Intégration auditive des modulations temporelles : effets du vieillissement et de la perte auditive

Nicolas Wallaert

▶ **To cite this version:**

Nicolas Wallaert. Intégration auditive des modulations temporelles : effets du vieillissement et de la perte auditive. Sciences cognitives. Université Paris sciences et lettres, 2017. Français. NNT : 2017PSLEE087 . tel-02071797

HAL Id: tel-02071797
https://theses.hal.science/tel-02071797

Submitted on 18 Mar 2019

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 de recherche Paris Sciences et Lettres
PSL Research University

à l'Ecole normale supérieure

INTEGRATION AUDITIVE DES MODULATIONS TEMPORELLES :

Effets du vieillissement et de la perte auditive

Ecole doctorale n°158

ECOLE DOCTORALE CERVEAU COGNITION COMPORTEMENT (ED3C)

Spécialité Sciences Cognitives

COMPOSITION DU JURY :

M. SANTURETTE Sébastien
Technical University of Denmark,
Rapporteur

M. GRIMAULT Nicolas
Lyon Neuroscience Research Center
CRNL, Rapporteur

M. OXENHAM Andrew
University of Minnesota, Membre du jury

M. DEMANY Laurent
Université de Bordeaux, Membre du jury

M. MOORE Brian C. J.
University of Cambridge, Président du jury

Soutenue par Nicolas WALLAERT
le 28 novembre 2017

Dirigée par **Christian LORENZI**

ENS ÉCOLE NORMALE SUPÉRIEURE

PSL RESEARCH UNIVERSITY PARIS

Cette thèse de doctorat est le fruit de travaux de recherche ayant été réalisés au sein de :

L'Equipe 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 Prof. Christian Lorenzi.

Ce travail de thèse a été soutenu financièrement par un contrat de recherche industriel (société Neurelec - Oticon Medical) et par l'Agence Nationale de la Recherche (ANR) (ANR HEARFIN et ANR HEART). .

Tous les travaux impliquant des sujets malentendants ont été conduits en collaboration avec le Centre Hospitalier Universitaire (CHU) de Reims (France) et le laboratoire Auditis de Reims (France).

Résumé

Les signaux de communication, dont la parole, contiennent des modulations d'amplitude et de fréquence relativement lentes qui jouent un rôle capital dans l'identification et la discrimination des sons. Le but de ce programme de recherche doctorale est de comprendre plus finement les mécanismes impliqués dans la perception de l'AM et de la FM, et de clarifier les effets du vieillissement et de la perte auditive neurosensorielle sur ceux-ci. Les seuils auditifs de détection d'AM et de FM sont mesurés pour des sujets normo-entendants (NE) jeunes et âgés, ainsi que pour des sujets malentendants (ME) âgés, à basse fréquence porteuse (500 Hz) et à basse cadence de modulation (2 et 20 Hz). Le nombre de cycles de modulation, *N*, varie entre 2 et 9. Les seuils de détection de FM sont mesurés en présence d'une AM interférente, de façon à contraindre l'utilisation des indices d'enveloppe temporelle. Pour l'ensemble des groupes, les seuils de détection d'AM et de FM sont meilleurs à 2 qu'à 20 Hz. La sensibilité à l'AM et la FM s'améliore lorsque *N* augmente, démontrant une intégration temporelle des indices d'AM et de FM. Pour l'AM, les effets du vieillissement et de la perte auditive sont antagonistes: aux deux cadences de modulation, la sensibilité à l'AM décline avec l'âge, tandis qu'elle s'améliore en présence d'une perte auditive. L'intégration temporelle est similaire pour les deux groupes de NE, tandis que l'intégration temporelle est améliorée chez les sujets ME. Pour la sensibilité à la FM, l'âge dégrade sélectivement les seuils de détection de FM à basse cadence de modulation, tandis que la perte auditive a un effet délétère global aux deux cadences de modulation. L'intégration temporelle est similaire pour l'ensemble des groupes. Deux modèles computationnels (mono-bande et multi-bandes) utilisant un banc de filtres de modulation et un processus d'appariement de gabarit sont développés pour rendre compte des données. Pris ensemble, les données psychophysiques et de modélisation suggèrent que: 1) pour des cadences de modulation rapides, la détection d'AM et de FM sont encodés par un mécanisme commun, probablement basé sur les indices d'enveloppe temporelle. A l'inverse, à basse cadence de modulation, l'AM et la FM sont encodés par des mécanismes distincts, utilisant probablement et respectivement des indices d'enveloppe temporelle et de structure temporelle fine; 2) le vieillissement dégrade la sensibilité à l'AM et la FM (ie. les indices d'enveloppe temporelle et de structure temporelle fine), mais affecte plus fortement ces derniers; 3) la perte auditive n'affecte pas la sensibilité à l'AM (indices d'enveloppe temporelle), mais dégrade la sensibilité à la FM aux deux cadences de modulation; 4) Les processus décisionnels et mnésiques impliqués dans l'intégration temporelle d'AM et de FM sont préservés par le vieillissement. En présence d'une perte auditive, l'intégration temporelle d'AM est améliorée, probablement en raison de la perte de compression cochléaire, tandis que l'intégration temporelle en FM reste préservée. Toutefois, certains aspects de l'efficacité de traitement (modélisés par un bruit interne) déclinent avec l'âge et encore plus fortement avec une perte auditive. Les implications de ces resultats pour la définition, le diagnostic et la réhabilitation de la presbyacousie sont discutés.

Mots clés: enveloppe temporelle, intégration temporelle, modulation d'amplitude, modulation de fréquence, perte auditive, structure temporelle fine, vieillissement

Abstract

Communication sounds, including speech, contain relatively slow (<5-10 Hz) patterns of amplitude modulation (AM) and frequency modulation (FM) that play an important role in the discrimination and identification of sounds. The goal of this doctoral research program was to better understand the mechanisms involved in AM and FM perception and to clarify the effects of age and hearing loss on AM and FM perception. AM and FM detection thresholds were measured for young and older normal-hearing (NH) listeners and for older hearing-impaired (HI) listeners, using a low carrier frequency (500 Hz) and low modulation rates (2 and 20 Hz). The number of modulation cycles, *N*, varied between 2 to 9. FM detection thresholds were measured with and without an interfering AM to disrupt temporal-envelope cues. For all groups of listeners, AM and FM detection thresholds were lower for the 2-Hz than for the 20-Hz rate. AM and FM sensitivity improved with increasing *N*, demonstrating temporal integration for AM and FM detection. As for AM thresholds, opposite effects of age and hearing loss were observed: AM sensitivity declines with age, but improves with hearing loss at both modulation rates. Temporal integration of AM cues was similar across NH listeners, but better for HI listeners. As for FM sensitivity, ageing degrades FM thresholds at the low modulation rate only, whereas hearing loss has a deleterious effect at both modulation rates. Temporal integration of FM cues was similar across all groups. Two computational models (a single-band and a multi-band version) using the modulation filterbank concept and a template-matching decision strategy were developed in order to account for the data. Overall, the psychophysical and modeling data suggest that: 1) at high modulation rates, AM and FM detection are coded by a common underlying mechanism, possibly based on temporal-envelope cues. In contrast, at low modulation rates, AM and FM are coded by different mechanisms, possibly based on temporal-envelope cues and temporal-fine-structure cues, respectively. 2) Ageing reduces sensitivity to both AM and FM (i.e., both temporal-envelope and temporal-fine-structure cues), but more so for the latter. 3) Hearing loss does not affect sensitivity to AM (temporal-envelope cues) but impairs FM sensitivity at both rates. 4) The memory and decision processes involved in the temporal integration of AM and FM cues are preserved with age. With hearing loss, the temporal integration of AM cues is enhanced, probably due to the loss of amplitude compression, while the temporal integration of FM cues remains unchanged. Still, some aspects of processing efficiency (as modeled by internal noise) decline with age and even more following cochlear damage. The implications for the definition, diagnosis and rehabilitation of presbyacysis are discussed.

Key words: ageing, amplitude modulation, frequency modulation, hearing loss, processing efficiency, temporal envelope, temporal fine structure, temporel integration

Abréviations

AM: Modulation d'Amplitude

AMDL: Amplitude Modulation Detection Limens

ANOVA : Analysis Of Variance

BMLD = Binaural masking level difference

BW : Bandwidth

CASP : Computational Auditory Signal-processing and Perception

dB HL : dB Hearing Level

dB SL : dB Supra-Threshold

dB SPL : dB Sound pressure level

EFR : Envelope Following Response

ENV = Enveloppe Temporelle

EPSM : Envelope-Power-Spectrum Model

ERB : Equivalent Rectangular Bandwidth

Fc : Fréquence porteuse

FM: Modulation de fréquence

Fm = Cadence de modulation

FMDL = Frequency Modulation Detection Limens

FMFR : Frequency Modulation Following response

IPD : Différences interaurales de phase

LNN : Low noise Noise

MDI : interférence pour la détection de modulation

MTF: Fonction de transfert de modulation

PEMO : Perception Model

PTA : Pure-Tone Average

RMS : Root mean Square

STI : Speech transmission index

TFS : Temporal Fine Structure

Sommaire

Résumé **3**
Abstract **4**
Abréviations **5**
Sommaire **6**

CHAPITRE 1: INTRODUCTION

1. La parole comme signal modulé **9**
Une nouvelle dichotomie en audition 11

2. Rôle des indices d'AM et de FM dans la perception de parole **13**
2.1 Intelligibilité de la parole vocodée 15
2.1.1 Rôle des indices d'AM 15
2.1.2 Rôle des indices de FM 18
2.1.3 Limitations des études par vocodeurs 22
2.2 Etudes corrélationnelles: quels liens entre l'intelligibilité de la parole et la sensibilité à l'AM et à la FM? 23

3. Perception des indices d'AM et de FM chez le sujet normo- entendant **31**
3.1 Perception de l'AM 32
3.1.1 Effet des paramètres expérimentaux sur la MTF 34
3.1.2 Sélectivité fréquentielle du système auditif pour l'AM 37
3.1.3 Modèle computationnel du traitement de l'AM 41
3.1.4 Etudes développementales 47
3. 2 Perception de la composante de FM 50
3.2.1 Un double code ? 50
3.2.2 Modèle computationnel du traitement de la FM 58
3.2.3 Etudes développementales 60

4. Effet du vieillissement sur la perception de l'AM et de la FM **62**
4.1 Effet du vieillissement sur la détection d'AM 62
4.2 Effet du vieillissement sur la détection de FM 71

5. Effets d'une lésion cochléaire sur la perception de l'AM et de la FM **77**
5.1 Effet d'une perte auditive sur la détection d'AM 78
5.2 Effet d'une perte auditive neurosensorielle sur la détection de FM 85
5.3 Causes possibles d'une dégradation de la perception des indices temporels 91

6. Intégration temporelle **94**
6.1 Intégration temporelle de l'AM 94
6.1.1 Intégration en détection d'AM 94
6.1.2 Intégration en discrimination d'AM 101
6.1.3 Modélisation de l'intégration temporelle en AM 106
6.2 Intégration temporelle de la FM 117

7. Objectifs principaux et plan de thèse **124**
7.1 Problématique 124
7.2 Plan de thèse 126

CHAPITRE 2: EFFETS DU VIELLISSEMENT

SUR LA DETECTION D'AM ET DE FM

8. Comparing the effects of age on amplitude modulation and frequency modulation detection 130
8.1 Résumé: 130
8.2 Article: 130

CHAPITRE 3: EFFETS DE LA PERTE AUDITIVE

SUR LA SENSIBILITE ET L'INTEGRATION TEMPORELLE EN AM

9. Sensorineural hearing loss enhances auditory sensitivity and temporal integration for amplitude modulation 153
9.1 Résumé: 153
9.2 Article: 153

CHAPITRE 4: EFFETS DE LA PERTE AUDITIVE

SUR L'INTEGRATION TEMPORELLE EN AM ET EN FM

10. Differential effects of sensorineural hearing loss on temporal integration of amplitude- and frequency-modulation cues 178
10.1 Résumé: 178
10.2 Article: 178

CHAPITRE 5:DISCUSSION ET CONCLUSIONS

11. Discussion et conclusion 210
11.1. Synthèse sur le traitement de l'AM: 210
11.1.1. Effet de la cadence de modulation 210
11.1.2. Effet du nombre de cycles 211
11.1.3. Effet du vieillissement 211
11.1.4. Effet de la perte auditive 212
11.2. Synthèse sur le traitement de FM: 213
11.2.1. Effet de la cadence de modulation 213
11.2.2. Effet du nombre de cycles 214
11.2.1. Effet du vieillissement 214
11.2.1. Effet de la perte auditive 215
11.2.1. Ce que la modélisation nous a appris 215
11.3. Comparaison des traitements d'AM et de FM 216
11.4. Implications en audiologie clinique 220

12. Références 223

13. Annexe 242

CHAPITRE 1

INTRODUCTION

1. La parole comme signal modulé

Au cours des dernières décennies, de nombreuses études ont amené à considérer et ainsi étudier la parole comme un signal modulé (e.g., Plomp, 1983; Houtgast & Steeneken, 1973; Rosen, 1992)

Ainsi, Plomp (1983) et Rosen (1992) proposent de considérer la parole comme un signal complexe large bande modulé continuellement sur le plan temporel à trois niveaux différents, illustrés en Figure 1 : i) la fréquence vibratoire des cordes vocales varie constamment, produisant ainsi des modulations de hauteur fondamentale de la voix; ii) l'enveloppe temporelle du signal, c'est à dire l'ensemble des variations d'intensités relativement lentes est modulée en élargissant ou rétrécissant le tractus vocal localement, à l'aide de la langue et des lèvres; iii) la langue et les lèvres, en combinaison avec le pharynx, le larynx, les cavités nasales et buccales, déterminent le spectre de puissance de la parole, qui peut être considéré comme une modulation sur un axe fréquentiel audio.

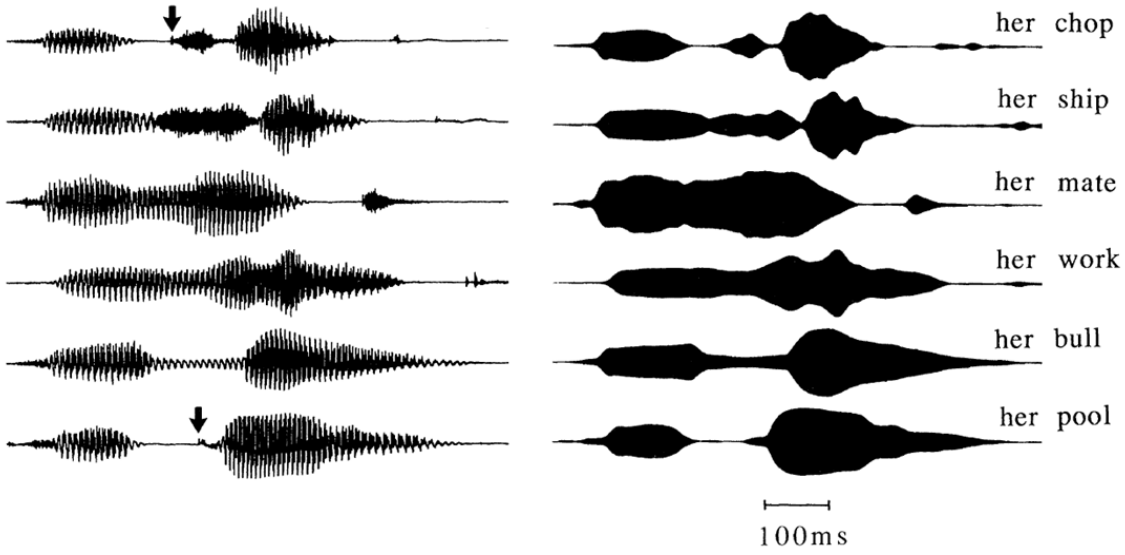


Figure 1. a. *Ondes du signal acoustique de 6 mots prononcés par un locuteur masculin; b. Ondes des signaux acoustiques précédents, après redressement et filtrage passe-bas à 20 Hz. Ce pré-traitement préserve les informations d'AM, en éliminant toutefois ses composantes hautes fréquences. Extrait de Rosen (1992).*

D'une façon analogue, Houtgast et Steeneken (1972, 1980, 1984) considèrent la parole comme un signal audio présentant des modulations d'amplitude (AM) saillantes, évoluant en fonction du temps et de la bande de fréquence audio considérée. Plus avant, Houtgast et Steeneken proposent de mesurer systématiquement la fonction de transfert de modulation (MTF) afin de caractériser tout système de transmission de parole (une prothèse auditive ou un auditorium, par exemple). Cette MTF compare les modulations d'amplitude du signal de parole original (l'entrée du système) à celles récupérées après diffusion de ce signal (la sortie du système) pour une situation acoustique donnée. Dans certaines conditions (e.g., réverbération, bruit ambiant, écrêtage, compression, etc.), les auteurs observent une atténuation importante (un lissage sur le plan temporel) des composantes d'AM, et ce particulièrement pour des cadences de modulation rapides. Houtgast et Steeneken (1984) proposent alors d'utiliser ces MTFs pour calculer un indice, le STI (Speech Transmission Index), quantifiant le degré global d'atténuation des modulations les plus saillantes du signal de parole et prédisant l'intelligibilité d'un signal de parole pour différentes conditions acoustiques.

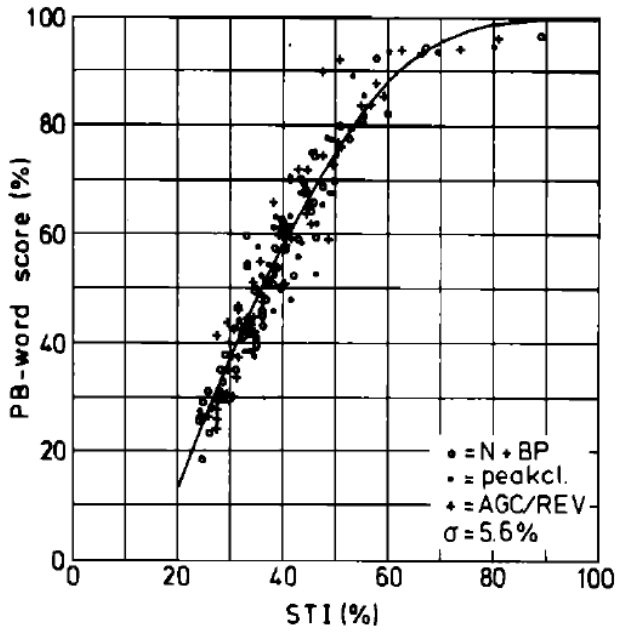


Figure 2. *Relation entre l'indice STI et les scores d'identification de mots pour différentes conditions expérimentales: bruit masquant, limitation de la bande passante, écrêtage, contrôle automatique de gain ou situations réverbérantes. La courbe de tendance représente l'ajustement le plus fin pour l'ensemble des données. Extrait de Steeneken et Houtgast (1980).*

Le pouvoir prédictif de l'indice STI est étudié en mesurant la relation liant les scores d'intelligibilité (performance d'identification ou seuil de réception de la parole) et les valeurs STI

obtenues pour un grand nombre de situations expérimentales. Dans une première étude dont les résultats sont représentés en Figure 2., Houtgast et Steeneken (1983) démontrent la validité de l'indice STI pour un grand nombre de conditions expérimentales (présence de bruit masquant, réverbération) et pour différentes populations de sujets (sujets normo-entendants, sujets malentendants). Ces travaux pionniers sont ensuite étendus et la portée de cette approche est précisée par de nombreuses études portant sur l'effet d'autres formes de distorsion du signal de parole produites par des vocodeurs ou des processeurs vocaux de prothèse (e.g., Drullman, 1995 ; Noordhoek & Drullman, 1997), des bruits masquants (e.g., Dubbelboer & Houtgast, 2007), des systèmes de rehaussement de la parole (e.g., van Buuren et al., 1999), ou encore des variations dans les conditions de production de la parole (e.g., Krause & Braida, 2004).

Pris ensemble, ces nombreux travaux suggèrent que: i) les indices de modulation sont l'un des traits essentiels du signal de parole, véhiculant des informations critiques pour l'identification de la parole; ii) la qualité de la transmission et réception de ces indices peut être utilisée pour prédire l'intelligibilité de la parole dans de nombreuses situations d'écoute.

Une nouvelle dichotomie en audition

En sortie des filtres cochléaires, les sons larges bandes tels que la parole sont décomposés en une série de signaux à bande étroite. Au vu de l'approche présentée ci-dessus, deux composantes élémentaires peuvent être extraites des signaux à bande étroite issus du filtrage cochléaire et utilisées par le système auditif central:

i. Les composantes d'AM correspondent aux fluctuations lentes de l'intensité de la parole, produite par l'alternance entre les parties fortement énergétiques et faiblement énergétiques du signal (Koenig et al., 1946; Miller et al., 1991; cf. Figure 3.)
ii. Les composantes de modulation de fréquence (FM) correspondent aux variations rapides de la fréquence instantanée du signal, induite par le fondamental laryngé (fo) et la structure harmonique du signal.

De nombreux travaux ont étudié l'importance respective des composantes d'AM et de FM de la parole (pour une revue récente, voir Shamma & Lorenzi, 2013). Cette analyse peut être menée à partir des spectres d'AM et de FM du signal de parole, représentant l'importance relative de chaque cadence de modulation d'amplitude ou de fréquence pour un corpus donné (Steeneken & Houtgast, 1980 ; Sheft et al., 2012). Les spectres d'AM et de FM s'avèrent être très similaires pour un grand nombre de langues (i.e., pour les langues syllabiques, à stress mais également et à ton ; Varnet et al., 2017 ; voir également Arai & Greenberg, 1997). L'analyse de ces spectres démontre également que les modulations temporelles de la parole sont dominées par des cadences relativement lentes, comprises entre quelques Hertz et quelques dizaines de Hertz. Dans le cas de la FM, les modulations les plus saillantes sont inférieures à 5-10 Hz. Ces modulations lentes semblent principalement liées à l'organisation rythmique (e.g., au débit syllabique dans le cas du Français) de la parole et syntaxique de la langue (Varnet et al., 2017).

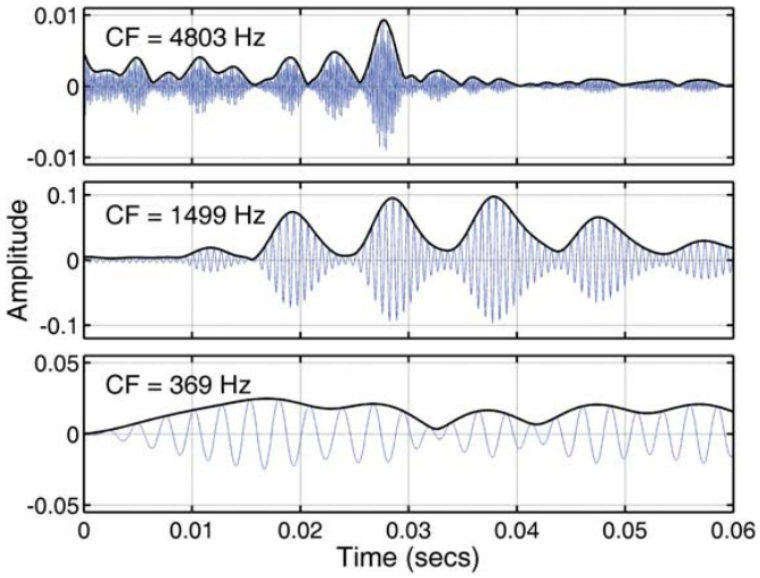


Figure 3. *Les courbes bleues représente l'onde acoustique en sortie de filtres passe bande centré à 369, 1499 et 4803 Hz. (de bas en haut) en réponse au phonème 'en'. La bande passante de chaque filtre est fixée de façon à égaler la bande passante des filtres auditifs (un ERB). Les formes d'ondes observées correspondraient ainsi à ce qui pourait être observé en différents points de la membrane basilaire. Pour chaque fréquence centrale, le signal peut être considéré comme une variation lente de son amplitude (AM ou enveloppe temporelle acoustique, traits noirs), et des variations plus rapides de sa fréquence instantanée (FM ou structure temporelle fine acoustique, traits bleus) - Extrait de Moore (2014).*

2. Rôle des indices d'AM et de FM dans la perception de parole

De nombreuses études ont démontré le rôle crucial joué par ces modulations temporelles dans le traitement robuste de la parole (e.g., Dudley, 1938; Licklider & Pollack, 1948; Schroeder, 1966; Flanagan, 1972, 1980; Rosen, 1992; Shannon et al., 1995; Drullman, 1995; Arai et al., 1996; Drullman et al., 1994). Les études les plus récentes suggèrent que les composantes d'AM et de FM, malgré la similarité de leurs spectres de modulation, véhiculent des indices complémentaires (e.g., Sheft et al., 2008). La figure 4 précise la terminologie utilisée dans la suite de ce travail de recherche.

La composante d'AM joue un rôle primordial dans l'intelligibilité de la parole (syllabes, mots ou phrases), et ce particulièrement dans le silence (e.g., Van Tasell et al., 1987; Shannon et al., 1995; Qin & Oxenham, 2003). Les premières études suggèrent que les indices de FM jouent principalement un rôle dans la perception des informations mélodiques (contour intonatif), l'intelligibilité des langues tonales ou encore dans la localisation spatiale (Smith et al., 2002 ; Xu & Pfingst, 2003). Toutefois, d'autres travaux plus récents suggèrent que, chez les sujets normo-entendants, les indices de FM jouent également un rôle important dans l'intelligibilité de la parole, mais ce principalement dans des situations d'écoute complexe, comme par exemple en présence de bruit de fond masquant, de parole concurrente ou d'autres formes de distorsion (e.g., filtrage sévère) conduisant à une diminution de la redondance du signal (Fullgrabe et al., 2006; Qin & Oxenham, 2003, 2006; Stickney et al. 2005; Lorenzi et al., 2006, 2009; Xu & Zheng, 2007).

Deux approches psychophysiques ont permis d'étudier la contribution respective des composantes d'AM et de FM pour la perception de la parole dans de multiples situations d'écoute (silence, bruit masquant, parole concurrente, brouhaha, réverbération, compression, écrêtage, etc.):

i. étude de l'intelligibilité de parole vocodée: différentes techniques de traitement du signal permettent de séparer les composantes d'AM et de FM de la parole. Pour des vocodeurs multi-bandes, le signal de parole est filtré par bande de fréquence. Les signaux à bandes étroites sont manipulés de façon à dégrader sélectivement les indices d'AM ou les indices de FM. Les signaux modulés issus de chaque bande sont ensuite recombinés, et l'intelligibilité du signal de parole vocodé est évaluée chez différents auditeurs.

ii. étude corrélationnelle: des études ont évalué la relation statistique entre l'intelligibilité de la parole mesurée dans différentes conditions d'écoute et la sensibilité auditive (e.g ., les seuils de détection) aux indices d'AM et de FM, évaluée au moyen de stimuli non linguistiques (des modulations sinusoïdales, des patterns de modulations complexes, des bruits de modulation).

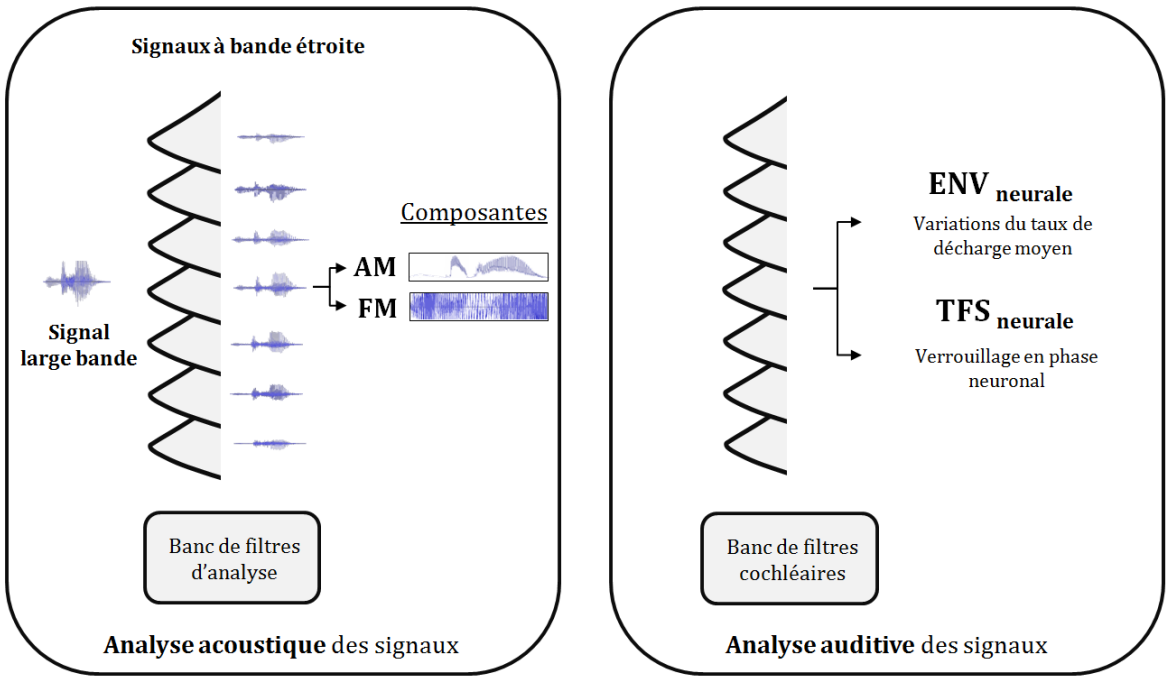


Figure 4. *Schématisation de la décomposition en sortie du banc de filtre d'analyse du signal acoustique en ses deux composantes élémentaires: l'AM et de FM. En sortie des filtres auditifs, deux grands indices sont utilisés pour l'analyse auditive des signaux: l'enveloppe temporelle (ENV) neurale et la structure temporelle fine (TFS) neurale.*

2.1 Intelligibilité de la parole vocodée

2.1.1 Rôle des indices d'AM

La composante d'AM d'un signal peut être extraite de différentes manières:

i. par redressement du signal puis filtrage passe bas avec une fréquence de coupure inférieure à la fréquence centrale de la bande (Rosen & Howell, 2010);
ii. par la transformée de Hilbert (Hilbert, 1904; Bracewell, 1986).

Pour des signaux à bande étroite, l'AM extraite par redressement et filtrage passe bas est similaire à l'enveloppe de Hilbert. Pour des signaux à bande large, la décomposition entre AM et FM est beaucoup plus complexe: l'AM n'est pas bien définie et peut être représentée à différentes échelles temporelles (Turner & Sahani, 2011). Pour des stimuli à bande large, les deux méthodes précédentes produisent des estimations très différentes de l'enveloppe temporelle . Toutefois, dans une oreille saine, un signal large bande est décomposé en une série de signaux à bandes étroites par le filtrage cochléaire (Patterson, 1976, 1982, 1995). Ainsi, il apparaît préférable de réaliser la décomposition d'un signal de parole en composantes d'AM et de FM à l'aide d'un banc de filtres sélectifs et étroits, respectant la sélectivité et la résolution fréquentielle cochléaire humaine (Moore, 2014).

Dudley (1939) fournit l'une des démonstrations les plus anciennes du fait qu'un signal de parole peut être parfaitement intelligible malgré une dégradation substantielle de son contenu spectral audio. Après avoir filtré la parole en 10 bandes de fréquence, l'AM est extraite par redressement et filtrage passe bas à 20 Hz. La parole est ensuite re-synthétisée à partir d'un signal d'excitation (un sifflement) filtré via le même banc de filtres, auquel est appliqué les informations d'enveloppe précédemment extraites. La parole re-synthétisée s'avère hautement intelligible. Dudley (1939) en conclut que la plupart des informations spectrales de la parole sont redondantes.

Plus récemment, Shannon et al. (1995) démontrent qu'une excellente intelligibilité (environ 90% d'identification de syllabes ou phrases) peut être obtenue en utilisant un nombre limité de bandes spectrales larges. Dans ce cas, l'intelligibilité ne repose que sur les seuls indices d'AM lents véhiculés par ces bandes spectrales . Plus précisément, les patterns d'AM (< 16-50 Hz) sont extraits au sein d'une à quatre bandes de fréquences, puis utilisés pour moduler un bruit de même bande passante, supprimant ainsi les informations spectrales fines dans chaque bande. Dorman et al. (1997), Loizou et al. (1999) et Lorenzi et al. (2006) obtiennent des résultats similaires pour des syllabes et phrases en utilisant des porteuses sinusoïdales (tone vocodeur).

