence, contrairement aux seuils de détection mesurés dans le bruit. Parallèlement, les mesures corticales indiquent une augmentation du gain cortical amplifiant à la fois le signal cible et le bruit de fond, susceptible d'expliquer pourquoi la détection dans le silence est préservée tandis que ce bénéfice n'est pas observé dans le bruit.

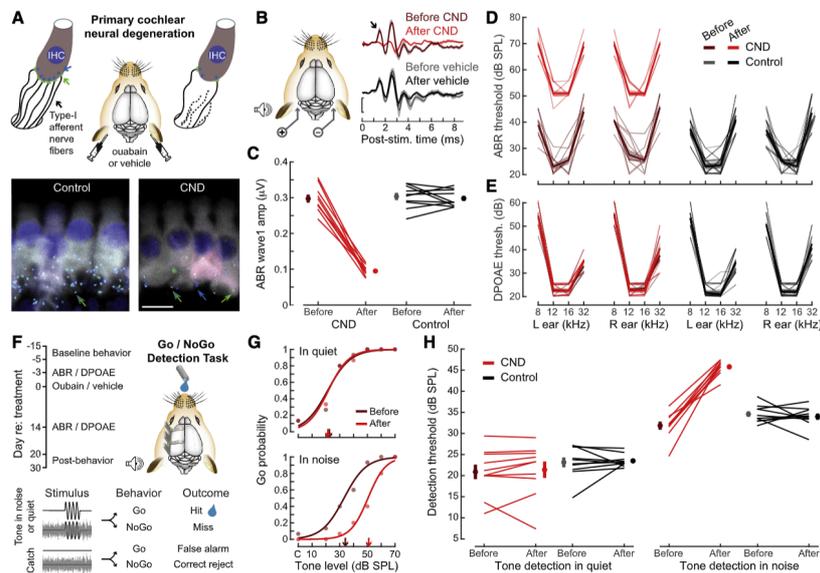


Figure 61. Résultats des tâches comportementales et mesures corticales démontrant que la dégénérescence neuronale des fibres auditives périphériques de type I dégrade la détection de sons purs dans le bruit mais pas dans le silence. (A) : le schéma du haut décrit l'approche utilisée pour éliminer les synapses afférentes des fibres du nerf auditif sur les cellules ciliées internes par application bilatérale d'ouabaïne sur la fenêtre ronde de la cochlée. Figure du bas : immunomarquage des zones présynaptiques et des terminaisons post-synaptiques des cochlées de souris avec et sans dégénérescence neuronale permettant d'identifier les synapses afférentes. La flèche bleue indique les éléments pré-synaptiques tandis que la flèche verte indique les récepteurs post-synaptiques glutamatergiques sur les terminaisons nerveuses des fibres de type I. (B) : schéma décrivant les mesures de potentiels évoqués auditifs en réponse à un son pur de fréquence égale à 12 kHz et d'intensité égale à 70 dB SPL. La flèche indique l'onde I. (C) : amplitude de l'onde I avant et après deux semaines d'administration d'ouabaïne (à gauche) ou d'une autre solution (à droite). Chaque ligne représente la moyenne des oreilles gauche et droite d'une souris stimulée par des sons purs de 12 kHz et de 70 dB SPL (10 souris sont traitées par ouabaïne et 10 souris forment un groupe témoin). (D) : seuils estimés à partir des potentiels évoqués auditifs et (E) des produits de distorsion (DPOAE) sur quatre fréquences avant et après traitement. (F) : déroulé des mesures comportementales et physiologiques pour chaque groupe. (G) : fonctions psychométriques pour la détection des sons purs dans le silence et dans le bruit. (H) : les seuils de détection avant et après traitement, dans le silence ou dans le bruit. Extrait de Resnik et Polley (2021).

Les enregistrements physiologiques révèlent que durant les essais où le signal - bien que présent - n'est pas détecté (essais « miss »), l'activité corticale apparaissant juste avant la présentation du signal cible en présence de bruit externe présente des bouffées d'activité neuronale aléatoire (voir figure 62 après CND dans le bruit, en bas à droite). **Cette activité neuronale précède le signal cible pourrait représenter une source de bruit interne responsable de déficits de détection dans le bruit chez les souris présentant une dégénérescence neuronale.** De plus, les résultats démontrent que ces activités corticales apparaissant juste avant la cible **prédissent les échecs de détection de la cible (Resnik & Polley, 2021).**

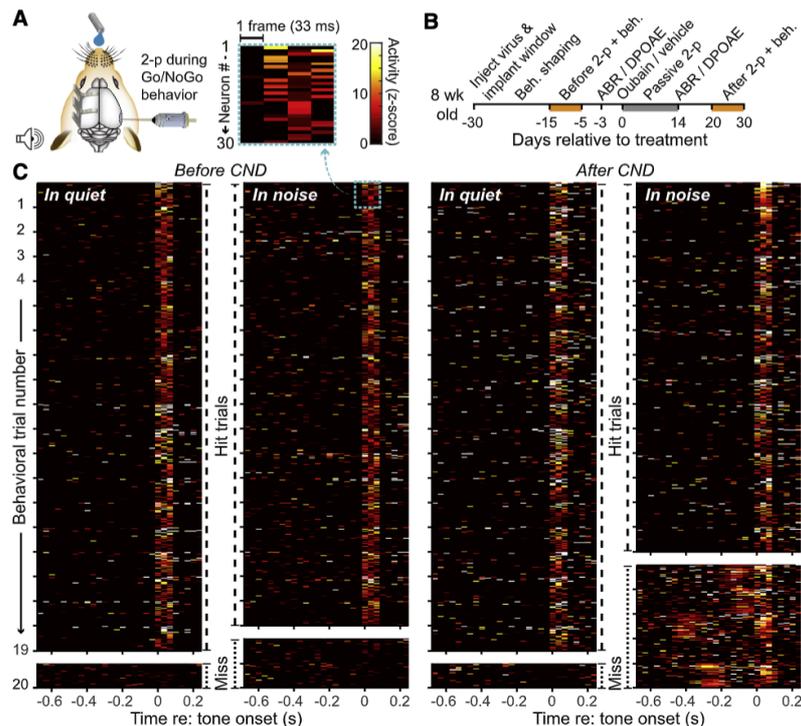


Figure 62. Dynamique d'ensemble de la couche corticale 2/3 pendant les essais comportementaux de type "hit and miss". (A) : réponses neuronales imagées pendant la tâche comportementale de type go/no-go. (B) : chronologie des mesures avant et après le traitement. (C) : neurogrammes d'une souris avec les changements d'activité trame par trame pour les essais comportementaux en réponse à un son cible de 50 dB SPL dans le silence ou en présence d'un bruit de fond continu de 50 dB SPL. Les réponses sont présentées séparément pour les essais réussis (hit) et manqués (miss). Les niveaux d'activité sont notés indépendamment pour chaque essai par rapport aux niveaux d'activité spontanée pendant la période totale de pré-stimulus. Extrait de Resnik et Polley (2021).

Il est à noter que différentes stratégies sont mises en place par le système nerveux central pour minimiser d'autres effets plus néfastes :

- i) le **moyennage** apparaît en présence d'informations sensorielles redondantes dès la première étape de traitement et permet de contrer l'effet du bruit interne lorsque différentes unités véhiculent un même signal et que chacune est affectée par différentes sources de bruit indépendantes (Faisal et al. , 2008). Ce procédé agit notamment sur les fluctuations aléatoires dans le mouvement des stéréocils des cellules ciliées internes par un couplage mécanique des canaux ioniques dans le processus de transduction (Kozlov et al., 2007).

- ii) la **divergence** apparaît lorsqu'un neurone établit plusieurs synapses. Quand les signaux électriques sont transmis sur de longues distances par l'intermédiaire d'axones myélinisés, ce processus intervient en envoyant le même signal de manière redondante sur différents axones et en combinant par la suite ces signaux une fois arrivés à destination. Ce phénomène apparaît dans le comportement des cellules ciliées internes qui fournissent une entrée divergente sur 10 à 30 cellules ganglionnaires grâce à une synapse en ruban (Glowatzki & Fuchs, 2002).
- iii) la **connaissance préalable** de la structure du signal permet de compenser ces inefficacités lors d'une tâche de reconnaissance d'un signal acoustique. La TDS montre qu'un "observateur idéal" soumis à un bruit interne additif peut être obtenu en adaptant les paramètres de cet observateur à ceux du signal à détecter. Ce phénomène intervient dans l'étape de prise de décision d'un observateur (Turin, 1960). Ceci correspond au principe de « filtre ajusté » où - sur le plan neurobiologique - les champs récepteurs neuronaux incorporent des connaissances préalables sur les entrées attendues et permettent donc aux neurones d'atténuer l'impact du bruit interne.

I.4.c. Méthodes d'estimation du bruit interne

Différentes méthodes psychophysiques ont été développées pour mesurer le bruit interne (Swets et al., 1959 ; Watson, 1963 ; Green, 1964 ; Spiegel & Green, 1981 ; Robinson & Berg, 1986 ; Berg, 1987 ; Lutfi, 1989 ; Dai, 1994 ; Lentz & Richards, 1997 ; Murray et al., 2002). Ces estimations de bruit interne découlent souvent d'une comparaison entre la performance d'un observateur idéal et celle d'un observateur réel à l'aide d'un modèle basé sur la TDS (Green, 1960b). La performance des observateurs humains est généralement 5 ou 6 dB moins bonne que celle d'un observateur idéal et cette différence, qui est relativement indépendante des propriétés du bruit externe, peut être attribuée à des inefficacités dans le processus de décision (Tanner & Birdsall, 1958) ou à d'autres sources de bruit interne (Swets et al., 1959 ; Green, 1960a ; de Boer, 1966). Les principales méthodes d'estimation sont présentées ci-dessous. D'autres méthodes ont été développées plus récemment, comme par exemple celle basée sur l'estimation de la sensibilité à une irrégularité dans l'enveloppe temporelle du signal (Moore & Vinay, 2019). Cette approche sera détaillée dans le chapitre portant sur les applications de ces méthodes à l'étude du développement, de l'apprentissage, du vieillissement ou de pathologies.

I.4.c.i. Méthode du bruit équivalent

A l'origine, la méthode du bruit équivalent a été développée par des ingénieurs afin de mesurer les propriétés de réponse des amplificateurs électroniques dont l'efficacité se

caractérise par la présence d'un **gain élevé** et d'un **bruit intrinsèque additif faible** (Friis, 1944). Ainsi, pour mesurer le bruit intrinsèque en sortie d'un amplificateur, un signal bruité est présenté en entrée en utilisant une étendue de valeurs d'amplitude (*i.e.*, d'écart-types) du bruit externe relativement large. Le rapport signal-sur-bruit est ensuite analysé en sortie de l'amplificateur en fonction du niveau de bruit externe: lorsque le bruit externe domine le bruit interne, le rapport signal-sur-bruit dépend principalement du bruit externe (il diminue à mesure que le bruit externe augmente) tandis que lorsque le bruit intrinsèque de l'amplificateur est plus fort que le bruit externe, le rapport signal-sur-bruit est déterminé majoritairement par le bruit intrinsèque et ne varie pas lorsque le niveau de bruit externe augmente (Lu & Doshier, 2008). Il s'ensuit que le niveau du bruit intrinsèque du système – pour sa composante additive- est équivalent au niveau du bruit externe présenté en entrée et estimé au coude de la fonction reliant l'amplitude du signal d'entrée aux écarts-types de bruit externe (voir figure 63 ci-dessous). Il est donc obtenu au point de transition entre les deux régimes (à l'endroit où le niveau de bruit externe et de bruit interne sont équivalents).

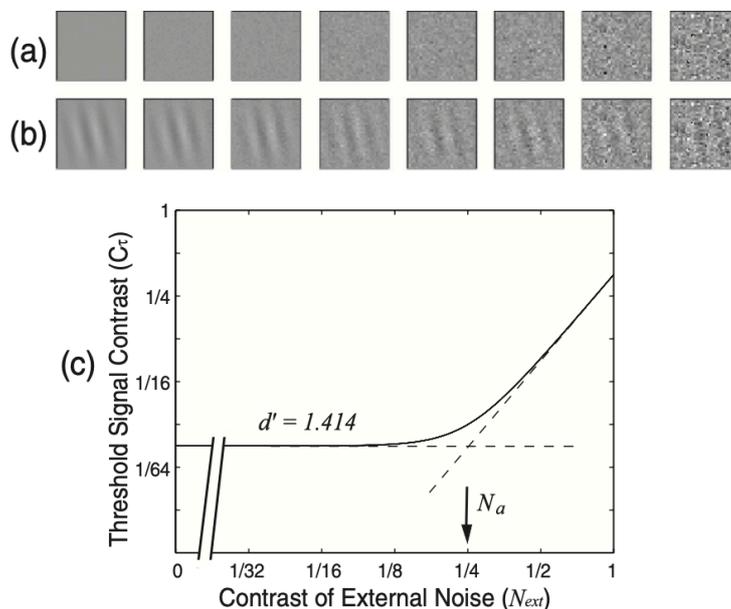
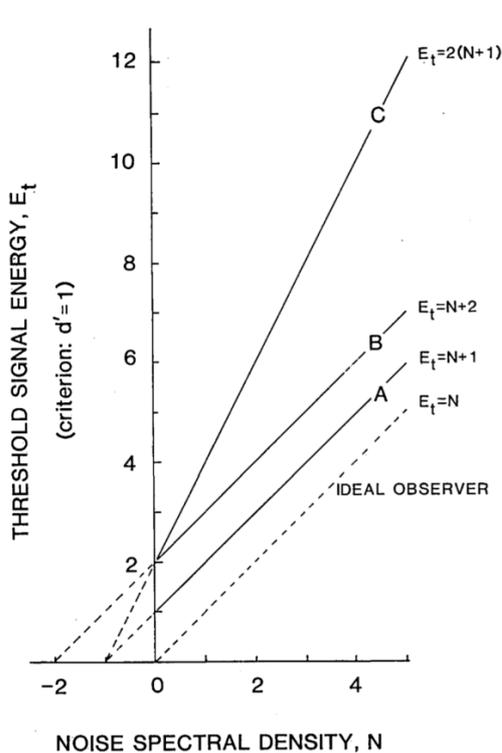


Figure 53. (a) : De gauche à droite : huit images de bruit visuel externe présentant une valeur de contraste croissante (N_{ext}); (b) :représentation du signal cible (un Gabor) additionné à chaque niveau de bruit externe. (c): Amplitude du signal d'entrée requise pour maintenir la performance de l'observateur au seuil de détection (niveau de sensibilité $d' = 1.414$) en fonction du contraste (N_{ext}) du bruit externe. Le seuil est constant jusqu'à un niveau du bruit externe N_a indiqué par la flèche verticale, au-delà duquel il augmente avec le contraste du bruit externe. Extrait de Lu et Doshier (2008).

Nagaraja (1964) propose d'utiliser cette méthode afin de mesurer le bruit interne dans les systèmes perceptifs. Cette méthode est couramment utilisée dans des tâches de détection et de discrimination auditive (Ahumada & Lovell, 1971 ; Bos & de Boer, 1966 ; Eijkman et al., 1966 ; Osman, 1971 ; Moore, 1975 ; Hartmann & Pumplin, 1988 ; Humes & Jesteadt, 1989 ; Richards et al., 1991) et encore plus largement dans des tâches visuelles (Rose , 1948 ; Tanner & Birdsall, 1958 ; Hay & Chesters, 1972 ; Pelli, 1981 ; Burgess et al., 1981 ; Legge et al., 1987 ; Geisler, 1989 ; Gegenfurtner & Kiper, 1992 ; D'Zmura & Knoblauch, 1998). L'idée de base est

très simple puisque l'estimation du niveau de bruit interne et des caractéristiques des systèmes perceptifs peut être obtenue en étudiant la façon dont la performance est affectée par un bruit externe manipulé directement par l'expérimentateur dans une tâche perceptive donnée.

D'un point de vue méthodologique, le seuil de détection ou de discrimination du signal cible est généralement mesuré à l'aide d'une méthode à deux intervalles et deux alternatives en choix forcé avec une procédure adaptative convergeant vers un niveau de sensibilité d' fixé (souvent à 1) (Legge et al., 1987 ; Lu & Doshier, 1998). Le signal cible et le signal de comparaison sont respectivement présentés dans un ordre aléatoire d'essai en essai. Ces deux signaux sont masqués par un bruit externe de variance fixe, synthétisé dans chaque intervalle d'observation. Le seuil de détection est mesuré pour des variances croissantes de bruit externe dans le domaine d'intérêt (intensité, modulation, fréquence, luminosité, etc.) (Legge et al., 1987). Cette méthode mesure donc les seuils de détection ou de discrimination sans faire d'hypothèse sur l'origine ou le locus du bruit interne (Nagaraja, 1964). Dans ce contexte, la mesure du bruit équivalent est une mesure de bruit interne qui est exprimée dans les mêmes unités que le bruit externe (Pardhan, 2004).



Équation utilisée dans le modèle de bruit équivalent :
 $E_t = k(N + N_{eq}) = kN + kN_{eq}$ (1)
 E_t : Energie du signal cible au seuil
 N : Densité spectrale de bruit externe
 k : Pente liée à l'efficacité avec laquelle l'observateur échantillonne les informations disponibles sur le stimulus
 N_{eq} : Interception avec l'axe des abscisses liée au niveau de bruit interne

Observateur idéal :
 Performance limitée uniquement par l'énergie du signal et le bruit externe
 $E_t = kN$ (1)
 $d'_c = \sqrt{E_t/N}$ (2)
 ou $E_t = (d'_c)^2 N$ (3)
 d'_c : Critère de détectabilité

Observateur réel :
 Performance limitée par un bruit interne additif et une inefficacité d'échantillonnage
 $d'_c = \sqrt{(kE)/(N + N_{eq})}$ (4)

Efficacité statistique :
 Quantification de la distance entre l'observateur idéal et l'observateur réel.
 $F = (d'_c(\text{observateur réel})/d'_c(\text{observateur idéal}))^2$ (5)

Figure 64. Prédictions issues du modèle linéaire obtenues à partir de l'équation (1). L'énergie du signal au seuil est tracée en fonction de la densité spectrale du bruit, tous deux étant exprimés dans les mêmes unités. Le seuil est défini pour un niveau de sensibilité (d') de 1. L'observateur idéal (courbe en pointillés) présente une efficacité d'échantillonnage (estimée par la pente de la courbe) égale à 1 et un bruit équivalent nul (correspondant à l'ordonnée à l'origine de la courbe). La courbe A présente une efficacité d'échantillonnage de 1 unité et un bruit équivalent de 1 unité. La courbe B présente une efficacité d'échantillonnage de 1 unité et un bruit équivalent de 2 unités. La courbe C présente une efficacité d'échantillonnage de 0.5 unités et un bruit équivalent de 1 unité. Extrait de Legge et al. (1987).

Ce paradigme du bruit équivalent permet d'évaluer séparément le niveau de **bruit interne additif** et **l'efficacité d'échantillonnage** de l'observateur à partir de la fonction empirique reliant le seuil de discrimination de l'observateur à l'énergie (ou la « magnitude ») du bruit externe (Pardhan, 2004). Cette fonction prend souvent le nom de fonction « seuil versus contraste (TvC) ». Ici, l'énergie du signal cible au seuil E_t , peut être reliée à la densité spectrale de bruit par une fonction linéaire du type : $E_t = k(N+N_{eq})$ (1) où la pente k (proportionnelle à l'efficacité d'échantillonnage) et l'intercept N_{eq} (le paramètre relié au bruit interne) sont estimés à partir des données empiriques. Plus ces valeurs sont grandes, plus les seuils sont élevés et par conséquent la performance est mauvaise (Legge et al., 1987). Le terme « bruit équivalent » est utilisé car N_{eq} est exprimé en unités de densité spectrale de bruit externe et peut par conséquent être conçu comme équivalent à une source de bruit additionnée au stimulus (Pelli, 1981). Une valeur d'efficacité d'échantillonnage inférieure à 1 unité reflète l'incapacité de l'observateur de collecter toute l'information disponible de manière optimale. Les fonctions « seuils vs. contraste » exprimées en unité d'énergie au seuil ($d' = 1$) dans le bruit (densité spectrale de bruit) sont présentées en figure 64 pour un observateur idéal (pointillés) et réel (traits plein). L'ordonnée à l'origine (négative) permet d'extraire la valeur (négative) du bruit interne équivalent et l'inverse de la pente permet d'extraire l'efficacité d'échantillonnage (Legge et al., 1987). Il est intéressant de noter que l'énergie au seuil en absence de bruit externe pour les courbes B et C est égale à deux fois celle de la courbe A en raison d'un bruit équivalent plus fort pour la courbe B et d'une efficacité d'échantillonnage plus petite pour la courbe C.

I.4.c.ii. Masquage de piédestal

Une alternative à la méthode du bruit équivalent consiste à mesurer puis **modéliser** les données de **discrimination** pour une modalité sensorielle spécifique dans laquelle un seul paramètre d'intérêt est varié. De fait, cette méthode réalise une estimation *indirecte* du bruit interne (Vilidaite & Baker, 2017).

Le principe de cette méthode consiste à mesurer les seuils de discrimination pour une entrée sensorielle donnée à l'aide d'une procédure adaptative à n alternatives en choix forcé (nAFC). Le paramètre d'intérêt du stimulus de base (« piedestal ») est soit augmenté soit diminué (à l'aide d'un pas spécifique échantillonnant une gamme de valeurs à tester dans un parmi n intervalles d'observation, et ce d'essai en essai. Ceci permet d'obtenir des seuils de discrimination pour plusieurs valeurs du paramètre d'intérêt (*i.e.*, plusieurs niveaux du « piedestal »). Généralement le signal cible et le (les) signau(x) de comparaison sont masqués. Dans une tâche visuelle de discrimination de contraste, la fonction résultante prend la forme d'une courbe en « U » (« dipper ») (Nachmias & Sansbury, 1974) dans laquelle on observe une amélioration des seuils à des niveaux de contraste relativement faibles et une élévation des seuils à des niveaux de contraste plus élevés. Cette fonction de contraste empirique peut être simulée à l'aide de l'équation de Naka-Rushton (Naka & Rushton, 1966) qui tient compte de

la non-linéarité du transducteur (Legge & Foley, 1980) : $resp = C^p / (Z + C^q) + \sigma_{int}$ où $resp$ est la réponse interne de l'observateur pour chaque intervalle, C représente le contraste du stimulus, p et q sont les exposants produisant une réponse accélératrice aux bas niveaux de contraste et compressive aux niveaux plus élevés, Z est la constante de saturation (paramètre qui contrôle le gain) et σ_{int} est proportionnel au bruit interne de chaque observateur. Pour simuler les expériences de discrimination par niveau de contraste, une réponse ($resp$) est calculée pour chacun des deux intervalles avec un bruit Gaussien de moyenne nulle et l'intervalle présentant la plus grande valeur de $resp$ est sélectionné. Cette modélisation permet donc de dissocier les effets du gain (Z) et du bruit interne additif (σ_{int}) (voir figure 65 ci-dessous).

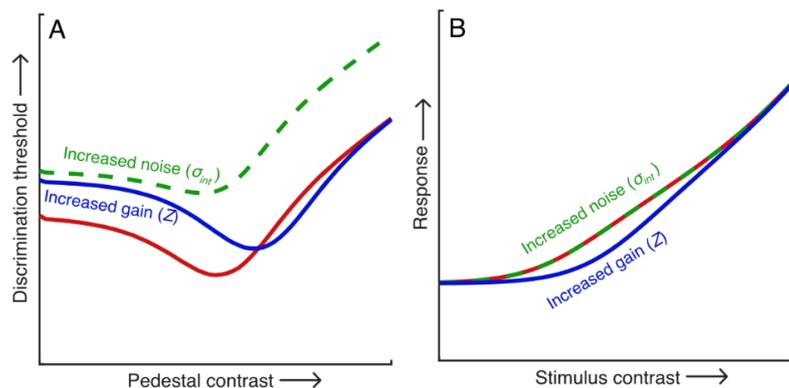


Figure 65. A : prédictions du modèle pour une tâche de discrimination de contraste avec différents paramètres du modèle. La courbe rouge correspond à la courbe de référence pour des paramètres $\sigma_{int} = 0.2$ et $Z=8$. Courbe verte : l'ensemble de la fonction est décalé verticalement lorsque le bruit interne est augmenté d'un facteur 3.5. Courbe bleue : l'ensemble de la fonction est décalé diagonalement lorsque le paramètre de gain est augmenté par un facteur 4 (aux faibles niveaux de contraste du piédestal, les seuils augmentent puis ils convergent aux valeurs de contraste élevées). B : Courbes de réponse aux valeurs de contrastes correspondantes. Les lignes rouges et vertes se chevauchent, indiquant que les changements de σ_{int} ne produisent pas de décalage de la fonction tandis qu'une augmentation du gain, Z , produit un décalage vers la droite. Extrait de Vilidaite et Baker (2017).

Glasberg et al. (2001) utilisent la méthode de masquage de piédestal pour étudier l'influence du **bruit interne** et du **bruit externe** sur la discrimination d'intensité pour un son pur faisant office de « piédestal ». La fréquence du son pur est égale à 250, 1000 ou 4000 Hz. L'intensité du signal cible peut augmenter (incrément) ou diminuer (décrément). La durée de l'incrément et du décrément varie systématiquement entre 5 et 200 ms. Deux conditions sont testées : i) en l'absence de bruit externe où seul le bruit interne est supposé affecter la performance ; ii) en présence d'un bruit externe large-bande supposé avoir une influence considérable sur la performance. La figure 66 présente les données obtenues par Glasberg et al. (2001) pour trois participants normo-entendants. Elles indiquent qu'en présence du bruit externe, les seuils de discrimination d'incrément et de décrément d'intensité diminuent (*i.e.* s'améliorent) lorsque la fréquence du piédestal augmente. Par contre, en l'absence de bruit externe, les seuils de discrimination d'incrément ou de décrément d'intensité ne varient pas significativement avec la fréquence du piédestal. Ce résultat suggère que **le bruit interne**

limitant la performance de discrimination d'intensité présente des caractéristiques similaires quelle que soit la fréquence. Les auteurs attribuent l'effet de la fréquence sur les seuils d'incrément et de décrément masqués aux fluctuations du bruit externe en sortie des filtres cochléaires centrés sur la fréquence du signal (à savoir, celle du piédestal). Ces fluctuations sont plus lentes dans les basses fréquences audio car les filtres cochléaires sont plus étroits, et ces fluctuations lentes - parce qu'elles sont audibles - ont un effet plus important (plus masquant) sur la discrimination d'incrément ou de décrément d'intensité.

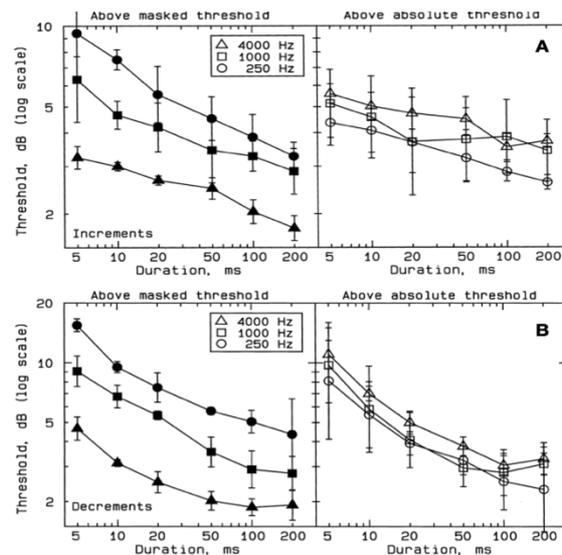


Figure 66. Seuils de discrimination moyens des 3 participants en fonction de la durée de l'incrément ou du décrément d'intensité (figure A : condition incréments ; figure B : condition décréments). Les seuils sont mesurés en présence de bruit masquant (colonne de gauche) ou en l'absence de bruit masquant (colonne de droite). Les courbes sont tracées à l'aide d'une échelle logarithmique. Le paramètre est la fréquence audio du stimulus. Les barres d'erreurs indiquent un écart-type de +/-1 entre les sujets. Extrait de Glasberg et al. (2001).

I.4.c.iii. Méthode de double-passe

En l'absence de variabilité dans le stimulus, la majeure partie de la variabilité des réponses de l'observateur est susceptible d'être produite par le bruit interne (Vilidaitte & Baker, 2017). Par conséquent, **il devrait être possible d'estimer le bruit interne en présentant une séquence de signaux physiques strictement identiques plusieurs fois de suite puis en mesurant la cohérence des réponses (« response consistency ») entre ces répétitions** (Lu & Doshier, 2008). Cette idée est exploitée par Green (1964) sous la forme du paradigme dit de « double-passe ». Cette méthode est considérée comme mesurant *directement* la quantité totale de bruit interne (additif et multiplicatif) par rapport au bruit externe, pour chaque condition de stimulus. Elle est habituellement réalisée à l'aide de deux passes successives (Green, 1964 ; Ahumada, 1967 ; Spiegel & Green, 1981 ; Burgess & Colborne, 1988 ; Lu & Doshier, 2008) mais peut être implémentée avec un nombre de passes supérieur à deux (paradigme n-pass).

La méthode de double-passe a rarement été utilisée en audition (Green, 1964 ; Spiegel & Green, 1981 ; Jones et al., 2013). Elle a beaucoup plus souvent été utilisée en modalité visuelle (e.g., Burgess & Colborne, 1988 ; Hasan et al., 2012 ; Lu & Doshier, 2008) ou pour des tâches plus cognitives (Diependaele et al., 2012). Dans l'étude princeps menée par Green (1964), une tâche de détection d'un son pur est effectuée à l'aide d'un paradigme à deux intervalles et deux alternatives en choix forcé 2I, 2AFC à stimuli constant sur une centaine d'essais. Chaque essai est divisé en deux intervalles d'observation successifs dans lequel le signal cible et le signal de comparaison sont respectivement présentés. Le signal cible est présenté au seuil (correspondant à une sensibilité d' fixée généralement à 1) tel que mesuré avant l'expérience. Les deux signaux sont masqués par un bruit externe de variance fixe, re-synthétisé dans chaque intervalle d'observation. Dans la moitié des essais et de manière aléatoire, le signal cible n'est pas présenté, mais le sujet n'est pas prévenu de l'absence de la cible. Dans ce cas, la réponse du sujet est uniquement déterminée par la variance du bruit interne.

D'un point de vue opérationnel, les signaux utilisés dans cette première session faisant office de « première passe » sont enregistrés et présentés à nouveau dans une nouvelle session faisant office de « deuxième passe ». Ces signaux sont donc strictement identiques sur le plan acoustique. Les réponses du sujet sont enregistrées lors de chaque essai et à chaque passe. Le **pourcentage d'accord** (PA) mesurant le fait que la réponse est identique ou non entre les deux passes successives et le pourcentage correct (PC) sont estimés pour chaque condition de test. Comme la performance, ce pourcentage d'accord varie donc entre 50 et 100 %. Un niveau d'accord de 100 % d'accord signifie que l'observateur donne toujours la même réponse aux essais correspondant des deux passes.

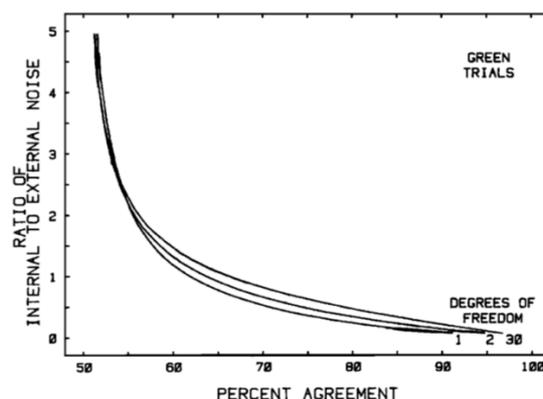


Figure 67. Prédiction du modèle probabiliste de décision dans le paradigme de double-passe proposé par Spiegel et Green (1981). Ce graphique représente la relation prédite entre le pourcentage d'accord (pour les essais où le signal est absent) entre les deux passes (présenté en abscisses) et le rapport des variances du bruit interne et du bruit externe (présenté en ordonnées). Comme indiqué dans cette figure, un pourcentage d'accord de 66 % (comme observé dans les expériences psychophysiques sur les sujets humains) est obtenu pour un rapport de variance de 1. Cette relation non linéaire entre le pourcentage d'accord du modèle et le rapport des variances (bruit interne/bruit externe) permet donc déterminer le niveau de bruit interne pour un observateur réel donné à partir des réponses mesurées empiriquement. Extrait de Spiegel et Green (1981).

Dans cette étude, Green (1964) considère seulement les essais où le signal est absent pour estimer le bruit interne. Seul le pourcentage d'accord, PA, est donc pris en compte pour l'estimation du bruit interne car PC n'est pas calculable. Le pourcentage d'accord observé empiriquement sur les participants est comparé au pourcentage d'accord prédit par un modèle probabiliste de décision extrêmement simple, faisant l'hypothèse qu'à chaque essai, la réponse du sujet (c'est-à-dire, le choix de l'intervalle cible) est uniquement déterminée par la variance des bruits externe et interne (le sujet choisissant l'intervalle d'observation pour lequel l'énergie dans la dimension pertinente est la plus grande), sachant que : i) ces deux bruits sont additionnés et ii) la distribution de ces bruits est Gaussienne et de moyenne nulle. Les prédictions de ce modèle probabiliste sont présentées en figure 67 dans le cas des expériences de détection menées par Green (1964). D'après ce modèle, un pourcentage d'accord moyen de 66 % est obtenu pour une variance de bruit interne égale à la variance de bruit externe. Ce niveau de pourcentage d'accord est observé empiriquement sur les sujets humains ayant participé aux tâches de détection. *Pour cette tâche, le bruit interne est donc aussi fort que le bruit externe ...*

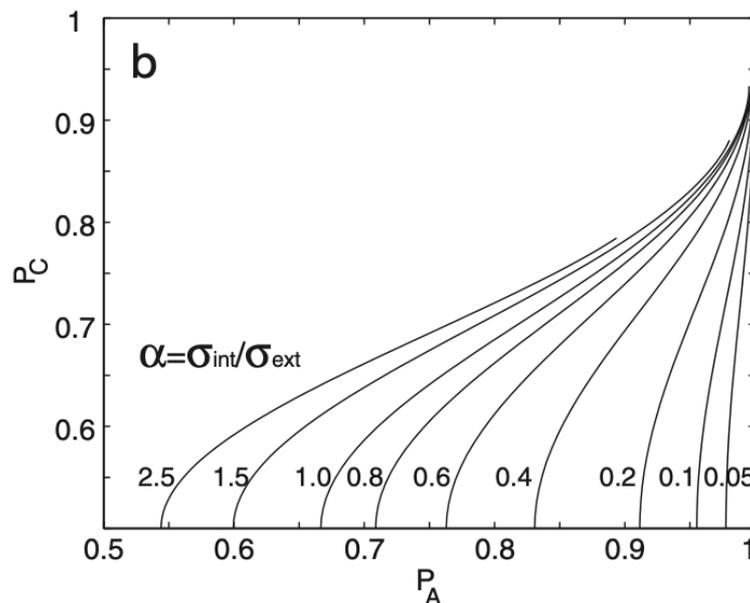


Figure 68. Courbes théoriques reliant le pourcentage correct PC au pourcentage d'accord PA pour une gamme de rapports entre le bruit interne σ_{int} et le bruit externe σ_{ext} . Extrait de Lu et Doshier (2008).

Dans le cas où le signal est totalement absent au sein des essais pris en compte, la variance du bruit interne s'estime donc directement à partir du pourcentage d'accord entre passes successives. Dans le cas où le signal est présent dans l'un des deux intervalles de chaque essai, la mesure du bruit interne dépend à la fois de la précision des réponses (le pourcentage de réponses correctes, PC) et de la cohérence des réponses entre passes successives (le pourcentage d'accord, PA) de l'observateur réel. Les estimations de pourcentage d'accord sont alors généralement obtenues pour différentes conditions d'amplitude du signal (Lu & Doshier, 2008) et donc pour différents niveaux de performance, PC. Pour chaque niveau de bruit externe σ_{ext} , la forme de la fonction « PC versus PA » est alors déterminée par le rapport

des variances $\alpha = \sigma_{int}/\sigma_{ext}$. Ces fonctions PC versus PA sont simulées en figure 68 pour différentes valeurs de α à l'aide des équations de base de la TDS appliquées à la double-passe. Le niveau de bruit interne σ_{int} (estimé en unités de bruit externe σ_{ext}) des sujets réels est obtenu en comparant les courbes PC versus PA empiriques aux courbes PC versus PA prédites (Lu & Doshier, 2008 ; Spilioti et al. , 2016 ; Vilidaite & Baker, 2017). Grâce à cette méthode, le bruit interne total - pouvant être de nature additive et/ou multiplicative - est comparé à une quantité physique contrôlée par l'expérimentateur (Lu & Doshier, 2008), à savoir l'amplitude du bruit externe.

Les estimations du bruit interne à l'aide du paradigme de double-passe indiquent que le ratio $\alpha = \sigma_{int}/\sigma_{ext}$ est généralement compris entre 1.0 et 1.4 en audition (Green, 1964 ; Swets et al., 1959) et entre 0.65 et 1 en vision (Burgess & Colborne, 1988; Gold et al., 1999 ; Levi & Klein, 2003). D'après Green (1964), ces estimations représentent des limites inférieures du rapport α car elles ne considèrent pas le biais de réponse qui a pour effet d'augmenter l'accord des réponses entre les deux passes et par conséquent est susceptible de diminuer l'estimation du bruit interne. Deux études récentes démontrent également que ce ratio n'est pas affecté par l'apprentissage perceptif (Chung et al., 2005 ; Gold et al., 1999).

Hasan et al. (2012) testent plusieurs facteurs susceptibles d'affecter les estimations de bruit interne à partir des données empiriques de double-passe : i) l'**ordre** de présentation des signaux aux essais correspondant entre les deux passes ; ii) le **retour sur les réponses (feedback)** du participant d'une passe à l'autre pour les essais où les signaux sont identiques (important à étudier car le feedback de chaque participant diffère d'une passe à l'autre et cette différence peut être attribuée au stimulus externe présenté au participant durant les deux passes, même si cette différence ne porte pas sur le stimulus auditif ou visuel en soi), ce qui va perturber l'ensemble de l'approche et enfin iii) le **nombre d'essais** présentés au sein de chaque condition de test (la résolution de l'estimation du bruit interne). Pour tester les effets de ces trois facteurs sur l'estimation du bruit interne via un paradigme de double-passe, des tâches de discrimination auditive (fréquentielle) et visuelles (discrimination de l'orientation d'un Gabor) sont respectivement effectuées sur six et dix participants. Dans un premier temps, différentes stratégies sont utilisées pour délivrer la seconde-passe : i) la séquence des signaux est présentée dans le même ordre en passe 1 et 2 ; ii) les signaux sont présentés dans un ordre aléatoire en passe 2 ; iii) chaque stimulus est répété immédiatement dans une seconde passe. Dans un second temps, l'effet du feedback est évalué pour chaque participant en conservant l'ordre des signaux présentés en passe 1 et en passe 2 en présentant exactement le même feedback du participant aux essais correspondants en passe 1 et 2. Ensuite, le bruit interne est extrait à partir des mesures empiriques individuelles de PA et de PC pour chaque condition à l'aide d'un algorithme minimisant les erreurs quadratiques entre le couple (PA ; PC) mesuré et prédit par un modèle dérivé de la TDS. Les résultats obtenus pour les trois stratégies visant à tester l'effet de l'ordre démontrent que **le choix spécifique d'une stratégie de double-passe**

ainsi que le feedback du participant d'une passe à l'autre n'affectent pas les estimations de bruit interne.

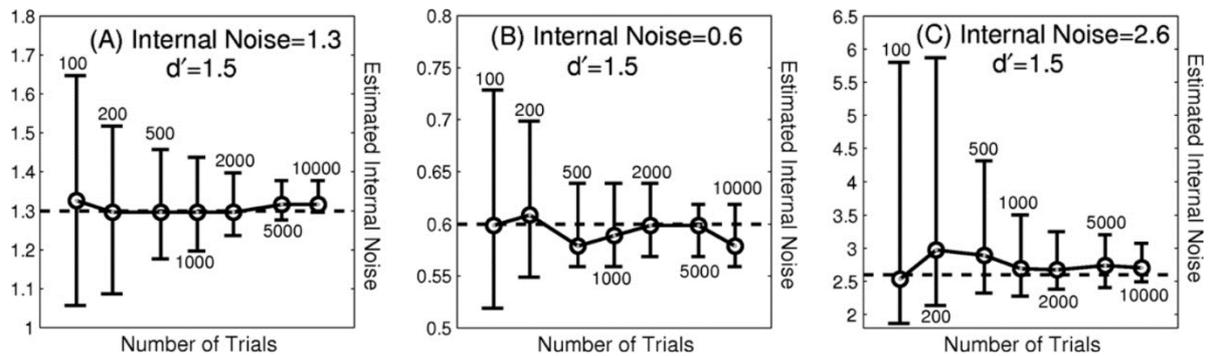


Figure 69. Effet du nombre d'essais sur l'estimation du bruit interne par double-passe. Les figures A à C montrent les résultats simulés pour trois valeurs de bruit interne distinctes (1.3, 0.6, 2.6) ; la même valeur de sensibilité (d') de 1.5 est utilisée pour les trois simulations. Les symboles représentent la valeur médiane du bruit interne estimé pour 100 simulations et les barres d'erreurs représentent la fourchette du 25^è et 75^è percentile. Les lignes en pointillés indiquent la valeur simulée. Le nombre d'essais indiqué ne concerne qu'un seul passage ; le nombre d'essais total qui serait nécessaire pour l'expérience de double-passe est donc le double du chiffre indiqué ci-dessus. Extrait de Hasan et al. (2012).

Enfin, l'effet de la résolution de l'estimation du bruit interne est étudié en simulant une tâche à deux alternatives en choix forcé avec un niveau de bruit interne fixe pour un nombre d'essais croissant allant de 100 à 10 000 et pour une sensibilité d' égale à 1.5. Les résultats représentés en figure 69 indiquent que **la variabilité de l'estimation du bruit interne diminue à mesure que le nombre d'essais augmente et ce quelle que soit la valeur de bruit interne**. Ces résultats indiquent qu'au moins 10 000 essais sont nécessaires afin d'obtenir une variabilité de l'estimation de bruit interne inférieure à 15% (Hasan et al., 2012).

I.4.c.iv. Estimation à partir de la pente de la fonction psychométrique

La fonction psychométrique associée à une tâche perceptive donnée fournit probablement le meilleur moyen d'estimer la distribution du bruit interne bien qu'il n'y ait pas de consensus concernant la forme exacte de cette fonction psychométrique (e.g., Buus & Florentine, 1991 ; Moore et al., 1999). La figure 70 ci-dessous illustre parfaitement les effets du bruit interne (dont la variance est variée) sur la **pente** de la fonction psychométrique dans une tâche psychophysique donnée. Notons qu'une augmentation de bruit interne est associée à une diminution de pente dans les tâches oui/non et les tâches en choix forcé, mais qu'elle est également associée à une augmentation de seuil dans les tâches en choix forcé.

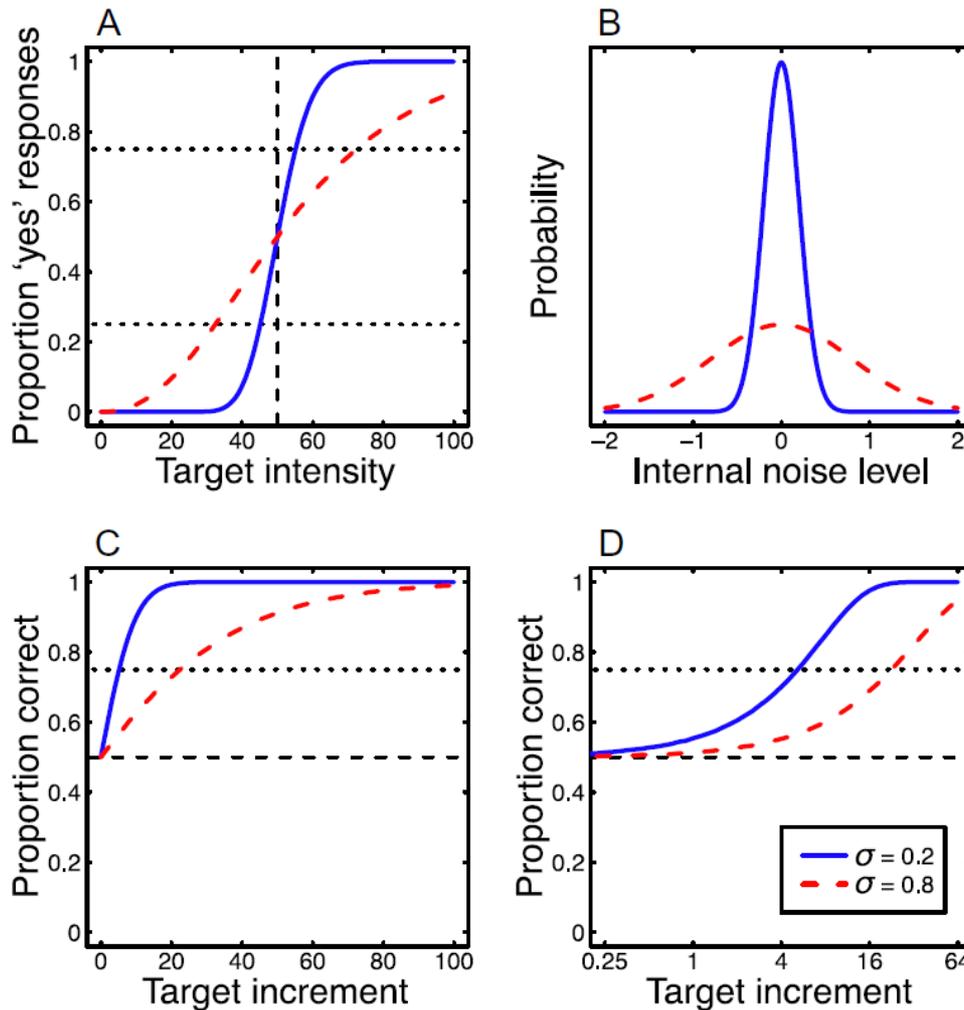


Figure 70. Fonctions psychométriques prédites dans une tâche oui/non (figure en haut à gauche) et dans une tâche en choix forcé (figures en bas à gauche [échelle linéaire] et à droite [échelle logarithmique]) lorsque la variance du bruit interne augmente par un facteur 4 (figure en haut à droite). Extrait de Vilidaite et Baker (2017).

Cette méthode d'extraction du bruit interne à partir de la pente de la fonction psychométrique a été exploitée dans des tâches de discrimination d'intensité en audition (Lufti, 1989 ; Buus & Florentine, 1991 ; Moore et al., 1999 ; Jesteadt et al., 2003) et est particulièrement bien décrite dans les travaux de Jesteadt et al. (2003). La fonction psychométrique est généralement construite à partir des données obtenues avec une méthode à stimuli constants. Notons que la fonction psychométrique peut également être reconstruite à partir des données obtenues avec une méthode adaptative. Dans ce cas, la pente de la fonction psychométrique est alors estimée à partir de l'écart-type des estimations de seuils pendant les dernières inversions de l'escalier psychophysique.

I.4.c.iv. Méthode de classification de limite

Cette méthode est appliquée par Jones et al. (2013) à une tâche de discrimination fréquentielle. Ici, la tâche psychophysique est conçue comme un problème de classification binaire : l'espace de décision est décrit comme un espace à deux dimensions où chaque axe

correspond à la valeur du paramètre d'intérêt tel qu'il est présenté dans l'intervalle en question (1 ou 2). Dans ce problème de classification, la « frontière » idéale est une droite passant par la valeur centrale du paramètre sur lequel est appliqué une variation dans les stimuli d'essai en essai. Dans la tâche de discrimination fréquentielle décrite par Jones et al. (2013), le signal cible et le signal de comparaison sont présentés dans un ordre aléatoire d'essai en essai entre les intervalles 1 et 2. Les deux signaux sont présentés symétriquement autour d'une certaine fréquence (X). Les fréquences du signal cible et de comparaison sont égales à $X \pm \sigma_{Hz}$ (où σ_{Hz} est variée avec un certain pas spécifique dans une gamme de valeurs permettant de couvrir l'ordre de grandeur du bruit interne dans la tâche en question). Par conséquent, la « frontière de classification » idéale passe par X et dans cet exemple présente une pente égale à 1 passant par l'origine (voir figure 71 ci-dessous).

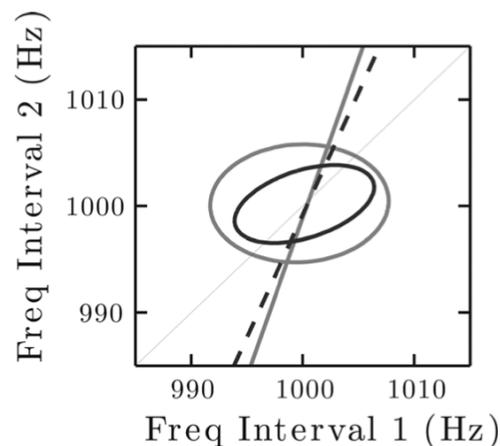


Figure 71. Ajustement de modèle de classification sur un seul auditeur. Les lignes de classification sont représentées par des lignes en traits gras et les écarts-types des erreurs par rapport aux limites (ellipses) indiquent le bruit interne. Une petite ellipse correspond à une valeur de bruit interne faible. Une limite de classification proche de la fonction d'identité (diagonale passant par l'origine) indique une stratégie d'encodage efficace. Extrait de Jones et al. (2013).

Dans cette méthode, chaque pente (ou « limite de classification ») est estimée en trouvant la fonction linéaire qui prédit le mieux les réponses des participants en fonction des fréquences des signaux présentés. A partir de l'ajustement de ces pentes, trois facteurs sont estimés : i) l'**efficacité d'échantillonnage** définie comme l'angle entre la pente observée et la pente idéale; ii) le **bruit interne** estimé à partir de l'écart entre les erreurs de classification (prédites vs. observées); iii) le **biais de réponse** représentant la distance euclidienne de la limite de classification par rapport au point de fréquence centrale X (ici, défini par les coordonnées X (1000 Hz ; 1000 Hz)).

I.4.c.v. Méthode d'estimation basée sur le gel des fluctuations externes

Ewert et Dau (2004) réalisent la première étude visant à caractériser la contribution relative du bruit interne et du bruit externe dans une tâche de détection d'AM. Ici, les seuils de détection d'une AM sinusoïdale de 4, 16, 64 et 256 Hz sont mesurés en présence d'un masque d'AM obtenu en utilisant des bandes de bruit étroites de largeur 8, 32, 128, 512 et 2048 Hz comme porteuses de la modulation cible. Dans une condition expérimentale, la porteuse (la bande de bruit) est rafraîchie entre stimuli (« condition aléatoire »). Dans une autre condition, la porteuse est « gelée » (c'est-à-dire, strictement identique entre stimuli). Le spectre de modulation du bruit externe est donc identique (et triangulaire) entre les deux conditions. Les effets de masquage dans le domaine de l'AM sont donc identiques dans les deux conditions. En condition « masque gelé », la stochasticité du masque est éliminée et les décisions de l'observateur ne sont donc influencées que par le bruit interne. En condition « masque aléatoire », le masque est stochastique (il varie de stimuli en stimuli) et les décisions de l'observateur sont donc influencées par deux sources de variabilité distinctes : la variabilité externe et à laquelle s'ajoute celle produite par le bruit interne. Les résultats sont représentés en figure 72 ci-dessous.

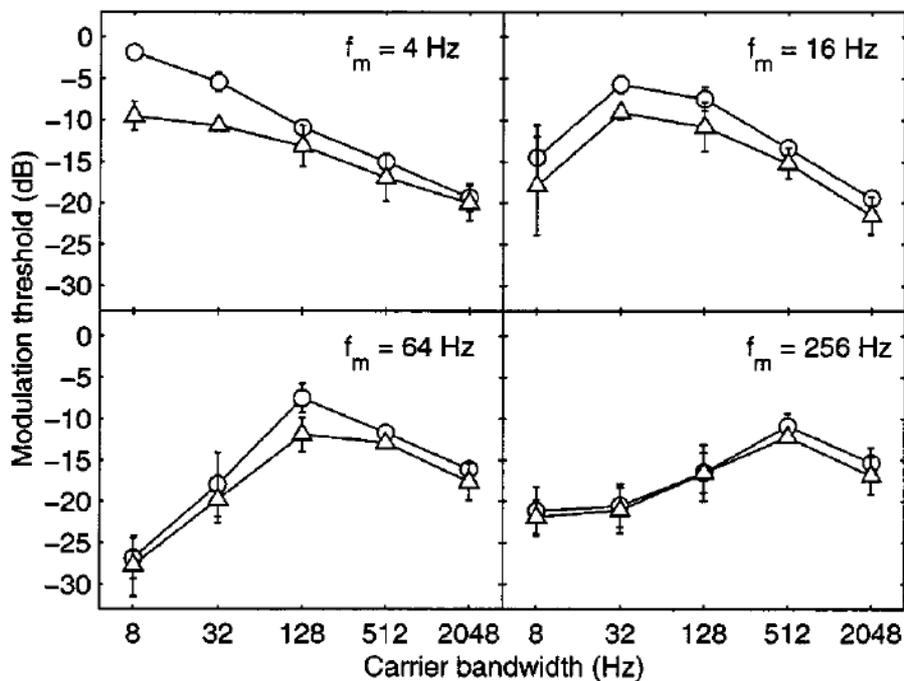


Figure 72. Seuils de détection d'une AM sinusoïdale de cadence f_m en fonction de la bande passante du bruit faisant office de porteuse. Chaque figure correspond à une valeur de f_m donnée (4, 16, 64 ou 256 Hz). Les triangles correspondent aux données mesurées en condition « masque gelé » et les cercles aux données mesurées en condition « masque aléatoire ». Extrait d'Ewert et Dau (2004).

Les seuils de détection d'AM sont généralement identiques en condition « masque gelé » et « masque aléatoire », à l'exception de la condition où la cadence d'AM cible est de 4 Hz et la bande de bruit porteuse est la plus étroite (8 et 32 Hz). Ainsi, dans les conditions où les fluctuations d'amplitude du bruit masquant et celles de la cible sont relativement rapides (*e.g.*, > 8-32 Hz), le bruit externe n'a que peu ou pas d'influence sur les seuils et la performance est essentiellement influencée par le bruit interne. ***La magnitude de ce dernier est donc plus forte que celle du bruit externe introduit par la porteuse. Ce résultat est important car il suggère que le bruit interne peut être extrêmement fort dans une tâche de détection d'AM (en tout cas, plus fort que le bruit externe).*** Dans les conditions où les fluctuations d'amplitude du bruit masquant et de la cible sont relativement lentes et résolues auditivement par les auditeurs, les seuils de détection s'améliorent de 5 à 10 dB en condition « masque gelé ». Cet effet de démasquage est inhérent à la méthode utilisée et reflète probablement l'usage d'un masque dont la structure temporelle est perceptible mais totalement identique entre stimuli. L'effet de démasquage résulterait donc de la mémorisation de la structure temporelle du masque par l'auditeur, permettant à l'auditeur de le « soustraire » en quelque sorte lors de la comparaison des intervalles cible et de comparaison au sein de chaque essai.

Dans le modèle EPSM, les décisions sont basées seulement sur l'augmentation de la variable de décision. Cela résulte en une même prédiction pour les conditions bruit masquant aléatoire gelé et non gelé. Le second modèle, PEMO, tout en surestimant la performance de détection dans les deux conditions, est capable de prédire une différence de seuil entre l'état gelé et aléatoire. Ceci peut s'expliquer par une plus grande quantité d'information reflétée dans la variable de décision au sein du modèle PEMO par rapport au modèle EPSM, indépendamment des différences de bruit interne. Cette observation contredit donc le fait que l'étape de décision d'un observateur dans une tâche de détection en 3AFC ne se base pas sur les phénomènes de mémorisation du stimulus comparé dans les trois intervalles au sein de chaque essai. Le modèle EPSM de Ewert et Dau (2000) ne tenant pas compte des phénomènes de mémorisation dans la prise de décision ne simule donc pas correctement le processus de décision d'auditeurs réels.

I.4.c.vi. Ajustement de modèles

Le bruit interne peut être estimé grâce à un modèle computationnel de traitement perceptif, en ajustement la variance du bruit interne de ce modèle de façon à ce que les données empiriques (*e.g.*, des mesures de seuils ou de performance d'un sujet) correspondent aux données simulées dans des conditions expérimentales identiques. Le ***bruit interne*** est souvent considéré comme uniforme à travers l'ensemble des conditions (*e.g.*, Patterson, 1976 ; Plack & Oxenham, 1998). Néanmoins, de multiples facteurs peuvent parfois être associés au bruit interne et certains sont susceptibles de varier en fonction des ***propriétés physiques des signaux*** ou de la ***procédure expérimentale*** (Jesteadt et al., 2003). Ces différentes sources de

bruit sont modélisables dans les modèles computationnels et ceci permet d'apporter une interprétation sur les mécanismes sous-jacents.

Par exemple, les travaux de Durlach et Braida (1969) sur la résolution auditive en intensité proposent que cette dernière est affectée par différentes sources de bruit associées au sujet, au paradigme expérimental ou aux signaux. Le modèle théorique de Durlach et Braida (1969) vise à prédire la résolution auditive en intensité et se décompose en une étape décisionnelle (distinguant la sensibilité du biais de réponse), un bruit interne dit « de sensation » un bruit interne dit « mnésique » et central : le **bruit de sensation** est supposé refléter l'inefficacité de codage à des étages de traitement précoces du système auditif et ne devrait pas être affecté par des changements dans le protocole expérimental, tandis que les sources de bruits **mnésiques et centraux** ne devraient pas dépendre du stimulus en lui-même mais plutôt des paramètres expérimentaux tels que la durée de l'intervalle entre les stimuli ou encore les conditions de présentation des stimuli entre les intervalles (Berliner & Durlach, 1973). La **source de bruit dominante** dépend des conditions de test et il est donc possible que différents paradigmes conduisent à des dépendances différentes vis-à-vis des paramètres testés comme la durée du stimulus, sans pour autant impliquer l'existence de mécanismes différents (Jesteadt et al., 2003).

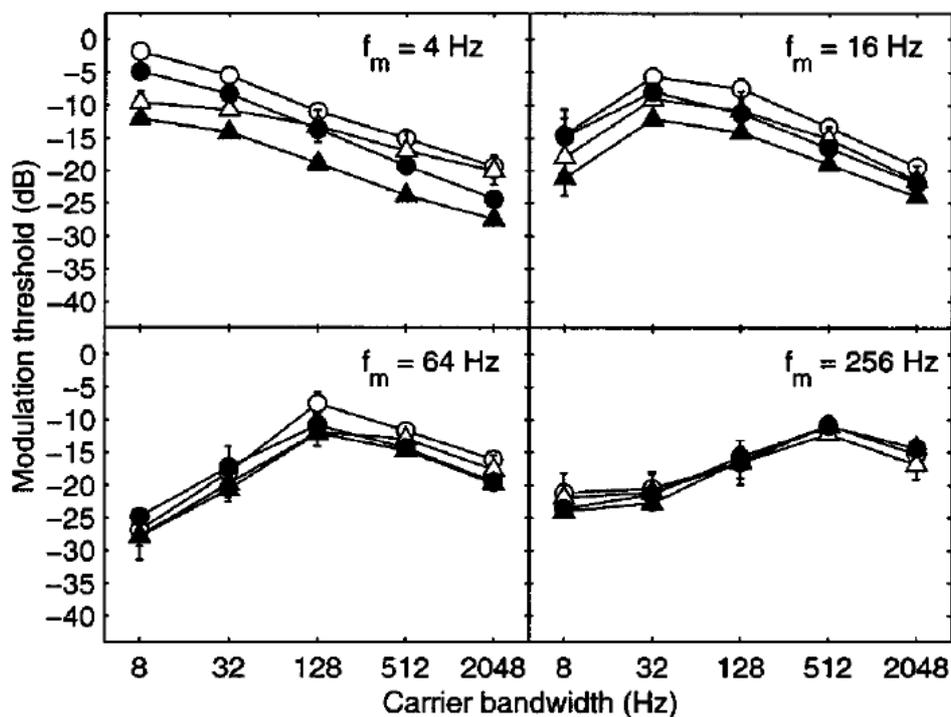


Figure 73. Seuils de détection d'une AM sinusoïdale simulés par le modèle EPSM (symboles pleins) Les résultats sont comparés aux données empiriques (symboles ouverts). Les triangles correspondent aux données mesurées en condition « masque gelé » et les cercles aux données mesurées en condition « masque aléatoire ». Extrait d'Ewert et Dau (2004).

Ewert et Dau (2004) ont tenté de simuler leurs données de détection d'AM obtenues en condition « masque gelé » et « masque aléatoire » (cf. figure 72 plus haut) à l'aide de différentes versions du modèle du « banc de filtres de modulation ». Le modèle « EPSM » (pour : *Modèle de spectre de puissance d'enveloppe*) élaboré par Ewert et Dau (2000) contient quatre étapes : *filtrage cochléaire, extraction d'enveloppe par filtrage passe-bande centré sur la fréquence de modulation du signal cible, calcul de puissance en sortie de chaque filtre de modulation passe-bande et prise de décision par comparaison de la puissance entre intervalles d'observation*. Les limitations de la résolution en intensité - et donc, de performance dans la tâche de détection d'AM - sont simulées par l'ajout d'un bruit interne (un bruit qualifié d'« additif ») appliqué indépendamment en sortie de chaque filtre et de variance constante (sa variance est donc indépendante du signal). A cela s'ajoute un bruit « multiplicatif » dont la variance est proportionnelle à la puissance calculée en sortie des filtres de modulation (sa variance augmente avec la puissance du signal). Ce bruit multiplicatif est nécessaire afin de reproduire la loi de Weber observée en discrimination de profondeur d'AM. Une autre version de ce modèle, dénommée *PEMO* ou « modèle de perception » (Dau et al., 1997), ne prend pas en compte de bruit multiplicatif (ce modèle n'utilise donc qu'un bruit additif), et l'étape de décision correspond à un mécanisme d'appariement de gabarit. Ici, le modèle calcule la fonction de corrélation croisée entre l'enveloppe extraite par les filtres de modulation (et bruitée par le bruit additif) et une représentation interne ou gabarit stocké en mémoire. Les résultats des simulations du modèle EPSM sont représentés en figure 73 ci-dessus. Ceux du modèle PEMO sont représentés en figure 74 ci-dessous.

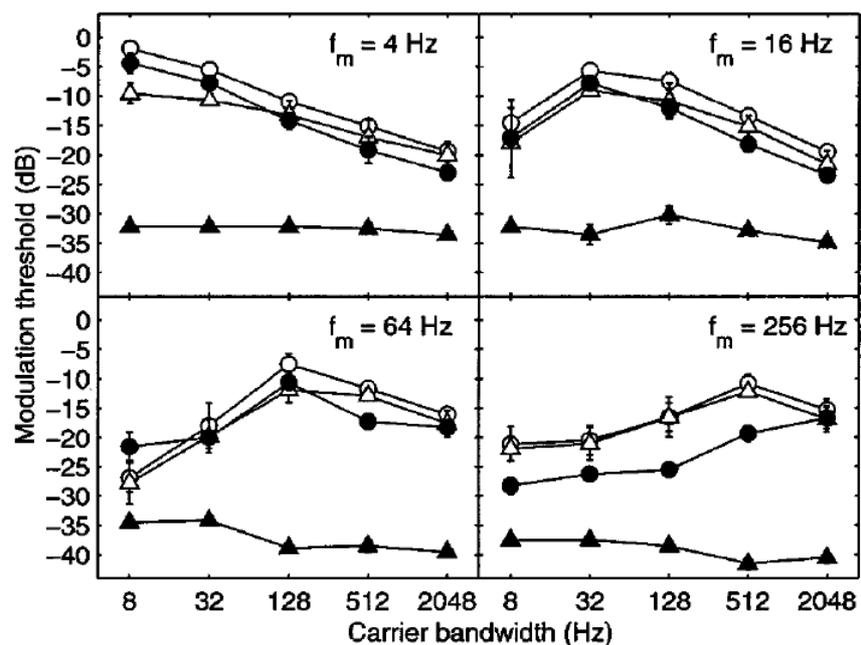


Figure 74. Seuils de détection d'une AM sinusoïdale simulés par le modèle PEMO (symboles pleins) Les résultats sont comparés aux données empiriques (symboles ouverts). Les triangles correspondent aux données mesurées en condition « masque gelé » et les cercles aux données mesurées en condition « masque aléatoire ». Extrait d'Ewert et Dau (2004).

Les simulations obtenues avec le modèle EPSM sont très proches des résultats empiriques (les seuils sont similaires en condition « masque aléatoire » et « masque gelé »), à l'exception des conditions où les fluctuations d'amplitude du masque et de la cible sont lentes. Les simulations obtenues avec le modèle PEMO sont très proches des résultats empiriques mais ce uniquement en condition « masque aléatoire ». En condition « masque gelé », les seuils simulés sont constants et bien trop bons : ils atteignent la valeur plancher fixée par le bruit interne additif. Par contre, le modèle PEMO simule correctement le fait que les seuils sont meilleurs en condition « masque gelé » qu'en condition « masque aléatoire ». Le succès du modèle EPSM est lié au fait qu'il utilise un bruit interne additif et un bruit interne multiplicatif, ce dernier n'étant pas incorporé dans le modèle PEMO. L'incapacité du modèle EPSM à simuler l'effet de démasquage observé lorsque les fluctuations d'amplitude du masque et de la cible sont relativement lentes est lié au fait que ce modèle n'a pas de mémoire. *A contrario*, le modèle PEMO, parce qu'il mémorise un gabarit du signal, peut s'affranchir de l'influence délétère du masque lors de la prise de décision par corrélation croisée entre gabarit et sortie de filtres de modulation.

I.4.d. Effets du bruit interne en perception auditive et visuelle

Les méthodes d'estimation du bruit interne ont largement été employées dans des tâches auditives et visuelles afin de répondre à de nombreuses questions que nous abordons dans cette section en faisant le parallèle entre quelques études. Ces méthodes se sont avérées utiles afin de *caractériser* finement le développement, l'apprentissage perceptif, les effets du vieillissement et certaines pathologies sensorielles et/ou cognitives en termes de bruit interne et d'efficacité d'échantillonnage. Ces mêmes méthodes ont également permis de *dissocier* les mécanismes sensoriels et cognitifs impliqués dans des tâches perceptives variées.

I.4.d.i. Trajectoires développementales

En modalité visuelle, un grand nombre d'études démontre que la **perception du mouvement** s'améliore progressivement au cours du développement (Bogfjellmo et al., 2014), arrivant à maturité entre 7 ans et 12 ans (Elleberg et al., 1999; Parrish et al., 2005). La sensibilité aux patterns plus complexes de mouvement global tels que l'expansion et la rotation se développe plus tard chez les enfants (Baumberger & Fluckiger, 2004). Des changements sont observés jusqu'à l'adolescence, ceci dépendant du stimulus et des paramètres utilisés (Falkenberg et al., 2010). L'utilisation de paradigmes psychophysiques décrits dans la section précédente est utilisée afin de déterminer dans quelle mesure ces trajectoires développementales résultent d'une évolution du **bruit interne**, de l'**efficacité d'échantillonnage**, ou des deux (Falkenberg et al., 2014).