Smith et al. (2002) étendent l'étude de Shannon et al. (1995) en évaluant l'importance perceptive respective des indices d'AM et de FM. Ils synthétisent des stimuli, appelés chimères auditives, présentant l'AM d'un son et la FM d'un autre son au sein de bandes de fréquences correspondantes entre les deux sons. Les scores d'identification de phrases en AM sont excellents pour un nombre de bandes d'analyses élevé (16)[1], et apparaissent robustes malgré l'interférence potentielle produite par la FM (cf. Figure 5). Smith et al (2002) en concluent que les indices d'AM sont capitaux pour l'identification de la parole, tandis que les indices de FM seraient principalement utilisés pour la localisation spatiale et la perception de la mélodie et du timbre.

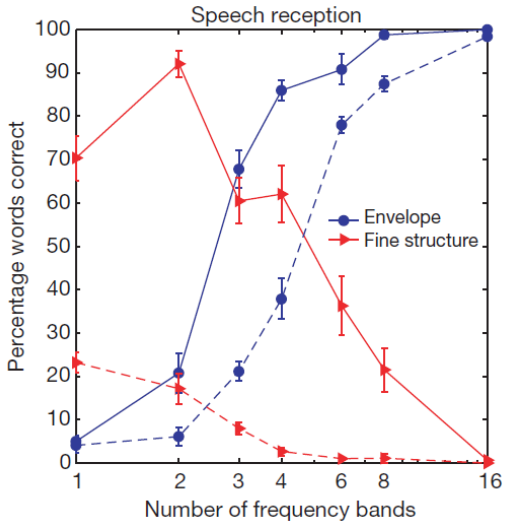


Figure 5. *Scores d'identification de phrases pour les chimères auditives contenant les composantes d'AM et de FM. Les lignes continues symbolise les scores d'identification obtenus pour les chimères auditives ne contenant que les composantes d'AM et de FM. Les lignes en pointillée représentent les scores d'identification pour les chimères contenant la composante d'AM d'un son et la composante de FM d'un autre son.*

Extrait de Smith et al. (2002)

[1] L'intelligibilité de la parole d'enveloppe augmente lorsque le nombre de bande de traitement de signal augmente (Drullman, 1995; Shannon et al., 1995; Loizou et al., 1999; Lorenzi et al., 2006).

Toutefois, Qin et Oxenham (2003) et Nelson et Jin (2004) montrent que la dégradation des indices de FM par vocodeur a un effet délétère plus important sur l'intelligibilité de phrases lorque les signaux de parole sont masqués par un bruit fluctuant plutôt qu'un bruit stationnaire. Ils en concluent que, malgré une résolution spectrale respectant la résolution fréquentielle humaine, les indices d'AM, à eux seuls, ne permettent pas de rendre compte de l'intelligibilité de la parole en présence de bruit fluctuant. D'autres études corroborent ce résultat, en montrant que l'intelligibilité d'un signal de parole initialement vocodé afin de ne conserver que les indices d'AM est fortement réduite en présence de bruit masquant ou de locuteur interférent (Nelson et al., 2003; Stone & Moore, 2003; Zeng et al., 2004). Ces résultats indiquent par ailleurs l'existence d'un effet de masquage de modulation entre les composantes d'AM du signal de parole cible et celles (parfois très saillantes) introduites par le bruit masquant en sortie des filtres cochléaires (Stone et al., 2011a; 2011b). Ce phénomène de masquage de modulation pourrait être responsable des différences d'intelligibilité produites par un bruit masquant stationnaire et fluctuant. Pris ensemble, ces résultats indiquent que les indices d'AM véhiculent des informations phonétiques et lexicales nécessaires à l'identification de la parole. Ils révèlent toutefois que les indices de FM véhiculent des informations complémentaires susceptibles de contribuer à la segmentation de la parole. Ces informations joueraient ainsi un rôle important dans la séparation de la parole du bruit masquant (Lunner et al., 2012; Moore et al., 2012).

Dans une série d'études par vocodeur, Hopkins et Moore (2008, 2010b) mesurent l'intelligibilité de phrases présentées dans le bruit et traitées de façon à ne conserver que les indices d'AM dans une région spectrale donnée, limitée aux basses ou hautes fréquences audio (le reste du signal demeurant intact, avec préservation des indices d'AM et de FM). Les auteurs font varier le nombre de bandes de fréquence pour lequel le vocodeur d'AM est utilisé (en d'autres termes, il font varier la fréquence de coupure à partir de laquelle les indices de FM sont à nouveau disponibles). Les résultats montrent que l'intelligibilité de la parole vocodée dans le bruit augmente à mesure que les indices de FM sont rendus disponibles, et ce que ce soit en basses fréquences ou en hautes fréquences. Hopkins et Moore (2010b) confirment le rôle des indices

d'AM en reconnaissance de phrases, mais concluent par ailleurs que les indices de FM jouent un rôle dans la capacité à identifier la parole en environnement bruyant (et plus précisément, en présence de locuteurs concurrents).

2.1.2 Rôle des indices de FM

La composante de FM peut être extraite par différents traitements du signal (compression infinie, clipping, ...). Pour des signaux à bande étroite, l'estimation de l'AM (par la transformée de Hilbert) est relativement bonne. Ainsi, la composante de FM du signal à bande étroite peut être estimée en divisant le signal par l'estimation de l'AM précédemment réalisée. Cette stratégie propose une estimation correcte de la FM, mais ceci uniquement lorsque l'AM est supérieure à zéro (Moore, 2014).

Les travaux de Licklider et Pollack (1948), Drullman et al. (1994a ; 1994b) et Smith et al (2002) ont déterminé successivement l'impact d'une suppression des indices d'AM sur l'intelligibilité de la parole tout en préservant, autant que possible, les indices de FM. Ces travaux démontrent que la suppression des indices d'AM n'affecte pas drastiquement l'intelligibilité de la parole dans le silence lorsque le nombre de bandes d'analyse du vocodeur est relativement faible (1-6 bandes d'analyse larges).

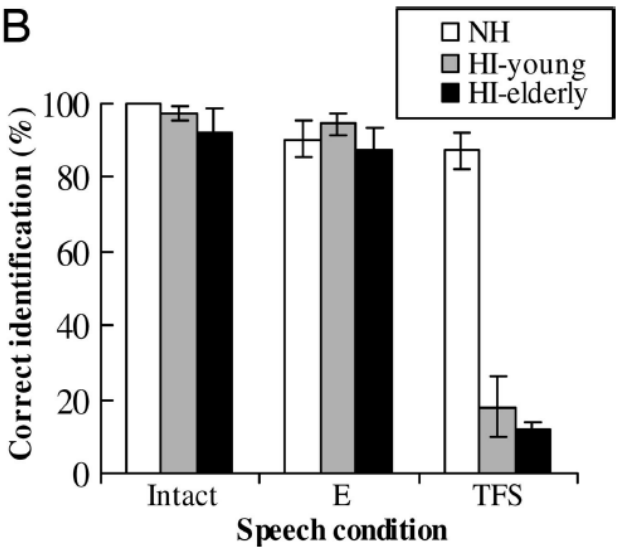
D'une manière similaire à Smith et al. (2002), Xu et Pfingst (2003) et Wang et al. (2011) évaluent l'importance relative des composantes d'AM et de FM pour l'intelligibilité d'une langue tonale, le mandarin. Contrairement aux résultats obtenus par Smith et al. (2002) pour une langue non tonale (l'anglais), ces deux études démontrent que l'intelligibilité de la parole dans le silence repose majoritairement sur les indices de FM.

De même, Gilbert et Lorenzi (2006) démontrent qu'une compréhension quasi parfaite de syllabes peut être obtenue sur la base des indices de FM quel que soit le nombre de bande d'analyse utilisé (1 à 16 bandes).

Lorenzi et al. (2006) et Sheft et al. (2008) mesurent l'intelligibilité de la parole vocodée de façon à préserver uniquement les indices d'AM ou de FM chez des sujets normo-entendants. Après entraînement, les sujets normo-entendants obtiennent des scores d'identification de syllabes relativement similaires pour les indices d'AM et de FM (cf. Figure 6), suggérant que les sujets normo-entendants peuvent également utiliser les indices de FM pour l'identification de la parole dans le silence en français, une langue à organisation syllabique (non tonale).

Figure 6. *Scores d'identification (en %) pour de la parole vocodée de façon à ne préserver que la composante d'AM, de FM ou contenant les deux composantes pour différents groupes de sujets (normo-entendants, malentendants jeunes et malentendants âgés. Les barres d'erreurs indiquent l'écart type pour chaque condition. Pour des sujets normo-entendants, les scores d'identification sont identiques pour la composante d'AM et de FM après entrainement.*

Extrait de Lorenzi et al. (2006)



Gilbert et al. (2007) évaluent l'intelligibilité des indices d'AM et de FM de syllabes extraits de 16 bandes, puis soumis à des interruptions périodiques (voir également Nelson & Jin, 2004 pour une manipulation similaire). Les signaux de parole d'AM et de FM s'avèrent compréhensibles pour l'ensemble des conditions expérimentales. Toutefois, l'effet des interruptions périodiques diffère sensiblement entre la parole d'AM et celle de FM, suggérant que l'AM et la FM véhiculent des indices phonétiques distincts.

L'estimation de la composante de FM d'un signal à bande étroite à partir de la division de ce signal par l'estimation de son enveloppe de Hilbert est plus problématique lorsque l'estimation de l'enveloppe est proche de 0. Ceci génère un changement de phase brusque et une variation importante de la fréquence instantanée (Stone et al., 2008; Hopkins et al., 2010) pour une partie du signal ayant initialement une importance modérée.

De façon à réduire ces excursions de fréquence dans le signal de FM à bande étroite, Hopkins et al. (2010) ajoutent un bruit de type « low noise noise » (LNN = bruit gaussien dont les phases sont modifiées de façon à ce que les fluctuations d'AM soient réduites ; Pumplin, 1985) dans chaque bande d'analyse. Pour un vocodeur utilisant 12 bandes d'analyse, l'intelligibilité du signal est significativement plus élevée lorsque le LNN est ajouté lors de la synthèse de la parole de FM, par comparaison avec de la parole de FM synthétisée sans LNN. Les auteurs en concluent que les indices de FM contribuent grandement à l'intelligibilité de la parole, lorsqu'une partie des artefacts introduits par l'extraction des indices de FM est masquée.

Toutefois, des travaux de modélisation réalisés par Ghitza (2001) suggèrent que les indices d'AM ayant été dégradés pour synthétiser le signal de parole de FM peuvent être reconstruits en sortie des filtres auditifs (cf. Figure 7), l'amplitude et la fréquence instantanée d'un signal n'étant pas indépendantes pour les signaux à bande large (Voelcker, 1996; Rice, 1973; Logan 1977).

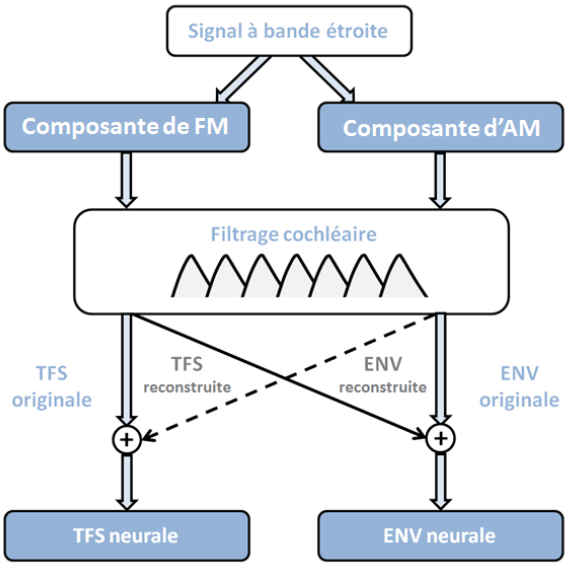


Figure 7. *La composante acoustique de FM peut être reconvertie en sortie du filtrage cochléaire en indices d'enveloppe neuraux, qui s'ajoutent aux indices de TFS neuraux. Cette conversion FM/AM survient en raison du filtrage à bande étroite cochléaire - Reproduit de Heinz & Swaminathan (2009)*

La contribution d'informations issues de la reconstruction des indices d'AM à partir des indices de FM en sortie des filtres auditifs a été confirmée par Zeng et al. (2004). Ces auteurs évaluent la contribution des indices d'AM reconstruits dans une tâche d'identification de parole. Dans un premiers temps, les indices d'AM de phrases sont supprimés en utilisant la décomposition de

Hilbert au sein d'une bande d'analyse large. Après filtrage du signal de FM précédemment obtenu par un banc de filtres cochléaires (banc de filtres dits "gammashirp"), les indices d'AM reconstruits à partir du signal de FM sont extraits, puis utilisés pour moduler en amplitude des bandes de bruit dont la bande passante est similaire à celle du signal d'origine. Les signaux résultants sont ensuite présentés à des sujets normo-entendants pour identification. Les indices d'AM reconstruits s'avèrent intelligibles (40% d'identification).

De nombreuses études ont ensuite tenté de préciser la contribution des indices d'AM reconstruits à partir des indices de FM à l'intelligibilité de la parole (e.g., Zeng et al., 2004; Gilbert & Lorenzi, 2006; Sheft et al., 2008; Ardoint & Lorenzi, 2010; Ardoint et al., 2010; Heinz & Swaminathan, 2009; Hopkins et al., 2010 ; Lorenzi et al., 2012; Won et al., 2014).

Ces différentes études suggèrent que:

a. La contribution des indices d'AM reconstruits en sortie des filtres auditifs dépend du rapport entre la bande passante des filtres d'analyse utilisés pour l'extraction des indices de FM et la bande passante des filtres auditifs, ie. les excursions dynamiques de la FM en dehors du filtre auditif sont converties en variation de l'amplitude du signal.
b. les indices de FM de la parole fournissent des informations phonétiques importantes, qui ne peuvent être uniquement la conséquence d'une reconstruction des indices d'AM.

Swaminathan et al. (2014) évaluent la contribution des indices d'AM reconstruits à partir du signal de parole de FM chez des auditeurs normo-entendants. L'intelligibilité est évaluée pour les indices de FM et les indices d'AM reconstruits (extraits avec une résolution importante, ie. 40 bandes). Les performances et confusions phonétiques des auditeurs sont similaires pour les deux conditions expérimentales. Ceci indique que les indices d'AM ne sont pas abolis dans la parole de FM et que l'identification de la parole de FM et d'AM repose sur des indices identiques, (voir également Léger et al., 2015, pour une démonstration similaire).

2.1.3 Limitations des études par vocodeurs

Les études utilisant des vocodeurs reposent sur un même principe : la dégradation sélective des indices acoustiques de FM ou d'AM au sein d'un certain nombre de bandes de fréquence d'analyse permet d'évaluer la contribution respective de chacune de ces deux composantes. Cette hypothèse d'indépendance de l'AM et de la FM n'est toutefois pas vérifiée, ce qui limite fortement l'interprétation des résultats de ces études.

Tout d'abord, les vocodeurs visant à dégrader sélectivement les indices de FM tout en préservant les indices d'AM se heurtent à deux limitations:

a. les indices de FM sont certes altérés, mais ils ne sont jamais totalement abolis de sorte qu'une information de FM, même appauvrie, peut affecter l'intelligibilité du signal. En effet, le signal de parole vocodé contient toujours une information de FM correspondant à celle véhiculée par la porteuse (e.g., des bandes de bruit).
b. pour des vocodeurs utilisant un bruit comme signal porteur, les indices d'AM initialement contenus dans le signal d'origine peuvent être masqués. En effet, les fluctuations aléatoires intrinsèques de l'amplitude du bruit s'ajoutent à l'AM du signal de parole cible et en gênent l'identification (Stone et al., 2008a)

De façon analogue, pour les vocodeurs visant à dégrader sélectivement les indices d'AM, la préservation des seuls indices de FM est contrainte par deux facteurs:

a. les indices de FM du signal d'origine peuvent être dégradés, notamment en raison d'artefacts tels que ceux produits par les excursions de fréquence instantanée produites dans les minima de l'enveloppe du signal d'origine.
b. les indices d'AM du signal d'origine sont certes altérés, mais ils peuvent être reconstruits par filtrage cochléaire sous certaines conditions.

Plus généralement, il apparaît que, lorsqu'un signal large bande tel qu'un signal de parole est décomposé par un banc de filtre passe bande suffisamment étroits (e.g., les filtres cochléaires), des informations d'AM peuvent être récupérées (reconstruites) à partir de la FM, et réciproquement, des informations de FM peuvent être récupérées à partir de l'AM (Heinz & Swaminathan, 2009; Sondergaard et al. 2012).

Dans une étude de modélisation, Shamma et Lorenzi (2013) évaluent dans quelle mesure la représentation centrale d'un signal de parole (son spectrogramme « interne » ou « auditif ») diffère lorsqu'elle est construite séparément à partir des indices d'AM, de FM, ou des deux informations combinées (AM+FM). Un réseau d'inhibition latéral (Shamma, 1985) est utilisé pour construire ces spectrogrammes auditifs à partir des signaux d'AM, de FM, et d'AM+FM. Les spectrogrammes auditifs construits à partir des indices de FM et d'AM sont similaires et sont corrélés de façon égale (corrélation de 0.7) avec le spectrogramme auditif construit à partir du signal AM+FM. Des résultats identiques sont obtenus pour différents types de vocodeur, variant par exemple par le nombre de bandes d'analyse. Les auteurs en concluent que l'AM et la FM véhiculent des informations phonétiques comparables et indépendantes du type de vocodeur utilisé.

2.2 Etudes corrélationnelles: quels liens entre l'intelligibilité de la parole et la sensibilité à l'AM et à la FM?

Une autre approche permettant d'apprécier la contribution relative des composantes d'AM et de FM à la perception de la parole consiste à évaluer la corrélation entre des mesures de sensibilité aux indices d'AM ou de FM et des mesures d'intelligibilité de la parole réalisées dans le silence ou en environnement bruyant.

Cazals et al. (1994) réalisent l'une des premières études estimant la corrélation entre scores d'intelligibilité et mesures de sensibilité aux indices de modulation temporelle. Ils mesurent la

capacité de sujets porteurs d'un implant cochléaire: i) à détecter une AM sinusoïdale en fonction de la cadence de modulation; ii) à identifier des voyelles et des consonnes. Une corrélation significative est observée entre la pente de la fonction reliant les seuils de détection d'AM (AMDLs) à la cadence d'AM (la fonction de transfert de modulation temporelle, Viemeister, 1979) et les scores d'intelligibilité chez les sujets implantés. Ces résultats suggèrent que la capacité à détecter une AM rapide (une mesure de la résolution temporelle du système auditif) conditionne la capacité des sujets implantés à utiliser l'information phonétique d'un signal de parole. Une étude réalisée par Fu (2002) confirme ce résultat, et obtient une corrélation significative bien plus importante : 71% et 97% de la variance des performances d'identification de voyelles et de consonnes sont respectivement expliquées par la sensibilité à une AM rapide de 100 Hz. Gnansia et al. (2014) étudient, également dans quelle mesure la capacité à identifier des consonnes et des voyelles dans le silence ou en présence de bruit stationnaire et fluctuant dépend de leur sensibilité à une AM lente de 8 Hz. Les performances d'identification sont modestement mais significativement corrélées à la sensibilité à l'AM dans le silence et dans un bruit stationnaire. Dans un bruit fluctuant, aucune corrélation n'est observée. En accord avec les études réalisées avec des vocodeurs, ces études corrélationnelles confortent l'idée selon laquelle les informations d'AM jouent principalement un rôle dans l'identification de parole dans le silence.

Buss et al. (2004) évaluent l'effet d'une perte auditive neurosensorielle sur la capacité à utiliser les indices d'AM et de FM et sur l'intelligibilité de parole intacte ou filtrée en passe bas avec une fréquence coupure de 1800 Hz. Ces auteurs mesurent l'intelligibilité de mots et les seuils de détection de FM (FMDLs) (Fm=2 Hz; fréquence porteuse Fc=0.5 ; 1 kHz). Les sujets malentendants sont déficitaires pour l'ensemble des tâches comparativement aux sujets normo-entendants. Chez les sujets malentendants, les scores d'intelligibilité et les FMDLs mesurés pour les deux porteuses sont significativement corrélés. Les auteurs concluent que: i) une lésion cochléaire est associée à une dégradation de la sensibilité aux indices de FM lents; ii) ce déficit temporel contribue au déficit d'intelligibilité observé chez les patients malentendants. Toutefois,

il est important de noter que cette étude présente deux limitations: i) la tâche de détection de FM est réalisée sans AM interférente: or, des indices d'AM peuvent être reconstruits à partir de la FM en sortie des filtres cochléaires; ii) la corrélation entre les seuils de détection de FM et les scores d'intelligibilité n'est obtenue que pour les seuils de détection de FM mesurés avec une porteuse de 1 kHz lorsque l'effet de la perte tonale (i.e., le seuils audiométriques tonaux à 0.5, 1 et 2 kHz) est contrôlé. Ces résultats suggèrent que ces deux mesures pourraient être influencées par un mécanisme commun, lié à la perte auditive per se.

Strelcyk et Dau (2009) réalisent la première étude examinant le lien entre intelligibilité et sensibilité aux indices de FM. Ils évaluent la relation entre la perception de la parole en présence de bruit (stationnaire ou fluctuant), la sélectivité fréquentielle (estimée à 750 Hz), différentes tâches supposées estimer les capacités de traitement temporel auditif [tâche de détection de FM à différentes fréquences porteuses (Fc=125; 750; 1500 Hz) et différentes cadences de modulation (Fm=2; 16 Hz); tâche de latéralisation (mesure de la sensibilité à des différences interaurales de phase : IPD); mesure du démasquage binaural (Binaural masking level difference : BMLD)].

Les auteurs font l'hypothèse que ces différentes tâches temporelles permettent d'évaluer préférentiellement la capacité du système auditif à encoder:

i) la composante de FM au niveau neural, dénommée structure temporelle fine (TFS) ; cette composante serait estimée par la tâche de détection de FM à basse porteuse et basse cadence de modulation (Sek & Moore, 1995; Moore & Sek, 1996) ainsi que les tâches de latéralisation et de démasquage binaural (Moore, 2014)

ii) la composante d'AM au niveau neural, dénommée enveloppe temporelle (ENV) ; cette composante serait estimée par la tâche de détection de FM en hautes fréquences audio ou aux basses fréquences audio mais ce pour une cadence de modulation élevée (Sek & Moore, 1995; Moore & Sek , 1996).

Six sujets normo-entendants et dix sujets malentendants présentant des seuils audiométriques tonaux normaux jusqu'à 1 kHz sont testés pour l'ensemble des tâches. Strelcyk et Dau (2009) observent que: i) les mesures binaurales (mesure de sensibilité aux indices de TFS) sont corrélées aux mesures monaurales de sensibilité à la FM; ii) les mesures reposant sur les indices de TFS et la sélectivité fréquentielle ne sont pas corrélées (en accord avec Hopkins et Moore, 2011), suggérant que le déficit temporel des sujets malentendants ne peut s'expliquer par l'élargissement des filtres cochléaires généralement associé à une perte auditive neurosensorielle; iii) les mesures d'identification de parole et de sensibilité aux indices de TFS sont corrélées, mais ce pour seulement certains des signaux interférents testés (à savoir, un mélange de deux voix) suggérant que la capacité auditive à séparer des signaux de parole concurrents est déterminée par la perception d'indices de TFS.

Cette étude présente également deux limitations: i) les effets d'âge et les effets de perte auditive sont confondus; ii) l'absence de corrélation entre les mesures de TFS et les seuils d'identification de parole pour les autres types de bruits reste inexpliquée.

Ruggles et al. (2011) mesurent l'intelligibilité de la parole en présence de sources de parole concurrentes et spatialisées pour différentes situations acoustiques (environnement anéchoïque ou réverbérant) chez 42 sujets âgés entre 18 et 55 ans. La capacité à détecter une FM lente est évaluée (Fm=2 Hz; Fc=750 Hz) pour les sujets du premier et du dernier quartile dans la tâche d'identification de la parole. L'intelligibilité en situation anéchoïque est négativement corrélée aux FMDLs. Ceci suggère que la capacité d'encodage des indices de FM contraint la capacité à suivre une source de parole en présence de sources de parole interférentes situées à des positions spatiales différentes.

Hopkins et Moore (2011) mesurent la sélectivité fréquentielle, la compréhension dans le bruit (stationnaire ou présentant des fluctuations spectro-temporelles) et la sensibilité aux indices de TFS évaluée par deux tests différents: i) le TFS1 - un test mesurant la capacité à discriminer une variation de hauteur causée par le décalage des composantes spectrales d'un son complexe

harmonique (Moore & Sek, 2009a, 2009b); ii) le "TFS-LF test" (Hopkins & Moore, 2010a) – un test évaluant la sensibilité aux changements de la phase interaurale d'un son pur de basse fréquence (Fréquence = 250, 500 et 750 Hz). Trois groupes de sujets sont testés: des sujets normo-entendants jeunes (20-35 ans), âgés (63-69 ans), et des sujets malentendants (d'âge variable). Les sujets âgés obtiennent des scores plus faibles aux deux tâches supposées mesurer la sensibilité auditive aux indices de TFS, malgré une sélectivité fréquentielle normale (en accord avec Lutman et al., 1991; Peters & Moore, 1992). Le vieillissement dégraderait donc la sensibilité à la TFS indépendamment de la perte audiométrique et de la sélectivité fréquentielle. Les scores d'intelligibilité dans le bruit sont corrélés aux seuils obtenus au test TFS1 mais pas aux seuils obtenus au test TFS-LF, lorsque l'on contrôle l'effet des seuils audiométriques. Hopkins et Moore (2011) en concluent que la dégradation des indices de TFS explique partiellement les difficultés de compréhension observées chez les sujets âgés et malentendants.

Neher et al. (2012) mesurent la sensibilité à la TFS (test TFS-LF; Fréquence = 250, 500, 750 Hz), l'intelligibilité de la parole dans le bruit (avec compensation de la perte d'audibilité) et les capacités cognitives de dix-sept auditeurs malentendants. Les scores obtenus au test TFS-LF sont corrélés avec l'âge, mais pas avec les seuils audiométriques, suggérant que l'âge a un effet délétère sur la perception des indices de TFS indépendamment de la perte audiométrique. Les seuils d'intelligibilité dans le bruit sont corrélés avec l'âge, avec les seuils TFS-LF, et avec les capacités cognitives des sujets. Ces résultats suggèrent que la dégradation de la perception des indices de TFS et des capacités cognitives pourrait expliquer, pour partie, les troubles d'intelligibilité en situation "cocktail party". Aucune corrélation n'est observée entre les scores d'intelligibilité et les seuils obtenus au test TFS-LF, ainsi qu'entre les scores d'intelligibilité et les capacités cognitives lorsque l'on contrôle les effets d'âge. Ces résultats suggèrent que les mesures de sensibilité à la TFS et les mesures des capacités cognitives sont influencées par un mécanisme commun, lié au vieillissement.

Mehraei et al. (2014) mesurent les capacités de huit sujets normo-entendants (24-60 ans) et douze malentendants (70-87 ans) à détecter des modulations spectro-temporelles (détection

modulation spectro-temporelles appliquée à une bande de bruit; Fréquence centrale = [500-4000] Hz; cadence des fluctuations temporelles = [4-32] Hz; densité d'ondulation spectrale = [0.5–4] cycles/octave). Ils montrent que la sensibilité aux modulations spectro-temporelles des sujets malentendants est, dans certaines conditions (basses cadences de modulation spectro-temporelle), corrélée à leurs seuils d'identification de parole, même lorsque l'effet de la perte auditive tonale est contrôlé.

Füllgrabe et al. (2015) mesurent la sensibilité aux indices de TFS (TFS1-test et TFS-LF test), la sensibilité aux indices d'ENV (tâche de détection d'AM), la perception de la parole (tâche d'identification de voyelles et consonnes réalisée dans le silence et dans le bruit) et les capacités cognitives (mémoire à court terme, mémoire de travail, attention, vitesse psychomotrice, recherche visuelle, raisonnement abstrait non verbal, perception spatiale et résolution de problèmes) chez des sujets normentendants jeunes (18-27 ans) et âgés (60-79 ans). La sensibilité à la TFS et à l'ENV est significativement réduite chez les sujets âgés. Par ailleurs, les seuils agrégés de TFS et ceux d'ENV sont significativement corrélés avec les scores d'identification de parole (cf. Figure 8). Ces résultats suggèrent que la sensibilité aux indices d'ENV et de TFS décline avec l'âge, et ce malgré l'absence de perte audiométrique, en accord avec les résultats de Moore et al. (2012). Par ailleurs, le déficit d'intelligibilité observé avec l'âge est similaire quel que soit le type de bruit masquant (stationnaire ou fluctuant), suggérant que l'utilisation des indices de TFS n'est pas spécifiquement associée à un déficit d'écoute dans les vallées du bruit, comme le suggéraient Lorenzi et al. (2006). Enfin, Füllgrabe et al. (2015) montrent que les meilleurs prédicteurs de l'intelligibilité sont les mesures des capacités cognitives et les mesures de sensibilité à la TFS.

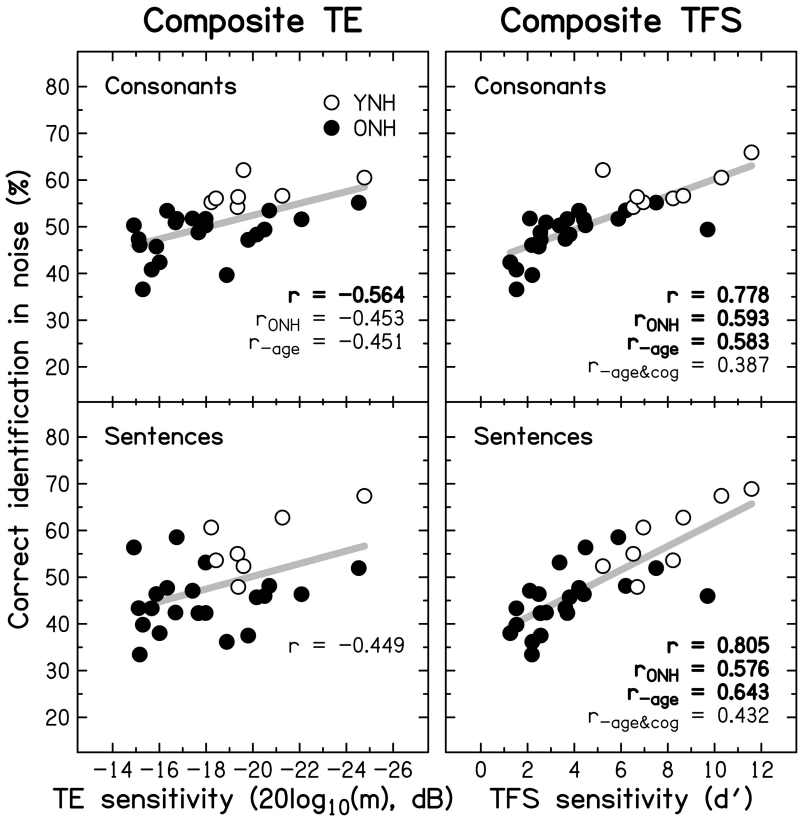


Figure 8 . *Représentation des seuils agrégés de TFS (figures de droite) et d'ENV (figures de gauche) en fonction des scores d'identification de consonnes (figures du haut) et de phrases (figures du bas). Les seuils agrégés de TFS et ceux d'ENV sont significativement corrélés avec les scores d'identification de parole - Reproduit de Füllgrabe et al.(2015).*

Johannesen et al. (2016) évaluent les liens entre la dysfonction de la mécanique cochléaire, le traitement temporel auditif (détection de FM avec AM interférente; Fc=1500 Hz; Fm= 2 Hz), l'âge et l'intelligibilité en environnement bruyant pour une large cohorte d'auditeurs (n=68 sujets) malentendants. Les analyses statistiques indiquent que d'une part, les seuils de détection de FM sont corrélés avec l'intelligibilité de la parole pour les deux types de signaux masquants testés (parole inversée temporellement et speech-shaped noise (SSN), ie. un bruit stationnaire présentant le spectre moyen à long terme de la parole) et d'autre part, l'intelligibilité se dégrade chez les personnes âgées.

En conclusion, malgré de nombreuses limitations, les résultats des études psychophysiques utilisant des vocodeurs afin d'estimer le rôle respectif des composantes d'AM et de FM de la parole suggèrent que ces indices contribuent tous deux à l'intelligibilité de la parole. La composante d'AM véhicule des informations permettant l'identification de la parole dans le silence, mais ne peut expliquer à elle seule les scores d'identification en environnement bruyant. La composante de FM semble contribuer à la capacité à séparer des signaux de parole concurrents. Le vieillissement et la perte auditive neurosensorielle affectent fortement la capacité à utiliser ces indices de FM. Les études corrélationnelles estimant la relation entre l'intelligibilité de la parole et la sensibilité aux indices d'AM et de FM confirment les conclusions précédentes. Elles suggèrent que la capacité à utiliser les informations de FM pour reconnaître la parole masquée ou filtrée est fortement affectée par le vieillissement et par la perte auditive neurosensorielle.

3. Perception des indices d'AM et de FM chez le sujet normo-entendant

De nombreuses recherches psychoacoustiques, neurophysiologiques ou de modélisation ont étudié les capacités de détection des indices d'AM et de FM chez l'homme et l'animal afin de caractériser les mécanismes de codage de l'information de modulation, à un niveau périphérique ou central. Ce chapitre porte principalement sur les études psychoacoustiques et de modélisation.

Les études psychophysiques utilisent différentes méthodes et paradigmes expérimentaux pour évaluer et dans certains cas, dissocier les mécanismes de codage de l'AM et la FM:

i. Etudes paramétriques: Ces études évaluent l'effet de la variation systématique d'un paramètre expérimental (fréquence porteuse, cadence de modulation, durée de stimulation, intensité de stimulation, ...) sur la perception de l'AM et la FM (e.g., Ernst & Moore, 2010; Jorgensen et al., 2013).

ii. Etudes des effets de masquage de modulation ou d'interférence: Ces études évaluent la perception d'une AM en présence d'une FM et vice versa (e.g., Sek & Moore, 1994; Paraouty et al., 2016).

iii. Etudes de populations: Ces études évaluent les capacités à percevoir les composantes d'AM et de FM chez différentes populations (comparaison de sujets jeunes et âgés, normo-entendants ou malentendants, sujets présentant une neuropathie auditive, ...): (e.g., Schoof & Rosen, 2014; Füllgrabe et al., 2015; Paraouty et al., 2016).

iv. Etudes développementales: Ces études caractérisent l'évolution des capacités perceptives de sujets depuis la naissance jusqu'à l'âge adulte. La majorité de ces études portent sur l'enfance (e.g., Hall & Grose, 1994; Buss et al., 1999; Lorenzi et al., 2000; Banai et al., 2011).

v. Etudes corrélationnelles: Ces études mesurent les seuils ou performances d'auditeurs dans différentes tâches auditives de détection ou discrimination de FM ou d'AM sur de grandes cohortes de sujets (e.g., Strelcyk & Dau, 2009; Ruggles et al., 2011; Bharadwaj et al., 2015; Whiteford & Oxenham, 2015; Otsuka et al., 2016; Paraouty & Lorenzi, 2017).

3.1 Perception de l'AM

La sensibilité humaine à l'AM a été étudiée de manière approfondie au cours des dernières décennies, et ce grâce à de multiples méthodes expérimentales, à savoir la mesure des capacités de: i) détection d'une interruption du signal (gap detection), ii) détection d'un incrément ou décrément d'intensité, iii) détection d'AM simple ou complexe. Ce chapitre se concentre sur les mesures de détection d'AM, à savoir des mesures de sensibilité à des variations dynamiques de l'intensité des sons.

Les mesures de sensibilité à l'AM diffèrent par le type de signal utilisé comme porteuse: i) un bruit blanc ou une bande de bruit modulée en amplitude (e.g., Viemeister, 1979 ; Eddins, 1999; Dau et al., 1997a, 1997b); ii) des sons complexes modulés temporellement (e.g., Yost & Moore, 1987; van Zanten & Senten, 1983) ; ou iii) une porteuse sinusoïdale (e.g., Zwicker, 1952; Zwicker & Kaiser, 1952; Zwicker, 1953; Zwicker & Graf, 1987 ; Maiwald, 1967; Paraouty et al., 2016). Le choix du type de porteuse n'est pas indifférent. En effet, l'application d'une AM à un bruit blanc ou à un bruit large bande n'affecte pas son spectre de puissance (audio) à long terme (e.g. Burns & Viemeister, 1981). A l'inverse, l'application d'une AM sur une porteuse sinusoïdale de fréquence Fc engendre des bandes spectrales latérales (de fréquence Fc-Fm et Fc+Fm lorsque l'AM est sinusoïdale et de cadence Fm). Ces bandes spectrales, lorsqu'elles sont résolues, c'est à dire lorsque la cadence de modulation est élevée et que les bandes spectrales latérales sont éloignées de la fréquence porteuse, peuvent contribuer à la détection de l'AM. Par ailleurs, une bande de bruit étroite véhicule des fluctuations aléatoires intrinsèques, à savoir des composantes d'AM saillantes dont la cadence est déterminée par la bande passante du bruit

(Lawson & Uhlenbeck, 1950) et pouvant engendrer des effets de masquage dans le domaine de l'AM (Dau et al., 1997a, 1997b).

Les mesures de sensibilité à l'AM diffèrent également par le type de signal de modulation (une sinusoïde, la somme de plusieurs sinusoïdes, une bande de bruit étroite). Généralement, une modulation sinusoïdale est imposée au signal porteur, de sorte que le signal modulé s'écrit sous la forme:

$$x(t) = [1 + m \cos (2\pi F_m t + Ø)] \text{ porteuse}(t)$$

avec : porteuse (t): la fonction caractérisant le signal porteur: bruit large bande de distribution gaussienne, bruit blanc uniforme, sinusoïde, somme de sinusoïdes, train de clics…

F_m: la cadence de modulation en Hz,

m: la profondeur de modulation en %,

Ø: la phase de modulation en radians ou degrés.

On définit la sensibilité à la modulation d'amplitude par la valeur minimale 20 log(1/m) permettant de discriminer un stimulus modulé en amplitude d'un stimulus non modulé.

La fonction représentant le seuil de détection d'AM (ou AMDL pour AM difference limen) en fonction de la cadence de modulation est dénommée fonction de transfert de modulation temporelle (TMTF ou MTF ; Viemeister, 1979).

De nombreux auteurs utilisent une bande de bruit comme signal porteur pour évaluer la MTF de sujets normo-entendants (e.g., Rodenburg, 1972, 1977; Viemeister, 1977, 1979; van Zanten, 1980; Bacon & Viemeister, 1985; Ozimek & Sek, 1988; Formby, 1985, 1987; Wakefield & Viemeister, 1990; Takahashi & Bacon, 1992; Bacon & Gleitman, 1992). Les MTFs mesurées chez des sujets normo-entendants correspondent approximativement à la fonction de transfert d'un filtre passe bas du premier ordre. En effet, la sensibilité à l'AM est relativement constante entre quelques Hertz et quelques dizaines de Hertz, puis décline (le seuil s'élève) au dessus de 50-100

Hz. La pente de la MTF est toutefois plus proche de 3-4 dB/octave que des 6 dB/octave d'un filtre du premier ordre (cf. Figure 9). La constante de temps τ du filtre passe bas, qui correspond au temps minimum d'intégration du système auditif, est déterminée par la relation $\tau = \frac{1}{2\pi fc}$, avec fc la fréquence de coupure à -3 dB du filtre passe bas. Pour des signaux larges bandes, la constante de temps τ est similaire (2-3ms) entre les différentes études.

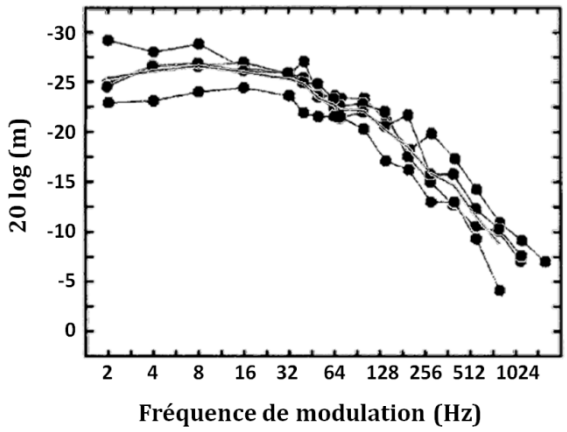


Figure 9. *AMDLs obtenus en fonction de la cadence de modulation pour une bande de bruit continue chez quatre sujets normo-entendants. Le niveau de puissance spectral est fixé à 30dB. Les seuils AMDLs moyens sont représentés par une ligne continue - La MTF présente une caractéristique passe bas, avec une pente d'environ -3dB/octave - Reproduit de Bacon & Viemeister (1985)*

3.1.1 Effet des paramètres expérimentaux sur la MTF

1/ Niveau de stimulation, fréquence porteuse et contribution des composantes latérales:

Pour des bandes de bruit modulées en AM et chez des sujets normo-entendants, les paramètres caractérisant la MTF (seuils, coupure et pente) s'avèrent indépendants du niveau de stimulation, sauf à très faible intensité (niveau spectral inférieur à 20 dB SPL) (Viemeister, 1979). Maiwald (1967) et Kohlrausch et al. (2000) comparent les MTFs obtenues pour une porteuse sinusoïdale et/ou pour des bandes de bruit. Lorsque le niveau de présentation augmente (20 à 100 dB SPL), la sensibilité à l'AM s'améliore fortement pour une porteuse sinusoïdale par comparaison avec un bruit.

Kohlrausch et al. (2000) mesurent les MTFs de six sujets normo-entendants pour des sinusoïdes porteuses comprises entre 1 et 10 kHz. Les AMDLs s'améliorent généralement à mesure que le niveau de présentation augmente, et ce malgré une grande variabilité interindividuelle. Pour une porteuse de 2 kHz ou plus, les AMDLs demeurent constants jusqu'à 100-150 Hz et se dégradent

(augmentent) au-delà avec une pente de 5-8 dB/octave: la MTF présente ainsi une caractéristique passe bas « typique ». Pour des cadences de modulation plus hautes (>150 Hz), l'AMDL décroit (la sensibilité s'améliore), reflétant la capacité des sujets à utiliser les composantes spectrales latérales créées par l'AM et résolues par les filtres cochléaires. Kolrausch et al. (2000) évaluent la contribution de ces composantes spectrales latérales à la détection de l'AM grâce à une méthode de masquage. L'asymétrie de masquage indique que lorsque les bandes latérales sont résolues par le système auditif, la bande latérale contribuant à la détection d'AM est la bande basse (celle centrée en dessous de la fréquence du signal porteuse). Les résultats indiquent également que lorsque les bandes latérales ne sont pas résolues, la détection d'AM est basée sur la sortie des filtres cochléaires centrés au dessus de la fréquence centrale du signal porteuse.

2/ Effet du type de bruit utilisé et puissance d'enveloppe du stimulus: Dau et al. (1999) ont mesuré les AMDLs pour différents types de porteuses modulées sinusoïdalement en amplitude (bruit gaussien, bruit de type LNN, ou bruit "multiplicatif" obtenu en multipliant une sinusoïde de 5 kHz par d'un bruit gaussien filtré en passe bas). Les AMDLs sont similaires pour le bruit gaussien et pour le bruit "multiplicatif", même si la sensibilité à l'AM mesurée avec le bruit "multiplicatif" est légèrement meilleure que celle mesurée avec le bruit gaussien . La sensibilité mesurée avec le bruit LNN est, quant à elle, bien meilleure que celle mesurée avec les deux autres types de bruit. Ces différences perceptives sont expliquées en termes de différences de puissance des fluctuations aléatoires intrinsèques de l'enveloppe temporelle dans la représentation auditive (c'est-à-dire, interne) des bruits gaussiens, multiplicatifs et LNN.

3/ Largeur de bande, fréquence centrale: Lorsque la bande passante du bruit utilisé comme signal porteur augmente, la fréquence de coupure de la MTF augmente. En d'autres termes, la constante de temps τ diminue pour atteindre un minimum de 2-3ms. De plus, les AMDLs s'améliorent à mesurent que la bande passante du bruit augmente (Rodenburg, 1972; Viemeister, 1979; Van Zanten, 1980; Formby & Muir, 1988). Dans l'ensemble des études citées ci-dessus, une augmentation de la bande passante s'accompagne systématiquement d'une

augmentation de la fréquence centrale du bruit, empêchant ainsi d'étudier la contribution respective de chacun des paramètres.

Néanmoins, Eddins (1993) étudie séparément l'impact de: i) la bande passante du bruit utilisé comme porteuse (BW $\in$ [200;1600] Hz) et; ii) de la fréquence centrale du bruit sur la MTF. Les résultats montrent que pour des bruits à bande étroite, la pente de la MTF est: i) relativement constante quelle que soit la région fréquentielle testée; ii) déterminée principalement par la bande passante du stimulus. Lorsque la bande passante du bruit diminue, la fréquence de coupure de la MTF diminue proportionnellement à la racine carré de la bande passante du stimulus. La sensibilité à la modulation augmente (les AMDLs s'améliorent) d'approximativement 3 dB lorsque la bande passante du bruit est doublée. Dau et al. (1997a, 1997b, 1999) démontrent que cet effet de la bande passante résulte de modifications du spectre de puissance d'enveloppe « interne » des signaux.

4/ Durée de stimulation, nombre de cycles: La durée du stimulus affecte la sensibilité à l'AM. Pour une cadence de modulation donnée, le nombre de cycles de modulation présentés dans l'intervalle d'observation diminue lorsque la durée du stimulus est réduite. L'intégration temporelle, c'est-à-dire la capacité du système auditif à bénéficier d'une augmentation du nombre de cycles de modulation, est plus largement décrite dans la section II de cette introduction.

Brièvement, les différents travaux ayant étudié l'effet de la durée du stimulus démontrent que: i) la sensibilité à l'AM se dégrade (les AMDLs s'élèvent) à mesure que la durée du stimulus diminue (Viemeister, 1979; Sheft & Yost, 1990; Viemeister & Wakefield, 1991; Dau et al., 1997a,b); ii) la durée de stimulation affecte essentiellement la sensibilité à l'AM aux cadences de modulation basses (en dessous de 250 Hz, mais ce principalement en dessous de 16 Hz environ) et pour les stimulus relativement brefs (< 500 ms) (Viemeister, 1979). Viemeister (1979) interprète ces résultats par le fait que lorsque la durée du signal diminue, le sujet a moins d'opportunités

d'échantillonner les fluctuations d'enveloppe temporelle du signal, en d'autres termes, le nombre d'observations indépendantes est réduit.

3.1.2 Sélectivité fréquentielle du système auditif pour l'AM

De nombreuses études psychoacoustiques ont suggéré l'existence de canaux de modulation sélectif chez l'homme (e.g., Kay & Matthews, 1972; Tansley & Suffield, 1983; Houtgast, 1989 ; Bacon & Grantham, 1989 ; Dau et al., 1997a, 1997b; Lorenzi et al., 1997 ; Richards et al., 1997; Wojtczak & Viemeister, 2003; Ewert & Dau, 2000). L'hypothèse de canaux de modulation est également attestée par l'existence d'une sélectivité neuronale à la cadence d'AM pour les neurones du tronc cérébral (e.g., le noyau cochléaire, le colliculus inférieur) et du cortex auditif primaire (pour une revue détaillée des données animales, voir Joris et al., 2004 ; pour une démonstration chez l'humain par neuro-imagerie (IRM-f) ou stéréo-électro-encéphalographie, voir Giraud et al., 2000 ; Liégeois-Chauvel et al., 2004). Une telle sélectivité fréquentielle dote le système auditif d'une capacité de décomposition (et de représentation) du signal à plusieurs échelles de temps. Elle permet de traiter indépendamment l'information temporelle au sein de chaque canal et ainsi, d'améliorer le rapport signal-sur-bruit interne en sortie de chaque canal, conduisant à une diminution des effets de masquage de modulation entre signaux concurrents (e.g., entre deux voix simultanées).

Deux paradigmes expérimentaux ont permis de corroborer l'existence de canaux d'AM: i) l'adaptation sélective (e.g., Kay & Matthews, 1972; Tansley & Suffield, 1983; Richards *et al.*, 1997; Wojtczak & Viemeister, 2003; voir toutefois Moody et al., 1984 et Wakefield & Viemeister, 1984); et ii) le masquage simultané (Bacon & Grantham, 1989; Houtgast, 1989 ; Strickland & Viemeister, 1996; Dau et al., 1997a, 1997b; Lorenzi et al., 1997).

Adaptation sélective: Il s'agit ici d'exposer un sujet à un signal modulé en amplitude (adaptateur) pendant une certaine durée (Fm et Fc fixes). L'AMDL du sujet est ensuite mesuré pour une cadence de modulation identique. Wojtczak et Viemeister (2003), Richards et al.

(1997) et Tansley et Suffield (1983) observent une dégradation de l'AMDL (Fc et Fm identiques à celles de l'adaptateur), consécutivement à une présentation extrêmement longue de l'adaptateur (<10 min). Cet effet (7-9 dB) est maximal vers 16 Hz. L'adaptation est généralement interprétée comme résultant d'une fatigue de neurones sélectivement accordés sur les paramètres du stimulus adaptateur.

L'adaptation est dite sélective dans le domaine des modulations (Kay, 1982) lorsque, suite à une période d'adaptation (à savoir, la présentation prolongée du stimulus adaptateur), le seuil de détection d'une modulation donnée (son test) augmente (i.e., se dégrade) à mesure que la distance (le long d'une dimension donnée) entre ce dernier et le stimulus adaptateur diminue.

Kay et Matthews (1972) mesurent les AMDLs et FMDLs (Fm = 8 Hz) en présence d'un stimulus adaptateur (AM ou FM, m=100%; Fc =250 Hz; Fm = 8 Hz). L'AMDL est multiplié par trois lorsque le signal d'AM test est présenté dans les 3 ou 4 secondes suivant le stimulus adaptateur d'AM ou de FM. En revanche, le FMDL n'est pas modifié lorsque le stimulus adaptateur est modulé en AM, mais reste multiplié par trois si le stimulus adaptateur est modulé en FM. Kay et Matthews (1972) démontrent ainsi l'existence d'une adaptation sélective tant en AM qu'en FM. Les effets rapportés sont également compatibles avec l'idée que la FM engendre des fluctuations d'AM « reconstruites » et perceptibles en sortie des filtres cochléaires, causant un effet d'adaptation de la FM sur l'AM. L'absence d'effet inverse (l'absence d'effet d'adaptation de l'AM sur la FM) suggère une forme d'indépendance entre les mécanismes de détection de l'AM et de la FM, qui sera discutée plus loin. Un effet d'adaptation similaire est également reporté par Tansley et Suffield dans des conditions expérimentales proches (m=10xAMDL; Fc =500 Hz; Fm =[2;32] Hz). Ces résultats pourraient suggérer l'existence de canaux auditifs sensibles à l'AM et à la FM opérant à un niveau central.

Certains résultats viennent nuancer l'existence d'une telle adaptation. En effet, Moody et al. (1984), Wakefield et Viemeister (1984) et Bruckert et al. (2006) ne retrouvent pas l'effet d'adaptation produit par l'AM sur les AMDLs lorsque les sujets ont été suffisamment entrainés

et/ou lorsqu'un bruit masquant est ajouté de façon à contraindre l'écoute hors fréquence. Les auteurs concluent que le phénomène d'adaptation est d'origine cognitive (attentionnelle) et non sensorielle. Alternativement, il est possible que l'effet d'adaptation soit réduit par apprentissage (la sélectivité à l'AM étant affectée par l'exposition prolongée à des signaux AM, cf. Joosten et al., 2016).

***Masquage simultané**:* L'existence de canaux de modulation a été essentiellement corroborée par les données issues d'expériences de masquage simultané dans le domaine de l'AM.

Masquage au sein d'une bande de fréquence: Bacon et Grantham (1989) mesurent les AMDLs (Fm=[2-512] Hz) d'un bruit large-bande en présence d'une AM masquante sinusoïdale (Fm=4, 16, 64 Hz; m=0.25, 0.5) chez trois sujets normo-entendants. L'effet de masquage est calculé par soustraction entre l'AMDL mesuré sans masque et l'AMDL mesuré en présence du masque. Globalement, le pattern de masquage présente un profil passe bande, l'effet de masquage étant d'autant plus grand que les cadences de modulation du signal et du masque sont proches. L'effet de masquage augmente lorsque la profondeur de modulation du masque augmente également, et culmine aux alentours de 10 dB.

Simultanément, Houtgast (1989), puis plus récemment, Strickland et Viemeister (1996), Lorenzi et al. (1997), Ewert et Dau (2000) et Ewert et al. (2002) répliquent ces données en utilisant une bande de bruit ou une sinusoïde comme porteuse de l'AM masque. Les patterns de masquage obtenus présentent également une allure passe bande, avec une bande passante relativement large, de l'ordre d'une octave.

Dans une autre série d'études de masquage simultané, Dau et al. (1997a, 1997b) mesurent les seuils de détection d'une AM sinusoïdale appliquée à une bande de bruit chez 5 sujets normo-entendants. La porteuse est une bande de bruit centrée sur 5 kHz, et sa bande passante est de 3, 31 ou 314 Hz. Dans tous les cas, la bande passante du bruit est plus faible ou proche de la bande passante du filtre cochléaire stimulé. Ainsi, toutes les composantes spectrales sont traitées au

sein d'un même canal audio. La détection d'AM repose alors principalement sur des indices temporels (des indices d'enveloppe temporelle) plutôt que sur des indices spectraux. Les données montrent que la forme de la MTF dépend de la bande passante du bruit. La MTF présente une caractéristique passe haut lorsque la bande passante du bruit est de 3 et 31 Hz, et une caractéristique passe bas lorsque la bande passante du bruit est de 314 Hz. Les AMDLs se dégradent ainsi pour des cadences de modulation spécifiques, correspondant à la bande passante du bruit. Ces effets de masquage sélectifs révèlent que le seuil de détection d'un signal d'AM masqué est principalement déterminé par le spectre de modulation de la porteuse (Lawson & Uhlenbeck, 1950), et plus précisément par son spectre de modulation « interne » ou auditif, tel qu'estimable en faisant l'hypothèse de filtres de modulation sélectifs et largement accordés.

Masquage entre bandes de fréquence distinctes: Simultanément aux travaux de Bacon et Grantham (1989) et Houtgast (1989), Yost et Sheft (1989) mesurent l'AMDL pour un son pur (Fc =1 kHz) auquel est appliqué une AM sinusoïdale (Fm= 10 Hz). Ils évaluent l'effet de masquage produit par l'adjonction d'un second son pur de fréquence distincte (Fc =4 kHz), également modulé en amplitude (Fm= 10 Hz; m=1). Malgré l'espacement fréquentiel des deux porteuses (les deux sons purs sont espacés d'au moins 2 octaves et les effets de masquage dit « énergétique » (i.e., cochléaire) entre ces deux sons sont ainsi minimes), l'AMDL est fortement dégradé (d'environ 10 dB) par l'adjonction de la seconde modulation. Cet effet de masquage est appelé MDI (interférence pour la détection de modulation). Yost et Sheft (1989) expliquent leurs résultats en supposant que les sorties de plusieurs bandes critiques sont groupées à un niveau central lorsqu'elles présentent des cadences de modulation similaires.

Yost et al. (1989) étendent leur paradigme initial en faisant varier la cadence de modulation du deuxième son pur. L'effet MDI diminue lorsque les cadences de modulation s'éloignent, de façon analogue aux données issues du masquage de modulation, suggérant une contribution de mécanismes centraux et sélectifs dans le domaine de l'AM.

Globalement, ces différents résultats révèlent que: i) la détection de modulation est sujette à des phénomènes d'interférence et/ou de masquage de nature non-énergétique (et donc, post-cochléaire) et; ii) l'existence d'une sélectivité fréquentielle pour l'AM. La largeur de la bande passante des filtres de modulation est estimée approximativement à 2-3 octaves chez Bacon et Grantham (1989) et une octave chez Houtgast (1989). Le facteur de qualité (Q) de ces filtres d'AM est finalement estimé à 1, voire moins (Ewert & Dau, 2000 ; Lorenzi et al., 2001)

3.1.3 Modèle computationnel du traitement de l'AM

Parallèlement aux mesures psychophysiques, de nombreux modèles computationnels ont été proposés de façon à: i) rendre compte de la résolution temporelle auditive et; ii) étudier séparément la contribution de chaque facteur aux performances de détection d'une AM.

Zwicker (1956, 1970) propose le premier modèle permettant de rendre compte des données d'AM. Dans la version dite « mono-bande » (single channel) de ce modèle, Zwicker postule que : i) les sujets calculent un « pattern d'excitation » représentant les variations d'activité sensorielle (e.g., neurale) en sortie des filtres auditifs périphériques en réponse aux signaux modulés, ii) l'AM engendre des fluctuations (expansion/réduction) en plusieurs points de ce pattern d'excitation ; et (iii) l'AM est détectée à partir des variations du niveau d'excitation dans la région haute fréquence du pattern. En effet, la pente basse fréquence du pattern d'excitation étant plus raide que sa pente haute fréquence, la magnitude des variations d'excitation induites par l'AM est plus importante sur la partie haute fréquence que sur la partie basse fréquence du pattern d'excitation. De plus, les filtres accordés aux fréquences plus élevées sont moins compressifs que celui accordé à la fréquence centrale (Zwicker, 1956, 1970; Robles et al., 1986; Nelson & Schroder, 1997; Moore & Oxenham, 1998). La détection d'AM est donc plus efficace lorsqu'elle est réalisée par les filtres dont la fréquence centrale est plus élevée que la fréquence porteuse du signal. La figure 10 schématise la théorie du modèle de pattern d'excitation, appliqué à l'AM.

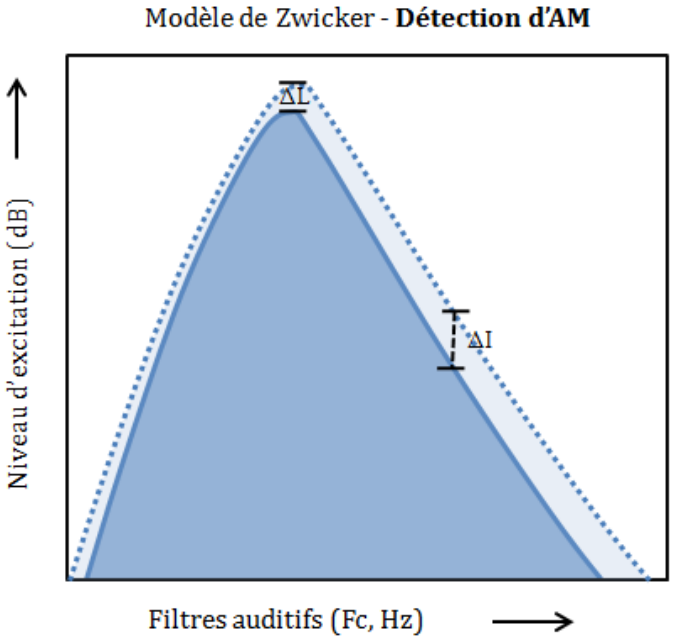


Figure 10. *Illustration schématique du modèle de pattern d'excitation appliqué à la détection d'AM. Les patterns d'excitation de deux sinusoïdes de même fréquence mais d'intensité différente sont représentés. La différence d'intensité, ΔI, peut, d'après le modèle de Zwicker, être détectable si le changement du niveau d'excitation (ΔL) dépasse une valeur seuil. Le changement du niveau d'excitation survient sur la partie haute fréquence du pattern d'excitation - Retracé à partir de Demany et Semal, 1986*

Florentine et Buus (1981) proposent une version « multi-bandes » (multi-channel) du modèle de Zwicker, combinant de manière optimale les informations issues de chacune des bandes critiques.

Plusieurs modèles computationnels de détection d'AM ont été développés à la suite de ces modèles « pionniers » mono et multi-bandes basés sur les patterns d'excitation, afin de mieux rendre compte des aspects dynamiques associés à la perception de patterns temporels produits par des AM simples et complexes. En général, ces modèles computationnels de détection d'AM comportent quatre étapes de traitement successives: i) un filtrage initial passe bande ou filtre de pré-détection; ii) un opérateur non-linéaire; iii) un intégrateur temporel (ou mécanisme de lissage temporel) ou un banc de filtres d'AM et; iv) un mécanisme de décision.

Dans le modèle de Viemeister (1979) ou modèle du détecteur linéaire d'enveloppe, le filtrage passe bande initial est analogue au filtrage cochléaire. L'opérateur non-linéaire rend compte du fait que les fibres du nerf auditif ne répondent qu'à la partie positive du signal acoustique (Rose et al., 1971). Il convertit le signal d'entrée en une grandeur purement positive, via une rectification mono-alternance. L'intégrateur temporel est modélisé par un simple filtrage passe

bas. La constante de temps et la pente du filtre passe-bas déterminent la résolution temporelle du système. La statistique de décision choisie par Viemeister (1979) repose sur un calcul de la variance du signal en sortie du filtre passe bas, sachant que la présence d'une AM accroit la variance de l'enveloppe du signal extraite par ce filtre.

Forrest et Green (1987) et Green et Forrest (1988) proposent un modèle similaire à celui de Viemeister, mais dont la statistique de décision correspond au rapport entre le maximum et le minimum (rapport max/min) en sortie du filtre passe bas. Comme la variance de l'enveloppe, le rapport max/min augmente avec la présence d'une AM. L'avantage principal de cette statistique de décision est qu'elle est indépendante du niveau de stimulation. Le modèle permet de prédire la pente de -3dB/octave de la MTF. Toutefois, cette statistique de décision a plus récemment été invalidée par Lorenzi et al. (1999), dans une expérience psychoacoustique de masquage, manipulant la phase relative des composantes d'AM (et de fait, le rapport max/min) d'un pattern de modulation complexe. D'autres statistiques de décision ont été proposées (le facteur de crête, l'asymétrie (skewness) ou la kurtosis de l'enveloppe (e.g., Strickland & Viemeister, 1996 ; Lorenzi et al., 1999). L'asymétrie de l'enveloppe semble mieux rendre compte des effets perceptifs produits par des changements de la phase relative des composantes d'une AM complexe (Lorenzi et al., 1999).

Dans les modèles précédents, la bande passante du filtre passe bande initiale est très largement supérieure à celle d'un filtre auditif (2 kHz pour Viemeister, 1979; 4 kHz pour Forrest et Green, 1987; 6 kHz pour Strickland et Viemeister, 1996). Ceci suggère que le filtrage périphérique n'a pas une influence considérable sur la sensibilité aux modulations d'amplitude et que le système auditif peut combiner (à un niveau central) des informations auditives en provenance de différents filtres cochléaires.

L'existence de canaux de modulation, suggérée initialement par les études d'adaptation et de masquage, est remise en cause par Strickland et Viemeister (1996). Leurs travaux de simulation basés sur le modèle du détecteur linéaire d'enveloppe (Viemeister, 1977, 1979) et sur la

statistique de décision max/min (Forrest & Green, 1987; Green & Forrest, 1988) permettent de rendre compte - tout au moins qualitativement - des courbes de masquage passe bande. Ces conclusions peuvent toutefois être invalidées par l'étude de Lorenzi et al. (1999), par les études psychoacoustiques de Dau et al. (1997a, 1997b) et par leurs résultats de simulation (Dau et al., 1996a, 1996b, 1997a, 1997b) d'un nouveau modèle basé sur l'hypothèse d'une sélectivité du système auditif à l'AM. Ce modèle est communément appelé PEMO (Perception Model). De façon similaire au modèle du détecteur linéaire d'enveloppe développé par Viemeister (1979), le modèle PEMO comporte plusieurs étapes de pré-traitement visant à extraire les informations d'enveloppe : i) un banc de filtres linéaires dits "gammatones" (Patterson et al., 1995) simulant le filtrage cochléaire; ii) une rectification mono alternance suivie d'un filtrage passe bas (détection d'enveloppe); iii) des circuits d'adaptation non linéaires simulant le phénomène d'adaptation nerveuse à court terme dans le nerf auditif.