Afin de caractériser finement la maturation des mécanismes impliqués dans la perception globale du mouvement, Bogfjellmo et al. (2014) mesurent l'*efficacité d'échantillonnage* et le

bruit interne dans une tâche de discrimination de mouvement à l'aide du **paradigme de bruit équivalent**. Dans ce contexte, la variabilité externe constitue une variabilité ajoutée dans chaque direction prise par les points durant la tâche de discrimination (Dakin et al., 2005). Cent trois participants âgés entre 6 et 17 ans sont testés à deux vitesses de déplacement différentes égales à 2.8 et 9.8 degrés par seconde. La procédure expérimentale employée consiste en une tâche à un intervalle dans laquelle deux types de signaux sont présentés simultanément. Chaque observateur fixe un point central coloré parmi un ensemble de points en mouvement et précise si la direction du mouvement global des points est dans le sens des aiguilles d'une montre ou dans le sens inverse. Le seuil de discrimination de la direction permettant de déduire l'efficacité d'échantillonnage et le bruit interne à partir de l'équation générale décrite précédemment. Ce seuil correspond au plus petit angle requis pour discriminer la direction du mouvement dans 75 % des essais en l'absence de bruit directionnel (0 degrés).

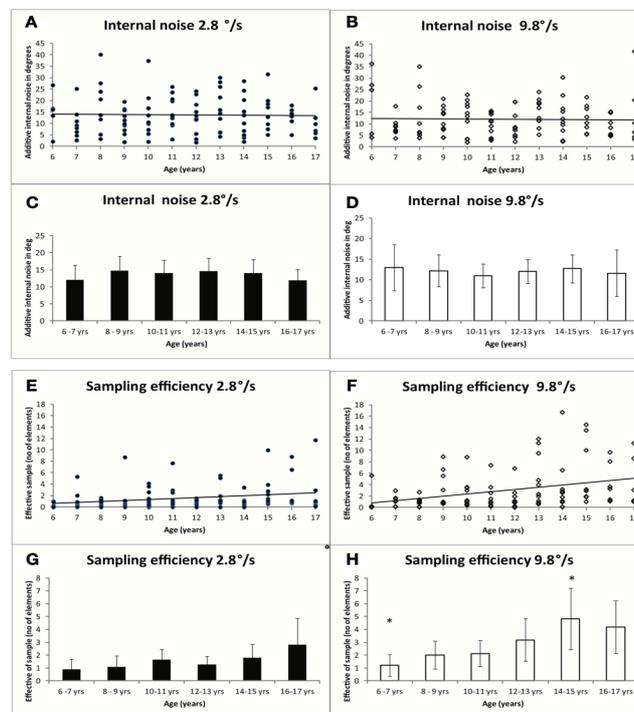


Figure 75. Évolution du bruit interne (figures A à D) et de l'efficacité d'échantillonnage (E à H) en fonction de l'âge à deux vitesses de déplacement des points différentes. Les figures A,C,E,G présentent les données obtenues pour une vitesse de 2.8 degrés par seconde tandis que les figures B,D,F,H présentent les résultats pour une vitesse de déplacement égale à 9.8 degrés par seconde. Les figures A,B,E,F représentent les données individuelles tandis que les figures C,D,G,H représentent les données moyennes. Les barres d'erreurs représentent les intervalles de confiance de +/- 95 %. Extrait de Bogfjellmo et al. (2004).

Les résultats présentés en figure 75 ci-dessus révèlent que durant le développement, la capacité à discriminer la direction d'un mouvement global se développe progressivement et que les enfants les plus jeunes ont besoin d'un angle plus grand pour discriminer la direction du mouvement. L'analyse des résultats démontre que l'amélioration de la discrimination avec

l'âge résulte d'une **amélioration de l'efficacité d'échantillonnage** mais **pas d'une réduction du bruit interne** et que la performance est meilleure lorsque la vitesse de déplacement est plus rapide (9.8 degrés par seconde).

Ces résultats sont compatibles avec ceux d'Hadad et al. (2011) qui démontrent que la sensibilité à la direction du mouvement global reste immature jusqu'à l'âge de 12 ans, et ce pour des vitesses de déplacement de 4 degrés par seconde et 18 degrés par seconde. Ils corroborent également ceux de Banton et al. (2001) qui suggèrent que si la bande passante des mécanismes de détection du mouvement (sélectifs directionnellement) se rétrécit avec l'âge, la sensibilité à la discrimination de la direction augmente. Une bande passante plus large signifie que la précision de la discrimination de la direction est pauvre car des petites différences de direction conduiront à des réponses non distinguables de la part d'un détecteur (Banton et al., 2001), rendant l'observateur incertain vis-à-vis de la direction locale d'éléments individuels dans le stimulus.

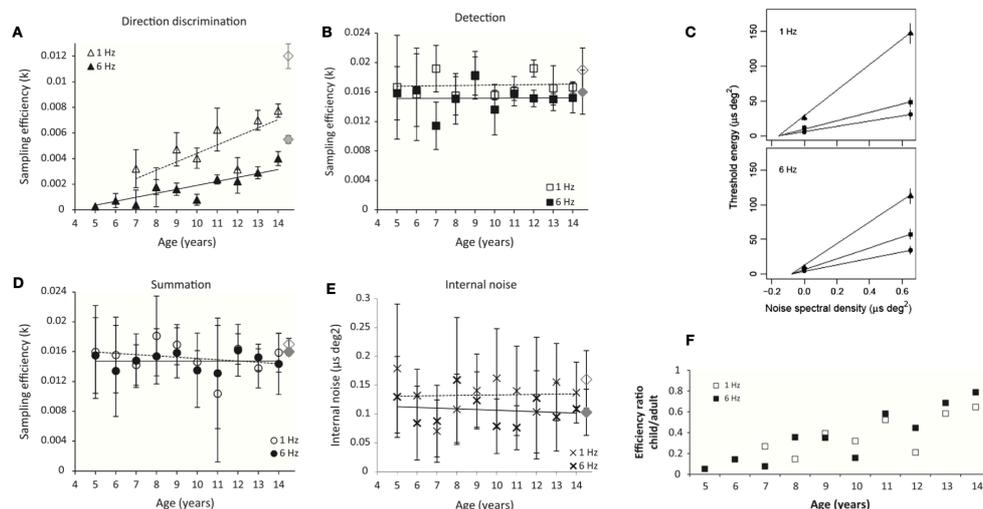


Figure 76. Estimations de l'efficacité d'échantillonnage moyenne pour (A) la discrimination de la direction (triangles) ; (B) la détection (carrés) et (D) la sommation (cercles), représentée en fonction de l'âge et de la vitesse. Les diamants gris dans tous les panneaux indiquent les valeurs moyennes pour les adultes. La figure (E) présente les estimations du bruit interne pour les trois tâches combinées en fonction de l'âge et de la vitesse. Les symboles ouverts représentent les valeurs moyennes à 1 Hz et les symboles pleins les valeurs moyennes à 6 Hz. Les barres d'erreur indiquent les intervalles de confiance à 95 %. Les lignes ajustées correspondent à des régressions linéaires estimées par les moindres carrés. A 1 Hz, l'efficacité d'échantillonnage pour la discrimination n'a pas pu être estimée avant l'âge de 7 ans. La figure (C) représente le seuil de contraste moyen (n=15 adultes) pour les tâches de détection d'un réseau allant vers le haut (carrés), de la somme d'un réseau allant vers le haut ou vers le bas (cercles) et pour la discrimination d'un réseau allant vers le haut et vers le bas (triangles) en fonction de la densité spectrale de puissance du bruit blanc Gaussien ajouté. La figure F représente les rapports d'efficacité relative des enfants par rapport aux adultes pour la discrimination de direction. Extrait de Falkenberg et al. (2014).

Falkenberg et al. (2014) utilisent également le **paradigme du bruit équivalent** pour étudier le développement de la **perception du mouvement** sur une large cohorte de 166 enfants âgés entre 5 et 14 ans et un plus petit groupe d'adultes (n=15 adultes) d'âge moyen de 30 ans.

Trois tâches psychophysiques sont effectuées dans un ordre aléatoire sur chaque participant : i) une tâche de **détection** dans laquelle le stimulus représente un réseau de points se déplaçant vers le haut ou un espace visuel vide, ii) une tâche de **sommation** où un réseau de points clignotant en contre phase (somme de deux réseaux en sens inverse avec la même probabilité d'occurrence) est présenté aux sujets, et iii) de **discrimination** (réseau de points se déplaçant soit vers le haut soit vers le bas avec la même probabilité d'occurrence et ce, à deux vitesses : 1 Hz et 6 Hz). La figure 76 présente l'ensemble des résultats obtenus pour les trois tâches ainsi que les estimations de bruit interne et d'efficacité d'échantillonnage. Les résultats montrent que, même en l'absence de variabilité externe, la discrimination de la direction est plus difficile que la détection, surtout à 1 Hz. De plus, seule la discrimination de la direction s'améliore avec l'âge et continue de s'améliorer au-delà de 14 ans.

La méthode du **bruit équivalent** permet de déterminer les causes de cette amélioration. La figure 76 illustre les seuils de contraste en présence et en l'absence de bruit externe pour les observateurs adultes. Quand le bruit externe est ajouté aux stimuli, les seuils augmentent, et ce surtout pour la tâche de discrimination de direction. Le modèle de bruit équivalent décrit une augmentation linéaire des seuils de discrimination avec le niveau de bruit externe et ce, pour chaque condition. Les estimations d'efficacité d'échantillonnage s'améliorent durant l'enfance pour les deux vitesses (1 et 6 Hz) mais de manière plus rapide lorsque la vitesse est lente (la pente étant plus raide). Le bruit interne, quant à lui, n'est pas affecté par la nature de la tâche ou par l'âge. En revanche, les enfants sont beaucoup moins efficaces que les adultes lorsqu'il s'agit d'extraire l'information utile et ce quelle que soit la vitesse. L'efficacité d'échantillonnage pour la perception du mouvement s'améliore donc pendant l'enfance. Même à l'âge de 14 ans, l'efficacité reste inférieure à celle des adultes, contrairement aux capacités de détection et de sommation où l'on n'observe pas de différences significatives avec l'âge. Ainsi, la détection de la présence de patterns visuels en mouvement et la perception de la direction du mouvement sont limitées à la fois par le bruit interne et par l'efficacité d'échantillonnage, mais leurs trajectoires développementales respectives sont bel et bien différentes. Tandis que la discrimination de la direction n'est pas encore stabilisée à 14 ans, la détection est déjà au niveau adulte à partir de l'âge de 5 ans. L'utilisation de la méthode du bruit équivalent démontre donc que ***l'amélioration de la performance avec l'âge pour la discrimination de la direction du mouvement est due à une amélioration de l'efficacité d'échantillonnage sans modification du bruit interne***. Ces différences de trajectoires développementales pour la discrimination de la direction et pour la détection du mouvement suggèrent que ces deux capacités font appel à des mécanismes distincts dont les périodes de maturation diffèrent.

En modalité auditive, de nombreuses études démontrent que les adultes sont plus performants que les enfants dans plusieurs tâches comme par exemple, la détection d'un son pur (Fior, 1972 ; Maxon & Hochberg, 1982 ; Allen & Wightman, 1994), la discrimination fréquentielle (Maxon & Hochberg, 1982 ; Jensen & Neff, 1993), la détection d'intervalles

(Wightman et al., 1989 ; Trehub et al., 1995), la discrimination d'intensité (Fior, 1972 ; Maxon & Hochberg, 1982 ; Jensen & Neff, 1993) ou des tâches auditives plus complexes comme la reconnaissance de la parole dans des conditions d'écoute difficiles (Elliott, 1979 ; Nábělek & Robinson, 1982) et de masquage informationnel (Allen & Wightman, 1995 ; Hall et al., 2005).

L'interprétation de tels résultats n'est pas aisée car la performance peut être affectée par des facteurs d'origine non-sensorielle plus difficiles à identifier (Hirsh & Watson, 1996 ; Talcott et al., 2002 ; Sutcliffe & Bishop, 2005 ; Buss et al., 2009) : des facteurs **cognitifs** (un bruit interne central probablement lié aux fluctuations aléatoires de l'attention, une instabilité du critère de décision ou encore des variations dans les capacités mnésiques) (e.g., Allen et al., 1998 ; Wightman et al., 1989 ; Allen & Nelles, 1996), ou des facteurs plus **spécifiques au traitement auditif** comme une incapacité à combiner de manière optimale l'information sensorielle (Allen et al., 1989) ; une réduction de l'efficacité d'écoute sélective (Hall & Grose, 1991 ; 1994) ; une diminution de l'attention temporelle sélective (Sutcliffe & Bishop, 2005) ou encore une combinaison relativement inefficace de l'information distribuée dans le temps (Schneider & Trehub, 1992).

Schneider et al. (1989, 1992) suggèrent que ces différences développementales sont associées à une augmentation de bruit interne (défini ici comme une *augmentation de la variabilité dans la représentation neuronale de l'entrée sensorielle*). Ces auteurs prédisent qu'un bruit interne important chez l'enfant devrait être associé des fonctions psychométriques pentues dans une **tâche de détection d'énergie masquée** car l'énergie en sortie du filtre auditif centré sur la fréquence centrale du signal devrait être plus variable. Schneider et al. (1989) testent cette hypothèse en estimant la performance de détection de sons purs masqués par un bruit dont la bande passante est d'1/3 d'octave sur des enfants âgés entre 6 mois et des adultes de 20 ans. Contrairement à cette hypothèse, la pente des fonctions psychométriques n'est pas différente entre les enfants et les adultes.

Allen et Wightman (1994) répliquent et étendent cette première investigation en mesurant les fonctions psychométriques individuelles de 18 enfants âgés entre 3 et 5 ans et de 14 adultes âgés entre 19 et 37 ans pour une **tâche de détection d'un son pur** de fréquence égale à 501, 1000 et 2818 Hz masqué par un bruit. Les résultats obtenus avec un paradigme en choix forcé indiquent que les seuils de détection des enfants sont significativement plus élevés que ceux des adultes. Par ailleurs, les **pentés des fonctions psychométriques** - quantifiées ici comme la variation de niveau nécessaire pour produire une augmentation de performance de 20 points - sont plus douces chez les enfants que chez les adultes (voir figure 77 ci-dessous). **La diminution significative de la pente des fonctions psychométriques chez les enfants s'explique par une augmentation de bruit interne pour la détection auditive masquée.** Ceci joue probablement un rôle dans l'élévation des seuils de détection masqués des enfants. Cependant, les auteurs notent de fortes différences inter-individuelles dans les performances

de détection, résultant probablement de différences dans les stratégies d'écoute (à savoir, des différences de contrôle de l'attention sélective liées à l'âge).

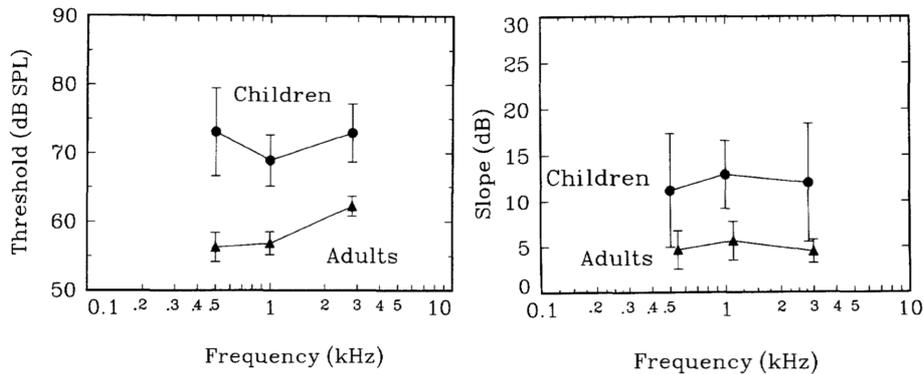


Figure 77. Seuils de détection moyens (à gauche) et pentes moyennes de la fonction psychométrique (à droite) pour la détection d'un son pur de fréquence fixe présenté dans un bruit masquant, en fonction de la fréquence du son pur. Les cercles indiquent les résultats des enfants tandis que les triangles indiquent les résultats des adultes. Les barres d'erreur indiquent les limites de confiance à 95 %. Le paramètre de « pente » (slope) correspond à la variation de niveau du son nécessaire pour produire une augmentation de performance de 20 points. Une valeur plus élevée signifie que pente de la fonction psychométrique est plus douce. Extrait de Allen et Wightman (1994).

Buss et coll. mettent en œuvre un programme de recherche exhaustif permettant d'évaluer l'évolution du bruit interne dans le **traitement auditif de l'information d'intensité**. Ce programme combine des protocoles inspirés de la méthode du bruit équivalent et des estimations basées sur la pente de la fonction psychométrique. Dans une première série d'études, Buss et al. (2006) testent l'hypothèse selon laquelle les effets développementaux observés dans des tâches psychoacoustiques dépendant de l'intensité des sons s'expliquent en partie par des changements de bruit interne en manipulant la variabilité externe chez des enfants et des adultes.

Dans une première expérience basée sur un protocole de détection, les auteurs testent l'hypothèse suivante : si le bruit interne est le facteur limitant la détection d'un son pur en présence d'un bruit à bande étroite, les enfants devraient moins bénéficier d'une réduction de l'amplitude des fluctuations du masque que les adultes, pour lesquels le bruit externe joue un rôle plus important sur la performance. En d'autres termes, **des niveaux élevés de bruit interne devraient réduire la capacité des enfants à utiliser efficacement les indices auditifs pertinents permettant de réaliser la tâche de détection dans un bruit dont les fluctuations sont faibles par rapport à un bruit Gaussien dont les fluctuations sont plus fortes**. Huit enfants âgés entre 5 et 10.5 ans (âge moyen : 7 ans) et dix adultes (âge moyen : 28 ans) sont testés. Les seuils de détection du son pur dans le bruit sont obtenus à l'aide d'une procédure adaptative en choix forcé. Au total, six conditions expérimentales sont testées (voir figure 78 ci-dessous pour plus de détails sur les conditions). Les résultats de cette première expérience indiquent que globalement, les seuils des enfants sont moins bons (plus élevés) que ceux des adultes. Par ailleurs, les adultes bénéficient d'une diminution de l'amplitude des fluctuations du bruit externe (leurs seuils s'améliorent) alors que les seuils des enfants sont comparables

pour les deux conditions de bruit externe. **Ce résultat est cohérent avec l'hypothèse selon laquelle le bruit interne en détection masquée est plus élevé chez les enfants que chez les adultes**, limitant la capacité des premiers à bénéficier d'une diminution de l'amplitude du bruit externe.

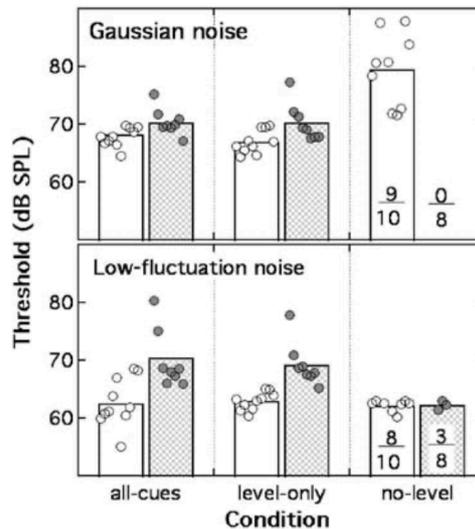


Figure 78. Seuils de détection moyens de l'expérience 1 représentés pour le bruit Gaussien (panneau supérieur) et le bruit dont les fluctuations sont plus faibles que celles du bruit Gaussien (Low-fluctuation noise ; panneau inférieur). Les barres indiquent les données moyennes des enfants (barres hachurées) et des adultes (barres vides), et les cercles associés à chaque barre indiquent les données individuelles. Les conditions du stimulus sont indiquées le long de l'axe des abscisses : i) « all-cues » : le signal est un son pur introduit dans la partie centrale du stimulus ; les interactions entre le signal et le masque produisent deux types d'indices : des indices de niveau sonore et des indices d'enveloppe temporelle; ii) « level-only » : le signal consiste en une augmentation du niveau sonore du masque, définie comme l'augmentation qui aurait été produite si un son pur avait été ajouté. Dans cette condition, il n'y a pas de changement dans les statistiques de l'enveloppe temporelle pendant la présentation du signal; iii) « no-level » : le niveau global de la partie centrale du stimulus a été réduit après l'ajout du signal de sorte qu'un changement dans les statistiques de l'enveloppe temporelle au cours de la présentation du signal indique la présence du signal, mais sans augmentation du niveau sonore. Les ratios à droite de chaque panneau indiquent la proportion d'auditeurs contribuant à chaque moyenne dans les conditions sans indice de niveau : les données individuelles sont omises de la moyenne si elles sont considérées comme indiscernables de la performance plafond. Extrait de Buss et al. (2006).

Dans une deuxième expérience basée sur un **protocole de discrimination**, le niveau sonore est varié aléatoirement (roving) au sein des essais et entre essais. Ici, 15 enfants âgés entre 5 et 10 ans (âge moyen : 8 ans) sont testés et 12 adultes âgés entre 17 et 50 ans (âge moyen de 29 ans) dans une tâche adaptative en choix forcé estimant le seuil de discrimination d'intensité d'un son pur (un signal ne présentant pas de variabilité au sein de l'intervalle d'observation) ou d'une bande de bruit étroite de 20 Hz (un signal présentant une variabilité au sein de l'intervalle d'observation). Le niveau sonore est par ailleurs varié aléatoirement entre essais afin d'augmenter le bruit externe dans le domaine de l'intensité.

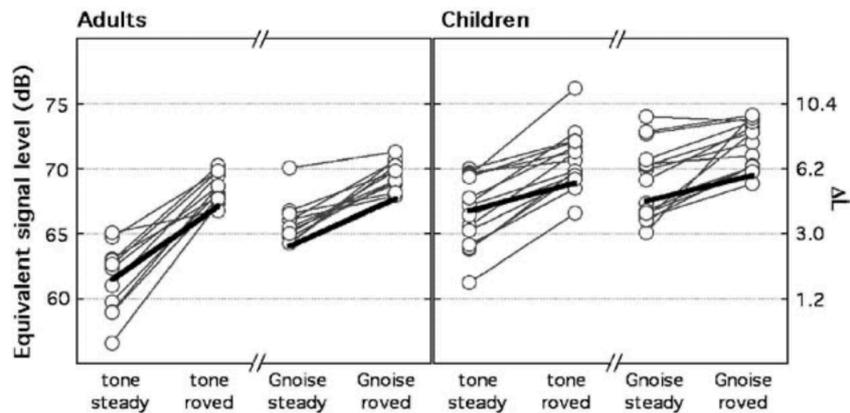


Figure 79. Données individuelles de l'expérience 2. Les ordonnées à droite montrent les seuils de discrimination ΔL en dB. Les panneaux de gauche et de droite montrent respectivement les résultats des adultes et des enfants. Les conditions de stimulation sont indiquées en abscisse : i) « tone steady » : son pur sans variabilité dans le niveau sonore ; « tone roved » : son pur dont le niveau sonore varie aléatoirement entre essais ; « Gnoise-steady » : bruit Gaussien, sans variabilité dans le niveau sonore ; « Gnoise-roved » : bruit Gaussien dont le niveau sonore varie aléatoirement entre essais. Les lignes épaisses indiquent les ajustements aux données. Extrait de Buss et al. (2006).

Les résultats de l'expérience de discrimination sont présentés en figure 79 ci-dessus. Ils indiquent que la variabilité intra- et inter-essais du stimulus a moins d'effet sur les seuils de discrimination d'intensité des enfants que sur ceux des adultes, conformément à l'hypothèse selon laquelle **les enfants présentent un bruit interne plus élevé que les adultes pour la discrimination d'intensité**. En modélisant la sensibilité d' comme le rapport entre le seuil de discrimination moyen et la racine carré de la somme du carré de la variabilité externe et interne, il est possible d'estimer le bruit interne associé à la tâche. Pour les adultes, l'écart-type du bruit interne est de l'ordre de 0.99 unités tandis que chez les enfants, l'écart-type du bruit interne est beaucoup plus important et de l'ordre de 2.46 unités.

Dans une étude complémentaire, Buss et al. (2009) testent les conclusions précédentes en mesurant la *pente de la fonction psychométrique pour la discrimination d'intensité* chez un nouveau groupe d'enfants. Cette étude inclut 16 participants jeunes âgés entre 4.9 et 9.4 ans (âge moyen : 7 ans) et 11 participants adultes âgés entre 19 et 55 ans (âge moyen : 27 ans). La tâche de discrimination est basée sur un paradigme en choix forcé. Le signal cible (un son pur de fréquence égale à 500 Hz) est présenté dans un ordre aléatoire d'essai en essai. Dans l'intervalle de comparaison, le son pur est présenté à une intensité de 65 dB SPL tandis que dans l'intervalle cible, celui-ci est plus intense. Le participant sélectionne d'essai en essai, l'intervalle contenant le signal le plus fort. La procédure est effectuée en deux étapes : i) dans la première, une procédure adaptative est utilisée pour estimer le seuil (m) et l'écart-type standard (s) des inversions. Ces estimations sont retenues pour fixer les conditions de test de la deuxième étape ; ii) dans la seconde, la méthode des stimuli constants est utilisée afin de mesurer la performance à 5 niveaux d'intensités fixés à : $[m-2s ; m-s ; m ; m+s ; m+2s]$ sur 10 blocs de 40 essais.

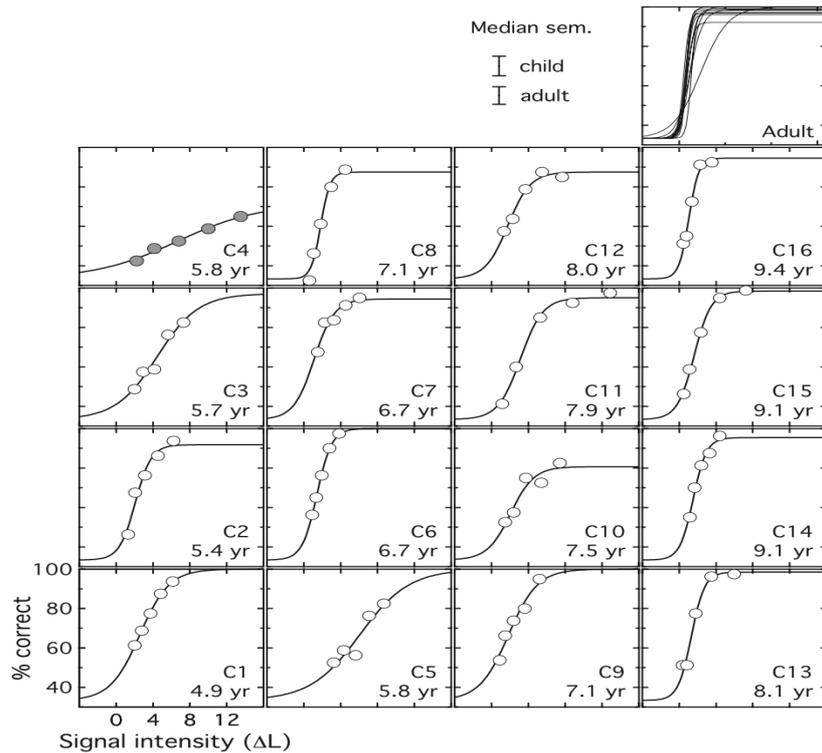


Figure 80. Fonctions psychométriques individuelles pour la discrimination d'intensité d'un son pur de 500 Hz. Chaque fonction psychométrique indique l'évolution de la performance (% correct) en fonction de l'intensité du signal exprimée en unités ΔL . Les cercles indiquent les données empiriques et les lignes représentent les ajustements aux données. L'âge de chaque observateur est indiqué dans chaque panneau, et les panneaux sont disposés de telle sorte que l'âge de l'enfant augmente de gauche à droite. Tous les ajustements pour les adultes sont présentés dans le panneau supérieur droit. Extrait de Buss et al. (2009).

La méthode des moindres carrés est utilisée afin d'ajuster les données de performance (le pourcentage de discrimination correcte) de chaque observateur à l'aide d'une fonction logit du type :

$$p(x) = \frac{n}{3} + (1 - n) \left[\frac{1}{3} + \frac{\frac{2}{3}}{\left(1 + e^{-\frac{x-\mu}{k}}\right)} \right]$$

où p représente le pourcentage correct compris entre 0 et 1, n est la probabilité d'inattention dans un essai donné, x est le niveau du signal exprimé en pas ΔL , μ est le point médian de la fonction, k est la pente, sachant que des valeurs plus grandes représentant des fonctions psychométriques moins inclinées (Buss et al., 2009). La pente de la fonction psychométrique (k) exprimée en unités de ΔL est utilisée pour estimer le bruit interne (σ_{int}) et émettre des prédictions de seuils.

Les estimations de bruit interne sont de 0.76 dB pour les adultes et de 1.91 dB pour les enfants. Ces estimations sont conformes à celles réalisées par Buss et al. (2006). Les seuils prédits à partir des estimations de bruit interne s'ajustent assez bien aux données empiriques malgré les fortes différences individuelles observées entre enfants (*cf.*, Wightman & Allen, 1992). Ces

résultats corroborent l'hypothèse selon laquelle **le bruit interne est plus important chez les enfants que chez les adultes, ceci expliquant pourquoi les seuils de discrimination d'intensité s'améliorent avec l'âge**. La stabilité relative de la performance des enfants suggère par ailleurs que cet effet de l'âge n'est pas lié à une plus grande variabilité dans la stratégie d'écoute ou la motivation.

Les estimations de bruit interne peuvent également s'avérer utiles afin de *dissocier les trajectoires développementales* de plusieurs capacités perceptives, et ainsi *préciser et distinguer les mécanismes sous-jacents* pour chaque capacité. Banai et al. (2011) estiment les trajectoires développementales pour la **détection d'AM et de FM** en évaluant deux indices de performance sur un même groupe d'auditeurs (des enfants de 8-10 ans, 11-12 ans et des adultes jeunes): le seuil de détection moyen (performance moyenne) et la variabilité de la performance intra-individuelle autour du seuil (« performance consistency »). Cet indice représente la variabilité intra-individuelle estimée par l'écart-type calculé sur les inversions d'une seule série ou bien sur l'ensemble des séries de l'escalier psychophysique. *Comme indiqué plus haut, cet indice est directement relié au bruit interne dans la tâche de détection de modulation*. Les résultats de cette étude indiquent que pour chaque enfant, les seuils de détection moyens en détection d'AM et de FM ne sont pas corrélés (*cf.* Fig.31). Il en va de même pour la cohérence de la performance. Dans l'ensemble, ces données suggèrent que les trajectoires développementales pour les tâches de détection d'AM et de FM sont partiellement distinctes et que par conséquent, **différents mécanismes (associés à différentes sources de bruit interne) sont responsables de la maturation de la capacité à détecter l'AM et la FM**. D'autre part, ces données indiquent que la capacité à détecter les modulations se développe indépendamment de la capacité à les percevoir de manière cohérente.

Plus récemment, Buss et al. (2019) mesurent les seuils de **détection d'AM** sur des enfants âgés entre 5 et 11 ans et des adultes en utilisant une porteuse de 4300 Hz modulée en amplitude à 3 cadences (16, 64 et 256 Hz). Les résultats de cette étude indiquent que les seuils de détection d'AM tendent à s'améliorer avec l'âge, l'effet développemental étant plus important à 64 et 256 Hz qu'à 16 Hz. Dans l'ensemble, ces données indiquent que certains aspects du traitement d'AM arrivent à maturité pendant l'enfance, mais cette étude ne permet pas de préciser si les changements observés reflètent une évolution du bruit interne ou de l'efficacité d'échantillonnage. Cabrera et al. (2019) complètent l'étude de Buss et al. (2019) en évaluant la sensibilité à l'AM sur une cohorte beaucoup plus large de 80 enfants répartis en trois groupes d'âge différents (5-6 ans ; 7-8 ans ; 10-11 ans) et un groupe de 32 adultes. La tâche (adaptative en choix forcé) consiste à détecter une AM sinusoïdale appliquée à une porteuse correspondant soit à un son pur, soit à une bande de bruit étroite dont la variance est faible ou forte. Ces manipulations de porteuse permettent ainsi de faire varier le niveau du bruit externe dans le domaine de l'enveloppe temporelle. Un modèle computationnel basé sur un banc de filtres cochléaires, un banc de filtres de modulation, un bruit interne Gaussien additif et un mécanisme d'appariement de gabarit est ensuite utilisé

afin de simuler les données empiriques de détection d'AM. Trois éléments sont manipulés dans le modèle : 1) le **facteur de qualité (Q)** des filtres de modulation, contrôlant la sélectivité fréquentielle (centrale) à l'AM ; 2) l'**écart-type du bruit interne additif ($\sigma_{int,add}$)** ; 3) la **stratégie de décision**, à savoir la capacité à utiliser l'information de canaux cochléaires hors-fréquence (stratégie optimale : utilisation des sorties de tous les filtres cochléaires ; stratégie sous-optimale : utilisation du canal accordé uniquement sur la fréquence porteuse).

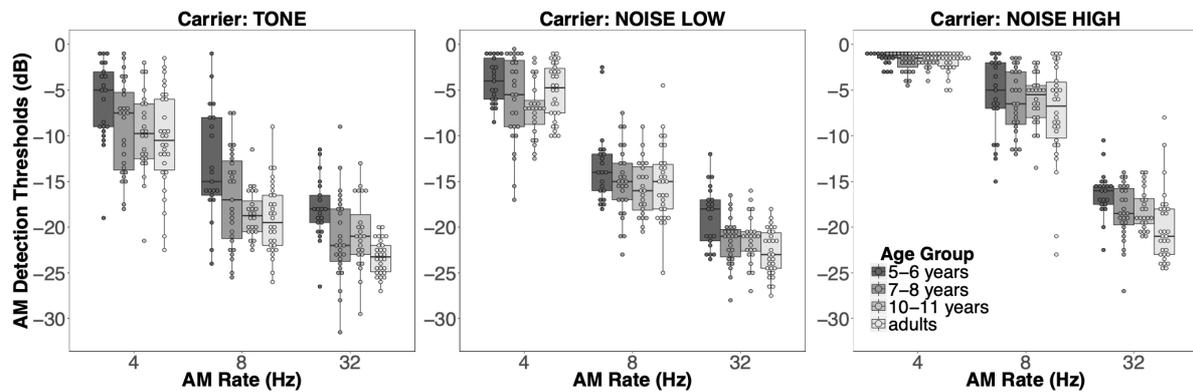


Figure 81. Seuils de détection d'AM en fonction de la cadence de modulation exprimée en Hertz. Les seuils sont présentés par groupe d'âge (dégradé de gris) et par condition de porteuse (TONE, NOISE-LOW, et NOISE-HIGH). Les seuils de détection sont exprimés sous la forme $20\log(m/100)$. La condition « TONE » correspond au cas où la porteuse est un son pur de 1027 Hz. La condition « NOISE » correspond au cas où la porteuse est un bruit à bande étroite (4 Hz) présentant des fluctuations faibles « NOISE LOW » ou fortes « NOISE HIGH ». Extrait de Cabrera et al. (2019).

Conformément aux études précédentes (e.g., Hall & Grose, 1994), les résultats présentés en figure 81 indiquent que les seuils de détection d'AM s'améliorent avec l'âge, et ce quelle que soit la condition testée. De plus, l'effet de la cadence d'AM et de la porteuse sur la détection d'AM est similaire pour tous les groupes d'âge : les seuils s'améliorent aux hautes cadences d'AM et ils se dégradent lorsque la porteuse présente des fluctuations aléatoires intrinsèques d'enveloppe (cette dégradation étant plus importante lorsque la variance du bruit porteuse est maximale). Ceci suggère que les facteurs purement sensoriels contraignant la détection d'AM (e.g., intégration temporelle et masquage de modulation) sont déjà à maturité à partir de 5 ans. Les résultats de l'étude par simulation indiquent que le meilleur ajustement des réponses du modèle computationnel aux données est obtenu lorsque **la magnitude du bruit interne est 10 fois plus grande chez les enfants les plus jeunes (5-6 ans) que chez les enfants les plus âgés (10-11 ans) et les adultes**. Ainsi, cette étude permet d'attribuer le développement de la sensibilité non masquée et masquée de l'AM à une **réduction progressive du bruit interne dans le domaine d'AM** avec l'âge, et non à une amélioration de sélectivité fréquentielle dans le domaine de l'AM ou à une optimisation de la stratégie d'écoute. Enfin, cette étude suggère que la diminution de la variance du bruit interne entre 5 et 11 ans est drastique !

I.4.d.ii. Apprentissage perceptif

L'apprentissage perceptif correspond à l'amélioration de la performance dans une tâche sensorielle résultant de la répétition de cette dernière (Jones et al., 2013). L'apprentissage perceptif est souvent spécifique aux stimuli présentés durant l'apprentissage (Fiorentini & Berardi, 1980 ; Ball & Sekuler, 1987 ; Poggio et al., 1992 ; Fahle & Morgan, 1996) indiquant que l'entraînement modifie les caractéristiques de réponses de certains neurones (Gold et al., 1999). D'autres mécanismes sous-jacents peuvent être liés à ce phénomène de plasticité pour une modalité sensorielle spécifique. En modalité visuelle, plusieurs études font l'hypothèse que l'entraînement améliore la discrimination en réduisant le bruit interne (Doshier & Lu, 1998) ou en améliorant l'efficacité d'encodage de l'entrée sensorielle (Gold et al., 1999). En modalité auditive, l'apprentissage perceptif conduirait à une réduction du bruit interne lors de la discrimination fréquentielle (Jones et al., 2013). La section suivante passe en revue certaines de ces études.

En modalité visuelle, l'apprentissage perceptif reflète une forme de plasticité dans le système visuel de l'adulte, notamment dans le cadre de l'apprentissage perceptif de la texture (Karni & Sagi, 1991) ou encore de l'orientation du mouvement (Fiorentini & Berardi, 1980 ; 1981; Fahle & Edelman, 1993; Saarinen & Levi, 1995). En évaluant l'apprentissage perceptif en présence d'une quantité variable de bruit externe et en répétant la tâche plusieurs fois, il est possible de préciser la nature de cette plasticité. En d'autres termes, un apprentissage perceptif combiné au paradigme du bruit équivalent et au paradigme de double-passe (Doshier & Lu, 1998 ; Lu & Doshier, 1998) permet d'identifier les mécanismes sous-jacents à l'apprentissage perceptif (*e.g.*, *exclusion du bruit externe*, *réduction du bruit interne additif*).

Doshier et Lu (1998) testent quatre participants dans une **tâche de discrimination de contraste** de Gabors en présence d'un bruit visuel. Ce dernier est tiré aléatoirement à partir d'une distribution Gaussienne parmi 8 niveaux de bruit externe et les seuils de discrimination de contraste sont estimés à l'aide d'une méthode adaptative ciblant 79.3 % et 70.3% de réponses correctes. Les résultats de cette étude permettent de caractériser la fonction « seuil versus contraste » attendue par le **paradigme du bruit équivalent** (Pelli, 1981 ; Ahumada & Watson, 1985) : à des niveaux élevés de bruit externe (où ce dernier représente le facteur limitant), le seuil de discrimination de contraste augmente tandis qu'aux niveaux faibles (où le bruit interne représente le facteur limitant), le seuil devient quasiment indépendant du bruit externe. Les résultats indiquent une amélioration de la performance de discrimination de l'orientation avec l'entraînement, quelle que soit la valeur du bruit externe (amélioration du seuil moyen d'un facteur 2.8 entre le premier et le dernier jour de l'expérience).

Un modèle d'observateur linéaire est utilisé afin de simuler les données empiriques. Ce modèle - illustré ci-dessous en figure 82 - est composé : i) d'un **filtre accordé** sur les caractéristiques physiques du signal (ce dernier décrit les caractéristiques de filtrage de

l'observateur dans une tâche spécifique et correspond à une combinaison pondérée d'entrées visuelles); ii) d'un **transducteur non linéaire** (traduisant les non-linéarités du système visuel) ; et de facteurs d'inefficacité de traitement, à savoir : iii) un **bruit interne multiplicatif** décrivant les conséquences des systèmes de contrôle du gain, puis iv) un **bruit interne additif** reflétant l'existence d'un seuil absolu en l'absence de bruit externe. Enfin, le modèle aboutit à une **règle de décision** basée sur le calcul d'une statistique donnée. Les résultats des différentes simulations - illustrés également en figure 82 - démontrent que **la réduction du bruit interne additif a pour effet d'améliorer les seuils de discrimination dans la région où le bruit externe est faible (cf. figure 82b) tandis que l'exclusion du bruit externe entraîne l'amélioration des seuils de discrimination dans la région où le bruit externe est élevé (cf. figure 82C).**

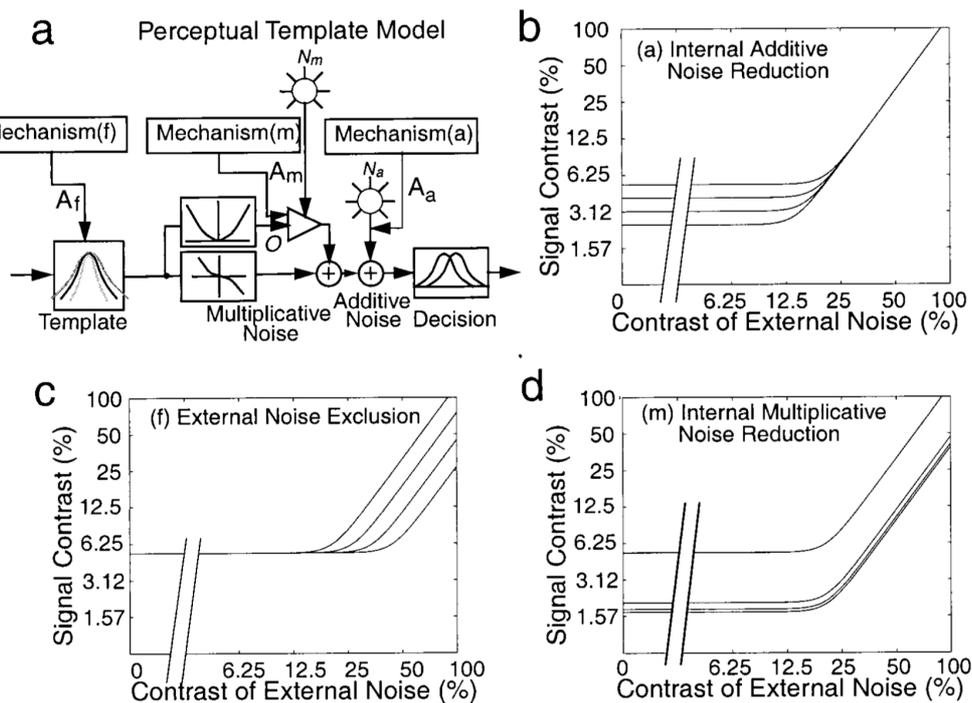


Figure 82. Simulations de la performance de discrimination de contraste pendant l'apprentissage perceptif. L'amélioration de la discrimination du signal est obtenue par i) réduction de bruit interne additif, ii) exclusion du bruit externe, ou iii) réduction de bruit interne multiplicatif. La figure (a) illustre les différentes étapes du modèle du gabarit perceptif (perceptual template model). La figure (b) illustre le fait que l'entraînement améliore uniquement la discrimination du stimulus dans la région limitée par le bruit interne additif (i.e., pour les valeurs de bruit externe faibles) lorsque seul le paramètre de bruit interne additif est modifié. La figure (c) illustre le fait que l'entraînement améliore la capacité à exclure le bruit externe, entraînant une amélioration des seuils aux niveaux de bruit externe élevés. La figure (d) illustre le fait que l'entraînement permet de réduire le bruit multiplicatif interne, entraînant une amélioration de la performance pour toutes les valeurs de bruit externe. Extrait de Doshier et Lu (1998).

Combinées aux données empiriques, les simulations démontrent que **l'apprentissage perceptif pour la discrimination de contraste résulte à la fois d'une réduction du bruit interne additif et de l'exclusion du bruit externe, visibles à partir de deux jours d'entraînement**. Par contre, elles ne corroborent pas l'hypothèse d'une réduction du bruit interne multiplicatif et d'une modification des non-linéarités (Lu & Doshier, 1998).

Gold et al. (1999) utilisent cette fois-ci le paradigme du **bruit équivalent** et le paradigme de **double-passe** afin de décrire les mécanismes impliqués dans la **discrimination de visages et de texture**. Dans cette étude, la tâche consiste à identifier un visage parmi dix visages non familiers présentés à chaque essai au participant. Ces signaux visuels sont masqués par un bruit Gaussien bidimensionnel statique dont la densité spectrale est variée. Dans cette tâche, une stratégie optimale consiste à réaliser une comparaison entre les dix modèles de visage et le stimulus bruité présenté à chaque essai. La performance d'un tel observateur est donc limitée par la quantité de bruit externe ajoutée au signal et par le processus d'appariement de gabarit. La tâche de discrimination de texture est implémentée de la même manière que la tâche précédente et elle consiste à identifier une texture de bruit Gaussien filtré en passe-bande parmi dix textures différentes. Les résultats de ces deux expériences démontrent que les seuils de discrimination de visage et de texture augmentent (se dégradent) avec le bruit externe et diminuent (s'améliorent) avec l'entraînement. L'efficacité et le bruit interne sont représentés en figure 83. Cette dernière indique que l'entraînement améliore l'efficacité (ce qui cohérent avec l'idée selon laquelle l'apprentissage permet d'extraire plus d'information dans le stimulus) mais il n'a pas d'impact sur le bruit interne. Cependant, la méthode du bruit équivalent suppose un bruit interne *additif*. Or, *une diminution du bruit interne multiplicatif a le même effet qu'une augmentation de l'efficacité d'échantillonnage*. Cette méthode présente donc une limite : il n'est pas possible de savoir si les effets d'amélioration associés à l'apprentissage perceptif sont liés à une réduction du bruit interne multiplicatif ou à une augmentation d'efficacité d'échantillonnage du signal. Le paradigme de double-passe qui estime une quantité globale de bruit interne (additif et multiplicatif) permet de répondre à cette question.

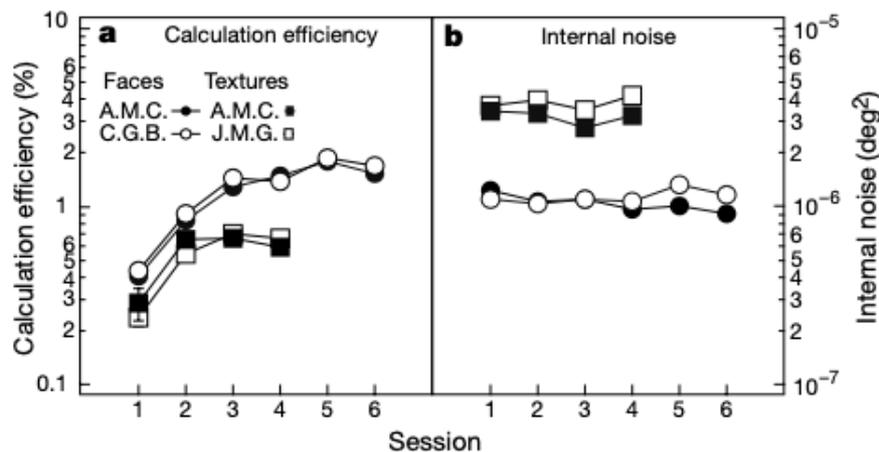


Figure 83. Estimations de l'efficacité (a) et du bruit interne (b) en fonction des sessions d'entraînement pour chaque participant (symboles ouverts/pleins). Les cercles représentent les données de la tâche de discrimination de visage tandis que les symboles carrés représentent les données de la tâche de discrimination de texture. Extrait de Gold et al. (1999)

Quatre nouveaux participants sont testés dans une tâche **d'identification du visage** en présence d'un bruit externe de faible magnitude ou de magnitude élevée, au sein de **deux passes** dans lesquelles les mêmes signaux et bruit sont présentés durant trois sessions de test. L'hypothèse est la suivante : si l'apprentissage réduit le bruit interne, le pourcentage d'accord

(« consistency ») devrait diminuer avec l'entraînement. Les résultats indiquent que quelle que soit la condition (bruit externe faible ou élevé), le bruit interne ne change pas au cours de l'apprentissage perceptif (voir figure 84). Ceci est compatible avec l'idée que dans ces tâches, **le bruit interne est probablement de nature additive et non multiplicative**.

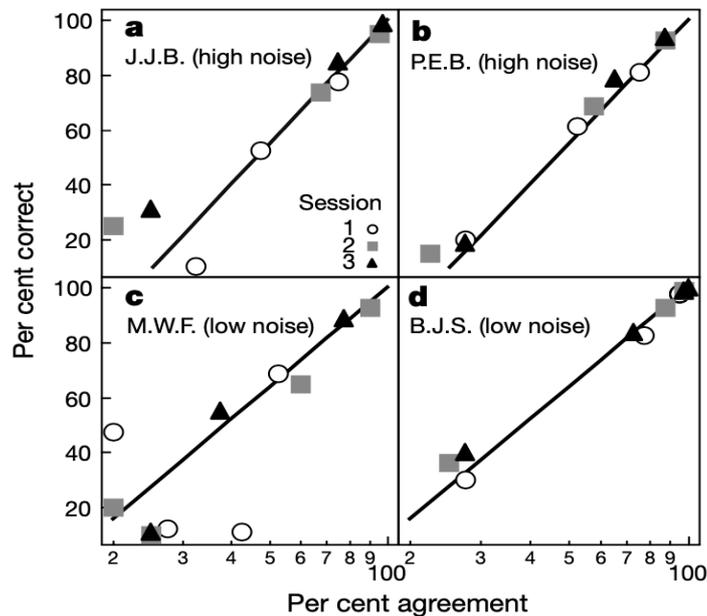


Figure 84. Données empiriques de double-passe (pourcentage d'accord vs. pourcentage correct) en présence d'un bruit externe faible et élevé dans une tâche d'identification de visage. Chaque point correspond à un observateur. Tous les observateurs présentent le même degré d'amélioration au fil des sessions de test que celui constaté lors de la première expérience. Les lignes pleines montrent les prédictions pour un observateur avec un rapport de bruit interne/externe de 1,0 (a, b) et 2,0 (c, d). Les symboles indiqués en légende indiquent la session d'entraînement. Extrait de Gold et al. (1999).

La littérature visuelle tend donc à conclure que les changements dans l'efficacité d'encodage sont généralement responsables de l'apprentissage perceptif (Gold et al., 1999 ; Gold et al., 2004 ; Lu & Doshier, 2009). L'apprentissage auditif se comporte-t-il comme l'apprentissage visuel ? C'est à cette question que tente de répondre l'étude menée par Jones et al. (2013).

Cette étude est remarquable par son exhaustivité, car elle évalue quatre types de facteurs associés à l'apprentissage perceptif dans des tâches de **discrimination fréquentielle**: le **bruit interne**, l'**efficacité d'encodage**, le **biais de réponse** et l'**inattention**, et ce à l'aide de quatre méthodes différentes : i) l'ajustement d'un modèle d'observateur dérivé de la TDS; ii) la méthode dite de « limite de classification » ; iii) la fonction psychométrique ; et iv) le paradigme de double-passe. Les tâches de discrimination fréquentielle sont menées sur 16 auditeurs normo-entendants jeunes (moyenne d'âge 25 ans).

Dans toutes les conditions, les stimuli sont des sinusoïdes de durée égale à 300 ms présentées à 70 dB SPL. Une procédure adaptative à deux alternatives en choix forcé (2I, 2AFC) est utilisée pour estimer les capacités de discrimination fréquentielle. Au sein de chaque essai, une paire de sons purs est présentée aléatoirement à partir de distributions Gaussiennes avec un écart-

type commun σ_{Hz} et des moyennes de $1000 \pm \sigma_{Hz}$ (σ_{Hz} est compris entre 0.5 et 5.5 Hz et varie avec un pas de 0.5 Hz). Cette gamme de valeurs est choisie à partir des données pilotes de telle sorte qu'elle couvre la gamme susceptible de valeurs du bruit interne. Les distributions (signal cible vs. signal de comparaison) sont séparées d'une distance égale à $\Delta Hz = 2\sigma_{Hz}$. Chaque procédure expérimentale est reproduite dans quatre sessions différentes.

Les résultats indiquent que **la sensibilité augmente en fonction des essais, démontrant un apprentissage perceptif pour la discrimination fréquentielle**. Quelle que soit la méthode d'analyse utilisée, **le bruit interne diminue significativement avec l'entraînement tandis que l'efficacité d'encodage ne change pas** (cf. figure 85). Une variation significative du biais de réponse avec les sessions est révélée grâce à la méthode de classification de limite mais pas à l'aide de la fonction psychométrique (le biais étant mesuré par l'erreur constante). Les données révèlent également que la méthode de la double-passe a tendance à estimer des valeurs de bruit interne plus grandes dans 88 % des cas.

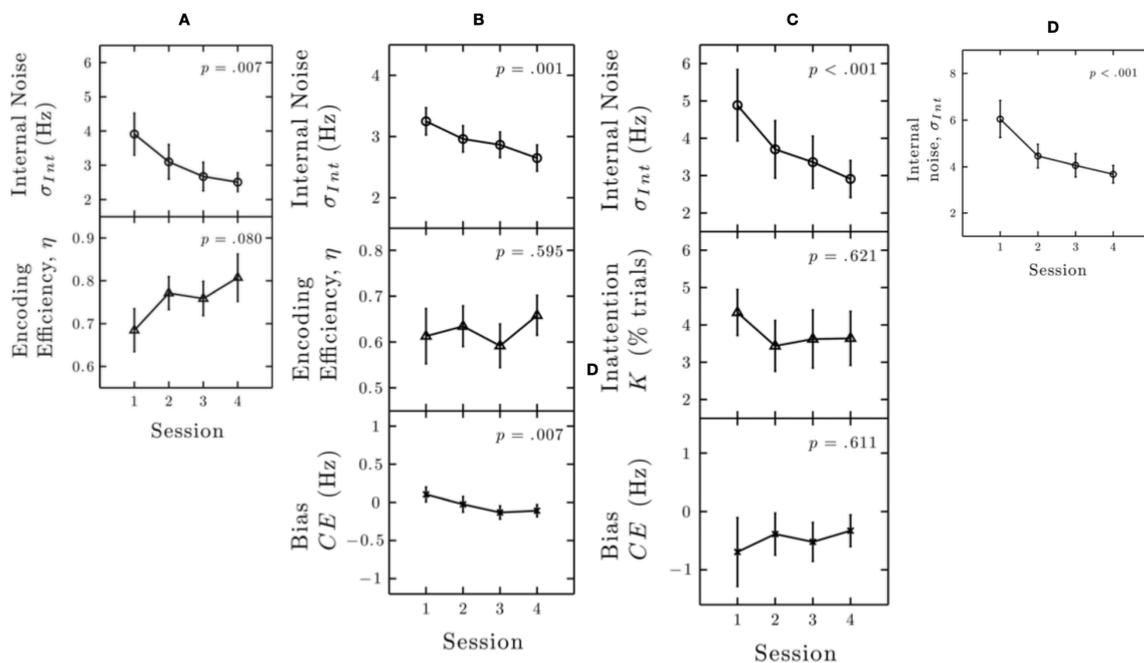


Figure 85. Estimations des différents facteurs (efficacité, bruit interne, biais, inattention) en fonction de la session pour chaque méthode d'analyse : (A) : Ajustement de modèle ; (B) : Classification de limites ; (C) : Fonction psychométrique; (D) : Double-passe. Extrait de Jones et al. (2013).

Ainsi cette étude démontre que l'amélioration de la discrimination fréquentielle suite à un apprentissage perceptif est accompagnée d'une **diminution significative du bruit interne**. Cette diminution du bruit interne est probablement liée à une repondération des premières informations sensorielles (présentées en session 1). Ces quatre mesures sont fortement corrélées ce qui renforce cette interprétation. Cependant, contrairement aux études visuelles, **peu de changements dans l'efficacité d'encodage, le biais de réponse et l'inattention sont mis en évidence avec l'apprentissage**.

I.4.d.iii. Vieillessement et troubles sensoriels liés à l'âge

Tout comme le système auditif, le système visuel se dégrade naturellement sous l'effet du vieillissement. Ceci provoque des modifications neuronales (Weale, 1975) et perceptives (Owsley et al., 1983 ; Hemenger, 1984 ; Wright & Drasdo, 1985 ; Sturr et al., 1988 ; Burton et al., 1993). Les modifications neuronales liées au vieillissement se produisent dans les voies neuronales entre la rétine et le cortex (Spear, 1993 ; Gao & Hollyfield, 1992). Elles peuvent être liées à une réduction de la densité de neurones se produisant essentiellement entre l'âge de 20 ans et 87 ans dans le cortex visuel (Devaney & Johnson, 1980) ou à une réduction qualitative du fonctionnement des neurones, susceptible d'affecter l'efficacité de traitement central (Walsh, 1976). Sur le plan perceptif, ces changements anatomiques et physiologiques affectent majoritairement l'accoutumance rétinienne (Owsley et al., 1983). Des recherches récentes démontrent qu'un système visuel d'une personne de 60 ans transmet approximativement 30% de lumière par comparaison avec le système visuel d'une personne de 20 ans. Ceci produit une diminution de sensibilité au contraste aux fréquences spatiales élevées (Kulikowski, 1971). Comme nous allons le voir ci-dessous, la mesure du bruit interne et de l'efficacité associées aux tâches de discrimination de contraste sur des participants âgés permet de séparer finement les facteurs perceptifs et neuronaux liés au vieillissement.

Pardhan (2004) exploite la **méthode du bruit équivalent** pour estimer la **sensibilité au contraste** en présence d'un bruit externe, ce afin de mesurer l'efficacité d'échantillonnage et les niveaux de bruit interne équivalents sous l'effet du vieillissement (Pardhan & Gilchirst, 1991). L'objectif de cette étude est de répliquer une étude précédente (Pardhan et al., 1996) afin d'estimer si ces facteurs dépendent de la fréquence spatiale. Dans cette étude, 8 participants jeunes (17-22 ans) et 8 participants plus âgés (60-72 ans) sont testés. La méthode à stimuli constants est combinée à une procédure en choix forcé et les participants sont tenus de détecter une image masquée par un bruit. Chaque session consiste en 50 blocs de stimulus, chaque bloc contenant 50 essais par fréquence spatiale et par niveau de bruit externe. Les résultats présentés en figure 86 indiquent : i) qu'en l'absence de bruit externe, la sensibilité au contraste est réduite dans le groupe de participants âgés par rapport au groupe de participants jeunes (en accord avec des études précédentes) ; ii) qu'en présence de bruit externe, **l'efficacité d'échantillonnage estimée est toujours plus grande chez les participants plus jeunes quelle que soit la fréquence spatiale** ; l'efficacité d'échantillonnage dépend de la fréquence spatiale : elle diminue lorsque la fréquence spatiale augmente. Le bruit équivalent est également dépendant de la fréquence spatiale : lorsque cette dernière est élevée (10 c/deg), une augmentation significative de bruit équivalent est observée chez les participants âgés (Pardhan, 2004).

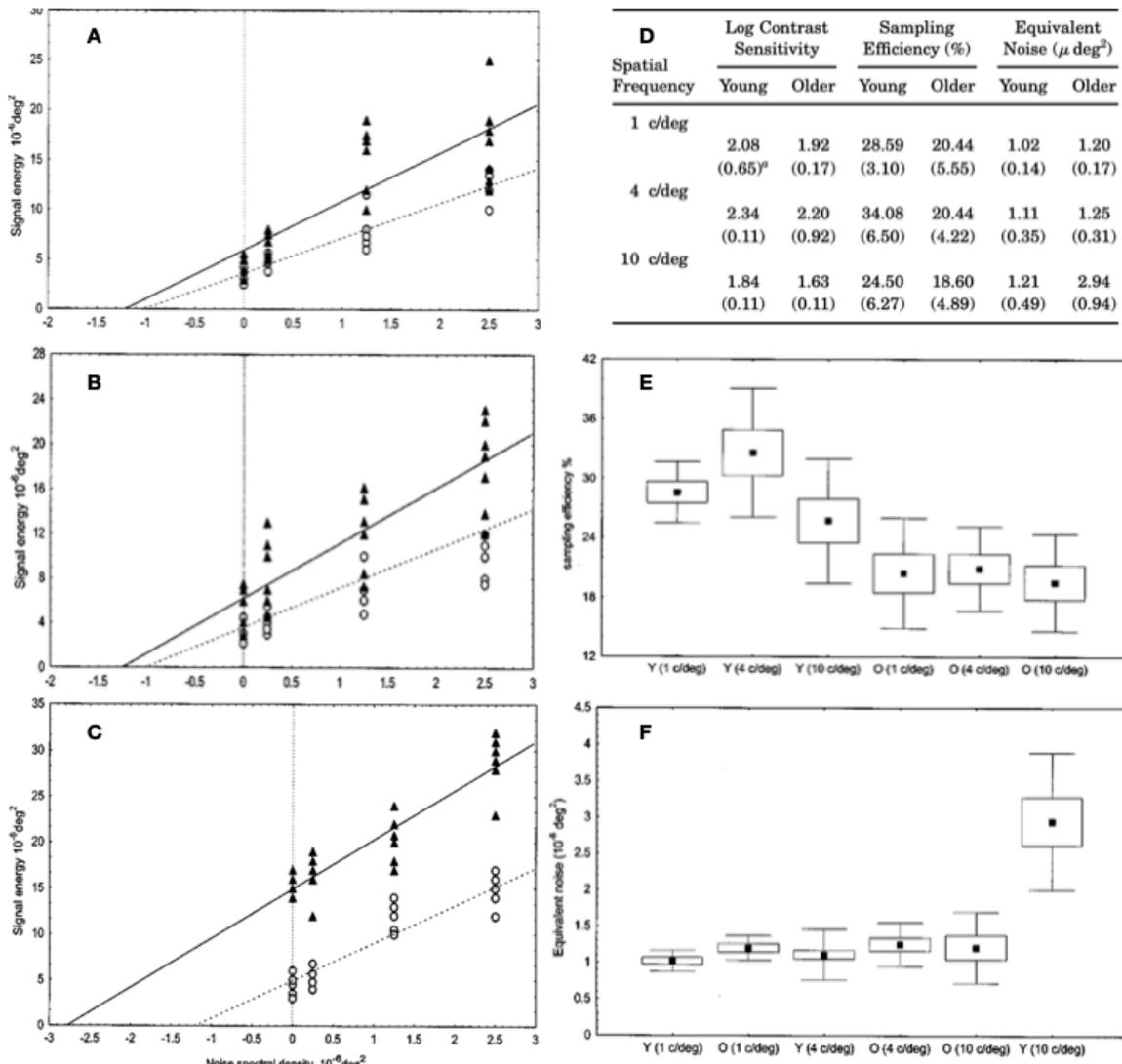


Figure 86. Résultats de la méthode du bruit équivalent sur deux groupes de participants (jeunes et âgés). Les figures (A), (B), (C) présentent les fonctions de détection de contraste dans le bruit pour chaque participant âgé (ligne pleine avec des triangles) et chaque participant jeune (ligne en pointillés avec des cercles ouverts). La fréquence spatiale est respectivement égale à 1c/deg ; 4c/deg et 10c/deg. Les figures (E) et (F) présentent les mesures d'efficacité d'échantillonnage et de bruit équivalent en fonction de la fréquence spatiale et ce pour chaque groupe (Y : jeunes ; O : âgés). La figure (D) présente les résultats des estimations d'efficacité d'échantillonnage et de bruit équivalent par groupe de participant et par fréquence spatiale. Extrait de Pardhan et al. (2004).

Dans l'ensemble, les résultats de cette étude démontrent que **des facteurs neuronaux et optiques sont responsables d'une diminution de la sensibilité au contraste sous l'effet du vieillissement** : aux fréquences spatiales basses et intermédiaires (entre 1 et 4 c/deg), la perte de sensibilité au contraste est principalement liée à une diminution de l'efficacité d'échantillonnage quel que soit l'âge des participants, tandis qu'aux fréquences spatiales élevées (10 d/deg), la perte de sensibilité au contraste est liée à une augmentation du bruit interne chez les participants âgés. Néanmoins, cette méthode ne permet pas de localiser ces facteurs d'inefficacité de traitement (Pardhan, 2004).

En modalité auditive et comme nous l'avons indiqué plus haut dans ce document, des études post-mortem sur des os temporaux humains démontrent que le vieillissement est associé à une perte de synapses entre les cellules ciliées internes et les fibres afférentes de type I, un phénomène dénommé « synaptopathie cochléaire » (Viana et al., 2015). Par conséquent, la perte de neurones véhiculant l'information du système auditif périphérique au système auditif central est susceptible d'altérer l'efficacité du codage neuronal des propriétés physiques des sons, comme par exemple l'information d'enveloppe temporelle ENV_p (Moore & Vinay, 2019). Néanmoins et comme indiqué dans un autre chapitre de ce document, les études récentes indiquent que les seuils de détection d'AM ne sont que très peu affectés par le vieillissement (Schoof & Rosen, 2014 ; Füllgrabe et al., 2015 ; Wallaert et al., 2016). Plusieurs raisons peuvent expliquer cela : i) il est possible que la synaptopathie cochléaire n'engendre pas une déafférentation suffisante pour créer des effets perceptifs détectables d'après des prédictions basées sur un modèle dérivé de la TDS (Oxenham, 2016) ; ii) il est possible que la synaptopathie cochléaire n'engendre pas distorsions temporelles suffisantes pour altérer la périodicité de l'enveloppe temporelle ; iii) le recrutement de sonie et la magnification perceptive d'enveloppe qui lui est associée pourraient compenser les effets délétères de la synaptopathie sur les seuils de détection d'AM (Ernst & Moore, 2012).

Sur la base de ces trois éléments, Moore et Vinay (2019) développent une **mesure de sensibilité à l'irrégularité de l'enveloppe temporelle du signal d'AM** pouvant fournir des preuves empiriques d'une augmentation de bruit interne dans le codage de l'enveloppe temporelle. Dans cette étude, la capacité à discriminer une AM régulière de cadence égale à 8 Hz d'une AM de même cadence mais irrégulière du point de vue de la profondeur et de la cadence est mesurée sur 10 sujets normo-entendants jeunes âgés entre 20 et 27 ans (âge moyen : 24 ans) et sur 10 sujets plus âgés (entre 64 et 80 ans ; âge moyen : 70 ans) et ce à deux niveaux d'intensité (20 et 80 dB SPL). Dans cette tâche de discrimination de l'irrégularité d'enveloppe dénommée « **ERD** », le degré d'irrégularité est varié afin de déterminer un seuil. Le raisonnement est le suivant : une AM régulière peut sembler irrégulière si le système auditif est affecté par une source de bruit interne en raison d'une synaptopathie cochléaire, ce qui se traduit par des seuils ERD plus élevés (moins bons). En d'autres termes, *en présence de synaptopathie cochléaire, le degré d'irrégularité dans le stimulus devrait être plus élevé pour surmonter les effets de l'irrégularité dans la représentation interne de l'enveloppe temporelle causée par un bruit interne accru*. Cette hypothèse, si validée, ouvrirait la voie à l'élaboration de tests cliniques permettant de diagnostiquer une perte auditive cachée.

Les résultats présentés en figure 87 indiquent que : i) Les seuils de discrimination d'ERD sont significativement meilleurs à 80 dB SPL qu'à 20 dB SPL (d'un facteur 2.8 dB) mais que cet effet du niveau est similaire quel que soit l'âge des participants ; ii) conformément à l'hypothèse selon laquelle le bruit interne augmente avec l'âge, les sujets âgés présentent des seuils ERD plus élevés que les sujets plus jeunes.