Les fluctuations d'amplitude précédemment extraites sont ensuite analysées par un banc de filtres de modulation. Dans l'intervalle 0-10 Hz, les filtres de modulations ont une bande passante constante (5 Hz). Entre 10 et 1000 Hz, la bande passante est ajustée sur une échelle logarithmique avec un facteur de qualité Q initialement fixé à 2. Des résultats plus récents (Ewert & Dau, 2000; Ewert et al., 2002; Joosten et al., 2016; Lorenzi et al., 2001; Sek & Moore, 2002) suggèrent qu'une valeur de Q égale à 1 est plus adaptée pour simuler les données psychophysiques de masquage de modulation et de perception de patterns d'AM complexes. Point important, les informations de phase de l'AM sont supprimées en sortie du banc de filtre de modulation pour des cadences de modulation élevées (>10 Hz). Ceci est simulé en ne conservant que l'enveloppe de Hilbert de la sortie des filtres de modulation accordés au dessus de 10 Hz. Cette perte d'information temporelle est introduite dans le modèle pour reproduire les résultats d'une expérience psychoacoustique de mesure de la sensibilité à la phase de départ d'une AM sinusoïdale (Dau, 1996), montrant que les performances de détection d'un changement de phase d'AM s'effondrent au-delà de 5-10 Hz.

La statistique de décision diffère également de celle du détecteur linéaire d'enveloppe. La statistique de décision du modèle PEMO repose sur un simple calcul de corrélation, dénommé « appariement de gabarit » (template matching). Un gabarit est construit initialement pour une valeur supraliminaire du stimulus cible (généralement, pour les premiers stimuli d'un escalier psychophysique donné). Il s'agit de la représentation interne du signal "à détecter" qui est stockée en mémoire pour être ensuite comparée par corrélation croisée (cross corrélation) à la sortie du modèle pour chaque nouveau stimulus. Le gabarit est donc un vecteur de poids perceptifs, correspondant aux « attentes » du modèle dans la tâche de détection.

Lorsque la bande passante de la porteuse est plus élevée que la bande critique du filtre auditif, l'analyse du modèle est réalisée en parallèle au sein de différents filtres périphériques. A l'étape de décision, les sorties de l'ensemble des filtres de modulation sont combinées linéairement et pondérées de façon optimale. Le modèle permet de rendre compte du fait que: i) les constantes de temps associées aux MTFs pour des bandes de bruit ne varient pas avec la fréquence centrale de la bande; ii) la constante de temps associée à la MTF diminue lorsque la bande passante du signal augmente.

Pour intégrer l'information au cours du temps, une stratégie "d'observations multiples" est implémentée à l'étape de décision. Cette stratégie permet au modèle de rendre compte des constantes de temps relativement longues observées dans les expériences d'intégration temporelle (discutées plus loin), sans avoir à introduire une « réelle » intégration à long terme. La longue constante de temps effective résulte ici d'une combinaison de l'information entre différentes observations via un échantillonnage multiple et une sommation de probabilité.

Pour des signaux déterministes (e.g., des sinusoïdes), la sensibilité du modèle est limitée par l'adjonction d'un bruit interne indépendant à la sortie de tous les filtres de modulation et dans tous les filtres périphériques. Dans ce cas, la pente des filtres périphériques joue un rôle critique pour les stimuli dont la détection est basée sur l'audibilité des bandes spectrales latérales créées par la modulation.

Ce modèle prédit parfaitement les effets de bande passante d'un bruit faisant office de porteuse sur l'allure de la MTF, tels que décrits par Dau et al. (1997a, 1997b). Le modèle prédit également correctement l'allure passe bas de la MTF pour un bruit porteuse large bande et les effets de masquage d'AM rapportés par Bacon et Grantham (1989), Houtgast (1989) ou Strickland et Viemeister (1996).

Ewert et Dau (2000) proposent une version simplifiée de PEMO. Ce modèle, dénommé EPSM (Envelope-Power-Spectrum Model), utilise également un banc de filtre de modulation, mais se distingue du modèle PEMO par l'absence de filtrage périphérique et de processus simulant l'adaptation et surtout par la perte complète de l'information liée à la phase d'AM. L'hypothèse fondamentale de l'EPSM repose sur le fait que la détection d'AM est avant tout liée à la puissance totale de l'enveloppe à la sortie du filtre de modulation centré sur la cadence d'AM du signal. Pour déterminer un seuil de détection d'AM à partir de la puissance d'enveloppe, les auteurs postulent qu'une augmentation d'un dB de la puissance de l'enveloppe en sortie du filtre de modulation accordé à la cadence de modulation du signal est nécessaire pour détecter le signal de modulation.

En 2004, ces mêmes auteurs évaluent la contribution relative de la variabilité "externe" du signal (le bruit dit « externe ») et des limites de la résolution interne (le bruit dit « interne ») du système auditif pour la détection et la discrimination d'une AM. Ils comparent également les prédictions issues des modèles PEMO et EPSM dans différentes conditions expérimentales utilisant des bruits porteuses gelés (frozen noise) ou rafraichis pour chaque stimulus (random noise). Seul le modèle EPSM permet de rendre compte de l'ensemble des données mais ceci est principalement lié au fait que le modèle EPSM est initialement amélioré par l'adjonction d'un bruit interne multiplicatif, dont la variance est proportionnelle à la profondeur de l'enveloppe en sortie des canaux de modulation, afin de rendre compte de la loi de Weber observée pour la discrimination de profondeur d'AM (Wakefield & Viemeister, 1990 ; Ewert & Dau, 2004).

De nombreux autres modèles ont été proposés de façon à améliorer et/ou compenser les limites conceptuelles des modèles susnommés. Jepsen et al. (2008) reprennent le modèle PEMO, mais utilisent un banc de filtres non linéaires à double résonance (Meddis et al., 2001) en lieu et place du banc de filtres "gammatones", et ce de façon à rendre compte des non-linéarités de la membrane basilaire. Ils y ajoutent, entre autre, une modélisation de la fonction de transfert de l'oreille externe et moyenne. Cette version remaniée du modèle PEMO est dénommée CASP (Computational Auditory Signal-processing and Perception). De la même façon, Biberger et Ewert (2016) ou encore Jorgensen et al. (2013) proposent des variantes du modèle EPSM de façon à rendre compte de l'intelligibilité de la parole en présence de bruits fluctuants et/ou prédire les données psychophysiques pour un grand nombre de conditions expérimentales.

L'ensemble des modèles proposés (modèles perceptifs ou fonctionnels) ci-dessus sont basés sur le fait que la détection d'une modulation temporelle repose principalement sur une statistique de décision unique. Toutefois, d'autres modèles computationnels ont également été développés à partir de données neurophysiologiques (e.g., Siebert, 1965, 1970; Heinz et al., 2001a, 2001b, Colburn et al., 2003): ces modèles s'appuient sur l'hypothèse d'un double code pour le traitement de l'AM dans le système auditif, les fluctuations lentes étant encodées par le pattern de synchronisation des réponses neurales sur l'enveloppe temporelle, tandis que les fluctuations rapides sont encodées par une variation du taux de décharge moyen des neurones auditifs.

3.1.4 Etudes développementales

Différentes méthodes ont permis d'évaluer le traitement des indices d'AM et leur maturation au cours du temps. Pour des stimuli à bandes étroites, les seuils de détection d'un intervalle de silence (gap detection) chez l'adulte sont de l'ordre de 1 à 3 ms. Une moins bonne résolution temporelle est observée chez le jeune enfant de 3, 6, et 12 mois (e.g., Trehub et al., 1995; Werner et al., 1992). Toutefois, des mesures électrophysiologiques du seuil de détection d'un intervalle de silence indiquent une maturation plus précoce (survenant entre le 3ème et le 6ème mois) du traitement des indices d'enveloppe temporelle (Trainor et al., 2001; Werner et al., 2001). De

plus, une maturation précoce du traitement des indices d'enveloppe temporelle a également été démontrée pour des tâches de détection d'AM (e.g., Levi & Werner, 1996) ou pour des tâches de masquage (e.g., Levi & Werner, 1996; Werner, 1996, 1999). Une seule étude d'électrophysiologie a évalué la maturité du traitement de l'AM chez le jeune enfant (Levi et al., 1995). Les auteurs de cette étude suggèrent que le traitement des indices d'AM est immature chez le nourrisson d'un mois.

Les études réalisées chez l'enfant concluent que la résolution temporelle auditive évolue avec l'âge. Cependant, les résultats de ces études peuvent être fortement affectés par les méthodes utilisées pour évaluer la résolution temporelle. Chez l'enfant, les seuils de détection d'un intervalle de silence indiquent une maturation de la résolution temporelle vers l'âge de 5-7 ans (Diedler et al., 2007; Trehub et al., 1995; Wightman et al., 1989) tandis que les seuils obtenus pour une tâche de détection d'AM indiquent une maturation plus précoce, vers l'âge de 4 ans (e.g., Buss et al., 1999; Hall & Grose, 1994).

La contribution relative des facteurs sensoriels et des facteurs non sensoriels à cette maturation est encore sujet à débat (Werner, 2007). Hall et Grose (1994) mesurent la MTF chez des enfants âgés de 4 à 10 ans pour caractériser le développement de l'acuité temporelle auditive. Conformément aux données des sujets adultes du groupe contrôle, la sensibilité à l'AM décroit lorsque la cadence de modulation augmente. La fréquence de coupure de la MTF est similaire pour l'ensemble des groupes mais la sensibilité absolue à l'AM s'améliore progressivement avec l'âge. Hall et Grose (1994) en concluent que : i) l'encodage périphérique de l'AM est adulte dès l'âge de 4 ans (les constantes de temps entre enfant et adulte sont similaires); ii) l'*efficacité de traitement*, c'est-à-dire la capacité à utiliser de façon optimale les indices d'AM disponibles est réduite chez l'enfant de moins de 10 ans (la sensibilité absolue à l'AM est réduite chez l'enfant). Lorenzi et al. (2000) ainsi que Banai et al. (2011) confirment ces résultats et démontrent que la sensibilité absolue à l'AM s'améliore avec l'âge pour devenir identique à l'adulte vers l'âge de 11-12 ans.

Buss et al. (2006, 2008) montrent que les performances d'enfants d'âge scolaire dans une tâche de discrimination d'intensité sont limitées par le bruit interne (la pente des fonctions psychométriques de discrimination d'intensité est plus faible pour les enfants que les adultes). Buss et al. (2006, 2008) confirment ainsi l'hypothèse selon laquelle l'efficacité de traitement centrale de l'AM pourrait être réduite chez l'enfant, indépendamment de sa capacité à encoder une variation d'intensité au niveau périphérique.

En conclusion, il apparait que la résolution temporelle auditive est proche de celle observée chez l'adulte dès l'âge de 6 mois. Toutefois, la sensibilité à l'AM s'améliore entre la naissance et l'âge adulte, et le rôle des facteurs sensoriels et non sensoriels impliqués dans cette maturation du traitement d'AM est encore méconnu. Par ailleurs, les résultats des études développementales peuvent être affectés par la complexité de la tâche ou du son utilisé (Buss et al., 1999).

3. 2 Perception de la composante de FM

3.2.1 Un double code ?

La détection de FM chez les auditeurs normo-entendants a fait l'objet d'un grand nombre d'études utilisant des approches expérimentales différentes: études de balayage de fréquence (e.g., Sergeant & Harris, 1962; Pollack, 1968; Nabelek & Hirsh, 1969; Tsumura et al., 1973; Arlinger et al., 1977; Tyler et al., 1983; Schouten, 1985; Dooley & Moore, 1988; Madden & Fire, 1996) ou études de la détection de modulations sinusoïdales de fréquence (e.g., Hartmann & Hnath, 1982; Demany & Semal, 1989; Moore & Glasberg, 1989; Edwards & Viemeister, 1994a, 1994b; Sek & Moore, 1995; Moore & Sek, 1996).

Chronologiquement, deux théories ont successivement été proposées afin de rendre compte des capacités de détection et de discrimination de FM.

La première, la théorie du pattern d'excitation (également appelée "théorie de la place") se base sur le fait que des fréquences différentes mettent en mouvement différents points de la membrane basilaire, et stimulent ainsi des neurones ayant des fréquences caractéristiques différentes. La théorie du pattern d'excitation suggère que la fréquence d'un son est reliée au pattern d'excitation produit par ce son: pour un son pur, la fréquence est estimée à partir du pic d'excitation de la membrane basilaire. Lorsque la fréquence d'un son varie, le changement de fréquence pourrait être détecté en analysant le changement du niveau d'excitation en un point du pattern d'excitation (Zwicker, 1952, 1956; Maiwald, 1967; Siebert, 1968; Moore & Sek, 1994; Zwislocki & Nguyen, 1999) ou sur l'intégralité du pattern d'excitation (Moore & Sek, 1994).

Une théorie alternative, appelée théorie temporelle, suggère que la fréquence d'un son est estimée à partir de la distribution temporelle des potentiels d'action évoqués par ce son, et plus particulièrement par la TFS du signal à bande étroite en sortie des filtres cochléaires. En effet, les potentiels d'actions surviennent à une phase particulière de la structure temporelle fine de l'onde acoustique. Les intervalles entre deux potentiels d'actions sont approximativement des

multiples entiers de la période de l'onde acoustique, et peuvent donc être utilisés pour en inférer sa fréquence (Goldstein & Srulovicz, 1977; Wakefield & Nelson, 1985). Ce mécanisme est appelé verrouillage en phase ou "phase locking" sur la TFS du signal à bande étroite en sortie des filtres cochléaires (cf. Figure 11).

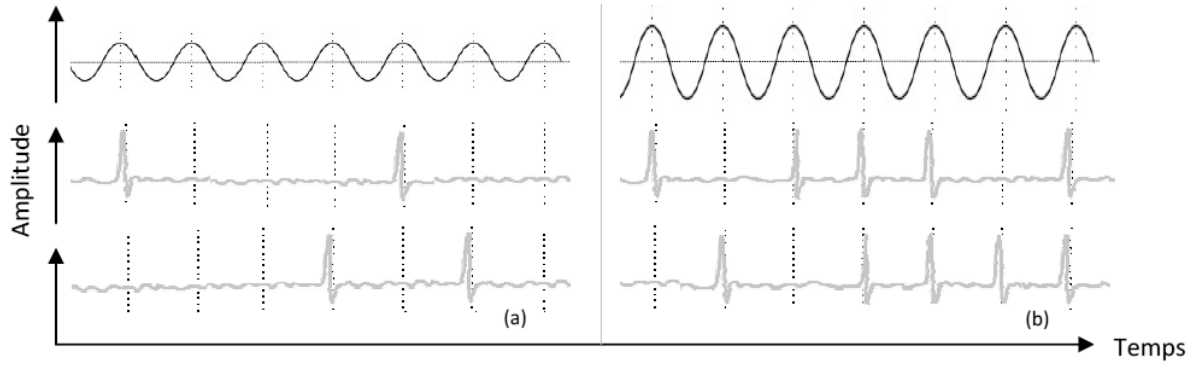


Figure 11. *Représentation du codage par phase locking pour une intensité faible (a) et une intensité plus élevée (b). Le potentiel d'action reste synchrone à la fréquence d'émission.*

Toutefois, cette théorie temporelle ne peut s'appliquer qu'aux basses fréquences audio, car pour la plupart des mammifères (e.g., Palmer & Russel, 1986), la fidélité du verrouillage en phase sur la TFS diminue à mesure que la fréquence audio augmente, sans que la limite exacte des capacités de verrouillage en phase ne soit clairement établie pour l'homme (Joris & Verschooten, 2013). Cette estimation de la fréquence pourrait être utilisée pour détecter une composante de FM (Moore & Glasberg, 1989, Moore & Sek, 1995, 1996; Sek & Moore, 1995).

En accord avec la théorie du pattern d'excitation, un changement de fréquence provoque un léger décalage du pattern d'excitation. Ce changement est supposé être détectable chaque fois que le niveau d'excitation en un certain point du pattern d'excitation dépasse une valeur seuil (environ 1dB d'après Zwicker). Ainsi, le modèle de Zwicker (1956) suggère que la détection d'une variation de fréquence est analogue à la détection d'une variation d'intensité, elle-même conditionnée par la pente du pattern d'excitation. La pente du pattern d'excitation est relativement constante lorsqu'elle est exprimée sur une échelle perceptive basée sur l'ERBn

estimant la bande passante des filtres auditifs (Equivalent Rectangular Bandwidth, ou bande passante rectangulaire équivalente pour un niveau de stimulation modéré et un sujet normo-entendant, Moore, 2007). Ainsi, le modèle de Zwicker prédit que le seuil de détection de FM (ou FMDL) devrait être une fraction constante de la valeur de l'ERBn à la fréquence considérée.

En accord avec le modèle Zwicker (1956), les FMDLs correspondent à une proportion constante de l'ERB pour des cadences de modulation supérieures à 10 Hz (Sek & Moore, 1995; Moore & Sek, 1995). Toutefois, les FMDLs dépendent de la cadence de modulation (Fm) et de la fréquence porteuse (Fc) (e.g. Demany & Semal, 1986, cf. Figure 12). A basse cadence de modulation, le rapport FMDL/ERBn n'est pas constant, et augmente lorsque la fréquence centrale augmente. Glasberg et Moore (1989) et Sek et Moore (1995) démontrent que ces résultats ne sont pas compatibles avec les prédictions du modèle de Zwicker et suggèrent l'existence d'un autre mécanisme permettant la détection d'une FM.

Moore et Sek (1995) réalisent une étude paramétrique et comparent l'évolution des seuils de détection d'AM et de FM en fonction de la fréquence porteuse, de la cadence de modulation et du type de modulation utilisée. Ils démontrent: i) des similarités entre les seuils de détection d'AM et de FM pour des fréquences porteuses élevées, pour lesquelles le "phase locking" est supposé être aboli chez l'humain; ii) des différences entre les seuils de détection d'AM et la FM à basses porteuses, pour lesquelles le "phase locking" est supposé être présent. Moore et Sek (1995) proposent l'idée que la détection de FM en basses fréquences dépendrait en partie de la capacité à encoder la fréquence instantanée par "phase locking".

Moore et Sek (1995, 1996) et Sek et Moore (1995) montrent qu'en basse fréquence audio, les FMDLs sont plus petits (c'est-à-dire, meilleurs) pour une cadence de modulation lente (Fm=2 Hz) que rapide (Fm=10 Hz). En haute fréquence (au-delà de Fc=4 kHz), le pattern des FMDLs est inversé. Moore et Sek proposent que la détection de FM à basse cadence de modulation et pour des porteuses inférieures à 5 kHz est basée sur la détection de changements dans le verrouillage en phase (codage temporel). En d'autres termes, la fréquence instantanée pourrait être

déterminée sur de courts intervalles temporels, en utilisant des informations issues du verrouillage de phase. Une variation dans l'estimation de la fréquence au fil du temps indique la présence de la FM.

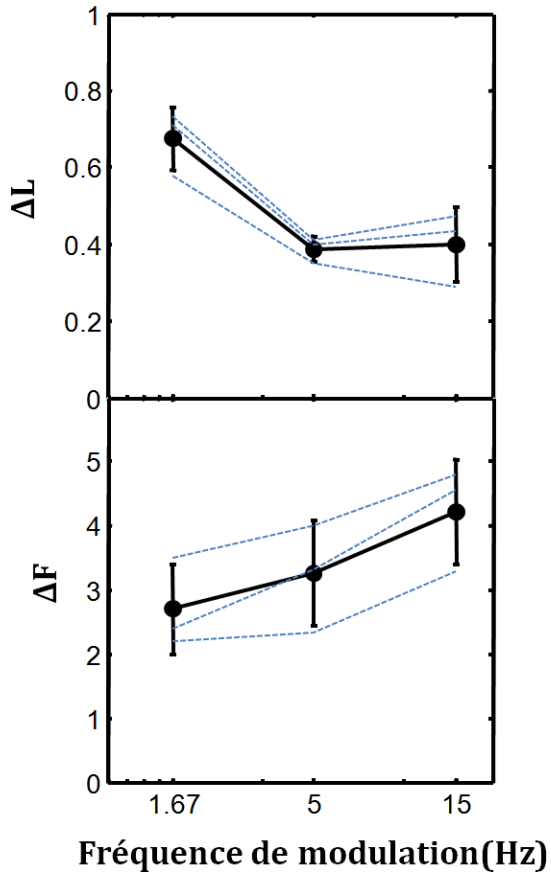


Figure 12 . *AMDLs (graphe supérieur) et FMDLs (graphe inférieur) mesurés chez trois sujets normo-entendants en fonction de la cadence de modulation. Pour un son pur de 1500Hz à une intensité de 65 dB, les AMDLs et FMDLs obtenus varient différemment avec la cadence de modulation pour des cadences de modulation lentes (Fm<5Hz). Les auteurs suggèrent que les mécanismes impliqués dans la détection d'une modulation avec une cadence de modulation rapide (15Hz) pourraient différer de ceux impliqués lorsque la cadence de modulation est plus lente (2Hz). Retracé à partir de Demany & Semal, 1986.*

Moore et Sek (1996) évaluent l'importance relative des indices temporels et des indices de pattern d'excitation dans la détection d'une FM. De façon à dégrader les indices de pattern d'excitation potentiellement utilisables pour la détection de la FM, ils mesurent les FMDLs en présence d'une AM interférente. L'AM interférente (de cadence de modulation identique, et de phase de départ aléatoires) introduit des variations d'intensité supraliminaires, non informatives pour la détection de la FM.

Les FMDLs obtenus avec et sans AM interférente sont comparés chez des sujets normo-entendants (cf. Figure 13). L'effet délétère de l'AM sur les seuils de détection de FM s'accroît lorsque l'on augmente la cadence de modulation pour une fréquence porteuse inférieure à 4

kHz. A partir de ces résultats, Moore et Sek (1996) concluent que: i) la contribution des indices de pattern d'excitation à la détection de la FM croît à mesure que la cadence de modulation augmente; ii) les mécanismes impliqués dans le décodage (la "lecture" des informations issues) du verrouillage de phase sur la TFS du signal sont lents (« sluggish »), et nécessitent d'échantillonner le stimulus pendant un certain temps afin de pouvoir estimer sa fréquence.

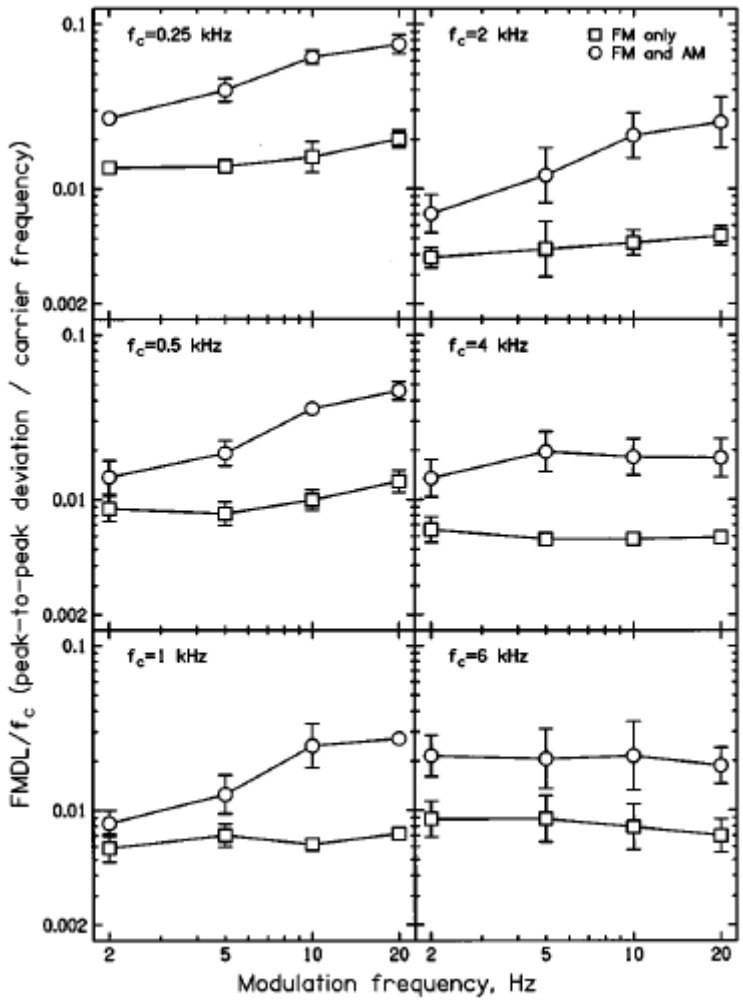


Figure 13. *Seuils de détection de FM (FMDLs) en fonction de la cadence de modulation. Les FMDLs correspondent à la déviation en fréquence (valeur crête-crête), divisée par la fréquence centrale. Les carrés montrent les FMDLs obtenus sans AM interférente, les ronds correspondent aux FMDLs obtenus en présence d'une AM interférente. Chaque graphique représente les résultats moyennés entre les sujets pour chaque fréquence porteuse testée. Les barres d'erreurs représentent l'erreur standard. Extrait de Moore & Sek (1996).*

Une telle lenteur dans l'analyse des informations encodées par "phase locking" est également suggérée par Moore et Glasberg (1987a) et plus récemment par Gockel et al. (2001). Ainsi, les mécanismes temporels ne joueraient qu'un rôle limité pour le traitement des cadences de modulation élevées, même si la fréquence porteuse appartient à l'intervalle pour lequel le verrouillage de phase reste précis. A 6 kHz, l'effet délétère de l'AM obtenu par Moore et Sek

(1996) est similaire quelle que soit la cadence de modulation, corroborant l'idée que seuls les indices de pattern d'excitation peuvent être utilisés en haute fréquence.

Les études réalisées par Moore et Sek (1995, 1996) suggèrent donc que les mécanismes temporels impliqués dans la détection de la FM sont lents (« sluggish ») et que la détection de la FM dépend du temps que le stimulus passe aux fréquences extrêmes. Pour tester cette hypothèse, Sek et Moore (2000) mesurent les seuils de détection pour une FM et une AM quasi trapézoïdale. Pour des cadences de modulation supérieures à 10 Hz, les FMDLs s'améliorent lorsque le temps que le stimulus passe aux fréquences extrêmes augmente, en accord avec la suggestion de Moore et Sek (1995, 1996). L'effet bénéfique d'une augmentation de la durée passée aux fréquences extrêmes sur les FMDLs est observé en basse fréquence (1 kHz) et non en haute fréquence (6 kHz). Cet effet différentiel de la fréquence corrobore l'hypothèse d'un double code pour la détection de FM, et plus particulièrement l'implication d'un code de nature purement temporelle, opérant principalement en basse fréquence audio (<1-4 kHz) et pour des cadences de FM extrêmement lentes, inférieures à 5-10 Hz. L'importance « écologique » de ce mécanisme temporel est évidente puisque le domaine d'application de ce dernier correspond parfaitement aux cadences de FM les plus représentées (les plus saillantes) dans la parole continue, et ce pour de nombreuses langues du monde (e.g., Sheft et al., 2012 ; Varnet et al., 2017). On notera toutefois que si la limite d'application de ce mécanisme temporel dans le domaine audio correspond bien à la limite du verrouillage en phase neural (1-4 kHz) pour la plupart des mammifères, la limite d'application de ce mécanisme temporel dans le domaine des modulations (5-10 Hz) ne correspond à aucune donnée neurophysiologique connue. L'hypothèse de « lenteur » (sluggishness) du mécanisme temporel reste donc à confirmer tant sur le plan psychophysique, computationnel qu'électrophysiologique.

D'autres études portant sur la détection de FM corroborent l'existence de deux codes. Chen et Zeng (2004) mesurent les FMDLs chez des auditeurs normo-entendants (n=3) et chez des sujets porteurs d'un implant cochléaire (n=3) de façon à évaluer la contribution relative des indices temporels et spectraux impliqués dans la perception d'une FM. L'étude des capacités des

patients implantés offre une opportunité unique de démontrer l'existence d'un deuxième code non-tonotopique, car dans le cas d'une stimulation électrique directe réalisée sur une seule électrode du dispositif implanté, la détection du patient ne peut être basée que sur des indices temporels de phase locking. Les résultats de Chen et Zeng (2004) montrent que les patients implantés peuvent détecter une FM, et confirment donc l'existence d'un mécanisme temporel indépendant pour le codage de la FM, prédominant aux basses fréquences audio et aux faibles cadences de modulation.

Ernst et Moore (2010) répliquent l'étude de Moore et Sek (1996) pour deux niveaux de stimulation différents (20 et 60 dB SL). A 60 dB SL, leurs résultats confirment ceux obtenus par Moore et Sek (1996), corroborant l'existence d'un double code. A faible intensité (20 dB SL), l'effet délétère de l'AM est similaire quelle que soit la cadence de modulation ou la fréquence porteuse, suggérant une contribution majeure des indices de pattern d'excitation et ce même en basse fréquence audio et pour de faibles cadences de modulation. Ernst et Moore (2010) et Moore (2014) interprètent ces résultats par le fait que: i) les filtres auditifs deviennent plus étroits à faible intensité (e.g., Moore & Glasberg, 1987b; Baker & Rosen, 2006; Oxenham & Simonson, 2006); ii) les mécanismes d'encodage par verrouillage de phase sont moins précis à faible intensité (Johnson, 1980; Palmer & Russell, 1986; Erulkar et al., 1968).

En conclusion, pour des stimuli sinusoïdaux, de nombreux aspects de la perception du sujet normo-entendant se dégradent lorsque la fréquence audio augmente. Ceci est compatible avec la diminution de la précision du verrouillage de phase neuronal sur la TFS du signal pour des fréquences supérieures à 1-4 kHz. Pour des fréquences supérieures à 1 kHz, une transition surviendrait entre le mécanisme temporel basé sur les indices neuronaux de TFS et le mécanisme basé sur les indices de pattern d'excitation, qui dépendent du filtrage cochléaire.

En haute fréquence audio, les seuils de détection de FM correspondent à une fraction constante de la bande passante des filtres auditifs à la fréquence considérée, en accord avec le modèle du pattern d'excitation (Zwicker, 1956; Zwicker & Fastl, 1990; Moore & Sek,1994, 1995; Saberi & Hafter, 1995; Sek & Moore, 1995). Toutefois, les seuils de détection de FM aux basses cadences de modulation augmentent comparativement à la valeur de la bande passante des filtres auditifs. Ceci suggère que les seuils de détection de FM sont déterminés principalement par un mécanisme temporel opérant en basse fréquence audio (Fc<4 kHz) mais ce uniquement pour les cadences de modulation basses, inférieures à 5-10 Hz (Moore & Sek, 1995, 1996; Plack & Carlyon, 1995; Sek & Moore, 1995). En d'autres termes, le mécanisme temporel est supposé être « lent », utilisant une fenêtre d'analyse (une fenêtre de calcul de la hauteur instantanée) de plusieurs dizaines de millisecondes.

3.2.2 Modèle computationnel du traitement de la FM

Différents modèles ont été proposés pour rendre compte de la sensibilité à la FM. Le modèle de Zwicker (1952, 1956) décrit précédemment suppose que la détection de FM est uniquement réalisée à partir des variations dans le pattern d'excitation (cf. Figure 14), de façon analogue à la détection d'une AM. Toutefois, contrairement à la détection d'une AM qui reposerait sur l'écoute du ou des filtres auditifs dans la région haute fréquence du pattern d'excitation, la détection d'une FM serait réalisée en écoutant le ou les filtres auditifs dans la région basse fréquence du pattern d'excitation (Maiwald, 1967a; Demany & Semal, 1986; Moore & Sek, 1994). Ce modèle repose donc sur un processus de « conversion de la FM en AM » (ou plus précisément de la FM en enveloppe temporelle), processus dont l'importance a été démontrée pour la reconnaissance robuste de la parole, lorsque cette dernière est soumise à des distorsions sévères telles qu'une compression d'amplitude ou un écrêtage (pour une revue, voir Shamma & Lorenzi, 2013). En effet, lorsqu'un signal de FM passe au travers un filtre auditif à bande étroite, les excursions dynamiques de la fréquence instantanée sont converties en variation dynamique du niveau de sortie (Saberi & Hafter, 1995).

Hartmann et Klein (1980) proposent un modèle distinct du modèle de Zwicker, permettant de rendre compte des performances de détection d'une FM quelle que soit la cadence de modulation. Pour des cadences de modulation rapides, le modèle d'Hartmann (1978, 1979) se comporte comme le modèle de pattern d'excitation proposé par Zwicker. Toutefois, pour les cadences de modulation lentes, Hartmann et Klein (1980) supposent que la détection d'une FM est basée sur un échantillonnage du signal et une estimation de sa hauteur au sein de chaque échantillon. Pour détecter une FM, un sujet compare donc l'estimation de la hauteur réalisée sur deux échantillons successifs. Si cette différence entre estimations successives est suffisamment importante, le sujet décide que le signal présente une FM. Hartmann et Klein (1980) calculent ensuite la corrélation entre les estimations successives de la hauteur et un "modèle" de la forme

d'onde du signal modulé mémorisé au préalable par le sujet pour améliorer l'efficacité du modèle.

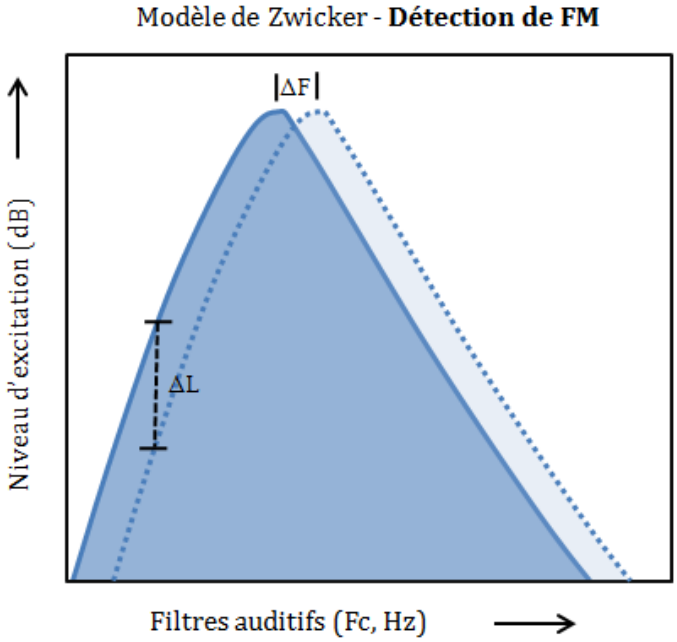


Figure 14*. Illustration schématique du modèle de pattern d'excitation appliqué à la détection de FM. Les patterns d'excitation de deux sinusoïdes de fréquences proches sont représentés. La différence de fréquence ΔF, peut être détectable si le changement du niveau d'excitation (ΔL) dépasse une valeur seuil (modèle de Zwicker). Le changement du niveau d'excitation survient sur la partie basse fréquence du pattern d'excitation (pente plus abrupte). Retracé à partir de Demany & Semal, 1986 et Moore, 2007.*

Demany et Semal (1986) proposent que la détection de l'AM et de la FM est réalisée par deux processus distincts, permettant d'extraire d'une part la puissance instantanée du signal (sa sonie) et d'autre part sa fréquence instantanée (sa hauteur). Ces processus impliquent d'échantillonner le signal sur des fenêtres temporelles plus courtes que la période d'un cycle de modulation. Ce fenêtrage temporel, pour être optimal, doit être ajusté en fonction de la cadence de modulation et de la fréquence porteuse. Demany et Semal (1986) suggèrent que la fenêtre temporelle d'analyse est d'environ 25 ms de façon à être optimale pour la perception de la voix humaine.

Ernst et Moore (2010) développent un modèle « multi-bande » basé sur le pattern d'excitation pour modéliser leurs données de FM en présence d'une AM interférente. Deux filtres auditifs

(filtres gammatones, Patterson et al., 1995) répartis de part et d'autre de la fréquence porteuse du stimulus, sont simulés. Ernst et Moore (2010) supposent que lorsque la FM est présente, les variations du niveau d'excitation induites par la FM au sein de chacun des filtres sont en opposition de phase. Les informations de chaque filtre sont ensuite combinées en calculant le rapport entre l'enveloppe du signal en sortie de chacun des deux filtres. Ce rapport permet de supprimer l'effet de masquage introduit par l'AM interférente, dans la mesure où les indices d'enveloppe introduits par l'AM interférente sont en phase. La statistique de décision se base sur l'écart-type de ce rapport, de sorte que plus l'écart-type est important, plus le signal a de chance d'être modulé. Un bruit interne est enfin ajouté à la variable de décision de façon à rendre compte des données expérimentales. Ce modèle, basé uniquement sur les indices dynamiques issus du pattern d'excitation, permet de rendre compte fidèlement des données expérimentales pour des cadences de modulations élevées et pour des porteuses élevées. En revanche, il ne permet pas de rendre compte du masquage produit par l'AM interférente (dépendant de la porteuse et de la cadence de modulation). Les auteurs en concluent qu'un modèle utilisant uniquement les indices issus du pattern d'excitation ne permet pas de modéliser correctement les données de FM et que la détection de FM ne peut résulter des seuls indices issus du pattern d'excitation. L'étude de modélisation réalisée par Paraouty et al. (2016) à l'aide d'une version multi-bande du modèle de traitement d'enveloppe EPSM (utilisant 2 filtres auditifs accordés de part et d'autre de la fréquence porteuse) aboutit aux mêmes conclusions qu'Ernst et Moore (2010) pour la détection de FM en présence d'AM interférente.

3.2.3 Etudes développementales

L'étude du développement auditif du traitement des modulations peut constituer une source précieuse d'informations quant à la nature du/des codes sensoriel(s) et des facteurs cognitifs impliqués dans la perception de l'AM et de la FM (voir la thèse de doctorat de L. Cabrera (2013) pour une revue détaillée des effets développementaux pour la perception des modulations). Au

niveau neural, les données animales collectées chez le chaton suggèrent que le verrouillage en phase des fibres du nerf auditif pour des sons purs s'améliore pendant les premiers jours de vie (Kettnet et al., 1985). Très peu d'études psychophysiques humaines ont toutefois évalué les capacités de détection d'une FM chez le nourrisson et l'enfant. Colombo et Horowitz (1986) montrent que des nourrissons âgés de 4 mois sont capables de discriminer deux balayages en fréquence différents (allant de 150 à 550 Hz et de 150 à 275 Hz sur une période d'une seconde). Aslin (1989) démontre toutefois que les nourrissons de 6 mois ont besoin de transitions fréquentielles plus importantes que les adultes pour détecter une FM. Chez l'enfant, la performance de détection d'une FM s'améliore entre 6 et 10 ans, la sensibilité aux basses cadences de modulation (Fm=2 Hz) étant particulièrement mauvaise jusqu'à l'âge de 9 ans (Dawes & Bishop, 2008). Ces premières études indiquent donc que la perception de FM évolue pendant les premières années de vie.

Banai et al. (2011) comparent les trajectoires développementales respectives pour la détection d'une AM et d'une FM chez des enfants âgés de 8 à 10 ans et de 11 à 12 ans. Pour la détection de FM, la performance moyenne et la disparité des réponses sont identiques à celles des adultes dès l'âge de 8-10 ans. A l'inverse, chez les mêmes sujets, les performances des enfants diffèrent sensiblement de celles obtenues chez l'adulte pour l'AM. Pour l'ensemble des sujets, aucune corrélation n'est retrouvée entre le degré de maturation pour la détection d'AM ou pour la détection de FM. Ces différences de trajectoires développementales, de performances moyennes et de disparité des réponses entre les deux types de modulation suggèrent l'existence de processus distincts impliqués dans le développement de la détection d'AM et de FM.

Ces résultats suggèrent que les jeunes enfants sont capables de percevoir une FM. Toutefois, il semblerait que le traitement auditif des indices de FM se développe pendant l'enfance, tout au moins jusqu'à 9-10 ans. La comparaison des trajectoires développementales pour la détection d'AM et de FM corrobore l'existence de deux codes pour la détection de FM.

4. Effet du vieillissement sur la perception de l'AM et de la FM

Comme l'étude du développement, l'étude des effets du vieillissement peut également constituer une source d'informations importante sur la nature du code sensoriel et des mécanismes cognitifs impliqués dans la perception de l'AM et de la FM. A l'inverse des études développementales, les études portant sur l'effet du vieillissement sont nombreuses, mais elles portent sur des capacités auditives temporelles extrêmement variées. Des déficits significatifs potentiellement causés par le vieillissement ont été décrits: i) dans des tâches de détection d'un intervalle de silence (Strouse et al., 1998; Phillips et al., 2000; Snell et al., 2002; Helfer & Vargo, 2009); ii) pour une tâche de détection d'AM ou de FM sinusoïdale (e.g., He et al., 2007, 2008); iii) dans des tâches mesurant les effets de masquage postérieur (Gifford et al., 2007); iv) dans une tâche de discrimination d'une différence interaurale de temps ou de phase (Divenyi & Haupt, 1997; Grose & Mamo, 2010; Hopkins & Moore, 2011) ou encore; v) dans une tâche de discrimination de rythme (Trainor & Trehub, 1989; Humes & Christopherson, 1991; Gordon-Salant & Fitzgibbons, 1999; Fitzgibbons et al., 2006).

Toutefois, l'interprétation des résultats de ces études demeure complexe, notamment en raison de l'importante variabilité interindividuelle observée dans ces études, de la petite taille des cohortes, et enfin d'effets potentiels de l'entrainement aux tâches auditives. Par ailleurs, il est souvent difficile de séparer l'effet de l'âge en soi sur les capacités sensorielles temporelles de l'effet de la perte auditive (la perte d'audibilité mesurée par l'audiogramme tonal) et/ou du déclin des capacités cognitives (mnésiques, attentionnelles et décisionnelles) généralement associés au vieillissement.

4.1 Effet du vieillissement sur la détection d'AM

Takahashi et Bacon (1992) réalisent la première étude évaluant l'effet de l'âge sur la détection de l'AM. Ils mesurent des MTFs (porteuse: large bande de bruit ; Fm= [2-1024] Hz; niveau de présentation de 35dB SPL) chez 10 sujets normo-entendants jeunes et 30 sujets plus âgés,

présentant une audition normale (des seuils audiométriques tonaux inférieurs à 20 dB HL) jusqu'à 2 kHz. Les AMDLs sont mesurés sans entrainement, et moyennés sur 3 estimations de seuil. Les MTFs présentent des fréquences de coupure et pentes similaires pour les deux groupes de sujets. Pour des cadences de modulation lentes (Fm=2 Hz), les sujets âgés présentent des AMDLs plus élevés que ceux des sujets jeunes, conférant à la MTF une allure passe bande. Aucune différence significative n'est observée entre la sensibilité à l'AM des sujets jeunes et des sujets plus âgés, malgré le fait que la sensibilité à l'AM apparaisse globalement moins bonne chez les sujets âgés. Une corrélation significative est toutefois retrouvée entre la sensibilité à l'AM et l'âge des sujets, et ce lorsque l'effet des seuils audiométriques est contrôlé. Ces premiers résultats suggèrent que l'âge dégrade la sensibilité à l'AM sans pour autant modifier l'acuité temporelle (estimée par la fréquence de coupure et la pente de la MTF). Toutefois, l'usage d'une bruit large bande comme signal porteuse introduit une différence de « bande passante audible » (listening bandwidth) entre sujets jeunes et âgés, ces derniers présentant une perte auditive tonale au delà de 2 kHz et donc, un accès limité aux informations de modulation en haute fréquence audio. La sensibilité à l'AM s'améliorant à mesure que la bande passante de la porteuse s'accroît (Eddins, 1993), l'effet délétère de l'âge sur la sensibilité à l'AM observé par Takahashi et Bacon (1992) pourrait donc résulter d'une perte d'audibilité en haute fréquence plutôt que d'un réel déficit de traitement d'enveloppe temporelle.

Purcell et al. (2004) évaluent l'effet de l'âge sur l'acuité temporelle via une tâche de détection d'AM, une tâche de détection d'un intervalle de silence et des enregistrements de potentiels évoqués par un signal d'AM (Envelope Following Response, EFR). Quatorze sujets normo-entendants jeunes (18-43 ans) et treize sujets plus âges (68-78 ans) présentant une audition normale ou une perte auditive légère (PTA=21 dB HL pour les fréquences comprises entre 0.5 et 4 kHz; 23% des sujets âgés présentent des seuils audiométriques tonaux supérieurs à 35 dB HL) ont été testés. Pour la tâche de détection d'AM, les auteurs estiment la cadence de modulation maximale pour laquelle les sujets arrivent à détecter une AM avec une profondeur de 25 % et appliquée à une bande de bruit (cadence de modulation fixée au départ à 50 Hz ou 10 Hz pour

les sujets ayant des difficultés pour réaliser la tâche). Les auditeurs sont testés à 60 dB SPL, ou à un niveau jugé confortable. La plus haute cadence de modulation pour laquelle l'AM à 25% demeure détectable varie significativement entre les sujets jeunes (Fm=567±95 Hz) et les sujets âges (Fm=264±131 Hz). Ces résultats, combinés à ceux obtenus dans la tâche de détection d'un intervalle de silence et aux enregistrements d'EFR, conduisent les auteurs à conclure que le vieillissement dégrade l'acuité temporelle auditive. Malheureusement, Purcell et al. (2004) ne peuvent établir la contribution relative du vieillissement et de la perte auditive au déficit d'acuité temporelle observé, les deux facteurs demeurant confondus.

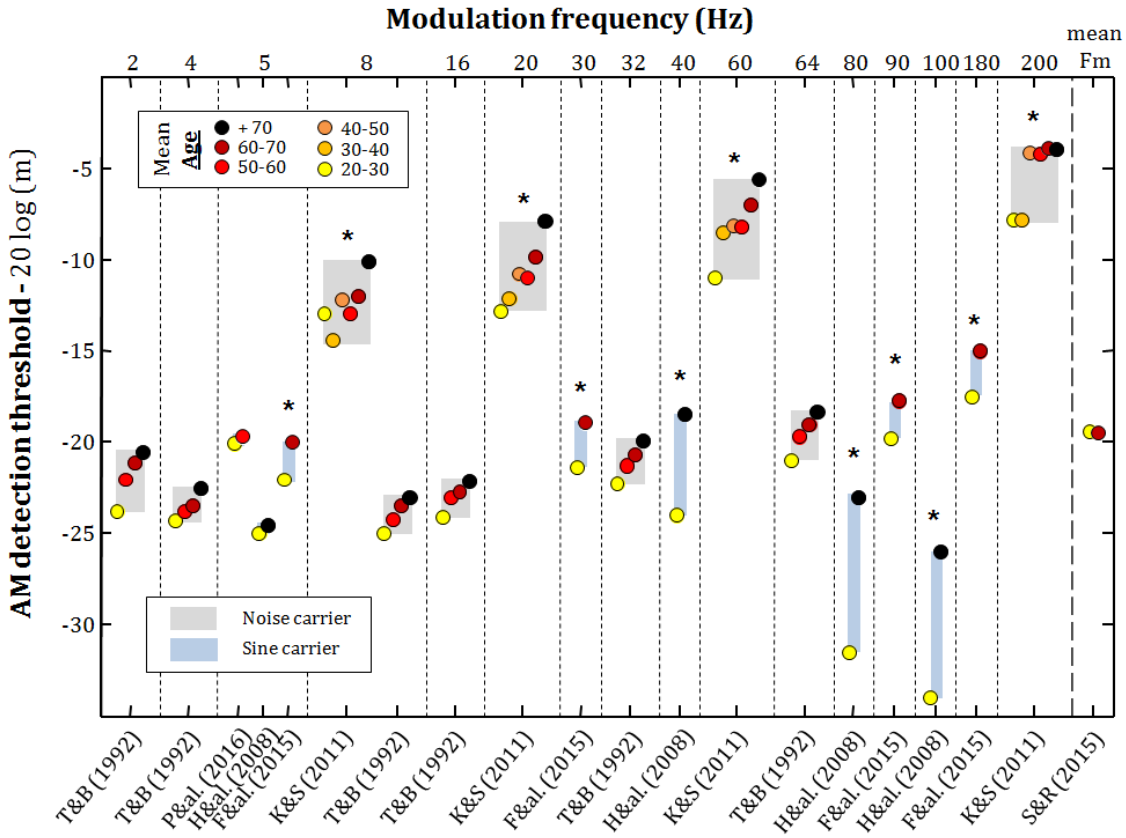


Figure 15: *AMDLs moyens obtenus par différentes études évaluant l'effet du vieillissement sur la perception de l'AM. La couleur des cercles indique l'âge moyen des sujets. Les barres verticales précisent l'intervalle min/max pour chaque distribution. La couleur des barres indique le type de porteuse utilisée: gris pour une porteuse bruit et bleu pour une porteuse sinusoïdale. Les AMDLs sont représentés en fonction de la cadence de modulation. Les seuils de détection d'AM pour Schoof & Rosen (2015) correspondent aux AMDLs moyennés pour les différentes cadences de modulation testées. Les différentes études sont symbolisées par: T&B (1992) - Takahashi & Bacon (1992); P&al. (2016) - Paraouty et al. (2016); H&al. (2008) - He et al. (2008); K&S (2011) - Kumat & Sangamanatha (2011); F&al. (2015) - Fullgrabe et al. (2015); S&R (2015) - Schoof & Rosen (2015). Les étoiles au dessus des données indiquent les conditions pour lesquelles le vieillissement dégrade significativement les AMDLs.*

He et al. (2008) mesurent les MTFs (Fm = [5-250] Hz pour Fc=0.5 kHz; Fm=[5-1000] Hz pour Fc=4 kHz; niveau de présentation de 75 dB SPL) chez 8 sujets normo-entendants jeunes (18-26 ans) et 8 sujets plus âgés (60-80 ans) présentant une audiométrie tonale normale jusqu'à 4 kHz, et des seuils audiométriques plus élevés que ceux des sujets normo-entendants au delà de 4 kHz (PTA_{6kHz}≈31dB HL). Point important, les sujets sont initialement entraînés jusqu'à stabilisation de leurs AMDLs (1 à 2h) dans les conditions correspondant à une cadence de modulation, Fm, de 5 Hz. A l'issue de l'entraînement, les AMDLs mesurés pour Fm=5 Hz ne diffèrent pas significativement entre groupes d'âge, quelle que soit la fréquence porteuse utilisée. Ceci permet de s'assurer du fait que tous les sujets ont compris la procédure expérimentale et sont à même d'utiliser les indices auditifs qui leurs sont présentés dans chaque tâche. Les MTFs mesurées pour les deux groupes et pour les deux fréquences porteuses présentent la même allure : les AMDLs augmentent (i.e., se dégradent) jusqu'à une cadence de modulation donnée (30 Hz pour Fc=0.5 kHz; 400 Hz pour Fc=4 kHz), à partir de laquelle les AMDLs s'améliorent fortement. Au-delà de cette cadence de modulation « de transition », la détection de l'AM se base principalement sur la détection des bandes spectrales latérales, désormais résolues par les filtres cochléaires (Kohlrausch et al., 2000). Les cadences de modulation « de transition » pour lesquelles les AMDLs s'améliorent sont comparables pour les deux groupes d'âge, conformément à l'idée que le vieillissement n'affecte pas la sélectivité fréquentielle auditive (Lutman et al., 1991; Gifford & Bacon, 2005). Pour la porteuse basse fréquence (Fc=0.5 kHz), le vieillissement dégrade significativement les AMDLs (effet de 3-8 dB) aux cadences de modulation comprises entre 5 et 250 Hz (il est intéressant de noter que les AMDLs, bien que s'améliorant au-delà de la cadence de modulation « de transition », sont toujours plus élevés (i.e., moins bons) chez les sujets âgés que chez les sujets jeunes). Pour la porteuse haute fréquence (Fc = 4 kHz), le vieillissement dégrade également les AMDLs de manière significative pour les cadences de modulation comprises entre 5 et 400 Hz. L'effet du vieillissement est globalement de même amplitude que celui observé pour Fc=0.5 kHz. Pour des cadences de modulation plus élevées (Fm>400 Hz), aucune différence significative n'est observée entre les deux groupes. He et al.

(2008) concluent que le vieillissement dégrade la détection d'AM rapides (> 5 Hz), et donc le traitement d'enveloppe temporelle, quelle que soit la porteuse utilisée. Point intéressant, les auteurs interprètent l'effet délétère du vieillissement sur la détection d'AM observé aux cadences supérieures à 30 Hz lorsque la porteuse est basse (Fc=0.5 kHz) comme réflétant un déficit de verrouillage en phase des réponses neuronales sur la porteuse (la TFS) du signal (et donc, sur la détectabilité des bandes spectrales latérales en basse fréquence). Toutefois, les données de He et al. (2008) ne permettent pas de conclure fermement quant à un effet du vieillissement sur le traitement sensoriel des indices temporels véhiculés par l'AM, cet effet n'étant plus observé lorsque les sujets sont suffisamment entraînés (i.e., à Fm=5 Hz).

Kumar et Sangamanatha (2011) réalisent la première étude des effets de l'âge sur la sensibilité à l'AM sur une large cohorte (n=176) d'auditeurs normo-entendants d'âge variable (20-85 ans) . Les AMDLs mesurés pour différentes conditions expérimentales (porteuse: bruit gaussien; Fm = 8, 20, 60, 200 Hz; niveau de présentation de 80 dB SPL) sont estimés à partir des fonctions psychométriques. Les résultats indiquent que les AMDLs se dégradent significativement avec l'âge pour l'ensemble des conditions expérimentales. L'effet délétère du vieillissement semble s'accroitre à mesure que la cadence de modulation augmente: les AMDLs sont significativement plus élevés pour des cadences de modulation élevées (Fm>60 Hz) à partir de 40 ans; les AMDLs ne se dégradent significativement que plus tardivement (vers 60 ans) pour des cadences de modulation plus lentes (Fm<20 Hz). Dans cette étude, Kumar et Sangamanatha (2011) démontrent également que l'âge est un très bon prédicteur de la sensibilité à l'AM,.

Schoof et Rosen (2014) mesurent les MTFs (porteuse: large bande bruit [1-4 kHz]; Fm = 10, 20, 40, 80, 160 Hz; niveau de présentation de 70 dB SPL) chez 19 sujets jeunes (19-29 ans) et 19 sujets âgés (60-72 ans) ayant une audition normale (seuils audiométriques tonaux < 25 dB HL) jusqu'à 4 kHz. Ces mesures auditives sont complétées par des mesures de la mémoire de travail et de la rapidité d'exécution. Les MTFs obtenues après entrainement des auditeurs présentent une caractéristique passe bas pour l'ensemble des sujets. Toutefois, contrairement aux études antérieures, Schoof et Rosen (2014) n'observent aucune différence significative entre les deux

groupes de sujets, tant pour la fréquence de coupure de la MTF que pour la sensibilité absolue à l'AM des sujets. Les tests cognitifs révèlent un déficit de mémoire de travail et de rapidité d'exécution chez les sujets âgés. Schoof et Rosen (2014) n'observent donc pas d'effet délétère du vieillissement sur le traitement d'AM..

Füllgrabe et al. (2015) mesurent les MTFs (Fc = 4 kHz; Fm = 5, 30, 90, 180 Hz; niveau de présentation de 30 dB SL) après entraînement chez 9 sujets jeunes (18-27 ans) et 21 sujets plus âgés (60-79 ans). Les sujets jeunes et âgés sont normo-entendants et sont appariés sur le plan audiométrique et éducatif (QI, nombre d'année d'étude). Les AMDLs des sujets les plus âgés sont significativement plus élevés quelle que soit la cadence de modulation. Toutefois, les pentes des MTFs sont comparables entre les deux groupes. Les auteurs en concluent que l'effet du vieillissement sur la sensibilité à l'AM résulte d'avantage d'un déficit d'efficacité de traitement (un déficit cognitif de mémoire de travail et/ou d'attention sélective), que d'un déficit sensoriel temporel stricto sensu.

Whiteford et al. (2017) évaluent également l'effet du vieillissement sur les AMDLs (Fc = 0.5 kHz; Fm = 1 ou 20 Hz; niveau de présentation de 60 dB SL) sur une large cohorte de sujets (n=85). Les sujets de moins de 50 ans présentent des seuils audiométriques tonaux inférieurs à 20 dB HL entre 0.25 et 8 kHz, tandis que les sujets de plus de 50 ans présentent des seuils audiométriques tonaux normaux jusqu'à 1 kHz. Les stimuli sont présentés de façon diotique ou dichotique, et les seuils sont mesurés après un entraînement préalable. Whiteford et al. (2017) n'obtiennent aucune corrélation significative entre l'âge et les AMDLs après correction de Bonferroni, sauf pour la condition utilisant une AM présentée de manière dichotique pour Fm=20 Hz. Les auteurs en concluent que l'âge n'affecte pas substantiellement la sensibilité à l'AM.

Enfin, Paraouty et Lorenzi (2017) mesurent les AMDLs (Fc = 0.5 kHz; Fm = 5 Hz; niveau de présentation de 40 dB SL) sur une large cohorte de sujets (n=70; 20-70 ans) présentant des seuils audiométriques tonaux inférieurs à 20 dB HL jusqu'à 2 kHz. Paraouty et Lorenzi observent une légère dégradation des AMDLs lorsque l'âge augmente. Toutefois, aucune corrélation

significative n'est obtenue entre l'âge et l'AMDL. L'absence d'effet du vieillissement sur la sensibilité à l'AM est conforme aux résultats de Takahashi et Bacon (1992), Schoof et Rosen (2014) et Paraouty et al. (2016). Ce résultat contraste toutefois avec les résultats des études de He et al. (2008) et Füllgrabe et al. (2015) pour lesquelles les seuils audiométriques des participants ont été soigneusement contrôlés et où tous les auditeurs ont été entraînés.

La figure 15 représente les AMDLs obtenus pour l'ensemble des études précédentes, en fonction de l'âge moyen des sujets, du type de porteuse et de la cadence de modulation. Les principaux paramètres expérimentaux et résultats des différentes études évaluant l'effet du vieillissmeent sur la détection d'AM sont présentés dans le tableau 1.

En conclusion, les résultats des différentes études évaluant l'effet du vieillissement sur la perception de l'AM sont fortement contrastés: les données de He et al. (2008), Füllgrabe et al. (2015), Kumar et Sangamanatha (2011) et Purcell et al. (2004) indiquent que le vieillissement dégrade significativement la sensibilité à l'AM. Paraouty et Lorenzi (2017) et Takahashi et Bacon (1992), quant à eux, observent un léger déficit de détection d'AM, mais le déclin observé n'est pas significatif. Enfin, Paraouty et al. (2016), Schoof et Rosen (2014) et Whiteford et al. (2017) ne trouvent aucun effet significatif de l'âge sur la sensibilité à l'AM. L'effet délétère de l'âge sur la sensibilité à l'AM, lorsqu'il est observé, demeure modeste (1 à 8 dB au maximum). La dégradation de la sensibilité à l'AM associée au vieillissement semble essentiellement intervenir aux cadences de modulation les plus rapides (He et al., 2008; Kumar & Sangamanatha, 2011; Purcell et al., 2004).

Par ailleurs, la dégradation de la sensibilité à l'AM pourrait également résulter d'un déficit dans le traitement central de l'information sensorielle disponible . L'origine de ce déficit « d'efficacité de traitement » correspondant à une utilisation sous-optimale des informations sensorielles disponibles demeure encore incertaine. Il pourrait correspondre à une augmentation du bruit interne, une diminution des capacités de stockage de la mémoire échoïque et/ou de la mémoire de travail, une altération des processus attentionnels, ou une dégradation des processus décisionnels et des stratégies d'écoute). L'impact de ces facteurs de haut niveau pourrait également expliquer la grande variabilité des résultats et l'influence d'effets d'entraînement et/ou d'effets procéduraux (réduits par entraînement).

Auteurs	Année	Porteuse	Fm (Hz)	Nombre de sujets	Tranches d'âge	Effet du vieillissement
Takahashi & Bacon	1992	Bande de Bruit	2-1024 par octave	40	21-33 50-59 60-69 70-76	**NON**
Purcell et al.	2004	Bruit	20-100; 100-600	38	18-43 60-78	**OUI**
He et al.	2008	Fc= 0.5, 4kHz	5, 40, 80, 100, 200, 250, 400, 500, 1000	16	18-26 60-80	**OUI**
Kumar & Sangamanatha	2011	Bruit gaussien	8, 20, 60, 200	176	20-30 31-40 41-50 51-60 61-70 71-85	**OUI**
Schoof & Rosen	2014	Bande de bruit (3kHz)	10, 20, 40, 80, 160	38	19-29 60-72	**NON**
Fullgrabe et al.	2015	Fc= 4kHz	5, 30, 90, 180	30	18-27 60-79	**OUI**
Paraouty et al.	2016	Fc= 0.5kHz	5	20	19-40 46-67	**NON**
Paraouty & Lorenzi	2017	Fc= 0.5kHz	5	70	20-70	**NON**
Whiteford et al.	2017	Fc= 0.5kHz	1, 20	85	20-79	**NON**

Tableau 1: *Tableau de synthèse présentant les paramètres expérimentaux des différentes études ayant évaluées l'effet du vieillissement sur la capacité à détecter une AM. Pour l'ensemble des études (à l'exception de Schoof & Rosen, 2015), le vieillissement tend à dégrader la sensibilité à l'AM. Toutefois, un effet significatif du vieillissement n'est observé que pour quatre d'entres elles. Ainsi, Takahashi & Bacon (1992) n'observent aucune différence significative entre la sensibilité à l'AM les sujets jeunes et les sujets plus âgés, mais relèvent une corrélation significative entre la sensibilité à l'AM et l'âge des sujets. He et al. (2008) observent un effet significatif du vieillissmeent sur la sensibilité à l'AM, qui tend à augmenter lorsque la cadence de modulation augmente. Cependant, l'effet délétère de l'âge semble pouvoir être "annulé" à basse cadence de modulation après entrainement. Fullgrabe et al. (2015) observent une dégradation globale de la sensibilité à l'AM, quelque soit la cadence de modulation. Ils en concluent que l'effet de l'âge sur la perception de l'AM pourrait résulter d'un déficit d'efficacité de traitement. Whiteford et al. (2017) n'observent aucun effet significatif de l'âge, hormis pour une cadence de 20Hz lorsque le signal est présenté de façon dichotique.*

4.2 Effet du vieillissement sur la détection de FM

Les études des effets du vieillissement sur la détection d'une FM sont peu nombreuses et ont été réalisées au cours de la dernière décennie.

Boettcher et al. (2002) réalisent l'une des premières études évaluant l'effet de l'âge sur la sensibilité à la FM. Ils enregistrent les potentiels évoqués en réponse à des stimuli de FM (appelés FMFR) pour différents paramètres expérimentaux (Fc=500 Hz; Fm=[4-38] Hz; Δf=[0;80]%) chez des sujets d'âge variable. L'effet de l'âge sur l'amplitude de la FMFR dépend de la cadence de modulation et de la profondeur de modulation. Pour des cadences de modulation lentes (4 Hz), l'amplitude de la FMFR est plus faible chez les sujets âgés. Pour des cadences de modulations plus rapides et pour des profondeurs de modulation plus importantes (40 à 70%), les résultats sont inversés. Ces données suggèrent que le vieillissement affecte le traitement de FM. Toutefois, en l'absence de mesures psychophysiques réalisées sur les mêmes sujets, il est impossible de savoir dans quelle mesure ces données neurophysiologiques rendent compte d'une éventuelle dégradation de la détection auditive de FM avec l'âge.

He et al. (2007) réalisent la première étude psychophysique mesurant les FMDLs (Fm = 5 Hz ; Fc =[500; 4000] Hz; modulation sinusoïdale ou quasi-sinusoïdale) pour huit sujets jeunes (âge moyen: 20 ans) et huit sujets âgés (71 ans) ayant des seuils audiométriques normaux (<20 dB HL) jusqu'à 4 kHz. Les FMDLs se dégradent avec l'âge quelle que soit le type de modulation. L'effet du vieillissement, mesuré par le rapport entre les FMDL des sujets âgés et jeunes, est plus élevé à 0.5 kHz qu'à 4 kHz. Les auteurs en concluent qu'à basse fréquence porteuse, la détection de FM se base sur les indices purement temporels issus du verrouillage de phase, et que l'âge affecte fortement ce mécanisme temporel.