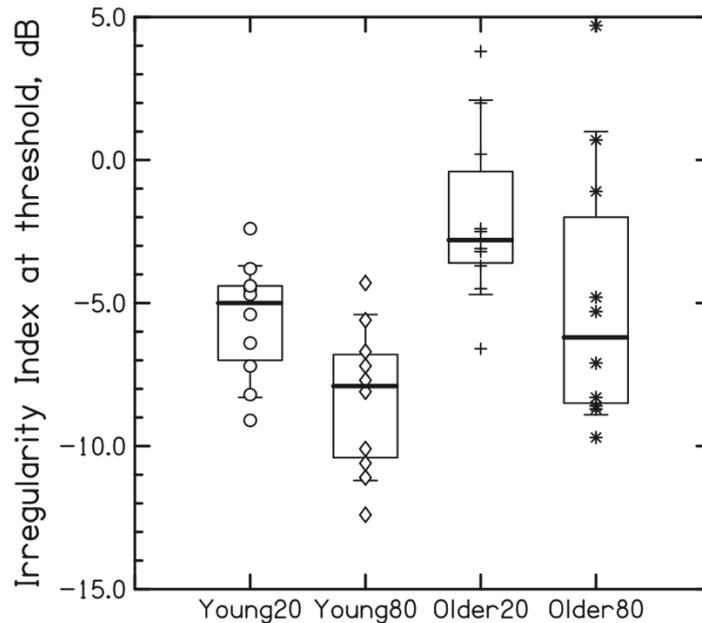


Figure 87. Boîtes à moustaches présentant les seuils d'irrégularité d'enveloppe (ERD) par groupe de participant (jeune ou âgé) et par niveau (20 ou 80 dB SPL). Les traits gras horizontaux indiquent les médianes, tandis que les bords inférieurs et supérieurs des boîtes indiquent les premiers et troisièmes quartiles et les moustaches indiquent les 10èmes et 90èmes percentiles. Les seuils de discrimination individuels sont représentés par des symboles. Extrait de Moore et Vinay (2019).

Bien que prometteuse, cette mesure de sensibilité à l'irrégularité de l'enveloppe temporelle ne permet pas de déterminer dans quelle mesure cette dégradation de performance est liée à des effets plus centraux liés à des aptitudes cognitives réduites (Wingfield, 1996 ; Füllgrabe et al., 2015) où à un accroissement du bruit interne dans le domaine de l'enveloppe temporelle.

Les travaux de modélisation visant à simuler les **capacités de détection d'AM et de FM** chez des sujets normo-entendants âgés et des sujets présentant une perte auditive neurosensorielle âgés suggèrent que le bruit interne peut augmenter sous l'effet du vieillissement (Paraouty et al., 2016 ; Wallaert et al., 2017 ; 2018 ; Ewert & Dau, 2000) et de dommages cochléaires (Derleth et al., 2001 ; Ives et al., 2014 ; Paraouty et al., 2016 ; Wallaert et al., 2017, 2018 ; Ewert et al., 2020). Parmi ces travaux, plusieurs études menées avec le même paradigme expérimental, les mêmes stimuli mais des modèles de structure différente indiquent que : **i) l'augmentation de la variance du bruit interne additif causée par le vieillissement en soi est relativement modeste et ii) celle causée par la presbycusie (la perte auditive neurosensorielle liée à l'âge) semble substantielle** (Paraouty et al., 2016 ; Wallaert et al., 2017, 2018 ; Ewert et al., 2020). Il est tout de même possible que les moins bonnes capacités de détection et de discrimination constatées chez les sujets âgés présentant une audition normale et les sujets presbycusiques résultent d'une altération des mécanismes (déterministes) de codage temporel (pour une revue, voir Gordon-Salant et al., 2010 ; Pichora-Fuller & Mac Donald, 2008) plutôt que d'une véritable augmentation de la variabilité neuronale dans les étapes périphériques et/ou centrales de traitement auditif.

D'autres investigations basées sur des paradigmes différents sont donc nécessaires afin de vérifier dans quelle mesure une augmentation de bruit interne est réellement responsable des modifications de la sensibilité à l'AM et à la FM observées chez les sujets presbycousiques.

I.4.d.iii. Pathologies cognitives

Une augmentation de la variabilité neuronale et comportementale peut parfois être à l'origine de pathologies affectant les traitements cognitifs (linguistiques et attentionnels) : Certains **enfants dyslexiques** présentent des troubles de traitement auditif (Hornickel & Kraus, 2013) altérant le développement normal de la parole et du langage (Bradlow et al., 2003 ; Ziegler et al., 2009 ; Bogliotti et al., 2008) ainsi que l'apprentissage de la lecture (Kraus et al., 1996 ; Wible et al., 2002 ; Abrams et al., 2006 ; Cao et al., 2006 ; Bonte et al., 2007 ; Gou et al., 2011). Une étude menée sur une large cohorte de 100 enfants par Hornickel et Kraus (2013) mesure la « cohérence neuronale » en réponse à des syllabes. Cette étude teste l'hypothèse selon laquelle ***une « instabilité neuronale » serait responsable de l'incapacité à faire le lien entre la phonologie et l'orthographe lors de la lecture chez l'enfant dyslexique.*** La cohérence des réponses neuronales est calculée séparément pour des portions reflétant la transition de formants (7-60 ms) et la voyelle à l'état stable (60-180 ms) en comparant les potentiels évoqués auditifs (PEAs) au début et à la fin de l'enregistrement. Les résultats démontrent que les PEAs des enfants présentant des troubles de la lecture sont plus variables que les enfants lisant correctement et que cette variabilité est présente d'essai en essai. Cette signature neurale typique de la dyslexie correspond à une ***augmentation de bruit interne***, donnant lieu à une forte instabilité neuronale des réponses aux sons (Hornickel & Kraus, 2013). Ces résultats confirment ceux d'études antérieures menées en modalité visuelle, suggérant que les enfants dyslexiques présentent un bruit interne plus élevé (Sperling et al., 2005) et entraînant des difficultés d'exclusion du bruit externe (Bradlow et al., 2003 ; Anderson et al., 2010). Les effets néfastes de ces facteurs s'accumuleraient depuis la naissance, expliquant pourquoi le traitement phonologique est moins efficace. Enfin, il est envisageable que l'instabilité de la représentation auditive élaborée au niveau du tronc cérébral soit responsable d'une inefficacité dans l'apprentissage des statistiques sonores chez les enfants dyslexiques, conduisant à des déficits dans le traitement neuronal combinés à des déficits de l'attention sélective. Cette étude permet donc de conclure que ***les enfants dyslexiques présentent une instabilité neuronale au sein du système auditif et que l'augmentation du bruit interne serait responsable de troubles auditifs et visuels associés au trouble d'acquisition du langage*** (Hornickel & Kraus, 2013).

Une variabilité intra-individuelle élevée des réponses comportementales pourrait représenter une caractéristique omniprésente et importante des ***troubles du déficit de l'attention avec ou sans hyper-activité (ADHD)***, (Castellanos & Tannok, 2002 ; Castellanos et al., 2005). Des travaux récents démontrent également une variabilité intra-individuelle significativement plus grande dans les temps de réponse d'un groupe de participants présentant des troubles

développementaux d'une autre nature : l'**autisme (ASD)**. Cette variabilité est observée dans une simple tâche à deux alternatives en choix forcé comparant un groupe de participants témoins (non pathologique) et un groupe de participants présentant un trouble du déficit de l'attention avec ou sans hyper-activité (Geurts et al., 2008) (voir courbes jaunes et rouges, figure 88 ci-dessous). La mesure de variabilité intra-individuelle représente par conséquent un indicateur important du traitement physiopathologique permettant d'expliquer certains comportements pour une catégorie de patients présentant un trouble spécifique (Milne, 2011). L'augmentation de la variabilité intra-individuelle est également observée dans les troubles développementaux tels que l'ADHD, l'ASD, ou encore le syndrome de Tourette (Geurts et al., 2004 ; Goldberg et al., 2005 ; Happé et al., 2006 ; Corbett & Constantine, 2006). Plusieurs études supposent que l'augmentation de la variabilité intra-individuelle résulte d'une réduction de myélinisation dans le cas de l'ADHD (Russel et al., 2006), de l'ASM (Cherkassky et al., 2006) et du syndrome de Tourette (Kates et al., 2002), affectant ainsi la connectivité fonctionnelle entre circuits cérébraux (Geurts et al., 2008). De nombreux auteurs suggèrent que l'augmentation du bruit interne neuronal révélée par une augmentation de la variabilité individuelle d'essai en essai dans des mesures d'EEG réalisées sur des participants autistes devrait contribuer à une dégradation de la performance dans des tâches cognitives (e.g., Rubenstein & Merzenich, 2003 ; Baron-Cohen & Belmonte, 2005 ; Dakin & Frith, 2005 ; Simmons et al., 2009). Par conséquent, **une augmentation de bruit interne pourrait être responsable des nombreuses conséquences perceptives et comportementales observées chez les enfants neuro-développementaux** (Milne, 2011 ; Geurts et al., 2008 ; Dinstein et al., 2012). Cette mesure de variabilité pourrait également permettre de dissocier ces troubles développementaux, comme l'indiquent les données de Geurts et al. (2008).

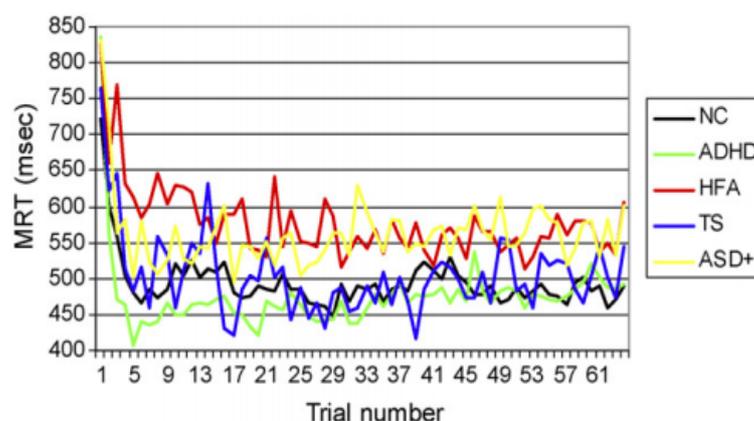


Figure 88. Temps de réaction (ms) d'essai en essai par type de trouble neuropsychologique (NC : trouble du comportement ; ADHD : déficit de l'attention avec ou sans hyperactivité ; HFA : Autisme à haut niveau de fonctionnement ; TS : Syndrome de Tourette ; ASD+ : Autistes). Extrait de Geurts et al. (2008).

En conclusion, ces études indiquent que :

Le concept de *bruit interne* désigne des processus neuronaux aléatoires intervenant à un niveau précoce et/ou tardif des traitements perceptifs, et à un niveau sensoriel et/ou cognitif. La variance de ce bruit interne peut être constante et indépendante du stimulus dans le cas d'un « bruit additif » ou proportionnelle à la magnitude du stimulus dans le cas d'un « bruit multiplicatif ». Ce bruit interne est susceptible de jouer un rôle à différentes échelles (neuronale, comportementale), et il peut avoir des conséquences néfastes (dégrader la performance de détection, discrimination ou reconnaissance) ou bénéfiques (*e.g.*, améliorer la performance par résonance stochastique) sur le traitement de l'information sensorielle. Le bruit interne peut être caractérisé par des mesures de variabilité neuronale ou comportementale.

De nombreux paradigmes psychophysiques ont été développés afin d'estimer la variance du bruit interne dans une tâche perceptive donnée. Cette estimation dépend toujours d'un modèle computationnel de traitement perceptif ou d'un modèle théorique dérivé de la théorie de la détection du signal. L'association de plusieurs paradigmes psychophysiques conduit à une mesure plus fiable du bruit interne permettant parfois de préciser sa nature (additive ou multiplicative). Par contre, il n'est pas possible d'identifier le locus du bruit interne sur la simple base de ces données comportementales.

En modalité visuelle, la technique du bruit équivalent est majoritairement utilisée tandis qu'en modalité auditive, l'estimation du bruit interne à partir de la pente de la fonction psychométrique est plus fréquente. La méthode de double-passe est plus rarement utilisée. Ces paradigmes psychophysiques peuvent être particulièrement utiles dans le domaine clinique. Ils permettent de compléter le diagnostic de certaines pathologies sensorielles ou cognitives, de dissocier certaines pathologies présentant des manifestations physiques semblables et enfin de suivre l'évolution d'une pathologie ou d'une capacité de traitement sensoriel ou cognitif au cours du temps.

Les études citées plus haut suggèrent que le traitement d'AM et de FM ferait appel à des mécanismes distincts associés à des sources de bruit interne indépendantes (Banai et al., 2011). Un modèle computationnel de traitement de ces informations de modulation dit « à double voie » implémente cette hypothèse (Ewert et al., 2020). Ces travaux indiquent également que l'altération des capacités de détection des informations d'AM et de FM associée au vieillissement et à des lésions cochléaires résulterait - pour partie - d'une augmentation du bruit interne dans le domaine temporel causée par une synaptopathie cochléaire (Wallaert et al., 2018 ; Moore & Vinay, 2019). Des travaux de modélisation concluent à la possibilité d'une augmentation *massive* du bruit interne dans le cas de la presbyacousie (*e.g.*, Wallaert et al., 2018). Ces résultats - potentiellement importants sur le plan théorique et clinique - demandent à être vérifiés à l'aide d'une méthode psychophysique évaluant directement le bruit interne pour la tâche en question.

I.5. Objectifs et plan de thèse

I.5.a. Problématique

A travers ces premières pages, nous avons présenté les principales connaissances dans le domaine de la perception auditive des modulations temporelles (AM et FM) véhiculées par les sons complexes tels que la parole, les sons musicaux ou encore les sons environnementaux. Nous avons ainsi caractérisé ces indices de modulation pour les trois catégories de sons cités ci-dessus, résumé nos connaissances quant aux capacités auditives humaines de détection, discrimination et reconnaissance de sons sur la base de ces indices, identifié les mécanismes sensoriels et cognitifs impliqués dans perception de ces indices d'AM et de FM, puis précisé les effets du vieillissement, du développement et d'une perte cochléaire sur la perception d'AM et de FM. Enfin, nous nous sommes concentrés sur le concept de « bruit interne » et son rôle dans les traitements sensoriels visuels et auditifs. Nous avons alors détaillé plusieurs méthodes psychophysiques d'estimation du bruit interne ainsi que les effets potentiels du vieillissement, du développement, de l'apprentissage et de pathologies auditives et visuelles sur le bruit interne en modalités auditive et visuelle. Ces travaux suggèrent que i) des sources de bruit interne distinctes contraignent le traitement et donc la sensibilité auditive à l'AM et à la FM (Ewert et al., 2020), et que ii) le vieillissement et la perte auditive neurosensorielle sont susceptibles d'augmenter - peut-être fortement - leur magnitude (*e.g.*, Wallaert et al., 2017 ; 2018 ; Moore & Vinay, 2019). Le premier constat a des conséquences théoriques importantes, car il demande de repenser l'architecture du traitement monaural des informations auditives de nature temporelle. Le deuxième constat a des fortes implications cliniques, car il engage à repenser les conséquences neurosensorielles du vieillissement et de lésions cochléaires en termes d'augmentation de variabilité neuronale et à développer de nouveaux tests visant des capacités auditives supraliminaires dépendant des manifestations spécifiques de ce bruit interne.

L'objectif général de la présente thèse consiste à *estimer - pour la première fois - le bruit interne dans des tâches auditives de détection d'AM et de FM chez des auditeurs humains à l'aide d'un paradigme psychophysique conçu pour mesurer directement les effets du bruit interne dans la dimension considérée*. Parmi les multiples méthodes d'estimation du bruit interne (*cf.* chapitre 4), nous avons donc choisi d'utiliser le **paradigme de double-passe** développé par Green (1964). Ce paradigme est original au sens où il permet de caractériser la perception d'AM et de FM à l'aide d'une nouvelle métrique : la **cohérence** des réponses (consistency). Le choix de ce paradigme est motivé par le fait qu'il estime le bruit interne de manière directe et intuitive. En effet, une augmentation du bruit interne est obligatoirement associée à une diminution de l'accord (la cohérence) dans les réponses d'un observateur à des signaux identiques. Par ailleurs, cette méthode s'avère être aussi efficace que les autres et elle présente l'avantage d'estimer un bruit interne global (additif et multiplicatif).

Ainsi, à partir de données empiriques collectées sur des sujets normo-entendants jeunes et âgés puis des sujets malentendants âgés, plusieurs **modèles** sont utilisés afin d'estimer la variance du bruit interne dans des tâches de détection d'AM et de FM. Plus précisément, les objectifs fixés visent à : i) déterminer dans quelle mesure la variance du bruit interne diffère entre des tâches de **détection d'AM et de FM lente et rapide** appliquées à une porteuse basse fréquence ; ii) mesurer les **effets du vieillissement** sur la variance de ce(s) bruit(s) interne(s) ; et enfin iii) évaluer les **effets d'une perte cochléaire** sur la variance de ce(s) bruit(s) interne(s).

Ce programme de recherche doctorale se développe autour de trois études distinctes :

La *première étude* évalue la cohérence des réponses et la performance sur un groupe de quinze sujets normo-entendants jeunes dans une tâche auditive de détection d'AM et de FM. Deux cadences de modulation (2 vs. 20 Hz) sont testées pour une fréquence porteuse de 500 Hz. Les signaux sont présentés dans un bruit externe à deux niveaux de variabilité distincts. Les données de double-passe sont simulées à l'aide du modèle dit du « banc de filtres de modulation » en ajustant un unique paramètre : la variance d'une source de bruit interne Gaussien et additif. L'hypothèse de travail principale est que les mécanismes spatiaux (tonotopiques) et temporels (verrouillage en phase sur la structure temporelle fine du signal) introduisent des sources de variabilité interne distinctes (Ewert et al., 2020), toutes deux étant susceptibles d'affecter différemment la cohérence des jugements des sujets lors de la détection d'AM et de FM. Cette hypothèse peut être raffinée suivant la cadence de modulation de telle sorte que : i) l'hypothèse d'un mécanisme sensoriel unique pour le traitement de l'AM et de la FM est corroborée si la variance du bruit interne ajustée aux données est identique aux cadences lente (2 Hz) et rapide (20 Hz) ; ii) l'hypothèse de deux mécanismes distincts pour le traitement de l'AM et de la FM est corroborée si la variance du bruit interne ajustée aux données est différente entre d'un côté la FM lente et de l'autre côté l'AM lente et rapide et la FM rapide. Par ailleurs, rien n'exclut que le traitement d'AM lente et rapide soit réalisé par des mécanismes distincts (Wright & Dai, 1998) compte tenu des différences de percepts évoqués par une AM lente (percept de fluctuations de sonie) et une AM rapide (percept de rugosité et de hauteur). Si tel est le cas, l'hypothèse de deux mécanismes distincts pour le traitement de l'AM lente vs. rapide est corroborée si la variance du bruit interne ajustée aux données est différente entre d'un côté l'AM lente et de l'autre côté l'AM rapide.

La *deuxième étude* étend la première étude à un deuxième groupe de quinze sujets normo-entendants âgés. Ici, l'estimation de la variance du bruit interne est réalisée à l'aide d'un modèle dérivé de la théorie de la détection du signal. L'hypothèse testée dans cette deuxième étude est que le vieillissement entraîne une augmentation de bruit interne en détection d'AM et de FM. Cette hypothèse repose sur la base de données récentes sur la synaptopathie cochléaire qui suggèrent que la perte massive de fibres auditives périphériques associée au vieillissement pourrait s'accompagner d'une augmentation de variabilité neuronale

(Goodman et al., 2018). Des études de modélisation récentes indiquent qu'une augmentation du bruit interne est nécessaire pour simuler correctement les données de détection d'AM et de FM de sujets âgés normo-entendants (Wallaert et al., 2017 ; 2018). Enfin, en considérant que la détection d'AM et de FM implique deux sources distinctes de bruit interne susceptibles d'affecter l'accord des réponses, il est probable que l'efficacité du mécanisme de verrouillage en phase sur la structure temporelle fine de la FM soit affectée sous l'effet du vieillissement.

La *troisième étude* compare les données de double-passe de deux groupes de sujets normo-entendants jeunes (n=19) et malentendants âgés (n=10). Les sujets sont testés au même niveau SPL pour une AM et une FM lente de cadence lente égale à 2 Hz et une porteuse sinusoïdale de 500 Hz. Comme pour l'étude précédente, l'estimation de la variance du bruit interne est réalisée à l'aide d'un modèle dérivé de la théorie de la détection du signal. L'objectif de cette troisième étude est de tester l'hypothèse selon laquelle la perte auditive neurosensorielle provoque une augmentation du bruit interne en détection d'AM et de FM. Cette hypothèse est basée sur plusieurs travaux de modélisation simulant la détection et la discrimination d'AM et de FM et suggérant que la variance du bruit interne augmente - peut-être substantiellement - à la suite de lésions cochléaires (Derleth et al., 2001 ; Ives et al., 2014 ; Parouty et al., 2016 ; Wallaert et al., 2017 ; 2018 ; Ewert et al., 2020).

I.5.b. Plan de thèse

Les trois études de ce programme de recherche sont présentées dans les chapitres II à IV. Les résultats révèlent que :

- i) Quelle que soit la cadence de modulation, les indices d'AM et de FM sont détectés par des mécanismes sensoriels distincts possédant leur propre source de bruit interne additif. Les simulations d'un modèle à banc de filtres de modulation reproduisant à la fois les données empiriques de cohérence et de performance pour un grand nombre de conditions expérimentales confirment sa capacité à offrir une approche unifiée de la perception d'AM lente et rapide. Toutefois, les résultats suggèrent qu'un bruit multiplicatif affecte également le traitement d'AM.
- ii) Avant l'âge de soixante ans, le bruit interne (estimé cette fois-ci à l'aide d'un algorithme basé sur la théorie de la détection du signal) n'augmente pas quelle que soit le type de modulation et la cadence de modulation.
- iii) Une perte auditive neurosensorielle associée au vieillissement n'altère pas le bruit interne dans la détection d'AM et de FM lente.

Le dernier chapitre de thèse évalue ces résultats à la lumière des études antérieures. Il discute les limites et voies d'amélioration de ce programme de recherche, ainsi que l'intérêt clinique de l'estimation du bruit interne pour le diagnostic de certaines pathologies auditives.

Chapitre II :

Application du paradigme de double-passe à la détection d'AM et de FM chez des sujets normo-entendants jeunes

II.1. Double-pass consistency for amplitude- and frequency-modulation detection in normal-hearing listeners

II.1.a. Résumé

Amplitude and frequency modulation (AM, FM) provide crucial auditory information. If FM is encoded as AM, it should be possible to give a unified account of AM and FM perception both in terms of response consistency and performance. These two aspects of behavior were estimated for normal-hearing participants using a constant-stimuli, forced-choice detection task repeated twice with the same stimuli (double pass). Sinusoidal AM or FM with rates of 2 or 20 Hz were applied to a 500-Hz pure-tone carrier and presented at detection threshold. All stimuli were masked by a modulation noise. Percent agreement of responses across passes (PA) and percent-correct detection (PC) for the two passes were used to estimate consistency and performance, respectively. These data were simulated using a model implementing peripheral processes, a central modulation filterbank, an additive internal noise and a template-matching device. Different levels of internal noise were required to reproduce AM and FM data but a single level could account for the 2- and 20-Hz AM data. The level of internal noise yielding best predictions increased with the level of the modulation-noise masker. These results suggest that distinct sources of internal variability are involved for AM and FM detection at low audio frequencies.

II.1.b. Article

L'article suivant a été accepté par *The Journal of the Acoustical Society of America* en Août 2021 puis publié le 15 Novembre 2021.

I. INTRODUCTION

The auditory perception of amplitude and frequency modulation (AM and FM) has received substantial interest over the last decades because of the repeated demonstration of the crucial role played by AM and FM cues in robust speech perception (*e.g.*, Shannon *et al.*, 1995; Zeng *et al.*, 2005) and in environmental sound perception (Singh and Theunissen, 2003; Thoret *et al.*, 2020). It is generally assumed that auditory processing of AM and FM overlaps because of “FM-to-AM conversion” in the cochlea (the frequency-dependent attenuation of the FM caused by the tuned cochlear filters resulting in AM cues; Zwicker, 1952; Saberi and Hafter, 1995). Although questions remain as to how slow FM is detected by humans at low audio frequencies (Whiteford and Oxenham, 2015; Paraouty *et al.*, 2016, 2017; Whiteford *et al.*, 2017; Ewert *et al.*, 2018; Wallaert *et al.*, 2018; King *et al.*, 2019; Moore *et al.*, 2019; Parthasarathy *et al.*, 2020; Whiteford *et al.*, 2020), the consensus that has emerged posits that the perception of AM and fast FM is mediated by AM processing through a cascade of processing stages including: bandpass (cochlear) filtering, fast-acting amplitude compression, half-wave rectification, short-term adaptation with fast time constants, demodulation achieved by lowpass filtering followed by a modulation filterbank, and a decision stage using either a template-matching strategy or time-averaged statistics such as the envelope power (Viemeister, 1979; Strickland and Viemeister, 1996; Dau *et al.*, 1997a,b; Ewert and Dau, 2000; McDermott and Simoncelli, 2011).

This general model (hereafter, referred to as the “modulation-filterbank model”; for recent implementations, see Jepsen *et al.*, 2008; Jørgensen *et al.*, 2013; Biberger and Ewert, 2016, 2017; Wallaert *et al.*, 2017, 2018; King *et al.*, 2019; Cabrera *et al.*, 2019) postulates that temporal-modulation cues in sounds are transformed into so-called “neural temporal-envelope” cues (*i.e.*, fluctuations in mean firing rate in auditory neurons) and that fine-timing “temporal fine structure” (TFS) cues (*i.e.*, carrier information) are discarded after demodulation achieved by central (post-cochlear) processes. This type of model incorporates an important source of “inefficiency” in temporal-modulation processing: a Gaussian *internal noise* added to the representation of temporal-envelope cues at the output of modulation filters. Internal noise (Green and Swets, 1966) refers to the accumulation of several sources of variability such as the stochastic nature of neuronal firing, the internal state of the observer’ organism, or fluctuations in attention (Javel and Viemeister, 2000; Faisal *et al.*, 2008; Amitay

et al., 2013). Internal noise is initially included in this model of modulation perception to limit the resolution of the observer and thus auditory sensitivity to temporal modulation, that otherwise would be perfect. This implies that the decision processes postulated for temporal-modulation detection use a degraded representation of temporal-envelope cues extracted by modulation filters (*i.e.* some envelope information is lost along the auditory pathway before final decision making). Surprisingly, little effort has been made to characterize internal noise in temporal-modulation detection, with the exception of the work by Ewert and Dau (2004). This study assessed the contribution of internal noise to AM detection by comparing AM detection thresholds for random- versus frozen-noise carriers. Threshold patterns obtained for the frozen-noise carriers were found to be quite similar to those obtained for the random-noise carriers when broadband noise was used as a carrier. The absence of any variability in the AM stimulus in the case of the frozen-noise carrier had little impact if any on the listener's decisions. This led the authors to conclude that AM detection performance is affected by a *large amount of internal noise* (relative to external variability) in the temporal-envelope processing pathway. This outcome is consistent with previous work conducted for other perceptual dimensions in the auditory and visual modalities, concluding that the "specific choice generated by an observer on a given trial depends, to a large extent, on a loud source of variability that is not under direct experimental control" (*e.g.*, Barlow, 1956; Green, 1964; Burgess and Colborne, 1988; Pelli, 1990; Neri, 2010).

However, it is important to note that internal noise not only limits observer sensitivity: it is also a source of neural variability that determines the *consistency* of the auditory judgments made by the observer in a given perceptual task. In other words, for a given stimulus, a high level of internal noise is expected to yield more variable responses from trial to trial (Green, 1964; Lu & Doshier, 2008). To quote Green (1964) on this issue, "internal noise is the limiting factor in a trial-by-trial prediction of the subject's responses". This operational definition of internal noise led to the "double-pass" method (Green, 1964), a psychophysical paradigm aiming at inferring the variance of internal noise for a given task from the agreement of individual responses between two successive trials using identical stimuli. The double-pass technique has been used extensively to assess internal noise for a wide range of perceptual tasks and different sensory modalities (*e.g.*, Green, 1964; Spiegel and Green, 1981; Burgess and Colborne, 1988; Lu and Doshier, 2008; Neri, 2010; Hasan *et al.*, 2012; Vilidaite and Baker, 2016). It also proved useful to characterize perceptual learning in the auditory (Jones *et al.*, 2013) and visual modalities (Lu and Doshier, 2004, 2006, 2008).

Here, the double-pass method was used to assess and compare behavioral consistency - and thus the influence of internal noise - for slow and fast AM and FM detection. According to the modulation-filterbank model described above, AM detection should be mediated by (neural) temporal-envelope cues *irrespective of modulation rate*. This implies that both slow (<10 Hz) and fast AM detection should be limited by the same sources of inefficiency. Consistent with this assumption, the modulation-filterbank model uses a single source (and

variance) of internal noise for temporal-envelope processing irrespective of temporal-envelope rate (Dau *et al.*, 1997a, b). Still, some doubts remain as to whether slow and fast AM are processed by the same mechanisms. For instance, Wright and Dai (1998) showed that the detectability of sinusoidal AM at unexpected rates differs for slow and fast modulation rates. Therefore, it is still unknown whether the same source of internal noise constrains AM processing at slow and fast rates.

For FM, it is generally assumed that detection is also mediated by temporal envelope (rate/place) information at fast modulation rates (higher than 5-10 Hz) or at high carrier frequencies (higher than 4-10 kHz) (Moore and Sek, 1994; Paraouty *et al.*, 2018; Whiteford *et al.*, 2020). As indicated above, this overlap between AM and FM processing is a consequence of “FM-to-AM conversion” in the cochlea: frequency-dependent attenuation of time-varying instantaneous frequency in the FM stimulus caused by the tuned cochlear filters results in AM (*i.e.*, temporal envelope cues) that is encoded *via* fluctuations in the mean discharge rate of auditory-nerve fibers. Still, doubts remain as to how slow (<10 Hz) FM is detected at low carrier frequencies (< 2-4 kHz). In that case, FM information is potentially conveyed by changes over time in the pattern of neural phase locking of auditory-nerve fibers or neurons in the low brainstem (Moore and Sek, 1996; Paraouty *et al.*, 2018; Wallaert *et al.*, 2018; Moore and Sek, 2019; Moore *et al.*, 2019; Parthasarathy *et al.*, 2020) because the precision of neural phase locking is constant for frequencies up to about 0.6-2 kHz for most mammals (Palmer and Russell, 1986). Some psychophysical studies also pointed out that TFS cues are not used to detect FM with modulation rates above about 10 Hz, arguing that the central mechanism using neural phase-locking information is sluggish (*e.g.*, Moore and Sek, 1994). This implies that at low carrier frequencies, slow FM detection should be limited by sources of inefficiency distinct from those limiting fast FM and slow and fast AM detection. Ewert *et al.* (2018) implemented this assumption in a “two-path” model of AM and FM detection using different sources (and variances) of internal noise for temporal-envelope (rate/place cues) and TFS (phase-locking cues) processing.

The general goal of the present study was to test whether the modulation-filterbank model described above (*i.e.*, a model of modulation processing using a single source of additive internal noise) was able to account for double-pass consistency data in an AM and FM detection task. More specifically, this study aimed to: i) assess double-pass consistency of auditory judgments in sinusoidal AM and FM detection using a slow (2 Hz) and a fast (20 Hz) modulation rate, a low (500 Hz) carrier frequency and a bandpass (AM or FM) noise masker (weak or strong in magnitude), and ii) compare the results with the predictions of the modulation-filterbank model using the same tasks and stimuli.

The double-pass method was used to estimate the consistency of auditory judgments in each experimental condition. This method (Green, 1964; Burgess and Colborne, 1988; Lu and Doshier, 2008) is typically performed using two passes. Stimuli presented in each pass are masked by an external source of noise, as double-pass consistency is determined by the ratio

of internal to external noise (Green, 1964): if external variability outweighs internal variability, then observer's consistency should be high. Conversely, if internal variability outweighs external variability, then observer's consistency should be low. To estimate double-pass consistency for modulation detection, a 2-interval, 2-alternative forced-choice (2I, 2AFC) detection-in-noise task was run twice with identical sequences of external noise in the two successive passes. External noises used for AM and FM detection were generated in the AM and FM domains, respectively (external variability was therefore introduced in the dimension - AM or FM - relevant for the task). Two levels of the variance of external (modulation) noise were used for each type of noise. The modulation-filterbank model simulated each experimental condition. The bandwidth of the simulated cochlear filters was set to 1 equivalent rectangular bandwidth (ERB_N); envelope phase was preserved at the output of modulation filters tuned below 10 Hz, and for each modulation filter centered at and above 10 Hz, only the Hilbert envelope of the output was passed on. For each experimental condition, the model was run for target modulation strengths set to the mean detection threshold across real participants, and the magnitude of the additive internal noise was varied systematically in order to find the value that minimized the model's prediction errors estimated in terms of performance and consistency. A unified account of AM and FM perception would be achieved if a single level of the variance of internal noise could minimize the model's prediction errors estimated in terms of performance *and* consistency in all experimental conditions.

II. METHODS

A. Listeners

Fifteen young normal-hearing listeners aged between 18 and 30 years (mean = 23 years; standard deviation, SD = 3 years) participated in the experiments. Only two of them had previous experience in psychoacoustic measurements. They were recruited through the logistic platform RISC ("*Relais d'Information sur les Sciences de la Cognition, UMS CNRS 332*") at Ecole normale supérieure (Paris). Each listener had pure-tone audiometric thresholds ≤ 20 dB HL between 0.25 and 4 kHz (ANSI, 1996) in both ears and the audiometric thresholds at the tested frequency (0.5 kHz) ranged between -5 and 20 dB HL (mean = 6.2 dB HL; SD = 5.3 dB HL).

All listeners were fully informed about the goal of the study and provided written consent before their participation. This study was approved by the local ethical committee of University Paris Descartes (CERES, N° IRB: 20143200001072).

B. Stimuli

All stimuli were generated digitally at a sampling rate of 48 kHz using the MATLAB environment (release 2013b). They were presented to the right ear of each participant at a sensation level (SL) of 40 dB using an external soundcard (DAC audio USB Audioengine D3, 24-bit resolution) and Sennheiser HD 280 pro headphones 64 ohms (Old Lyme, CT).

The whole series of tests was conducted in an IAC double-wall soundproof booth and all levels were calibrated using a Bruel and Kjaer 2250 sound-level meter. The latter was also used to calibrate the headphone with an artificial ear (4153 6 cc coupler, IEC standard-60318-1). Calibration was performed between 0.125 and 4 kHz.

All experiments were based on a 2I, 2AFC paradigm. Each trial was divided into two successive observation intervals containing a 500-Hz pure tone modulated (in amplitude or in frequency) by a modulation-noise masker. Onset and offset of the modulation-noise masker were simultaneous with each 500-Hz tone. In the target interval, a sinusoidal signal at a fixed rate (2 or 20 Hz) was added to the modulation-noise masker before being applied to the 500-Hz pure-tone carrier. The sinusoidal signal had the same duration as the pure-tone carrier. The target and comparison intervals were presented in random order. The two observation intervals were separated by a 400-ms silent interval. The duration of each stimulus was set to 1 sec. A 100-ms raised-cosine ramp was applied to the beginning and the end of each stimulus. All stimuli were normalized in terms of their root-mean-square (RMS) power.

1. AM detection

A modulation noise was generated by filtering a 1-sec Gaussian white noise $n(t)$ with zero mean below 80 Hz using a 4th-order lowpass Butterworth filter. The standard deviation of the modulation noise, $\sigma_{ext, AM}$, was set to two values: 0.07 or 0.14 modulation-depth units (m.d.u.). The comparison stimuli were sinusoidal carriers (with a frequency $f_c = 500$ Hz) with instantaneous amplitude modulated using the bandpass noise masker, $n(t)$. The target stimuli were sinusoidal carriers (with a frequency $f_c = 500$ Hz) with instantaneous amplitude modulated by the sum of a sinusoidal signal (of rate f_m and magnitude m) and the bandpass noise masker, $n(t)$.

Equation [1] describes the target stimulus $T_{AM}(t)$:

$$T_{AM}(t) = [1 + m \sin(2\pi f_m t + \varphi_{AM}) + \sigma_{ext, AM} n(t)] \sin(2\pi f_c t + \varphi_c) \quad [1]$$

Equation [2] describes the comparison stimulus $C_{AM}(t)$:

$$C_{AM}(t) = [1 + \sigma_{ext, AM} n(t)] \sin(2\pi f_c t + \varphi_c) \quad [2]$$

where t is the time expressed in seconds (s), m is the AM depth (ranging between 0 and 100 %) of the target signal, f_m is the modulation rate of the target signal expressed in Hertz (2 or 20 Hz), φ_{AM} and φ_c are the starting phases of the AM target signal and the carrier, respectively. φ_{AM} and φ_c were randomly and independently chosen (between 0 and 2π radians) for each stimulus in the setup experiment (measure of AM detection thresholds) and in the first pass of the main experiment (measure of double-pass consistency for AM detection).

2. FM detection

As for the AM task, a modulation noise was generated by filtering a 1-sec Gaussian white noise $n(t)$ with zero mean below 80 Hz using a 4th-order lowpass Butterworth filter. The standard deviation of the modulation noise, $\sigma_{ext, FM}$, was set to two values: 0.07 or 0.14 modulation-index units (m.i.u.). The comparison stimuli were sinusoidal carriers (with a frequency $f_c= 500$ Hz) with instantaneous frequency modulated using the bandpass noise masker, $n(t)$. The target stimuli were sinusoidal carriers (with a frequency $f_c= 500$ Hz) with instantaneous frequency modulated by the sum of a sinusoidal signal (of rate f_m and magnitude $\beta = \Delta f/f_m$) and the bandpass noise masker, $n(t)$.

Equation [3] describes the target stimulus $T_{FM}(t)$:

$$T_{FM}(t) = \sin [2\pi f_c t + \varphi_c + \beta \sin (2\pi f_m t + \varphi_{FM}) + \sigma_{ext, FM} n(t)] \quad [3]$$

Equation [4] describes the comparison stimulus $C_{FM}(t)$:

$$C_{FM}(t) = \sin[2\pi f_c t + \varphi_c + \sigma_{ext, FM} n(t)] \quad [4]$$

where t is the time expressed in seconds (s), $\beta = \Delta f/f_m$ where Δf is the frequency excursion of the target signal expressed in Hertz, f_m is the modulation rate of the target signal expressed in Hertz (2 or 20 Hz), φ_{FM} and φ_c are the starting phases of the FM target signal and the carrier, respectively. φ_{FM} and φ_c were randomly and independently chosen (between 0 and 2π radians) for each stimulus in the septup experiment (measure of FM detection thresholds) and in the first pass of the main experiment (measure of double-pass consistency for FM detection).

3. Modulation-noise maskers

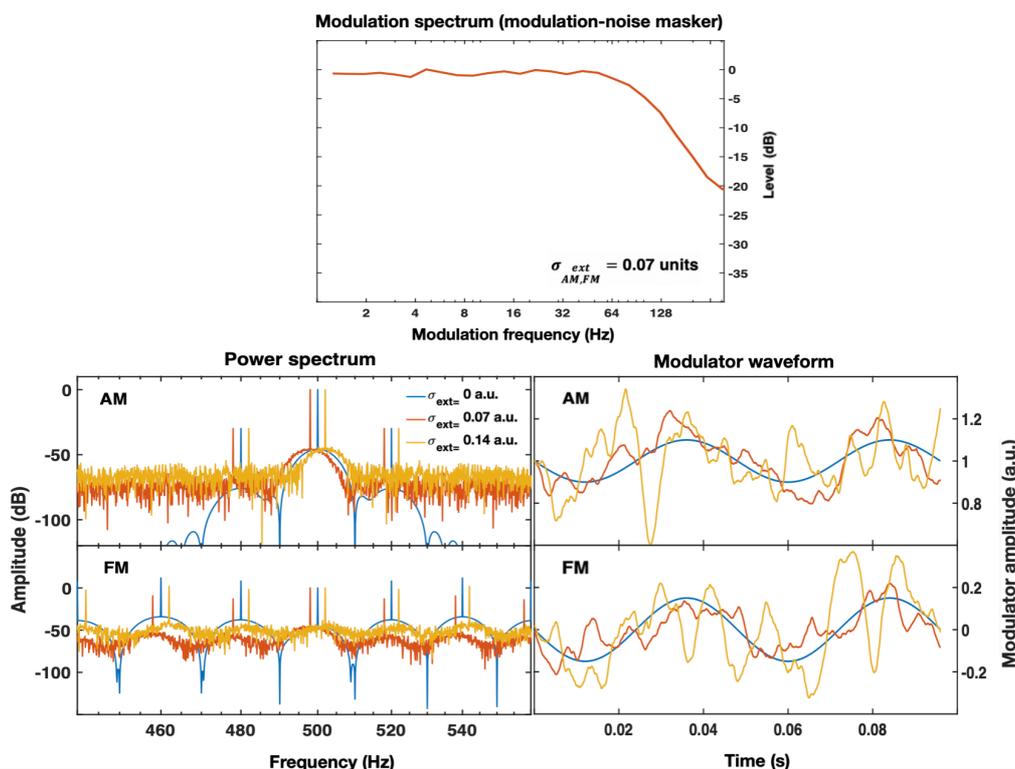


Figure 1. Stimuli: Modulation-noise maskers and AM/FM targets. *Top panel: Long-term modulation spectrum of AM and FM modulation-noise maskers. The standard deviation of the masker, $\sigma_{AM,FM}^{ext}$, was set to 0.07 units. Modulation spectra were averaged over 100 stimuli. Bottom panels: a) left panels: long-term power spectra of a 20-Hz sinusoidal AM (top panel) and FM (bottom panel) target with (orange and yellow lines) and without (blue lines) AM and FM modulation-noise maskers. Power spectra were offset (+/- 2 Hz) for clarity. Blue lines (no masker) are shown here to illustrate the detrimental effects caused by modulation-noise maskers; AM and FM stimuli were always presented against a modulation-noise maskers in the setup and main experiments. Here, $\sigma_{AM,FM}^{ext}$ was set to three values: 0 (no masker), 0.07 or 0.14 units (m.d.u. or m.i.u.). For AM target stimuli, modulation depth was set to 10 %. For FM target stimuli, frequency excursion was set to 3 Hz. Power spectra were averaged over 10 stimuli. b) right panels: modulator waveforms of the 20-Hz sinusoidal AM (top panel) and FM (bottom panel) target. Modulator waveforms were obtained for independent samples of modulation noise.*

For both AM- and FM-detection tasks, the modulation-noise masker, $n(t)$, was refreshed between intervals and trials in the setup experiment (measure of modulation-detection thresholds), and in the first pass of the main experiment (measure of double-pass consistency). The modulation-noise maskers used in the second pass of the main experiment were exactly the same as those used in the first pass of the main experiment (in corresponding trials). Modulation-noise maskers were different across participants (each participant was assigned a unique set of randomly selected modulation-noise maskers, but the statistics of modulation noises remained globally the same for all participants). Fig. 1 (top panel) shows that the long-term modulation spectrum of the modulation-noise maskers is lowpass in shape,

cutting around 80 Hz. The effects of the modulation noise on the AM and FM targets are illustrated in the bottommost panels of Fig. 1, that show the long-term power spectra (left panels) and modulator waveforms (right panels) of the AM and FM target stimuli used in the experiments. Power spectra and modulator waveforms are shown for a 20-Hz sinusoidal AM (bottommost, top panels) and FM (bottommost, bottom panels) signal. The bottommost left panels show that the AM and FM noise maskers introduce audio-frequency components on either side of the sidebands of the target signal, and that the level of these masking audio-frequency components increases when the standard deviation of the modulation noise increases from 0.07 to 0.14 units (m.d.u. or m.i.u.). The bottommost right panels illustrate the corruption caused by the modulation noise on the target sinusoidal modulators in the time domain.

C. Procedure

Two experiments were completed in the same order for each participant: in the first (so-called “setup”) experiment, an adaptive procedure was used to measure individual AM and FM detection thresholds corresponding to three distinct levels of performance. In the second (so-called “main”) experiment, a double-pass consistency task was used to measure the percentage of agreement in the same detection tasks at a fixed level of performance based on individual detection thresholds measured in the first experiment. The whole series of experiments were based on one or two hours per session with several sessions spread out over a 1-4 week period for each participant. Altogether, the experiments lasted on average 14 hours per participant. For each task and for each trial, the participant was instructed to choose the interval containing the target sinusoidal modulation.

Each participant was tested for the following 24 experimental conditions: two types of modulation [AM vs. FM]; two modulation rates [2 vs. 20 Hz]; two levels of the standard deviation of the (external) modulation-noise masker [0.07 vs. 0.14 units (m.d.u. or m.i.u.)]; three levels of (targeted) performance [64 % ($d'=0.5$), 76 % ($d'=1$), and 84 % ($d'=1.5$) correct detection]. The two levels of the modulation-noise masker standard deviation were determined in a pilot experiment to yield distinct levels of detectability of modulation targets. These values were also selected to prevent over-modulation. The individual AM and FM detection thresholds were then used to set the modulation strength (*i.e.*, the modulation depth or frequency excursion, respectively) of the target signals in the subsequent double-pass experiments. Hence, the AM and FM detection tasks in the double-pass consistency paradigm were performed at same level of difficulty.

The double-pass consistency task was conducted with the purpose of estimating PA in the participant responses for corresponding trials in two successive passes. PC was also calculated across the two successive passes for each participant.

1. Setup experiment: Detection thresholds

AM detection thresholds (modulation depth, m) and FM detection thresholds (frequency excursion, Δf) were first measured for each participant with adaptive staircases that attempted to converge on 64% ($d'=0.5$), 76% ($d'=1$) and 84% ($d'=1.5$) correct detection using a 2I, 2AFC paradigm and a weighted 1-up, 1-down method (Garcia-Pérez, 1998).

Each block consisted of a variable number of trials. Each trial was divided into two successive observation intervals containing the target and comparison stimuli, presented in random order. The bandpass noise masker was refreshed for each interval and trial. Each observation interval was marked by a red light on the computer screen that was synchronized with the stimulus' presentation. Participants were instructed to report the observation interval containing the target stimulus. Visual feedback was given to the participant after each trial. The dependent variables for these measures were the modulation depth m for the AM detection task and the frequency excursion Δf for the FM detection task. At the beginning of each staircase, the modulation strength of AM and FM targets was set to a supra-threshold level: The AM target was presented at a modulation depth of 80 % (-2 dB when expressed as $20\log m$) and the FM target was presented at a frequency excursion of 8 Hz. At the beginning of each staircase, the tracking variable (the modulation strength of AM and FM targets) changed by a step (a factor) up of 4 after a wrong response and by a step (a factor) down of 2 after a correct response. After the first 4 reversals, the 1-down 1-up rule was used up to sixteen reversals, using different steps to reach 64 %, 76% and 86 % of correct responses (Garcia-Pérez, 1998). These twenty reversals were performed for each series of measures and the mean threshold was taken as the geometric mean of the tracking variable during the last sixteen reversals. The order of the conditions was randomized across runs.

AM and FM detection thresholds at 2 and 20 Hz were first measured once in the absence of any modulation noise masker in order to familiarize participants with the detection tasks. AM and FM detection thresholds were then measured twice in each condition (*i.e.*, forty-eight staircases). Two measures of detection threshold were thus carried out in each experimental condition and the final detection threshold for each condition was taken as the mean of these two estimates.

2. Main experiment: Double-pass consistency

The consistency of auditory judgments in AM and FM detection was then measured in each experimental condition and for each participant using the double-pass paradigm developed by Green (1964). Here, AM and FM detection were measured using a 2I, 2AFC paradigm with constant stimuli. In total, each participant was tested over $2 \times 100 = 200$ trials in each of the 24 experimental conditions. In other words, each participant was run twice with identical stimuli, making two successive “passes” of 100 trials in each experimental condition. However, for each trial, the presentation order of the target and comparison stimuli was randomized (and was thus different across the two passes) to avoid contextual effects. The 24 experimental conditions were tested in random order. Thus, 4800 trials were run per participant. These 4800 trials were split into 10 (shorter) series of double-pass experiments in which all (24) conditions were completed. Each series was therefore composed of 480 trials: it consisted in a first block of 240 trials in which all (24) conditions were completed 10 times in random order, repeated twice with exactly the same stimuli in a different trial order in a second block of 240 trials (pass 2) after a 5 min break. Short breaks were given to participants between each series of double-pass experiments.

Each block therefore consisted of 240 trials. Five training trials (including random conditions) where modulation targets were presented at a modulation strength yielding a d' of 3 were inserted at the beginning of each block to familiarize the participant with the task and stimuli. For each block, each trial was divided into two successive observation intervals containing the target and comparison stimuli, presented in random order. Each observation interval was marked by a red light on the computer screen that was synchronized with stimulus' presentation. Participants were instructed to report the observation interval containing the target stimulus. Participants were not given any visual feedback after each trial, except for the first five training trials. Participants' responses to these training trials were not taken into consideration in the assessment of double-pass consistency and percent-correct detection.

For each experimental condition, double-pass consistency was calculated as the percentage of agreement between the participants' responses to corresponding trials within the first and the second pass. The accuracy of responses was also calculated as the percentage of correct responses over the 200 trials [100 trials (pass 1) + 100 trials (pass 2)] corresponding to a given experimental condition.

III. RESULTS

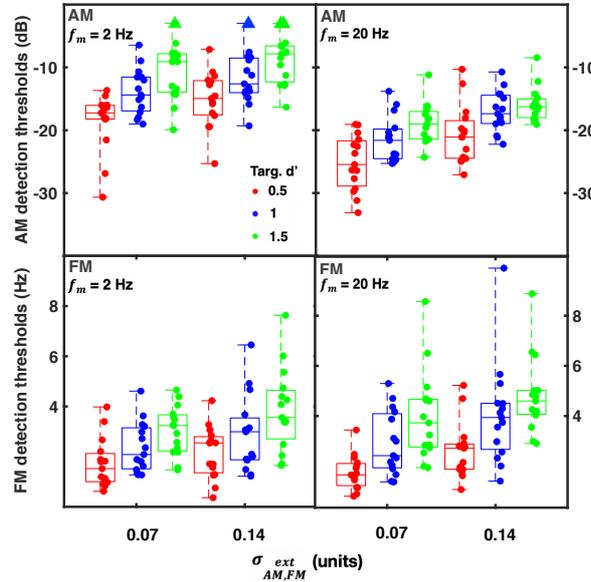


Figure 2. Individual detection thresholds as a function of the standard deviation of the modulation-noise masker ($\sigma_{AM,FM}^{ext}$). Top and bottom panels show AM and FM detection thresholds, respectively.

Left and right panels show the 2-Hz and 20-Hz data, respectively. For each panel, detection thresholds are shown for each level of sensitivity targeted by the adaptive procedure (red: targeted $d'=0.5$; blue: targeted $d'=1$; green: targeted $d'=1.5$). Upward triangles show conditions where participants were not able to perform the AM-detection task. On each box, horizontal lines indicate the median threshold; boxes span the 25th-75th percentiles and vertical lines span the minimum and maximum thresholds.

A. Setup experiment (adaptive tasks)

Individual detection thresholds are plotted as a function of the standard deviation of the modulation-noise masker ($\sigma_{AM,FM}^{ext}$) in each panel of Fig. 2. Top and bottom panels show AM and FM detection thresholds, respectively. Left and right panels show the 2-Hz and 20-Hz data, respectively. For each panel, detection thresholds are shown for each level of sensitivity targeted by the adaptive procedure (red: $d'=0.5$; blue: $d'=1$; green: $d'=1.5$).

Linear mixed-effect models were run with R (R Core Team, 2020) using the lme4 package (Bates et al., 2004). Independent variables tested were: modulation rate [f_m : 2 or 20 Hz], standard deviation of the modulation-noise [$\sigma_{AM,FM}^{ext}$: 0.07 or 0.14 units (m.d.u. or m.i.u.)], and targeted level of sensitivity [d' : 0.5, 1 or 1.5]. The dependent variable was: i) AM detection thresholds (expressed in dB, $20 \log m$), or ii) log-transformed FM detection thresholds. In the

statistical model, fixed effects corresponded to f_m , $\sigma_{AM,FM}^{ext}$ and targeted d' , with an interaction term between these variables (to test all significant interactions). Random effects corresponding to individual participants were introduced in the model to take into account individual variability. All categorical variables were expressed using contrast coding. There were no noticeable deviations from homoscedasticity or normality. p -values were obtained by means of Type III analysis of variance (ANOVA) on linear-mixed models. The significance level was set at 0.05. The results of these statistical analyses are detailed in Tables I- and II of the Appendix section.

1. AM detection thresholds

The analysis revealed, as expected: i) a significant main effect of f_m on AM detection thresholds [$\beta=6.6$; standard error (SE)= 0.8; $\chi^2(1) = 59.8$; $p < 0.0001$], indicating that AM detection thresholds were significantly lower (better) at 20 Hz: AM detection thresholds were on average 7 dB lower at 20 Hz compared to 2 Hz; ii) a significant main effect of σ_{AM}^{ext} on AM detection thresholds [$\beta=4.6$; SE= 0.8; $\chi^2(1) = 4.6$; $p < 0.0001$], indicating that AM detection thresholds increased (worsened) with σ_{AM}^{ext} : AM detection thresholds were on average 7 dB higher at $\sigma_{AM}^{ext}=0.14$ m.d.u. than at $\sigma_{AM}^{ext}=0.07$ m.d.u.; iii) a significant main effect of targeted sensitivity level on AM detection thresholds [$\beta=6.3$; SE= 0.8; $\chi^2(2) = 53.8$; $p < 0.0001$]. There were no significant interactions between independent variables f_m , σ_{AM}^{ext} and d' (all $p > 0.05$).

AM-detection thresholds	χ^2	df	p
Intercept	656.28	1	2.20e-16
σ_{ext}	28.4702	1	9.51e-08
f_m	59.8814	1	1.00e-14
d'	53.8709	2	2.00e-12
$\sigma_{AM}^{ext} * f_m$	0.8539	1	0.35
$\sigma_{AM}^{ext} * d'$	2.0002	2	0.37
$f_m * d'$	1.9334	2	0.38
$\sigma_{AM}^{ext} * f_m * d'$	0.5336	2	0.77

Table I. Analysis of deviance (Type III Wald chi-square tests) showing p -values from lmer models (lme4 package) including AM-detection thresholds as dependent variable, and modulation rate f_m , targeted sensitivity level (d'), $\sigma_{AM,FM}^{ext}$ as independent variables. df stands for degree of freedom.

2. FM detection thresholds

The analysis revealed, as expected: i) a significant main effect of σ_{ext}^{FM} on FM detection thresholds [$\beta=9.7$; $SE=1.9$; $\chi^2(1)=16.2$; $p<0.0001$], indicating that FM detection thresholds increased (worsened) with σ_{ext}^{FM} : FM detection thresholds were on average 1.8 times higher at $\sigma_{ext}^{FM}=0.14$ m.i.u. than at $\sigma_{ext}^{FM}=0.07$ m.i.u; ii) a significant main effect of targeted sensitivity level on FM detection thresholds [$\beta=8$; $SE=1.9$; $\chi^2(2)=67.8$; $p<0.0001$]. There was no significant main effect of f_m on FM detection thresholds and no significant interactions between independent variables f_m , σ_{ext}^{FM} and d' (all $p>0.05$).

FM detection thresholds	χ^2	df	P
Intercept	18.8811	1	1.39e-05
$\sigma_{AM,FM}^{ext}$	16.2447	1	5.67e-05
f_m	0.5465	1	0.46
d'	67.8274	2	1.86e-15
$\sigma_{FM}^{ext} * f_m$	1.9553	1	0.16
$\sigma_{FM}^{ext} * d'$	1.4967	2	0.47
$f_m * d'$	1.5699	2	0.45

Table II. Analysis of deviance table (Type III Wald chi-square tests) showing p-values from lmer models (lme4 package) including FM-detection thresholds as dependent variable, and modulation rate f_m , targeted sensitivity level (d'), $\sigma_{AM,FM}^{ext}$ as independent variables. df stands for degree of freedom.

B. Main experiment (double-pass consistency)

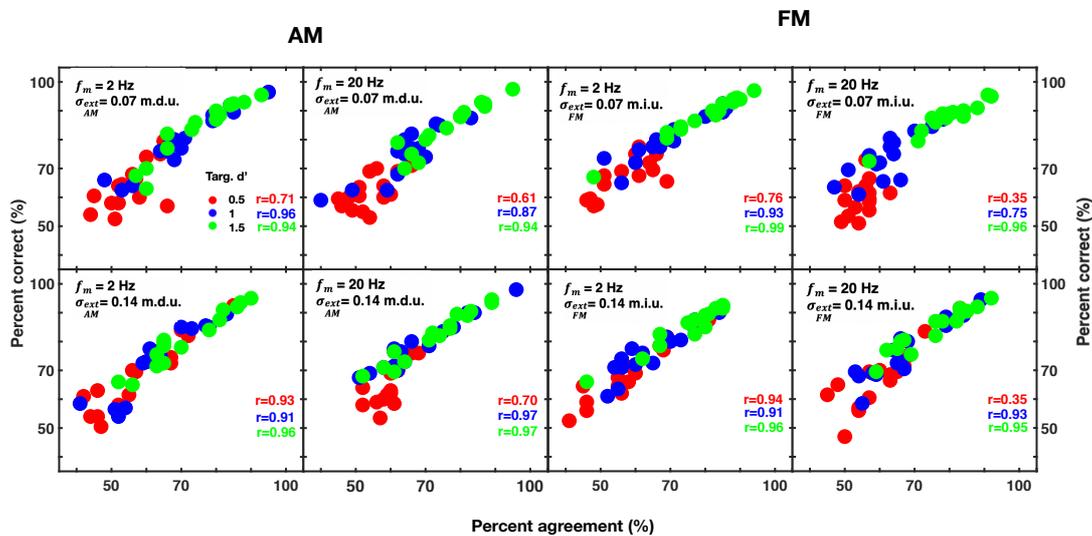


Figure 3. PC versus PA curves showing double-pass consistency (i.e., percent agreement, PA) and performance (i.e., percent correct, PC) data for each participant. The top and bottom panels show the data collected using σ_{ext} of 0.07 and 0.14 units (m.d.u. or m.i.u.), respectively. For each panel, the individual data are shown for each level of sensitivity targeted by the initial adaptive procedure (red: targeted $d'=0.5$; blue: targeted $d'=1$; green: targeted $d'=1.5$).

During the adaptive tasks, three participants (referred to as A, B, C) reached AM detection thresholds above -5 dB in some experimental conditions, suggesting that they were not able to perform the task. They are shown by upward triangles in Fig. 2. More precisely, one participant (A) could not perform the task in two experimental conditions (2-Hz AM detection, $\sigma_{ext} = 0.07$ and 0.14 m.d.u., targeted $d'=1.5$) and two participants (B, C) could not perform the task in a single experimental condition (participant B: 2-Hz AM detection, $\sigma_{ext} = 0.14$ m.d.u., targeted $d'=1$; participant C: 2-Hz AM detection, $\sigma_{ext} = 0.14$ m.d.u., targeted $d'=1.5$). For these three participants, modulation depth was set to -3 dB in the subsequent (double-pass) constant-stimuli experiment in these corresponding conditions. These participants could perform the task in the double-pass experiment well above chance level (PC >75%) and PA scores were not affected. All participants could perform the adaptive task in the FM detection experiments.

Figure 3 shows the double-pass consistency and performance data for each participant. The left and right panels show the 2- and 20-Hz data, respectively. The top and

bottom panels show the data collected using $\sigma_{AM,FM}^{ext}$ of 0.07 and 0.14 units (m.d.u. or m.i.u.), respectively. For each panel, the individual data are shown for each level of sensitivity targeted by the initial adaptive procedure (red: targeted $d'=0.5$; blue: targeted $d'=1$; green: targeted $d'=1.5$). For each experimental condition, a positive correlation was found between PC and PA data: as expected, higher detection performance is associated with higher consistency in listeners' judgments : Bravais-Pearson r ranged between 0.6 and 0.99, except for 20-Hz FM detection where $r=0.35$ for $\sigma_{FM}^{ext} = 0.07$ m.i.u. and targeted $d'=0.5$. Mean PA and PC calculated for each modulation type (AM, FM) and rate (2, 20 Hz) are shown in Table III.

CONDITIONS		AM		FM	
		$f_m= 2$ Hz	$f_m= 20$ Hz	$f_m= 2$ Hz	$f_m= 20$ Hz
$\sigma_{AM,FM}^{ext} = 0.07$ u.	$d'=0.5$	56 (65)	54 (62)	57 (68)	55 (60)
	$d'=1$	70 (79)	64 (75)	69 (80)	62 (74)
	$d'=1.5$	75 (83)	77 (85)	80 (88)	79 (88)
$\sigma_{AM,FM}^{ext} = 0.14$ u.	$d'=0.5$	58 (68)	60 (65)	57 (68)	60 (68)
	$d'=1$	64 (74)	68 (79)	65 (76)	67 (76)
	$d'=1.5$	71 (81)	73 (82)	75 (85)	75 (84)

Table III. Mean percent-agreement (PA, in %) in each condition. Percent correct (PC, in %) is shown between parentheses. Data averaged across the 15 participants.

A statistical analysis of double-pass consistency data was conducted using a linear mixed-effect model with R (R Core Team, 2020) using the lme4 package for each dependent variable (PC or PA converted into rationalized arcsine units). Note that this statistical analysis made on empirical PA and PC data was purely descriptive (*i.e.*, it was not testing a specific hypothesis). Modulation type [AM or FM], f_m [2 or 20 Hz], $\sigma_{AM,FM}^{ext}$ [0.07 or 0.14 units (m.d.u. or m.i.u.)] and targeted d' level [0.5,1 or 1.5] were included as fixed effects together with all interactions between factors; the participants were included as random intercepts. Significance level was set at 0.05 and p -values were obtained using a Type III ANOVA on linear-mixed models.

The analysis conducted on PA scores showed that the main effects of modulation type,

f_m and $\sigma_{AM,FM}^{ext}$ were not significant (all $p > 0.05$). The analysis only showed a significant main effect of targeted d' on PA [$\beta = 10$; $SE = 1.9$; $\chi^2(2) = 55.3$; $p < 0.0001$]. There was also a significant interaction between targeted d' and $\sigma_{AM,FM}^{ext}$ [$\beta = -3$; $SE = 1.9$; $\chi^2(1) = 2.7$; $p = 0.046$] indicating that PA was significantly increased when $\sigma_{AM,FM}^{ext}$ was highest at targeted d' levels of 0.5 and 1. The analysis did not reveal any significant interactions between independent variables modulation type or f_m (all $p > 0.05$).

The analysis conducted on PC scores confirmed that targeted d' had a significant main effect on PC data [$\beta = 15$; $SE = 2.9$; $\chi^2(2) = 72$; $p < 0.0001$]. The main effects of modulation type, f_m and σ_{ext} were not significant (all $p > 0.05$). This was expected as listeners were tested at their individual detection threshold targeting a d' of 0.5, 1 and 1.5 in each experimental condition. The analysis did not reveal any significant interaction between independent variables (all $p > 0.05$). The results of these statistical analyses are detailed in the Table IV of the Appendix section. The outcome of these statistical analyses was not affected by removing the three participants (A, B and C) that could not perform the task in the setup experiment.

IV. MODEL SIMULATIONS

The general goal of the present study was to test whether the modulation-filterbank model was able to account for the double-pass consistency data collected for young normal-hearing participants in the AM and FM detection tasks described above.

A. Model specifications

The model structure was similar to that used by Wallaert *et al.* (2017, 2018), King *et al.* (2019) and Cabrera *et al.* (2019). This model implemented a cascade of processing stages including: peripheral (*i.e.*, cochlear) filtering, half-wave rectification, amplitude compression, short-term adaptation, demodulation (lowpass filtering followed by a modulation filterbank), and a decision stage using a template-matching strategy:

1. A bank of five, 1- ERB_N wide gammatone filters, one centered at the carrier frequency f_c of the stimulus, and the remaining four centered at 1 and 2 Cams (units of the ERB_N number scale; Glasberg and Moore, 1990) above and below the f_c of the stimulus.
2. A 'broken-stick' input-output function for the output of the gammatone filter tuned to the f_c of the stimulus; the function was linear up to a knee-point of 40 dB SPL and compressive (using a power law with an exponent of 0.3) above; the

remaining four gammatone filters were not submitted to this non-linear transform.

3. Half-wave rectification of the output of the five gammatone filters.
4. High-pass filtering (1st order, 6 dB/oct roll-off, 3-Hz cut-off) of the output of all channels to simulate short-term adaptation (cf. Tchorz and Kollmeier, 1999).
5. The resulting signal at the output of each channel was passed to a filterbank (Butterworth filters; -6 dB/oct rolloff) with 10 logarithmically-spaced channels between 2 and 120 Hz (Moore *et al.*, 2009), each with a Q factor of 1 (Lorenzi *et al.*, 2001) to decompose the modulations of the processed signals, producing 50 channels; envelope-phase information was preserved at the output of each bandpass modulation filter. For each modulation filter centered below 10 Hz, the waveform at the output of the filter was passed on for further processing, while for each filter centered at and above 10 Hz only the Hilbert envelope of the output was passed on.
6. Independent Gaussian noises were added to the output of each of the 50 channels; this noise had a constant standard deviation σ_{int} expressed in “model” units, m.u. (Wallaert *et al.*, 2017, 2018; King *et al.*, 2019).
7. The final decision stage was based on a template-matching strategy (Wallaert *et al.*, 2017). The model generated an “optimal” template at the start of each pass with the modulation strength (the dependent variable: m or Δf) set at a starting value of -2 dB for AM detection and 8 Hz for FM detection, and without any added external and internal noise. The template was calculated as the difference between the internal representations of the target and reference stimuli, channel by channel. On each trial, the target and reference stimulus intervals were correlated (channel by channel) with the template. The interval with the largest correlation coefficient (summed across channels) was selected by the model.