Grose et Mamo (2012) mesurent les FMDLs (Fc≈500 Hz; Fm= 2 Hz; niveau de présentation de 65 dB SPL; présentation monaurale et binaurale) chez trois groupes de douze sujets normo-entendants: jeunes (19-29 ans), « modérément » âgés (43-57 ans) et âgés (65-77 ans). Pour

l'ensemble des conditions expérimentales, les FMDLs se dégradent avec l'âge. Le vieillissement dégrade les FMDLs monauraux par un facteur 1.4 chez les sujets modérément âgés et par un facteur 1.8 chez les sujets les plus âgés. L'effet délétère de l'âge semble légèrement plus faible que celui rapporté par He et al. (2007), ces derniers observant une dégradation des FMDLs par un facteur 2.4. Contrairement à He et al. (2008), Grose et Mamo (2012) introduisent une variation aléatoire de la fréquence porteuse entre les intervalles d'observation de façon à limiter l'utilisation des indices de pattern d'excitation. Cette variation aléatoire de la porteuse, si elle limite la possibilité de détecter la FM par comparaison de l'excitation estimée en un même point du pattern d'excitation entre les différents intervalles d'observation d'un même essai, n'empêche pas les sujets d'utiliser des indices de pattern d'excitation au sein d'un intervalle d'observation donné. L'effet délétère de l'âge sur la détection de la FM pourrait donc résulter d'une utilisation imparfaite des indices de pattern d'excitation par les sujets âgés, tout comme dans l'étude de He et al. (2007).

Sheft et al. (2012) évaluent les capacités de discrimination de patterns complexes de FM chez des sujets normo-entendants jeunes et âgés. Les patterns de FM sont synthétisés en modulant en fréquence un son pur de 1 kHz par une modulation obtenue en filtrant un bruit passe bas à 5 Hz. Les résultats montrent un effet significatif de l'âge sur la discrimination de la FM et suggèrent que l'âge est associé à une dégradation des capacités à discriminer des patterns de FM aux basses cadences de modulation.

Schoof et Rosen (2014) mesurent les FMDLs (Fc = 1000 Hz; Fm = 2 Hz; 70 dB SL) chez des auditeurs jeunes (19-29 ans) et âgés (60-72 ans). Ils ne retrouvent aucune différence significative entre les deux groupes d'âge. Cette absence d'effet d'âge contraste avec les études précédentes.

Kortlang et al. (2015) étudient l'effet du vieillissement sur les performances de sujets dans différentes tâches psychoacoustiques. Ils mesurent les FMDLs (modulation = bande de bruit filtré entre 1 et 4 Hz; Fc = [0.5;2;6] kHz) chez des sujets normo-entendants jeunes (n=6) et âgés

(n=6) avec et sans AM interférente, dans le silence et en présence de bruit masquant. A basse fréquence uniquement (500 Hz), les FDMLs des sujets normo-entendants âgés diffèrent significativement de ceux des sujets jeunes.

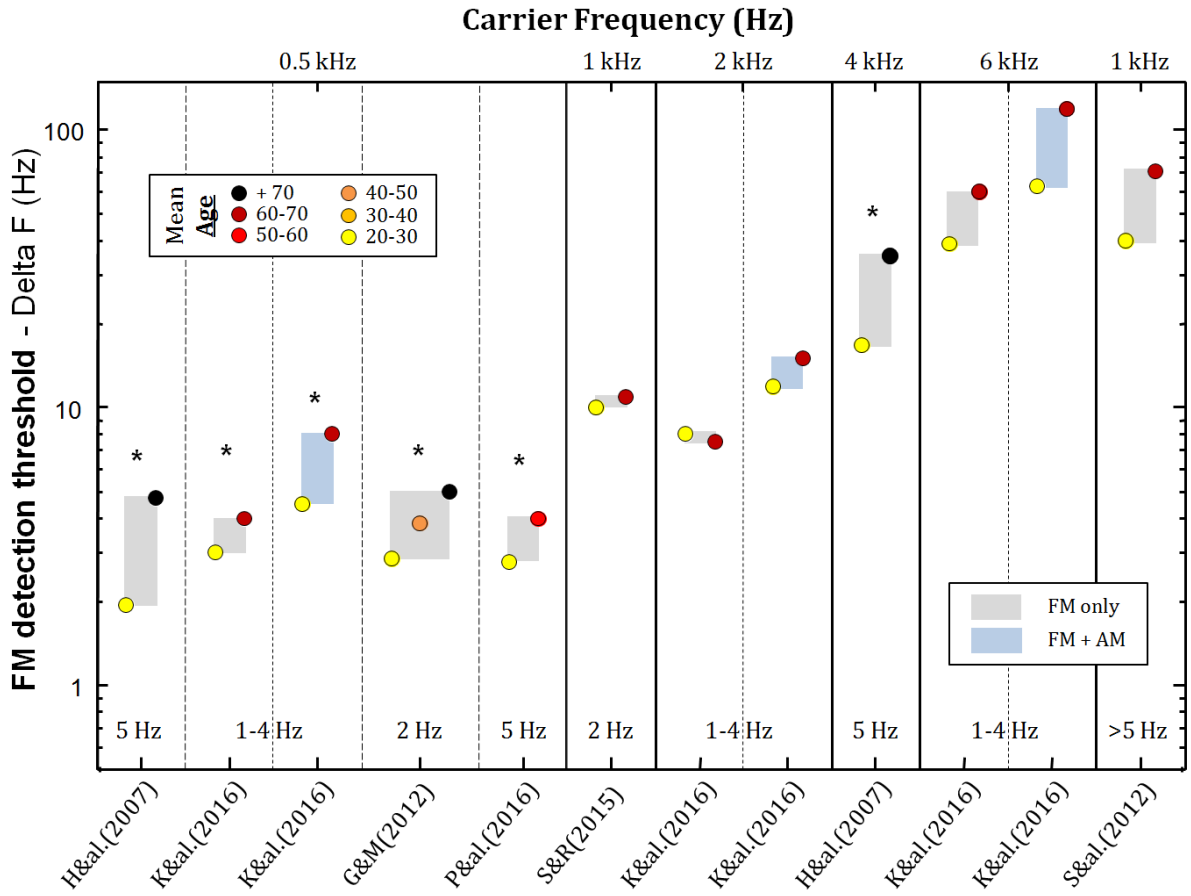


Figure 16: *FMDLs moyens obtenus par différentes études évaluant l'effet du vieillissement sur la perception de la FM. La couleur des cercles indique l'âge moyen des sujets. Les barres verticales précisent l'intervalle min/max pour chaque distribution. La couleur des barres indique la présence d'une AM interférente, introduite de façon à limiter l'utilisation des indices de pattern d'excitation. La cadence de modulation pour chaque condition est indiquée au dessus de l'axe des abscisses, la fréquence porteuse est précisée au dessus de chaque panel. Les seuils de détection de FM sont significativement affectés par l'âge à basse cadence de modulation et pour une porteuse rekativement basse. Les différentes études sont symbolisées par: H&al. (2007) - He et al. (2007); P&al. (2016) - Paraouty et al. (2016); H&al. (2008) - K&al. (2016); Kortlang et al. (2016); G&M (2012) - Grose & Mamo (2012); S&R (2015) - Schoof & Rosen (2015); S&al. (2012) - Sheft et al. (2012). Les étoiles au dessus des données indiquent les conditions pour lesquelles le vieillissement dégrade significativement les FMDLs.*

Johannesen et al. (2016) mesurent les FMDLs (Fc=1500 Hz; Fm=2 Hz; 30 dB SL) chez 68 sujets malentendants (25 à 82 ans) avec le même paradigme que Strelcyk et Dau (2009). L'effet de la perte auditive sur les FMDLs est décrit ultérieurement dans la section 5.2. Johannesen et al.

(2016) ne trouvent aucune corrélation entre l'âge des sujets et les FMDLs, suggérant que l'âge en soi n'influence pas les FMDLs chez les sujets malentendants.

Paraouty et Lorenzi (2017) ont mesuré les FMDLs (Fc=500 Hz; Fm =5 Hz) avec et sans AM interférente sur un large échantillon d'auditeurs normo-entendants (20-70 ans). Les auteurs obtiennent une corrélation significative entre les FMDLs et l'âge et ce avec et sans AM interférente, suggérant que la sensibilité à la FM se dégrade avec l'âge même lorsque les auditeurs sont contraints à utiliser des indices autres que les indices de pattern d'excitation. La corrélation entre les FMDLs (avec et sans AM interférente) et l'âge n'est toutefois plus significative lorsque l'on contrôle l'effet des seuils audiométriques tonaux. Cependant, lorsque l'étude corrélationnelle est restreinte à un sous-groupe de sujets présentant de grandes similitudes audiométriques (une perte tonale inférieure à 15 dB HL), une corrélation significative entre FMDLs et âge est obtenue lorsque les FMDLs sont mesurés avec une AM interférente.

Enfin, Whiteford et al. (2017) mesurent les FMDLs (Fc=500 Hz; Fm = [1;20] Hz; 60 dB SPL; présentation diotique et dichotique) chez 85 sujets d'âge variable (20-79 ans) ayant des seuils audiométriques normaux en dessous de 1 kHz. Les FMDLs sont corrélés avec l'âge pour l'ensemble des conditions expérimentales (diotique et dichotique; cadences lentes et rapides). Seule la corrélation entre l'âge et les FMDLs mesurés pour une basse cadence de modulation et en mode de présentation diotique demeure significative lorsque la sensibilité à l'AM est contrôlée dans l'analyse. Combinés aux résultats obtenus en détection d'AM (voir plus haut), ces résultats suggèrent que le vieillissement affecte davantage la sensibilité à la FM qu'à l'AM.

La figure 16 représente les FMDLs obtenus pour l'ensemble des études précédentes, en fonction de l'âge moyen des sujets, de la présence ou de l'absence d'AM interférente, de la cadence de modulation et de la fréquence porteuse. Les principaux paramètres expérimentaux et résultats des différentes études évaluant l'effet de l'âge sur la détection de FM sont présentés dans le tableau 2.

En grande majorité, les études rapportent un effet délétère de l'âge sur la détection de FM (He et al., 2007; Grose & Mamo, 2012; Grose et al., 2015; Sheft et al., 2012; Kortlang et al., 2015; Paraouty & Lorenzi, 2017; Whiteford et al., 2017). Seules quelques études (Paraouty & Lorenzi, 2017; Whiteford et al., 2017, et dans une certaine mesure, Grose & Mamo, 2012) ont examiné la contribution relative des indices temporels et des indices de pattern d'excitation au déficit de FM. Il en ressort que le vieillissement affecte la détection de FM plus fortement à basse fréquence, et pour des cadences de modulation relativement lentes (c'est à dire lorsque le code temporel est supposé prédominer). Chez les sujets malentendants, l'âge n'est jamais corrélé aux FMDLs (Johannesen et al., 2016; Strelcyk & Dau, 2009), suggérant que les effets du vieillissement et de la perte auditive sur les FMDLs ne sont pas de même nature.

Auteurs	Année	Porteuse	Masque d'AM	Fm (Hz)	Nombre de sujets	Tranches d'âge	Effet du vieillissement
He et al.	2007	Fc= 0.5, 4kHz	NON	5	20	19.8 70.8	**OUI**
Sheft et al.	2012	Fc= 1kHz	NON	noise LP at 5Hz	36	22-33 61-72	**NON**
Schoof & Rosen	2014	Fc= 1kHz	NON	2	38	18-26 60-80	**NON**
Kortlang et al.	2015	Fc= 0.5, 2, 6kHz	OUI	noise BP bw [1;4]	12	25-30 65-77	**OUI**
Johanessen et al.	2016	Fc= 1.5kHz	OUI	2	68	25-82	**NON**
Grose & Mamo	2012	Fc= [0.46;0.54]kHz	NON	2	36	19-29 43-57 65-77	**OUI**
Whiteford et al.	2017	Fc= 0.5kHz	NON	1, 20	85	20-79	**OUI**
Paraouty et al.	2017	Fc= 0.5kHz	OUI et NON	5	70	20-70	**OUI**

Tableau 2: *Tableau de synthèse présentant les paramètres expérimentaux des différentes études ayant évaluées l'effet du vieillissement sur la capacité à détecter une FM. Pour l'ensemble des études (à l'exception de Schoof & Rosen, 2015 et de Johannessen et al, 2016), le vieillissement tend à dégrader significativement la sensibilité à la FM. Toutefois, Sheft et al. (2012) n'observent pas d'effet du vieillissement sur la perception d'une FM dans le silence, mais obtiennent une dégradation des FMDLs dans le bruit. L'absence d'effet significatif de l'âge sur les FMDLs chez Schoof & Rosen (2015) et l'absence de corrélation entre l'âge des sujets et les FMDLs demeurent inexpliquée. Au global un effet délétère de l'âge est mesurable: i) tant dans le silence qu'en environnement bruyant; ii) avec ou sans un masque d'AM, venant contraindre l'utilisation des indices de pattern d'excitation. Par ailleurs, de nombreuses données évaluant la sensibilité aux indices de TFS confirment l'existence d'un effet délétère de l'âge (Füllgrabe et al., 2013; Füllgrabe et al., 2015; Hopkins & Moore, 2011)..*

5. Effets d'une lésion cochléaire sur la perception de l'AM et de la FM

La perte auditive peut avoir différentes étiologies chez l'homme. Une perte auditive est fréquemment associée à une atteinte cochléaire, communément appelée perte cochléaire. Généralement, elle survient consécutivement à l'exposition à de fortes expositions sonores, à l'administration de substances ototoxiques, à la survenue d'infections, des troubles métaboliques, des maladies auto-immunes ou génétiques. La cochlée et le système auditif peuvent également être affectés par le vieillissement (Willott et al., 2003). La susceptibilité aux pertes auditives induites par l'âge - la presbyacousie - est également influencée par des facteurs environnementaux et génétiques (Gates et al., 1999; Kujawa & Liberman, 2006; Van Eyken et al., 2007). Une perte auditive peut également être associée à un dysfonctionnement des synapses des cellules ciliées internes, ou à une dégénérescence des fibres du nerf auditif (Kujawa & Libermann, 2006, 2009). Quand la perte auditive associe une atteinte cochléaire et une atteinte neurale, le terme de perte auditive neurosensorielle est alors utilisé. Ce travail de recherche porte principalement sur les effets engendrés par une perte auditive cochléaire et/ou neurosensorielle.

Une perte cochléaire affecte substantiellement le fonctionnement des cellules ciliées externes et internes, produisant de nombreuses conséquences fonctionnelles: i) une diminution de la sensibilité aux sons faibles, l'intensité sonore devant être augmentée pour stimuler de façon similaire la membrane basilaire (Sellick et al. 1982); ii) un élargissement des courbes d'accord (Evans, 1975) et; iii) une disparition progressive de la non linéarité cochléaire (Robles et Ruggero, 2001). Sur le plan auditif, ceci se manifeste par une perte d'audibilité (légère, modérée, sévére, ou profonde), une perte de sélectivité fréquentielle (les filtres cochléaires sont élargis par un facteur 2-5), et un phénomène de recrutement de sonie, à savoir un accroissement anormalement rapide de l'intensité perçue en fonction de l'intensité physique (pour une présentation détaillée, voir Moore, 2007).

5.1 Effet d'une perte auditive sur la détection d'AM

En théorie, la sensibilité à l'AM chez les sujets présentant une surdité neurosensorielle devrait être influencée par : i) la perte de sélectivité fréquentielle, la réponse impulsionnelle d'un filtre étant d'autant plus courte que la bande passante du filtre augmente ; ii) la perte de la compression d'amplitude instantanée causée par la contraction des cellules ciliées externes. Ainsi, la sensibilité à l'AM et la résolution temporelle (la plus haute cadence d'AM détectable) des sujets atteints d'une perte neurosensorielle pourraient être paradoxalement meilleure que celle des sujets normo-entendants.

Formby (1982, 1987) mesurent les MTFs sur l'oreille saine et atteinte chez des patients présentant une surdité unilatérale (syndrome de Ménière). La porteuse est un bruit large bande. La sensibilité à l'AM est identique sur les deux oreilles. La MTF de l'oreille atteinte présente une fréquence de coupure de 100 Hz, et une pente de -6 dB/octave. Par comparaison, la MTF de l'oreille saine présente une fréquence de coupure similaire, mais une pente de -3dB/octave. Formby (1986) en conclut que les patients présentant une maladie de Ménière peuvent détecter normalement une AM à basse cadence de modulation, mais que leur sensibilité à l'AM s'amoindrie plus rapidement lorsque la cadence de modulation s'accroît.

Bacon et Viemeister (1985) mesurent la MTF chez quatre sujets normo-entendants (20-31 ans) et six malentendants (19-68 ans mais 4 sujets ont un âge<35 ans) présentant une perte auditive en haute fréquence. La porteuse est un bruit large bande. Les MTFs indiquent que: i) les sujets malentendants sont moins sensibles à l'AM que les sujets normo-entendants; ii) la sensibilité à l'AM varie en fonction du niveau de présentation : les AMDLs se dégradent lorsque l'intensité de stimulation diminue, principalement aux hautes cadences de modulation (>100 Hz), contrairement aux résultats obtenus par Viemeister (1979) sur des sujets normo-entendants. La fréquence de coupure des MTFs est inférieure à 40 Hz (la constante de temps ou acuité temporelle associée est donc $\tau \approx 4$ ms), et la pente du filtre est d'environ -10 dB/octave. Bacon et

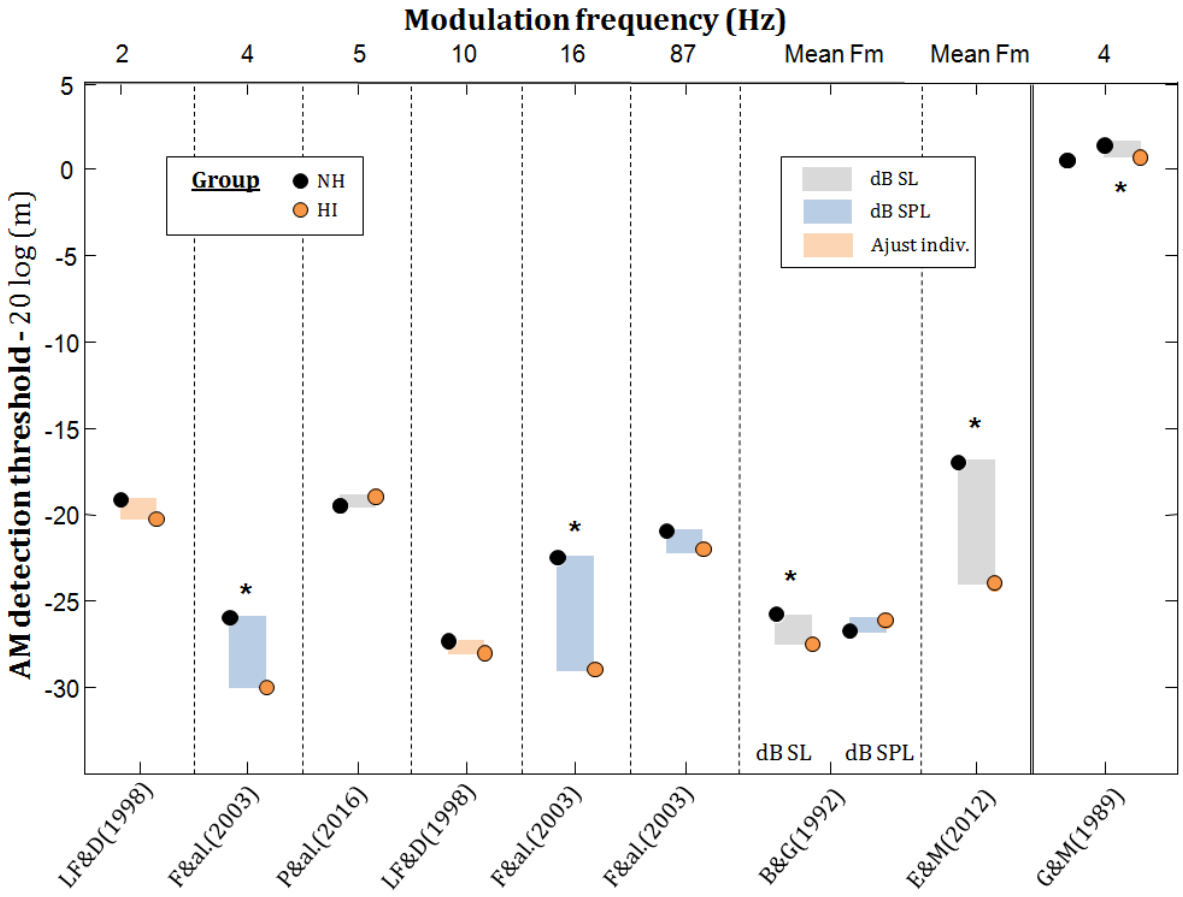


Figure 17: *AMDLs moyens obtenus par différentes études évaluant l'effet de la perte auditive sur la perception de l'AM. La couleur des cercles indique le groupe des sujets. Les barres verticales précisentla différence entre normo-entendants et malentendants. La couleur des barres indiquele niveau de stimulation utilisé (égalisaton en dB SL, présnetation à dB SPL fixe, ...). La cadence de modulation pour chaque condition est indiquée au dessus de l'axe des abscisses, la fréquence porteuse est précisée au dessus de chaque panel. Les seuils de détection d'AM sont significativement meilleurs en présence d'une lésion cochléaire pour la plupart des études. Les différentes études sont symbolisées par: LF&D. (1998) - Lacher-Fougeres & Demany (1998); P&al. (2016) - Paraouty et al. (2016); F&al. (2003) - Füllgrabe et al. (2003); G&M (1989) - Glasberg & Moore (1989); E&M (2012) - Ersnt & Moore (2012); B&G (1992) - Bacon & Gleitmann (1992). Les étoiles au dessus des données indiquent les conditions pour lesquelles la perte auditive améliore significativement les AMDLs.*

Viemeister (1985) mesurent également les MTFs dans des conditions similaires chez des sujets normo-entendants, mais en filtrant en passe bas à 1.6 kHz le bruit modulé, de façon à simuler la perte auditive en haute fréquence audio. Les MTFs des sujets normo-entendants obtenues dans cette condition filtrée en passe bas sont similaires à celles mesurées chez les sujets malentendants en bruit large bande, suggérant que le déficit temporel relevé chez les sujets

malentendants résulte d'une diminution de la bande passante effective (la bande passante « audible »), en accord avec les données d'Eddins (1993). Malheureusement, les effets d'âge et les effets de perte auditive sont confondus dans cette étude.

Glasberg et Moore (1989) mesurent les AMDLs (Fc= 0.5, 1 ou 2 kHz; Fm = 4 Hz ; mesures à niveau égal en dB SPL ou en dB SL) de 9 sujets (43-73 ans) présentant une surdité unilatérale et de 6 sujets (19-70 ans) présentant une perte auditive bilatérale. Les AMDLs (moyenne sur les différentes porteuses) des auditeurs normo-entendants et des auditeurs malentendants sont similaires lorsque le niveau de stimulation est égalisé en dB SPL. A même niveau SL, les AMDLs des auditeurs malentendants sont significativement meilleurs que ceux des auditeurs contrôles.

Bacon et Gleitman (1992) mesurent les MTFs (porteuse: bande de bruit large; Fm = [2-1024] Hz; égalisation en dB SPL ou en dB SL) de 5 sujets normo-entendants (22-29 ans) et 8 sujets malentendants (11-63 ans) présentant des pertes audiométriques relativement plates. Quel que soit le niveau de présentation (égalisé en dB SPL ou dB SL), les MTFs des auditeurs normo-entendants et malentendants sont similaires. Pour un niveau de présentation faible (20 dB SL), les MTFs des auditeurs normo-entendants et malentendants diffèrent sensiblement, les auditeurs malentendants présentant des AMDLs bien meilleurs que ceux des auditeurs normo-entendants, quelle que soit la cadence de modulation, et ce en accord avec Glasberg et Moore (1990).

Moore et al. (1992) comparent les MTFs (porteuse: bande de bruit centrée sur 2 kHz; Fm = [4-512] Hz; égalisation du niveau en dB SPL ou en dB SL) de l'oreille saine et de l'oreille atteinte chez trois sujets (59-76 ans) présentant une perte auditive unilatérale. Les AMDLs obtenus aux deux oreilles sont similaires quelle que soit l'intensité de présentation du signal, suggérant que la lésion cochléaire n'affecte pas significativement la sensibilité à l'AM lorsque la contribution des facteurs de haut niveau est contrôlée (l'efficacité de traitement est identique entre mesures sur oreille saine et malentendante, chaque sujet étant ici son propre contrôle). Moore et al. (1992) testent également trois sujets malentendants supplémentaires (68-74 ans) présentant

une perte auditive bilatérale dans des conditions expérimentales similaires. A nouveau, les MTFs sont équivalentes à celles des sujets précédents (oreilles saines et malentendantes).

Moore et al. (1996) réalisent l'une des premières études évaluant l'impact d'une atteinte neurosensorielle sur le sensibilité à l'AM, en mesurant parallèlement i) la fonction de sonie (de façon à évaluer l'impact du recrutement consécutif à la perte du mécanisme actif cochléaire); ii) la sensibilité à l'AM. Ils évaluent la perception de sons purs modulés en AM (Fc= 1 kHz; Fm = 4, 8, 16 et 32 Hz) chez trois sujets (âges non communiqués) présentant une surdité unilatérale. La tâche d'ajustement consiste à égaliser la saillance des fluctuations produites par l'AM entre les deux oreilles. Les résultats indiquent que la profondeur de modulation sur l'oreille saine doit être significativement augmentée comparativement à l'oreille malentendante, quelle que soit la cadence de modulation. Ces résultats suggèrent que l'amélioration de la sensibilité à l'AM généralement observée chez les sujets malentendants peut être expliquée par le phénomène de recrutement de sonie et la magnification perceptive des fluctuations d'enveloppe causée par la perte de la compression cochléaire.

Grant et al. (1998) mesurent les seuils de détection et de discrimination d'une AM (porteuse: bande de bruit; Fm = 80, 160, 320 Hz) chez des sujets normo-entendants et malentendants. Les seuils de détection et de discrimination sont fortement affectés par la perte cochléaire. Toutefois, ces différences de traitement temporel entre les sujets normo-entendants et malentendants peuvent s'expliquer, au moins pour partie, par les différences d'audibilité du signal, et ce en particulier dans la région haute fréquence du bruit modulé, conformément aux résultats de Bacon et Vieimeister (1985).

Füllgrabe et al (2003) mesurent les AMDLs (Fc= 2 kHz; Fm = 4, 16, 64, 87 Hz; niveau de présentation de 40 dB SL) chez 9 sujets normo-entendants (âge moyen=22 ans) et 7 sujets malentendants (53-77 ans; âge moyen=66 ans) présentant une perte de 40-45 dB HL à la fréquence testée. En moyenne, les AMDLs obtenus chez les sujets malentendants sont meilleurs (de 3-8 dB) que ceux des sujets normo-entendants aux cadences de modulation les plus lentes (4

et 16 Hz). Pour des cadences de modulation plus rapides (87 Hz), les AMDLs sont globalement similaires entre les deux groupes. Les auteurs concluent que la meilleure sensibilité à l'AM des sujets malentendants peut être expliquée par le recrutement de sonie.

Enfin, Paraouty et al. (2016) mesurent les AMDLs avec ou sans FM interférente (Fc= 0.5 kHz; Fm = 5 Hz; niveau de présentation de 40 dB SL) chez 10 sujets normo-entendants (46-70 ans) et 10 sujets malentendants (46-64 ans) présentant une surdité légère à moyenne. Les AMDLs ne diffèrent pas significativement entre les groupes, quelle que soit la profondeur de modulation de la FM interférente.

Les résultats d'études électrophysiologiques réalisées chez le chinchilla par Heinz et coll. fournissent plusieurs corrélats neurophysiologiques périphériques du rehaussement perceptif des indices d'AM observé chez les patients malentendants. Kale et Heinz (2010, 2012) observent une amélioration significative de la représentation neurale de l'AM dans les patterns de décharge des fibres du nerf auditif consécutivement à une perte auditive produite par un traumatisme sonore. Kale et Heinz (2010, 2012) suggèrent que cette amélioration de la représentation des indices d'enveloppe s'explique par: i) la perte de compression cochléaire au niveau de la membrane basilaire; ii) l'augmentation de la pente de la fonction entrée-sortie des cellules ciliées internes (Heinz & Young, 2004). D'autres résultats indiquent qu'un changement de l'adaptation nerveuse à court terme dans les fibres du nerf auditif pourrait également être à l'origine du rehaussement de la représentation neurale des indices d'enveloppe temporelle (Scheidt et al., 2010).

La figure 17 représente les AMDLs obtenus pour l'ensemble des études précédentes, en fonction de la cadence de modulation, du niveau de stimulation utilisé et du statut audiométrique des sujets. Les principaux paramètres expérimentaux et résultats des différentes études évaluant l'effet de la perte auditive sur la détection d'AM sont présentés dans le tableau 3.

En conclusion, les études précédentes démontrent que la perte auditive d'origine neurosensorielle: i) préserve la sensibilité à l'AM lorsque la comparaison entre sujets normo-entendants et malentendants est effectuée à même niveau SPL (Moore et al., 1992; Glasberg & Moore, 1990; Bacon & Viemeister, 1985); ii) améliore la sensibilité à l'AM lorsque la comparaison est réalisée au même niveau SL, et ce d'autant plus que le niveau de stimulation est faible (Füllgrabe et al., 2003; Moore et al., 1996; Bacon & Gleitman, 1992; Glasberg & Moore, 1990). L'amélioration de la sensibilité à l'AM associée à une atteinte cochléaire semble principalement s'expliquer par la perte du mécanisme actif cochléaire et par sa manifestation auditive : le recrutement de sonie (e.g. Füllgrabe et al., 2003; Moore et al., 1996). Cette interprétation est corroborée par des résultats d'études électrophysiologiques menées dans le nerf auditif d'animaux cochléo-lésés par traumatisme acoustique. Une modification de la transduction et de l'adaptation nerveuse à court terme pourrait également être responsable du rehaussement de la perception des indices d'AM chez les sujets humains malentendants.

Auteurs	Année	Porteuse	Fm (Hz)	Niveau de présentation	Nombre de sujets	Seuils audiométriques (dBHL)	Effet d'une lésion cochléaire
Bacon & Gleitmann	1992	Bande de Bruit	2-1024 par octave	Même SL Même SPL	13	PTA =[-8.3;1.7] PTA = [18.3;45]	**OUI (+/SL)**
Moore	1992	Bande de bruit (2 kHz)	4-512 par octave	Même SL Même SPL	6	PTA <25 PTA = [53.7;62.5]	**NON**
Bacon & Viemeister	1985	Bande de Bruit	2-1024 par octave	Même puissance spectrale	10	PTA <20 PTA = [28.7;50]	**OUI (-)**
Lacher-Fouchère & Demany	1998	Fc= 0.5 kHz	1, 10	Ajustement indiv.	11	PTA <15 PTA = [41;72]	**NON**
Glasberg & Moore	1989	Fc= 0.5, 1, 2 kHz	4	Même SL Même SPL	15	Thres =[0;45] Thres =[30;80]	**OUI (+/SL)**
Moore et al.	1996	Fc= 1 kHz	4, 8, 16, 32	Ajustement indiv.	3	PTA <27 PTA >60	**OUI (+)**
Grant et al.	1998	Bruit gaussien	80, 160, 320	Même SPL	12	PTA <15 Thres =[25;95]	**OUI (-)**
Füllgrabe et al.	2003	Fc= 2 kHz	4, 16, 67, 87	70-85 dBSPL	16	PTA <20 PTA = [27;55]	**OUI (+/fm<16Hz)**
Ernst & Moore	2012	Fc= 1, 4, 6 kHz	2, 10	Même SL Même SPL	10	PTA <15 PTA = 42	**OUI (+/SL)**
Paraouty et al.	2016	Fc= 0.5kHz	5	Même SL	20	PTA = 12 (+/-3) PTA = 35 (+/-8)	**NON**

Tableau 3: *Tableau de synthèse présentant les paramètres expérimentaux des différentes études ayant évaluées l'effet d'une lésion cochléaire sur la capacité à détecter une AM. Dans la colonne "Effet d'une lésion cochléaire", un signe positif indique un effet bénéfique de la perte auditive sur la sensibilité à l'AM, tandis qu'un signe négatif indique un effet délétère. Les conditions pour lesquelles un effet positif est observé sont précisées entre parenthèses. Au final, la sensibilité à l'AM n'est pas significativement affectée par la lésion cochléaire lorsque les AMDLs sont mesurés à même niveau SPL. Lorsque la comparaison est réalisée en égalisant le niveau de présentation en dB SL, la perte auditive améliore la sensibilité à l'AM.*

5.2 Effet d'une perte auditive neurosensorielle sur la détection de FM

L'effet d'une lésion cochléaire sur la détection d'une FM a fait l'objet de nombreux travaux au cours des dernières décennies (e.g., Zurek & Formby, 1981; Moore & Glasberg, 1986; Grant, 1987; Lacher-Fougere & Demany, 1998; Moore & Skrodzka, 2002; Buss et al., 2004; Strelcyk & Dau, 2009; Ernst & Moore, 2012; Kortlang et al., 2016; Johannesen et al., 2016; Paraouty et al., 2016).

Zurek et Formby (1981) réalisent l'une des premières études évaluant l'effet d'une lésion cochléaire sur la perception de la FM. Les FMDLs sont mesurés chez 8 sujets malentendants (Fm=3 Hz; Fc=[125;4000] Hz) à une intensité de 25 dB SL. Les FMDLs se dégradent lorsque la perte auditive s'accroit. A perte auditive comparable, la dégradation des FMDLs est plus importante en basse fréquence qu'en haute fréquence audio. Zurek et Formby (1981) expliquent cet effet différentiel de la perte auditive sur les FMDLs par l'existence de deux codes sensoriels pour la FM et suggèrent que la perte auditive affecterait plus fortement le code temporel opérant essentiellement en basse fréquence, que celui utilisant le pattern d'excitation et opérant essentiellement en haute fréquence.

Moore et Glasberg (1986) obtiennent des résultats similaires chez des sujets présentant une surdité unilatérale. Ils mesurent les FMDLs pour une cadence de 4 Hz à 80 dB SPL (10 dB SL à minima). Sur l'oreille malentendante, les FMDLs augmentent fortement dans la région basse fréquence (dégradation par un facteur 3.8 à 0.5 kHz versus 1.5 à 2 kHz) par comparaison avec les FMDLs de l'oreille saine, et ce malgré une perte auditive (telle qu'estimée par l'audiométrie tonale) similaire aux deux fréquences. Les FMDLs des sujets malentendants sont relativement bien simulés par le modèle de Zwicker, indiquant que les sujets malentendants utilisent uniquement des indices de pattern d'excitation lorsqu'ils détectent une FM.

Grant (1987) réalise la seule estimation connue de FMDLs (Fc=[100;1000] Hz; Fm=3 Hz; modulation triangulaire) chez 3 sujets présentant une surdité profonde (3 sujets normo-

entendants en contrôle). Les sujets normo-entendants sont testés à 30 dB SL, tandis que les sujets malentendants sont testés à un niveau "jugé confortable" (110 à 135 dB SPL). La perte auditive profonde altère les FMDLs par un facteur 9.5. Les FMDLs sont également mesurés en présence d'une AM interférente (AM de bruit filtré passe bas à 3 Hz), de façon à restreindre l'utilisation d'indices de pattern d'excitation. Conformément aux prédictions du modèle de Zwicker, l'introduction de l'AM interférente dégrade les FMDLs. L'effet délétère de l'AM interférente est plus important chez les sujets malentendants que chez les sujets normo-entendants (dégradation des FMDLs par un facteur 16 en présence de l'AM interférente).

Lacher-Fougere et Demany (1998) mesurent les FMDLs (Fc= 500 Hz; Fm = 2 ; 10 Hz) à un niveau jugé confortable chez cinq normo-entendants et sept sujets malentendants présentant des seuils audiométriques tonaux compris entre 30 et 75 dB HL à la fréquence testée. Les FMDLs sont 2 fois plus élevés chez les sujets présentant des pertes inférieures à 45 dB HL que chez les sujets normo-entendants. Les sujets porteurs de pertes plus importantes présentent des FMDLs jusqu'à dix fois supérieurs à la normale. L'effet de la perte auditive est similaire aux deux cadences de modulation. Lacher-Fougere et Demany (1998) suggèrent que la perte auditive dégrade tant les indices de pattern d'excitation que les indices temporels évoqués par la FM.

Moore et Skrodzka (2002) répliquent l'étude de Grant (1987) pour un grand nombre de fréquences porteuses (Fc = 0.25;0.5;1; 2; 4 et 6 kHz) et différentes cadences de modulation (Fm = 2; 5; 10 et 20 Hz). Ils mesurent les FMDLs à 85 dB SPL chez trois sujets normo-entendants jeunes, et quatre sujets malentendants âgés, avec et sans AM interférente. Les sujets malentendants présentent une perte de sélectivité fréquentielle. La perte auditive est globalement associée à une élévation des FMDLs. L'effet délétère de l'AM interférente sur les seuils FMDLs: i) est plus élevé chez les sujets malentendants que les sujets normo-entendants, en accord avec les résultats de Grant (1987); ii) reste relativement stable lorsque la cadence de modulation augmente, contrairement aux résultats obtenus sur les sujets normo-entendants. Moore et Skrodzka (2002) suggèrent que les sujets malentendants n'utilisent pas (ou peu) les indices temporels de verrouillage en phase pour la détection d'une FM.

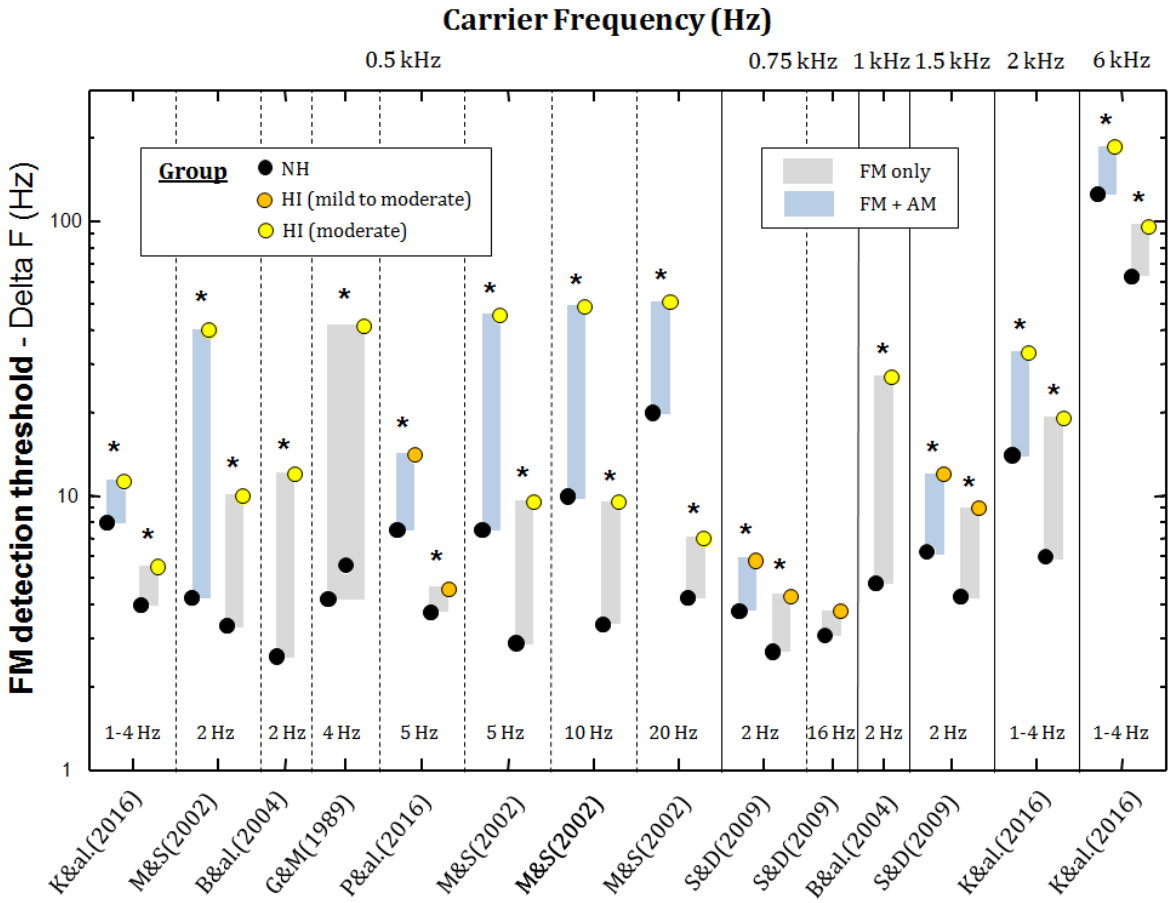


Figure 18: *FMDLs moyens obtenus par différentes études évaluant l'effet de la perte auditive sur la perception de la FM. La couleur des cercles indique les seuils audiométriques moyens des sujets. Les barres verticales précisent l'effet de la perte auditive. La couleur des barres indique la présence d'une AM interférente, introduite de façon à limiter l'utilisation des indices de pattern d'excitation. La cadence de modulation pour chaque condition est indiquée au dessus de l'axe des abscisses, la fréquence porteuse est précisée au dessus de chaque panel. Les seuils de détection de FM sont significativement affectés par la lésion cochléaire, quelque soit la cadence de modulation. Les différentes études sont symbolisées par: M&S (2002) - Moore & Skrodzka (2002); P&al. (2016) - Paraouty et al. (2016); G&M (1989) - Glasberg & Moore (1989); K&al. (2016); Kortlang et al. (2016); S&D (2009) - Strelcyk & Dau (2009); B&al. (2004) - Buus et al. (2004). Les étoiles au dessus des données indiquent les conditions pour lesquelles la perte auditive dégrade significativement les FMDLs.*

Strelcyk et Dau (2009) mesurent les FMDLs (Fm = [2;16] Hz; Fc = [125;750;1500] Hz) à 30 dB SL et 70 dB SPL chez six sujets normo-entendants et dix sujets malentendants présentant une perte audiométrique en haute fréquence. Les FMDLs sont similaires aux deux niveaux de présentation, et sont 1.5 fois plus élevés (moins bons) chez les sujets malentendants que chez les sujets normo-entendants. Les FMDLs sont également mesurés en présence d'une AM interférente.

Contrairement aux résultats de Moore et Skrodzka (2002), l'AM a un effet délétère sur les FMDLs similaire pour les deux groupes.

Ernst et Moore (2012) montrent que les FMDLs et AMDLs mesurés chez cinq sujets malentendants sont meilleurs à 10 Hz qu'à 2 Hz au delà de 4 kHz. A 1 kHz, les AMDLs restent meilleurs à 10 Hz qu'à 2 Hz, tandis que les FMDLs sont similaires aux deux cadences. Pris ensemble, ces résultats suggèrent que: i) la détection de FM pour des fréquences porteuses élevées est principalement basée sur des indices de pattern d'excitation pour les sujets malentendants; ii) la détection de FM pour des fréquences porteuses basses et des cadences de modulation lentes est basée sur des indices temporels de verrouillage en phase, utilisables par les sujets malentendants lorsque la perte est relativement légère (<40 dB HL).

Kortlang et al. (2016) étudient également les effets d'une lésion cochléaire sur la détection de FM. Les sujets testés sont soit jeunes et normo-entendants (n=6 ; 25-30 ans), âgés et normo-entendants (n=6 ; 65-77 ans) ou âgés et malentendants (n=6 ; 60-82 ans). Les FMDLs sont mesurés pour un bruit de FM dont les composantes sont comprises entre 1 et 4 Hz. Pour l'ensemble des conditions expérimentales (porteuses de 0.5, 2 et 6 kHz, présence ou absence d'AM interférente, présence ou absence de bruit masquant), les sujets malentendants âgés obtiennent des FMDLs significativement plus élevés que ceux de sujets normo-entendants jeunes et âgés.

Enfin, Johannesen et al. (2016) mesurent les FMDLs (Fc=1500 Hz; Fm=2 Hz; 30 dB SL) chez 68 sujets malentendants en utilisant le paradigme de Strelcyk et Dau (2009). Les FMDLs sont similaires à ceux de Strelcyk et Dau (2009), et confirment l'altération des FMDLs sur une plus grande cohorte de sujets malentendants. Les auteurs en concluent que ce déficit pourrait résulter de: i) la perte de sélectivité fréquentielle associée à une atteinte cochléaire, qui elle-même dégraderait la représentation périphérique des indices de pattern d'excitation; ii) un déficit dans l'encodage.

La figure 18 représente les FMDLs obtenus pour l'ensemble des études précédentes, en fonction de la perte auditive moyenne des sujets, de la présence d'une AM interférente, de la cadence de modulation et de la fréquence porteuse. Les principaux paramètres expérimentaux et résultats des différentes études évaluant l'effet de la perte auditive sur la détection de FM sont résumés dans le tableau 4.

En conclusion, une perte auditive neurosensorielle semble affecter les deux codes sensoriels impliqués dans la détection d'une FM. Chez les sujets malentendants, la détection de FM aux cadences de modulation rapides (supérieures à 10 Hz) dépend fortement des indices de pattern d'excitation. Ceci semble résulter de l'élargissement des filtres auditifs associé à la perte auditive. Pour des cadences de modulation plus lentes, la détection de la FM semble dépendre d'informations temporelles de TFS encodées par verrouillage en phase. Toutefois, la capacité à utiliser ces indices temporels semble fortement contrainte par la sévérité de la perte auditive, et pourrait être abolie chez certains sujets malentendants.

Auteurs	Année	Porteuse (kHz)	Masque d'AM	Fm (Hz)	Nombre de sujets malentendants (normo)	Seuils audiométriques	Effet d'une perte auditive
Zurek & Formby	1981	0.125, 4	NON	3	8 (2)	Thres =[0;20] Thres =[5;90]	**OUI***
Moore & Glasberg	1986	0.5, 1, 2	NON	4	17 (10)	PTA <20 PTA ≈ [40;60]	**OUI**
Grant	1987	0.1, 1	NON/OUI	3	3 (3)	PTA <20 PTA ≈ [101;-]	**OUI**
Lacher-Fougere & Demany*	1998	0.5	NON	2, 10	7 (5)	PTA <15 PTA = [43;72]	**OUI**
Moore & Skrodzka*	2002	0.25, 0.5, 1, 2, 4, 6	NON/OUI	2, 5, 10, 20	4 (3)	PTA <20 PTA = [41;57]	**OUI**
Buss et al.	2004	0.5, 1	NON	2	12 (14)	PTA <15 PTA = [35;67]	**OUI**
Strelcyk & Dau*	2009	0.125, 0.75, 1.5	NON/OUI	2, 16	10 (6)	Thres =[0;20] Thres =[5;70]	**OUI***
Ernst & Moore*	2012	1, 4, 6	NON/OUI	2, 10	5 (5)	PTA <15 PTA ≈ 41	**OUI**
Paraouty et al.	2016	0.5	NON/OUI	5	10 (10)	PTA <15 PTA = 43	**OUI**
Kortlang et al.	2016	0.5, 1, 6	NON/OUI	[1;4]	6 (6)	PTA <15 PTA = [41;52]	**OUI**
Johanessen et al.	2016	1.5	NON	2	68	-	**NON**

Tableau 4: *Tableau de synthèse présentant les paramètres expérimentaux des différentes études ayant évaluées l'effet de la perte auditve sur la capacité à détecter une FM. Pour l'ensemble des études (à l'exception de Johanessen et al., 2016), la perte auditive tend à dégrader significativement la sensibilité à la FM. Les astérisques derrière les noms d'auteurs précisent les études pour lesquelles les effets du vieillissement et les effets de la perte auditive sont confondus. La perte auditive dégrade la sensibilité à la FM, que ce soit avec ou sans AM interférente. L'effet de la perte auditive n'est parfois significatif qu'à basse fréquence porteuse (Zurek & Formby, 1981) ou pour des cadences de modulation relativement lentes (Sterlcyk & Dau, 2009).*

5.3 Causes possibles d'une dégradation de la perception des indices temporels

Les mécanismes impliqués dans l'encodage et le décodage (la « lecture ») des informations de TFS véhiculées par la FM semblent affectés par la perte auditive et par le vieillissement. Toutefois, ceci fait encore l'objet de nombreux débats. Différentes hypothèses ont été avancées au cours des dernières décennies pour expliquer la dégradation des informations de TFS (pour une revue détaillée, voir également Moore, 2014):

1. Les informations de TFS pourraient être décodées grâce à une comparaison (un processus de corrélation croisée) entre les trains de potentiels d'action émis par des neurones du nerf auditif présentant des fréquences caractéristiques différentes (Shamma, 1985; Carney et al., 2002; Carlyon et al., 2012). Les différences de phase dans les réponses temporelles des neurones en différents points de la membrane basilaire peuvent être utilisées pour estimer la fréquence du stimulus à partir de ce processus post-cochléaire de corrélation. En modifiant la réponse temporelle en différents points de la membrane basilaire (Ruggero, 1994), les dommages cochléaires pourraient ainsi affecter le processus de décodage central des indices temporels de TFS. Heinz et al. (2010) démontrent que la propagation de l'onde acoustique le long de la membrane basilaire est plus rapide à la suite d'une perte auditive. Ceci altère fortement les différences de phase dans les réponses temporelles en différents points de la membrane basilaire, et pourrait rendre inefficaces les mécanismes plus centraux chargés du décodage des informations de TFS.
2. Les informations de TFS pourraient être décodées par un mécanisme « dépendant de la place ». En d'autres termes, le codage temporel et spatial de la fréquence pourraient être liés. En effet, dans une oreille saine, il existe une correspondance entre le pic d'excitation sur la membrane basilaire (et ainsi, les informations d'enveloppe) et les informations de TFS encodées par verrouillage en phase en ce point précis de la membrane basilaire. En

raison de la perte du mécanisme actif, la perte auditive provoque une modification de la tonotopie (la cartographie fréquentielle sur la membrane basilaire dans la cochlée) et un déplacement du pic d'excitation en réponse au son (Sellick et al., 1982; Liberman & Dodds, 1984). Ceci modifie la bonne correspondance entre codage temporel et spatial de la fréquence, ou plus généralement, entre codage de TFS et codage d'enveloppe (Henry & Heinz, 2012; 2013). Une telle distorsion (une perte de « tonotopicité ») pourrait également rendre inefficaces les mécanismes centraux chargés du décodage des informations fréquentielles.

3. En sortie des filtres cochléaires, les signaux à bandes étroites contiennent une information de FM variant relativement lentement. La perte du mécanisme actif cochléaire, en provoquant une perte de sélectivité fréquentielle (Glasberg & Moore, 1986; Moore, 2007), modifie la bande passante en sortie des filtres cochléaires et engendre des variations plus rapides dans les patterns de FM en chaque point de la membrane basilaire. Moore et Sek (1995) suggèrent que le mécanisme de décodage des informations de TFS est lent (« sluggish »). Ainsi, le mécanisme de décodage des indices de TFS pourrait devenir inefficace à la suite de dommages cochléaires.
4. Les mécanismes de décodage des indices de TFS pourraient dépendre d'un processus d'inhibition latérale entre des neurones de fréquences adjacentes (Shamma, 1985; Gleich et al., 2003). Salvi et al. (2000) suggèrent que des changements centraux surviennent consécutivement à une perte auditive et conduisent à une réduction des mécanismes d'inhibition. Ce changement pourrait affecter le décodage central des informations de TFS.
5. En haute fréquence, les neurones du nerf auditif encodent principalement les informations d'enveloppe temporelle tandis qu'en basse fréquence, ils encodent principalement les informations de TFS. A la suite d'un traumatisme acoustique, la capacité des neurones du nerf auditif à être entraînés par l'enveloppe ou la TFS du signal est modifiée au bénéfice de l'enveloppe (Kale & Heinz, 2010, 2012), et ce probablement

en raison de la perte de compression sur la membrane basilaire ou d'une modification de la fonction entrée-sortie des cellules ciliées internes. Ceci engendre un déficit « relatif » du codage de la TFS des signaux acoustiques.

6. La perte auditive est associée à une réduction du nombre de cellules ciliées internes, de fibres afférentes et de neurones auditifs (Kujajwa & Libermann, 2009; Makary et al. 2011; Sergeyenko et al., 2013 ; Lobarinas et al., 2013). La réduction du nombre de neurones transmettant les informations de TFS par verrouillage en phase pourrait se traduire par une augmentation de la stochasticité neuronale globale (une augmentation du bruit neural interne) qui, à son tour, pourrait affecter la représentation neuronale de la TFS à un niveau plus central.
7. La gaine de myéline du nerf auditif peut être modifiée par le vieillissement (Bartzokis, 2004). Ceci devrait modifier le temps de conduction des potentiels d'action et altérer les processus de décodage des informations de TFS.

6. Intégration temporelle

L'intégration temporelle (ou sommation temporelle) se réfère au fait que les seuils de détection auditive s'améliorent à mesure que la durée de stimulation augmente. L'intégration temporelle est observée pour un grand nombre de signaux acoustiques et pour des conditions expérimentales très diverses (pour une revue, voir Gerken et al., 1990; Sharf, 1978; Oxenham et al., 1996). Cette intégration temporelle apparait utile pour la perception de la texture des sons (McDermott & Simoncelli, 2011) et ainsi, pour l'identification des sources sonores composants notre environnement acoustique. Les phénomènes d'intégration temporelle suggèrent l'existence de fenêtres de traitement auditif relativement longues, de l'ordre de plusieurs dizaines voire centaines de millisecondes, contrastant avec les fenêtres plus courtes postulées pour les tâches de résolution temporelle auditive. Différents modèles ont été proposés pour décrire ce phénomène. L'intégration temporelle auditive étudiée dans le cadre de cette thèse se limite au cas de l'AM et de la FM.

6.1 Intégration temporelle de l'AM

6.1.1 Intégration en détection d'AM

Viemeister (1970) étudie l'effet de la durée de stimulation dans une tâche de détection de battements produits par la somme de sons purs. Pour une fréquence de battement de 4 Hz, les performances de détection augmentent lorsque la durée de stimulation augmente de 125 à 500 ms. Par ailleurs, lorsque le nombre total de battements est fixé (en réduisant la durée de stimulation et en augmentant l'écart en fréquence entre les composantes fréquentielles du son (Δf), ou vice versa), le seuil de détection reste constant. Viemeister (1970) conclue que les variations des seuils de détection observées en fonction de Δf sont déterminées principalement par le nombre de battements et non par une atténuation réalisée par le système auditif (une intégration auditive de l'information temporelle). Pour des basses fréquences de battement,

Viemeister montre que la détectabilité des battements augmente en fonction de la racine carrée du nombre de battements.

En 1979, Viemeister caractérise l'atténuation réalisée par le système auditif à l'aide de la MTF du sujet mesurée pour une tâche de détection d'AM. Viemeister suppose que la MTF reflète la réponse temporelle du système auditif à une AM en "régime permanent". Cette hypothèse impose que, pour des durées de stimulation suffisamment grandes comparativement à la constante de temps du système auditif, la MTF: i) doit être insensible à un changement de la durée de stimulation; ii) ne doit pas dépendre du mode de stimulation (porteuse continue ou limitée temporellement à la durée pendant laquelle la modulation est appliquée).

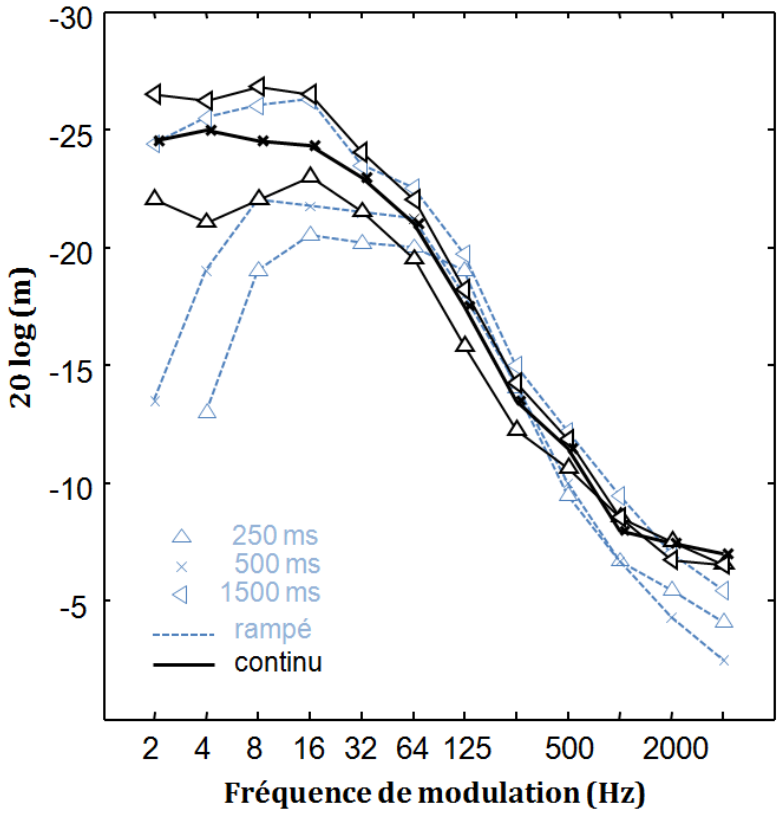


Figure 19. *Seuils AMDLs moyens obtenus chez 2 sujets normo-entendants pour différentes durées de stimulation. L'axe des ordonnées représente la profondeur de modulation nécéssaire pour obtenir 70,7% de bonnes réponses. Le niveau de puissance spectrale est fixé 40 dB SPL. Les lignes bleues représentent les AMDL dans les conditions où la porteuse est limitée par des rampes, tandis que les traits noirs représentent les AMDLs pour une porteuse présentée de façon continue. Retracé à partir des données de Viemeister, 1979.*

Afin de vérifier ces hypothèses, Viemeister (1979) mesure la MTF en fonction de la durée de stimulation (durée = [250-1500] ms) et en fonction du type de porteuse utilisée (bruit continu versus limité par des rampes; Fm = [2-4000] Hz; porteuse large bande de bruit; 2 sujets normo-entendants), comme le montre la figure 19.

Les résultats montrent que:

i) les AMDLs s'améliorent lorsque la durée de stimulation augmente, et ce principalement pour des cadences de modulation faibles et pour des durées relativement courtes. Pour des cadences de modulation supérieures à 250 Hz, l'amélioration des AMDLs en fonction de la durée de stimulation est négligeable pour des signaux continus.

ii) les AMDLs se dégradent à basse cadence de modulation et pour les durées de stimulation les plus faibles lorsque des rampes sont appliquées à la porteuse. La MTF présente alors une caractéristique passe-bande, en accord avec Forrest et Green (1987). Pour des porteuses continues, la MTF conserve une forme de type passe bas.

L'amélioration globale des performances observée lorsque la durée de stimulation augmente est compatible (au moins qualitativement) avec l'hypothèse d'un mécanisme d'observations multiples appelé "multiple-looks", hypothèse formulée en 1970. Pour une cadence de modulation fixée, lorsque la durée de stimulation croît, le sujet bénéficie de l'augmentation du nombre de cycles de modulation. Il a ainsi davantage d'opportunités d'échantillonner la modulation. L'augmentation du nombre d'échantillons engendre alors une amélioration de sa performance (d' augmente selon $\sqrt{n}$, pour n échantillons indépendants, si et seulement si les échantillons sont associés à une sensibilité d' constante).

Cette hypothèse "multiple-looks" est compatible avec le fait que:

i) l'intégration temporelle est plus importante à basse cadence de modulation;

ii) un nombre de cycle de modulation fixe (et donc, un nombre d'observation fixe) donne des AMDLs relativement constants lorsque la porteuse est limitée par des rampes à 250 et 500 ms.

En revanche, l'hypothèse "multiple-looks" ne permet pas d'expliquer les différences de seuils observées entre la MTF obtenue pour une porteuse continue et pour une porteuse dont la durée est limitée par des rampes.

Viemeister (1979) explique cette différence d'AMDLs entre porteuse continue et porteuse bornée comme résultant de l'adaptation nerveuse à court terme, un mécanisme neural transitoire apparaissant dès le niveau du nerf auditif. Il suggère que ce processus d'adaptation introduit au début du stimulus une forte variation dans sa représentation interne, comparativement aux variations induites par la modulation. Les fluctuations dans la représentation interne provoquées par l'adaptation interfèrent alors avec les fluctuations induites par la modulation. La dégradation des AMDLs résulterait donc d'un phénomène de masquage de modulation entre fluctuations d'amplitude dépendant du signal et celles causées par l'adaptation nerveuse.

Sheft et Yost (1990) réalisent par la suite une étude systématique de l'intégration temporelle pour la détection d'AM. Ces auteurs supposent que le processus d'intégration temporelle est affecté par la quantité d'information sensorielle disponible pour réaliser la tâche de détection. Toutefois, différentes études suggèrent que l'information sensorielle issue des indices de modulation dépend de l'écart temporel entre le début du signal porteur et le début de la modulation: i) en effet, la profondeur de modulation de la réponse nerveuse mesurée dans le nerf auditif diminue durant la phase d'adaptation à court terme (Smith et al., 1985); ii) Viemeister (1979) démontre qu'à faible cadence de modulation, les AMDLs sont plus élevés pour des stimuli présentant des rampes comparativement à des stimuli constants.

De façon à étudier conjointement: i) l'effet de la durée de modulation et; ii) l'influence de l'écart temporel entre le début du signal porteur et le début de la modulation, Sheft et Yost (1990) mesurent les AMDLs pour différentes durées de modulation et pour différentes conditions expérimentales (Fm = [2,5-640] Hz; porteuse: large bande de bruit; durée = [6,25-400] ms, 4 sujets normo-entendants). La porteuse est soit i) limitée temporellement par des rampes de

début/fin, l'AM étant appliquée sur l'ensemble de l'intervalle; ii) présentée de façon continue; iii) présentée durant une durée de 500 ms préalablement à l'application de l'AM.

La figure 20 représente les AMDLs obtenus par Sheft et Yost (1990) en fonction de la cadence de modulation.