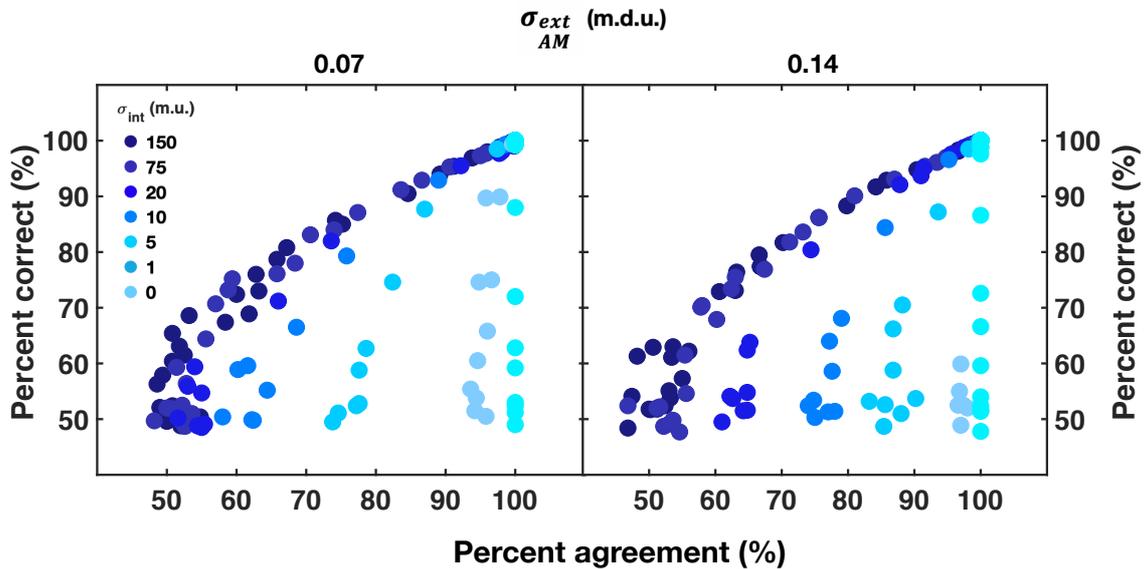


Figure 4. Predictions of the modulation-filterbank model for the “double-pass” AM detection task. A 2-Hz sinusoidal AM was used as the target to illustrate the model’s behavior for this task. The standard deviation of the modulation-noise masker, σ_{ext}^{AM} , was set to 0.07 m.d.u. (left panel) and 0.14 m.d.u. (right panel). The model was simulated for several levels (i.e., standard deviation, expressed in model units, m.u.) of the additive internal noise. Each dot was calculated from 200 trials. For each level of internal noise (as indicated by the hue code, dark blue corresponding to the highest level), double-pass consistency (percent agreement, PA) and performance (percent-correct detection, PC) were simulated for AM depth varying between -40 dB (1%) and -3 dB (70 %) in 1-5 dB discrete steps. For a given PC level, PA decreases with increasing levels of internal noise and it increases with increasing levels of external modulation noise.

Double-pass consistency was simulated with this general model. The model was initially calibrated with a pure tone presented at 40 dB SPL to map “physical dB SPL” into model units (m.u.) and to set the knee point of the amplitude-compression function (see model description, stage 2). At the beginning of each pass, an internal template was generated in the absence of (external) modulation noise masker and by setting the variance of internal noise to zero. The template was generated using supra-threshold values of the modulation depth for AM and the excursion frequency for FM (-2 dB and 8 Hz, respectively). These supra-threshold values corresponded to those used at the beginning of the adaptive staircases to assess modulation detection thresholds for real participants in the setup experiment. The model was always run with the same stimuli and psychophysical procedure used for real participants except that i) the starting phase of target modulations was set to π (instead of being randomized for real participants) because the decision device of the model calculated

the correlation between the internal template and the output of modulation filters as in Wallaert *et al.* (2017); ii) the number of trials that was necessary to obtain reliable estimates of PC and PA scores was increased to 1500 trials per pass (instead of 100 for real participants). Simulations were run using the same Matlab environment (AFC psychoacoustics software, version 1.40.1, Ewert, 2013).

Figure 4 illustrates the model's behavior for this task, and more specifically the effects of external and internal noise on PC and PA in two experimental conditions: 2-Hz AM detection with σ_{AM}^{ext} set to 0.07 m.d.u. (left panel) and 0.14 m.d.u. (right panel). In each experimental condition, the model was simulated for several levels of the internal noise ranging from 0 to 150 m.u. For each level of internal noise, PC and PA were simulated for AM depth varying between -40 dB (1%) and -3 dB (70 %) in 1-5 dB discrete steps. Figure 4 shows that PC and PA are strongly related (Lu and Doshier, 2008): a perfect or near-perfect performance is necessarily associated with perfect or near-perfect agreement; in addition, high PC values cannot be associated with low PA values (if $PA < 100\%$, there must be one of the two passes where the observer makes incorrect responses and therefore $PC < 100\%$). Figure 4 also shows that for a given PC level, PA decreases with increasing levels of internal noise: At high levels of internal noise, the PC versus PA curves are slanted; as internal noise decreases, the PC versus PA curves become steeper. Finally, comparison between the two panels of Fig. 4 shows that PA increases with increasing levels of external noise σ_{AM}^{ext} . In summary, these preliminary simulations indicate that PA is governed by the ratio between internal-noise and external-noise standard deviations: the higher this ratio, the lower PA.

The model was then run for the 24 experimental conditions described above, and for target modulation strengths (m or Δf) set to the mean detection threshold across participants. For each experimental condition, the model was simulated with a range of values of the standard deviation of the additive internal noise, σ_{int} (between 80 and 300 m.u. for AM detection and between 20 and 160 m.u. for FM detection). For each value of σ_{int} , the absolute error between the model and the participants' data (the absolute value of the difference in PA and PC between simulated and real data, referred to the "prediction error" below) was calculated.

B. Simulation results

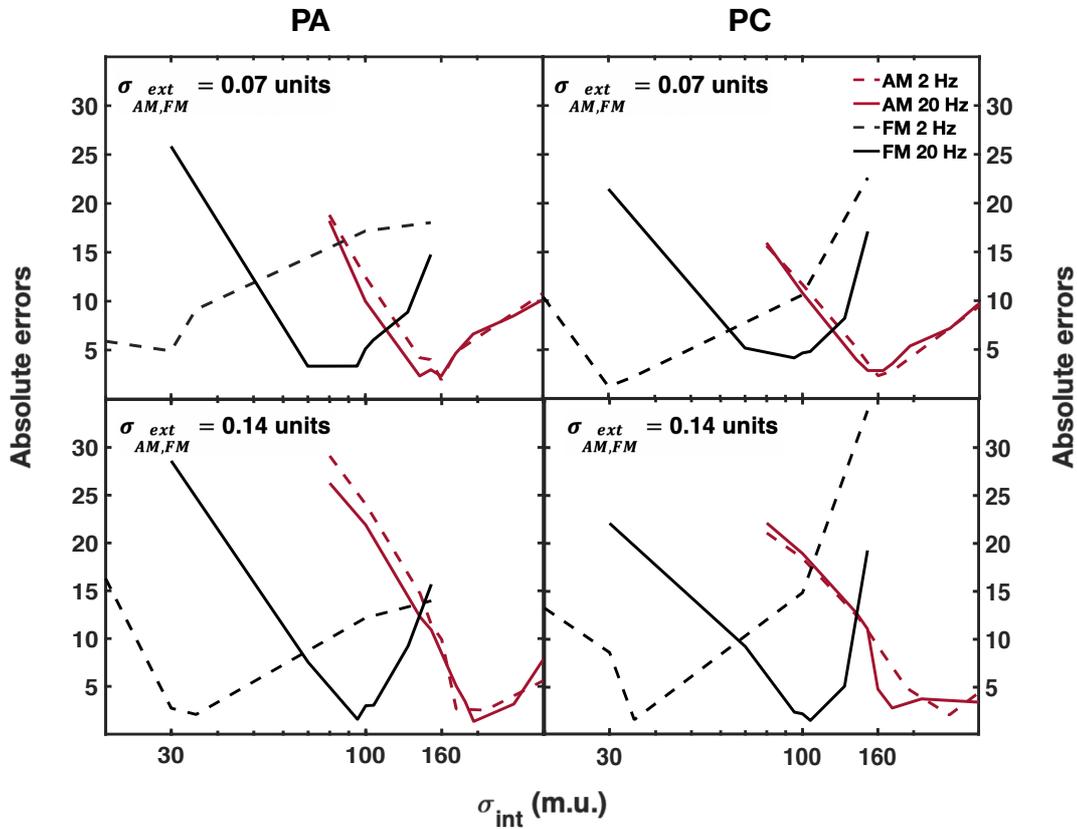


Figure 5. Prediction errors (average across targeted d' levels) of the modulation-filterbank model as a function of σ_{int} for PA (left panel) and PC (right panel) scores. In each panel, prediction errors are shown for 2-Hz AM (dashed dark-red lines), 20-Hz AM (continuous dark-red lines), 2-Hz FM (dashed black lines) and 20-Hz FM (continuous black lines). Top and bottom panels show the model's prediction errors for $\sigma_{AM,FM}^{ext} = 0.07$ and 0.14 units (m.d.u. or m.i.u.), respectively. The x-axis is logarithmic.

Figure 5 shows the model's prediction errors (absolute errors between empirical and simulated data) averaged across targeted d' levels as a function of σ_{int} for PA (left panel) and PC (right panel) scores. In each panel, prediction errors are shown for 2-Hz AM (dashed dark-red lines), 20-Hz AM (continuous dark-red lines), 2-Hz FM (dashed black lines) and 20-Hz FM (continuous black lines). Top and bottom panels show the model's prediction errors for $\sigma_{AM,FM}^{ext} = 0.07$ and 0.14 units (m.d.u. or m.i.u.), respectively. Figure 5 indicates that in each experimental condition, prediction errors calculated for both PA and PC were substantially affected by changes in σ_{int} . In each experimental condition, the model did not predict PA and PC scores equally well (error curves did not overlap perfectly for PA and PC) but prediction

errors dropped below 2-5 points of percentage for both PA and PC for specific values of σ_{int} . The level of σ_{int} that yielded the lowest prediction errors changed with modulation type, modulation rate, and level of the modulation-noise masker. Simulations were run ten times using these values of σ_{int} to check that the model predictions did not change substantially across simulations. The coefficient of variation (CV) was calculated for PA and PC in each experimental condition. CV was generally less than 0.5, indicating that model predictions were quite reproducible. Overall, the level of σ_{int} that yielded the lowest prediction errors for PA and PC was lower for FM than AM, and lower for 2-Hz FM than for 20-Hz FM; for each type of modulation, it was generally lower at the lowest level of the modulation-noise masker. For each experimental condition, values of σ_{int} lower than this level were associated with simulated PA and PC scores higher than empirical scores. Conversely, values of σ_{int} higher than this level were associated with simulated PA and PC scores lower than empirical scores.

The following lines describe the effects of the modulation-noise masker on simulated PA and PC scores. When $\sigma_{ext_{AM}}$ was lowest (0.07 m.d.u.), prediction errors were lower than 5 percentage points for *both* PA and PC for 2- and 20-Hz AM when σ_{int} =160 m.u. When $\sigma_{ext_{AM}}$ was increased to 0.14 m.d.u., higher levels of internal noise were required to reach a minimum in prediction errors for both PA and PC for 2- and 20-Hz AM (σ_{int} =175 and 250 m.u., respectively). When $\sigma_{ext_{FM}}$ was lowest (0.07 m.i.u.), a level of internal noise of 30 m.u. was associated with prediction errors lower than 5 points of percentage for both PA and PC for 2-Hz FM; in comparison, a level of 100 m.u. was required to reach a minimum in prediction errors for both PA and PC for 20-Hz FM. When $\sigma_{ext_{FM}}$ was increased to 0.14 m.i.u., a higher level of internal noise was required for 2-Hz FM only (σ_{int} =35 m.u.). It is noteworthy that an increase in $\sigma_{ext_{AM}}$ was associated with an over-estimation of PA scores for AM; an increase in $\sigma_{ext_{FM}}$ was associated with an over-estimation of PC for FM. Therefore, the modulation-noise masker had more impact on simulated PA and PC scores than on real scores.

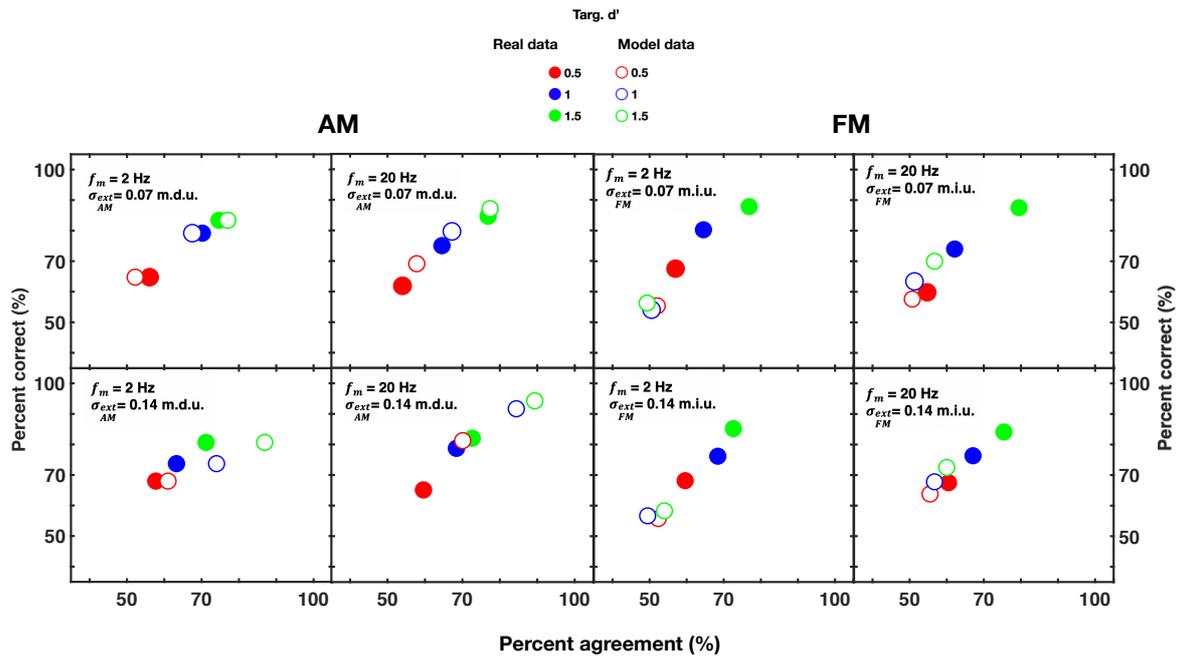


Figure 6. PC versus PA data for AM and FM detection. Each panel shows the simulation data obtained with $\sigma_{int} = 160$ m.u. plotted along with the empirical data (mean across participants; closed circles). In each panel, the simulation (closed circles) and real data (open circles) are shown for each level of sensitivity targeted by the initial adaptive procedure (red: targeted $d' = 0.5$; blue: targeted $d' = 1$; green: targeted $d' = 1.5$).

In Fig. 5, model's prediction errors were calculated with a range of values of the standard deviation of the additive internal noise, σ_{int} , and averaged across the three d' levels targeted by the initial adaptive procedure. Figure 6 shows the simulated PA and PC scores for each targeted d' level (red: targeted $d' = 0.5$; blue: targeted $d' = 1$; green: targeted $d' = 1.5$) when a fixed level of internal noise was used, that is for $\sigma_{int} = 160$ m.u. In each panel, the simulated data (open circles) are plotted along with the empirical data (mean across participants; closed circles). As indicated above, the model predicted accurately PA and PC scores for 2- and 20-Hz AM when σ_{ext} was lowest (0.07 m.d.u.). For 2-Hz AM, increasing σ_{ext} to 0.14 m.d.u. yielded accurate predictions for PC scores, but PA scores were over-estimated (*i.e.*, better than observed in real participants). For 20-Hz AM, both PA and PC scores were over-estimated when $\sigma_{ext} = 0.14$ m.d.u. The model failed to account for FM data. For 2-Hz FM, the model predicted PA and PC scores at or just above chance level (<60 %). For 20-Hz FM, the predicted PC scores increased up to about 70 % for a targeted d' of 1.5; however, predicted PA scores remained close to chance level, inconsistent with the real data. Thus, the model failed when the internal noise was fixed. In other words, different levels of internal noise were required to fit all the data.

Taken together, these results indicate that the modulation-filterbank model was able to reproduce accurately double-pass consistency data for both slow and fast AM at the lowest level of the modulation-noise masker (here, $\sigma_{ext}^{AM}=0.07$ m.d.u.) but it was more sensitive to changes in the level of the modulation-noise masker than the real participants. Still, the modulation-filterbank model could not account for double-pass consistency data for both AM and FM detection using a single source (and level) of additive internal noise.

V. DISCUSSION

1. Separate source(s) of internal variability for AM and FM detection at low carrier frequencies

a. Separate sources of internal noise for AM and FM detection. Our crude implementation of the “modulation-filterbank model” using a single source (and thus, a single level) of additive internal noise gave a good account of consistency *and* detection performance of listeners’ judgements (mean PA and PC data across listeners) for both slow (2 Hz) and fast (20 Hz) AM when external variability conveyed by the modulation-noise masker was relatively low ($\sigma_{ext}^{AM}=0.07$ m.d.u.). The model could also account for slow and fast FM detection but for levels of the internal noise smaller than that yielding best model predictions for AM. Moreover, different levels of internal noise were required to predict performance and consistency of listeners’ judgements for slow and fast FM detection.

Taken together, these results are consistent with the idea that the same sources of sensory and cognitive variability constrain slow and fast AM processing. These results add to those of recent studies demonstrating the capacity of the modulation-filterbank model to account for a large set of detection, discrimination or identification data collected in psychophysical tasks where listeners are expected to rely on temporal-envelope cues (*e.g.*, Biberger and Ewert, 2016). It is important to note that the present study tested for the first time the ability of the modulation-filterbank model to reproduce the *consistency* of human auditory judgements in AM detection tasks in addition to accuracy.

Our model could not provide a unified account of AM *and* FM detection at a low audio frequency (500 Hz). These results are therefore inconsistent with the notion that slow and fast FM are detected via a simple frequency-to-place (FM-to-AM) conversion mechanism operating at the cochlear level (Whiteford *et al.*, 2017, 2020). More precisely, these results show that FM-to-AM conversion cues resulting from cochlear filtering cannot be used by a modulation filterbank to detect either slow or fast FM. Other mechanisms and thus, other sources of internal variability make a noticeable contribution to FM detection irrespective of FM rate. Recent electrophysiological data collected on guinea pigs revealed that both slow

and fast FM are well encoded in the phase-locked discharge patterns of cochlear-nucleus neurons (Paraouty *et al.*, 2018). Therefore, these sources of variability may be related to those constraining the coding of temporal fine-structure (TFS) cues *via* neural phase locking in auditory-nerve fibers and brainstem neurons (*e.g.*, Moore and Sek, 1992). Alternatively, these sources of variability may be related to those constraining a more *central* mechanism distinct from those involved in monaural AM detection. It is still possible that the latter may be using out-of-phase temporal-envelope cues at different cochlear place, but they would require a neural machinery different from that used to detect AM (this issue is discussed further below). The effects of the mechanism responsible for FM detection on behavioral variability - whether it corresponds to neural phase locking to TFS cues or not - may be estimated within the present framework in terms of an “equivalent” internal noise in the temporal-envelope domain. The level of this “equivalent” internal noise is smaller (by a factor 1.5-5) than that constraining AM detection. This suggests that this mechanism is *more efficient* than that operating on temporal-envelope fluctuations at the output of cochlear filters.

b. Mechanisms of FM detection at low carrier frequencies. It is perhaps surprising that the modulation-filterbank model proved unable to predict FM detection scores (PC) above chance level in the present experimental conditions, knowing that FM detection thresholds at fast (20 Hz) rate could be predicted relatively well in previous work (Wallaert *et al.*, 2018; King *et al.*, 2019) using a comparable implementation of the modulation-filterbank model. The reason for this might be the addition of the (FM) modulation-noise masker in the present FM detection experiments, a crucial aspect of the double-pass design (Green, 1964). Previous work used either no masker or a deterministic (sinusoidal) AM masker when measuring and predicting FM detection thresholds. It follows that the stochastic fluctuations introduced by the FM-noise masker prevented the template-matching decision module of the current model to make efficient use of the temporal-envelope cues resulting from FM-to-AM conversion at the output of cochlear filters. To examine this possibility, an additional model simulation was carried out. The current implementation of the modulation-filterbank model was simulated in the absence of external modulation noise (that is by setting σ_{ext} to 0 units) in the adaptive tasks (*cf.* setup experiment). Here, we used the level of variance of internal noise $\sigma_{int}=160$ m.u. that minimized the model's prediction errors for the double-pass experiment for 2-Hz and 20-Hz AM in the presence of low external modulation noise ($\sigma_{ext} = 0.07$ m.d.u.; *cf.* Fig. 5). When there was no external noise, the model accounted for all conditions except slow FM. When external noise was applied, the model was successful for AM but not for either FM condition. More precisely, our model reproduced unmasked AM-detection thresholds at 2 and 20 Hz very well (-18.2 and -25.8 dB, respectively, when targeted $d'=1$; for comparison: -17 and -24 dB for real listeners in King *et al.* (2019) in comparable conditions) and unmasked FM-detection thresholds at 20 Hz (3.3 Hz when targeted $d'=1$; for comparison: 2.3 Hz for real listeners in King *et al.* (2019) in comparable conditions). As in King *et al.* (2019), our model underestimated real unmasked 2-Hz FM-detection thresholds by a large amount (7.4 Hz when targeted $d'=1$; for comparison: 1.8 Hz for real listeners in King *et al.* (2019) in comparable

conditions). Our model was also run using 2 and 20-Hz AM and FM masked by low external modulation noise ($\sigma_{AM,FM}^{ext} = 0.07$ units) using $\sigma_{int} = 160$ m.u. Our model reproduced masked AM-detection thresholds at 2 and 20 Hz very well (-14.2 and -20.7 dB, respectively, when targeted $d' = 1$; for comparison: -13.8 and -21.4 dB for the real listeners tested here). Not surprisingly, our model underestimated masked FM-detection thresholds by a substantial amount at 2 and 20 Hz (12 and 6 Hz, respectively, when targeted $d' = 1$; for comparison: 2.4 and 3 Hz, respectively for the real listeners tested here). These simulation data are shown in Table V (simulation data are also shown for targeted d' of 0.5 and 1.5).

In summary, the modulation-filterbank model as implemented here could successfully account for slow and fast-AM detection measured in the presence or absence of modulation maskers with a random structure. The modulation-filterbank model could also account for fast-FM detection in the absence of masker or in the presence of a deterministic masker (*e.g.*, a sinusoidal AM) on the sole basis of FM-to-AM conversion cues resulting from cochlear filtering. In other words, the model could use the out-of-phase envelope patterns elicited by FM at the output of cochlear filters tuned to frequencies below and above the carrier frequency to detect fast FM at a performance level similar to that of real listeners. The modulation-filterbank model therefore gives a parsimonious account of the fast-FM data measured in these conditions: no additional mechanism - such as a cross-channel mechanism - is required to account for these behavioral data. However, the model clearly fails when fast FM is presented against a masker with random structure, indicating that FM-to-AM conversion does not convey robust cues for FM detection in noisy conditions, and that an additional mechanism distinct from that involved in AM detection is required to account for performance and sensitivity of real listeners in this case.

It has been assumed that even at low carrier frequencies (<2-4 kHz) where phase locking information is available, FM detection at high rates is mediated by temporal-envelope cues resulting from FM-to-AM conversion at the output of cochlear filters because auditory processing of TFS cues may be too sluggish (*e.g.*, Moore and Sek, 1994). Still, electrophysiological (*i.e.*, single-unit) data collected on guinea pigs indicate that fast FM carried by low-frequency tones is accurately encoded in terms of neural phase-locking to TFS cues in the low brainstem, and more precisely, in the ventral cochlear nucleus (Paraouty *et al.*, 2018). Thus, it might be the case that at threshold, FM detection at high rates is mainly mediated by (salient) temporal-envelope cues resulting from FM-to-AM conversion when FM signals are presented in quiet or masked by deterministic (that is, non-random) sounds such as a sinusoidal AM masker (as in Wallaert *et al.* (2018) or King *et al.*, 2019); however, neural TFS cues - although weaker because of sluggish processing at more central stages - may come into play (in addition to temporal-envelope cues) when fast-FM signals are corrupted by concurrent sounds showing random fluctuations in instantaneous amplitude and frequency, as in the present study. In other words, temporal-envelope cues resulting from FM-to-AM conversion may be useful to detect fast FM but they may not be very useful in the presence

of masking sounds with random structure. This interpretation is compatible with the idea that processing of TFS cues at a relatively central stage of the auditory system is sluggish: it mainly suggests that neural TFS cues evoked by fast FM - although weaker because of such central limitations - may help listeners when temporal-envelope cues cannot be used efficiently by decision-making mechanisms.

Still, it is not possible to exclude the possibility that a central “cross-channel” mechanism using out-of-phase envelope cues at different cochlear place is responsible for both slow and fast FM detection. Neurophysiological and modelling work is required to determine the plausibility of such a mechanism.

2. Characterization of internal noise in the temporal-envelope domain

a. A Gaussian internal noise ? Internal noise was modelled as additive and it followed a Gaussian distribution. Still, it may be the case that other statistical distributions could yield better predictions. For instance, Neri (2013) showed that internal noise for visual detection tasks followed a leptokurtic distribution (a distribution with kurtosis larger than that of a normal distribution) instead of a normal one. Such a distribution might result from slow changes in the variance of the distribution over time (P. Neri, personal communication; see also Goris et al., 2014, for the neural basis of this phenomenon). Further work is thus warranted to explore this possibility.

b. Nature of internal noise. In the model, internal noise was added at the output of modulation filters, before the decision device. From the current experiments, it is unfortunately impossible to infer whether this source of internal variability is of sensory or cognitive (*i.e.*, attentional, memory) origin, and whether it results from a single source of neural variability (*e.g.*, spontaneous activity in central neurons tuned to specific temporal-envelope rates) or from many separate sources located at different (peripheral and central) stages of the auditory system.

The double-pass procedure was developed to assess the influence of the total amount of internal noise, both additive and multiplicative, relative to external noise, on perceptual processes (Green, 1964; Spiegel and Green, 1981; Lu and Doshier, 2008). The model used an additive (that is a fixed-variance) internal noise only. It may be that the relative inability of the model to account for the effects of external noise on PA and PC scores using a single level of additive internal noise resulted from the absence of multiplicative internal noise in the model, *i.e.* an additional source of internal variability whose variance would be proportional to the envelope power of the stimuli. This would be consistent with the fact that the standard deviation of internal noise that yielded best model predictions increased with the standard deviation of the modulation-noise masker. Indeed, Ewert and Dau (2004) previously showed that a multiplicative noise constrains temporal-envelope processing and is required to

simulate the Weber's law behavior for AM-depth discrimination and the absence of certain carrier effects (*i.e.*, random versus frozen broadband noise carriers) on AM detection. Additional work is thus necessary to characterize further the effects of external variability on AM detection and the contribution of a multiplicative internal-noise source to consistency in AM detection.

3. Limitations and novelty of the present approach

As pointed out in the Results section (see Fig. 4), response consistency and performance co-vary (Lu and Doshier, 2008). Therefore, the same conclusion (*i.e.*, separate sources of internal noise, and thus separate mechanisms constrain AM and FM detection) would have been reached by measuring performance only, as indicated by comparable (although not identical) performance and consistency curves in Fig. 5. This could be interpreted as an important limitation of the present approach based on the double-pass paradigm. Still, it is crucial to keep in mind that the primary goal of this study was to give a *unified* account of AM and FM perception *both* in terms of response consistency and performance. This is why consistency and performance curves shown in Fig. 5 should not be read separately but in conjunction. The important result here is that, for a given level of external noise, the same level of internal noise (*e.g.*, 160 m.u.) yields best predictions for *both* consistency and performance data. It was not clear from the outset that the modulation-filterbank model could account for both performance and consistency of human judgments in slow and fast AM detection tasks, as pointed out by the simulation results shown in Fig. 4. Demonstrating this was an important step confirming further the validity of this modelling approach for AM processing, adding to the outcome of previous studies demonstrating the capacity of the modulation-filterbank model to account for a large set of psychophysical data collected in tasks where listeners are expected to rely on temporal-envelope cues (*e.g.*, Biberger and Ewert, 2016). This demonstration was also a prerequisite when interpreting the incapacity of the modulation-filterbank model to account for FM detection (consistency and performance) data for slow and fast FM presented against noisy maskers. In particular, it was not anticipated that the modulation-filterbank model would not be able to account for consistency and performance data in the fast FM condition, because previous simulation studies of Wallaert *et al.* (2018) and King *et al.* (2019) showed that this model could predict successfully fast FM detection thresholds in the presence or absence of a sinusoidal AM masker.

4. Implications for the assessment of the perceptual consequences of ageing and cochlear damage

The present study was conducted with young normal-hearing participants. Some studies aiming to model AM detection and discrimination for elderly listeners and listeners with sensorineural hearing loss suggest that the variance of the additive internal noise may increase substantially (up to a factor 6-10) because of ageing and cochlear lesions (Derleth *et al.*, 2001; Ives *et al.*, 2014; Paraouty *et al.*, 2016; Wallaert *et al.*, 2017, 2018). From a behavioral perspective, this increase in internal noise should yield a decrease in the consistency of the responses of elderly and hearing-impaired persons, as measured in the current double-pass paradigm. However, to the best of our knowledge, this consistency has never been investigated in these two populations. The degree of consistency of responses and the estimation of internal noise could be of clinical interest by complementing the performance and sensitivity measures traditionally used to characterize suprathreshold deficits caused by sensorineural hearing loss and presbycusis. These measures may be related to the degree of neural de-afferentation (or “synaptopathy”) caused by ageing and noise exposure (Lopez-Poveda, 2014; Marmel *et al.*, 2015), which is supposed to affect mainly the processing of high sound amplitudes, and consequently the neural coding of temporal-envelope cues (Furman *et al.*, 2013; Bharadwaj *et al.*, 2014). Consistent with this hypothesis, a recent modelling study by Goodman *et al.* (2017) indicates that massive loss of auditory-nerve fibers results in an increase in neural variability and alterations of AM coding at the low brainstem (cochlear nucleus) level.

In that respect, it is possible that a better understanding of the impact of aging and cochlear lesions on double-pass consistency and internal noise would allow for a more accurate estimation of the limits of signal amplification and speech-enhancement systems applied by current digital hearing aids.

VI. SUMMARY AND CONCLUSIONS

Response consistency and performance in a modulation-detection task were estimated for young, normal-hearing participants using a double-pass paradigm and sinusoidal AM and FM targets masked by a modulation-noise masker. The double-pass data were simulated using a computational model of temporal-envelope processing implementing a peripheral filterbank, half-wave rectification, instantaneous amplitude compression, adaptation, a filterbank of bandpass modulation filters, an additive internal noise and a template-matching decision device.

The results showed that :

- 1) Distinct levels (*i.e.*, standard deviations) of additive internal noise were required to reproduce AM and FM data but a single level of internal noise could account for the 2- and 20-Hz AM data.
- 2) The level of internal noise that yielded best model predictions increased with the standard deviation of the modulation-noise masker.
- 3) The level of internal noise yielding best model predictions for FM detection was smaller (by a factor 1.5-5) than the one yielding best model predictions for AM detection.

These results suggest that distinct sources of internal variability are involved for AM and FM detection at low audio frequencies. They also suggest that AM detection may be constrained by an additional source of internal noise whose variance is proportional to the envelope power of the stimuli, and that the mechanism involved in FM detection is more efficient than the one operating on temporal-envelope fluctuations at the output of cochlear filters.

In conclusion, the current study based on the double-pass paradigm showed that the modulation filterbank-model is able to give a unified account of slow and fast AM detection both in terms of response consistency and accuracy. This demonstrates further the capacity of this modelling approach to account for a large set of perceptual data. However, an additional mechanism distinct from that involved in AM detection must be considered to account for slow and fast FM detection at low carrier frequencies.

ACKNOWLEDGMENTS

This work was supported by two grants from the Collège National des Audioprothésistes (France) and Entendre SAS. This work was also supported by ANR-17-EURE-0017, ANR-16-CE28-0016 and ANR-19-CE28-0010-01 . The authors wish to thank warmly Peter Neri for helpful discussions about the double-pass paradigm, J Bernstein and two anonymous reviewers for very constructive feedback on the present article.

VII. REFERENCES

- Amitay, S., Guiraud, J., Sohoglu, E., Zobay, O., Edmonds, B. A., Zhang, Y. X., and Moore, D. R. (2013). "Human decision making based on variations in internal noise: An EEG study," *PloS one*, **8**, e68928. doi: 10.1371/journal.pone.0068928.
- ANSI, S. (1996). *Determination of Occupational Noise Exposure and Estimation of Noise-Induced Hearing Impairment*. New York: American National Standards Institute.
- Bharadwaj, H.M., Verhulst, S., Shaheen, L., Liberman, M.C., and Shinn-Cunningham, B.G. (2014). "Cochlear neuropathy and the coding of supra-threshold sound," *Frontiers Syst. Neurosci.*, **21**, 8-26.
- Barlow, H. B. (1956). "Retinal noise and absolute threshold," *J. Opt. Soc. Am.*, **46**, 634-639.
- Bates, D., Maechler, M., and Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes* [Software]. Retrieved from CRAN. R-project. org/package=lme4.
- Biberger, T., and Ewert, S. D. (2016). "Envelope and intensity based prediction of psychoacoustic masking and speech intelligibility," *J. Acoust. Soc. Am.*, **140**, 1023-1038.
- Biberger, T., and Ewert, S. D. (2017). "The role of short-time intensity and envelope power for speech intelligibility and psychoacoustic masking," *J. Acoust. Soc. Am.*, **142**, 1098-1111.
- Burgess, A. E., and Colborne, B. (1988). "Visual signal detection. IV. Observer inconsistency," *J. Opt. Soc. Am. A*, **5**, 617-627.
- Cabrera, L., Varnet, L., Buss, E., Rosen, S. and Lorenzi, C. (2019). "Development of temporal auditory processing in childhood: Changes in efficiency rather than temporal-modulation selectivity," *J. Acoust. Soc. Am.*, **146**, 2415-2429.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997a). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," *J. Acoust. Soc. Am.*, **102**, 2892-2905.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997b). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," *J. Acoust. Soc. Am.*, **102**, 2906-2919.
- Derleth, R. P., Dau, T., and Kollmeier, B. (2001). "Modeling temporal and compressive properties of the normal and impaired auditory system," *Hear. Res.*, **159**, 32-149.
- Ewert, S.D. (2013). "AFC - A modular framework for running psychoacoustic experiments and computational perception models," in *Proceedings of the International Conference on Acoustics AIA-DAGA 2013*, Merano, Italy, pp. 1326-1329.
- Ewert, S. D., and Dau, T. (2000). "Characterizing frequency selectivity for envelope fluctuations," *J. Acoust. Soc. Am.*, **108**, 1181-1196.
- Ewert, S. D., and Dau, T. (2004). "External and internal limitations in amplitude-modulation processing," *J. Acoust. Soc. Am.*, **116**, 478-490.
- Ewert, S. D., Paraouty, N., and Lorenzi, C. (2018). "A two-path model of auditory modulation detection using temporal fine structure and envelope cues," *Eur. J. Neurosci.* doi: 10.1111/ejn.13846.
- Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008). "Noise in the nervous system," *Nature Rev. Neurosci.*, **9**, 292-303.
- Furman, A. C., Kujawa, S.G., and Liberman, M.C. (2013). "Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates," *J. Neurophysiol.*, **110**, 577-586.
- Garcia-Pérez, M. A. (1998). "Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties," *Vis. Res.*, **38**, 1861-1881.
- Glasberg, B. R., and Moore, B. C. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.*, **47**, 103-138.

- Goodman, D. F. M., Winter, I. M., Léger, A.C., de Cheveigné, A., and Lorenzi, C. (2017). "Modelling firing regularity in the ventral cochlear nucleus: Mechanisms, and effects of stimulus level and synaptopathy," *Hear. Res.*, **358**, 98-110.
- Goris, R. L., Movshon, J. A., and Simoncelli, E. P. (2014). "Partitioning neuronal Variability," *Nature Neurosci.*, **17**, 858.
- Green, D. M. (1964). "Consistency of auditory detection judgments," *Psychol. Rev.*, **71**, 392-407.
- Green, D. M., and Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). New York: Wiley.
- Hasan, B. A. S., Joosten, E., and Neri, P. (2012). "Estimation of internal noise using double-passes: Does it matter how the second pass is delivered?" *Vis. Res.*, **69**, 1-9.
- Ives, D. T., Kalluri, S., Strelcyk, O., Sheft, S., Miermont, F., Coez, A., ... and Lorenzi, C. (2014). "Effects of noise reduction on AM perception for hearing-impaired listeners.," *J. Assoc. Res. Otolaryngol.*, **15**, 839-848.
- Javel, E., and Viemeister, N. F. (2000). "Stochastic properties of cat auditory nerve responses to electric and acoustic stimuli and application to intensity discrimination," *J. Acoust. Soc. Am.*, **107**, 908-921.
- Jepsen, M. L., Ewert, S. D., and Dau, T. (2008). "A computational model of human auditory signal processing and perception," *J. Acoust. Soc. Am.*, **124**, 422-438.
- Jones, P. R., Moore, D. R., Amitay, S., and Shub, D. E. (2013). "Reduction of internal noise in auditory perceptual learning," *J. Acoust. Soc. Am.*, **133**, 970-981.
- Jørgensen, S., Ewert, S. D., and Dau, T. (2013). "A multi-resolution envelope-power based model for speech intelligibility," *J. Acoust. Soc. Am.*, **134**, 436-446.
- Joris, P. X., Schreiner, C. E., and Rees, A. (2004). "Neural processing of amplitude modulated sounds," *Physiol. Rev.* **84**, 541-577.
- King, A., Varnet, L., and Lorenzi, C. (2019). "Accounting for masking of frequency modulation by amplitude modulation with the modulation filter-bank concept," *J. Acoust. Soc. Am.*, **145**, 2277-2293.
- Lopez-Poveda, E. A. (2014). "Why do I hear but not understand? Stochastic undersampling as a model of degraded neural encoding of speech," *Frontiers Neurosci.*, **30**; 8:348. doi: 10.3389/fnins.
- Lorenzi, C., Soares, C., and Vonner, T. (2001). "Second-order temporal modulation transfer functions," *J. Acoust. Soc. Am.*, **110**, 1030-1038.
- Lu, Z. L., and Doshier, B. A. (2004). "Perceptual learning retunes the perceptual template in foveal orientation identification," *J. Vis.*, **4**, 5-5.
- Lu, Z. L., Chu, W., and Doshier, B.A. (2006). "Perceptual learning of motion direction discrimination in fovea: Separable mechanisms," *Vis. Res.*, **46**, 2315-2327.
- Lu, Z. L., and Doshier, B.A. (2008). "Characterizing observers using external noise and observer models: assessing internal representations with external noise," *Psychol. Rev.*, **115**, 44.
- Marmel, F., Rodríguez-Mendoza, M.A., and Lopez-Poveda, E.A. (2015). "Stochastic undersampling steepens auditory threshold/duration functions: implications for understanding auditory deafferentation and aging," *Frontiers Aging Neurosci.*, **15**; 7:63. doi: 10.3389/fnagi.
- McDermott, J. H., and Simoncelli, E. P. (2011). "Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis," *Neuron*, **71**, 926-940.
- Moore, B. C., Füllgrabe, C., and Sek, A. (2009). "Estimation of the center frequency of the highest modulation filter," *J. Acoust. Soc. Am.*, **125**, 1075-1081.

- Moore, B. C., & Sek, A. (1992). Detection of combined frequency and amplitude modulation. *The Journal of the Acoustical Society of America*, **92**, 3119-3131.
- Moore, B. C. J., and Sek, A. P. (1994). "Discrimination of modulation type (amplitude modulation or frequency modulation) with and without background noise," *J. Acoust. Soc. Am.*, **96**, 726-732.
- Moore, B. C. J., and Sek, A. P. (1996). "Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking," *J. Acoust. Soc. Am.*, **100**, 2320-2331.
- Moore, B. C. J., Şek, A., Vinay, and Füllgrabe, C. (2019). "Envelope regularity discrimination," *J. Acoust. Soc. Am.*, **145**, 2861-2870.
- Moore, B. C. J., and Şek, A. (2019). "Discrimination of the phase of amplitude modulation applied to different carriers: Effects of modulation rate and modulation depth for young and older subjects," *J. Acoust. Soc. Am.* **146**, 1696-1704.
- Moore, B. C. J., Mariathasan, S., and Şek, A.P. (2019). "Effects of Age and Hearing Loss on the Discrimination of Amplitude and Frequency Modulation for 2- and 10-Hz Rates," *Trends Hear.* doi: 10.1177/2331216519853963.
- Neri, P. (2010). "How inherently noisy is human sensory processing?" *Psych. Bull. Rev.*, **17**, 802-808.
- Neri, P. (2013). "The statistical distribution of noisy transmission in human sensors," *J. Neural Eng.* **10**, doi: 10.1088/1741-2560/10/1/016014.
- Palmer, A. R., and Russell, I. J. (1986). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," *Hear. Res.*, **24**, 1-15.
- Paraouty, N., Ewert, S. D., Wallaert, N., and Lorenzi, C. (2016). "Interactions between amplitude modulation and frequency modulation processing: Effects of age and hearing loss," *J. Acoust. Soc. Am.*, **140**, 121-131.
- Paraouty, N., and Lorenzi, C. (2017). "Using individual differences to assess modulation-processing mechanisms and age effects," *Hear. Res.*, **344**, 38-49.
- Paraouty, N., Stasiak, A., Lorenzi, C., Varnet, L., and Winter, I.M. (2018). "Dual coding of frequency modulation in the ventral cochlear nucleus," *J. Neurosci.*, **38**, 4123-4137.
- Pelli, D. G., and Blakemore, C. (1990). "The quantum efficiency of vision," *Vision: Coding and efficiency*, 3-24.
- Parthasarathy, A., Hancock, K. E., Bennett, K., DeGruttola, V., and Polley, D. B. (2020). "Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing," *Elife*, **9**, e51419.
- R. Core Team. (2018). *R Foundation for Statistical Computing; Vienna, Austria: 2015. R: A language and environment for statistical computing*, 2013.
- Saberi, K., and Hafter, E. R. (1995). "A common neural code for frequency-and amplitude-modulated sounds," *Nature*, **374**, 537-539.
- Shannon, R.V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science*, **270**, 303-304.
- Singh, N. C., and Theunissen, F. E. (2003). "Modulation spectra of natural sounds and ethologic theories of auditory processing," *J. Acoust. Soc. Am.*, **114**, 3394-3411.
- Spiegel, M. F., and Green, D. M. (1981). "Two procedures for estimating internal noise," *J. Acoust. Soc. Am.*, **70**, 69-73.
- Strickland, E. A., and Viemeister, N.F. (1996). "Cues for discrimination of envelopes," *J. Acoust. Soc. Am.*, **99**, 3638-3646.

- Tchorz, J., and Kollmeier, B. (1999). "A model of auditory perception as front end for automatic speech recognition," *J. Acoust. Soc. Am.*, **106**, 2040-2050.
- Thoret, E., Varnet, L., Boubenec, Y., F  rri  re, R., Le Tourneau, F. M., Krause, B., and Lorenzi, C. (2020). "Characterizing amplitude and frequency modulation cues in natural soundscapes: A pilot study on four habitats of a biosphere reserve," *J. Acoust. Soc. Am.*, **147**, 3260-3274.
- Viemeister, N. F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am.*, **66**, 1364-1380.
- Vilidaite, G., and Baker, D. H. (2017). "Individual differences in internal noise are consistent across two measurement techniques," *Vis. Res.*, **141**, 30-39.
- Wallaert, N., Varnet, L., Moore, B. C., and Lorenzi, C. (2018). "Sensorineural hearing loss impairs sensitivity but spares temporal integration for detection of frequency modulation," *J. Acoust. Soc. Am.*, **144**, 720-733.
- Whiteford, K. L., and Oxenham, A. J. (2015). "Using individual differences to test the role of temporal and place cues in coding frequency modulation," *J. Acoust. Soc. Am.*, **138**, 3093-3104.
- Whiteford, K. L., Kreft, H. A., and Oxenham, A. J. (2017). "Assessing the role of place and timing cues in coding frequency and amplitude modulation as a function of age," *J. Assoc. Res. Otolaryngol.*, **18**, 619-633.
- Whiteford, K. L., Kreft, H. A., and Oxenham, A. J. (2020). "The role of cochlear place coding in the perception of frequency modulation," *Elife*. 30(9): e58468. doi: 10.7554/eLife.58468.
- Wright, B. A., and Dai, H. (1998). "Detection of sinusoidal amplitude modulation at unexpected rates," *J. Acoust. Soc. Am.*, **104**, 2991-2996.
- Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargava, A. and Cao, K. (2005). "Speech recognition with amplitude and frequency modulations," *Proc. Nat. Acad. Sci. USA*, **102**, 2293-2298.
- Zwicker, E. (1952). "Die Grenzen der H  rbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones," *Acta Acust. united with Acust.*, **2**, 125-133.

Chapitre III:

Application du paradigme de double-passe à la détection d'AM et de FM chez des sujets normo-entendants jeunes et âgés

III.1. No effect of age on sensitivity and double-pass consistency for amplitude-and frequency-modulation detection

III.1.a. Résumé

Several studies suggest that ageing impairs auditory perception of amplitude and frequency-modulation (AM, FM). If these deficits result from increased internal noise caused by cochlear synaptopathy or central changes, not only sensitivity but also response consistency should decrease with age in modulation-detection tasks. These two aspects of behavior were estimated for normal-hearing subjects aged between 40 and 62 years, using 2 and 20-Hz sinusoidal AM or FM applied to a 500-Hz pure-tone carrier. Target AM and FM were masked by an AM and FM noise, respectively. Modulation-detection thresholds were first measured using an adaptive, forced-choice procedure. Double-pass consistency for modulation detection was then measured using a constant-stimuli, forced-choice task repeated twice with the same stimuli. Subjects were tested at their individual modulation-detection threshold. Percent agreement of responses across passes was used to estimate consistency. The data were compared to those reported previously [Attia, King, Varnet, Ponsot, and Lorenzi. *J. Acoust. Soc. Am.*] for a group of younger normal-hearing subjects aged between 18 and 30 years. There was no significant effect of group on modulation-detection thresholds and percent-agreement data. Taken together with analyses derived from signal-detection theory, these results suggest that below 60 years, internal noise does not increase with age for temporal-modulation detection tasks.

III.1.b. Article

I. INTRODUCTION

Current models of auditory perception include internal variability - usually referred to as “internal noise” (Green and Swets, 1966) - in order to limit detection, discrimination or identification performance that otherwise would be perfect. This is the case for computational models developed to account for amplitude-modulation (AM) and frequency-modulation (FM) perception, two acoustic features playing a key role in speech, music and natural-soundscape perception (e.g., Zeng *et al.*, 2005; Varnet *et al.*, 2017; Thoret *et al.*, 2020, 2021). These features are generally assumed to be encoded as temporal-envelope cues and/or temporal fine-structure cues in the early auditory system *via* fluctuations in the mean discharge rate of auditory neurons and changes over time in the pattern of neural phase locking, respectively (Moore and Sek, 1996). In these models, internal noise is implemented as additive and multiplicative noise sources decreasing the fidelity of temporal-envelope and temporal fine-structure coding (e.g., Dau *et al.*, 1997; Paraouty *et al.*, 2016; Wallaert *et al.*, 2017, 2018; King *et al.*, 2019; Cabrera *et al.*, 2019; Ewert *et al.*, 2020). Within this framework, internal noise refers to the accumulation of several sources of neural variability located along the auditory pathway such as the stochastic nature of neuronal firing, the internal state of the observer organism, or fluctuations in attention (Javel and Viemeister, 2000; Faisal *et al.*, 2008; Amitay *et al.*, 2013).

Computational modelling studies attempting to simulate detection of sinusoidal AM and FM and discrimination of complex AM and FM patterns for elderly subjects with normal hearing and subjects with sensorineural hearing loss suggest that the variance of these internal-noise sources may increase because of ageing (Paraouty *et al.*, 2016; Wallaert *et al.*, 2017, 2018; Ewert *et al.*, 2020) and cochlear damage (Derleth *et al.*, 2001; Ives *et al.*, 2014; Paraouty *et al.*, 2016; Wallaert *et al.*, 2017, 2018; Ewert *et al.*, 2020). Among them, several studies conducted with the same experimental paradigm, the same stimuli but different model structures indicate that the increase of the variance of internal-noise sources caused by ageing *per se* may be relatively modest and that caused by presbycusis (the age-related sensorineural hearing loss) may be substantial (Paraouty *et al.*, 2016; Wallaert *et al.*, 2017, 2018; Ewert *et al.*, 2020). Still, it is possible that the poorer detection and discrimination capacities found in elderly subjects with normal hearing and presbycusis subjects resulted from alterations of (deterministic) temporal-coding mechanisms (for a review, see Gordon-Salant *et al.*, 2010; Pichora-Fuller and Mac Donald, 2008) rather than from a genuine increase in neural variability at peripheral and/or central stages of the auditory system.

The goal of the present study is to test the hypothesis that ageing *per se* causes an increase in internal noise for AM and FM detection. Animal and human studies have revealed that ageing is associated with neural de-afferentation subsequent to “cochlear synaptopathy”, that is the loss of synapses between hair cells and auditory-nerve fibers (Sergeyenko et al., 2013; Viana *et al.*, 2015; Wu *et al.*, 2019). For humans, a temporal-bone study of “normal-ageing” subjects using autopsy material showed a massive loss (> 50 %) of peripheral axons in a number of subjects over 50 years (Wu *et al.*, 2019). In responses to sounds, individual auditory-nerve fibers show a stochastic behavior whose detrimental effects on sensory coding of acoustic information can be overcome by pooling the responses of many neural units at more central stages (Lopez-Poveda, 2014; Wever, 1949; Wever and Bray, 1930). The increase in internal noise suggested by the modelling studies cited above may then be related to the degree of neural de-afferentation caused by ageing which is supposed to affect mainly the processing of high sound amplitudes, and consequently the neural coding of temporal-envelope cues (Furman *et al.*, 2013; Bharadwaj *et al.*, 2014; Parthasarathy *et al.*, 2016; Parthasarathy and Kujawa, 2018; Grose *et al.*, 2019; however, see Carcagno and Plack, 2020). Consistent with this hypothesis, a modelling study by Goodman *et al.* (2017) indicates that the loss of auditory-nerve fibers causes an increase in neural variability and alterations of temporal-envelope coding at the low brainstem (*i.e.*, cochlear nucleus) level. Similar effects may be considered for temporal fine-structure coding since synaptopathy also affects auditory-nerve fibers with high spontaneous rates and low firing thresholds (Paul *et al.*, 2017) that may be involved in neural phase locking to temporal fine-structure cues. Alternatively, increased internal noise may result from changes at more central levels of the auditory system (*e.g.*, Frisina, 2010; Gordon-Salant *et al.*, 2010), part of which may be inherited from peripheral changes such as those caused by cochlear synaptopathy. Two sets of neurophysiological studies are consistent with this hypothesis. First, spontaneous activity of cortical neurons increases with age in non-human primates, presumably because of an age-related imbalance of neural excitation and inhibition in the auditory cortex (Ng and Recanzone, 2018; Recanzone, 2018; Ramamurthy and Recanzone, 2020). Second, cochlear synaptopathy induced pharmacologically in adult mice yields random surges of synchronized activity in cortical neurons caused by cortical hyperexcitability that in turn results from plasticity in cortical excitatory and inhibitory neurons (Resnick and Polley, 2021).

It is important to note that AM and FM detection thresholds worsen only slightly (Takahashi and Bacon, 1992; He *et al.*, 2007, 2008; Grose and Mamo, 2012; Füllgrabe *et al.*, 2015; Wallaert *et al.*, 2016; Whiteford *et al.*, 2017; Carcagno and Plack, 2021), if at all (Schoof and Rosen, 2014; Paraouty *et al.*, 2016; Grose *et al.*, 2019), with increasing age. This may relate to the fact that massive loss (*e.g.*, 50%) of auditory-nerve fibers associated with ageing is not expected to yield changes in thresholds that are sufficiently large to be reliably measured, according to predictions derived from signal-detection theory (Oxenham, 2016). Alternative approaches are therefore required to provide a clearer perceptual basis for the increase in

internal noise. Moore and Vinay (2019) developed an AM discrimination task that may provide empirical evidence for an increase in internal noise for temporal-envelope coding. In this study, the ability to discriminate regular 8-Hz sinusoidal AM from AM that had the same mean rate but was irregular in AM depth and AM rate was measured for normal-hearing subjects aged between 20 and 27 years and older normal-hearing subjects aged between 64 and 80 years. In this envelope-regularity discrimination (ERD) task, the amount of irregularity was varied to determine a "threshold". It was reasoned that regular AM may sound irregular if the auditory system is noisy as a result of cochlear synaptopathy, and this - in turn - would result in higher (poorer) ERD thresholds. In other words, for an auditory system with synaptopathy, the amount of irregularity in the stimulus would have to be higher to overcome the effects of the irregularity in the internal representation of the temporal envelope caused by increased internal noise. Consistent with the assumption that internal noise increases with age, the older subjects showed poorer ERD thresholds than the younger subjects.

Here, a different approach is taken to assess the impact of ageing on internal noise for both temporal-envelope and temporal-fine structure coding. This approach relies on the notion that internal noise does not only limit sensitivity (and thus, performance in detection and discrimination tasks); it also limits *consistency* in auditory judgments. In other words, for a given stimulus and a given task, a high level of internal noise is expected to cause more variable responses from trial to trial and a decrease in the consistency of the subjects' responses. The "double-pass" method (Green, 1964) is a psychophysical paradigm aiming at assessing the consistency of subjects' responses from the agreement of individual responses between two trials using identical stimuli. An estimate of the variance of internal noise for the task under study can then be obtained *via* auditory modelling (Attia *et al.*, Accepted) or through analyses based on signal-detection theory (Green, 1964; Spiegel and Green, 1981). In a companion study (Attia *et al.*, Accepted), the double-pass paradigm was applied to slow (2 Hz) and fast (20 Hz) AM and FM detection for a group of young normal-hearing subjects. The present study tested the hypothesis that double-pass consistency should be poorer - and thus, internal noise should be higher - for a group of older subjects, even when pure-tone audiograms of the older subjects are within the normal range at the tested audio-frequency, because of increased neural variability in the auditory structures involved in AM and FM coding. The participants included in the present study were 40 years of age or older as evidence accumulates that deficiencies in auditory temporal processing emerge relatively early (> 40 years) in the ageing process (Grose *et al.*, 2006; Ross *et al.*, 2007; Helfer and Vargo, 2009; Grose and Mamo, 2010, 2012). They were no older than 65 years of age in order to avoid any cognitive deficits associated with ageing (Wingfield, 1996; Füllgrabe *et al.*, 2015). In the main experiment of the present research, double-pass consistency of auditory judgments was assessed in these subjects for sinusoidal AM and FM detection using a slow (2 Hz) and a fast (20 Hz) modulation rate, a low (500 Hz) carrier frequency and a bandpass (AM or FM) noise masker (weak or strong in magnitude). To estimate double-pass consistency for modulation detection, a 2-interval, 2-alternative forced-choice (2I, 2AFC) detection-in-noise

task was run *twice* with identical sequences of (external) modulation-noise maskers in the two successive passes. Modulation noises used for AM and FM detection were generated in the AM and FM domains, respectively. Two levels of the variance of modulation noise were used for each type of noise. All subjects were tested at their individual modulation-detection threshold as measured in a setup experiment. The double-pass data measured in the main experiment were compared to those collected by Attia *et al.* (Accepted) for younger (18-30 years) normal-hearing subjects using exactly the same stimuli and experimental procedures. Analyses of double-pass data based on signal-detection theory were then conducted to estimate internal noise in each experimental condition and each group of subject (Green, 1964; Spiegel and Green, 1981; Ponsot *et al.*, 2021).

II. METHODS

A. Participants

Fifteen subjects (referred to as NHo subjects) aged between 45 and 62 years (mean=53 years; standard deviation (SD)=5 years) participated in the psychoacoustical experiments. All subjects were recruited *via* the logistic platform RISC (“Relais d’Information sur les Sciences de la Cognition, UMS CNRS 332”) and tested at the Ecole normale supérieure (Paris, France). Subjects were included if they presented no history of hearing disorder, acoustic trauma or tinnitus.

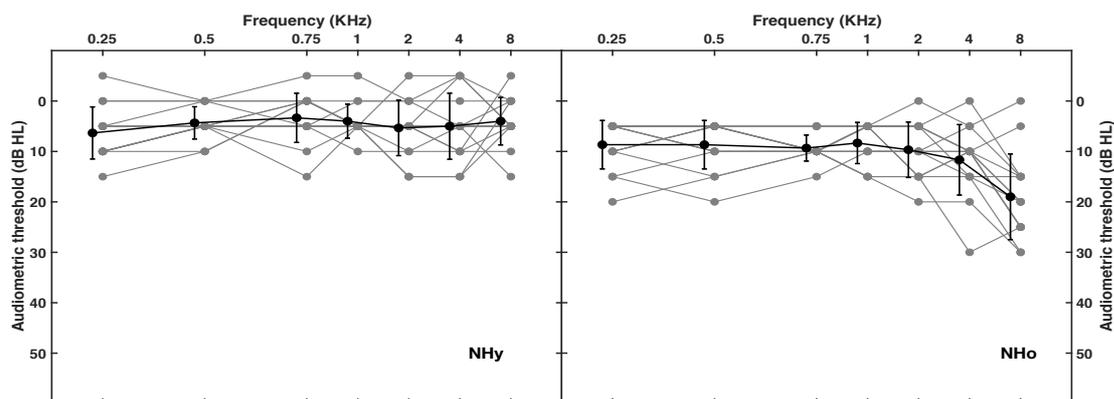


Figure 1. Individual (grey lines) and mean (black lines) audiometric thresholds for the 15 normal-hearing subjects (NH; left panel) and the 15 older normal-hearing subjects (NHo; right panel) for the (tested) right ear. Error bars indicate ± 1 standard deviation of the me

Their pure-tone audiogram was assessed for octave frequencies between 0.25 and 8

kHz using an Otometrics Madsen Itera II audiometer manufactured in Denmark equipped with supra-aural TDH39 headphones. Based on pure-tone audiometry, their hearing was classified as normal below 4 kHz at both ears (Goodman, 1965) and symmetric below 8 kHz: All NHo subjects had audiometric (pure-tone) thresholds lower than or equal to 20 dB hearing level (HL) between 0.25 and 2 kHz in each ear, and the difference between audiometric thresholds measured in the left and right ears was strictly lower than 10 dB HL between 0.25 and 8 kHz. All subjects were tested in their right ear. Three subjects had audiometric thresholds of 25 and 30 dB HL at 4 and 8 kHz in this ear, indicating that they had a mild hearing loss above 2 kHz. Pure-tone average (PTA) estimates calculated for audiometric frequencies between 0.5 and 4 kHz for each individual NHo subject ranged between 3.8 and 16.2 dB HL (mean PTA = 9.6 dB HL; SD=3.6 dB HL) in the right ear.

The psychophysical data of these 15 NHo subjects were compared to those collected on the 15 younger normal-hearing (NH_y) subjects aged between 18 and 30 years (mean=23 years; SD=3 years) who participated in a companion study by Attia *et al.* (Accepted). The NH_y subjects had audiometric thresholds below or equal to 20 dB HL between 0.25 and 8 kHz in both ears. Their PTAs ranged between 1.3 and 6.3 dB HL (mean PTA = 4.5 dB HL; SD=3.6 dB HL) in the right ear.

All subjects were tested in their right ear using 0.5-kHz pure tones modulated in amplitude or frequency. Pure-tone audiometric thresholds (mean and individual data) for the right ear are shown in the left and right panels of Fig. 1 for the 15 NH_y and NHo subjects, respectively. The audiometric thresholds at 0.5 kHz in the right ear ranged between -5 and 20 dB HL (mean= 6.2 dB HL; SD= 5.3 dB HL) for NH_y subjects, and between 5 and 20 dB HL (mean: 8.6 dB HL; SD=4.8 dB HL) for NHo subjects. Audiometric thresholds were slightly but significantly higher (that is, poorer) for NHo subjects than for NH_y subjects (on average, the difference in absolute thresholds at 500 Hz was about 2 dB HL), as shown by a Student *t*-test conducted on audiometric thresholds at 0.5 kHz for NH_y and NHo groups [$t(28)=2.9$; $p<0.05$]. A Student *t*-test conducted on PTAs also showed a significant difference between NH_y and NHo groups [$t(28)=3.8$; $p<0.05$].

NHo subjects carried out the psychoacoustical experiments under exactly the same conditions (*i.e.*, stimuli and experimental procedures described below) as NH_y subjects. Before starting the experiment, subjects were informed about the purpose and conditions of the study and provided written consent. This study was approved by the local ethical committee of Université de Paris (CERES, N° IRB: 20143200001072). All NHo subjects were paid 10 Euros per hour for their participation in this study.

B. Stimuli

All stimuli were generated digitally at a sampling rate of 44.1 kHz using the MATLAB environment (release 2013b). They were presented to the right ear of each subject at a sensation level (SL) of 40 dB using an external soundcard (ECHO Indigo DJ 1-2 Virtual Out- 16-bit resolution) and delivered via Sennheiser (Wedemark, Germany) HD 280 pro monaurally (right ear). All measurements took place in a double-wall soundproof cabin. Levels were calibrated between 0.125 and 4 kHz with a Bruel and Kjaer 2250 sound-level meter. Headphone was calibrated with this same sound-level meter and an artificial ear (4153 6-cc coupler, IEC standard-60318-1).

The whole experimental session was divided into two successive phases referred to as “setup” and “main” experiments. Both experiments consisted of AM and FM detection tasks based on a 2 interval, 2 alternative forced-choice (2IAFC) paradigm. During the setup experiment, AM and FM detection thresholds were measured on each individual subject using an adaptative procedure. During the main experiment, detection performance and double-pass consistency for AM and FM detection were measured on each individual subject tested at his/her own AM or FM detection thresholds using a constant-stimuli procedure. In both experiments, each trial was divided into two successive observation intervals separated by a silent inter-stimulus interval (ISI). The target and the comparison stimuli were presented in a random order in the two successive observation intervals. The duration of each stimulus was 1 sec. The duration of the ISI was 400 ms. All stimuli were modulated in amplitude or in frequency by a modulation noise (a white Gaussian noise lowpass filtered with a cutoff frequency of 80 Hz in the AM or FM domain). More precisely, the modulation noise was generated by filtering a 1-sec Gaussian white noise $n(t)$ with zero mean below 80 Hz using a 4th-order Butterworth filter. The standard deviation of the modulation noise, $\sigma_{AM,FM}^{ext}$, was set to two values: 0.07 or 0.14 stimulus units [modulation depth units (m.d.u.) for the AM noise, or modulation index units (m.i.u.) for the FM noise]. The target signal was a sinusoidal AM or FM at a fixed rate (2 or 20 Hz) that was added to the modulation-noise masker before being applied to the 500-Hz pure-tone carrier. A 100-ms raised-cosine ramp was applied to the beginning and the end of each stimulus. All stimuli were normalized in terms of their root-mean-square (RMS) power.

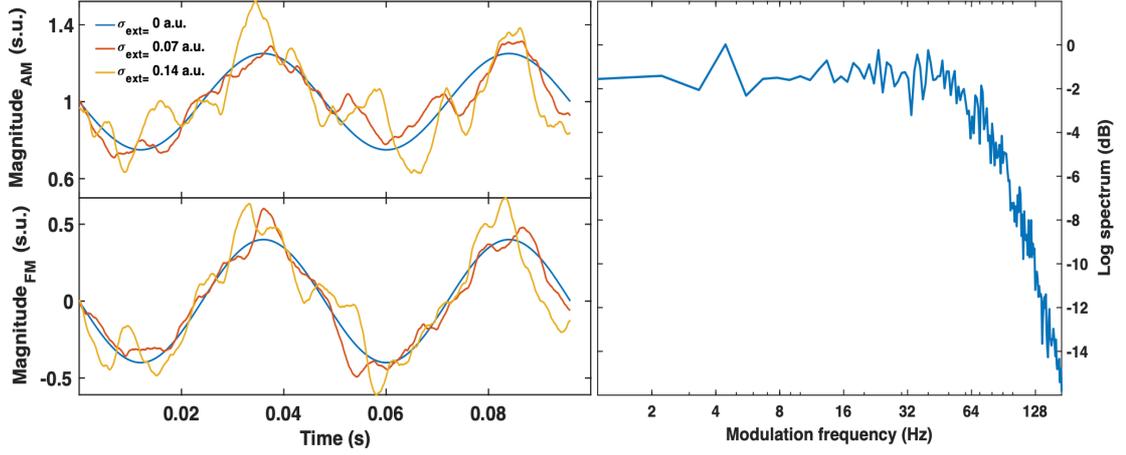


Figure 2. Left panels: Time course of instantaneous amplitude of the 20-Hz sinusoidal AM (top panel) and FM (bottom panel) target modulators. Modulators were obtained for independent samples of modulation noise. The standard deviation of the masker, $\sigma_{ext}^{AM,FM}$, was set to two values: 0, 0.07 or 0.14 units (m.d.u. or m.i.u.). Right panel: Long-term power spectrum of modulation noise with a standard deviation set to 0.07 units. Modulation spectra were averaged over 100 (1-sec long) samples.

Therefore, the comparison stimuli (C_{AM} or C_{FM}) were sinusoidal carriers (with a frequency $f_c=500$ Hz) whose instantaneous amplitude or frequency was modulated using the bandpass noise masker, $n(t)$. The target stimuli (T_{AM} or T_{FM}) were sinusoidal carriers (with a frequency $f_c=500$ Hz) whose instantaneous amplitude or frequency were modulated by the sum of a sinusoidal signal and the bandpass noise masker, $n(t)$. The sinusoidal signal had a rate f_m of 2 or 20 Hz and magnitude m for AM stimuli or $\beta = \Delta f/f_m$ for FM stimuli.

Equation [1] describes the target stimulus $T_{AM}(t)$:

$$T_{AM}(t) = [1 + m \sin(2\pi f_m t + \varphi_{AM}) + \sigma_{ext}^{AM} n(t)] \sin(2\pi f_c t + \varphi_c) \quad [1]$$

Equation [2] describes the comparison stimulus $C_{AM}(t)$:

$$C_{AM}(t) = [1 + \sigma_{ext}^{AM} n(t)] \sin(2\pi f_c t + \varphi_c) \quad [2]$$

Equation [3] describes the target stimulus $T_{FM}(t)$:

$$T_{FM}(t) = \sin [2\pi f_c t + \varphi_c + \beta \sin(2\pi f_m t + \varphi_{FM}) + \sigma_{ext}^{FM} n(t)] \quad [3]$$

Equation [4] describes the comparison stimulus $C_{FM}(t)$:

$$C_{FM}(t) = \sin[2\pi f_c t + \varphi_c + \sigma_{ext} n(t)] \quad [4]$$

where t is the time expressed in seconds (s), m is the AM depth (ranging between 0 and 1) of the AM target signal, $\beta = \Delta f / f_m$ where Δf is the frequency excursion of the FM target signal expressed in Hz, f_m is the modulation rate of the target signal expressed in Hz (2 or 20 Hz), φ_{AM} and φ_c are the starting phases of the AM target signal and the carrier, respectively. φ_{AM} and φ_c were randomly and independently chosen (between 0 and 2π radians) for each stimulus in the setup experiment (measure of AM detection thresholds) and in the first pass of the main experiment (measure of double-pass consistency for AM detection). φ_{FM} and φ_c are the starting phases of the FM target signal and the carrier, respectively. φ_{FM} and φ_c were randomly and independently chosen (between 0 and 2π radians) for each stimulus in the setup experiment (measure of FM detection thresholds) and in the first pass of the main experiment (measure of double-pass consistency for FM detection). For both AM- and FM-detection tasks, the modulation-noise masker, $n(t)$, was refreshed between intervals and trials in the setup experiment, and in the first pass of the main experiment. The modulation-noise maskers used in the second pass of the main experiment were identical to those used in the first pass of the main experiment (in corresponding trials). Modulation-noise maskers were different across subjects.

Figure 2 shows the time course of instantaneous amplitude of 20-Hz AM and FM target modulators and the long-term average modulation spectrum of the modulation noise used in the experiments. The time course of instantaneous amplitude was obtained for 100-ms independent samples of modulation noise. The spectrum of the modulation noise was calculated for 100 independent, 1-sec long samples of modulation noise. The left panels illustrate the corruption caused by the modulation noise on the target sinusoidal modulators in the temporal domain. This detrimental effect increases when the standard deviation of the modulation noise increases from 0.07 to 0.14 units (m.d.u. or m.i.u.). The right panel shows that, as expected, the modulation power of the modulation-noise masker used in the experiments drops above 80 Hz.

II. PROCEDURE

The setup and main experiments were completed in the same order for each subject: In the setup experiment, an adaptive staircase was used to measure individual AM and FM detection thresholds corresponding to three distinct levels of performance: 64 % ($d'=0.5$), 76 % ($d'=1$), and 84 % ($d'=1.5$). In the main experiment, a double-pass consistency task was used to measure the percentage of agreement for the same detection tasks at a fixed level of performance based on individual detection thresholds measured in the setup experiment. The

whole series of experiments was based on one or two hours per session with several sessions spread out over a 1-4 week period for each subject. Altogether, the experiments lasted on average 14 hours per subject. For each task and for each trial, the subject was instructed to choose the interval containing the target sinusoidal AM or FM.

Each subject was tested for the following 24 experimental conditions: two types of modulation (AM vs. FM); two modulation rates (2 vs. 20 Hz); two levels of the standard deviation of the modulation-noise masker [0.07 vs. 0.14 units (m.d.u. or m.i.u.)]; three levels of (targeted) performance: 64 % ($d'=0.5$), 76 % ($d'=1$), and 84 % ($d'=1.5$) correct detection. The individual AM and FM detection thresholds were then used to set the modulation strength (*i.e.*, the modulation depth or frequency excursion, respectively) of the target signals in the subsequent double-pass experiments. Hence, the AM and FM detection tasks in the double-pass consistency paradigm were performed at same level of difficulty.

The double-pass consistency task was conducted with the purpose of estimating percent agreement in the subjects' responses for corresponding trials in two successive passes. Percent-correct detection was also calculated across the two successive passes for each subject.

1. Setup experiment: Detection thresholds

AM detection thresholds (modulation depth, m) and FM detection thresholds (frequency excursion, Δf) were first measured for each subject with adaptive staircases that attempted to converge on 64% ($d'=0.5$), 76% ($d'=1$) and 84% ($d'=1.5$) correct detection using a 2I, 2AFC paradigm and a weighted 1-up, 1-down method (Garcia-Pérez, 1998).

Each block consisted of a variable number of trials. Each trial was divided into two successive observation intervals containing the target and comparison stimuli, presented in random order. The bandpass noise masker was refreshed for each interval and trial. Each observation interval was marked by a red light on the computer screen that was synchronized with stimulus' presentation. Subjects were instructed to report the observation interval containing the target stimulus. Visual feedback as to the correct answer was given to the subject after each trial. The dependent variables for these measures were the modulation depth m for the AM detection task and the frequency excursion Δf for the FM detection task. At the beginning of each staircase, the modulation strength of AM and FM targets was set to a supra-threshold level: The AM target was presented at a modulation depth of 80 % (-2 dB when expressed as $20\log m$) and the FM target was presented at a frequency excursion of 8 Hz. Before the first four reversals, the tracking variable (the modulation strength of AM and FM targets) changed by a step (a factor) up of 4 after a wrong response and by a step (a factor) down of 2 after a correct response. After the first 4 reversals, the 1-down 1-up rule was used up to sixteen reversals, using different steps to reach 64 %, 76% and 86 % of correct responses

(Pérez-García, 1998). Up to twenty reversals were obtained each series of measures and the mean threshold was taken as the geometric mean of the tracking variable during the last sixteen reversals. The order of the conditions was randomized across runs.

AM and FM detection thresholds at 2 and 20 Hz were first measured once in the absence of any modulation noise masker in order to familiarize subjects with the detection tasks. AM and FM detection thresholds were then measured twice in each condition (*i.e.*, 48 staircases). Two measures of detection threshold were thus carried out in each experimental condition and the final detection threshold for each condition was taken as the mean of these two estimates.

2. Main experiment: Double-pass consistency

The consistency of subjects' responses in AM and FM detection was then measured in each experimental condition and for each subject using the double-pass paradigm (Green, 1964; Hasan *et al.*, 2012; Vilidaite and Baker, 2016). Here, AM and FM detection were measured using a 2I, 2AFC paradigm with constant stimuli. In total, each subject was tested over $2 \times 100 = 200$ trials in each experimental condition (AM or FM; modulation rate of 2 or 20 Hz; modulation-noise masker with a standard deviation of 0.07 or 0.14 units (m.d.u. or m.i.u. for AM and FM, respectively); target modulation strength set to the mean individual detection threshold corresponding to a d' of 0.5, 1 or 1.5, as assessed in a setup experiment). In other words, each subject was run twice with identical stimuli, making two successive "passes" of 100 trials in each experimental condition. However, for each trial and for each participant, the presentation order of the target and comparison stimuli was randomized (and was thus different across the two passes) to avoid contextual effects. The 24 experimental conditions were tested in random order. Thus, 4800 trials were run per participant. These 4800 trials were split into 10 (shorter) series of double-pass experiments in which all (24) conditions were completed. Each series was therefore composed of 480 trials: It consisted in a first block of 240 trials in which all (24) conditions were completed 10 times in random order, repeated twice with exactly the same stimuli in a second block of 240 trials (pass 2) after a 5-min break. Short breaks were given to participants between each series of double-pass experiments.

Each block therefore consisted of 240 trials. Five training trials where modulation targets were presented at a modulation strength yielding a d' of 3 were inserted at the beginning of each block to familiarize the participant with the task and stimuli. For each block, each trial was divided into two successive observation intervals containing the target and comparison stimuli, presented in random order. Each observation interval was marked by a red light on the computer screen that was synchronized with stimulus' presentation. Subjects were instructed to report the observation interval containing the target stimulus. Subjects were not given any visual feedback after each trial, except for the first five training trials. Subjects' responses to these training trials were not taken into consideration in the

assessment of double-pass consistency and percent-correct detection.

For each experimental condition, double-pass consistency was calculated as the percentage of agreement between the subjects' responses to corresponding trials within the first and the second pass. Detection performance was also calculated as the percentage of correct responses over the 200 trials [100 trials (pass 1) +100 trials (pass 2)] corresponding to a given experimental condition.

III. RESULTS

A. Modulation-detection thresholds

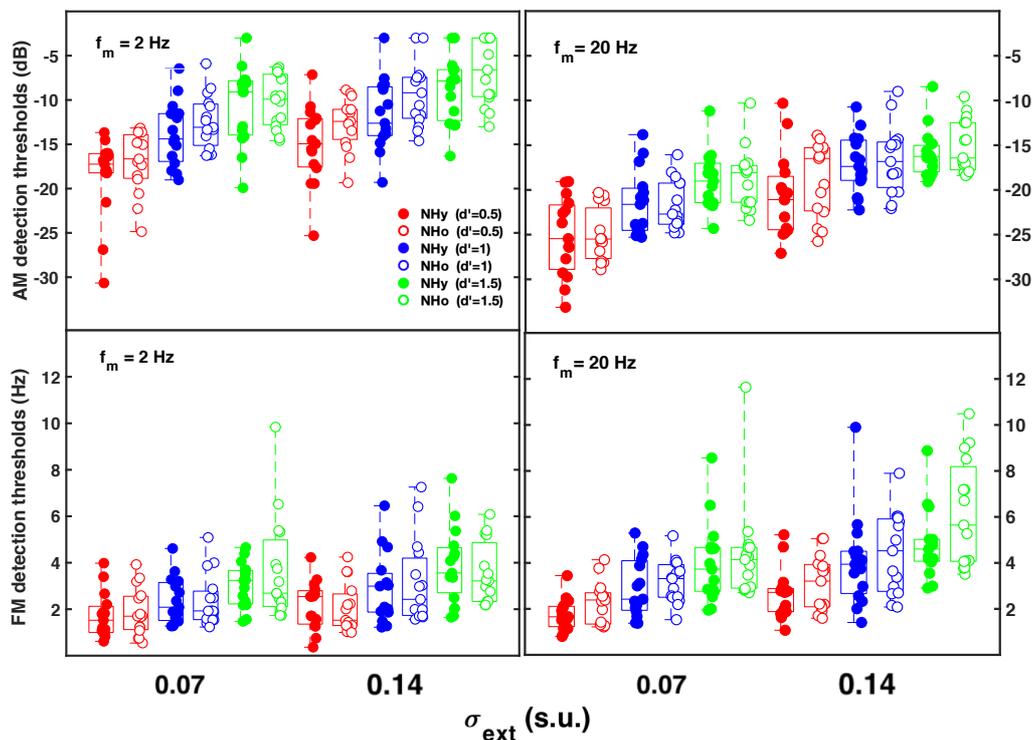


Figure 3. Individual modulation-detection thresholds as a function of the standard deviation of the modulation-noise masker (σ_{ext}) for NHy subjects (closed symbols) and NHo subjects (open symbols). NHy data are replotted from Attia et al. (Accepted). Top and bottom panels show AM and FM detection thresholds, respectively. Left and right panels show the 2-Hz and 20-Hz data, respectively. For each panel, detection thresholds are shown for each level of sensitivity targeted by the adaptive procedure (red: targeted $d'=0.5$; blue: targeted $d'=1$; green: targeted $d'=1.5$). On each box, horizontal lines indicate the median threshold; boxes span the 25th-75th percentiles and vertical lines span the minimum and maximum thresholds.

The individual modulation-detection thresholds of NHo subjects are plotted as a function of the standard deviation of the modulation-noise masker ($\sigma_{AM,FM}^{ext}$) in each panel of Fig. 3. Top and bottom panels show AM and FM detection thresholds, respectively. Left and right panels show the 2-Hz and 20-Hz data, respectively. For each panel, detection thresholds are shown for each level of sensitivity targeted by the adaptive procedure (red: $d'=0.5$; blue: $d'=1$; green: $d'=1.5$). In each panel, the individual NHo data are shown by open circles. NHo data are plotted along with the data collected by Attia *et al.* (Accepted) for NHy subjects using exactly the same stimuli and adaptive procedure. NHy data shown by filled circles.

Separate linear mixed models were run in R (R Core Team version 4.0.2) on AM- and FM-detection threshold data. For each analysis, the independent variables were: Modulation rate (f_m : 2 or 20 Hz), standard deviation of the modulation-noise [$\sigma_{AM,FM}^{ext}$: 0.07 or 0.14 units (m.d.u. or m.i.u.)], targeted level of sensitivity (d' : 0.5, 1 or 1.5), and group (NHy vs. NHo). The dependent variable was: i) AM-detection thresholds (expressed in dB, $20 \log m$), or ii) log-transformed FM-detection thresholds. In the statistical model, fixed effects corresponded to f_m , $\sigma_{AM,FM}^{ext}$, targeted d' and group, with interaction term between these four variables. Random effects corresponding to the identification number of subjects were introduced in the model to take into account individual variability. All categorical variables were expressed using contrast coding. There were no noticeable deviations from homoscedasticity or normality. p -values were obtained by means of Type III analysis of variance (ANOVA) on linear-mixed models. The significance level was set at 0.05.

1. AM-detection thresholds

The results of statistical analyses for AM-detection thresholds are detailed in Table I. The analysis showed: i) A significant main effect of $\sigma_{AM,FM}^{ext}$ on AM-detection thresholds ($\beta=5.5$, $SE=0.93$; $\chi^2(1)=41.1$; $p<0.0001$) indicating that AM-detection threshold increased (worsened) with σ_{AM}^{ext} (mean AM-detection thresholds were 7 dB higher at $\sigma_{AM}^{ext} = 0.14$ m.d.u. than at $\sigma_{AM}^{ext} = 0.07$ m.d.u); ii) a significant main effect of f_m on AM-detection thresholds ($\beta=7.6$, $SE=0.86$; $\chi^2(1)=78.3$; $p<0.0001$) indicating that AM-detection thresholds decreased (improved) with f_m (AM-detection thresholds were on average 9 dB lower at 20 Hz than at 2 Hz); iii) as expected, a significant main effect of targeted sensitivity level on AM-detection thresholds ($\beta=2.6$, $SE=0.86$; $\chi^2(1)=49.8$; $p<0.0001$). There was no significant main effect of group on AM-detection thresholds ($p>0.05$) and no significant interaction between independent variables (all $p>0.05$).

AM-detection thresholds (NHy vs. NHo)	χ^2	df	p
Intercept	718.72	1	2.2e-16
σ_{AM}^{ext}	41.17	1	1.39e-10
f_m	78.33	1	2.20e-16
d'	49.86	2	1.44e-11
group	0.070	1	0.79
$\sigma_{AM}^{ext} * f_m$	2.67	1	0.10
$\sigma_{AM}^{ext} * d'$	5.024	2	0.081
$f_m * d'$	2.46	2	0.29
$\sigma_{AM}^{ext} * \text{group}$	0.62	1	0.43
$f_m * \text{group}$	0.68	1	0.41
$d' * \text{group}$	0.99	2	0.61
$\sigma_{AM}^{ext} * f_m * d'$	4.44	2	0.11
$\sigma_{AM}^{ext} * f_m * \text{group}$	0.25	1	0.61
$\sigma_{AM}^{ext} * d' * \text{group}$	0.38	2	0.82
$f_m * d' * \text{group}$	1.14	2	0.56

Table 1. Analysis of deviance (Type III Wald chi-square tests) showing p-values from lmer models (lme4 package) including AM-detection thresholds as dependent variable, and modulation rate f_m , targeted sensitivity level (d'), σ_{AM}^{ext} and group (NHy vs. NHo) as independent variables. df stands for degree of freedom.

Overall, these results indicate that the AM noise applied to each stimulus produces substantial modulation masking, whose magnitude increases with the standard deviation of the AM noise. Modulation masking is well explained by changes in signal-to-noise ratio at the output of (central) modulation filters tuned to the target AM rates of 2 and 20 Hz (Dau *et al.*, 1997; King *et al.*, 2019; Wallaert *et al.*, 2018). The improvement in AM sensitivity at higher modulation rate is well explained by temporal integration of AM information (*e.g.*, Dau *et al.*, 1997; Wallaert *et al.*, 2016): Stimulation duration being fixed (duration=1 sec), the number of AM cycles is higher at 20 Hz than at 2 Hz. Temporal integration is currently assumed to reflect decision processes involving “multiple looks” (Viemeister & Wakefield, 1991) or a template-matching process (Dau *et al.*, 1997; Wallaert *et al.*, 2018). The absence of age effects on AM sensitivity is discussed in the final section of this paper.

2. FM-detection thresholds

The results of statistical analyses for FM-detection thresholds are detailed in Table II. The analysis showed: i) A significant main effect of σ_{FM}^{ext} on FM-detection thresholds ($\beta=6.1$, $SE=1.9$; $\chi^2(1) = 9.4$; $p < 0.01$) indicating that FM-detection thresholds increased (worsened) with σ_{FM}^{ext} (FM detection thresholds were on average 1.5 times higher at $\sigma_{FM}^{ext}=0.14$ m.i.u. than at $\sigma_{FM}^{ext}=0.07$ m.i.u); ii) no significant main effect of f_m on FM-detection thresholds ($p = 0.074$); iii) a significant main effect of targeted sensitivity level on FM-detection thresholds ($\beta=7.9$, $SE=1.9$; $\chi^2(1) = 51.2$; $p < 0.0001$). There was no significant main effect of group on FM-detection thresholds ($p > 0.05$) and no significant interaction between independent variables (all $p > 0.05$).

FM-detection thresholds (NH _y vs. NH _o)	χ^2	df	p
Intercept	42.64	1	6.57e-11
σ_{FM}^{ext}	9.41	1	0.0022
f_m	3.20	1	0.074
d'	51.11	2	7.7e-12
group	2.36	1	0.12
$\sigma_{FM}^{ext} * f_m$	1.23	1	0.26
$\sigma_{FM}^{ext} * d'$	0.30	2	0.86
$f_m * d'$	0.39	2	0.82
$\sigma_{FM}^{ext} * \text{group}$	0.46	1	0.49
$f_m * \text{group}$	0.55	1	0.46
$d' * \text{group}$	0.66	2	0.71
$\sigma_{FM}^{ext} * f_m * d'$	0.92	2	0.63
$\sigma_{FM}^{ext} * f_m * \text{group}$	0.041	1	0.84
$\sigma_{FM}^{ext} * d' * \text{group}$	0.91	2	0.64
$f_m * d' * \text{group}$	0.87	2	0.65

Table II. Analysis of deviance (Type III Wald chi-square tests) showing p-values from lmer models (lme4 package) including FM-detection thresholds as dependent variable, and modulation rate f_m , targeted sensitivity level (d'), σ_{FM}^{ext} and group (NH_y vs. NH_o) as independent variables. df stands for degree of freedom.

The present results indicate that the FM noise applied to each stimulus produces substantial modulation masking, whose magnitude increases with the standard deviation of the FM noise. This form of modulation masking may reflect AM masking resulting from the conversion of FM cues evoked by the FM target and FM noise into AM cues at the output of cochlear filters (Moore et al., 1991; King *et al.*, 2019). As for AM, FM masking may then reflect changes in signal-to-noise ratio at the output of (central) modulation filters tuned to the target FM rates of 2 and 20 Hz (Dau *et al.*, 1997; King *et al.*, 2019; Wallaert *et al.*, 2018). Still, the contribution of FM-to-AM conversion to FM detection remains debated and the observed FM masking effect may reflect other processes involving the use of temporal-fine structure conveyed by FM (Paraouty *et al.*, 2016). As for AM, the improvement in FM sensitivity at higher modulation rate is well explained by temporal integration of FM information (*e.g.*, Wallaert *et al.*, 2016, 2018; King *et al.*, 2019) and reflect decision processes involving “multiple looks” or a template-matching process. The absence of age effects on FM sensitivity is discussed in the final section of this paper.

B. Double-pass consistency

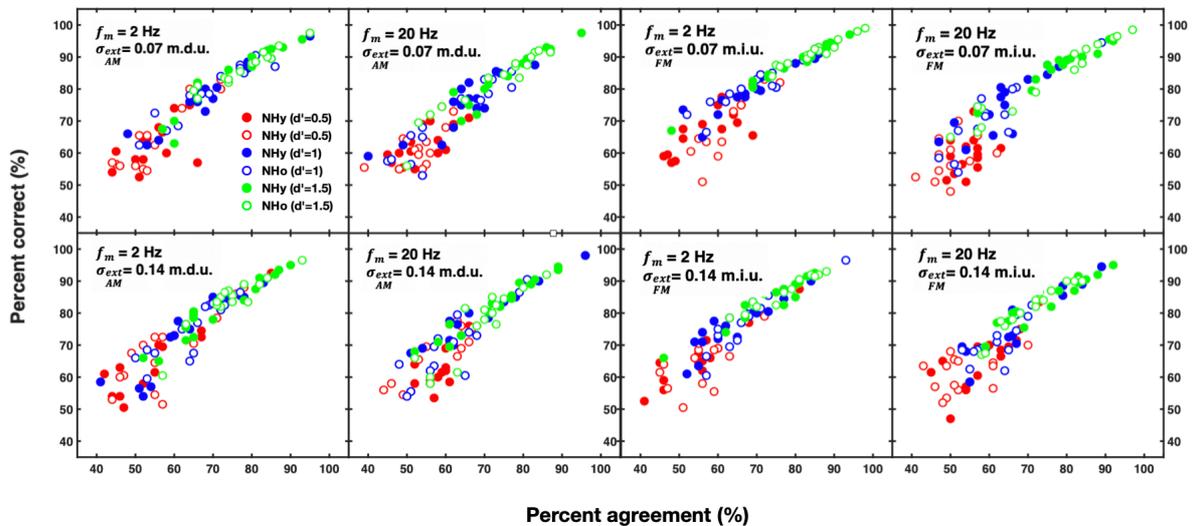


Figure 4. Performance versus consistency curves showing percent correct and percent agreement data for each subject (closed symbols: NHy subjects; open symbols: NHo subjects). The left and right panels show the 2- and 20-Hz data, respectively. The top and bottom panels show the data collected using $\sigma_{ext}^{AM,FM}$ of 0.07 and 0.14 units (m.d.u. or m.i.u.), respectively. For each panel, the individual data are shown for each level of sensitivity targeted by the initial adaptive procedure (red: targeted $d'=0.5$; blue: targeted $d'=1$; green: targeted $d'=1.5$).

Three NHy subjects (*cf.* Attia *et al.*, Accepted) and seven NHo participants were unable to complete threshold runs in some conditions of the AM-detection task. For these NHy and NHo subjects and for these specific conditions, AM depth was set to -3 dB in the double-pass

experiments (this corresponds to 9 data points in Fig. 4). These subjects could perform the detection task in the double-pass experiment well above chance level (performance >75%). All subjects could perform the adaptive task in the FM-detection experiments.

Figure 4 shows the double-pass consistency (*i.e.*, percent agreement) and performance data for each individual NHo (open circles) subject. The data collected by Attia *et al.* (Accepted) for individual NHy subjects using exactly the same stimuli and double-pass procedure are shown by filled circles. The left and right panels show the 2- and 20-Hz data, respectively. The top and bottom panels show the data collected using $\sigma_{AM,FM}^{ext}$ of 0.07 and 0.14 units (m.d.u. or m.i.u.), respectively. For each panel, the individual NHy and NHo data are shown for each level of sensitivity targeted by the initial adaptive procedure (red: targeted $d'=0.5$; blue: targeted $d'=1$; green: targeted $d'=1.5$).

For each experimental condition and each group, a positive correlation was found between performance and percent-agreement data: For each experimental condition and for both NHy and NHo subjects, higher detection performance was associated with higher consistency in subjects' responses. For NHy subjects, Bravais-Pearson r ranges between 0.6 and 0.99 except for 20-Hz FM detection where $r=0.35$ for targeted $d'=0.5$. For NHo subjects, Bravais-Pearson r ranges between 0.45 and 0.97 except for 20-Hz FM detection where $r=0.3$ for targeted $d'=0.5$. Such a correlation was expected because performance and percent agreement are mathematically related (Lu and Doshier, 2008). This is obvious at extreme points: A perfect or near-perfect performance is necessarily associated with perfect or near-perfect agreement. In addition, high performance cannot be associated with low consistency values: If percent agreement is below 100%, there must be one of the two passes where the observer makes incorrect responses and therefore performance is lower than 100%.

CONDITIONS		AM		FM	
		$f_m=2\text{ Hz}$	$f_m=20\text{ Hz}$	$f_m=2\text{ Hz}$	$f_m=20\text{ Hz}$
$\sigma_{AM,FM}^{ext} = 0.07\text{ u.}$	$d'=0.5$	56 (65)	54 (62)	57 (68)	55 (60)
	$d'=1$	70 (79)	64 (75)	69 (80)	62 (74)
	$d'=1.5$	75 (83)	77 (85)	80 (88)	79 (88)
$\sigma_{AM,FM}^{ext} = 0.14\text{ u.}$	$d'=0.5$	58 (68)	60 (65)	57 (68)	60 (68)
	$d'=1$	64 (74)	68 (79)	65 (76)	67 (76)
	$d'=1.5$	71 (81)	73 (82)	75 (85)	75 (84)

Table III. Mean percent-agreement (PA, in %) in each condition. Mean Percent correct (PC, in %) between parentheses. Data are averaged across the 15 NHy listeners.

CONDITIONS		AM		FM	
		$f_m=2\text{ Hz}$	$f_m=20\text{ Hz}$	$f_m=2\text{ Hz}$	$f_m=20\text{ Hz}$
$\sigma_{AM,FM}^{ext} = 0.07\text{ u.}$	$d'=0.5$	56 (66)	53 (62)	64 (71)	50 (58)
	$d'=1$	69 (80)	61 (71)	70 (81)	63 (73)
	$d'=1.5$	78 (87)	72 (82)	85 (91)	75 (83)
$\sigma_{AM,FM}^{ext} = 0.14\text{ u.}$	$d'=0.5$	59 (67)	58 (68)	57 (66)	53 (61)
	$d'=1$	65 (76)	62 (71)	67 (78)	62 (71)
	$d'=1.5$	75 (84)	69 (77)	76 (87)	72 (82)

Table IV. Mean percent-agreement (PA, in %) in each condition. Mean Percent correct (PC, in %) between parentheses. Data are averaged across the 15 NHo listeners.

Mean percent agreement and performance calculated for each modulation type (AM, FM) and rate (2, 20 Hz) are shown for NHy and NHo subjects in Tables III and IV, respectively. Two separate sets of statistical analyses were run on arcsine transformed percent-agreement scores and arcsine transformed percent-correct scores, respectively. For each analysis, a linear mixed effect model was run with transformed percent-agreement and percent-data as the dependent variables. For both models, $\sigma_{AM,FM}^{ext}$ (0.07 or 0.14 units), f_m (2 or 20 Hz), targeted sensitivity level (0.5, 1 or 1.5), modulation type (AM or FM) and group (NHy or NHo) were included as fixed effects altogether with all interactions between factors. All categorical variables were expressed using contrast coding. Random effects corresponding to the identification number of subjects were introduced in each model to take into account individual variability. Assumptions of homogeneity of residual variance, linearity and normality of random effects were checked before running statistical analyses. p -values were obtained by means of type III ANOVAs. Significance level testing was set to a p -value of 0.05.

1. Percent-agreement scores

The analysis conducted on percent-agreements data showed: i) No significant main effects of factors group, modulation type, f_m , and $\sigma_{AM,FM}^{ext}$; ii) a significant main effect of targeted sensitivity level ($\beta=8.14$, $SE=3.15$; $\chi^2(1)=36.8$; $p<0.0001$); iii) a significant interaction between modulation type and f_m ($\beta=8.98$, $SE=4.5$; $\chi^2(2)=40.6$; $p<0.0001$), indicating that percent-agreement was significantly higher at 2 Hz compared to 20 Hz, and that this effect was greater for FM (by 12.5 percentage points) than for AM (by 4 percentage points). The remaining interactions between fixed factors were not significant (all $p>0.05$).

2. Percent-correct scores

Statistical analyses conducted on percent-correct data showed: i) No significant main effects of factors group, modulation type, f_m , and $\sigma_{AM,FM}^{ext}$; (ii) a significant main effect of targeted d' level ($\beta=9.46$, $SE=21.86$; $\chi^2(2)=51.5$; $p<0.0001$); iii) a significant interaction between f_m and modulation type ($\beta=8.74$; $SE=4.32$; $\chi^2(1)=4.09$; $p<0.05$), indicating that performance was significantly higher at 2 Hz compared to 20 Hz, and that this effect was greater for FM (by 14 percentage points) than for AM (by 4 percentage points); iv) a significant interaction between the targeted d' level and $\sigma_{AM,FM}^{ext}$ ($\beta=-5.94$ for $d'=1$ and $\beta=-11.77$ for $d'=1.5$; $SE=4.32$; $\chi^2(2)=7.43$; $p<0.05$): Performance was higher (by 2 percentage points) when $\sigma_{AM,FM}^{ext}$ was the lowest except at $d'=0.5$ where performance was higher when $\sigma_{AM,FM}^{ext}$ was the highest (difference of 5 percentage points).

3. Assessment of internal noise

In order to confirm the robustness of our main analyses based on performance and agreement scores, we also converted these measurements into values of internal noise inferred from a signal-detection theory (SDT) model (Green, 1964; Burgess and Colborne, 1988). We followed the established procedure where the internal-to-external noise ratio is inferred from a typical SDT model fitted to the percentage of correct responses and percentage of agreement obtained across repeated presentation of the same trials (for a detailed description of this inference procedure, the reader is invited to refer to previous studies, *e.g.*, Joosten and Neri, 2012). Here, the model assumes no response bias, a choice justified by the fact that the influence of response bias on our measurements, if any, was controlled by the randomization of the presentation intervals across the two passes (see Procedure).

We first averaged the percent agreement and percent correct values over the three levels of targeted d' (0.5, 1 and 1.5) so that, for each individual and condition [modulation type (AM/FM) x modulation rate (2/20 Hz) x level of external noise (0.07/0.14 units)], estimates of internal noise were based on the repetition of 300 trials, which corresponds to the minimal resolution required for comparison across groups/conditions (Hasan *et al.*, 2012).

We followed this procedure to infer internal noise values for all these conditions, but only values that fell within the interpretable range [0.2; 5] returned by the SDT model (Neri, 2010) were kept for subsequent analyses. Overall, this approach allowed us to infer plausible internal noise values from ~67% of our dataset. Because the purpose of the present analysis was to corroborate the observations based on percent agreement and percent correct scores, we then aggregated the internal noise values across the different conditions (levels of external noise and modulation rates).

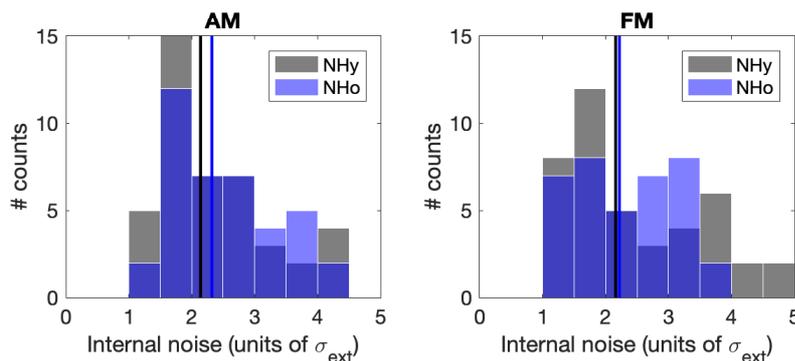


Figure 5. Distribution of internal noise values for AM and FM for the two groups of subjects tested (black color: NHy subjects; blue color: NHo subjects) inferred from percent correct and percent agreement data collected, aggregated across experimental conditions (see the text for details). Vertical lines show the corresponding median value of each distribution.

Figure 5 shows the distributions of internal noise estimates for AM and FM aggregated across conditions for the two groups tested. Internal noise values were very similar for AM and FM for both groups, with median values falling between 2.1 and 2.3 units of external noise, consistent with the range observed for other low-level auditory psychophysical tasks (Neri, 2010; Ponsot *et al.*, 2010). Wilcoxon rank-sum tests showed no significant differences between groups for both AM and FM conditions ($ps > .1$).

IV. DISCUSSION

1. No effect of age on AM and FM detection

We did not find any significant effect of group on AM and FM detection thresholds measured in 12 experimental conditions (2 rates x 2 levels of modulation noise x 3 targeted d' levels), respectively. This is at odds with previous results collected by our group using a similar psychophysical procedure and identical modulation rates (*e.g.*, Wallaert *et al.*, 2016). This discrepancy may be due to the fact that our older group (aged between 40 and 62 years) was somewhat younger than that used by Wallaert *et al.* (47-66 years). Still, it is important to note that the effects of age reported by He *et al.* (2008), Grose and Mamo (2012), Füllgrabe *et al.* (2015) and Wallaert *et al.* (2016) for subjects older than those used here were relatively small for both AM and FM detection (mean thresholds higher by about 2 dB for the older than for the younger subjects for AM; a factor of 1.7-1.8 higher for the older than for the younger subjects for FM). In a large cohort study ($n=102$ subjects), Carcagno and Plack (2021) estimated that AM-detection thresholds increase with age at a rate of only 0.1-1.5 dB per age decade at high presentation levels. Moreover, several studies (*e.g.*, Schoof and Rosen, 2014; Grose *et al.*, 2019) conducted with older subjects than those used here (Schoof and Rosen: 60-72 years; Grose *et al.*: 61-78 years) including one conducted by our group (Paraouty *et al.*, 2016: 46-64 years) did not find any effect of age on AM and FM detection thresholds at similar or higher rates. So, the current data are compatible with the notion that both AM and FM detection capacities worsen only slightly - if at all - with increasing age. As pointed out in the Introduction section, this negative outcome is not incompatible with an age-related cochlear synaptopathy, as predictions based on a model derived from signal-detection theory showed that massive loss of auditory-nerve fibers is not expected to cause changes in thresholds that are sufficiently large to be reliably measured (Oxenham, 2016). It might also be the case that the detrimental effects of increased internal noise on AM and FM detection were compensated by sub-clinical sensorineural hearing loss in our NHO subjects. Indeed, loss of outer hair cells and subsequent perceptual magnification of temporal-envelope fluctuations evoked by AM and FM at the output of cochlear filters (Moore *et al.*, 1996) might have counteracted the degradation in AM and FM coding caused by cochlear synaptopathy, leading to apparently normal AM and FM detection thresholds despite increased neural variability at more central stages. This explanation would be compatible with the slightly but significantly

poorer-than-normal audiometric thresholds shown by our NHo subjects at 500 Hz (8.6 dB HL on average for NHo subjects, compared to 6.2 dB for NHy subjects).

It is noteworthy that the current AM and FM detection thresholds were measured in the presence of a modulation-noise masker. Target modulation rates were 2 and 20 Hz, and the modulation noise had a modulation spectrum covering the 0-80 Hz range. AM and FM detection thresholds increased (worsened) significantly with the standard deviation of the modulation-noise masker. The absence of significant group effect on modulation-detection thresholds measured in the presence of this modulation-noise masker indicates that modulation masking in the AM and FM domains is not altered by age. Had the selectivity of tuned modulation filters (such as those assumed by Dau *et al.* (1997) for AM detection) been reduced for older subjects, the amount of modulation masking produced by the modulation-noise masker should have increased, especially at the highest level of the variance of the modulation noise. These data are consistent with modulation-masking data reported by Takahashi and Bacon (1992) and Sek *et al.* (2015), showing comparable modulation-masking patterns for young subjects with normal hearing or moderate hearing loss and older subjects with mild or moderate hearing loss when a target sinusoidal AM was presented against a masking sinusoidal AM. These findings are however inconsistent with brain-imaging (f-MRI) data suggesting that temporal selectivity of neural units declines in the auditory cortex of elderly human subjects (Erb *et al.*, 2020). The discrepancy between the present psychophysical data and these f-MRI data remains unexplained.

2. No effect of age on double-pass consistency

The goal of the present study was to test the hypothesis that ageing *per se* causes an increase in internal noise for AM and FM detection using a double-pass paradigm. The idea that ageing *per se* may cause an increase in neural variability is supported by vision studies demonstrating that age-related losses in contrast and motion sensitivity are well accounted for by higher internal noise (Allard *et al.*, 2013; Arena *et al.*, 2013; Yan *et al.*, 2020). Our investigation was especially motivated by the outcome of modelling studies of temporal auditory processes: (1) studies based on perceptual models showing that age-related deficits in AM and FM detection are successfully simulated by assuming that internal noise at the output of temporal-modulation filters is increased for older subjects compared to younger ones (Paraouty *et al.*, 2016; Wallaert *et al.*, 2017, 2018; Ewert *et al.*, 2020); (2) a study based on a phenomenological model of brainstem neural responses, demonstrating that the loss of auditory-nerve fibers subsequent to cochlear synaptopathy causes an increase in neural variability and alterations of temporal-modulation coding in cochlear-nucleus neurons (Goodman *et al.*, 2017). This investigation was also motivated by the outcome of neurophysiological studies showing that spontaneous activity of cortical auditory neurons increases with age in non-human primates (Ng and Recanzone, 2018; Recanzone, 2018), and that cochlear synaptopathy causes random surges of hyper-synchronized activity in cortical

neurons (Resnick and Polley, 2021).

We did not find any significant effect of group on double-pass consistency data for AM and FM detection at both slow (2 Hz) and fast (20 Hz) rates, and at low and high levels of the modulation-noise masker. An analysis based on inferred internal noise levels (expressed in units of external noise) using an SDT model yielded the same result. This outcome is not consistent with the notion that ageing leads to an increase in neural variability in the peripheral and central auditory structures involved in temporal-envelope and temporal fine-structure processing. Again, it is important to note that the older subjects (aged between 40 and 62 years) who participated in the present study were younger than those (aged between 64 and 80 years) who participated in the study conducted by Moore and Vinay (2019). Thus, we cannot rule out that internal noise may actually increase beyond about 60-65 years of age. A recent psychophysical study by Ponsot *et al.* (2020) explored auditory detection of spectro-temporal modulations for 10 young normal-hearing subjects (21-37 years) and 7 older subjects with moderate sensorineural hearing loss (59-67 years). The target stimulus had a temporal-modulation rate of 7.1 Hz and a spectral modulation rate of 1 cycle/octave. Several behavioral tasks were used, including a double-pass paradigm similar to the one used here. The authors did not find any significant difference in double-pass consistency across groups, providing additional evidence that internal noise for temporal-modulation coding may not be affected by ageing (at least below 67 years). Alternatively, it may be the case that changes in the variance of internal noise caused by ageing are too small to be reliably measured by the double-pass paradigm when subjects are below 67 years. Once performance and agreement scores derived from these studies are converted into internal noise levels, it became clear that our measurements exposed a substantial amount of inter-individual variability within each group, as can be seen in the histograms plotted in Fig. 5. Further work is therefore warranted to explore ageing effects using alternative paradigms that may show better sensitivity to internal noise (Allart *et al.*, 2013; Vilidaite and Baker, 2016; Yan *et al.*, 2020). Still, the present findings suggest that at least part of the deficits in AM and FM detection reported by Wallaert *et al.* (2016) for normal-hearing subjects aged between 47 and 66 years might not result from a genuine increase in neural variability, as proposed by Wallaert *et al.* (2017, 2018). It is possible that these deficits were partly caused by a decline in the general auditory ability to perform any task, perhaps related to impoverished cognitive abilities such as processing speed, attention or memory (Wingfield, 1996; Füllgrabe *et al.*, 2015). Alternatively, but not exclusively, they might be caused by systematic (that is, non-stochastic) perceptual distortions of temporal-envelope and temporal fine-structure cues resulting from alterations of the central temporal-coding mechanisms (*e.g.*, Overton and Recanzone, 2016; Recanzone, 2018). The latter may be modelled by altering certain aspects of modulation processing or decision making (*e.g.*, use of sub-optimal templates) in computational models of temporal-modulation perception (*e.g.*, Wallaert *et al.*, 2017, 2018; Ewert *et al.*, 2020).

V. CONCLUSION

Detection thresholds and double-pass consistency were estimated for AM and FM detection in a group of normal-hearing subjects aged between 40 and 62 years. AM and FM stimuli were masked by AM and FM noise, respectively. The data collected over 24 experimental conditions (2 modulation types x 2 modulation rates x 2 levels of modulation noise x 3 targeted d' levels) were compared to those reported previously by Attia *et al.* (Accepted) for a group of younger (18-20 years) normal-hearing subjects. AM and FM detection thresholds and double-pass consistency data for AM and FM detection did not differ significantly across groups. In addition, the variance of internal noise estimated from analyses based on signal-detection theory did not differ significantly across groups.

Altogether, these results do not provide evidence that below about 60 years, neural variability in peripheral and central structures of the auditory system increases with age. They also corroborate the notion that temporal-modulation processing worsens only slightly with increasing age, if at all, below 60 years.

ACKNOWLEDGMENTS

This work was supported by two grants from the Collège National des Audioprothésistes (France) and Entendre SAS. This work was also supported by ANR-17-EURE-0017.

VI. REFERENCES

- Allard, R., Renaud, J., Molinatti, S., and Faubert, J. (2013). "Contrast sensitivity, healthy aging and noise," *Vis. Res.*, **92**, 47-52.
- Amitay, S., Guiraud, J., Sohoglu, E., Zobay, O., Edmonds, B. A., Zhang, Y. X., and Moore, D R. (2013). "Human decision making based on variations in internal noise: An EEG study," *PloS one*, **8**, e68928. doi: 10.1371/journal.pone.0068928.
- Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., and Shinn-Cunningham, B. G. (2014). "Cochlear neuropathy and the coding of supra-threshold sound," *Front. Syst. Neurosci.*, **8**, 26.
- Burgess, A. E., and Colborne, B. (1988). "Visual signal detection. IV. Observer inconsistency," *J. Opt. Soc. Am. A*, **5**, 617-627.
- Cabrera, L., Varnet, L., Buss, E., Rosen, S. and Lorenzi, C. (2019). "Development of temporal auditory processing in childhood: Changes in efficiency rather than temporal-modulation selectivity," *J. Acoust. Soc. Am.*, **146**, 2415-2429.
- Carcagno, S., and Plack, C. J. (2020). "Effects of age on electrophysiological measures of cochlear synaptopathy in humans," *Hear. Res.*, **396**, 108068.
- Carcagno, S., and Plack, C. J. (2020). "Effects of age on psychophysical measures of auditory temporal processing and speech reception at low and high levels," *Hear. Res.*, **400**, 108117.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," *J. Acoust. Soc. Am.*, **102**, 2906-2919.
- Derleth, R. P., Dau, T., and Kollmeier, B. (2001). "Modeling temporal and compressive properties of the normal and impaired auditory system," *Hear. Res.*, **159**, 32-149.
- Erb, J., Schmitt, L. M., and Obleser, J. (2020). "Temporal selectivity declines in the aging human auditory cortex," *Elife*, **9**, e55300.
- Ewert, S. D., Paraouty, N., and Lorenzi, C. (2020). "A two-path model of auditory modulation detection using temporal fine structure and envelope cues," *Eur. J. Neurosci.*, **51**, 1265-1278.
- Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008). "Noise in the nervous system," *Nature Rev. Neurosci.*, **9**, 292-303.
- Frisina, R. D. (2009). "Age-related hearing loss: ear and brain mechanisms". *Ann. New York Acad. Sci.*, **1170**, 708-717.
- Füllgrabe, C., Moore, B. C., and Stone, M. A. (2015). "Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition," *Front. Aging Neurosci.*, **6**, 347.
- Furman, A. C., Kujawa, S. G., and Liberman, M. C. (2013). "Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates," *J. Neurophysiol.*, **110**, 577-586.
- García-Pérez, M. A. (1998). "Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties," *Vis. Res.*, **38**, 1861-1881.
- Goodman, D. F., Winter, I. M., Léger, A. C., de Cheveigné, A., and Lorenzi, C. (2018). "Modelling firing regularity in the ventral cochlear nucleus: Mechanisms, and effects of stimulus level and synaptopathy," *Hear. Res.*, **358**, 98-110.
- Gordon-Salant, S., Frisina, R. D., Fay, R. R., and Popper, A. (Eds.). (2010). "The aging auditory system" (Vol. 34). Springer Science & Business Media.

- Green, D. M. (1964). "Consistency of auditory detection judgments," *Psychol. Rev.*, **71**, 392-407.
- Green, D. M., and Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). New York: Wiley.
- Grose, J. H., Hall III, J. W., and Buss, E. (2006). "Temporal processing deficits in the presenescent auditory system," *J. Acoust. Soc. Am.*, **119**, 2305-2315.
- Grose, J. H., and Mamo, S. K. (2010). "Processing of temporal fine structure as a function of age," *Ear Hear.*, **31**, 755.
- Grose, J. H., and Mamo, S. K. (2012). "Frequency modulation detection as a measure of temporal processing: Age-related monaural and binaural effects," *Hear. Res.*, **294**, 49-54.
- Grose, J. H., Buss, E., and Elmore, H. (2019). "Age-related changes in the auditory brainstem response and suprathreshold processing of temporal and spectral modulation," *Trends Hear.*, **23**, doi: 10.1177/2331216519839615.
- Hasan, B. A. S., Joosten, E., and Neri, P. (2012). "Estimation of internal noise using double-passes: Does it matter how the second pass is delivered?," *Vis. Res.*, **69**, 1-9.
- He, N. J., Mills, J. H., and Dubno, J. R. (2007). "Frequency modulation detection: effects of age, psychophysical method, and modulation waveform," *J. Acoust. Soc. Am.*, **122**, 467-477.
- He, N. J., Mills, J. H., Ahlstrom, J. B., and Dubno, J. R. (2008). "Age-related differences in the temporal modulation transfer function with pure-tone carriers," *J. Acoust. Soc. Am.*, **124**, 3841-3849.
- Helfer, K. S., and Vargo, M. (2009). "Speech recognition and temporal processing in middle-aged women," *J. Amer. Acad. Audiol.*, **20**, 264-271.
- Ives, D. T., Kalluri, S., Strelcyk, O., Sheft, S., Miermont, F., Coez, A., ... and Lorenzi, C. (2014). "Effects of noise reduction on AM perception for hearing-impaired listeners," *J. Assoc. Res. Otolaryngol.*, **15**, 839-848.
- Javel, E., and Viemeister, N. F. (2000). "Stochastic properties of cat auditory nerve responses to electric and acoustic stimuli and application to intensity discrimination," *J. Acoust. Soc. Am.*, **107**, 908-921.
- Joosten, E. R., and Neri, P. (2012). "Human pitch detectors are tuned on a fine scale, but are perceptually accessed on a coarse scale," *Biol. Cyber.*, **106**, 465-482.
- King, A., Varnet, L., and Lorenzi, C. (2019). "Accounting for masking of frequency modulation by amplitude modulation with the modulation filter-bank concept," *J. Acoust. Soc. Am.*, **145**, 2277-2293.
- Kujawa, S. G., and Liberman, M. C. (2009). "Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss," *J. Neurosci.*, **29**, 14077-14085.
- Lopez-Poveda, E. A. (2014). Why do I hear but not understand? Stochastic undersampling as a model of degraded neural encoding of speech. *Front. Neurosci.*, **8**, 348.
- Lu, Z. L., and Doshier, B.A. (2008). "Characterizing observers using external noise and observer models: assessing internal representations with external noise," *Psychol. Rev.*, **115**, 44.
- Moore, B. C. J., Glasberg, B. R., Gaunt, T., and Child, T. (1991). "Across-channel masking of changes in modulation depth for amplitude- and frequency-modulated signals," *Q. J. Exp. Psychol.* **43A**, 327-347.

- Moore, B. C. J., and Sek, A. P. (1996). "Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking," *J. Acoust. Soc. Am.*, **100**, 2320-2331.
- Moore, B. C. J., Wojtczak, M., and Vickers, D. A. (1996). "Effect of loudness recruitment on the perception of amplitude modulation," *J. Acoust. Soc. Am.* **100**, 481-489.
- Moore, B. C., and Vinay (2019). "Effect of age on envelope regularity discrimination," *J. Acoust. Soc. Am.*, **146**, 1207-1211.
- Neri, P. (2010). "How inherently noisy is human sensory processing?," *Psychonom. Bull. Rev.*, **17**, 802-808.
- Ng, C. W., and Recanzone, G. H. (2018). "Age-related changes in temporal processing of rapidly-presented sound sequences in the macaque auditory cortex," *Cereb. Cortex*, **28**, 3775-3796.
- Overton, J. A., and Recanzone, G. H. (2016). "Effects of aging on the response of single neurons to amplitude-modulated noise in primary auditory cortex of rhesus macaque," *J. Neurophysiol.*, **115**, 2911-2923.
- Oxenham, A. J. (2016). "Predicting the perceptual consequences of hidden hearing loss," *Trends Hear.*, **20**, 2331216516686768.
- Paraouty, N., Ewert, S. D., Wallaert, N., and Lorenzi, C. (2016). "Interactions between amplitude modulation and frequency modulation processing: Effects of age and hearing loss," *J. Acoust. Soc. Am.*, **140**, 121-131.
- Parthasarathy, A., Lai, J., and Bartlett, E. L. (2016). "Age-related changes in processing simultaneous amplitude modulated sounds assessed using envelope following responses," *J. Assoc. Rese. Otolaryngol.*, **17**, 119-132.
- Parthasarathy, A., and Kujawa, S. G. (2018). "Synaptopathy in the aging cochlea: Characterizing early-neural deficits in auditory temporal envelope processing". *J. Neurosci.*, **38**, 7108-7119.
- Paul, B. T., Bruce, I. C., and Roberts, L. E. (2017). "Evidence that hidden hearing loss underlies amplitude modulation encoding deficits in individuals with and without tinnitus," *Hear. Res.*, **344**, 170-182.
- Pichora-Fuller, M. K., and MacDonald, E. (2009). "Sensory aging: hearing," in *Encyclopedia of Neuroscience*, edited by L.R. Squire (Elsevier), pp. 645-650.
- Ponsot, E., Varnet, L., Wallaert, N., Daoud, E., Shamma, S. A., Lorenzi, C., & Neri, P. (2021). Mechanisms of spectrotemporal modulation detection for normal-and hearing-impaired listeners. *Trends in hearing*, **25**, 2331216520978029.
- Ramamurthy, D. L., and Recanzone, G. H. (2020). "Age-related changes in sound onset and offset intensity coding in auditory cortical fields A1 and CL of rhesus macaques," *J. Neurophysiol.*, **123**, 1015-1025.
- Recanzone, G. (2018). "The effects of aging on auditory cortical function," *Hear. Res.*, **366**, 99-105.
- Resnik, J., and Polley, D. B. (2021). "Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise," *Neuron*, **109**, 1–13.
- Ross, B., Fujioka, T., Tremblay, K. L., and Picton, T. W. (2007). "Aging in binaural hearing begins in mid-life: evidence from cortical auditory-evoked responses to changes in interaural phase," *J. Neurosci.*, **27**, 11172-11178.
- Schoof, T., and Rosen, S. (2014). "The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners," *Front. Aging Neurosci.*, **6**, 307.

- Sek, A., Baer, T., Crinnion, W., Springgay, A., and Moore, B. C. (2015). "Modulation masking within and across carriers for subjects with normal and impaired hearing," *J. Acoust. Soc. Am.*, **138**, 1143-1153.
- Sergeyenko, Y., Lall, K., Liberman, M. C., and Kujawa, S. G. (2013). "Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline," *J. Neurosci.*, **33**, 13686-13694.
- Spiegel, M. F. and Green, D. M. (1981). "Two procedures for estimating internal noise," *J. Acoust. Soc. Am.*, **70**, 69-73.
- Takahashi, G. A., and Bacon, S. P. (1992). "Modulation detection, modulation masking, and speech understanding in noise in the elderly," *J. Speech Lang. Hear., Res.*, **35**, 1410-1421.
- Thoret, E., Varnet, L., Boubenec, Y., Ferriere, R., Le Tourneau, F.-M., Krause, B. and Lorenzi, C. (2020). "Characterizing amplitude and frequency modulation cues in natural soundscapes: A pilot study in four habitats of a biosphere reserve," *J. Acoust. Soc. Am.*, **147**, 3260-3274.
- Thoret, E., Caramiaux, B., Depalle, P., and Mc Adams, S. (2021). "Learning metrics on spectrotemporal modulations reveals the perception of musical instrument timbre," *Nat. Hum. Behav.*, **5**, 369-377.
- Varnet, L., Ortiz-Barajas, M. C., Erra, R. G., Gervain, J., and Lorenzi, C. (2017). "A cross-linguistic study of speech modulation spectra," *J. Acoust. Soc. Am.*, **142**, 1976-1989.
- Viana, L. M., O'Malley, J. T., Burgess, B. J., Jones, D. D., Oliveira, C. A., Santos, F., ... and Liberman, M. C. (2015). "Cochlear neuropathy in human presbycusis: Confocal analysis of hidden hearing loss in post-mortem tissue," *Hear. Res.*, **327**, 78-88.
- Viemeister, N. F., and Wakefield, G. H. (1991). "Temporal integration and multiple looks," *J. Acoust. Soc. Am.*, **90**, 858-865.
- Vilidaite, G., and Baker, D. H. (2017). "Individual differences in internal noise are consistent across two measurement techniques," *Vis. Res.*, **141**, 30-39.
- Wallaert, N., Moore, B. C. J., and Lorenzi, C. (2016). "Comparing the effects of age on amplitude modulation and frequency modulation detection," *J. Acoust. Soc. Am.* **139**, 3088-3096.
- Wallaert, N., Moore, B. C., Ewert, S. D., and Lorenzi, C. (2017). "Sensorineural hearing loss enhances auditory sensitivity and temporal integration for amplitude modulation," *J. Acoust. Soc. Am.*, **141**, 971-980.
- Wallaert, N., Varnet, L., Moore, B. C., and Lorenzi, C. (2018). "Sensorineural hearing loss impairs sensitivity but spares temporal integration for detection of frequency modulation," *J. Acoust. Soc. Am.*, **144**, 720-733.
- Wever, E. G. (1949). *Theory of Hearing*. New York: Wiley.
- Wever, E. G. and Bray, C. W. (1930). "The nature of acoustic response: the relation between sound frequency and frequency of impulses in the auditory nerve," *J. Exp. Psychol.* **13**, 373-87.
- Whiteford, K.L., Kreft, H.A., and Oxenham, A.J. (2017). "Assessing the role of place and timing cues in coding frequency and amplitude modulation as a function of age," *J. Assoc. Res. Otolaryngol.*, **18**, :619-633.
- Wingfield, A. (1996). "Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory". *Journal-American Academy of Audiology*, **7**, 175-182.

- Wu, P. Z., Liberman, L. D., Bennett, K., De Gruttola, V., O'Malley, J. T., and Liberman, M. C. (2019). "Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear," *Neurosci.*, **407**, 8-20.
- Yan, F. F., Hou, F., Lu, H., Yang, J., Chen, L., Wu, Y., ... and Huang, C. B. (2020). "Aging affects gain and internal noise in the visual system," *Scient. Rep.*, **10**, 1-10.
- Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A. and Cao, K. (2005). "Speech recognition with amplitude and frequency modulations," *Proc. Nat. Acad. Sci. USA*, **102**, 2293-2298.

Chapitre IV:

Application du paradigme de double-passe à la détection d'AM et de FM chez des sujets normo-entendants jeunes et malentendants âgés

IV.1. No effect of sensorineural hearing loss on double-pass consistency for amplitude- and frequency-modulation detection

IV.1.a. Résumé

Sensorineural hearing loss alters detection of amplitude and frequency modulation (AM, FM). If such changes result in part from increased internal noise in temporal envelope and fine-structure coding, not only sensitivity but also response consistency should be affected by hearing loss in modulation-detection tasks. These two aspects of behavior were estimated for a group of normal-hearing subjects (18-30 years) and a group of older subjects (45-80 years) with moderate sensorineural hearing loss, using 2-Hz sinusoidal AM or FM targets applied to a 500-Hz pure-tone carrier. AM and FM targets were masked by AM and FM noise, respectively. Each subject was tested at a sound-pressure level of 80 dB. Modulation-detection thresholds were first measured using an adaptive, forced-choice procedure. Double-pass consistency for modulation detection was then measured using a constant-stimuli, forced-choice paradigm repeated twice with the same stimuli. Subjects were tested at their individual modulation-detection threshold. Percent agreement of responses across passes was used to estimate consistency. Modulation-detection thresholds were significantly different across groups for FM only. Percent-agreement data for AM and FM were not significantly different across groups. Taken together with modelling results, these findings suggest that sensorineural hearing loss does not alter neural variability for temporal-modulation detection.

IV.1.b. Article

I. INTRODUCTION

Amplitude modulation (AM) and frequency modulation (FM) are two acoustic features playing a crucial role in the auditory perception of speech, music and environmental sounds (*e.g.*, Zeng *et al.*, 2005; Patil *et al.*, 2012; Varnet *et al.*, 2017; Thoret *et al.*, 2020). These features are generally assumed to be encoded as temporal-envelope cues and/or temporal fine-structure cues in the early auditory system *via* fluctuations in the mean discharge rate of auditory neurons and changes over time in the pattern of neural phase locking, respectively (*e.g.*, Moore and Sek, 1996; Joris *et al.*, 2004; Paraouty *et al.*, 2018).

Many studies indicate that sensorineural hearing loss has differential effects on AM and FM sensitivity. While sensorineural hearing loss impairs - sometimes substantially - the ability to detect FM (*e.g.*, Lacher-Fougere and Demany, 1998; Moore and Skrodzka, 2002; Buss *et al.*, 2004; Strelcyk and Dau, 2009; Ernst and Moore, 2012; Johannesen *et al.*, 2016; Kortlang *et al.*, 2016; Paraouty *et al.*, 2016; Wallaert *et al.*, 2018; Whiteford *et al.*, 2020), it spares or even improves the ability to detect AM, depending on whether AM-detection thresholds are measured at equal sound-pressure levels (*e.g.*, Bacon and Gleitman, 1992; Moore *et al.*, 1992; Moore and Glasberg, 2001; Grose *et al.*, 2016; Sek *et al.*, 2015; Schlittenlacher and Moore, 2016) or equal sensation levels (Bacon and Gleitman, 1992; Moore *et al.*, 1992; Moore *et al.*, 1996; Füllgrabe *et al.*, 2003; Wallaert *et al.*, 2017) for normal-hearing and hearing-impaired subjects.

The detrimental effect of sensorineural hearing loss on FM detection has been interpreted in different ways and the debate is not entirely resolved (for a review, see Moore, 2014). It may result from (i) reduced conversion of FM into temporal-envelope cues at the output of the cochlear filters, resulting from the loss of the active mechanism in the cochlea and the consequent broadening of cochlear filters (*e.g.*, Whiteford *et al.*, 2020; however see Wallaert *et al.*, 2018; Moore *et al.*, 2019; Attia *et al.*, Accepted, for a different conclusion) and/or (ii) a reduced ability to encode and/or use temporal-fine structure cues (Moore, 2008; Ewert *et al.*, 2020; Parthasarathy *et al.*, 2020). Several explanations have been put forward for this temporal fine-structure processing deficit (for a review, see Moore, 2014). It may be caused by a mismatch between neural temporal-fine structure information and the place on the basilar membrane that would “normally” respond to that information (Henry *et al.*, 2016, 2019; Parida and Heinz, 2021), or more speculatively by changes in the relative phase of response at different points along the basilar membrane (Ruggero, 1994), which would disrupt central mechanisms decoding temporal fine-structure cues based on correlation of the outputs of adjacent places (Shamma, 1985; Carney *et al.*, 2002). The beneficial effect of

sensorineural hearing loss on AM detection has been interpreted as resulting from either: (i) The loss of fast-acting amplitude compression in the cochlea as a consequence of damage to the active mechanism (*i.e.*, a loss of function of outer hair cells) and the subsequent perceptual magnification of temporal-envelope cues (Moore *et al.*, 1992; Moore *et al.*, 1996; Wallaert *et al.*, 2017; Jennings *et al.*, 2018) and/or (ii) enhanced neural phase locking to temporal-envelope cues in auditory-nerve fibers (Kale and Heinz, 2010, 2012), an effect possibly resulting from changes in short-term neural adaptation (Scheidt *et al.*, 2010).

Modelling studies aiming to simulate detection of sinusoidal AM and FM and discrimination of complex AM and FM patterns suggest that in addition to the systematic (*i.e.*, deterministic) deficits listed above, the variance of “internal noise” - that is sources of neural variability located along the auditory pathway (Javel and Viemeister, 2000; Faisal *et al.*, 2008; Amitay *et al.*, 2013) - may increase because of cochlear damage (Derleth *et al.*, 2001; Ives *et al.*, 2014; Paraouty *et al.*, 2016; Wallaert *et al.*, 2017, 2018; Ewert *et al.*, 2020) and contribute to the specific pattern of AM and FM sensitivity found in hearing-impaired people. Four studies conducted with the same experimental paradigm, the same stimuli and different model structures indicate that the increase of the variance of internal-noise sources caused by cochlear damage may be substantial (*e.g.*, Paraouty *et al.*, 2016; Wallaert *et al.*, 2017, 2018; Ewert *et al.*, 2020). In these models, internal noise was implemented as additive and multiplicative noise sources decreasing the fidelity of temporal-envelope and temporal fine-structure coding. For FM detection, an increase in neural variability in the auditory structures responsible for temporal fine-structure coding may explain why FM sensitivity worsens for hearing-impaired subjects (Ewert *et al.*, 2020). For AM detection, an increase in neural variability in the auditory structures responsible for temporal-envelope coding may counteract the beneficial effect of the loss of fast-acting amplitude compression in the cochlea on AM detection, that otherwise would be too good (Wallaert *et al.*, 2017, 2018).

In line with the notion that cochlear damage may cause an increase in internal noise, some hearing-impaired subjects report pure tones as sounding noise-like, and unmodulated tones appear to fluctuate (Huss and Moore, 2005a,b; Moore and Vinay, 2009). The noisy and fluctuating character of the auditory percept of unmodulated carriers may contribute to the alteration in AM and FM sensitivity found in hearing-impaired subjects. Several hypotheses can be put forward as to the origin of this increase in neural variability following cochlear damage. First, it may result from central alterations such as changes in neural excitation and inhibition at the level of the primary auditory cortex. Part of these central changes may be related to peripheral changes such as those caused by cochlear synaptopathy, the loss of synapses between hair cells and auditory-nerve fibers, and subsequent neural deafferentation (Kujawa and Liberman, 2009). Consistent with this hypothesis, neurophysiological studies conducted with guinea pigs, cats and mice show that noise or tone exposure (Basura *et al.*, 2015; Seki and Eggermont, 2003; Eggermont, 2017) or pharmacologically-induced cochlear synaptopathy (Resnick and Polley, 2021) result in

increased spontaneous firing rates and random surges of hyper-synchronized activity in cortical neurons. Second, a modelling study by Goodman *et al.* (2017) showed that massive loss of auditory-nerve fibers should cause an increase in neural variability at the low brainstem level (the cochlear nucleus), suggesting that cochlear synaptopathy may also yield an increase in internal noise in more peripheral auditory structures. It is noteworthy that the hearing-impaired subjects who participated in the studies conducted by Paraouty *et al.* (2016), Wallaert *et al.* (2017, 2018) and Ewert *et al.* (2020) were older than their normal-hearing controls (the age of hearing-impaired subjects ranged between 45 and 67 years). Ageing *per se* is also associated with cochlear synaptopathy (Sergeyenko *et al.*, 2013; Viana *et al.*, 2015; Wu *et al.*, 2019) and central changes (*e.g.*, Frisina, 2009), part of which may be related to peripheral changes, such as those caused by cochlear synaptopathy. As in the case of cochlear damage, spontaneous activity of cortical neurons increases with age in non-human animals, presumably because of an age-related imbalance of neural excitation and inhibition in the auditory cortex (Ng and Recanzone, 2018; Recanzone, 2018). Thus, the association of cochlear damage with ageing as in the case of presbycusis - the age-related sensorineural hearing loss - may explain why internal noise had to be increased substantially in simulation studies attempting to account for AM and FM detection thresholds for (relatively) old hearing-impaired subjects (Ewert *et al.*, 2020; Paraouty *et al.*, 2016; Wallaert *et al.*, 2017, 2018).

Still, it is possible that the abnormal detection and discrimination capacities found in hearing-impaired subjects only resulted from alterations of (deterministic) temporal-modulation coding mechanisms rather than from a genuine increase in neural variability at peripheral and/or central stages of the auditory system. The goal of the present study is to test the hypothesis that sensorineural hearing loss causes an increase in internal noise for AM and FM detection. Moore *et al.* (2019b) developed a psychophysical paradigm that may be used to assess the perceptual consequences of increased internal noise on modulation processing. This paradigm measured the ability of human subjects to discriminate regular 8-Hz sinusoidal AM from AM that had the same mean AM rate but was irregular in AM depth and AM rate. The amount of irregularity was varied to determine a threshold called “envelope-regularity discrimination” (ERD) thresholds. Moore *et al.* (2019b) reasoned that regular AM should sound irregular if the auditory system is noisy, and this - in turn - should yield higher (poorer) ERD thresholds. In other words, for subjects with higher internal noise, the amount of irregularity in the stimulus should be higher to overcome the effects of the irregularity in the internal representation of the temporal envelope caused by increased internal noise. Moore and Vinay (2019b) measured ERD thresholds for normal-hearing subjects aged between 20 and 27 years and older normal-hearing subjects aged between 64 and 80 years. Consistent with the idea that age is associated with increased internal noise in the modulation domain, the older subjects showed poorer ERD thresholds than the younger subjects.

Here, an alternative approach is proposed to explore the impact of cochlear damage

combined with ageing on internal noise for both temporal-envelope and temporal-fine structure coding. This approach relies on the idea that internal noise does not only limit sensitivity; it also limits *consistency* in auditory judgments. In other words, for a given stimulus and a given task, a high level of internal noise is expected to cause more variable responses from trial to trial and a decrease in the consistency of the subjects' responses. The "double-pass" method (Green, 1964) is a psychophysical paradigm aiming at estimating the consistency of subjects' responses from the agreement of individual responses between two trials using identical stimuli. Estimates of the variance of internal noise can then be obtained *via* simulations of models based on signal-detection theory (SDT, Green, 1964; Spiegel and Green, 1981) or models of temporal-modulation processing by the human auditory system (Attia *et al.*, Accepted). In two companion studies (Attia *et al.*, Accepted; Attia *et al.*, submitted), the double-pass paradigm was applied to slow (2 Hz) and fast (20 Hz) AM and FM detection for a group of young normal-hearing subjects (aged between 18 and 30 years) and a group of older normal-hearing subjects (aged between 40 and 62 years). Double-pass consistency data and estimates of internal noise were not significantly different across groups, providing no evidence that neural variability is affected by ageing *per se* below about 60 years. The present study tested the hypothesis that double-pass consistency will be poorer - and thus internal noise is stronger - for a group of hearing-impaired subjects compared to a group of younger normal-hearing subjects, because of increased neural variability in the auditory structures involved in AM and FM coding. In the main experiment, double-pass consistency was assessed in these subjects for sinusoidal AM and FM detection using a relatively slow (2-Hz) modulation rate, a low (500 Hz) carrier frequency and a bandpass (AM or FM) noise masker. To estimate double-pass consistency for modulation detection, a 2-interval, 2-alternative forced-choice (2I, 2AFC) detection-in-noise task was run *twice* with identical sequences of modulation noise in the two successive passes. Modulation noises used for AM and FM detection are generated in the AM and FM domains, respectively. A single level of the variance of modulation noise was used for each type of noise. All subjects were tested at their individual modulation-detection threshold as measured in a setup experiment. Analyses of double-pass data (*i.e.*, percent agreement scores) based on signal-detection theory were then conducted to estimate the variance of internal noise in each experimental condition and for each group of subjects (Green, 1964; Spiegel and Green, 1981; Ponsot *et al.*, 2021).

II. METHODS

A. Participants

Nineteen young normal-hearing subjects (referred to as NHy subjects) aged between 18 and 30 years (mean=23 years; standard deviation (SD)=4 years) and ten older hearing-impaired subjects (referred to as Hlo subjects) aged between 45 and 80 years (mean=65 years; SD=9 years) participated. All NHy listeners were recruited via classified ads. in public places (Rennes, France). Hlo subjects were recruited from the database of a hearing-care devices laboratory (Rennes) and from the database of the ENT clinic at Institut Arthur Verne (Paris, France). Participants with any of the following were excluded: psychiatric or attentional disorders, acoustic trauma, Ménière's disease, severe tinnitus, middle/external ear damage, or conductive hearing loss (asymmetry between the two ears >15 dB hearing level (HL) at audiometric frequencies between 0.75 and 1.5 kHz). Differences between air- and bone-conduction thresholds ranged from 0 to 10 dB HL across Hlo subjects, indicating that the hearing loss was of sensory origin. One NHy subject was excluded from this study because he/she could not complete the experiments. Five Hlo subjects were excluded because they failed at the performing the adaptive task (n=3) or because they could not participate anymore in the experiments because of the Covid-19 pandemic that occurred in 2020 (n=2).

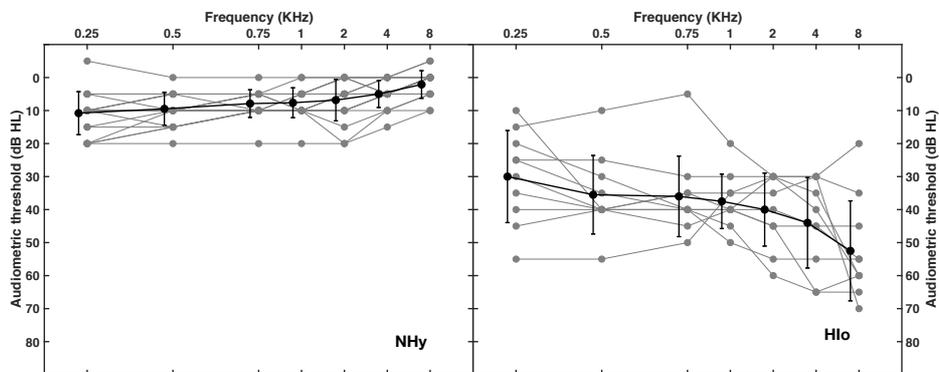


Figure 1. Pure-tone audiometric thresholds (mean and individual data) for the right ear for the 19 NHy (left panel) and 10 Hlo subjects (right panel). Error bars indicate ± 1 standard deviation of the mean.

The pure-tone audiogram was initially assessed at octave frequencies between 0.25 and 8 kHz using an Otosuite audiometer manufactured in Denmark equipped with supra-aural TDH39 headphones. Based on pure-tone audiometry, the hearing of NHy subjects was classified as normal below 4 kHz for both ears (Goodman, 1965) and symmetric below 8 kHz. All NHy subjects had audiometric thresholds lower than or equal to 20 dB HL between 0.25 and 2 kHz in each ear, and the difference between audiometric thresholds measured in the left and right ears was lower than 10 dB HL between 0.25 and 8 kHz. All subjects were tested in their right ear. Pure-tone average (PTA) estimates calculated for audiometric frequencies between 0.5 and 4 kHz for each individual Hlo subject ranged between 30 and 51 dB HL (mean PTA=42 dB HL; SD=7 dB) for the right ear.

Pure-tone audiometric thresholds (mean and individual data) for the right ear are shown in the left and right panels of Fig. 1 for the 19 NHy and 10 Hlo subjects, respectively. The audiometric thresholds at 0.5 kHz in the right ear ranged between 0 and 20 dB HL (mean=10 dB HL; SD=5 dB) for NHy subjects, and between 10 and 55 dB HL (mean=36 dB HL; SD=12 dB) for Hlo subjects. Audiometric thresholds were significantly higher (that is, poorer) for Hlo subjects than for NHy subjects, as shown by a Welch *t*-test (an adaptation of the Student *t*-test to compare two groups with different sizes) conducted on audiometric thresholds at 0.5 kHz for NHy and Hlo groups [$t(11)=6.6$; $p<0.05$]. A Welch *t*-test conducted on PTAs also showed a significant difference between NHy and Hlo groups [$t(10)=12.7$; $p<0.05$].

NHy and Hlo subjects were tested in the same conditions (except for the last three Hlo subjects who participated in this study because of the covid-19 pandemic). All subjects were fully informed about the purpose and the conditions of the study and provided written consent. All subjects were paid 10 Euros per hour to participate in this study. This study was approved by the local ethical committee at Paris University (CERES, N° IRB: 20143200001072).

B. Stimuli

All stimuli were generated digitally at a sampling rate of 44.1 kHz using the MATLAB environment (Mathworks, release 2017 b). They were presented to the right ear of each participant at 80 dB SPL using an external soundcard (Audioengine D3, 24-bit resolution) and delivered via Beyerdynamic DT 770 pro 250 Ohms headphone monaurally (right ear) calibrated beforehand with an artificial ear (4153 6-cc coupler, IEC standard-60318-1). Most experiments were conducted in an IAC double-wall soundproof booth and levels between 0.125 and 4 kHz were calibrated using a Bruel & Kjaer 2250 sound-level meter. For the last three Hlo subjects who participated in this study, tests were performed in a quiet cabin used for speech and hearing therapy.

The whole experimental session was divided into two successive phases referred to as setup and main experiments. Both experiments consisted of AM and FM detection tasks based on a

2 interval, 2 alternative forced-choice (2IAFC) paradigm. During the setup experiment, AM and FM detection thresholds were measured for each subject using an adaptive procedure. During the main experiment, detection performance and double-pass consistency for AM and FM detection were measured for each subject tested at his/her own AM or FM detection thresholds using a constant-stimuli procedure. In both experiments, each trial was divided into two successive observation intervals separated by a silent inter-stimulus interval (ISI). The target and comparison stimuli were presented in a random order in the two successive observation intervals. The duration of each stimulus was 1 sec. The duration of the ISI was 400 ms. All stimuli were modulated in amplitude or in frequency by a modulation noise (a white Gaussian noise filtered between 1 and 40 Hz in the AM or FM domain)¹. More precisely, the modulation noise was generated by filtering a 1-sec Gaussian white noise $n(t)$ with zero mean between 1 and 40 Hz using a 4th-order Butterworth filter. The standard deviation of the modulation noise, $\sigma_{AM,FM}^{ext}$, was set 0.07 stimulus units [modulation-depth units (m.d.u.) for the AM noise, or modulation-index units (m.i.u.) for the FM noise]. The target signal was sinusoidal AM or FM at a fixed rate of 2 Hz that was added to the modulation-noise masker before being applied to the 500-Hz pure-tone carrier. A 100-ms raised-cosine ramp was applied to the beginning and the end of each stimulus. All stimuli were normalized in terms of their root-mean-square (RMS) power.

¹ In this last study, we divided by half the bandwidth of the external noise to avoid that the internal representation of this external noise (and more precisely, its internal/perceived modulation spectrum) differs between NHy and Hlo subjects due to differences in cochlear frequency selectivity. Indeed, Hlo subjects have wider cochlear filters than NHy subjects (in the present case, their cochlear filters are probably 1.5 to 2 times wider than those of NHy subjects). Consequently, the faster modulation components of the external noise are less attenuated in Hlo subjects than in NHy subjects. An external modulation noise ranging between 0.5 and 80 Hz produces sidebands in the audio-frequency domain with the highest frequency components being +/- 80 Hz compared to 500 Hz. The total bandwidth of the sound modulated by this external noise reaches 160 Hz (2x80 Hz). However, the equivalent rectangular bandwidth (ERB) of the cochlear filter centered at 500 Hz is approximately 80 Hz at a moderate level of stimulation in NHy subjects. Temporal modulations above 40 Hz (half the ERB at 500 Hz) are therefore attenuated by this cochlear filter. These fast modulations should be less attenuated in Hlo subjects because the ERB of the cochlear filter tuned to 500 Hz can be as high as 160 Hz in some Hlo subjects. For this reason, the bandwidth of the external noise was reduced to about 40 Hz in this last study.

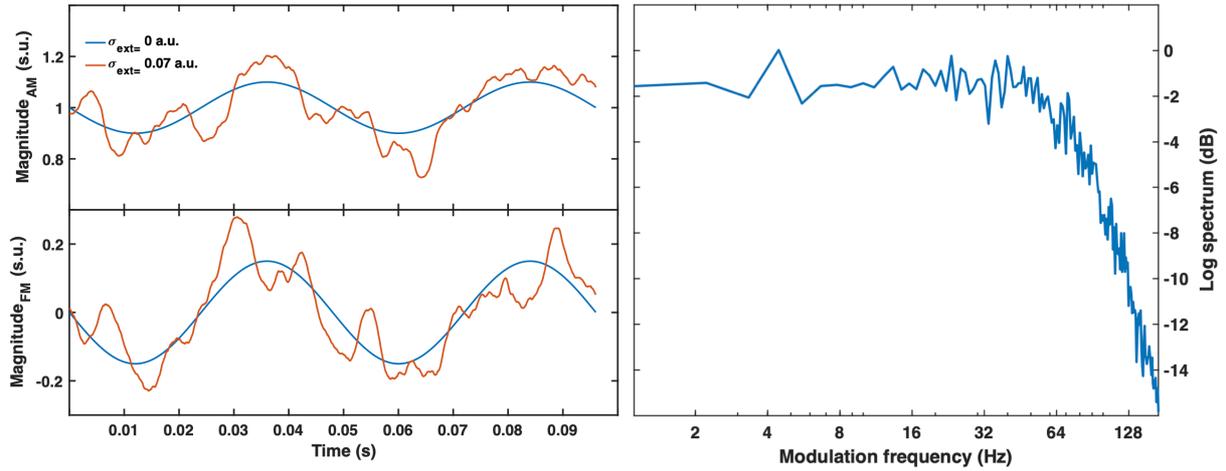


Figure 2. Left panels: Time course of instantaneous amplitude of the 2-Hz sinusoidal AM (top panel) and FM (bottom panel) target. Modulators were obtained for independent samples of modulation noise. The standard deviation of the masker, $\sigma_{ext}^{AM,FM}$, was set to two values: 0 and 0.07 units (m.d.u. or m.i.u.). Right panel: Long-term power spectrum of modulation noise with a standard deviation set to 0.07 units. Modulation spectra were averaged over 100 (1-sec long) samples.

The comparison stimuli (C_{AM} or C_{FM}) were sinusoidal carriers (with a frequency $f_c = 500$ Hz) whose instantaneous amplitude or frequency was modulated using the bandpass noise masker, $n(t)$. The target stimuli (T_{AM} or T_{FM}) were sinusoidal carriers (with a frequency $f_c = 500$ Hz) whose instantaneous amplitude or frequency were modulated by the sum of a sinusoidal signal and the bandpass noise masker, $n(t)$. The sinusoidal signal had a rate f_m of 2 Hz and modulation index m for AM stimuli or $\beta = \Delta f / f_m$ for FM stimuli.

Equation [1] describes the target stimulus $T_{AM}(t)$:

$$T_{AM}(t) = [1 + m \sin(2\pi f_m t + \varphi_{AM}) + \sigma_{ext}^{AM} n(t)] \sin(2\pi f_c t + \varphi_c) \quad [1]$$

Equation [2] describes the comparison stimulus $C_{AM}(t)$:

$$C_{AM}(t) = [1 + \sigma_{ext}^{AM} n(t)] \sin(2\pi f_c t + \varphi_c) \quad [2]$$

Equation [3] describes the target stimulus $T_{FM}(t)$:

$$T_{FM}(t) = \sin[2\pi f_c t + \varphi_c + \beta \sin(2\pi f_m t + \varphi_{FM}) + \sigma_{ext}^{FM} n(t)] \quad [3]$$

Equation [4] describes the comparison stimulus $C_{FM}(t)$:

$$C_{FM}(t) = \sin[2\pi f_c t + \varphi_c + \sigma_{ext}^{FM} n(t)] \quad [4]$$

where t is time expressed in seconds (s), m is the AM depth (ranging between 0 and 1) of the AM target signal, $\beta = \Delta f/f_m$ where Δf is the frequency excursion of the FM target signal expressed in Hz, f_m is the modulation rate of the target signal expressed in Hz (2 Hz), φ_{AM} and φ_c are the starting phases of the AM target signal and the carrier, respectively. φ_{AM} and φ_c were randomly and independently chosen (between 0 and 2π radians) for each stimulus in the setup experiment (measure of AM-detection thresholds) and in the first pass of the main experiment (measure of double-pass consistency for AM detection). φ_{FM} and φ_c are the starting phases of the FM target signal and the carrier, respectively. φ_{FM} and φ_c were randomly and independently chosen (between 0 and 2π radians) for each stimulus in the setup experiment (measure of FM-detection thresholds) and in the first pass of the main experiment (measure of double-pass consistency for FM detection). For both AM- and FM-detection tasks, the modulation-noise masker, $n(t)$, was refreshed between intervals and trials in the setup experiment, and in the first pass of the main experiment. The modulation-noise maskers used in the second pass of the main experiment were identical to those used in the first pass of the main experiment (in corresponding trials). Modulation-noise maskers were different across subjects.

Figure 2 shows the time course of instantaneous amplitude of 2-Hz AM and FM target modulators and the long-term average spectrum of the modulation noise used in the experiments. Modulators and log spectrum were obtained for 100 independent, 1-sec long independent samples of modulation noise. The left panels illustrate the corruption caused by the modulation noise on the target sinusoidal modulators in the temporal domain. The right panel shows that the modulation power of the modulation noise used in the experiments drops above 40 Hz.

III. PROCEDURE

The setup and main experiments were completed in the same order for each subject: In the setup experiment, an adaptive staircase was used to measure individual AM and FM detection thresholds corresponding to three levels of performance: 64 % ($d'=0.5$), 76 % ($d'=1$), and 84 % ($d'=1.5$) correct detection. In the main-experiment, a double-pass consistency task was used to measure the percentage of agreement for the same detection tasks at a fixed level of performance based on individual modulation-detection thresholds measured in the setup experiment. The whole series of experiments was based on one hour per session with three sessions spread out over a 1-2 week period for each subject. For each task and for each trial, the subject was instructed to choose the interval containing the target sinusoidal AM or FM.

Each subject was tested for the following 6 conditions: two types of modulation (AM vs. FM) and three levels of (targeted) performance: 64 % ($d'=0.5$), 76 % ($d'=1$), and 84 % ($d'=1.5$) correct detection. The individual AM and FM detection thresholds were then used to set the modulation strength (*i.e.*, the modulation depth or frequency excursion, respectively)

of the target signals in the subsequent double-pass experiments. Hence, the AM and FM detection tasks in the double-pass consistency paradigm were performed at same level of difficulty.

The double-pass consistency task was conducted with the purpose of estimating percent agreement in the subjects' responses for corresponding trials in two passes. Percent-correct detection was also calculated across the two passes for each subject.

1. Setup experiment: Detection thresholds

AM detection thresholds (modulation depth, m) and FM detection thresholds (frequency excursion, Δf) were first measured for each subject with adaptive staircases that attempted to converge on 64 % ($d'=0.5$), 76 % ($d'=1$) and 84 % ($d'=1.5$) correct detection using a 2I, 2AFC paradigm and a weighted 1-up, 1-down method (Garcia-Pérez, 1998).

Each block consisted of a variable number of trials. Each trial was divided into two successive observation intervals containing the target and comparison stimuli, presented in random order. The bandpass noise masker was refreshed for each interval and trial. Each observation interval was marked by a red light on the computer screen that was synchronized with stimulus' presentation. Subjects were instructed to report the observation interval containing the target stimulus. Visual feedback as to the correct answer was given to the subject after each trial. The dependent variables for these measures were the modulation depth m for the AM detection task and the frequency excursion Δf for the FM detection task. At the beginning of each staircase, the modulation strength of AM and FM targets was set to a supra-threshold level: The AM target was presented at a modulation depth of 0.8 (-2 dB when expressed as $20\log m$) and the FM target was presented at a frequency excursion of 8 Hz. Before the first four reversals, the tracking variable (the modulation strength of AM and FM targets) changed by a step (a factor) up of 4 after a wrong response and by a step (a factor) down of 2 after a correct response. The tracking variable then changed with the 1-down 1-up rule up to sixteen reversals, using different steps to reach 64 %, 76 %; and 86 % of correct responses (Garcia-Pérez, 1998). Twenty reversals were performed for each series of measures and the mean threshold was taken as the geometric mean of the tracking variable at the last sixteen reversals. The order of the conditions was randomized across runs.

AM- and FM-detection thresholds at 2 Hz were first measured once in the absence of any modulation-noise masker in order to familiarize subjects with the detection tasks. AM and FM detection thresholds were then measured three times in each condition (*i.e.*, six staircases). Three measures of detection threshold were thus carried out in each experimental condition and the final detection threshold for each condition was taken as the mean of these three estimates.

2. Main experiment: Double-pass consistency

The consistency of subjects' responses was then measured in each experimental condition and for each subject using the double-pass paradigm (Green, 1964; Hasan *et al.*, 2012; Vilidaitė and Baker, 2016). Here, AM and FM detection were measured using a 2I, 2AFC paradigm with constant stimuli. In total, each subject was tested over $2 \times 100 = 200$ trials in each condition (AM or FM; modulation rate of 2 Hz; modulation-noise masker with a standard deviation of 0.07 units (m.d.u. or m.i.u. for AM and FM, respectively); the target modulation strength was set to the mean individual detection threshold corresponding to d' 0.5, 1 or 1.5, as assessed in the setup experiment). In other words, each subject was run twice with identical stimuli, making two "passes" of 100 trials in each condition. Unlike Attia *et al.* (Accepted), the presentation order of the target and comparison stimuli was fixed across the two passes for each trial. The 6 experimental conditions were tested in random order. Thus, 1200 trials were run per subject. These 1200 trials were split into one series of double-pass experiments in which all 6 conditions were completed all at once. The second pass was conducted after a 10-min break. Short breaks were given to subjects if they needed it in the middle of the first pass and the middle of the second pass. Double-pass testing lasted about 40 min. Each block therefore consisted of 600 trials. Five training trials where modulation targets were presented at a modulation strength yielding a d' of 3 were inserted at the beginning of each block to familiarize the subjects with the task and stimuli. For each block, each trial was divided into two successive observation intervals containing the target and comparison stimuli, presented in random order. Each observation interval was marked by a red light on the computer screen that was synchronized with stimulus' presentation. Subjects were instructed to report the observation interval containing the target stimulus. Subjects were not given any visual feedback after each trial, except for the first five training trials. Subjects' responses to these training trials were not taken into consideration in the assessment of double-pass consistency and percent-correct detection. For each trial, the two possible responses ("interval 1" and "interval 2") were shown by two boxes on the computer screen. To limit any mistakes caused by reduced motor capacities (*e.g.*, limited hand mobility), the H10 subjects gave their responses by pointing to the appropriate box on the screen, and the experimenter provided the response using the computer mouse.

For each condition, double-pass consistency was calculated as the percentage of agreement (PA) between the subjects' responses to corresponding trials within the first and the second pass. The accuracy of responses was also calculated as the percentage of correct responses (PC) over the 200 trials [100 trials (pass 1) + 100 trials (pass 2)] corresponding to a given condition.

IV. RESULTS

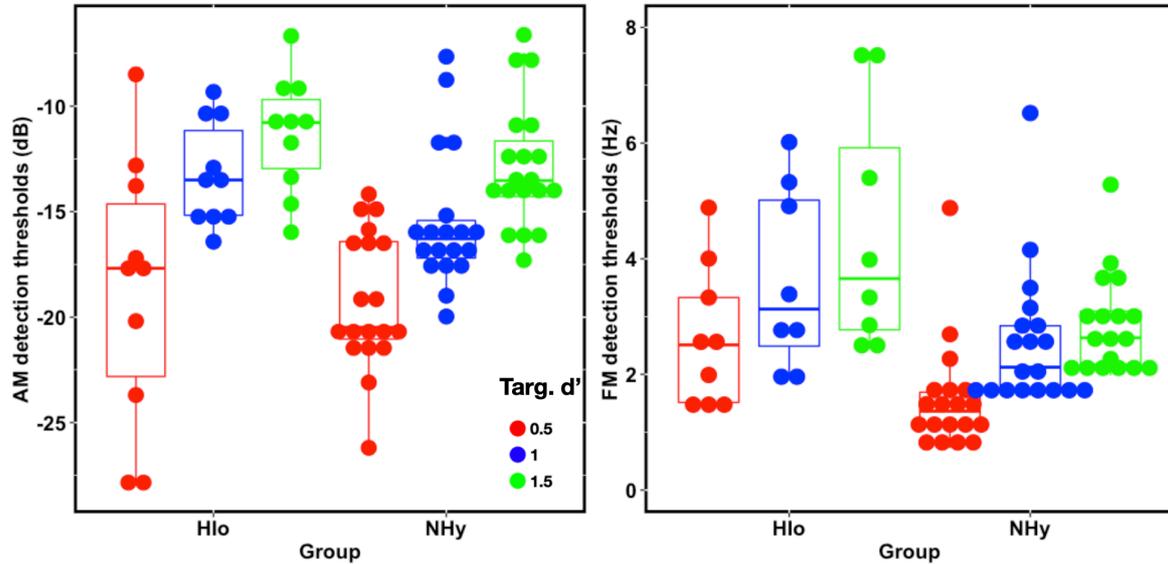


Figure 3. Individual modulation-detection thresholds of NHy and Hlo subjects. Left panel: AM detection thresholds; right panel: FM detection thresholds. For each panel, detection thresholds are shown for each level of sensitivity targeted by the adaptive procedure (red: $d'=0.5$; blue: $d'=1$; green: $d'=1.5$). On each box, horizontal lines indicate the median threshold; boxes span the 25th-75th percentiles and vertical lines span the minimum and maximum thresholds.

The individual AM- and FM-detection thresholds for the NHy and Hlo subjects are plotted in the left and right panels of Fig. 3, respectively. For each panel, detection thresholds are shown for each level of sensitivity targeted by the adaptive procedure (red: $d'=0.5$; blue: $d'=1$; green: $d'=1.5$).

Statistical analyses of modulation-detection thresholds and PA and PC scores for AM and FM detection tasks were conducted using linear mixed-effect models with R (R Core Team Version 4.0.2). These mixed-effect linear regressions were obtained using the lme4 library. One model was run per dependent variable. These four models included: i) AM-detection thresholds; ii) FM-detection thresholds; iii) PA scores and iv) PC scores as dependent variables. Targeted sensitivity level (d' of 0.5, 1 or 1.5) and Group (NHy or Hlo) were fixed effects for modulation-detection thresholds analyses. The modulation-type factor (AM or FM) was introduced in models used to assess significant effects for double-pass consistency data only. Subjects were included as a random factor and more specifically as a random intercept. All fixed variables were contrast coded.

A log transformation was applied to AM and FM detection thresholds and an arcsine transformation was applied to PA and PC scores. Linear mixed-model assumptions were validated using graphic diagnostic tools (i- residuals followed a normal distribution; ii-random

effects followed a multivariate normal distribution and iii- residual variance was homogenous). Significance level was assessed using type III ANOVAS with alpha set at 0.05. Data sets consisted of 29 subjects with 87 observations (3 observations per subject per modulation type condition) for adaptive tasks, and of 174 observations (6 observations per subject) for the double-pass consistency task.

1. AM-detection thresholds

As expected, the statistical conducted on AM-detection thresholds showed a significant main effect of targeted sensitivity level [$\beta=5.5$; $SE=5.5$; $\chi^2(2) =173$; $p<0.0001$]. However, there was no significant main effect of Group [$\beta=0.5$; standard error (SE)=1.1; $\chi^2(2) =1.5$; $p=0.2$] and no significant interaction between Group and targeted sensitivity level. Consistent with previous studies assessing AM-detection thresholds for normal-hearing and hearing-impaired subjects tested at the same SPL, the current study showed that AM sensitivity was not significantly affected by sensorineural hearing loss and ageing.

2. FM-detection thresholds

The analysis conducted on FM-detection thresholds revealed : i) A significant main effect of Group [$\beta=14.1$; $SE=3.9$; $\chi^2(1) =12.5$; $p<0.001$], indicating that FM-detection thresholds were significantly higher (poorer) for HIo subjects than for NHy subjects (HIo detection thresholds being a factor 2.2 higher on average than NHy thresholds); ii) as expected, a significant main effect of targeted sensitivity level [$\beta=8.6$; $SE=1.1$; $\chi^2(2) =128.1$; $p<0.0001$]. In line with previous studies comparing FM-detection thresholds for normal-hearing and hearing-impaired subjects tested at the same SPL or the same sensation level, the current study showed that FM sensitivity was significantly impaired by sensorineural hearing loss combined with ageing.

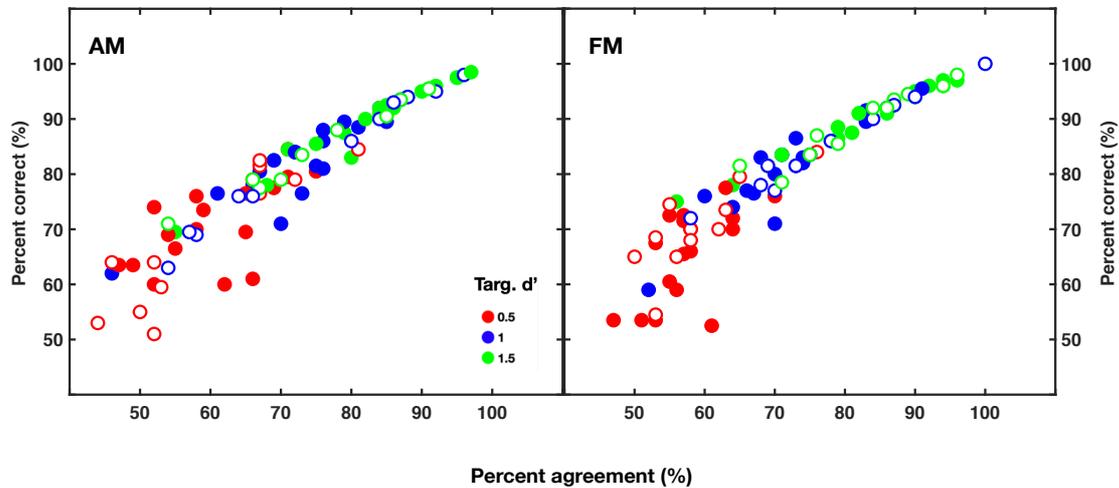


Figure 4. Performance versus consistency curves showing percent correct (PC) and percent agreement (PA) data for each subject (closed symbols: NHy subjects; open symbols : Hlo subjects). For each panel, the individual data are shown for each level of sensitivity targeted by the initial adaptive procedure (red: targeted $d'=0.5$; blue: targeted $d'=1$; green: targeted $d'=1.5$).

3. Percent-agreement data

Figure 4 shows the double-pass consistency (PA) and PC data for each NHy (filled circles) and Hlo (open circles) subject. The left and right panels show AM and FM data, respectively. For each panel, the individual NHy and Hlo data are shown for each level of sensitivity targeted by the initial adaptive procedure (red: targeted $d'=0.5$; blue: targeted $d'=1$; green: targeted $d'=1.5$).

All subjects could perform the adaptive task in the AM- and FM-detection experiments. For all experimental conditions, a positive correlation was found between PA and PC data: As expected, higher detection performance was associated with higher double-pass consistency (Bravais-Pearson r ; ($0.7 < r < 0.9$) for NHy and Hlo subjects).

For the PA scores there was: i) no significant main effect of Group [$\beta=2.70$; $SE=3.88$; $\chi^2(1)=0.48$; $p=0.49$]; ii) a significant main effect of targeted sensitivity level [$\beta=14.5$; $SE=3.49$; $\chi^2(2)=29.31$; $p<0.0001$]; iii) no significant main effect of modulation type [$\beta=1.1$; $SE=3.49$];

$\chi^2(1)=0.099$; $p=0.75$]; iv) no significant interactions between fixed effects (Group, sensitivity level, modulation type).

4. Percent-correct data

For the PC scores there was: i) no significant main effect of Group [$\beta=4.97$; $SE=3.23$; $\chi^2(1)=2.36$; $p=0.12$]; ii) a significant main effect of targeted sensitivity level with different slopes for d' of 1 [$\beta=14.25$; $SE=2.77$; $\chi^2(1)=48.31$; $p<0.0001$] and for d' of 1.5 [$\beta=18.40$; $SE=2.77$; $\chi^2(1)=48.31$; $p<0.0001$]; iii) No significant main effect of modulation type [$\beta=1.9$; $SE=2.7$; $\chi^2(1)=0.49$; $p=0.5$]; iv) no significant interactions between fixed effects (all $p>0.05$).

5. Assessment of internal noise

We also converted our measurements of performance and agreement into values of internal noise using a SDT model (Green, 1964; Burgess and Colborne, 1988). We followed the established procedure where the internal-to-external noise ratio is inferred from a typical SDT model fitted to the percentage of correct responses and percentage of agreement obtained across repeated presentation of the same trials (for a detailed description of this inference procedure, the reader is invited to refer to previous studies, *e.g.*, Joosten and Neri, 2012). Here, the model assumes no response bias, a choice justified by the fact that the influence of response bias on our measurements, if any, was controlled by the randomization of the observation intervals across the two passes (see Procedure).

We first averaged the percent agreement and percent correct values over the three levels of targeted d' (0.5, 1 and 1.5) so that, for each individual and modulation type (AM/FM), estimates of internal noise were based on the repetition of 300 trials, which corresponds to the minimal resolution required for comparison across groups/conditions (Hasan *et al.*, 2012). We followed this procedure to infer internal noise values for all these conditions, but only values that fell within the interpretable range [0.2; 5] returned by the SDT model (Neri, 2010) were kept for subsequent analyses. Overall, this approach allowed us to infer plausible internal noise values for 67% of our dataset.

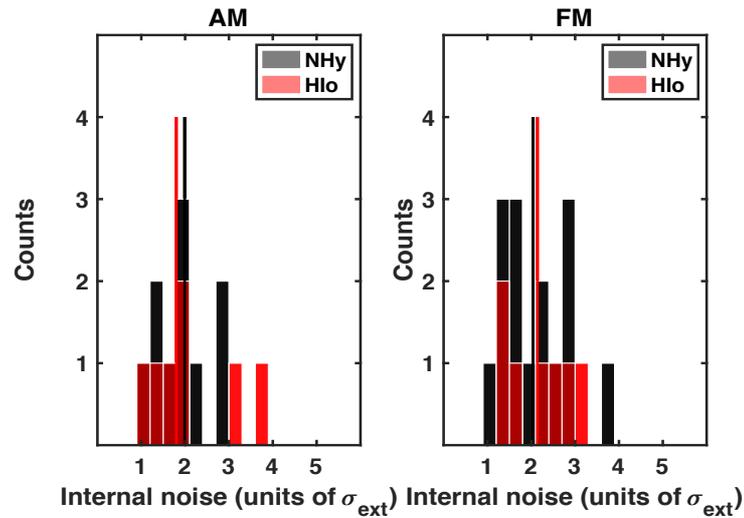


Figure 5. Distribution of internal noise values for AM and FM for the two groups of subjects tested (black color: NHy subjects; red color: Hlo subjects) inferred from percent correct and percent agreement data collected. Vertical lines show the corresponding median value of each distribution.

Figure 5 plots the distributions of internal noise estimates for AM and FM for the two groups. Internal noise values were very similar for AM and FM for both groups, with median values varying within a very narrow range ([1.8 – 2.1] units of ext. noise), consistent with the range observed for other low-level auditory psychophysical tasks (Neri, 2010; Ponsot *et al.*, 2021). Wilcoxon rank-sum tests showed no significant differences between groups for both AM and FM conditions ($p > .1$).

V. DISCUSSION

1. Differential effects of hearing loss on masked AM and FM detection

Modulation-detection thresholds were measured for slow sinusoidal AM or FM in the presence of a bandpass AM or FM noise masker, respectively, for two groups of subjects: young normal-hearing and older hearing-impaired subjects. We did not find any significant effect of group on AM-detection thresholds. This finding is consistent with the outcome of previous studies assessing AM sensitivity for normal-hearing and hearing-impaired subjects tested at the same SPL (*e.g.*, Bacon and Gleitman, 1992; Moore *et al.*, 1992; Moore and Glasberg, 2001; Sek *et al.*, 2015; Grose *et al.*, 2016; Schlittenlacher and Moore, 2016). Although these studies did not use a bandpass AM noise masker as here, some of them measured AM sensitivity using AM targets masked by a sinusoidal AM (Sek *et al.*, 2015) or applied sinusoidal AM targets to noise carriers (*e.g.*, Bacon and Gleitman, 1992; Moore *et al.*,

1992), the intrinsic envelope fluctuations of the noise carrier at the output of the cochlear filters therefore acting as a bandpass AM masker (Dau *et al.*, 1997). Our hearing-impaired subjects were much older (age ranged between 45 and 80 years) than our normal-hearing controls (age range 18 to 30 years). Some studies indicate that ageing *per se* might spare masked and unmasked AM sensitivity (Attia *et al.*, submitted; Schoof and Rosen, 2014; Paraouty *et al.*, 2016; Grose *et al.*, 2019). Still, other studies show that ageing is associated with a mild, though significant degradation in AM sensitivity (Takahashi and Bacon, 1992; He *et al.*, 2008; Füllgrabe *et al.*, 2015; Wallaert *et al.*, 2016), AM-detection thresholds increasing with age at a rate of about 0.1-1.5 dB per age decade at high presentation levels (Carcagno and Plack, 2020). The detrimental effect of suprathreshold deficits caused by age on AM sensitivity - if present - was probably compensated by loss of fast-acting amplitude compression in the cochlea as a consequence of damage to the active mechanism and the subsequent perceptual magnification of temporal-envelope cues.

This is, to our knowledge, the first study assessing the effects of sensorineural hearing loss on FM-detection thresholds measured in the presence of an FM noise. Our results indicate that FM-detection thresholds were significantly elevated for hearing-impaired subjects compared to normal-hearing subjects. This result is consistent with the outcome of many studies assessing FM sensitivity for unmasked FM targets or FM targets masked by sinusoidal AM for subjects with sensorineural hearing loss (*e.g.*, Lacher-Fougere and Demany, 1998; Moore and Skrodzka, 2002; Buss *et al.*, 2004; Strelcyk and Dau, 2009; Ernst and Moore, 2012; Johannesen *et al.*, 2016; Kortlang *et al.*, 2016; Paraouty *et al.*, 2016; Wallaert *et al.*, 2018; Whiteford *et al.*, 2020). Our hearing-impaired subjects being much older than our normal-hearing controls, it cannot be excluded that part of this FM-detection deficit might arise because of supra-thresholds deficits caused by ageing *per se*, as shown by previous studies comparing young and older subjects with normal hearing (He *et al.*, 2007; Grose and Mamo, 2012; Wallaert *et al.*, 2016; Whiteford *et al.*, 2017; however see Attia *et al.* (submitted) showing that ageing *per se* does not alter masked FM detection thresholds measured in the same conditions as in here).

Overall, the present data indicate that sensorineural hearing loss combined with ageing has differential effects on masked AM and FM sensitivity measured at a slow (2 Hz) modulation rate in the presence of a modulation noise: Sensorineural hearing loss combined with ageing degrades FM sensitivity while sparing AM sensitivity.

2. No effect of hearing loss on double-pass consistency

We did not find any significant effect of group on double-pass consistency data and estimates of internal noise for AM and FM detection. These findings are not consistent with the notion that sensorineural hearing loss associated with ageing yields an increase in internal noise, that is increased neural variability in the peripheral and central auditory structures

playing a role in AM and FM perception. A small cohort of hearing-impaired subjects ($n=10$) with audiometric thresholds ranging between 15 and 40 dB HL at 500 Hz (mean=35 dB HL) was included in the present study (9 out of 10 subjects had audiometric thresholds > 30 dB HL at 500 Hz). Thus, it cannot be excluded that changes in the variance of internal noise caused by cochlear damage are too small to be reliably measured by the double-pass paradigm when subjects have mild-moderate hearing loss at the tested (audio) frequency, and that internal noise may increase for hearing losses greater than 35-40 dB HL. Further work is therefore warranted to explore effects of cochlear damage using a larger cohort of subjects with more severe sensorineural hearing loss. Still, the present findings suggest that at least part of the alterations in AM and FM detection reported by Wallaert *et al.* (2017, 2018) for a group of 12 subjects (aged 50-64 years) showing moderate hearing loss (mean audiometric threshold: 40 dB HL at 500 Hz; min=30 dB HL; max=50 dB HL) might not result from a genuine increase in neural variability, as proposed by Wallaert *et al.* (2017, 2018). This interpretation is compatible with the fact that FM-detection thresholds were elevated by approximately the same amount (by a factor 1.5-2.2) for hearing-impaired subjects as for normal-hearing controls in both studies. It is possible that these alterations in modulation detection were mainly caused by (deterministic) perceptual distortions of temporal-envelope and temporal fine-structure cues resulting from outer and inner hair-cell damage, loss of auditory-nerve fibers or more central deficits resulting from peripheral damage. The latter may be modelled by altering decision statistics (template matching, envelope power) in computational models of temporal-modulation perception (*e.g.*, Wallaert *et al.*, 2017, 2018; Ewert *et al.*, 2020).

VI. CONCLUSIONS

Detection thresholds and double-pass consistency were estimated for sinusoidal AM and FM for a group of young normal-hearing subjects aged between 18 and 30 years and a group of older hearing-impaired subjects aged between 45 and 80 years. Modulation rate and carrier frequency were set to 2 and 500 Hz, respectively. AM and FM stimuli were masked by AM and FM noise, respectively. Importantly, all subjects were tested at the same SPL of 80 dB. Consistent with previous studies, AM-detection thresholds did not differ significantly across groups, but FM-detection thresholds were significantly poorer for hearing-impaired subjects. For both AM and FM detection, double-pass consistency data and estimates of internal noise derived from signal-detection analysis did not differ significantly across groups.

Overall, these results do not provide evidence that neural variability in peripheral and central structures of the auditory system increases with both sensorineural hearing loss and ageing.

ACKNOWLEDGMENTS

This work was supported by two grants from the Collège National des Audioprothésistes (France) and Entendre SAS. This work was also supported by ANR-17-EURE-0017.

VII. REFERENCES

- Amitay, S., Guiraud, J., Sohoglu, E., Zobay, O., Edmonds, B. A., Zhang, Y. X., and Moore, D R. (2013). "Human decision making based on variations in internal noise: An EEG study," *PloS one*, **8**, e68928. DOI: 10.1371/journal.pone.0068928.
- Bacon, S. P., and Gleitman, R. M. (1992). "Modulation detection in subjects with relatively flat hearing losses," *J. Speech Lang. Hear. Res.*, **35**, 642-653.
- Basura, G.J., Koehler, S.D., and Shore, S.E. (2015). "Bimodal stimulus timing-dependent plasticity in primary auditory cortex is altered after noise exposure with and without tinnitus," *J. Neurophysiol.*, **114**, 3064-3075.
- Buss, E., Hall III, J. W., and Grose, J. H. (2004). "Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss," *Ear Hear.*, **25**, 242-250.
- Burgess, A. E., and Colborne, B. (1988). "Visual signal detection. IV. Observer inconsistency," *J. Opt. Soc. Am A*, **5**, 617-627.
- Carcagno, S., and Plack, C. J. (2020). "Effects of age on psychophysical measures of auditory temporal processing and speech reception at low and high levels," *Hear. Res.*, **400**, 108117.
- Carney, L. H., Heinz, M. G., Evilsizer, M. E., Gilkey, R. H., and Colburn, H. S. (2002). "Auditory phase opponency: a temporal model for masked detection at low frequencies," *Acta Acustica United with Acustica*, **88**, 334-347.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," *J. Acoust. Soc. Am.*, **102**, 2906-2919.
- Derleth, R. P., Dau, T., and Kollmeier, B. (2001). "Modeling temporal and compressive properties of the normal and impaired auditory system," *Hear. Res.*, **159**, 32-149.
- Eggermont, J. J. (2017). "Acquired hearing loss and brain plasticity," *Hear. Res.*, **343**, 176-190.
- Ernst, S. M., and Moore, B. C. (2012). "The role of time and place cues in the detection of frequency modulation by hearing-impaired listeners," *J. Acoust. Soc. Am.*, **131**, 4722-4731.
- Ewert, S. D., Parag, N., and Lorenzi, C. (2020). "A two-path model of auditory modulation detection using temporal fine structure and envelope cues," *Eur. J. Neurosci.*, **51**, 1265-1278.
- Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008). "Noise in the nervous system," *Nature Rev. Neurosci.*, **9**, 292-303.
- Frisina, R. D. (2009). "Age-related hearing loss: ear and brain mechanisms," *Annals New York Acad. Sci.*, **1170**, 708-717.
- Füllgrabe, C., Meyer, B., and Lorenzi, C. (2003). "Effect of cochlear damage on the detection of complex temporal envelopes," *Hear. Res.*, **178**, 35-43.

- Füllgrabe, C., Moore, B. C., and Stone, M. A. (2015). "Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition," *Front. Aging Neurosci.*, **6**, 347.
- Goodman, D. F., Winter, I. M., Léger, A. C., de Cheveigné, A., and Lorenzi, C. (2018). "Modelling firing regularity in the ventral cochlear nucleus: Mechanisms, and effects of stimulus level and synaptopathy," *Hear. Res.*, **358**, 98-110.
- Green, D. M. (1964). "Consistency of auditory detection judgments," *Psychol. Rev.*, **71**, 392-407.
- Grose, J. H., and Mamo, S. K. (2012). "Frequency modulation detection as a measure of temporal processing: Age-related monaural and binaural effects," *Hear. Res.*, **294**, 49-54.
- Grose, J. H., Menezes, D. C., Porter, H. L., and Griz, S. (2016). "Masking period patterns and forward masking for speech-shaped noise: age-related effects," *Ear Hear.*, **37**, 48.
- Grose, J. H., Buss, E., and Elmore, H. (2019). "Age-related changes in the auditory brainstem response and suprathreshold processing of temporal and spectral modulation," *Trends Hear.*, **23**, DOI: 10.1177/2331216519839615.
- Hasan, B. A. S., Joosten, E., and Neri, P. (2012). "Estimation of internal noise using double passes: Does it matter how the second pass is delivered?," *Vis. Res.*, **69**, 1-9.
- He, N. J., Mills, J. H., Ahlstrom, J. B., and Dubno, J. R. (2008). "Age-related differences in the temporal modulation transfer function with pure-tone carriers," *J. Acoust. Soc. Am.*, **124**, 3841-3849.
- Henry, K. S., Kale, S., Scheidt, R. E., and Heinz, M. G. (2011). "Auditory brainstem responses predict auditory nerve fiber thresholds and frequency selectivity in hearing impaired chinchillas," *Hear. Res.*, **280**, 236-244.
- Henry, K. S., Kale, S., and Heinz, M. G. (2016). "Distorted tonotopic coding of temporal envelope and fine structure with noise-induced hearing loss," *J. Neurosci.*, **36**, 2227-2237.
- Henry, K. S., Sayles, M., Hickox, A. E., and Heinz, M. G. (2019). "Divergent auditory nerve encoding deficits between two common etiologies of sensorineural hearing loss," *J. Neurosci.*, **39**, 6879-6887.
- Huss, M., and Moore, B. C. (2005a). "Dead regions and pitch perception," *J. Acoust. Soc. Am.*, **117**, 3841-3852.
- Huss, M., and Moore, B. C. (2005a). "Dead regions and noisiness of pure tones," *Int. J. Audiol.*, **44**, 599-611.
- Ives, D. T., Kalluri, S., Strelcyk, O., Sheft, S., Miermont, F., Coez, A., ... and Lorenzi, C. (2014). "Effects of noise reduction on AM perception for hearing-impaired listeners," *J. Assoc. Rese. Otolaryngol.*, **15**, 839-848.
- Javel, E., and Viemeister, N. F. (2000). "Stochastic properties of cat auditory nerve responses to electric and acoustic stimuli and application to intensity discrimination," *J. Acoust. Soc. Am.*, **107**, 908-921.
- Jennings, S. G., Chen, J., Fultz, S.E., Ahlstrom, J.B., and Dubno, J. R. (2018). "Amplitude modulation detection with a short-duration carrier: Effects of a precursor and hearing loss," *J. Acoust. Soc. Am.*, **143**, 2232-2243.
- Johannesen, P. T., Pérez-González, P., Kalluri, S., Blanco, J. L., and Lopez-Poveda, E. A. (2016). "The influence of cochlear mechanical dysfunction, temporal processing deficits, and age on the intelligibility of audible speech in noise for hearing-impaired listeners," *Trends Hear.*, **20**, 2331216516641055.

- Joosten, E. R., and Neri, P. (2012). "Human pitch detectors are tuned on a fine scale, but are perceptually accessed on a coarse scale," *Biol. Cyber.* **106**, 465-482.
- Joris, P. X., Schreiner, C. E., and Rees, A. (2004). "Neural processing of amplitude-modulated sounds," *Physiol. Reviews*, **84**, 541-577.
- Kale, S., and Heinz, M. G. (2010). "Envelope coding in auditory nerve fibers following noise-induced hearing loss," *J. Assoc. Res. Otolaryngol.*, **11**, 657-673.
- Kale, S., and Heinz, M. G. (2012). "Temporal modulation transfer functions measured from auditory-nerve responses following sensorineural hearing loss," *Hear. Res.*, **286**, 64-75.
- Kortlang, S., Mauermann, M., and Ewert, S. D. (2016). "Suprathreshold auditory processing deficits in noise: Effects of hearing loss and age," *Hear. Res.*, **331**, 27-40.
- Lacher-Fougère, S., and Demany, L. (1998). "Modulation detection by normal and hearing-impaired listeners," *Audiol.*, **37**, 109-121.
- Moore, B. C. (2008). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," *J. Acoust. Soc. Am.*, **9**, 399-406.
- Moore, B. C. (2014). *"Auditory processing of temporal fine structure: Effects of age and hearing loss"*. World Scientific.
- Moore, B. C., and Glasberg, B. R. (2001). "Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.*, **110**, 1067-1073.
- Moore, B. C. J., and Sek, A. P. (1996). "Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking," *J. Acoust. Soc. Am.*, **100**, 2320-2331.
- Moore, B.C.J., Shailer, M.J., and Schooneveldt, G.P. (1992). "Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss," *Br. J. Audiol.* **26**, 229-237.
- Moore, B. C., and Skrodzka, E. (2002). "Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation," *J. Acoust. Soc. Am.*, **111**, 327-335.
- Moore, B. C., Mariathasan, S., and Sęk, A. P. (2019a). "Effects of age and hearing loss on the discrimination of amplitude and frequency modulation for 2-and 10-Hz rates," *Trends Hear.*, **23**, DOI: 10.1177/2331216519853963.
- Moore, B.C.J., Sęk, A.P., Vinay, and Füllgrabe, C. (2019b). "Envelope regularity discrimination," *J. Acoust. Soc. Am.*, **145**, 2861-2870.
- Moore, B. C., and Vinay, S. N. (2019). "Effect of age on envelope regularity discrimination," *J. Acoust. Soc. Am.*, **146**, 1207-1211.
- Moore, B.C.J., and Vinay (2009). "Amplitude modulation detection by listeners with unilateral dead regions," *J. Am. Acad. Audiol.*, **20**, 597-606.
- Ng, C. W., and Recanzone, G. H. (2018). "Age-related changes in temporal processing of rapidly-presented sound sequences in the macaque auditory cortex," *Cereb. Cortex*, **28**, 3775-3796.
- Neri, P. (2010). "How inherently noisy is human sensory processing?," *Psych. Bull. Rev.*, **17**(6), 802-808.
- Paraouty, N., Ewert, S. D., Wallaert, N., and Lorenzi, C. (2016). "Interactions between amplitude modulation and frequency modulation processing: Effects of age and hearing loss," *J. Acoust. Soc. Am.*, **140**, 121-131.

- Paraouty, N., Stasiak, A., Lorenzi, C., Varnet, L., and Winter, I. M. (2018). "Dual coding of frequency modulation in the ventral cochlear nucleus," *J. Neurosci.*, **38**, 4123-4137.
- Parida, S., & Heinz, M. G. (2021). Noninvasive Measures of Distorted Tonotopic Speech Coding Following Noise-Induced Hearing Loss. *Journal of the Association for Research in Otolaryngology*, **22**, 51-66.
- Parthasarathy, A., Hancock, K. E., Bennett, K., DeGruttola, V., and Polley, D. B. (2020). "Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing," *Elife*, **9**, e51419.
- Patil, K., Pressnitzer, D., Shamma, S., and Elhilali, M. (2012). "Music in our ears: the biological bases of musical timbre perception". *PLoS Comput. Biol.* **8**:e1002759. DOI: 10.1371/journal.pcbi.1002759
- Ponsot, E., Varnet, L., Wallaert, N., Daoud, E., Shamma, S. A., Lorenzi, C., and Neri, P. (2021). "Mechanisms of spectrotemporal modulation detection for normal-and hearing-impaired listeners," *Trends Hear.*, **25**, 1-19, DOI: 10.1177/2331216520978029.
- Recanzone, G. (2018). "The effects of aging on auditory cortical function". *Hear. Res.*, **366**, 99-105.
- Resnik, J., and Polley, D. B. (2021). "Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise," *Neuron*, **109**, 1–13.
- Ruggero, M. A. (1994). "Cochlear Delays and Traveling Waves: Comments on 'Experimental Look at Cochlear Mechanics,'" *Audiol.*, **33**, 131-142.
- Sek, A., Baer, T., Crinnion, W., Springgay, A., and Moore, B. C. (2015). "Modulation masking within and across carriers for subjects with normal and impaired hearing," *J. Acoust. Soc. Am.*, **138**, 1143-1153.
- Seki, S., and Eggermont, J. J. (2003). "Changes in spontaneous firing rate and neural synchrony in cat primary auditory cortex after localized tone-induced hearing loss," *Hear. Res.* **180**, 28-38.
- Scheidt, R. E., Kale, S., and Heinz, M. G. (2010). "Noise-induced hearing loss alters the temporal dynamics of auditory-nerve responses". *Hear. Res.*, **269**, 23-33.
- Schlittenlacher, J., and Moore, B. C. (2016). "Discrimination of amplitude-modulation depth by subjects with normal and impaired hearing," *J. Acoust. Soc. Am.*, **140**, 3487-3495.
- Schoof, T., and Rosen, S. (2014). "The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners". *Front. Aging Neurosci.*, **6**, 307.
- Sek, A., Baer, T., Crinnion, W., Springgay, A., and Moore, B. C. (2015). "Modulation masking within and across carriers for subjects with normal and impaired hearing," *J. Acoust. Soc. Am.*, **138**, 1143-1153.
- Sergeyenko, Y., Lall, K., Liberman, M. C., and Kujawa, S. G. (2013). "Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline," *J. Neurosci.*, **33**, 13686-13694.
- Shamma, S. A. (1985). "Speech processing in the auditory system II: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve," *J. Acoust. Soc. Am.*, **78**, 1622-1632.
- Spiegel, M. F. and Green, D. M. (1981). "Two procedures for estimating internal noise," *J. Acoust. Soc. Am.*, **70**, 69-73.
- Strelcyk, O., and Dau, T. (2009). "Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing," *J. Acoust. Soc. Am.*, **125**, 3328-3345.

- Takahashi, G. A., and Bacon, S. P. (1992). "Modulation detection, modulation masking, and speech understanding in noise in the elderly," *J. Speech Lang. Hear. Res.*, **35**, 1410-1421.
- Thoret, E., Varnet, L., Boubenec, Y., Ferriere, R., Le Tourneau, F.-M., Krause, B. and Lorenzi, C. (2020). "Characterizing amplitude and frequency modulation cues in natural soundscapes: A pilot study in four habitats of a biosphere reserve," *J. Acoust. Soc. Am.*, **147**, 3260-3274.
- Varnet, L., Ortiz-Barajas, M. C., Erra, R. G., Gervain, J., and Lorenzi, C. (2017). "A cross-linguistic study of speech modulation spectra," *J. Acoust. Soc. Am.*, **142**, 1976-1989.
- Viana, L. M., O'Malley, J. T., Burgess, B. J., Jones, D. D., Oliveira, C. A., Santos, F., ... and Liberman, M. C. (2015). "Cochlear neuropathy in human presbycusis: Confocal analysis of hidden hearing loss in post-mortem tissue," *Hear. Res.*, **327**, 78-88.
- Wallaert, N., Moore, B. C., and Lorenzi, C. (2016). "Comparing the effects of age on amplitude modulation and frequency modulation detection," *J. Acoust. Soc. Am.*, **139**, 3088-3096.
- Wallaert, N., Moore, B. C., Ewert, S. D., and Lorenzi, C. (2017). "Sensorineural hearing loss enhances auditory sensitivity and temporal integration for amplitude modulation," *J. Acoust. Soc. Am.*, **141**, 971-980.
- Wallaert, N., Varnet, L., Moore, B. C., and Lorenzi, C. (2018). "Sensorineural hearing loss impairs sensitivity but spares temporal integration for detection of frequency modulation," *J. Acoust. Soc. Am.*, **144**, 720-733.
- Whiteford, K. L., Kreft, H. A., and Oxenham, A. J. (2017). "Assessing the role of place and timing cues in coding frequency and amplitude modulation as a function of age," *J. Assoc. Res. Otolaryngol.*, **18**, 619-633.
- Whiteford, K. L., Kreft, H. A., and Oxenham, A. J. (2020). "The role of cochlear place coding in the perception of frequency modulation," *Elife*, **9**, e58468.
- Wu, P. Z., Liberman, L. D., Bennett, K., De Gruttola, V., O'Malley, J. T., and Liberman, M. C. (2019). "Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear," *Neuroscience*, **407**, 8-20.
- Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A. and Cao, K. (2005). "Speech recognition with amplitude and frequency modulations," *Proc. Nat. Acad. Sci. USA*, **102**, 2293-2298.
- Zwicker, E. (1952). "Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones," *Acta Acust. united with Acust.*, **2**, 125-133.

Chapitre V:

Discussion et conclusion

V.1. Précisions sur le paradigme expérimental

V.1.a. Choix du paradigme expérimental

Les protocoles psychophysiques traditionnels éliminent la variabilité des mesures en limitant les analyses à des données comportementales moyennées sur plusieurs jugements perceptifs individuels. Cette approche psychophysique - qualifiée de « molaire » - tente ainsi d'éliminer l'influence de paramètres autres que les paramètres physiques liés au stimulus. Or, toute mesure psychophysique est influencée par plusieurs types de facteur (sensoriels vs. cognitifs, déterministes vs. aléatoires). Les variables distinctes du stimulus et impliquées dans les jugements perceptifs sont identifiables à partir des multiples séquences de réponses, comme dans le cas du **paradigme de double-passe**. Cette approche alternative - qualifiée de « moléculaire » (Green, 1964) - s'avère utile car elle fournit l'opportunité de réaliser une analyse rigoureuse de la manière dont les **facteurs internes aléatoires** influencent les réponses individuelles.

Point notable, le paradigme de double-passe fournit une **méthode d'estimation directe** et très intuitive des manifestations comportementales du bruit interne : à un niveau de sensibilité donné, un faible niveau de bruit interne est directement associé à une forte valeur de « cohérence » dans les jugements auditifs. Le paradigme de double-passe présente certains avantages : il fournit des estimations similaires à celles fournies par d'autres méthodes, et il estime un bruit interne global (additif et multiplicatif), contrairement à d'autres méthodes comme celle du bruit équivalent qui estime uniquement un bruit interne de nature additive.

Enfin, certains modèles auditifs - de par leur structure - sont susceptibles de prédire les réponses individuelles des participants à chaque essai dans une tâche psychophysique. Néanmoins, la majorité des études de modélisation auditive porte sur des réponses moyennes (la performance ou le seuil estimé sur un bloc d'essai) plutôt que sur des réponses individuelles (Green, 1964). Le **modèle à banc de filtres de modulation** a rencontré un franc succès pendant ces dernières décennies, car il a démontré sa capacité à rendre compte d'un très grand nombre de données moyennes de détection, discrimination et identification en perception auditive chez le sujet jeune et âgé normo-entendant et chez le sujet malentendant.

Le paradigme de double-passe fournit l'opportunité de tester plus avant la capacité de ce modèle à prédire les jugements individuels, d'essai en essai, dans des tâches auditives de détection d'AM et de FM.

Pour toutes ces raisons, nous avons décidé de limiter notre investigation à ce paradigme expérimental afin de fournir des réponses aux trois questions centrales de cette thèse. Dans la section suivante, nous détaillons à nouveau nos choix expérimentaux et théoriques.

V.1.b. Choix des paramètres expérimentaux

V.1.b.i. Cadence de modulation

Dans la première étude, nous avons choisi d'utiliser des signaux d'AM et de FM de cadences de modulation lente (2 Hz) et rapide (20 Hz) appliqués à une fréquence porteuse basse (500 Hz), car ces paramètres de modulation et de porteuse permettent de tester l'hypothèse d'une double-code.

D'une manière générale, des mesures de bruit interne différentes entre deux tâches sont interprétées comme reflétant l'opération de mécanismes sensoriels et/ou cognitifs distincts (Ewert et al., 2020).

Plus précisément, et compte tenu de nos connaissances sur le traitement d'AM et de FM :

- i) l'hypothèse d'un mécanisme sensoriel unique pour le traitement d'AM et de FM est corroborée si les mesures de bruit interne sont identiques pour la détection d'AM et de FM aux deux cadences ;
- ii) l'hypothèse de deux mécanismes distincts pour le traitement de l'AM et de la FM est corroborée si les mesures de bruit interne diffèrent entre la détection de FM lente (2 Hz) et les autres conditions : la détection d'AM lente (2 Hz) et rapide (20 Hz) et la détection de FM rapide (20 Hz)

Par ailleurs, compte tenu des différences observées dans le traitement et la perception d'AM lente et rapide (*e.g.*, Wright & Dai, 1998) :

- i) l'hypothèse d'un mécanisme sensoriel unique pour le traitement d'AM lente et rapide est corroborée si les mesures de bruit interne sont identiques pour la détection d'AM aux deux cadences ;
- ii) l'hypothèse de deux mécanismes sensoriels distincts pour le traitement d'AM lente et rapide est corroborée si les mesures de bruit interne diffèrent pour la détection d'AM aux deux cadences.

Dans la seconde étude, ces paramètres ont été maintenus afin d'étendre cette recherche à une population de sujets normo-entendants plus âgés, testés dans les mêmes conditions expérimentales que dans la première étude.

Dans la troisième étude, nous nous sommes limités à une cadence de modulation lente. Il s'agit avant tout de réduire la durée des tests afin de minimiser les effets de fatigue pour les participants âgés et malentendants, sachant que la procédure implique un nombre conséquent d'essais. Le choix d'une cadence de modulation lente est dicté par les résultats d'études antérieures démontrant que l'effet délétère de la perte auditive neurosensorielle sur la détection de FM pourrait en partie résulter d'une capacité réduite à encoder les indices de TFS_N , résultant probablement d'une dégradation du mécanisme de verrouillage en phase des réponses des neurones périphériques. Enfin, cette condition expérimentale permet également de vérifier l'existence d'un double code pour le traitement d'AM et de FM.

V.1.b.ii. Caractéristiques du bruit externe

Dans ces trois études, nous avons choisi de synthétiser et appliquer une source de bruit externe dans la dimension d'intérêt compte tenu de la tâche, c'est-à-dire dans le *domaine de l'AM et de la FM*. Dans le paradigme de double-passe et pour un niveau de performance donné, l'augmentation du bruit externe devrait a priori produire une augmentation de la cohérence (mesurée par le pourcentage d'accord) des réponses des participants. Cette hypothèse est testée dans les deux premières études.

Les valeurs de bruit externe utilisées dans ces études sont basées sur les résultats d'une étude pilote estimant les seuils de détection masqués en AM et FM en fonction de l'écart-type du bruit externe de modulation (σ_{ext}) sur trois auditeurs normo-entendants jeunes. Nous avons utilisé une cadence de modulation lente (2 Hz) et rapide (20 Hz) et une procédure adaptative à deux intervalles et deux alternatives en choix forcé estimant la profondeur de modulation ou l'excursion de fréquence correspondant à 76 % de réponses correctes ($d'=1$) sur la fonction psychométrique (Levitt, 1971). Sept valeurs d'écart-type de bruit externe d'AM et de FM ont été testées dans cette expérience pilote. Celles-ci sont comprises entre 0 et 0.3 unités (unités de profondeur ou unités d'excursion). Les résultats sont présentés dans la figure 89 ci-dessus. Ces courbes correspondent aux courbes « seuils versus contraste » utilisées dans le paradigme du bruit équivalent. En AM comme en FM, les courbes sont plus plates à 20 Hz qu'à 2 Hz ; une inflexion (*i.e.*, une augmentation des seuils) est apparente autour de 0.05 unités à 20 Hz.

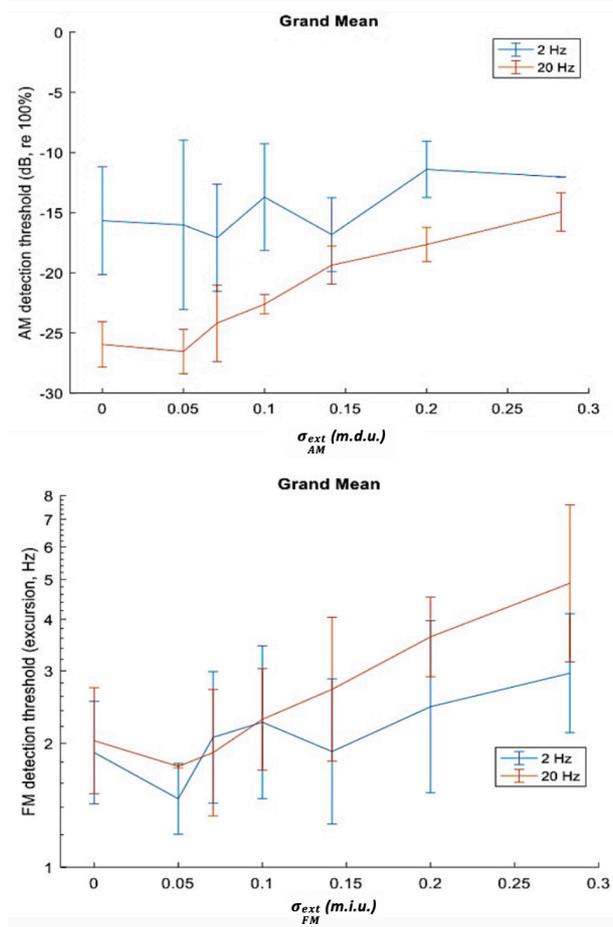


Figure 89. Seuils de détection d'AM (en haut) et de FM (en bas) masquées par un bruit de modulation (bruit d'AM pour la détection d'AM, bruit de FM pour la détection de FM). Seuils de détection moyens (sur deux répétitions) obtenus sur trois participants normo-entendants jeunes en fonction de l'écart-type du bruit externe exprimé en unité de profondeur pour l'AM et unités d'excursion pour la FM, avec pour paramètre la cadence de modulation (courbe bleue : $f_m = 2$ Hz ; courbe orange : $f_m = 20$ Hz). Les seuils de détection d'AM sont exprimés en dB [$20 \log(m)$]. Les seuils de détection de FM sont exprimés en Hz (échelle logarithmique). Les barres d'erreurs indiquent l'erreur type par rapport au seuil moyen.

Nous avons défini « l'écart-type de coupure » comme la valeur de σ_{ext} à partir de laquelle le seuil commence à se dégrader. Cette valeur de coupure a été déterminée visuellement sur chaque courbe. Nous avons ainsi déterminé deux valeurs d'écart-type de bruit externe à partir de ces mesures de seuils de détection d'AM et de FM masqués de sorte que :

- Celles-ci n'induisent pas de sur-modulation dans le stimulus;
- La plus basse valeur de σ_{ext} soit juste inférieure à la valeur de « l'écart-type de coupure » sur chacune des fonctions de masquage ;
- La plus haute valeur de σ_{ext} soit juste supérieure à la valeur de « l'écart-type de coupure » sur chacune des fonctions de masquage ;
- Les deux valeurs soient dans un rapport de 2 ($\sigma_{ext,1} = 2 \sigma_{ext,2}$).

Ces deux valeurs (0.07 et 0.14 unités de profondeur ou d'index de modulation) ont été retenues et utilisées dans les deux premières études afin d'estimer l'effet de ce paramètre sur la cohérence des réponses des sujets normo-entendants jeunes et âgés en détection d'AM et de FM. La dernière étude a dû être optimisée en termes de temps et de niveau de difficulté. En effet, les sujets malentendants âgés ont eu besoin d'un entraînement supplémentaire avant de démarrer la tâche. Une augmentation de la variabilité externe aurait certainement rendu la tâche impossible à réaliser. Nous avons donc sélectionné la valeur de variance de bruit externe la plus basse pour cette troisième étude ($\sigma_{ext}=0.07$ unités).

V.1.b.iii. Niveaux de sensibilité (performance) ciblés

Dans l'ensemble des études, trois niveaux de sensibilité (et donc, trois niveaux de performances sur les fonctions psychométriques sous-jacentes) ont été ciblés ($d'=0.5$, $d'=1$, $d'=1.5$) afin d'échantillonner au mieux la courbe reliant la performance (PC) à la cohérence (le pourcentage d'accord, PA). Il était attendu qu'une augmentation de la profondeur de modulation ou de l'index de modulation produise une amélioration de la performance (PC) dans les tâches de détection d'AM et de FM.

V.1.c. Choix théoriques

Dans la première étude, les estimations de bruit interne sont obtenues à partir d'un modèle computationnel testé dans les mêmes conditions que les sujets normo-entendants jeunes et dont les réponses sont ajustées aux données empiriques de double-passe. Ce modèle est similaire à celui utilisé par Wallaert et al. (2018). Il implémente un banc de filtres cochléaires, un mécanisme de transduction (rectification, compression, adaptation), un banc de filtres de modulation sélectifs ($Q=1$), un bruit additif interne Gaussien σ_{int} suivi d'un mécanisme de décision utilisant un mécanisme d'appariement de gabarit. Pour chaque condition expérimentale, le modèle est simulé pour une large gamme de valeurs d'écart-types du bruit interne additif σ_{int} . La valeur de σ_{int} retenue pour chaque condition expérimentale correspond à celle minimisant les erreurs entre le couple de données {cohérence (PA); performance (PC)} réelles (mesurées sur les sujets) et simulées par le modèle.

Cette étude par simulation d'un modèle de mécanismes auditifs physiologiquement et perceptivement plausibles est malheureusement extrêmement longue et fastidieuse. Dans les deux autres études, les estimations de bruit interne réalisées sur les sujets normo-entendants âgés et les sujets malentendants âgés sont cette fois-ci obtenues à partir d'un modèle de décision plus simple, dérivé de la théorie de la détection du signal (Green, 1964). Ce modèle est identique à celui de Neri (2010) et s'applique aux données individuelles d'accord et de performance. Il s'agit d'une technique précédemment adoptée par Burgess et Colborne

(1988) et appliquée à la détection de signaux visuels dans une double-passe. Pour une procédure à deux alternatives en choix forcé, on suppose que la représentation interne préalable à l'addition du bruit interne suit une distribution normale pour le signal cible (de moyenne d') et le signal de comparaison (de moyenne nulle). Chaque réponse interne dans ces intervalles respectifs est ensuite additionnée à une source de bruit Gaussienne d'écart-type σ_{int} régénérée au sein de chaque intervalle de présentation. Cette source de variabilité est la seule variation entre des présentations identiques de stimuli, répétées d'une passe à l'autre. Dans chaque essai, le modèle sélectionne le signal associé à la réponse la plus ample conformément à la TDS. Les valeurs d' et σ_{int} sélectionnées sont celles qui minimisent l'erreur entre les valeurs de cohérence et de performance prédites et réelles.

V.2. Principales conclusions de la première étude

Les courbes empiriques reliant la performance (PC) au pourcentage d'accord (PA) sont globalement similaires dans les 24 conditions expérimentales testées sur nos 15 sujets normo-entendants jeunes : la cohérence des réponses augmente avec la performance et le niveau de bruit externe σ_{ext} n'a que peu d'effet sur ces couples de données.

Les simulations sont réalisées à partir des profondeurs et index de modulation moyens au seuil de détection (3 niveaux : $d'=0.5$, 1 et 1.5) pour l'ensemble des sujets testés. Ainsi, les niveaux de pourcentage d'accord et de performance simulés sont représentatifs d'un participant « moyen » (au seuil de détection moyenné sur chaque condition expérimentale testée) dans les 24 conditions expérimentales testées. Comme indiqué plus haut, pour chaque condition expérimentale, le modèle est simulé pour une gamme de valeurs du bruit interne. La valeur de σ_{int} retenue pour chaque condition expérimentale est celle minimisant les erreurs entre le couple de données (PA ; PC) réelles et simulées.

Les résultats de cette première étude démontrent que :

- 1) Le modèle du banc de filtres de modulation simule correctement les données d'accord et de performance dans chaque condition expérimentale pour certaines valeurs de bruit interne Gaussien additif σ_{int} . Cette capacité du modèle n'avait jamais été évaluée auparavant.
- 2) Pour un niveau de bruit externe donné, σ_{ext} , une même valeur de bruit interne additif, σ_{int} , permet de rendre compte des données d'AM aux deux cadences de modulation testées (2 et 20 Hz). Le modèle du banc de filtres de modulation – tel qu'implémenté ici - fournit donc une *approche unifiée* rendant compte des données de *détection d'AM lente et rapide*.

- 3) Dans tous les cas, le niveau de bruit interne additif σ_{int} conduisant aux meilleurs ajustements augmente avec l'écart-type du bruit externe σ_{ext} appliqué aux modulations. L'effet du bruit externe σ_{ext} sur la qualité des ajustements des données empiriques aux réponses du modèle suggère que le bruit interne des sujets réels présente une *composante multiplicative*, conformément aux conclusions réalisées par Ewert et Dau (2004) avec un paradigme différent.
- 4) Le modèle du banc de filtres de modulation est capable de prédire les données d'accord et de performance en détection de FM lente et rapide, mais ce pour des valeurs de bruit interne additif σ_{int} plus faibles qu'en AM. En effet, le niveau de bruit interne (considéré comme un bruit équivalent dans le domaine de l'enveloppe temporelle) permettant de simuler au mieux les données empiriques en détection de FM lente et rapide est réduit d'un facteur 1.5-5 par rapport à celui simulant au mieux les données empiriques en détection d'AM lente et rapide.
- 5) Par ailleurs, des valeurs de bruit interne additif σ_{int} différentes sont nécessaires pour rendre compte des données d'accord et de performance en détection de FM lente (2 Hz) et rapide (20 Hz). Pris ensemble, les résultats (4) et (5) suggèrent que *des mécanismes neuronaux distincts de ceux implémentés par le modèle du banc de filtres de modulation sont donc à l'œuvre en détection de FM lente et rapide lorsque les modulations sont bruitées (i.e., lorsque le masque de modulation est stochastique)*.
- 6) Le modèle du banc de filtres de modulation ne fournit donc pas une approche unifiée rendant compte des données de détection d'AM et de FM (des valeurs très différentes de bruit interne additif σ_{int} étant requises).

L'ensemble de ces résultats corrobore l'idée selon laquelle les indices d'AM lents et rapides sont détectés par un mécanisme unique (présentant une source de variabilité interne propre, additive et multiplicative) utilisant un banc de filtres sélectifs dans le domaine de l'enveloppe temporelle et un mécanisme de décision par appariement de gabarit.

Par contre, ces résultats ne sont pas compatibles avec l'idée selon laquelle les indices de FM lents et rapides sont détectés par un simple mécanisme spatio-temporel (tonotopique) basé sur la conversion des indices de FM en AM. D'autres mécanismes et d'autres sources de variabilité interne doivent contribuer à la détection de FM quelle que soit la cadence de modulation.

V.3. Principales conclusions de la deuxième étude

Dans cette étude, l'effet de l'âge sur le bruit interne en détection d'AM et de FM lente et rapide est évalué en étendant ce même paradigme à un groupe de 15 sujets normo-entendants âgés entre 40 et 62 ans. Les données empiriques ont été estimées dans les mêmes conditions qu'auparavant. La seule différence entre la première et la deuxième étude porte sur la méthode d'estimation du bruit interne, qui cette fois, est évalué à partir d'un modèle mathématique de décision basé sur la théorie de la détection du signal (Green, 1964).

Afin d'obtenir la résolution minimale requise pour l'estimation du bruit interne (Hasan et al., 2012), les données individuelles d'accord et de performance sont moyennées sur les trois niveaux de sensibilité cibles (0.5, 1, 1.5). Le bruit interne est estimé pour chaque type de modulation (AM/FM) à partir de la répétition de 300 essais (par passe) puis comparé entre groupes et conditions.

Les résultats de cette étude démontrent que :

- 1) La sensibilité à l'AM masquée, préalablement évaluée dans une tâche adaptative précédant la mesure en double-passe, n'est pas affectée par le vieillissement et ce quelle que soit la cadence de modulation.
- 2) La sensibilité à la FM n'est pas non plus affectée par le vieillissement et ce quelle que soit la cadence de modulation.
- 3) La sensibilité à l'AM et à la FM se dégradent à mesure que l'écart-type du bruit externe σ_{ext} augmente. Cette dégradation ne dépend pas de l'âge. Pris ensemble, ces résultats suggèrent que l'effet de masquage produit par le bruit de modulation (un masque d'AM ou de FM stochastique) n'augmente pas avec l'âge.
- 4) Les estimations de bruit interne σ_{int} sont comprises entre 1,8 et 2,11 unités de bruit externe, en accord avec les valeurs mesurées dans d'autres études et pour d'autres tâches auditives (Neri, 2010 ; Ponsot et al., 2021). Ce bruit interne σ_{int} s'avère élevé (deux fois plus fort que le bruit externe).
- 5) En détection d'AM et quelle que soit la cadence, le bruit interne σ_{int} n'est pas affecté par le vieillissement.
- 6) En détection de FM et quelle que soit la cadence, le bruit interne σ_{int} n'est pas affecté par le vieillissement.

Pris ensemble, ces résultats ne confirment pas l'hypothèse selon laquelle le vieillissement et la synaptopathie cochléaire qui lui est associée conduisent à une augmentation de la variabilité neuronale dans le système auditif central au sein des structures dédiées au traitement de l' ENV_N et de TFS_N .

Avant 60 ans, le niveau du bruit interne σ_{int} en détection d'AM et de FM n'évolue donc pas, ou tout au moins d'une manière suffisante pour être détectée par le paradigme de double-passe.

Autre résultat notable de cette étude, quel que soit l'âge des sujets, le niveau du bruit interne σ_{int} en détection d'AM et de FM est élevé (deux fois plus fort que le bruit externe pour la tâche en question).

V.4. Principales conclusions de la troisième étude

Dans cette dernière étude, l'objectif est de tester l'hypothèse selon laquelle une perte auditive d'origine neurosensorielle associée au vieillissement entraîne une augmentation *substantielle* du bruit interne en détection d'AM et de FM, conformément aux résultats de simulation de Wallaert et al., (2017, 2018). Les données empiriques ont été obtenues sur 19 sujets normo-entendants jeunes âgés entre 18 et 30 ans, et 10 sujets malentendants âgés entre 45 et 80 ans. Cette fois-ci, la détection d'AM et de FM est estimée uniquement pour une cadence lente (2 Hz) et pour un niveau de bruit externe bas (0.07 unités). Point important, les mesures sont réalisées au même niveau SPL sur tous les participants. Les estimations de bruit interne sont obtenues à l'aide du modèle mathématique de décision dérivé de la TDS utilisé dans la deuxième étude.

Les résultats de cette étude démontrent que :

- 1) La sensibilité à l'AM lente, préalablement évaluée dans une tâche adaptative précédant la mesure en double-passe, n'est pas significativement altérée par la perte auditive neurosensorielle et le vieillissement. Ce résultat est conforme aux études antérieures réalisant ces mesures sur des participants normo-entendants et malentendants testés à même niveau SPL.
- 2) La sensibilité à la FM lente est significativement dégradée sous l'effet de la perte auditive neurosensorielle et du vieillissement.
- 3) Les estimations de bruit interne σ_{int} sont élevées car comprises entre 1.8 et 2.1 unités de bruit externe, en accord avec les valeurs de bruit interne rapportées dans des études antérieures (Neri, 2010 ; Ponsot et al., 2021).
- 4) Les estimations de bruit interne σ_{int} ne sont pas significativement différentes entre les deux groupes de participants, que ce soit en détection d'AM lente ou de FM lente.

Contrairement aux conclusions des études de modélisation menées par Wallaert et al. (2017, 2018) et d'Ewert et al. (2020), la perte auditive neurosensorielle associée au vieillissement ne semble donc pas affecter le niveau du bruit interne en détection d'AM et de FM lente, ou tout au moins d'une manière suffisante pour être détectée par le paradigme de double-passe.

La synaptopathie cochléaire supposée être associée au vieillissement et à des lésions cochléaires ne semble donc pas engendrer d'augmentation notable de la variabilité neuronale dans les structures auditives centrales impliquées dans le codage temporel.

Enfin, cette dernière étude confirme le fait que, quel que soit le statut et l'âge du sujet, le niveau du bruit interne σ_{int} en détection d'AM et de FM est élevé (deux fois plus fort que le bruit externe pour la tâche en question).

V.5. Discussion

Bruit interne en détection d'AM et de FM. La première étude de cette thèse démontre qu'en détection d'AM, il est possible de rendre compte des seuils masqués et de la cohérence des réponses à l'aide d'un même modèle de mécanismes de traitement temporel n'utilisant qu'une seule source de bruit interne Gaussien et additif. Un bruit multiplicatif doit toutefois être pris en compte afin de pouvoir reproduire les effets de la variabilité externe sur la cohérence des réponses. Nous avons tenté d'implémenter un tel bruit multiplicatif (un bruit dont la variance est proportionnelle à la puissance de modulation estimée à la sortie des filtres de modulation ; cf. Wallaert et al., 2017) mais les premières simulations ne sont pas probantes. Une étude de modélisation plus poussée s'impose. Cette première étude indique par ailleurs qu'en détection de FM, il n'est pas possible de rendre compte des seuils de détection masqués et de la cohérence des réponses avec ce même modèle de traitement temporel, suggérant que la détection de FM - lente comme rapide - résulte de l'opération de mécanismes distincts de ceux opérant en détection d'AM. Ces mécanismes pourraient utiliser des informations de TFS_N . Alternativement, ils pourraient utiliser les informations d' ENV_N évoquées par la FM dans des filtres cochléaires situés de part et d'autre de région encodant la fréquence porteuse des stimuli, mais ne peuvent être modélisés par la simple opération d'un banc de filtres de modulation. Une machinerie neuronale distincte de celle encodant l'AM est mise en œuvre pour la détection de FM. Des *détecteurs de changement fréquentiel* (frequency-shift detectors, FSD ; Semal & Demany, 2006 ; Demany & Semal, 2018) opérant automatiquement sur la sortie des filtres cochléaires pourraient être responsables de nos capacités de détection de FM lente. La durée de la fenêtre d'intégration de cette opération de détection de changement fréquentiel pourrait expliquer l'effet de la cadence sur la détection de FM. Dans tous les cas, cet autre mécanisme doit avoir une source de bruit interne distincte de celle contraignant la détection d'AM.

Point remarquable, il est possible de reproduire les estimations de seuils de détection de FM non masqués et ceux masqués par un masque déterministe (une AM sinusoïdale) avec notre modèle de banc de filtres de modulation lorsque la cadence de FM est rapide (cf. Wallaert et al., 2018 ; King et al., 2019). L'usage d'un bruit de modulation externe est donc à l'origine de l'incapacité du modèle à banc de filtres de modulation à reproduire ces seuils et ces mesures de cohérence. En d'autres termes, le modèle reproduit bien le phénomène de conversion de FM en AM à la sortie des filtres cochléaires, et il est bel et bien capable d'utiliser ces informations pour détecter correctement une FM rapide *en l'absence de stochasticité externe*. Le bruit de modulation externe utilisé dans le présent paradigme altère donc fortement cette capacité. Le/les mécanisme(s) impliqué(s) dans la détection de FM - qu'elle soit lente ou rapide - sont donc résistants vis-à-vis de la variabilité externe dans le domaine des modulations.

Effets du vieillissement. Pourquoi n'observons-nous aucun effet de l'âge sur le bruit interne en détection d'AM et de FM ? Tout d'abord, notons que nos mesures préliminaires de seuils de détection masqués ne diffèrent pas entre sujets normo-entendants jeunes et âgés. Ces résultats sont intéressants à deux titres. Tout d'abord, parce qu'ils démontrent que les effets de masquage dans le domaine de la modulation (AM comme FM) ne sont pas exacerbés par le vieillissement. Dans nos trois études, les effets de masquage sont produits par un bruit de modulation, à savoir des fluctuations aléatoires relativement lentes dans le domaine de l'AM et de la FM. Ni la puissance du masque, ni son caractère stochastique ne semblent donc produire de déficit spécifique chez le sujet âgé, suggérant que les mécanismes de codage central (e.g., la sélectivité des filtres de modulation, le mécanisme de décision) sont intacts chez les sujets âgés. Enfin, l'observation de seuils de détection similaires chez les sujets normo-entendants jeunes et âgés est tout à fait compatible avec l'absence de différence de cohérence entre les groupes dans l'étude en double-passe. Ces résultats n'étaient pas attendus au vu de la littérature (e.g., Wallaert et al., 2016, 2017, 2018). Par ailleurs, l'absence d'élévation de seuil chez les sujets âgés - même normo-entendants - ne signifie pas pour autant que le bruit interne est comparable entre les deux groupes. En effet, chez les sujets les plus âgés, une perte auditive légère (« subclinique ») causée par la perte de cellules ciliées externes pourrait engendrer une perte de compression d'amplitude périphérique, à savoir un phénomène de recrutement de sonie engendrant une magnification perceptive modeste des indices d' ENV_N évoqués par l'AM et la FM à la sortie des filtres cochléaires, et par conséquent, une amélioration de sensibilité aux modulations. Cette légère amélioration est peut-être compensée par une augmentation parallèle de la variabilité neuronale interne, résultant d'une synaptopathie cochléaire. Nous avons inspecté l'écart-type des estimations de seuil réalisées sur les dernières inversions des escaliers psychophysiques sachant que cette mesure est reliée à la pente de la fonction psychométrique sous-jacente, mais nous n'avons pas trouvé de différence d'écart-type entre groupes jeunes et âgés.

Nous pourrions être tentés de conclure que ces effets délétères ne se manifestent qu'au-delà de 60 ans (l'âge du participant normo-entendant le plus âgé de la deuxième étude). Toutefois,

notre dernière étude ayant inclus des sujets malentendants âgés entre 45 et 80 ans, cette interprétation ne semble pas la plus pertinente. Notre revue de littérature révèle que les effets du vieillissement en détection de modulation ne sont pas systématiquement observés. Ceci conforte donc l'idée selon laquelle les effets délétères du vieillissement en détection d'AM et de FM, si présents, sont probablement *modestes et dépendants du sujet* (e.g., de facteurs environnementaux et héréditaires). Si la synaptopathie cochléaire associée au vieillissement est bien réelle compte tenu des observations post-mortem réalisées sur l'humain, la perte d'un contingent important de fibres afférentes de type I ne semble pas suffisante pour modifier la variabilité au niveau comportemental. Il reste possible que la variabilité neuronale centrale s'accroisse modestement à mesure que s'accroît la perte neuronale, mais que le paradigme de double-passe n'ait pas la sensibilité suffisante pour distinguer un groupe de sujets jeunes d'un groupe de sujets âgés. Nous reviendrons plus bas sur les limites de notre paradigme psychophysique.

Effets d'une perte auditive neurosensorielle. Pourquoi n'observons-nous aucun effet de la perte auditive sur le bruit interne en détection d'AM et de FM ? Notons à nouveau que nos mesures préliminaires de seuils de détection d'AM masqués ne diffèrent pas entre sujets normo-entendants jeunes et sujets malentendants âgés. Ceci était anticipable sachant que les deux groupes étaient testés au même niveau SPL, suggérant qu'une perte cochléaire n'exacerbe pas les effets de masquage de modulation. Le même raisonnement que celui tenu plus haut s'applique alors aux sujets malentendants : peut-être présentent-ils un recrutement de sonie – cette fois, important au vu de la perte neurosensorielle – compensé par un bruit interne accru résultant d'une synaptopathie cochléaire. L'inspection de l'écart-type des estimations de seuil réalisées sur les dernières inversions des escaliers psychophysiques ne révèle aucune différence entre groupes. L'observation de seuils de détection similaires chez les sujets normo-entendants jeunes et les sujets malentendants âgés est donc tout à fait compatible avec l'absence de différence de cohérence entre les groupes dans l'étude en double-passe. A nouveau, ces résultats n'étaient pas attendus au vu des études de modélisation réalisées dans des conditions expérimentales comparables (e.g., Wallaert et al., 2017, 2018) – à l'exception du bruit de modulation externe. Ceci suggère que l'ajustement des réponses du modèle du banc de filtres de modulation aux données mesurées par Wallaert et al. (2017, 2018) par simple linéarisation de la fonction entrée-sortie du module de transduction (i.e., élimination de la compression périphérique) et augmentation du bruit interne n'était pas réaliste. Il est probable que les sujets malentendants de l'étude de Wallaert et al. (2017, 2018) conservaient un certain degré de compression cochléaire.

Notons maintenant que nos mesures préliminaires de seuils de détection de FM masqués sont plus élevées chez les sujets malentendants âgés que chez les sujets normo-entendants jeunes. Ceci est tout à fait compatible avec la littérature portant sur les effets de lésions cochléaires sur la détection de FM. Malgré cela, l'inspection de l'écart-type des estimations de seuil réalisées sur les dernières inversions des escaliers psychophysiques ne révèle aucune différence entre groupes, conformément à l'absence de différence de cohérence entre les

groupes dans l'étude en double-passe. Ainsi, la perte d'un contingent important de fibres afférentes de type I susceptible d'être associée à la perte cochléaire combinée au vieillissement ne semble pas suffisante pour modifier la variabilité au niveau comportemental. A nouveau, il reste possible que la variabilité neuronale centrale s'accroisse modestement à mesure que s'accroît la perte neuronale, mais que le paradigme de double-passe n'ait pas la sensibilité suffisante pour distinguer un groupe de sujets jeunes d'un groupe de sujets âgés. Ce point est discuté plus bas.

Méthodes d'estimation du bruit interne. Dans nos trois études, les estimations du bruit interne à partir des données empiriques de cohérence et de performance ont été obtenues soit à l'aide d'un modèle à banc de filtres de modulation (étude I), soit à l'aide d'un modèle de décision basé sur la théorie de la détection du signal (études II et III). Ces modèles s'opposent :

- i) par le **nombre et la spécificité des étapes de traitement**. En effet, le modèle à banc de filtres de modulation inclue plusieurs étapes de traitement spécifiques - linéaires et non-linéaires - avant la prise de décision, ce qui n'est pas le cas du modèle dérivé de la théorie de la détection du signal qui n'opère que sur la variable de décision;
- ii) par **l'étape d'ajout du bruit interne**. En effet, le modèle à banc de filtres de modulation intègre un bruit interne additif en sortie des filtres de modulation, appliqué juste *avant* l'étape de prise de décision par appariement de gabarit. Le modèle dérivé de la théorie de la détection du signal - quant à lui - applique un bruit interne additif à la variable de décision.

Au vu de cela, ces deux types de bruit interne sont-ils vraiment comparables ?

Notons tout d'abord que le bruit interne du modèle à banc de filtres de modulation est de nature purement sensorielle. Il simule toutes les sources de variabilité neuronale situées entre la cochlée et les centres nerveux réalisant la décomposition de l'enveloppe temporelle du signal. Cette variabilité influence toutefois la variable de décision du modèle, correspondant à la valeur de la corrélation entre le gabarit stocké en mémoire et la sortie bruitée des filtres de modulation. La variance du bruit interne extraite par la méthode dérivée de la théorie de la détection du signal - quant à elle - part de l'hypothèse que cette source de variabilité s'ajoute *directement* à la statistique de décision. Ce bruit interne purement « décisionnel » prend néanmoins et implicitement en compte des éléments de stochasticité du codage sensoriel. Il ne nous apparaît donc pas pertinent de distinguer ces bruits internes du point de vue de l'étape de traitement à laquelle ils sont introduits.

Notons maintenant que la *relation* entre la performance (le pourcentage de détections correctes) et la cohérence (le pourcentage d'accord des réponses) diffère fortement entre les deux approches de modélisation. En effet, du point de vue de la forme et de la pente, les courbes présentées en figure 4 de notre premier article (étude I) sont fortement influencées

par toutes les étapes de traitement du modèle du banc de filtres de modulation. Tout changement de la sélectivité des filtres cochléaires, du niveau de compression d'amplitude ou encore de la sélectivité des filtres de modulation altère substantiellement ces relations PC-PA. Ces dernières ne sont pas prises en compte par le modèle dérivé de la théorie de la détection du signal. Qui plus est, ces relations sont fortement modifiées en détection de FM, expliquant pourquoi le bruit interne permettant de simuler les données d'AM est trop élevé pour reproduire les données en détection de FM. ***Les effets du bruit interne sur la performance et la cohérence des réponses en détection de modulation ne sont donc pas modélisés de la même manière dans les deux approches.***

V.6. Limites

Ce projet de recherche doctorale a permis de mettre en place un protocole mesurant la cohérence des jugements auditifs en détection d'AM et de FM sur trois groupes de sujets : des sujets jeunes et normo-entendants, des sujets âgés et normo-entendants et enfin des sujets âgés et malentendants, afin de préciser la nature du traitement de ces modulations ainsi que les effets du vieillissement et de la perte auditive neurosensorielle sur le bruit interne en détection d'AM et de FM.

Néanmoins, il est important de mentionner certaines limites de notre protocole :

Sur le plan pratique, la résolution minimale requise pour estimer de manière robuste la variance du bruit interne requiert un nombre d'essais conséquent pour chaque condition expérimentale et chaque passe, ceci contraignant fortement la *durée* de passation des tests. La durée relativement importante du protocole complet est susceptible d'engendrer des effets de fatigue ainsi que des variations de vigilance de la part des participants, et ce avant tout chez les participants les plus âgés et malentendants. Ceci pourrait se traduire par une augmentation de la contribution de la composante *cognitive* du bruit interne tel qu'estimé par le paradigme de double-passe - à savoir des fluctuations aléatoires d'attention, des capacités mnésiques ou encore des stratégies d'écoute, minimisant ainsi la contribution - probablement plus faible - de la composante sensorielle du bruit interne. Ceci pourrait avoir limité fortement la possibilité d'observer des différences de bruit interne d'origine purement sensorielle entre sujets jeunes, âgés et malentendants.

Concernant l'inclusion des participants, il est possible que l'étendue et la moyenne d'âge relativement basse (étendue= [45-62 ans] ; moyenne=53 ans) des participants de l'étude II ainsi que le faible nombre de participants au sein du groupe de sujets malentendants âgés (n=10) de l'étude III limitent la puissance statistique de ces études, et expliquent l'absence d'effets du vieillissement et de la perte auditive neurosensorielle sur le bruit interne en détection d'AM et de FM. Ces études devraient donc être répliquées sur des cohortes plus importantes de sujets plus âgés.

Sur le plan méthodologique, les choix expérimentaux ont conduit à observer des valeurs de cohérence et de performance souvent proches de la *courbe théorique limite* au-delà de laquelle les données deviennent aberrantes (Neri et al., 2010). Ces courbes sont représentées ci-dessous avec les données de l'étude III. Cet « écrasement » des données empiriques dans cette région de l'espace PC-PA a certainement limité fortement la possibilité d'observer des différences entre conditions expérimentales et groupes de sujets. Des valeurs plus importantes de bruit externe auraient pu décaler les données empiriques vers des valeurs de PA plus importantes, accentuant la possibilité de séparer conditions et groupes. Toutefois, ceci aurait été associé à une forte élévation des seuils de détection, empêchant probablement les sujets âgés et malentendants de réaliser les tâches de détection d'AM et de FM. Il est donc difficile de trouver un compromis dans le cas présent ...

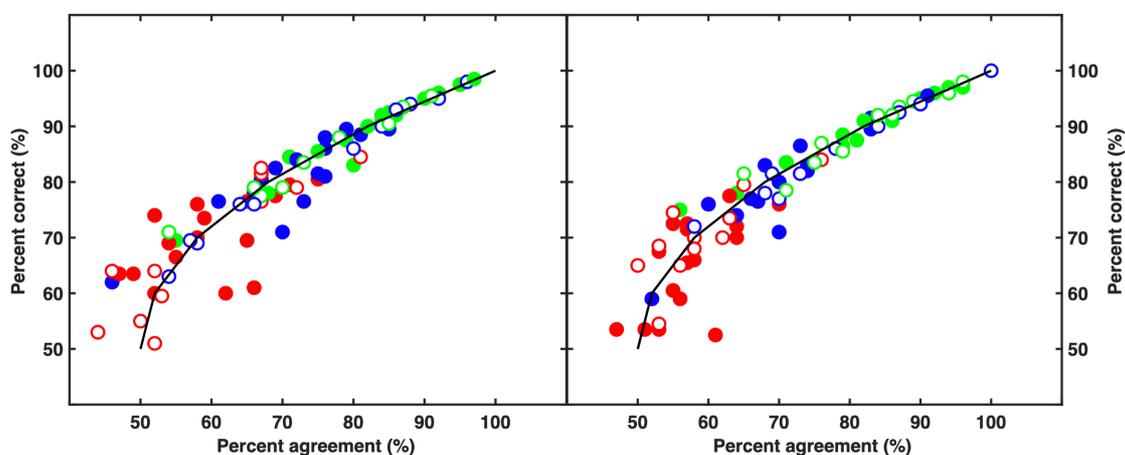


Figure 90. Données individuelles de double-passe par condition expérimentale testée dans l'étude III (symboles fermés : sujets NHy ; symboles ouverts : sujets Hlo). Les graphiques de gauche et de droite présentent respectivement les données en détection d'AM et de FM. Les données sont présentées par niveau de sensibilité ciblée à l'aide de la procédure adaptative initiale (rouge : $d'=0.5$; bleu : $d'=1$; vert : $d'=1.5$). Les courbes noires représentent les courbes théoriques limites définies à l'aide du modèle de la théorie de la détection du signal identique à Neri et al. (2010) délimitant les estimations de bruit interne cohérentes.

Plusieurs limites peuvent être mentionnées concernant le paradigme de double-passe et son implémentation : i) tout d'abord, l'estimation du bruit interne à l'aide d'un modèle dérivé de la théorie de la détection du signal ne permet pas de faire la distinction entre un bruit interne de nature additive et un bruit interne de nature multiplicative. Ceci aurait pu être possible en combinant ce paradigme à un deuxième paradigme (par exemple, celui du bruit équivalent) en détection d'AM et de FM ; ii) ce paradigme ne permet pas de localiser le bruit interne dans les voies auditives. Des mesures électrophysiologiques non invasives combinées aux mesures comportementales sont nécessaires pour préciser ce point. Les mesures d'ASSR en réponse à un stimulus modulé réalisées par Gransier et al. (2021) nous semblent prometteuses à cet égard. Enfin, iii) les tâches de détection d'AM et de FM (toutes conditions expérimentales

confondues) ont été réparties sur différentes sessions de test. D'autres configurations auraient pu être testées.

V.7. Applications cliniques

Si les études tentant de mesurer les effets du bruit interne sur le traitement des modulations temporelles sont rares, quelques études récentes visant à modéliser la détection et la discrimination d'AM chez les sujets normo-entendants âgés et les patients malentendants suggèrent que la variance du bruit interne additif s'accroît considérablement (jusqu'à un facteur 6) sous l'effet du vieillissement ou de lésions cochléaires (Derleth et al., 2001 ; Ives et al., 2014 ; Paraouty et al., 2016 ; Wallaert et al., 2007). Nous avons fait l'hypothèse que cette augmentation de bruit interne se traduit par une diminution de la cohérence des réponses des personnes âgées et malentendantes. Nous n'avons pas observé d'effets significatifs du vieillissement et de la perte auditive neurosensorielle sur nos estimations de bruit interne en détection d'AM et de FM, mais une implémentation de notre protocole prenant en compte les limites citées ci-dessus aurait peut-être permis de révéler de tels effets, même modestes. Notons par ailleurs que plusieurs études psychoacoustiques portant sur l'audition binaurale (*e.g.*, Bernstein & Trahiotis, 2008) suggèrent que le bruit interne augmente sous l'effet d'une synaptopathie cochléaire. De plus, des études de modélisation portant sur la perception de la parole démontrent qu'il est nécessaire de prendre en compte le bruit interne pour simuler correctement les effets de la perte auditive sur les seuils d'intelligibilité dans le bruit des auditeurs malentendants (*e.g.*, Vicente et al., 2020). Pour toutes ces raisons, nous continuons à supposer que ***l'estimation comportementale de la variance du bruit interne couplée à des mesures électrophysiologiques non-invasives de variabilité neuronale pourrait présenter un intérêt clinique*** en complétant les mesures de performance et de sensibilité traditionnellement utilisées en audiologie pour caractériser les pertes auditives d'origine neurosensorielle. Les études récentes sur la perception auditive et visuelle discutées dans le chapitre IV indiquent que l'estimation du bruit interne présente un véritable intérêt clinique et fournit un outil robuste pour le diagnostic de certaines pathologies sensorielles et cognitives et le suivi du développement perceptif de l'enfant. En modalité auditive, l'estimation du bruit interne pourrait notamment compléter les mesures supraliminaires actuelles visant à diagnostiquer une neuropathie auditive.

Nous concluons malgré tout sur une note relativement négative. En effet, le fait qu'un bruit interne de nature multiplicative contraigne le traitement auditif des informations d'AM est en soi une mauvaise nouvelle, fixant une **limite fondamentale à certaines méthodes de réhabilitation des troubles auditifs par prothèse ou implant** : *si les effets d'un bruit interne additif peuvent être a priori compensés par une stratégie de rehaussement des indices acoustiques d'AM comme cela est proposé par Koning et Wouters (2012), cette approche ne peut être bénéfique dans le cas d'un bruit multiplicatif dont la variance augmente avec la puissance des indices d'AM.* Ceci incite à concentrer les efforts de développement sur le débruitage du signal acoustique (e.g., grâce à des microphones directionnels) plutôt que sur le rehaussement des indices acoustiques pertinents par un traitement de signal spécifique.

Références

- Abrams, D. A., Nicol, T., Zecker, S. G., & Kraus, N. (2006). Auditory brainstem timing predicts cerebral asymmetry for speech. *Journal of Neuroscience*, 26(43), 11131-11137.
- Ahumada, A. J. (1967). Detection of tones masked by noise; a comparison of human observers with digital-computer-simulated energy detectors of varying bandwidths (Doctoral dissertation, University of California, Los Angeles).
- Ahumada, A.J., & Lovell, J. (1971). Stimulus Features in Signal Detection. *The Journal of the Acoustical Society of America*, 49(6B), 1751-1756. <https://doi.org/10.1121/1.1912577>
- Ahumada, A. J., & Watson, A. B. (1985). Equivalent-noise model for contrast detection and discrimination. *JOSA A*, 2(7), 1133-1139.
- Allen, J. B. (1989). Is Basilar Membrane Tuning the Same as Neural Tuning--Where Do We Stand?. In *Cochlear Mechanisms: Structure, Function, and Models* (pp. 453-463). Springer, Boston, MA.
- Allen, P., & Wightman, F. (1994). Psychometric functions for children's detection of tones in noise. *Journal of Speech, Language, and Hearing Research*, 37(1), 205-215.
- Allen, P., & Wightman, F. (1995). Effects of signal and masker uncertainty on children's detection. *Journal of Speech, Language, and Hearing Research*, 38(2), 503-511.
- Allen, P., & Nelles, J. (1996). Development of auditory information integration abilities. *The Journal of the Acoustical Society of America*, 100(2), 1043-1051.
- American National Standards Institute (1996). Specifications for audiometers (ANSI S3.6-1996). New York: Author.
- American National Standards Institute Terminology, A. P. (1973). S3. 20. *New York: American National Standards Institute*.
- Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010). Neural timing is linked to speech perception in noise. *Journal of Neuroscience*, 30(14), 4922-4926.
- Andrews, M. W., & Dowling, W. J. (1991). The Development of Perception of Interleaved Melodies and Control of Auditory Attention. *Music Perception*, 8(4), 349-368. <https://doi.org/10.2307/40285518>
- Arai, T., & Greenberg, S. (1997). The temporal properties of spoken Japanese are similar to those of English. In *Fifth european conference on speech communication and technology*.
- Ardoint, M., Lorenzi, C., Pressnitzer, D., & Gorea, A. (2008). Investigation of perceptual constancy in the temporal-envelope domain. *The Journal of the Acoustical Society of America*, 123(3), 1591-1601. <https://doi.org/10.1121/1.2836782>
- Artieda, J., Pastor, M. A., Lacruz, F., & Obeso, J. A. (1992). Temporal discrimination is abnormal in Parkinson's disease. *Brain*, 115(1), 199-210. <https://doi.org/10.1093/brain/115.1.199>

- Attias, H., & Schreiner, C. E. (1997). Coding of Naturalistic Stimuli by Auditory Midbrain Neurons. 7.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, 61(3), 183-193. <https://doi.org/10.1037/h0054663>
- Bacon, S. P., & Gleitman, R. M. (1992). Modulation Detection in Subjects With Relatively Flat Hearing Losses. *Journal of Speech, Language, and Hearing Research*, 35(3), 642-653. <https://doi.org/10.1044/jshr.3503.642>
- Bacon, S. P., & Grantham, D. W. (1989). Modulation masking : Effects of modulation frequency, depth, and phase. *The Journal of the Acoustical Society of America*, 85(6), 2575-2580. <https://doi.org/10.1121/1.397751>
- Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision research*.
- Banai, K., Sabin, A. T., & Wright, B. A. (2011). Separable developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. *Hearing Research*, 280(1-2), 219-227. <https://doi.org/10.1016/j.heares.2011.05.019>
- Banton, T., Dobkins, K., & Bertenthal, B. I. (2001). Infant direction discrimination thresholds. *Vision research*, 41(8), 1049-1056.
- Barlow, H. B. (1953). Action potentials from the frog's retina. *The Journal of Physiology*, 119(1), 58-68. <https://doi.org/10.1113/jphysiol.1953.sp004828>
- Barlow, H. B. (1977). Retinal and central factors in human vision limited by noise. *Vertebrate photoreception*, 337, C358.
- Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex differences in the brain: implications for explaining autism. *Science*, 310(5749), 819-823.
- Baumberger, B., & Flückiger, M. (2004). The development of distance estimation in optic flow. *Perception*, 33(9), 1081-1099.
- Benzi, R., Sutera, A., & Vulpiani, A. (1981). The mechanism of stochastic resonance. *Journal of Physics A: mathematical and general*, 14(11), L453.
- Berg, B. G. (1987). Internal noise in auditory decision tasks (Doctoral dissertation, Indiana University).
- Berg, B. G., & Green, D. M. (1990). Spectral weights in profile listening. *The Journal of the Acoustical Society of America*, 88(2), 758-766. <https://doi.org/10.1121/1.399725>
- Berg, B. G. (2004). A molecular description of profile analysis : Decision weights and internal noise). *J. Acoust. Soc. Am.*, 115(2), 8.
- Berger, K. W. (1964). Some Factors in the Recognition of Timbre. *The Journal of the Acoustical Society of America*, 36(10), 1888-1891. <https://doi.org/10.1121/1.1919287>
- Berliner, J. E., & Durlach, N. I. (1973). A Perceptual-Anchor Model for Context Coding in Intensity Perception. *The Journal of the Acoustical Society of America*, 54(1), 336-336.
- Bernstein, J. G., & Oxenham, A. J. (2003). Pitch discrimination of diotic and dichotic tone complexes : Harmonic resolvability or harmonic number? *The Journal of the Acoustical Society of America*, 113(6), 3323. <https://doi.org/10.1121/1.1572146>

- Bernstein, L. R., & Trahiotis, C. (2008). Binaural signal detection, overall masking level, and masker interaural correlation: Revisiting the internal noise hypothesis. *The Journal of the Acoustical Society of America*, 124(6), 3850-3860.
- Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., & Shinn-Cunningham, B. G. (2014). Cochlear neuropathy and the coding of supra-threshold sound. *Frontiers in Systems Neuroscience*, 8. <https://doi.org/10.3389/fnsys.2014.00026>
- Biberger, T., & Ewert, S. D. (2016). Envelope and intensity based prediction of psychoacoustic masking and speech intelligibility. *The Journal of the Acoustical Society of America*, 140(2), 1023-1038. <https://doi.org/10.1121/1.4960574>
- Boettcher, F. A., Poth, E. A., Mills, J. H., & Dubno, J. R. (2001). The amplitude-modulation following response in young and aged human subjects. *Hearing Research*, 153(1-2), 32-42. [https://doi.org/10.1016/S0378-5955\(00\)00255-0](https://doi.org/10.1016/S0378-5955(00)00255-0)
- Bogfjellmo, L.-G., Bex, P. J., & Falkenberg, H. K. (2014). The development of global motion discrimination in school aged children. *Journal of Vision*, 12.
- Bogliotti, C., Serniclaes, W., Messaoud-Galusi, S., & Sprenger-Charolles, L. (2008). Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls. *Journal of experimental child psychology*, 101(2), 137-155.
- Bonte, M. L., Poelmans, H., & Blomert, L. (2007). Deviant neurophysiological responses to phonological regularities in speech in dyslexic children. *Neuropsychologia*, 45(7), 1427-1437.
- Bos, C. E., & de Boer, E. (1966). Masking and Discrimination. *The Journal of the Acoustical Society of America*, 39(4), 708-715. <https://doi.org/10.1121/1.1909945>
- Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities.
- Burgess, A., Wagner, R., Jennings, R., & Barlow, H. (1981). Efficiency of human visual signal discrimination. *Science*, 214(4516), 93-94. <https://doi.org/10.1126/science.7280685>
- Burgess, A., & Barlow, H. B. (1983). The precision of numerosity discrimination in arrays of random dots. *Vision Research*, 23(8), 811-820. [https://doi.org/10.1016/0042-6989\(83\)90204-3](https://doi.org/10.1016/0042-6989(83)90204-3)
- Burgess, A. E., & Colborne, B. (1988). Visual signal detection. IV. Observer inconsistency. *JOSA A*, 5(4), 617-627.
- Burton, K. B., Owsley, C., & Sloane, M. E. (1993). Aging and neural spatial contrast sensitivity: photopic vision. *Vision research*, 33(7), 939-946.
- Buus, S. R., & Florentine, M. (1991). Psychometric functions for level discrimination. *The Journal of the Acoustical Society of America*, 90(3), 1371-1380.
- Buss, E., Hall, J. W., & Grose, J. H. (2004a). Temporal Fine-Structure Cues to Speech and Pure Tone Modulation in Observers with Sensorineural Hearing Loss: Ear and Hearing, 25(3), 242-250. <https://doi.org/10.1097/01.AUD.0000130796.73809.09>

- Buss, E., Hall, J. W., & Grose, J. H. (2004b). Temporal Fine-Structure Cues to Speech and Pure Tone Modulation in Observers with Sensorineural Hearing Loss: *Ear and Hearing*, 25(3), 242-250. <https://doi.org/10.1097/01.AUD.0000130796.73809.09>
- Buss, E., Hall, J. W., & Grose, J. H. (2009). Psychometric functions for pure tone intensity discrimination : Slope differences in school-aged children and adults. *The Journal of the Acoustical Society of America*, 125(2), 1050-1058. <https://doi.org/10.1121/1.3050273>
- Buss, E., Iii, J. W. H., & Grose, J. H. (2006). Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. *J. Acoust. Soc. Am.*, 120(5), 13.
- Buss, E., Lorenzi, C., Cabrera, L., Leibold, L. J., & Grose, J. H. (2019). Amplitude modulation detection and modulation masking in school-age children and adults. *The Journal of the Acoustical Society of America*, 145(4), 2565-2575. <https://doi.org/10.1121/1.5098950>
- Cabrera, L., Lorenzi, C., & Bertoincini, J. (2015). Infants Discriminate Voicing and Place of Articulation With Reduced Spectral and Temporal Modulation Cues. *Journal of Speech, Language, and Hearing Research*, 58(3), 1033-1042. https://doi.org/10.1044/2015_JSLHR-H-14-0121
- Cabrera, L., Varnet, L., Buss, E., Rosen, S., & Lorenzi, C. (2019). Development of temporal auditory processing in childhood : Changes in efficiency rather than temporal-modulation selectivity. *The Journal of the Acoustical Society of America*, 146(4), 2415-2429. <https://doi.org/10.1121/1.5128324>
- Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions : A confirmatory study using synthetic tones. *J. Acoust. Soc. Am.*, 118(1), 12.
- Cao, F., Bitan, T., Chou, T. L., Burman, D. D., & Booth, J. R. (2006). Deficient orthographic and phonological representations in children with dyslexia revealed by brain activation patterns. *Journal of child psychology and psychiatry*, 47(10), 1041-1050.
- Carcagno, S., & Plack, C. J. (2021). Effects of age on psychophysical measures of auditory temporal processing and speech reception at low and high levels. *Hearing Research*, 400, 108117. <https://doi.org/10.1016/j.heares.2020.108117>
- Carney, L. H. (1993). A model for the responses of low-frequency auditory-nerve fibers in cat. *The Journal of the Acoustical Society of America*, 93(1), 401-417.
- Caspary, D. M., Palombi, P. S., Backoff, P. M., Helfert, R. H., & Finlayson, P. G. (1993). GABA and glycine inputs control discharge rate within the excitatory response area of primary-like and phase-locked AVCN neurons. In *The Mammalian Cochlear Nuclei* (pp. 239-252). Springer, Boston, MA.
- Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nature Reviews Neuroscience*, 3(8), 617-628.

- Castellanos, F. X., Sonuga-Barke, E. J., Scheres, A., Di Martino, A., Hyde, C., & Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biological psychiatry*, 57(11), 1416-1423.
- Cherkassky, V. L., Kana, R. K., Keller, T. A., & Just, M. A. (2006). Functional connectivity in a baseline resting-state network in autism. *Neuroreport*, 17(16), 1687-1690.
- Chung, S. T. L., Levi, D. M., & Tjan, B. S. (2005). Learning letter identification in peripheral vision. *Vision Research*, 45(11), 1399-1412. <https://doi.org/10.1016/j.visres.2004.11.021>
- Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency : Evidence from human frequency-following response recordings. *Hearing Research*, 264(1-2), 48-55. <https://doi.org/10.1016/j.heares.2009.11.010>
- Colburn, H. S. (1973). Theory of binaural interaction based on auditory-nerve data: I. General strategy and preliminary results on interaural discrimination. *Journal of the Acoustical Society of America*, 54(6), 1458–1470.
- Colburn, H. S. (1977). Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. *The Journal of the Acoustical Society of America*, 61(2), 525-533.
- Collins, J. J., Chow, C. C., Capela, A. C., & Imhoff, T. T. (1996). Aperiodic stochastic resonance. *Physical Review E*, 54(5), 5575-5584. <https://doi.org/10.1103/PhysRevE.54.5575>
- Committee on Hearing, Bioacoustics, and Biomechanics (CHABA) (1988) Speech understanding and aging. *J Acoust Soc Am* 83:859–895.
- Coninx, P. (1978). The detection of combined differences in frequency and intensity. *Acta Acustica united with Acustica*, 39(3), 137-150.
- Corbett, B. A., & Constantine, L. J. (2006). Autism and attention deficit hyperactivity disorder: Assessing attention and response control with the integrated visual and auditory continuous performance test. *Child Neuropsychology*, 12(4-5), 335-348.
- Dakin, S., & Frith, U. (2005). Vagaries of visual perception in autism. *Neuron*, 48(3), 497-507.
- Dai, H. (1994). Signal-frequency uncertainty in spectral-shape discrimination: Psychometric functions. *The Journal of the Acoustical Society of America*, 96(3), 1388-1396.
- Daikoku, T., & Goswami, U. (2020). The Hierarchical Structure of Temporal Modulations in Music is Universal across Genres and matches Infant-Directed Speech. 36.
- Dakin, S. C., Mareschal, I., & Bex, P. J. (2005). Local and global limitations on direction integration assessed using equivalent noise analysis. *Vision research*, 45(24), 3027-3049.
- Dau, T., Püschel, D., & Kohlrausch, A. (1996). A quantitative model of the “effective” signal processing in the auditory system. II. Simulations and measurements. *The Journal of the Acoustical Society of America*, 99(6), 3623-3631. <https://doi.org/10.1121/1.414960>

- Dau, T., Kollmeier, B., & Kohlrausch, A. (1997). Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. *The Journal of the Acoustical Society of America*, 102(5), 2906-2919.
- Dau, T., Kollmeier, B., & Kohlrausch, A. (1997a). Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *The Journal of the Acoustical Society of America*, 102(5), 2892-2905.
- Dau, T., Kollmeier, B., & Kohlrausch, A. (1997b). Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. *The Journal of the Acoustical Society of America*, 102(5), 2906-2919.
- Dawes, P., & Bishop, D. V. M. (2008). Maturation of Visual and Auditory Temporal Processing in School-Aged Children. *Journal of Speech, Language, and Hearing Research*, 51(4), 1002-1015. [https://doi.org/10.1044/1092-4388\(2008/073\)](https://doi.org/10.1044/1092-4388(2008/073))
- de Boer, E. (1966). Intensity Discrimination of Fluctuating Signals. *The Journal of the Acoustical Society of America*, 40(3), 552-560. <https://doi.org/10.1121/1.1910118>
- Demany, L., and Semal, C. (1986). "On the detection of amplitude modulation and frequency modulation at low modulation frequencies". *Acta Acust. united Ac.*, 61, 243–255.
- Demany, L., & Semal, C. (2018). Automatic frequency-shift detection in the auditory system: A review of psychophysical findings. *Neuroscience*, 389, 30-40.
- Denk, W., Webb, W. W., & Hudspeth, A. J. (1989). Mechanical properties of sensory hair bundles are reflected in their Brownian motion measured with a laser differential interferometer. *Proceedings of the National Academy of Sciences*, 86(14), 5371-5375. <https://doi.org/10.1073/pnas.86.14.5371>
- Derleth, R. P., Dau, T., & Kollmeier, B. (2001). Modeling temporal and compressive properties of the normal and impaired auditory system. *Hearing Research*, 159(1-2), 132-149.
- Destexhe, A., Rudolph, M., Fellous, J.-M., & Sejnowski, T. J. (2001). Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons. *Neuroscience*, 107(1), 13-24. [https://doi.org/10.1016/S0306-4522\(01\)00344-X](https://doi.org/10.1016/S0306-4522(01)00344-X)
- Deutsch, K. M., & Newell, K. M. (2004). Changes in the structure of children's isometric force variability with practice. *Journal of experimental child psychology*, 88(4), 319-333.
- Devaney, K. O., & Johnson, H. A. (1980). Neuron loss in the aging visual cortex of man. *Journal of Gerontology*, 35(6), 836-841.
- Diependaele, K., Brysbaert, M., & Neri, P. (2012). How noisy is lexical decision?. *Frontiers in psychology*, 3, 348.
- Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, 81, 181-187. <https://doi.org/10.1016/j.neubiorev.2017.02.011>
- Dinstein, I., Heeger, D. J., Lorenzi, L., Minshew, N. J., Malach, R., & Behrmann, M. (2012). Unreliable Evoked Responses in Autism. *Neuron*, 75(6), 981-991. <https://doi.org/10.1016/j.neuron.2012.07.026>

- Divenyi, P. L., & Haupt, K. M. (1997). Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. I. Age and lateral asymmetry effects. *Ear and hearing*, 18(1), 42-61.
- Dosher, B. A., & Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences*, 95(23), 13988-13993.
- Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological review*, 85(4), 341.
- Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology: A Journal of Research in Music Cognition*, 1(1), 30-49. <https://doi.org/10.1037/h0094275>
- Drullman, R. (1995). Temporal Envelope and Fine Structure Cues for Speech Intelligibility. 8.
- Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America*, 95(2), 1053-1064. <https://doi.org/10.1121/1.408467>
- Dubno, J. R., Lee, F.-S., Matthews, L. J., Ahlstrom, J. B., Horwitz, A. R., & Mills, J. H. (2008). Longitudinal changes in speech recognition in older persons. *The Journal of the Acoustical Society of America*, 123(1), 462-475. <https://doi.org/10.1121/1.2817362>
- Dudley, H., Riesz, R. R., & Watkins, S. S. (1939). A synthetic speaker. *Journal of the Franklin Institute*, 227(6), 739-764.
- Duquesnoy, A. J. (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *The Journal of the Acoustical Society of America*, 74(3), 739-743. <https://doi.org/10.1121/1.389859>
- Durlach, N. I., & Braida, L. D. (1969). Intensity perception. I. Preliminary theory of intensity resolution. *The Journal of the Acoustical Society of America*, 46(2B), 372-383.
- D'Zmura, M., & Knoblauch, K. (1998). Spectral bandwidths for the detection of color. *Vision Research*, 38, 3117-3128.
- Eggermont, J. J. (1994). Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hearing Research*, 74(1-2), 51-66. [https://doi.org/10.1016/0378-5955\(94\)90175-9](https://doi.org/10.1016/0378-5955(94)90175-9)
- Eijkman, E., Thijssen, J. M., & Vendrik, A. J. H. (1966). Weber's Law, Power Law, and Internal Noise. 10.
- Elleberg, D., Lewis, T. L., Maurer, D., Hong Lui, C., & Brent, H. P. (1999). Spatial and temporal vision in patients treated for bilateral congenital cataracts. *Vision Research*, 39(20), 3480-3489. [https://doi.org/10.1016/S0042-6989\(99\)00078-4](https://doi.org/10.1016/S0042-6989(99)00078-4)
- Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *The Journal of the Acoustical Society of America*, 66(3), 651-653.
- Elliott, T. M., & Theunissen, F. E. (2009). The Modulation Transfer Function for Speech Intelligibility. *PLoS Computational Biology*, 5(3), e1000302. <https://doi.org/10.1371/journal.pcbi.1000302>

- Ernst, S. M. A., & Moore, B. C. J. (2012). The role of time and place cues in the detection of frequency modulation by hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 131(6), 4722-4731. <https://doi.org/10.1121/1.3699233>
- Escabí, M. A., Miller, L. M., Read, H. L., & Schreiner, C. E. (2003). Naturalistic Auditory Contrast Improves Spectrotemporal Coding in the Cat Inferior Colliculus. *The Journal of Neuroscience*, 23(37), 11489-11504. <https://doi.org/10.1523/JNEUROSCI.23-37-11489.2003>
- Ewert, S. D., & Dau, T. (2000). Characterizing frequency selectivity for envelope fluctuations. *The Journal of the Acoustical Society of America*, 108(3), 1181. <https://doi.org/10.1121/1.1288665>
- Ewert, S. D., & Dau, T. (2004). External and internal limitations in amplitude-modulation processing. *The Journal of the Acoustical Society of America*, 116(1), 478-490. <https://doi.org/10.1121/1.1737399>
- Ewert, S. D., Paraouty, N., & Lorenzi, C. (2020). A two-path model of auditory modulation detection using temporal fine structure and envelope cues. *European Journal of Neuroscience*, 51(5), 1265-1278. <https://doi.org/10.1111/ejn.13846>
- Fahle, M., & Edelman, S. (1993). Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision research*, 33 (3), 397-412.
- Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, 6 (3), 292-297.
- Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4), 292-303. <https://doi.org/10.1038/nrn2258>
- Falkenberg, H. K., Dutton, G. N., & Simpson, W. A. (2010). Does Motion Discrimination Sensitivity in Children Aged 5-14 Years Improve After 1 Year?. *Investigative Ophthalmology & Visual Science*, 51 (13), 1839-1839.
- Falkenberg, H. K., Simpson, W. A., & Dutton, G. N. (2014). Development of sampling efficiency and internal noise in motion detection and discrimination in school-aged children. *Vision Research*, 100, 8-17. <https://doi.org/10.1016/j.visres.2014.04.001>
- Farina, A., & Gage, S. H. (Éds.). (2017). *Ecoacoustics : The ecological role of sounds* (First edition). John Wiley & Sons, Inc.
- Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America*, 88(4), 1725-1736. <https://doi.org/10.1121/1.400247>
- Feth, L. L. (1972). Combinations of amplitude and frequency differences in auditory discrimination. *Acta Acustica united with Acustica*, 26(2), 67-77.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4(12), 2379. <https://doi.org/10.1364/JOSAA.4.002379>
- Fior, R. (1972). Physiological maturation of auditory function between 3 and 13 years of age. *Audiology*, 11(5-6), 317-321.

- Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature*, 287(5777), 43-44.
- Fiorentini, A., & Berardi, N. (1981). Learning in grating waveform discrimination: Specificity for orientation and spatial frequency. *Vision research*, 21(7), 1149-1158.
- Fitzgibbons, P. J., & Gordon-Salant, S. (1995). Age effects on duration discrimination with simple and complex stimuli. *The Journal of the Acoustical Society of America*, 98(6), 3140-3145. <https://doi.org/10.1121/1.413803>
- Fitzgibbons, P. J., Gordon-Salant, S., & Friedman, S. A. (2006). Effects of age and sequence presentation rate on temporal order recognition. *The Journal of the Acoustical Society of America*, 120(2), 991-999. <https://doi.org/10.1121/1.2214463>
- Flanagan, J. L. (1980). Parametric coding of speech spectra. *The Journal of the Acoustical Society of America*, 68(2), 412-419. <https://doi.org/10.1121/1.384752>
- Friis, H. T. (1944). Noise figures of radio receivers. *Proceedings of the IRE*, 32(7), 419-422.
- Formby, C. C. (1982). Differential sensitivity to tonal frequency and to the rate of modulation of broad-band noise by hearing-impaired listeners.
- Füllgrabe, C., Berthommier, F., & Lorenzi, C. (2006). Masking release for consonant features in temporally fluctuating background noise. *Hearing research*, 211(1-2), 74-84.
- Füllgrabe, C. (2013). Age-Dependent Changes in Temporal-Fine-Structure Processing in the Absence of Peripheral Hearing Loss. *American Journal of Audiology*, 22(2), 313-315. [https://doi.org/10.1044/1059-0889\(2013\)12-0070](https://doi.org/10.1044/1059-0889(2013)12-0070)
- Füllgrabe, C., Moore, B. C., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Frontiers in aging neuroscience*, 6, 347.
- Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *Journal of Neurophysiology*, 110(3), 577-586. <https://doi.org/10.1152/jn.00164.2013>
- Galvin, J. J., Fu, Q.-J., & Oba, S. (2008). Effect of instrument timbre on melodic contour identification by cochlear implant users. *The Journal of the Acoustical Society of America*, 124(4), EL189-EL195. <https://doi.org/10.1121/1.2961171>
- Galvin, J. J., Fu, Q.-J., & Shannon, R. V. (2009). Melodic Contour Identification and Music Perception by Cochlear Implant Users. *Annals of the New York Academy of Sciences*, 1169(1), 518-533. <https://doi.org/10.1111/j.1749-6632.2009.04551.x>
- Gao, H., & Hollyfield, J. G. (1992). Aging of the human retina. Differential loss of neurons and retinal pigment epithelial cells. *Investigative ophthalmology & visual science*, 33(1), 1-17.
- Garcia-Lazaro, J. A., Ahmed, B., & Schnupp, J. W. H. (2006). Tuning to Natural Stimulus Dynamics in Primary Auditory Cortex. *Current Biology*, 16(3), 264-271. <https://doi.org/10.1016/j.cub.2005.12.013>
- Gegenfurtner, K. R., & Kiper, D. C. (1992). Contrast detection in luminance and chromatic noise. *JOSA A*, 9(11), 1880-1888.
- Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discrim-

- inations. *Psychological Review*, 96, 267–314.
- Gelfand, S. A., Piper, N., & Silman, S. (1986). Consonant recognition in quiet and in noise with aging among normal hearing listeners. *The Journal of the Acoustical Society of America*, 80(6), 1589-1598. <https://doi.org/10.1121/1.394323>
- Geurts, H. M., Verté, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. A. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism?. *Journal of child psychology and psychiatry*, 45(4), 836-854.
- Geurts, H. M., Grasman, R. P., Verté, S., Oosterlaan, J., Roeyers, H., van Kammen, S. M., & Sergeant, J. A. (2008). Intra-individual variability in ADHD, autism spectrum disorders and Tourette's syndrome. *Neuropsychologia*, 46(13), 3030-3041.
- Gfeller, K., & Lansing, C. R. (1991). Melodic, Rhythmic, and Timbral Perception of Adult Cochlear Implant Users. *Journal of Speech, Language, and Hearing Research*, 34(4), 916-920. <https://doi.org/10.1044/jshr.3404.916>
- Gfeller, K., Witt, S., Mehr, M. A., Woodworth, G., & Knutson, J. (2002). Effects of Frequency, Instrumental Family, and Cochlear Implant Type on Timbre Recognition and Appraisal. *Annals of Otology, Rhinology & Laryngology*, 111(4), 349-356. <https://doi.org/10.1177/000348940211100412>
- Ghitza, O. (2001). On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception. *The Journal of the Acoustical Society of America*, 110(3), 1628-1640. <https://doi.org/10.1121/1.1396325>
- Gilbert, G., & Lorenzi, C. (2006). The ability of listeners to use recovered envelope cues from speech fine structure. *The Journal of the Acoustical Society of America*, 119(4), 2438-2444. <https://doi.org/10.1121/1.2173522>
- Gifford, R. (2007). *Environmental psychology: Principles and practice* (p. 372). Colville, WA: Optimal books.
- Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing : Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511-517. <https://doi.org/10.1038/nn.3063>
- Glasberg, B. R., Moore, B. C. J., & Peters, R. W. (2001). The influence of external and internal noise on the detection of increments and decrements in the level of sinusoids. *Hearing Research*, 13.
- Glowatzki, E., & Fuchs, P. A. (2002). Transmitter release at the hair cell ribbon synapse. *Nature neuroscience*, 5(2), 147-154.
- Gnansia, D., Jourdes, V., & Lorenzi, C. (2008). Effect of masker modulation depth on speech masking release. *Hearing Research*, 239(1-2), 60-68. <https://doi.org/10.1016/j.heares.2008.01.012>
- Gnansia, D., Péan, V., Meyer, B., & Lorenzi, C. (2009). Effects of spectral smearing and temporal fine structure degradation on speech masking release. *The Journal of the Acoustical Society of America*, 125(6), 4023-4033. <https://doi.org/10.1121/1.3126344>
- Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature*, 402(6758), 176-178.

- Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, 28(2), 167-207.
- Goldberg, M. C., Mostofsky, S. H., Cutting, L. E., Mahone, E. M., Astor, B. C., Denckla, M. B., & Landa, R. J. (2005). Subtle executive impairment in children with autism and children with ADHD. *Journal of autism and developmental disorders*, 35(3), 279-293.
- Goldstein, J. L. and P. Sruлович (1977), "Auditory-nerve spike intervals as an adequate basis for aural frequency measurement", *Psychophysics and physiology of hearing*, Evans and Wilson ed. Academic Press: London, 337- 345.
- Goodman, D. F. M., Winter, I. M., Léger, A. C., de Cheveigné, A., & Lorenzi, C. (2018). Modelling firing regularity in the ventral cochlear nucleus : Mechanisms, and effects of stimulus level and synaptopathy. *Hearing Research*, 358, 98-110. <https://doi.org/10.1016/j.heares.2017.09.010>
- Gordon, J. W. (1987). The perceptual attack time of musical tones. *The Journal of the Acoustical Society of America*, 82(1), 88-105.
- Gordon-Salant, S. (2005). Hearing loss and aging : New research findings and clinical implications. *The Journal of Rehabilitation Research and Development*, 42(4s), 9. <https://doi.org/10.1682/JRRD.2005.01.0006>
- Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal Factors and Speech Recognition Performance in Young and Elderly Listeners. *Journal of Speech, Language, and Hearing Research*, 36(6), 1276-1285. <https://doi.org/10.1044/jshr.3606.1276>
- Gordon-Salant, S., & Fitzgibbons, P. J. (1999). Profile of Auditory Temporal Processing in Older Listeners. *Journal of Speech, Language, and Hearing Research*, 42(2), 300-311. <https://doi.org/10.1044/jslhr.4202.300>
- Gordon-Salant, S., Frisina, R. D., Fay, R. R., & Popper, A. (Eds.). (2010). *The aging auditory system* (Vol. 34). Springer Science & Business Media.
- Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: converging perspectives?. *Laboratory Phonology*, 4(1), 67-92.
- Gou, Z., Choudhury, N., & Benasich, A. A. (2011). Resting frontal gamma power at 16, 24 and 36 months predicts individual differences in language and cognition at 4 and 5 years. *Behavioural brain research*, 220(2), 263-270.
- Gransier, R., Hofmann, M., van Wieringen, A., & Wouters, J. (2021). Stimulus-evoked phase-locked activity along the human auditory pathway strongly varies across individuals. *Scientific Reports*, 11(1), 143. <https://doi.org/10.1038/s41598-020-80229-w>
- Green, D. M. (1960a). Auditory Detection of a Noise Signal. *The Journal of the Acoustical Society of America*, 32(1), 121-131. <https://doi.org/10.1121/1.1907862>
- Green, D. M. (1960b). Psychoacoustics and Detection Theory. *The Journal of the Acoustical Society of America*, 32(10), 1189-1203. <https://doi.org/10.1121/1.1907882>
- Green, D. M. (1993). A maximum-likelihood method for estimating thresholds in a yes-no task. *The Journal of the Acoustical Society of America*, 93(4), 2096-2105.

- Green, D. M. (1964). Consistency of auditory detection judgments. *Psychological review*, 71(5), 392.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1, pp. 1969-12). New York: Wiley.
- Green, D. M. (1993). A maximum-likelihood method for estimating thresholds in a yes–no task. *The Journal of the Acoustical Society of America*, 93(4), 2096-2105. <https://doi.org/10.1121/1.406696>
- Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech—A syllable-centric perspective. *Journal of Phonetics*, 31(3-4), 465-485. <https://doi.org/10.1016/j.wocn.2003.09.005>
- Grose, J. H., Buss, E., & Elmore, H. (2019). Age-Related Changes in the Auditory Brainstem Response and Suprathreshold Processing of Temporal and Spectral Modulation. *Trends in Hearing*, 23, 233121651983961. <https://doi.org/10.1177/2331216519839615>
- Grose, J. H., Buss, E., & Hall, J. W. (2009). Within- and across-channel factors in the multiband comodulation masking release paradigm. *The Journal of the Acoustical Society of America*, 125(1), 282-293. <https://doi.org/10.1121/1.3023067>
- Grose, J. H., Hall, J. W., & Buss, E. (2006). Temporal processing deficits in the pre-senescent auditory system. *The Journal of the Acoustical Society of America*, 119(4), 2305-2315. <https://doi.org/10.1121/1.2172169>
- Grose, J. H., & Mamo, S. K. (2010). Processing of Temporal Fine Structure as a Function of Age. *Ear & Hearing*, 31(6), 755-760. <https://doi.org/10.1097/AUD.0b013e3181e627e7>
- Grose, J. H., & Mamo, S. K. (2012). Frequency modulation detection as a measure of temporal processing : Age-related monaural and binaural effects. *Hearing Research*, 294(1-2), 49-54. <https://doi.org/10.1016/j.heares.2012.09.007>
- Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *The Journal of the Acoustical Society of America*, 115(3), 1252-1265. <https://doi.org/10.1121/1.1635840>
- Hadad, B. S., Maurer, D., & Lewis, T. L. (2011). Long trajectory for the development of sensitivity to global and biological motion. *Developmental science*, 14(6), 1330-1339.
- Hajda, J. M. (2007). The Effect of Dynamic Acoustical Features on Musical Timbre. In J. W. Beauchamp (Éd.), *Analysis, Synthesis, and Perception of Musical Sounds* (p. 250-271). Springer New York. https://doi.org/10.1007/978-0-387-32576-7_7
- Hall, J. W., & Grose, J. H. (1991). Notched-noise measures of frequency selectivity in adults and children using fixed-masker-level and fixed-signal-level presentation. *Journal of Speech, Language, and Hearing Research*, 34(3), 651-660.
- Hall, J. W., & Grose, J. H. (1994). Development of temporal resolution in children as measured by the temporal modulation transfer function. *The Journal of the Acoustical Society of America*, 96(1), 150-154. <https://doi.org/10.1121/1.410474>

- Hall, J. W., Buss, E., & Grose, J. H. (2005). Informational masking release in children and adults. *The Journal of the Acoustical Society of America*, 118(3), 1605-1613. <https://doi.org/10.1121/1.1992675>
- Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature neuroscience*, 9(10), 1218-1220.
- Hart, H. C., Palmer, A. R., & Hall, D. A. (2003). Amplitude and frequency-modulated stimuli activate common regions of human auditory cortex. *Cerebral Cortex*, 13(7), 773-781.
- Hartmann, W. M., & Pumplin, J. (1988). Noise power fluctuations and the masking of sine signals. *The Journal of the Acoustical Society of America*, 83(6), 2277-2289. <https://doi.org/10.1121/1.396358>
- Harrison, R. V., & Evans, E. F. (1979). Some aspects of temporal coding by single cochlear fibres from regions of cochlear hair cell degeneration in the guinea pig. *Archives of oto-rhino-laryngology*, 224(1), 71-78.
- Hasan, B. A. S., Joosten, E., & Neri, P. (2012). Estimation of internal noise using double-passes: Does it matter how the second pass is delivered?. *Vision research*, 69, 1-9.
- Hay, G. A., & Chesters, M. S. (1972). Signal-transfer functions in threshold and suprathreshold vision. *Journal of the Optical Society of America*, 62, 990-998.
- He, N., Mills, J. H., & Dubno, J. R. (2007). Frequency modulation detection: effects of age, psychophysical method, and modulation waveform. *The Journal of the Acoustical Society of America*, 122(1), 467-477.
- He, N., Mills, J. H., Ahlstrom, J. B., & Dubno, J. R. (2008). Age-related differences in the temporal modulation transfer function with pure-tone carriers. *The Journal of the Acoustical Society of America*, 124(6), 3841-3849.
- Hecht, S., Schlaer, S., & Pirenne, M. H. (1942). Energy, quanta, and vision. *Journal of General Physiology*, 25(6), 819-840.
- Heinz, M. G., Zhang, X., Bruce, I. C., & Carney, L. H. (2001). Auditory nerve model for predicting performance limits of normal and impaired listeners. *Acoustics Research Letters Online*, 2(3), 91-96.
- Heinz, M. G., & Swaminathan, J. (2009). Quantifying Envelope and Fine-Structure Coding in Auditory Nerve Responses to Chimaeric Speech. *Journal of the Association for Research in Otolaryngology*, 10(3), 407-423. <https://doi.org/10.1007/s10162-009-0169-8>
- Helfer, K. S., & Vargo, M. (2009). Speech Recognition and Temporal Processing in Middle-Aged Women. *Journal of the American Academy of Audiology*, 20(04), 264-271. <https://doi.org/10.3766/jaaa.20.4.6>
- Hemenger, R. P. (1984). Intraocular light scatter in normal vision loss with age. *Applied Optics*, 23(12), 1972. <https://doi.org/10.1364/AO.23.001972>
- Henry, K. S., & Heinz, M. G. (2013). Effects of sensorineural hearing loss on temporal coding of narrowband and broadband signals in the auditory periphery. *Hearing Research*, 303, 39-47. <https://doi.org/10.1016/j.heares.2013.01.014>

- Henry, K. S., Kale, S., & Heinz, M. G. (2014). Noise-induced hearing loss increases the temporal precision of complex envelope coding by auditory-nerve fibers. *Frontiers in Systems Neuroscience*, 8. <https://doi.org/10.3389/fnsys.2014.00020>
- Henry, K. S., Kale, S., & Heinz, M. G. (2016). Distorted tonotopic coding of temporal envelope and fine structure with noise-induced hearing loss. *Journal of Neuroscience*, 36(7), 2227-2237.
- Hirsh, I. J., & Watson, C. S. (1996). AUDITORY PSYCHOPHYSICS AND PERCEPTION. *Annual Review of Psychology*, 47(1), 461-484. <https://doi.org/10.1146/annurev.psych.47.1.461>
- Hopkins, K., & Moore, B. C. J. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *The Journal of the Acoustical Society of America*, 125(1), 442-446. <https://doi.org/10.1121/1.3037233>
- Hopkins, K., & Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The Journal of the Acoustical Society of America*, 130(1), 334-349. <https://doi.org/10.1121/1.3585848>
- Hopkins, K., Moore, B. C. J., & Stone, M. A. (2010). The effects of the addition of low-level, low-noise noise on the intelligibility of sentences processed to remove temporal envelope information. *The Journal of the Acoustical Society of America*, 128(4), 2150-2161. <https://doi.org/10.1121/1.3478773>
- Hornickel, J., & Kraus, N. (2013). Unstable Representation of Sound : A Biological Marker of Dyslexia. *Journal of Neuroscience*, 33(8), 3500-3504. <https://doi.org/10.1523/JNEUROSCI.4205-12.2013>
- Houtgast, T. (1989). Frequency selectivity in amplitude-modulation detection. *The Journal of the Acoustical Society of America*, 85(4), 1676-1680. <https://doi.org/10.1121/1.397956>
- Houtgast, T., & Steeneken, H. J. M. (1973). The Modulation Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility. *The Journal of the Acoustical Society of America*, 54(2), 557-557. <https://doi.org/10.1121/1.1913632>
- Houtgast, T., & Steeneken, H. J. M. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *The Journal of the Acoustical Society of America*, 77(3), 1069-1077. <https://doi.org/10.1121/1.392224>
- Humes, L. E., & Jesteadt, W. (1989). Models of the additivity of masking. *The Journal of the Acoustical Society of America*, 85(3), 1285-1294. <https://doi.org/10.1121/1.397459>
- Humes, L. E., & Christopherson, L. (1991). Speech identification difficulties of hearing-impaired elderly persons: The contributions of auditory processing deficits. *Journal of Speech, Language, and Hearing Research*, 34(3), 686-693.
- Humes, L. E., Nelson, K. J., & Pisoni, D. B. (1991). Recognition of Synthetic Speech by Hearing-Impaired Elderly Listeners. *Journal of Speech, Language, and Hearing Research*, 34(5), 1180-1184. <https://doi.org/10.1044/jshr.3405.1180>

- Humes, L. E., & Roberts, L. (1990). Speech-Recognition Difficulties of the Hearing-Impaired Elderly : The Contributions of Audibility. 10.
- Ives, D. T., Kalluri, S., Strelcyk, O., Sheft, S., Miermont, F., Coez, A., ... & Lorenzi, C. (2014). Effects of noise reduction on AM perception for hearing-impaired listeners. *Journal of the Association for Research in Otolaryngology*, 15(5), 839-848.
- Javel, E., & Viemeister, N. F. (2000). Stochastic properties of cat auditory nerve responses to electric and acoustic stimuli and application to intensity discrimination. *The Journal of the Acoustical Society of America*, 107(2), 908-921.
- Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory discrimination in preschool children. *Psychological Science*, 4(2), 104-107.
- Jepsen, M. L., Ewert, S. D., & Dau, T. (2008). A computational model of human auditory signal processing and perception. *The Journal of the Acoustical Society of America*, 124(1), 422-438. <https://doi.org/10.1121/1.2924135>
- Jerger, J., Shedd, J. L., & Harford, E. (1959). On the detection of extremely small changes in sound intensity. *AMA archives of Otolaryngology*, 69(2), 200-211.
- Jesteadt, W., Nizami, L., & Schairer, K. S. (2003). A measure of internal noise based on sample discrimination. *The Journal of the Acoustical Society of America*, 114(4), 2147-2157. <https://doi.org/10.1121/1.1610456>
- John, M. S., Dimitrijevic, A., van Roon, P., & Picton, T. W. (2001). Multiple Auditory Steady-State Responses to AM and FM Stimuli. *Audiology and Neuro-Otology*, 6(1), 12-27. <https://doi.org/10.1159/000046805>
- Johnson, B. W., Weinberg, H., Ribary, U., Cheyne, D. O., & Ancill, R. (1988). Topographic distribution of the 40 Hz auditory evoked-Related potential in normal and aged subjects. *Brain Topography*, 1(2), 117-121. <https://doi.org/10.1007/BF01129176>
- Jones, P. R., Moore, D. R., Amitay, S., & Shub, D. E. (2013). Reduction of internal noise in auditory perceptual learning. *The Journal of the Acoustical Society of America*, 133(2), 970-981. <https://doi.org/10.1121/1.4773864>
- Joris, P. X., Carney, L. H., Smith, P. H., & Yin, T. C. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *Journal of Neurophysiology*, 71(3), 1022-1036. <https://doi.org/10.1152/jn.1994.71.3.1022>
- Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural Processing of Amplitude-Modulated Sounds. *Physiological Reviews*, 84(2), 541-577. <https://doi.org/10.1152/physrev.00029.2003>
- Kaasinen, V., Vilkmann, H., Hietala, J., Nagren, K., Helenius, H., Olsson, H., Farde, L., & O.Rinne, J. (2000). Age-related dopamine D2/D3 receptor loss in extrastriatal regions of the human brain. *Neurobiology of Aging*, 21(5), 683-688. [https://doi.org/10.1016/S0197-4580\(00\)00149-4](https://doi.org/10.1016/S0197-4580(00)00149-4)
- Kale, S., & Heinz, M. G. (2010). Envelope Coding in Auditory Nerve Fibers Following Noise-Induced Hearing Loss. *Journal of the Association for Research in Otolaryngology*, 11(4), 657-673. <https://doi.org/10.1007/s10162-010-0223-6>

- Kale, S., & Heinz, M. G. (2012). Temporal modulation transfer functions measured from auditory-nerve responses following sensorineural hearing loss. *Hearing Research*, 286(1-2), 64-75. <https://doi.org/10.1016/j.heares.2012.02.004>
- Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences*, 88(11), 4966-4970.
- Kates, W. R., Frederikse, M., Mostofsky, S. H., Folley, B. S., Cooper, K., Mazur-Hopkins, P., ... & Kaufmann, W. E. (2002). MRI parcellation of the frontal lobe in boys with attention deficit hyperactivity disorder or Tourette syndrome. *Psychiatry Research: Neuroimaging*, 116(1-2), 63-81.
- Kates, J. M. (2011). Spectro-temporal envelope changes caused by temporal fine structure modification. *The Journal of the Acoustical Society of America*, 129(6), 3981-3990. <https://doi.org/10.1121/1.3583552>
- King, A., Hopkins, K., & Plack, C. J. (2014). The effects of age and hearing loss on interaural phase difference discrimination. *The Journal of the Acoustical Society of America*, 135(1), 342-351. <https://doi.org/10.1121/1.4838995>
- King, A., Varnet, L., & Lorenzi, C. (2019). Accounting for masking of frequency modulation by amplitude modulation with the modulation filter-bank concept. *The Journal of the Acoustical Society of America*, 145(4), 2277-2293. <https://doi.org/10.1121/1.5094344>
- King, K., & Stephens, D. (1992). Auditory and Psychological Factors in 'Auditory Disability with Normal Hearing'. *Scandinavian Audiology*, 21(2), 109-114. <https://doi.org/10.3109/01050399209045990>
- Kohlrausch, A., Fassel, R., & Dau, T. (2000). The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *The Journal of the Acoustical Society of America*, 108(2), 723-734. <https://doi.org/10.1121/1.429605>
- Kong, Y.-Y., Cruz, R., Jones, J. A., & Zeng, F.-G. (2004). Music Perception with Temporal Cues in Acoustic and Electric Hearing: *Ear and Hearing*, 25(2), 173-185. <https://doi.org/10.1097/01.AUD.0000120365.97792.2F>
- Kong, Y.-Y., Mullangi, A., Marozeau, J., & Epstein, M. (2011). Temporal and Spectral Cues for Musical Timbre Perception in Electric Hearing. *Journal of Speech, Language, and Hearing Research*, 54(3), 981-994. [https://doi.org/10.1044/1092-4388\(2010/10-0196\)](https://doi.org/10.1044/1092-4388(2010/10-0196))
- Kong, Y.-Y., Mullangi, A., Marozeau, J., & Epstein, M. (2012). Temporal and Spectral Cues for Musical Timbre Perception in Electric Hearing. 24.
- Koning, R., & Wouters, J. (2012). The potential of onset enhancement for increased speech intelligibility in auditory prostheses. *The Journal of the Acoustical Society of America*, 132(4), 2569-2581.
- Kortlang, S., Mauermann, M., & Ewert, S. D. (2016). Suprathreshold auditory processing deficits in noise : Effects of hearing loss and age. *Hearing Research*, 331, 27-40. <https://doi.org/10.1016/j.heares.2015.10.004>

- Koumura, T., Terashima, H., & Furukawa, S. (2019). Cascaded Tuning to Amplitude Modulation for Natural Sound Recognition. *The Journal of Neuroscience*, 39(28), 5517-5533. <https://doi.org/10.1523/JNEUROSCI.2914-18.2019>
- Kozlov, A. S., Risler, T., & Hudspeth, A. J. (2007). Coherent motion of stereocilia assures the concerted gating of hair-cell transduction channels. *Nature neuroscience*, 10(1), 87-92.
- Kraus, N., McGee, T. J., Carrell, T. D., Zecker, S. G., Nicol, T. G., & Koch, D. B. (1996). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science*, 273(5277), 971-973.
- Krauss, P., Tziridis, K., Metzner, C., Schilling, A., Hoppe, U., & Schulze, H. (2016). Stochastic Resonance Controlled Upregulation of Internal Noise after Hearing Loss as a Putative Cause of Tinnitus-Related Neuronal Hyperactivity. *Frontiers in Neuroscience*, 10. <https://doi.org/10.3389/fnins.2016.00597>
- Kujawa, S. G., & Liberman, M. C. (2006). Acceleration of Age-Related Hearing Loss by Early Noise Exposure : Evidence of a Misspent Youth. *Journal of Neuroscience*, 26(7), 2115-2123. <https://doi.org/10.1523/JNEUROSCI.4985-05.2006>
- Kujawa, S. G., & Liberman, M. C. (2009). Adding Insult to Injury : Cochlear Nerve Degeneration after « Temporary » Noise-Induced Hearing Loss. *Journal of Neuroscience*, 29(45), 14077-14085. <https://doi.org/10.1523/JNEUROSCI.2845-09.2009>
- Kulikowski, J. J. (1971). Effect of eye movements on the contrast sensitivity of spatio-temporal patterns. *Vision research*, 11(3), 261-273.
- Kuwada, S., Batra, R., & Maher, V. L. (1986). Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones. *Hearing Research*, 21(2), 179-192. [https://doi.org/10.1016/0378-5955\(86\)90038-9](https://doi.org/10.1016/0378-5955(86)90038-9)
- Lacher-Fougère, S., & Demany, L. (1998). Modulation Detection by Normal and Hearing-impaired Listeners. *International Journal of Audiology*, 37(2), 109-121. <https://doi.org/10.3109/00206099809072965>
- Lacher-Fougère, S., & Demany, L. (2005). Consequences of cochlear damage for the detection of interaural phase differences. *The Journal of the Acoustical Society of America*, 118(4), 2519-2526. <https://doi.org/10.1121/1.2032747>
- Lamore, P. J. J., Verweij, C., & Brocaar, M. P. (1984). Reliability of auditory function tests in severely hearing-impaired and deaf subjects. *Audiology*, 23(5), 453-466.
- Lawson, J. L., & Uhlenbeck, G. E. (1950). Threshold Signals. 388.
- Léger, A. C., Ives, D. T., & Lorenzi, C. (2014). Abnormal intelligibility of speech in competing speech and in noise in a frequency region where audiometric thresholds are near-normal for hearing-impaired listeners. *Hearing Research*, 316, 102-109. <https://doi.org/10.1016/j.heares.2014.07.008>
- Léger, A. C., Moore, B. C. J., Gnansia, D., & Lorenzi, C. (05/2012c). Effects of spectral smearing on the identification of speech in noise filtered into low- and mid-frequency regions. *The Journal of the Acoustical Society of America*, 131(5), 4114-4123. <https://doi.org/10.1121/1.3699265>

- Léger, A. C., Moore, B. C. J., & Lorenzi, C. (2012). Abnormal speech processing in frequency regions where absolute thresholds are normal for listeners with high-frequency hearing loss. *Hearing Research*, 294(1-2), 95-103. <https://doi.org/10.1016/j.heares.2012.10.002>
- Léger, A. C., Moore, B. C. J., & Lorenzi, C. (12/2012b). Abnormal speech processing in frequency regions where absolute thresholds are normal for listeners with high-frequency hearing loss. *Hearing Research*, 294(1-2), 95-103. <https://doi.org/10.1016/j.heares.2012.10.002>
- Léger, A. C., Moore, B. C. J., & Lorenzi, C. (02/2012a). Temporal and spectral masking release in low- and mid-frequency regions for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 131(2), 1502-1514. <https://doi.org/10.1121/1.3665993>
- Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Josa*, 70(12), 1458-1471.
- Legge, G. E., Kersten, D., & Burgess, A. E. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America*, 4(A), 391-404.
- Leigh-Paffenroth, E. D., & Fowler, C. G. (2006). Amplitude-Modulated Auditory Steady-State Responses in Younger and Older Listeners. *Journal of the American Academy of Audiology*, 17(08), 582-597. <https://doi.org/10.3766/jaaa.17.8.5>
- Lentz, J. J., & Richards, V. M. (1997). Sensitivity to changes in overall level and spectral shape: An evaluation of a channel model. *The Journal of the Acoustical Society of America*, 101(6), 3625-3635.
- Lesica, N. A., & Grothe, B. (2008). Dynamic spectrotemporal feature selectivity in the auditory midbrain. *Journal of Neuroscience*, 28(21), 5412-5421.
- Levi, D. M., & Klein, S. A. (2003). Noise provides some new signals about the spatial vision of amblyopes. *Journal of Neuroscience*, 23(7), 2522-2526.
- Levi, D. M., Klein, S. A., & Chen, I. (2007). The response of the amblyopic visual system to noise. *Vision Research*, 47(19), 2531-2542. <https://doi.org/10.1016/j.visres.2007.06.014>
- Levin, J. E., & Miller, J. P. (1996). Broadband neural encoding in the cricket cercal sensory system enhanced by stochastic resonance. *Nature*, 380(6570), 165-168. <https://doi.org/10.1038/380165a0>
- Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2B), 467-477. <https://doi.org/10.1121/1.1912375>
- Lewicki, M. S. (2002). Efficient coding of natural sounds. *Nature Neuroscience*, 5(4), 356-363. <https://doi.org/10.1038/nn831>
- Liang, L., Lu, T., & Wang, X. (2002). Neural Representations of Sinusoidal Amplitude and Frequency Modulations in the Primary Auditory Cortex of Awake Primates. *Journal of Neurophysiology*, 87(5), 2237-2261. <https://doi.org/10.1152/jn.2002.87.5.2237>
- Lieberman, M. C., & Kiang, N. Y. (1978). Acoustic trauma in cats: cochlear pathology and auditory-nerve activity. *Acta oto-laryngologica*.

- Liberman, M. C. (2017). Noise-induced and age-related hearing loss : New perspectives and potential therapies. *F1000Research*, 6, 927. <https://doi.org/10.12688/f1000research.11310.1>
- Licklider, J. C. R. (1946). Effects of Amplitude Distortion upon the Intelligibility of Speech. *The Journal of the Acoustical Society of America*, 18(1), 249-249. <https://doi.org/10.1121/1.1902443>
- Licklider, J. C. R., & Pollack, I. (1948). Effects of Differentiation, Integration, and Infinite Peak Clipping upon the Intelligibility of Speech. *The Journal of the Acoustical Society of America*, 20(1), 42-51. <https://doi.org/10.1121/1.1906346>
- Lillywhite, P. G., & Laughlin, S. B. (1979). Transducer noise in a photoreceptor. *Nature*, 277(5697), 569-572.
- Lobarinas, E., Salvi, R., & Ding, D. (2016). Selective Inner Hair Cell Dysfunction in Chinchillas Impairs Hearing-in-Noise in the Absence of Outer Hair Cell Loss. *Journal of the Association for Research in Otolaryngology*, 17(2), 89-101. <https://doi.org/10.1007/s10162-015-0550-8>
- London, Michael, et al. "Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex." *Nature* 466.7302 (2010): 123-127.
- Lopez-Poveda, E. A., & Barrios, P. (2013). Perception of stochastically undersampled sound waveforms : A model of auditory deafferentation. *Frontiers in Neuroscience*, 7. <https://doi.org/10.3389/fnins.2013.00124>
- Lopez-Poveda, E. A., Eustaquio-Martín, A., Stohl, J. S., Wolford, R. D., Schatzer, R., Gorospe, J. M., Ruiz, S. S. C., Benito, F., & Wilson, B. S. (2017). Intelligibility in speech maskers with a binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. *Hearing Research*, 348, 134-137. <https://doi.org/10.1016/j.heares.2017.02.003>
- Lorenz, Edward N. "Deterministic nonperiodic flow." *Journal of atmospheric sciences* 20.2 (1963): 130-141.
- Lorenz, Edward N. "Atmospheric predictability as revealed by naturally occurring analogues." *Journal of Atmospheric Sciences* 26.4 (1969): 636-646.
- Lorenzi, C., Dumont, A., & Füllgrabe, C. (2000). Use of Temporal Envelope Cues by Children With Developmental Dyslexia. *Journal of Speech, Language, and Hearing Research*, 43(6), 1367-1379. <https://doi.org/10.1044/jslhr.4306.1367>
- Lorenzi, C., Dumont, A., & Füllgrabe, C. (2000). Use of temporal envelope cues by children with developmental dyslexia. *Journal of speech, language, and hearing research*, 43(6), 1367-1379.
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences*, 103(49), 18866-18869. <https://doi.org/10.1073/pnas.0607364103>
- Lorenzi, C., Micheyl, C., Berthommier, F., & Portalier, S. (1997). Modulation Masking in Listeners With Sensorineural Hearing Loss. *Journal of Speech, Language, and Hearing Research*, 40(1), 200-207. <https://doi.org/10.1044/jslhr.4001.200>

- Lorenzi, C., Soares, C., & Vonner, T. (2001). Second-order temporal modulation transfer functions. *The Journal of the Acoustical Society of America*, 110(2), 1030-1038. <https://doi.org/10.1121/1.1383295>
- Lorenzi, C., Wallaert, N., Gnansia, D., Leger, A. C., Ives, D. T., Chays, A., Garnier, S., & Cazals, Y. (2012). Temporal-Envelope Reconstruction for Hearing-Impaired Listeners. *Journal of the Association for Research in Otolaryngology*, 13(6), 853-865. <https://doi.org/10.1007/s10162-012-0350-3>
- Lu, Z.-L., & Doshier, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research*, 38, 1183–1198.
- Lu, Z. L., & Doshier, B. A. (2008). Characterizing observers using external noise and observer models : Assessing internal representations with external noise. *Psychological Review*, 115(1), 44-82. <https://doi.org/10.1037/0033-295X.115.1.44>
- Lu, Z. L., & Doshier, B. A. (2009). Mechanisms of perceptual learning. *Learning & perception*, 1(1), 19-36.
- Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behavior: Sensitivity and discrimination. *Ear and hearing*, 3(6), 301-308.
- Lutfi, R. A. (1989). Informational processing of complex sound. I: Intensity discrimination. *The Journal of the Acoustical Society of America*, 86(3), 934-944.
- Maiwald, D.: Ein Funktionsschema des Gehörs zur Beschreibung der Erkennbarkeit kleiner Frequenz- und Amplitudenänderungen. *Acustica* 18, 81-92 (1967a)
- Maiwald, D.: Die Berechnung von Modulationsschwellen mit Hilfe eines Funktionsschemas. *Acustica* 18, 193-207 (1967b)
- McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres : Common dimensions, specificities, and latent subject classes. *Psychological Research*, 58(3), 177-192. <https://doi.org/10.1007/BF00419633>
- McAlpine, D., Haywood, N., Undurraga, J., & Marquardt, T. (2016). Objective measures of neural processing of interaural time differences. In *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing* (pp. 197-205). Springer, Cham.
- McAnany, J. J., Alexander, K. R., Genead, M. A., & Fishman, G. A. (2013). Equivalent Intrinsic Noise, Sampling Efficiency, and Contrast Sensitivity in Patients With Retinitis Pigmentosa. *Investigative Ophthalmology & Visual Science*, 54(6), 3857. <https://doi.org/10.1167/iovs.13-11789>
- McDermott, H. J. (2004). Music Perception with Cochlear Implants : A Review. *Trends in Amplification*, 8(2), 49-82. <https://doi.org/10.1177/108471380400800203>
- McDermott, H. J., & Looi, V. (2004). Perception of complex signals, including musical sounds, with cochlear implants. *International Congress Series*, 1273, 201-204. <https://doi.org/10.1016/j.ics.2004.08.034>
- McDermott, J. H., & Simoncelli, E. P. (2011). Sound Texture Perception via Statistics of the Auditory Periphery : Evidence from Sound Synthesis. *Neuron*, 71(5), 926-940. <https://doi.org/10.1016/j.neuron.2011.06.032>

- McPherson, M. J., Dolan, S. E., Durango, A., Ossandon, T., Valdés, J., Undurraga, E. A., Jacoby, N., Godoy, R. A., & McDermott, J. H. (2020). Perceptual fusion of musical notes by native Amazonians suggests universal representations of musical intervals. *Nature Communications*, 11(1), 2786. <https://doi.org/10.1038/s41467-020-16448-6>
- McWalter, R., & Dau, T. (2017). Cascaded Amplitude Modulations in Sound Texture Perception. *Frontiers in Neuroscience*, 11, 485. <https://doi.org/10.3389/fnins.2017.00485>
- McWalter, R., & McDermott, J. H. (2018). Adaptive and Selective Time Averaging of Auditory Scenes. *Current Biology*, 28(9), 1405-1418.e10. <https://doi.org/10.1016/j.cub.2018.03.049>
- Meredith, D. (Ed.). (2016). *Computational music analysis* (Vol. 62). Heidelberg: Springer.
- Miller, J. M., Chi, D. H., O'Keefe, L. J., Kruszka, P., Raphael, Y., & Altschuler, R. A. (1997). Neurotrophins can enhance spiral ganglion cell survival after inner hair cell loss. *International Journal of Developmental Neuroscience*, 15(4-5), 631-643. [https://doi.org/10.1016/S0736-5748\(96\)00117-7](https://doi.org/10.1016/S0736-5748(96)00117-7)
- Millman, R. E., Lorenzi, C., Apoux, F., Füllgrabe, C., Green, G. G., & Bacon, S. P. (2002). Effect of duration on amplitude-modulation masking. *The Journal of the Acoustical Society of America*, 111(6), 2551-2554.
- Milne, E. (2011). Increased Intra-Participant Variability in Children with Autistic Spectrum Disorders : Evidence from Single-Trial Analysis of Evoked EEG. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00051>
- Moore, B. C.J., & Raab, D. H. (1975). Intensity discrimination for noise bursts in the presence of a continuous, bandstop background: Effects of level, width of the bandstop, and duration. *The Journal of the Acoustical Society of America*, 57(2), 400-405.
- Moore, M. J., & Caspary, D. M. (1982). Binaural properties of neurons in the chinchilla lateral superior olivary nucleus. *The Journal of the Acoustical Society of America*, 71(S1), S99-S99.
- Moore, B. C. J., & Glasberg, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *The Journal of the Acoustical Society of America*, 74(3), 750-753. <https://doi.org/10.1121/1.389861>
- Moore, B. C. J., & Glasberg, B. R. (1986). The role of frequency selectivity in the perception of loudness, pitch and time. *Frequency selectivity in hearing*, 251-308.
- Moore, B. C. J., & Glasberg, B. R. (1987). Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. *Hearing Research*, 28(2-3), 209-225. [https://doi.org/10.1016/0378-5955\(87\)90050-5](https://doi.org/10.1016/0378-5955(87)90050-5)
- Moore, B. C. J. (1992). Across-channel processes in auditory masking. *Journal of the Acoustical Society of Japan (E)*, 13(1), 25-37. <https://doi.org/10.1250/ast.13.25>
- Moore, B. C.J. (Ed.). (1995). *Hearing*. Academic Press.
- Moore, B. C. J., Peters, R. W., & Glasberg, B. R. (1992). Detection of temporal gaps in sinusoids by elderly subjects with and without hearing loss. *The Journal of the Acoustical Society of America*, 92(4), 1923-1932. <https://doi.org/10.1121/1.405240>

- Moore, B. C. J., & Glasberg, B. R. (1996). A Revision of Zwicker's Loudness Model. 12.
- Moore, B. C.J., Wojtczak, M., & Vickers, D. A. (1996). Effect of loudness recruitment on the perception of amplitude modulation. *The Journal of the Acoustical Society of America*, 100(1), 481-489.
- Moore, B. C. J., & Sek, A. (1996). Detection of frequency modulation at low modulation rates : Evidence for a mechanism based on phase locking. *The Journal of the Acoustical Society of America*, 100(4), 2320-2331. <https://doi.org/10.1121/1.417941>
- Moore, B. C., Peters, R. W., & Glasberg, B. R. (1999). Effects of frequency and duration on psychometric functions for detection of increments and decrements in sinusoids in noise. *The Journal of the Acoustical Society of America*, 106(6), 3539-3552.
- Moore, B. C. J., & Glasberg, B. R. (2001). Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 110(2), 1067-1073. <https://doi.org/10.1121/1.1385177>
- Moore, B. C. J., & Skrodzka, E. (2002). Detection of frequency modulation by hearing-impaired listeners : Effects of carrier frequency, modulation rate, and added amplitude modulation. *The Journal of the Acoustical Society of America*, 111(1), 327-335. <https://doi.org/10.1121/1.1424871>
- Moore, B. C. (2007). *Cochlear hearing loss: physiological, psychological and technical issues*. John Wiley & Sons.
- Moore, D. R., Cowan, J. A., Riley, A., Edmondson-Jones, A. M., & Ferguson, M. A. (2011). Development of Auditory Processing in 6- to 11-Yr-Old Children. *Ear & Hearing*, 32(3), 269-285. <https://doi.org/10.1097/AUD.0b013e318201c468>
- Moore, B. C. J. (2012). *An introduction to the psychology of hearing*. Bingley: Emerald.
- Moore, B. C. (2014). *Auditory processing of temporal fine structure: Effects of age and hearing loss*. World Scientific.
- Moore, B. C., & Vinay. (2019). Effect of age on envelope regularity discrimination. *The Journal of the Acoustical Society of America*, 146(2), 1207-1211.
- Moore, B. C. J., Mariathan, S., & Şek, A. P. (2019). Effects of Age and Hearing Loss on the Discrimination of Amplitude and Frequency Modulation for 2- and 10-Hz Rates. *Trends in Hearing*, 23, 233121651985396. <https://doi.org/10.1177/2331216519853963>
- Moss, F., Ward, L. M., & Sannita, W. G. (2004). Stochastic resonance and sensory information processing: a tutorial and review of application. *Clinical neurophysiology*, 115(2), 267-281.
- Mozley, P. D., Acton, P. D., Barraclough, E. D., Plessi, K., Gur, R. C., Alavi, A., Mathur, A., Saffer, J., & Kung, H. F. (1999). Effects of Age on Dopamine Transporters in Healthy Humans. 6.
- Müller, M. (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications*. Springer.
- Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2002). Optimal methods for calculating classification images: Weighted sums. *Journal of Vision*, 2(1), 6-6.

- Nábělek, A. K., & Robinson, P. K. (1982). Monaural and binaural speech perception in reverberation for listeners of various ages. *The Journal of the Acoustical Society of America*, 71(5), 1242-1248
- Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: discrimination may be better than detection. *Vision research*.
- Nagaraja, N. S. (1964). Effect of Luminance Noise on Contrast Thresholds. *Journal of the Optical Society of America*, 54(7), 950. <https://doi.org/10.1364/JOSA.54.000950>
- Naka, K. I., & Rushton, W. A. (1966). S-potentials from colour units in the retina of fish (Cyprinidae). *The Journal of physiology*, 185(3), 536-555.
- Nelken, I., Rotman, Y., & Yosef, O. B. (1999). Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*, 397(6715), 154-157. <https://doi.org/10.1038/16456>
- Neri, P. (2010). How inherently noisy is human sensory processing?. *Psychonomic Bulletin & Review*, 17(6), 802-808.
- Nimmons, G. L., Kang, R. S., Drennan, W. R., Longnion, J., Ruffin, C., Worman, T., Yueh, B., & Rubinstein, J. T. (2008). Clinical Assessment of Music Perception in Cochlear Implant Listeners. *Otology & Neurotology*, 29(2), 149-155. <https://doi.org/10.1097/mao.0b013e31812f7244>
- Noble, W., Byrne, D., & Lepage, B. (1994). Effects on sound localization of configuration and type of hearing impairment. *The Journal of the Acoustical Society of America*, 95(2), 992-1005. <https://doi.org/10.1121/1.408404>
- Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. *Trends in Cognitive Sciences*, 23(11), 913-926. <https://doi.org/10.1016/j.tics.2019.08.004>
- Oertel, D., Bal, R., Gardner, S. M., Smith, P. H., & Joris, P. X. (2000). Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. *Proceedings of the National Academy of Sciences*, 97(22), 11773-11779. <https://doi.org/10.1073/pnas.97.22.11773>
- Osman, E. (1971). A correlation model of binaural masking level differences. *The Journal of the Acoustical Society of America*, 50(6B), 1494-1511.
- Ostroff, J. M., McDonald, K. L., Schneider, B. A., & Alain, C. (2003). Aging and the processing of sound duration in human auditory cortex. *Hearing Research*, 181(1-2), 1-7. [https://doi.org/10.1016/S0378-5955\(03\)00113-8](https://doi.org/10.1016/S0378-5955(03)00113-8)
- Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research*, 23(7), 689-699. [https://doi.org/10.1016/0042-6989\(83\)90210-9](https://doi.org/10.1016/0042-6989(83)90210-9)
- Oxenham, A. J., & Buus, S. (2000). Level discrimination of sinusoids as a function of duration and level for fixed-level, roving-level, and across-frequency conditions. *The Journal of the Acoustical Society of America*, 107(3), 1605-1614.
- Oxenham, A. J. (2016). Predicting the Perceptual Consequences of Hidden Hearing Loss. *Trends in Hearing*, 20, 2331216516686768. <https://doi.org/10.1177/2331216516686768>

- Paraouty, N., Ewert, S. D., Wallaert, N., & Lorenzi, C. (2016). Interactions between amplitude modulation and frequency modulation processing : Effects of age and hearing loss. *The Journal of the Acoustical Society of America*, 140(1), 121-131. <https://doi.org/10.1121/1.4955078>
- Paraouty, N., & Lorenzi, C. (2017). Using individual differences to assess modulation-processing mechanisms and age effects. *Hearing Research*, 344, 38-49. <https://doi.org/10.1016/j.heares.2016.10.024>
- Pardhan, S., Gilchrist, J., Elliott, D. B., & Beh, G. K. (1996). A comparison of sampling efficiency and internal noise level in young and old subjects. *Vision Research*, 36(11), 1641-1648. [https://doi.org/10.1016/0042-6989\(95\)00214-6](https://doi.org/10.1016/0042-6989(95)00214-6)
- Pardhan, S., & Gilchrist, J. (1991). The importance of measuring binocular contrast sensitivity in unilateral cataract. *Eye*, 5(1), 31-35.
- Pardhan, S. (2004). Contrast sensitivity loss with aging : Sampling efficiency and equivalent noise at different spatial frequencies. *Journal of the Optical Society of America A*, 21(2), 169. <https://doi.org/10.1364/JOSAA.21.000169>
- Parrish, E. E., Giaschi, D. E., Boden, C., & Dougherty, R. (2005). The maturation of form and motion perception in school age children. *Vision Research*, 45(7), 827-837.
- Patel, A. D. (2010). *Music, language, and the brain*. Oxford university press.
- Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *The Journal of the Acoustical Society of America*, 59(3), 640-654.
- Pelli, D. G. (1981). *Effects of visual noise* (Doctoral dissertation, University of Cambridge).
- Peter, V., Wong, K., Narne, V. K., Sharma, M., Purdy, S. C., & McMahon, C. (2014). Assessing Spectral and Temporal Processing in Children and Adults Using Temporal Modulation Transfer Function (TMTF), Iterated Ripple Noise (IRN) Perception, and Spectral Ripple Discrimination (SRD). *Journal of the American Academy of Audiology*, 25(02), 210-218. <https://doi.org/10.3766/jaaa.25.2.9>
- Peters, R. W., Moore, B. C. J., & Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. 11.
- Phillips, S. L., Gordon-Salant, S., Fitzgibbons, P. J., & Yeni-Komshian, G. (2000). Frequency and Temporal Resolution in Elderly Listeners With Good and Poor Word Recognition. *Journal of Speech, Language, and Hearing Research*, 43(1), 217-228. <https://doi.org/10.1044/jslhr.4301.217>
- Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International Journal of Audiology*, 42(sup2), 11-16. <https://doi.org/10.3109/14992020309074638>
- Pichora-Fuller, M. K., & MacDonald, E. (2008). Auditory temporal processing deficits in older listeners: A review and overview. In *Auditory signal processing in hearing-impaired listeners: Proceedings of the First International Symposium on Audiological and Auditory Research (ISAAR 2007)* (pp. 297-306). Holbaek, Denmark: Centertryk A/S.

- Picton, T. W., Vajsar, J., Rodriguez, R., & Campbell, K. B. (1987). Reliability estimates for steady-state evoked potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 68(2), 119-131. [https://doi.org/10.1016/0168-5597\(87\)90039-6](https://doi.org/10.1016/0168-5597(87)90039-6)
- Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., Gage, S. H., & Pieretti, N. (2011). Soundscape Ecology : The Science of Sound in the Landscape. *BioScience*, 61(3), 203-216. <https://doi.org/10.1525/bio.2011.61.3.6>
- Plack, C. J., & Oxenham, A. J. (1998). Basilar-membrane nonlinearity and the growth of forward masking. *The journal of the acoustical society of America*, 103(3), 1598-1608.
- Plack, C. J., Barker, D., & Prendergast, G. (2014). Perceptual Consequences of “Hidden” Hearing Loss. *Trends in Hearing*, 18, 233121651455062. <https://doi.org/10.1177/2331216514550621>
- Plomp, R., & Mimpen, A. M. (1979). Speech-reception threshold for sentences as a function of age and noise level. *The Journal of the Acoustical Society of America*, 66(5), 1333-1342. <https://doi.org/10.1121/1.383554>
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows : Cerebral lateralization as ‘asymmetric sampling in time’. *Speech Communication*, 41(1), 245-255. [https://doi.org/10.1016/S0167-6393\(02\)00107-3](https://doi.org/10.1016/S0167-6393(02)00107-3)
- Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science*, 256(5059), 1018-1021.
- Pollak, G. D., & Park, T. J. (1993). The effects of GABAergic inhibition on monaural response properties of neurons in the mustache bat’s inferior colliculus. *Hearing Research*, 65(1-2), 99-117. [https://doi.org/10.1016/0378-5955\(93\)90205-F](https://doi.org/10.1016/0378-5955(93)90205-F)
- Ponsot, E., Varnet, L., Wallaert, N., Daoud, E., Shamma, S. A., Lorenzi, C., & Neri, P. (2021). Mechanisms of spectrotemporal modulation detection for normal-and hearing-impaired listeners. *Trends in hearing*, 25, 2331216520978029.
- Pressnitzer, D. (2005). Music to Electric Ears : Pitch and Timbre Perception by Cochlear Implant Patients. *Annals of the New York Academy of Sciences*, 1060(1), 343-345. <https://doi.org/10.1196/annals.1360.050>
- Purcell, D. W., John, S. M., Schneider, B. A., & Picton, T. W. (2004). Human temporal auditory acuity as assessed by envelope following responses. *The Journal of the Acoustical Society of America*, 116(6), 3581-3593. <https://doi.org/10.1121/1.1798354>
- Rammsayer, T., & Classen, W. (1997). Impaired temporal discrimination in Parkinson's disease: temporal processing of brief durations as an indicator of degeneration of dopaminergic neurons in the basal ganglia. *International Journal of Neuroscience*, 91(1-2), 45-55.
- Rappaport, J. M., Gulliver, J. M., Phillips, D. P., Van Dorpe, R. A., Maxner, C. E., & Bhan, V. (1994). Auditory temporal resolution in multiple sclerosis. *The Journal of otolaryngology*, 23(5), 307-324.
- Rees, A., Green, G. G. R., & Kay, R. H. (1986). Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hearing research*, 23(2), 123-133.

- Resnik, J., & Polley, D. B. (2021). Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise. *Neuron*, 109(6), 984-996.
- Rice, S. O. (1973). Distortion Produced by Band Limitation of an FM Wave. *Bell System Technical Journal*, 52(5), 605-626. <https://doi.org/10.1002/j.1538-7305.1973.tb01981.x>
- Richards, V. M., Heller, L. M., & Green, D. M. (1991). The detection of a tone added to a narrow band of noise: The energy model revisited. *The Quarterly Journal of Experimental Psychology*, 43(3), 481-501.
- Rickards, F. W., & Clark, G. M. (1982). Steady-state evoked potentials in humans to continuous amplitude modulated tones. *The Journal of the Acoustical Society of America*, 72(S1), S54-S54.
- Robinson, D. E., & Berg, B. G. (1986). A partitioned variance model for multiple observations. In *Nineteenth Annual Meeting of the Society for Mathematical Psychology*, Cambridge, MA.
- Rodriguez, F. A., Chen, C., Read, H. L., & Escabi, M. A. (2010). Neural Modulation Tuning Characteristics Scale to Efficiently Encode Natural Sound Statistics. *Journal of Neuroscience*, 30(47), 15969-15980. <https://doi.org/10.1523/JNEUROSCI.0966-10.2010>
- Rose, A. (1948). The sensitivity performance of the human eye on an absolute scale. *JOSA*, 38(2), 196-208.
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 336(1278), 367-373.
- Ross, B., Fujioka, T., Tremblay, K. L., & Picton, T. W. (2007). Aging in Binaural Hearing Begins in Mid-Life: Evidence from Cortical Auditory-Evoked Responses to Changes in Interaural Phase. *Journal of Neuroscience*, 27(42), 11172-11178. <https://doi.org/10.1523/JNEUROSCI.1813-07.2007>
- Roß, B., Borgmann, C., Draganova, R., Roberts, L. E., & Pantev, C. (2000). A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *The Journal of the Acoustical Society of America*, 108(2), 679-691.
- Roß, B., Picton, T. W., & Pantev, C. (2002). Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hearing research*, 165(1-2), 68-84.
- Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, 2(5), 255-267.
- Russell, H. F., Wallis, D., Mazzocco, M. M., Moshang, T., Zackai, E., Zinn, A. R., ... & Muenke, M. (2006). Increased prevalence of ADHD in Turner syndrome with no evidence of imprinting effects. *Journal of pediatric psychology*, 31(9), 945-955.
- Saarinen, J., & Levi, D. M. (1995). Perceptual learning in vernier acuity: what is learned?. *Vision research*, 35(4), 519-527.

- Saberi, K., & Haftert, E. R. (1995). A common neural code for frequency- and amplitude-modulated sounds. *Nature*, 374(6522), 537-539. <https://doi.org/10.1038/374537a0>
- Saldanha, E. L., & Corso, J. F. (1964). Timbre Cues and the Identification of Musical Instruments. 6.
- Samson, S., Zatorre, R. J., & Ramsay, J. O. (1997). Multidimensional scaling of synthetic musical timbre: Perception of spectral and temporal characteristics. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 51(4), 307.
- Saunders, G. H. (1989). Determinants of objective and subjective auditory disability in patients with normal hearing (Doctoral dissertation, University of Nottingham).
- Scheidt, R. E., Kale, S., & Heinz, M. G. (2010). Noise-induced hearing loss alters the temporal dynamics of auditory-nerve responses. *Hearing Research*, 269(1-2), 23-33. <https://doi.org/10.1016/j.heares.2010.07.009>
- Schlittenlacher, J., & Moore, B. C. (2016). Discrimination of amplitude-modulation depth by subjects with normal and impaired hearing. *The Journal of the Acoustical Society of America*, 140(5), 3487-3495.
- Schmiedt, R. A., Mills, J. H., & Boettcher, F. A. (1996). Age-related loss of activity of auditory-nerve fibers. *Journal of Neurophysiology*, 76(4), 2799-2803. <https://doi.org/10.1152/jn.1996.76.4.2799>
- Schneider, B. A., Trehub, S. E., Morrongiello, B. A., & Thorpe, L. A. (1989). Developmental changes in masked thresholds. *The Journal of the Acoustical Society of America*, 86(5), 1733-1742.
- Schneider, B. A., & Trehub, S. E. (1992). Sources of developmental change in auditory sensitivity.
- Schneider, B. (1997). Psychoacoustics and aging: implications for everyday listening. *Journal of Speech-Language Pathology and Audiology*.
- Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K. (2002). Listening in aging adults : From discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 56(3), 139-152. <https://doi.org/10.1037/h0087392>
- Schoof, T., & Rosen, S. (2014). The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Frontiers in Aging Neuroscience*, 6. <https://doi.org/10.3389/fnagi.2014.00307>
- Sek, A., Baer, T., Crinnion, W., Springgay, A., & Moore, B. C. (2015). Modulation masking within and across carriers for subjects with normal and impaired hearing. *The Journal of the Acoustical Society of America*, 138(2), 1143–1153.
- Sek, A., & Moore, B. C. J. (1995). Frequency discrimination as a function of frequency, measured in several ways. *The Journal of the Acoustical Society of America*, 97(4), 2479-2486. <https://doi.org/10.1121/1.411968>
- Sekhar, S. C., & Sreenivas, T. V. (2004). Novel approach to AM-FM decomposition with applications to speech and music analysis. 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2, ii-753-756. <https://doi.org/10.1109/ICASSP.2004.1326367>

- Semal, C., & Demany, L. (2006). Individual differences in the sensitivity to pitch direction. *The Journal of the Acoustical Society of America*, 120(6), 3907-3915.
- Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G. (2013). Age-Related Cochlear Synaptopathy : An Early-Onset Contributor to Auditory Functional Decline. *Journal of Neuroscience*, 33(34), 13686-13694. <https://doi.org/10.1523/JNEUROSCI.1783-13.2013>
- Shadlen, Michael N., and William T. Newsome. "The variable discharge of cortical neurons: implications for connectivity, computation, and information coding." *Journal of neuroscience* 18.10 (1998): 3870-3896.
- Shafiro, V. (2008). Identification of Environmental Sounds With Varying Spectral Resolution. *Ear & Hearing*, 29(3), 401-420. <https://doi.org/10.1097/AUD.0b013e31816a0cf1>
- Shamma, S. (2001). On the role of space and time in auditory processing. *Trends in Cognitive Sciences*, 5(8), 340-348. [https://doi.org/10.1016/S1364-6613\(00\)01704-6](https://doi.org/10.1016/S1364-6613(00)01704-6)
- Shamma, S. A. (1985). Speech processing in the auditory system II : Lateral inhibition and the central processing of speech evoked activity in the auditory nerve. *The Journal of the Acoustical Society of America*, 78(5), 1622-1632. <https://doi.org/10.1121/1.392800>
- Shamma, S., & Lorenzi, C. (2013). On the balance of envelope and temporal fine structure in the encoding of speech in the early auditory system. *The Journal of the Acoustical Society of America*, 133(5), 2818-2833. <https://doi.org/10.1121/1.4795783>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: The university of Illinois Press.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. *Science*, 270(5234), 303-304. <https://doi.org/10.1126/science.270.5234.303>
- Sheft, S., & Yost, W. A. (1990). Temporal integration in amplitude modulation detection. *The Journal of the Acoustical Society of America*, 88(2), 796-805. <https://doi.org/10.1121/1.399729>
- Shu, Y., Hasenstaub, A., Badoual, M., Bal, T., & McCormick, D. A. (2003). Barrages of Synaptic Activity Control the Gain and Sensitivity of Cortical Neurons. *The Journal of Neuroscience*, 23(32), 10388-10401. <https://doi.org/10.1523/JNEUROSCI.23-32-10388.2003>
- Siebert, W. M. (1968). Stimulus transformations in the peripheral auditory system. *Recognizing patterns*, 104, 133.
- Siebert, W. M. (1970). Frequency discrimination in the auditory system : Place or periodicity mechanisms? *Proceedings of the IEEE*, 58(5), 723-730. <https://doi.org/10.1109/PROC.1970.7727>
- Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision research*, 49(22), 2705-2739.

- Singh, N. C., & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *The Journal of the Acoustical Society of America*, 114(6), 3394-3411. <https://doi.org/10.1121/1.1624067>
- Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416(6876), 87-90. <https://doi.org/10.1038/416087a>
- Snell, K. B., & Frisina, D. R. (2000). Relationships among age-related differences in gap detection and word recognition. *The Journal of the Acoustical Society of America*, 107(3), 1615-1626. <https://doi.org/10.1121/1.428446>
- Snell, K. B., Mapes, F. M., Hickman, E. D., & Frisina, D. R. (2002). Word recognition in competing babble and the effects of age, temporal processing, and absolute sensitivity. *J. Acoust. Soc. Am.*, 112(2), 8.
- Sperling, A. J., Lu, Z. L., Manis, F. R., & Seidenberg, M. S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature neuroscience*, 8(7), 862-863.
- Spiegel, M. F., & Green, D. M. (1981). Two procedures for estimating internal noise. *The Journal of the Acoustical Society of America*, 70(1), 69-73. <https://doi.org/10.1121/1.386583>
- Spilioti, M., Vargesson, N., & Neri, P. (2016). Quantitative assessment of intrinsic noise for visually guided behaviour in zebrafish. *Vision research*, 127, 104-114.
- Spear, P. D. (1993). Neural bases of visual deficits during aging. *Vision research*, 33(18), 2589-2609.
- Stamatakis, S., Francis, H. W., Lehar, M., May, B. J., & Ryugo, D. K. (2006). Synaptic alterations at inner hair cells precede spiral ganglion cell loss in aging C57BL/6J mice. *Hearing Research*, 221(1-2), 104-118. <https://doi.org/10.1016/j.heares.2006.07.014>
- Stapells, D. R., Linden, D., Suffield, J. B., Hamel, G., & Picton, T. W. (1984). Human auditory steady state potentials. *Ear and hearing*, 5(2), 105-113.
- Steeneken, H. J. M., & Houtgast, T. (1980). A physical method for measuring speech-transmission quality. *The Journal of the Acoustical Society of America*, 67(1), 318-326. <https://doi.org/10.1121/1.384464>
- Stickney, G. S., Assmann, P. F., Chang, J., & Zeng, F.-G. (2007). Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences. *J. Acoust. Soc. Am.*, 122(2), 10.
- Story, B. H. (2014). Structure, Movement, Sound, and Perception. *Perspectives on Speech Science and Orofacial Disorders*, 24(1), 7-20. <https://doi.org/10.1044/ssod24.1.7>
- Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J. Acoust. Soc. Am.*, 125(5), 18.
- Strouse, A., Ashmead, D. H., Ohde, R. N., & Grantham, D. W. (1998). Temporal processing in the aging auditory system. *The Journal of the Acoustical Society of America*, 104(4), 2385-2399. <https://doi.org/10.1121/1.423748>

- Sturr, J. F., Church, K. L., & Taub, H. A. (1988). Temporal summation functions for detection of sine-wave gratings in young and older adults. *Vision Research*, 28(11), 1247-1253. [https://doi.org/10.1016/0042-6989\(88\)90040-5](https://doi.org/10.1016/0042-6989(88)90040-5)
- Sueur, J., & Farina, A. (2015). Ecoacoustics : The Ecological Investigation and Interpretation of Environmental Sound. *Biosemiotics*, 8(3), 493-502. <https://doi.org/10.1007/s12304-015-9248-x>
- Sutcliffe, P., & Bishop, D. (2005). Psychophysical design influences frequency discrimination performance in young children. *Journal of Experimental Child Psychology*, 91(3), 249-270. <https://doi.org/10.1016/j.jecp.2005.03.004>
- Swaminathan, J., & Heinz, M. G. (2012). Psychophysiological Analyses Demonstrate the Importance of Neural Envelope Coding for Speech Perception in Noise. *Journal of Neuroscience*, 32(5), 1747-1756. <https://doi.org/10.1523/JNEUROSCI.4493-11.2012>
- Swets, J. A., Shipley, E. F., McKey, M. J., & Green, D. M. (1959). Multiple Observations of Signals in Noise. *The Journal of the Acoustical Society of America*, 31(4), 514-521. <https://doi.org/10.1121/1.1907745>
- Takahashi, G. A., & Bacon, S. P. (1992). Modulation Detection, Modulation Masking, and Speech Understanding in Noise in the Elderly. *Journal of Speech, Language, and Hearing Research*, 35(6), 1410-1421. <https://doi.org/10.1044/jshr.3506.1410>
- Talcott, J. B., Witton, C., Hebb, G. S., Stoodley, C. J., Westwood, E. A., France, S. J., Hansen, P. C., & Stein, J. F. (2002). On the relationship between dynamic visual and auditory processing and literacy skills; results from a large primary-school study. *Dyslexia*, 8(4), 204-225. <https://doi.org/10.1002/dys.224>
- Tanner Jr, W. P., & Birdsall, T. G. (1958). Definitions of d' and η as psychophysical measures. *The Journal of the Acoustical society of America*, 30 (10), 922-928.
- Theunissen, F. E., & Elie, J. E. (2014). Neural processing of natural sounds. *Nature Reviews Neuroscience*, 15(6), 355-366. <https://doi.org/10.1038/nrn3731>
- Thoret, E., Varnet, L., Boubenec, Y., F erriere, R., Le Tourneau, F.-M., Krause, B., & Lorenzi, C. (2020). Characterizing amplitude and frequency modulation cues in natural soundscapes : A pilot study on four habitats of a biosphere reserve. *The Journal of the Acoustical Society of America*, 147(5), 3260-3274. <https://doi.org/10.1121/10.0001174>
- Tolhurst, David J., J. Anthony Movshon, and Andrew F. Dean. "The statistical reliability of signals in single neurons in cat and monkey visual cortex." *Vision research* 23.8 (1983): 775-785.
- Tomko, George J., and Donald R. Crapper. "Neuronal variability: non-stationary responses to identical visual stimuli." *Brain research* 79.3 (1974): 405-418.
- Trainor, L. J., & Trehub, S. E. (1989). Aging and auditory temporal sequencing : Ordering the elements of repeating tone patterns. *Perception & Psychophysics*, 45(5), 417-426. <https://doi.org/10.3758/BF03210715>
- Trehub, S. E., Schneider, B. A., & Henderson, J. L. (1995). Gap detection in infants, children, and adults. *The Journal of the Acoustical Society of America*, 98(5), 2532-2541.

- Turin G. An introduction to matched filters. *IEEE Trans. Information Theory*. 1960; 6:311–329
- Van Tasell, D. J., Soli, S. D., Kirby, V. M., & Widin, G. P. (1987). Speech waveform envelope cues for consonant recognition. *The Journal of the Acoustical Society of America*, 82(4), 1152-1161. <https://doi.org/10.1121/1.395251>
- van Rooij, J. C. G. M., & Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. II : Multivariate analyses. *The Journal of the Acoustical Society of America*, 88(6), 2611-2624. <https://doi.org/10.1121/1.399981>
- Varnet, L., Ortiz-Barajas, M. C., Erra, R. G., Gervain, J., & Lorenzi, C. (2017). A cross-linguistic study of speech modulation spectra. *The Journal of the Acoustical Society of America*, 142(4), 1976-1989. <https://doi.org/10.1121/1.5006179>
- Velligan, D. I., Bow-Thomas, C. C., Mahurin, R., Miller, A., Dassori, A., & Erdelyi, F. (1998). Concurrent and predictive validity of the Allen Cognitive Levels Assessment. *Psychiatry research*, 80(3), 287-298.
- Verschooten, E., Shamma, S., Oxenham, A. J., Moore, B. C. J., Joris, P. X., Heinz, M. G., & Plack, C. J. (2019). The upper frequency limit for the use of phase locking to code temporal fine structure in humans : A compilation of viewpoints. *Hearing Research*, 377, 109-121. <https://doi.org/10.1016/j.heares.2019.03.011>
- Viana, L. M., O'Malley, J. T., Burgess, B. J., Jones, D. D., Oliveira, C. A. C. P., Santos, F., Merchant, S. N., Liberman, L. D., & Liberman, M. C. (2015). Cochlear neuropathy in human presbycusis : Confocal analysis of hidden hearing loss in post-mortem tissue. *Hearing Research*, 327, 78-88. <https://doi.org/10.1016/j.heares.2015.04.014>
- Vicente, T., Lavandier, M., & Buchholz, J. M. (2020). A binaural model implementing an internal noise to predict the effect of hearing impairment on speech intelligibility in non-stationary noises. *The Journal of the Acoustical Society of America*, 148(5), 3305-3317.
- Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *The Journal of the Acoustical Society of America*, 66(5), 1364-1380. <https://doi.org/10.1121/1.383531>
- Viemeister, N. F., & Wakefield, G. H. (1991). Temporal integration and multiple looks. *The Journal of the Acoustical Society of America*, 90(2), 858-865.
- Vilidaite, G., & Baker, D. H. (2017). Individual differences in internal noise are consistent across two measurement techniques. *Vision Research*, 141, 30-39. <https://doi.org/10.1016/j.visres.2016.10.008>
- Voelcker, H. B. (1966). Toward a unified theory of modulation part I : Phase-envelope relationships. *Proceedings of the IEEE*, 54(3), 340-353. <https://doi.org/10.1109/PROC.1966.4695>
- Vogels, R., Spileers, W., & Orban, G. A. (1989). The response variability of striate cortical neurons in the behaving monkey. *Experimental Brain Research*, 77(2), 432-436. <https://doi.org/10.1007/BF00275002>
- Voss, R. F., & Clarke, J. (1975). '1/fnoise' in music and speech. *Nature*, 258(5533), 317-318. <https://doi.org/10.1038/258317a0>

- Walker, B. A., Gerhards, C. M., Werner, L. A., & Horn, D. L. (2019). Amplitude modulation detection and temporal modulation cutoff frequency in normal hearing infants. *The Journal of the Acoustical Society of America*, 145(6), 3667-3674. <https://doi.org/10.1121/1.5111757>
- Wallaert, N., Moore, B. C. J., Ewert, S. D., & Lorenzi, C. (2017). Sensorineural hearing loss enhances auditory sensitivity and temporal integration for amplitude modulation. *The Journal of the Acoustical Society of America*, 141(2), 971-980. <https://doi.org/10.1121/1.4976080>
- Wallaert, N., Moore, B. C. J., & Lorenzi, C. (2016). Comparing the effects of age on amplitude modulation and frequency modulation detection. *The Journal of the Acoustical Society of America*, 139(6), 3088-3096. <https://doi.org/10.1121/1.4953019>
- Wallaert, N., Varnet, L., Moore, B. C. J., & Lorenzi, C. (2018). Sensorineural hearing loss impairs sensitivity but spares temporal integration for detection of frequency modulation. *The Journal of the Acoustical Society of America*, 144(2), 720-733. <https://doi.org/10.1121/1.5049364>
- Walsh, D. A. (1976). Age differences in central perceptual processing: A dichoptic backward masking investigation. *Journal of Gerontology*, 31(2), 178-185.
- Watson, C. S. (1963). Signal detection and certain physical characteristics of the stimulus during the observation interval. Indiana University.
- Weale, R. (1995). Why does the human visual system age in the way it does? *Experimental Eye Research*, 60(1), 49-55. [https://doi.org/10.1016/S0014-4835\(05\)80083-1](https://doi.org/10.1016/S0014-4835(05)80083-1)
- Whiteford, K. L., Kreft, H. A., & Oxenham, A. J. (2017). Assessing the Role of Place and Timing Cues in Coding Frequency and Amplitude Modulation as a Function of Age. *Journal of the Association for Research in Otolaryngology*, 18(4), 619-633. <https://doi.org/10.1007/s10162-017-0624-x>
- Whiteford, K. L., Kreft, H. A., & Oxenham, A. J. (2020). The role of cochlear place coding in the perception of frequency modulation. *ELife*, 9, e58468. <https://doi.org/10.7554/eLife.58468>
- Whiteford, K. L., & Oxenham, A. J. (2015). Using individual differences to test the role of temporal and place cues in coding frequency modulation. *The Journal of the Acoustical Society of America*, 138(5), 3093-3104. <https://doi.org/10.1121/1.4935018>
- Wible, B., Nicol, T., & Kraus, N. (2002). Abnormal neural encoding of repeated speech stimuli in noise in children with learning problems. *Clinical Neurophysiology*, 113(4), 485-494.
- Wiesenfeld, K., & Moss, F. (1995). Stochastic resonance and the benefits of noise : From ice ages to crayfish and SQUIDs. *Nature*, 373(6509), 33-36. <https://doi.org/10.1038/373033a0>
- Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child development*, 611-624.
- Wightman, F., & Allen, P. (1992). Individual differences in auditory capability among preschool children.

- Wingfield, A. (1996). Cognitive factors in auditory performance: context, speed of processing, and constraints of memory. *Journal of the American Academy of Audiology*, 7(3), 175-182.
- Witton, C., Green, G. G. R., Rees, A., & Henning, G. B. (2000). Monaural and binaural detection of sinusoidal phase modulation of a 500-Hz tone. *The Journal of the Acoustical Society of America*, 108(4), 1826-1833. <https://doi.org/10.1121/1.1310195>
- Wojtczak, M., & Viemeister, N. F. (1999). Intensity discrimination and detection of amplitude modulation. *The Journal of the Acoustical Society of America*, 106(4), 1917-1924. <https://doi.org/10.1121/1.427940>
- Won, J. H., Lorenzi, C., Nie, K., Li, X., Jameyson, E. M., Drennan, W. R., & Rubinstein, J. T. (2012). The ability of cochlear implant users to use temporal envelope cues recovered from speech frequency modulation. *The Journal of the Acoustical Society of America*, 132(2), 1113-1119. <https://doi.org/10.1121/1.4726013>
- Won, J. H., Shim, H. J., Lorenzi, C., & Rubinstein, J. T. (2014). Use of Amplitude Modulation Cues Recovered from Frequency Modulation for Cochlear Implant Users When Original Speech Cues Are Severely Degraded. *Journal of the Association for Research in Otolaryngology*, 15(3), 423-439. <https://doi.org/10.1007/s10162-014-0444-1>
- Woolley, S. M. N., Fremouw, T. E., Hsu, A., & Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nature Neuroscience*, 8(10), 1371-1379. <https://doi.org/10.1038/nn1536>
- Wright, C. E., & Drasdo, N. (1985). The influence of age on the spatial and temporal contrast sensitivity function. *Documenta Ophthalmologica*, 59(4), 385-395. <https://doi.org/10.1007/BF00159172>
- Wright, B. A., & Dai, H. (1998). Detection of sinusoidal amplitude modulation at unexpected rates. *The Journal of the Acoustical Society of America*, 104(5), 2991-2996.
- Wu, P. Z., Liberman, L. D., Bennett, K., de Gruttola, V., O'Malley, J. T., & Liberman, M. C. (2019). Primary Neural Degeneration in the Human Cochlea : Evidence for Hidden Hearing Loss in the Aging Ear. *Neuroscience*, 407, 8-20. <https://doi.org/10.1016/j.neuroscience.2018.07.053>
- Yost, W. A., Tanis, D. C., Nielsen, D. W., & Bergert, B. (1975). Interaural time vs. Interaural intensity in a lateralization paradigm. *Perception & Psychophysics*, 18(6), 433-440. <https://doi.org/10.3758/BF03204117>
- Zeng, C., Yang, Z., Shreve, L., Bledsoe, S., & Shore, S. (2012). Somatosensory Projections to Cochlear Nucleus Are Upregulated after Unilateral Deafness. *Journal of Neuroscience*, 32(45), 15791-15801. <https://doi.org/10.1523/JNEUROSCI.2598-12.2012>
- Zeng, F.-G., Nie, K., Liu, S., Stickney, G., & Chen, H. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues) (L). 4.
- Zeng, F.-G., Nie, K., Stickney, G. S., Kong, Y.-Y., Vongphoe, M., Bhargave, A., Wei, C., & Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proceedings of the National Academy of Sciences*, 102(7), 2293-2298. <https://doi.org/10.1073/pnas.0406460102>

- Zeng, F.-G., Oba, S., Garde, S., Sininger, Y., & Starr, A. (1999). Temporal and speech processing deficits in auditory neuropathy: *NeuroReport*, 10(16), 3429-3435. <https://doi.org/10.1097/00001756-199911080-00031>
- Zhang, L., Xie, S., Li, Y., Shu, H., & Zhang, Y. (2020). Perception of musical melody and rhythm as influenced by native language experience. *The Journal of the Acoustical Society of America*, 147(5), EL385-EL390. <https://doi.org/10.1121/10.0001179>
- Zhao, W., & Evans, E. F. (1990). Pharmacological microiontophoretic investigation of receptive-field and temporal properties of units in the cochlear nucleus. *Br. J. Audiol.*, 24, 193.
- Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, 12(5), 732-745. <https://doi.org/10.1111/j.1467-7687.2009.00817.x>
- Zimmermann, C. E., Burgess, B. J., & Nadol Jr, J. B. (1995). Patterns of degeneration in the human cochlear nerve. *Hearing research*, 90(1-2), 192-201.
- Zurek, P. M., & Formby, C. (1981). Frequency-Discrimination Ability of Hearing-Impaired Listeners. 5.
- Zwicker, E. (1952). Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones. *Acta Acustica United With Acustica*, 2(3), 125-133

Les informations de modulation d'amplitude (AM) et de modulation de fréquence (FM) lentes jouent un rôle crucial dans l'analyse des scènes auditives et la communication. Le but de ce programme de recherche doctorale est de caractériser le rôle de la variabilité neuronale (le « bruit interne ») dans le traitement auditif d'AM et de FM chez le sujet humain jeune et âgé normo-entendant et le sujet âgé malentendant à partir de mesures comportementales et de simulations informatiques. Les mesures comportementales sont obtenues avec le paradigme dit de « double-passe ». La modélisation informatique utilise soit un banc de filtres cochléaires suivi d'un banc de filtres de modulation, d'un bruit Gaussien additif et d'un mécanisme de décision par appariement de gabarit, soit un modèle mathématique dérivé de la théorie de la détection du signal (TDS). Cette approche combinant mesures psychophysiques et modélisation permet de préciser les mécanismes du traitement de l'AM et de la FM chez le sujet entendant et les effets du vieillissement et de lésions cochléaires sur le bruit interne dans le domaine de l'AM et de la FM. Trois études sont menées au cours de ce programme de recherche: i) une première étude portant sur un groupe de sujets normo-entendants jeunes ; ii) une seconde comparant deux groupes de sujets normo-entendants jeunes et âgés ; iii) une troisième comparant deux groupes de sujets normo-entendants jeunes et malentendants âgées. Pour chaque étude, des mesures sont réalisées à l'aide de la méthode des stimuli constants et d'une procédure en choix forcé afin d'estimer la capacité des sujets à détecter une AM et une FM sinusoïdales appliquées à un son pur de 500 Hz. La première étude utilise une cadence d'AM et de FM de 2 et 20 Hz, alors que les deux autres études n'utilisent qu'une cadence de 2 Hz. Un bruit de modulation de variance donnée est appliqué à tous les stimuli. Les signaux cibles d'AM et de FM sont présentés à trois niveaux de performance. Tous les stimuli utilisés au cours de la première session (la première « passe ») sont stockés et présentés à nouveau dans une deuxième session (la deuxième « passe ») après quelques minutes. Le pourcentage d'accord entre les réponses des sujets est calculé pour les essais correspondants dans lesquels les signaux sont identiques entre la première et de la deuxième passe. Le bruit interne des sujets est estimé à partir de ces données empiriques soit en ajustant les réponses du modèle informatique de traitement de modulation par modification de la variance du bruit Gaussien ajouté à la sortie des filtres de modulation (première étude), soit à l'aide d'un algorithme basé sur la TDS (deuxième et troisième études). Les données empiriques et de modélisation de la première étude portant sur les sujets normo-entendants jeunes indiquent qu'AM et FM - quelle que soit leur cadence - sont détectées par des mécanismes sensoriels distincts possédant leur propre source de bruit interne. Le modèle du banc de filtres de modulation permet de rendre compte des données de détection d'AM lente et rapide, confirmant sa capacité à offrir une approche unifiée de la perception d'informations d'enveloppe temporelle. Les résultats de la deuxième étude indiquent qu'avant l'âge de soixante ans, le bruit interne n'augmente pas dans des tâches de détection d'AM et de FM. Enfin, les résultats de la dernière étude indiquent qu'une perte neurosensorielle n'altère pas le bruit interne dans la détection d'AM et de FM. La variabilité neuronale associée au codage et à la détection de modulations temporelles lentes et rapides ne semble donc pas augmenter avec l'âge et à la suite de lésions cochléaires. Le paradigme de double passe - malgré des limites pratiques évidentes liées à la durée des tests - apporte une nouvelle mesure des capacités supraliminaires susceptible de compléter le diagnostic des pertes neurosensorielles.

MOTS CLÉS

modulation d'amplitude, modulation de fréquence, bruit interne, vieillissement, perte auditive.

Slow amplitude modulation (AM) and frequency modulation (FM) information plays a crucial role in auditory scene analysis and communication. The aim of this doctoral research programme is to characterise the role of neural variability ("internal noise") in the auditory processing of AM and FM in young and elderly normal-hearing and hearing-impaired humans using behavioural measurements and computer simulations. The behavioural measurements are obtained using the so-called "double-pass" paradigm. Computational modelling uses either a cochlear filter bank followed by a modulation filter bank, additive Gaussian noise and a template matching decision mechanism, or a mathematical model derived from signal detection theory (SDT). This approach, combining psychophysical measurements and modelling, makes it possible to specify the mechanisms of AM and FM processing in normal-hearing subjects and the effects of ageing and cochlear lesions on internal noise in the AM and FM domain. Three studies are being conducted in this research programme: i) a first study involving a group of young normal-hearing subjects; ii) a second comparing two groups of young and elderly normal-hearing subjects; iii) a third comparing two groups of young and elderly normal-hearing subjects. For each study, measurements are made using the constant stimulus method and a forced choice procedure to estimate the ability of subjects to detect sinusoidal AM and FM applied to a pure tone of 500 Hz. The first study used an AM and FM rate of 2 and 20 Hz, while the other two studies used a rate of only 2 Hz. A modulation noise is applied to all stimuli. The AM and FM target signals are presented at three performance levels. All stimuli used in the first session (the first "pass") are stored and presented again in a second session (the second "pass") after a few minutes. The percentage of agreement between subjects' responses is calculated for the corresponding trials in which the signals are identical between the first and second passes. The subjects' internal noise is estimated from these empirical data either by adjusting the responses of the computational model by modifying the variance of the Gaussian noise added to the output of the modulation filters (first study) or by using a SDT-based algorithm (second and third studies). The empirical and modelling data from the first study of young normal-hearing subjects indicate that AM and FM - regardless of their rate - are detected by distinct sensory mechanisms with their own internal noise source. The modulation filterbank model is able to account for both slow and fast AM detection data, confirming its ability to provide a unified approach to the perception of temporal envelope information. The results of the second study indicate that before the age of sixty, internal noise does not increase in AM and FM detection tasks. Finally, the results of the last study indicate that sensorineural hearing loss does not alter internal noise in AM and FM detection. The neural variability associated with the encoding and detection of slow and fast temporal modulations therefore does not appear to increase with age and following cochlear damage. The double-pass paradigm - despite obvious practical limitations related to test duration - provides a new measure of suprathreshold abilities that may complement the diagnosis of sensorineural hearing loss.

KEYWORDS

amplitude modulation, frequency modulation, internal noise, ageing, hearing loss.