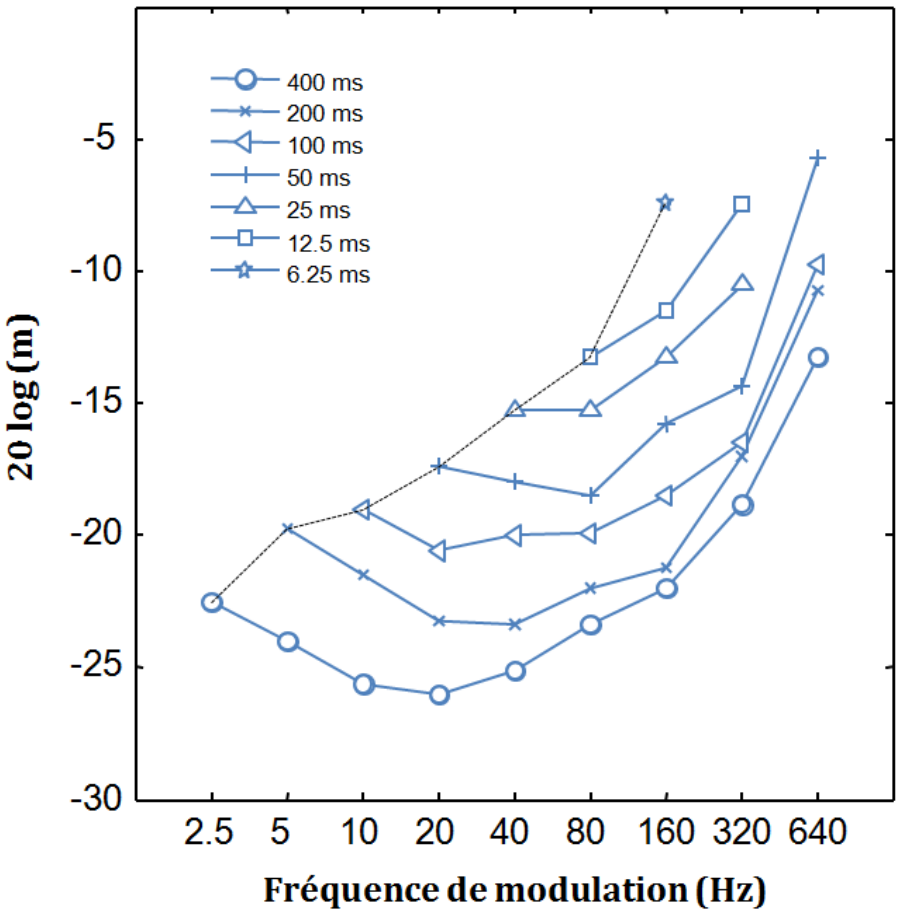


Figure 20. *Moyenne des AMDLs obtenus chez 2 sujets normo-entendants en fonction de la cadence de modulation pour un signal dont la porteuse est limitée par des rampes. Les AMDLs sont mesurés pour différentes durées de stimulation. La ligne en pointillée connecte les seuils obtenus pour un seul cycle de modulation. Retracé à partir des données de Sheft et Yost, 1990.*

Ces résultats obtenus pour des stimuli limités temporellement par des rampes de début/fin montrent que :

i) la MTF présente une caractéristique passe-bande pour des signaux dont la porteuse est limitée dans le temps par des rampes, en accord avec Viemeister (1979) et Forrest et Green (1987). La

plupart des études sur la détection d'AM fixe la durée de stimulation indépendamment de la cadence de modulation. Ainsi, le nombre de cycles présentés varie en fonction de la cadence de modulation. Sheft et Yost suggèrent que l'AMDL résulte tant des caractéristiques d'atténuation du système auditif dans le domaine temporel que de la capacité (centrale) des auditeurs à combiner l'information provenant des différentes observations. Lorsque le nombre de cycle de modulation est égalisé entre les différentes cadences de modulation (ligne en pointillés), les MTFs présentent une caractéristique passe-bas et non plus passe-bande. Les auteurs précisent que ces "MTFs" associent un niveau d'information sensoriel constant à un nombre de cycle constant (plutôt qu'à une durée de stimulation constante).

ii) une diminution de la durée où l'AM est présente s'accompagne d'une dégradation du seuil de détection. Viemeister (1979) observe des effets de la durée de stimulation pour des stimuli allant de 250 à 1500 ms. Il observe une intégration temporelle plus importante pour des cadences de modulation basses lorsque la porteuse est rampée. Sheft et Yost (1990) obtiennent des résultats similaires pour des durées de stimulation plus courtes (6.5-400 ms). Pour une cadence de modulation donnée, le taux auquel l'AMDL s'améliore lorsque la durée de stimulation augmente est une mesure de l'intégration temporelle pour la tâche en question. Pour les cadences de modulation supérieures à la partie passe-haut de la MTF, la pente de la fonction d'intégration est en moyenne de -7.46 dB par unité de durée sur une échelle log pour des stimuli limités temporellement par des rampes.

Les AMDLS obtenus pour des porteuses continues et pour des stimuli pour lesquels l'AM n'est appliquée que 500 ms après l'attaque de la porteuse montrent que:

i) Pour les signaux où la modulation démarre 500 ms après l'attaque du signal, les MTFs obtenues présentent à nouveau une caractéristique passe bande, avec toutefois une réduction de la pente de la fonction du coté passe-haut comparativement à la pente de la MTF mesurée avec des signaux rampés.

ii) Pour les signaux continus, les performances de détection de l'AM s'améliorent pour des cadences de modulation lentes. La MTF présente alors une caractéristique passe-bas, de façon analogue à Viemeister (1979).

Les différences de seuils obtenues entre les conditions où la porteuse est présentée de façon continue et celles où la porteuse démarre 500 ms avant l'application de l'AM suggèrent la contribution d'un facteur autre que l'adaptation nerveuse à court terme. Un processus d'adaptation plus lent pourrait être impliqué. En effet, comme l'ont précédemment relevé Moore (1976) puis Hartmann et Hnath (1982), les sujets rapportent une persistance de la perception de l'AM (une image consécutive) dans l'intervalle de comparaison de la tâche en choix forcé lorsque le stimulus est limité temporellement , et ce pour des cadences de modulation faibles. L'adjonction de 500 ms de signal stationnaire avant l'AM ne permet pas de s'affranchir totalement de cette AM rémanente, alors qu'elle disparait dans le cas où la porteuse est présentée de façon continue. L'origine de cette rémanence ou image consécutive reste encore relativement inexpliquée. Si l'AM est traitée par des canaux sélectifs aux modulations comme le suggèrent certains travaux neurophysiologiques (e.g. Moller, 1976; Rees & Palmer, 1989; Schreiner & Urbas, 1988) ou psychophysiques (e.g., Kay & Matthews, 1972; Tansley & Suffield, 1983 ; Bacon & Grantham, 1989 ; Houtgast, 1989 ; Dau et al., 1997), une image consécutive de l'AM pourrait être induite tant par l'activité spontanée des canaux de modulation que par l'adaptation de chaque canal.

Pour finir, l'intégration temporelle s'accroit de 2 dB lorsque la porteuse est présentée de façon continue ou démarre antérieurement à l'AM, pour s'établir à -9.3 dB par unité de durée sur une échelle log. Les auteurs proposent que cette plus forte intégration temporelle pourrait être due à une sur-intégration. Pour des durées de stimulation relativement faibles, une portion du signal non modulé pourrait être incluse dans le calcul de l'énergie du signal de modulation.

En conclusion, les résultats de Viemeister (1979) et Sheft et Yost (1990) montrent que les AMDLs s'améliorent lorsque la durée de stimulation augmente (6.5 à 1500 ms). Ce phénomène d'intégration temporelle est plus important à basse cadence de modulation lorsque la porteuse est limitée temporellement. L'intégration temporelle pourrait résulter d'une sommation d'information, à savoir la sommation des différentes opportunités de détection de l'AM (les variations au sein de chaque cycle de modulation étant considérées comme des opportunités de détecter la modulation).

Ces résultats indiquent également que la MTF mesurée pour des stimuli limités dans le temps présente une caractéristique passe bande et non passe bas. La dégradation des AMDLs à basse cadence de modulation pour des signaux limités temporellement pourrait s'expliquer par la contribution de l'adaptation nerveuse à court terme, un processus neural transitoire dont la constante de temps va de quelques millisecondes à quelques dizaines de millisecondes.

L'ajout de 500 ms de porteuse avant le démarrage de la modulation ne permet pas d'éliminer la portion passe-haut de la MTF. Ceci suggère que le traitement temporel auditif de l'AM n'atteint toujours pas son "régiment permanent", même 500 ms après le démarrage de la porteuse.

6.1.2 Intégration en discrimination d'AM

La seule mesure d'intégration temporelle pour une tâche de discrimination de profondeur d'AM est réalisée par Lee et Bacon (1997) (Fm=[10;20;40;80] Hz; Fc= 4000 Hz; $\Delta m_{standard}$= [0.1;0.18;0.3] %; niveau de présentation fixe de 60 dB SPL) chez 3 sujets normo-entendants (23-30 ans). La durée de présentation du signal est fixée de façon à égaliser le nombre de cycles de modulation entre les différentes conditions expérimentales (25, 50, 100, 200, 400, et 800 ms).

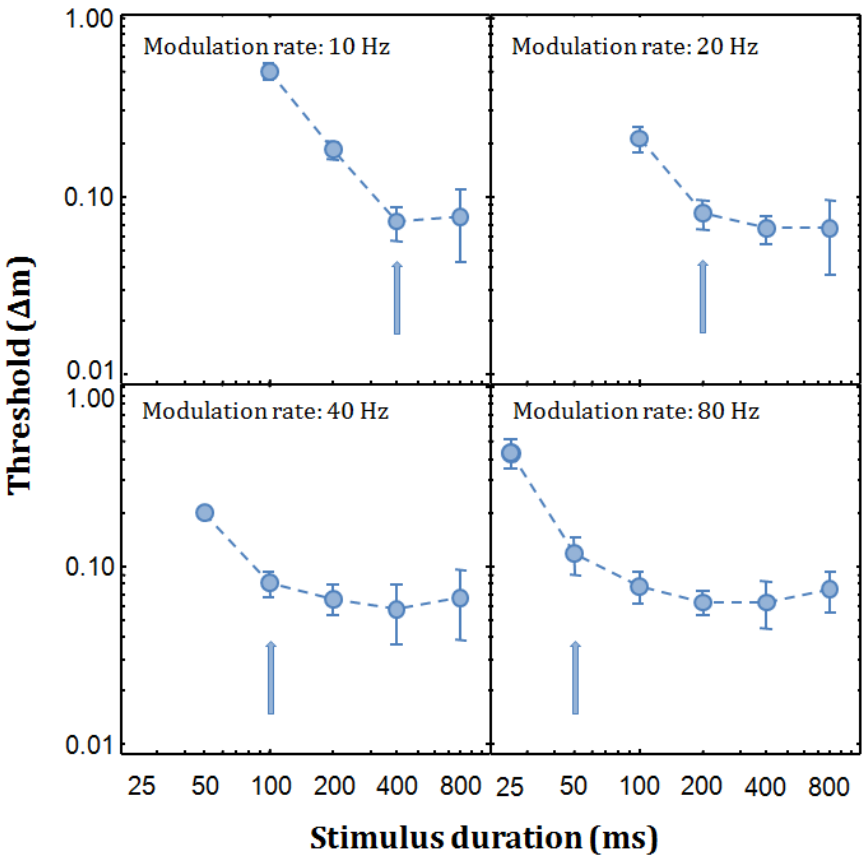


Figure 21 *Seuils de discrimination (Δm) d'AM en fonction de la durée de stimulation. Chaque figure représente les données pour une cadence de modulation (10, 20, 40 et 80 Hz). L'effet de la profondeur de modulation standard étant non significatif, les seuils de discrimination d'AM représentés ici correspondent la moyenne des seuils obtenus pour les différentes cadences de modulation standard. Les flèches indiquent les durées correspondant à quatre cycles de modulation. Retracé à partir de Lee et Bacon, 1997.*

Comme le montre la figure 21, pour toutes les cadences de modulation, les seuils de discrimination diminuent significativement lorsque la durée de stimulation augmente, et ce indépendamment de la profondeur de modulation standard utilisée. Les seuils de discrimination d'AM diminuent a minima d'un facteur deux lorsque la durée de stimulation est doublée pour les durées de stimulation les plus courtes. Toutefois, au delà d'une certaine durée de stimulation (environ 4 cycles de modulation), les seuils de discrimination restent stables ou ne diminuent que faiblement.

Ces données d'intégration temporelle en discrimination de profondeur d'AM sont compatibles avec les données d'intégration temporelle obtenues en discrimination de la cadence d'AM. En

effet, Lee (1994) montre que pour des cadences de modulation de 20 et 40 Hz, les seuils de discrimination diminuent rapidement (décroissance d'un facteur 2) jusqu'à une certaine durée de stimulation (environ 5 cycles de modulation), à partir de laquelle les seuils de discrimination ne diminuent plus que légèrement ou restent stables.

Toutefois, les données d'intégration temporelle de Lee et Bacon (1997) et Lee (1994) en discrimination de profondeur et de cadence d'AM ne sont pas compatibles avec celles observée en détection d'AM (Viemeister, 1979; Sheft et Yost, 1990), pour laquelle les seuils de détection décroissent progressivement lorsque la durée de stimulation augmente, sans réel point d'inflexion notable aux alentours de 4-5 cycles de modulation. Les données de détection de battements de Vieimeister (1970) ne montrent également aucun point d'inflexion lorsque la durée de stimulation augmente.

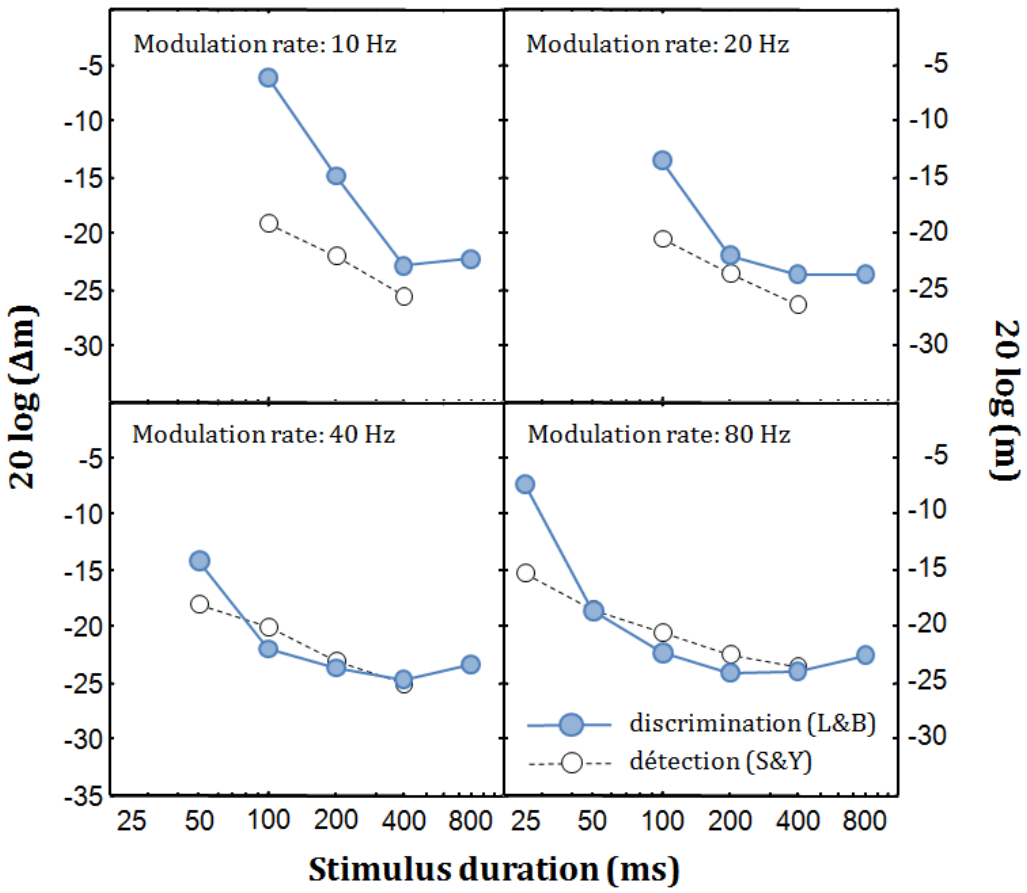


Figure 22. *Comparaison de l'effet de la durée de stimulation sur les seuils de détection d'AM (données issues de Sheft et Yost, 1990 : S&Y) et sur les seuils de discrimiantion de profondeur d'AM (Lee et Bacon, 1997 : L&B)*

Lee et Bacon (1997) comparent l'intégration temporelle observée en détection d'AM et en discrimination de profondeur d'AM. Comme le montre la figure 22, pour toutes les cadences de modulation, les seuils de détection et de discrimination d'AM évoluent de façon globalement similaire lorsque la durée de stimulation est supérieure à 4 cycles de modulation. Les seuils de détection et de discrimination divergent toutefois pour des durées de stimulation plus courtes.

Lee et Bacon (1997) interprètent cette différence entre les seuils de détection et de discrimination d'AM pour des durées de stimulation courtes comme reflétant la difficulté intrinsèque de chaque tâche. Pour une tâche de détection d'AM, le sujet doit simplement identifier lequel des deux stimulus présente une AM. Pour une tâche de discrimination d'AM, le sujet doit déterminer lequel des deux stimulus présente l'AM la plus saillante. Ceci pourrait nécessiter une meilleure représentation interne des stimuli, la représentation interne pouvant être sensiblement affectée pour des durées de stimulation courtes lorsque le nombre de cycle est réduit.

Les auteurs évaluent enfin dans quelle mesure le modèle "multiple looks" (Viemeister & Wakefield, 1991) permet de rendre compte des données d'intégration temporelle en discrimination de profondeur d'AM. En théorie, les seuils de discrimination de profondeur d'AM devraient diminuer proportionnellement à la racine carrée du nombre d'observations disponibles (Green et Swets, 1988) si toutefois les observations: i) sont indépendantes; ii) ont une sensibilité d' identique et; iii) sont combinées de façon optimale. Lee et Bacon (1997) définissent une observation comme étant un cycle de modulation "complet", c'est à dire un cycle de modulation n'ayant pas été affecté par les rampes de début et de fin. Le nombre d'observation est donc égal à n-1, n étant le nombre de cycles de modulation. Les auteurs justifient cette définition par les résultats d'une étude pilote, démontrant que le seuil de discrimination d'un sujet est globalement similaire qu'on lui présente 1.25, 1.5 ou 2 cycles de modulation pour une cadence de modulation de 20 Hz.

Les fonctions psychométriques sont mesurées chez deux des sujets précédents pour différentes durées de stimulation, de façon à comparer les niveaux de sensibilité d' obtenus expérimentalement et ceux prédits par le modèle "multiple looks". Les résultats sont représentés dans la figure 23.

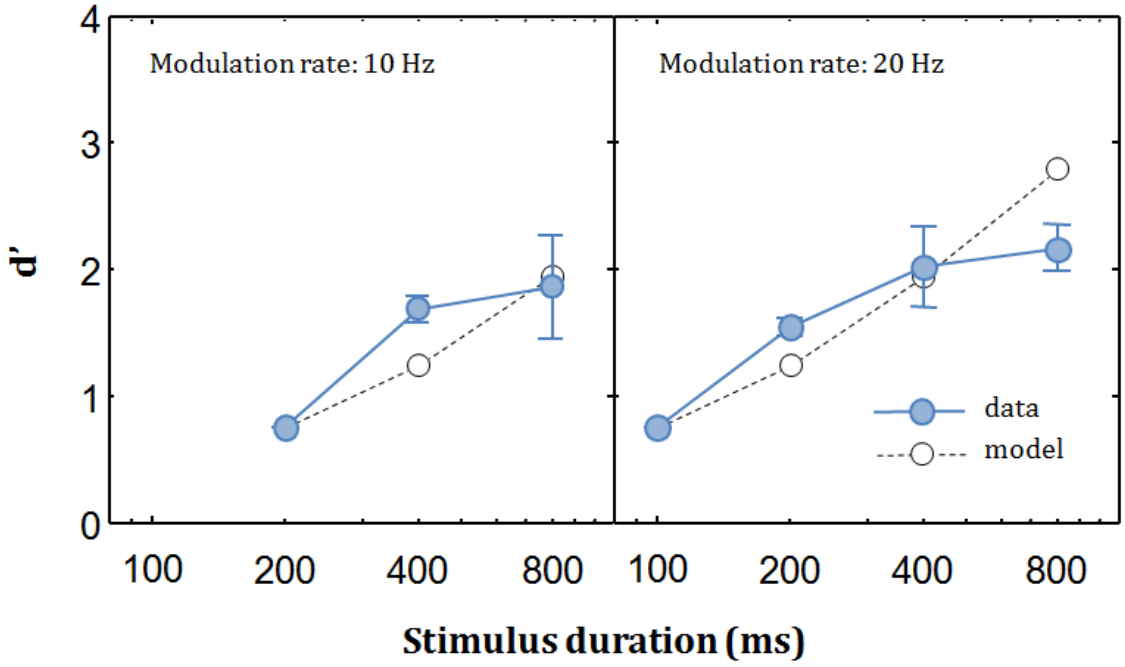


Figure 23. *Sensibilité d' obtenue expérimentalement (moyenne des 2 sujets) et prédite par le modèle "mutiple looks" en fonction de la durée de stimulation pour une cadence de modulation de 10 Hz (figure de gauche) et une cadence de modulation de 20Hz (figure de droite) - Retracé à partir de Lee et Bacon, 1997.*

Les auteurs concluent que le modèle "multiple looks" fournit une estimation raisonnable des seuils de discrimination d'AM en fonction de la durée de stimulation.

Toutefois, lorsque la durée de stimulation augmente (d'une à trois observations), le modèle estime mal l'intégration temporelle. La faible valeur de d' obtenue pour la première observation peut être expliquée par le processus d'adaptation ou l'incertitude liée à l'apparition du signal (Viemeister, 1979; Viemeister & Wakefield, 1991). L'adaptation à court terme réduit la sensibilité à la première observation comparativement aux observations suivantes, probablement parce que la réponse neurale initiale produite par l'apparition du stimulus est

relativement importante comparativement aux changements de la réponse neurale produits par l'AM.

Pour des durées de stimulation plus importantes, le modèle surestime les valeurs des indices de sensibilité. Au delà de 4 cycles de modulation, les seuils de discrimination comportementaux ne s'améliorent que faiblement, alors que les indices d' prédits par le modèle continuent d'augmenter. Cette dégradation de la performance des sujets pourrait être expliquée par les limitations mnésiques, la détectabilité des premières observations pouvant être réduite pour des stimulus longs (Viemeister & Wakefield, 1991). En résumé, ces résultats indiquent que la détectabilité de chacune des observations n'est pas identique et que des poids perceptifs différents doivent être appliqués à chaque observation pour permettre de rendre compte des données expérimentales avec le modèle "multiple- looks".

6.1.3 Modélisation de l'intégration temporelle en AM

L'intégration temporelle doit être prise en compte lorsque l'on modélise la détection et la discrimination d'un son et plus spécifiquement d'une AM. Le terme intégration temporelle semble peu judicieux, car il sous-entend l'existence d'un processus de sommation (au sens strict) au sein du système auditif. Par exemple, le fait que le seuil de détection d'un son s'améliore de 3 dB lorsque la durée de stimulation est doublée suggère que le système auditif se comporte comme s'il effectuait une intégration de puissance, mais ce n'est pas nécessairement le cas (Viemeister & Wakefield, 1991) comme le démontrent certains travaux de modélisation.

Historiquement, deux approches ont été proposées pour rendre compte du phénomène d'intégration temporelle:

i) la première suppose que le processus d'intégration temporelle est réalisé sur une échelle temporelle relativement longue, de l'ordre de quelques centaines de millisecondes (e.g., Green, 1960; Green & Swets, 1966; Munson, 1947; Plomp & Bouman, 1959; Zwislocki, 1960, 1969;

Jeffress, 1967, 1968). Par exemple, Green et Swets (1966) proposent un modèle se basant sur un échantillonnage du signal à court terme pour estimer différentes observations. Les informations provenant de chaque observation sont ici combinées. Dans ces modèles, la puissance du signal est intégrée (totalement ou partiellement) au sein d'un intervalle de temps défini.

ii) la seconde suppose que le processus d'intégration temporelle est réalisé sur une échelle temporelle plus courte. Ces modèles ont été développés pour rendre compte des différents phénomènes tels que la détection de modulation (Viemeister, 1979) ou d'intervalle de silence (Forrest & Green, 1987), ainsi que certains aspects liés au masquage non simultané (Moore et al., 1988).

Cependant, les principales différences entre ces modèles reposent sur le conflit entre constante de temps d'intégration et de résolution (de Boer, 1985). Pour la détection de modulation, Sheft et Yost (1990) obtiennent des constantes de temps en intégration temporelle largement supérieures à celles qui sont observées pour la résolution temporelle (celles-ci sont de l'ordre de 1-3ms). Le paradoxe entre constante de temps courtes et constantes de temps longues n'est pour autant pas insoluble:

i) la première hypothèse, peu parcimonieuse, consiste à considérer l'existence de deux systèmes distincts ayant chacun une constante de temps différente. Le système utilisé dépendrait de la tâche à réaliser. Ainsi, Viemeister (1979) suggère que la disparité entre les constante de temps d'intégration et de résolution temporelle n'est pas problématique. Les tâches demandées au système auditif sont très différentes et Viemeister propose que des stratégies différentes peuvent être utilisées. Une tâche de résolution temporelle semble faire appel à des processus principalement périphériques. A l'inverse, l'intégration temporelle pourrait faire appel à des processus de plus haut niveau de façon à réaliser un échantillonnage multiple et une sommation de probabilité.

ii) la seconde hypothèse suggère l'existence d'un système unique, mais dont la constante de temps pourrait être ajustée en fonction de la durée du signal.

Penner (1978) propose un modèle permettant de concilier constante de temps longue et constante de temps courte. Elle démontre qu'une compression d'amplitude suivie d'une constante de temps courte entraine une intégration "efficace" relativement longue. Ce modèle d'intégration se base sur une "observation unique''.

D'une façon générale, les modèles d'intégration temporelle susnommés ne rendent compte que très grossièrement des phénomènes d'intégration temporelle (Viemeister & Wakefield, 1991).

Ainsi, Viemeister (1979) propose le premier modèle tentant de rendre compte des données de détection d'AM et de l'intégration temporelle en AM (cf. section 3.1.3). Quoique de manière imparfaite, le modèle du détecteur linéaire d'enveloppe prédit l'amélioration des performances lorsque la durée de la modulation augmente. Par ailleurs, pour des signaux dont la porteuse est limitée par des rampes, le modèle prédit la partie passe haut de la MTF, mais ne rend pas compte de la pente et de la magnitude de celle-ci.

Sheft et Yost (1990) évaluent différents modèles computationnels (sommation des observations, détecteur linéaire ou non linéaire d'enveloppe associé à différentes stratégies de décision) de façon à rendre compte de leurs données d'intégration temporelle en AM.

Hypothèse d'une sommation des observations: Lorsque les informations de n observations indépendantes ayant des indices de sensibilité d' identiques sont combinées de façon optimale, d' augmente proportionnellement à la racine carrée de n (Green & Swets, 1974).

La variation du seuil de détection d'une modulation (d' constant) en fonction du nombre d'observation dépend de la relation entre d' et la puissance de la modulation. Avec une relation exponentielle entre d' et la puissance de la modulation, la détectabilité augmente en fonction de la racine carrée de n, ce qui permet de prédire qu'au seuil:

$$20 \log m = C - (5/k) \log (n)$$

où C est la profondeur de modulation au seuil lorsqu'un seul cycle de modulation est présenté.

Les fonctions psychométriques pour détecter une AM à basse cadence de modulation montrent que d' est schématiquement proportionnel à la puissance de modulation (Sheft & Yost, 1989, 1990) : la constante k est environ égale à 1. En remplaçant k par 1 dans l'équation précédente, cela suggère que le seuil de détection d'une AM devrait s'améliorer de 5 dB par unité de durée sur une échelle logarithmique, sous-estimant ainsi l'intégration temporelle obtenue dans les données expérimentales. Le principe de sommation des observations est donc rejeté.

La pente d'intégration temporelle de -5 dB par unité de durée sur une échelle logarithmique se base sur l'hypothèse que le d' associé à chaque observation est constant. Toutefois, les données de Viemeister (1979) suggèrent que la détection de la modulation au démarrage du signal est limitée par l'adaptation à court terme. Cette limitation induite par l'adaptation à court terme renforce le poids des observations plus tardives lorsque les différents échantillons sont recombinés. Sheft et Yost (1990) modélisent cette inefficacité lors du démarrage de la stimulation par une pondération exponentielle des valeurs de d' associées à chacune des observations.

Modèles de détection linéaire/non linéaire d'enveloppe: En sortie du détecteur d'enveloppe, différentes statistiques de décision ont été évaluées par Sheft et Yost: RMS d'enveloppe (Viemeister, 1979) ou ratio max/min (Forrest & Green, 1987).

Les résultats obtenus pour les différentes simulations sont représentés en figure 24 et montrent que:

i) les MTFS obtenues dans la condition où la porteuse démarre 500 ms préalablement à l'apparition de l'AM approchent les MTFs obtenues expérimentalement dans la condition où le signal est limité par des rampes;

ii) à faible niveau de stimulation, les MTFs présentent une caractéristique passe-bas. En utilisant le modèle de synapse de Meddis (1986,1988) simulant l'adaptation nerveuse à court terme dans le nerf auditif, les AMDLs sont légèrement plus élevés à basse cadence de modulation que les AMDLs obtenus en utilisant une simple rectification mono alternance. A plus forte intensité, la

sensibilité à l'AM décroit pour des cadences de modulation inférieures à 40 Hz. Cette dépendance au niveau de stimulation est caractéristique de la réponse neurale à l'AM et reflète un phénomène de saturation des réponses des fibres du nerf auditif (Smith & Brachman, 1980; Yates, 1987; Rees & Palmer, 1989). Elle contraste avec le faible effet du niveau de stimulation observé pour des mesures psychophysiques (Viemeister, 1979; Bacon & Viemeister, 1985).

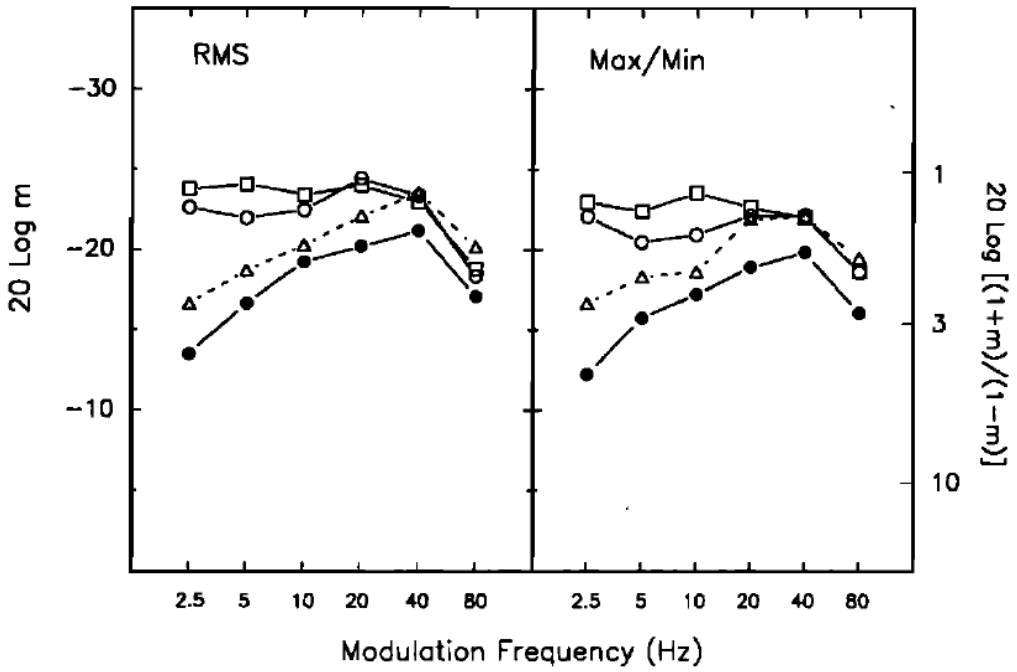


Figure 24. *MTFs simulées pour un signal modulé pendant 400ms. Deux statistiques de décision sont utilisées: puissance RMS (figure de gauche) et le ratio max/min (figure de droite). Les carrés ouverts représentent les seuils pour la condition où la porteuse est présentée 500 ms préalablement au démarrage de la modulation, avec rectification passe bas. Les autres conditions utilisent l'algorithme de Meddis comme opérateur non linéaire de façon à simuler la transduction neurale. Avec cet algorithme, la condition dans laquelle la porteuse est présentée 500 ms avant le début de la modulation a été simulée à deux niveaux d'intensité : à un niveau pour lequel les fibres saturent (triangle ouverts) et à un niveau plus faible (cercles ouverts). Les cercles pleins symbolisent les seuils simulés pour la condition où la porteuse est limitée par des rampes en deça du niveau de saturation. Reproduit de Sheft et Yost (1990).*

iii) les simulations obtenues avec l'algorithme de Meddis pour un signal ayant été limité temporellement par des rampes montrent que la MTF obtenue présente une caractéristique passe-bande, en accord avec les données expérimentales. Ces résultats confirment la prédiction de Viemeister (1979) imputant le segment passe haut de la MFT à l'adaptation à court terme.

iv) l'amélioration de la qualité des simulations lorsque la durée de la modulation augmente provient principalement : a/ d'une réduction de la variance des estimations de la puissance RMS; b/ d'une augmentation de l'écart entre les distributions moyennes obtenues en présence ou en l'absence de signal pour le ratio max/min, ainsi que d'une réduction de la variance. La réduction de la variance pour le ratio max/min est toutefois significativement plus faible que celle observée pour la puissance RMS.

v) avec le ratio max/min, l'intégration temporelle est inférieure à celle observée expérimentalement. En règle générale, la puissance RMS fournit de meilleures estimations des AMDLs que ceux obtenus avec le ratio max/min.

Sur les trois approches computationnelles proposées, le meilleur modèle pour des stimuli rampés est celui associant détecteur d'enveloppe et calcul de la puissance RMS. Les auteurs proposent de modifier le modèle "multiple looks" pour améliorer son caractère prédictif. Au lieu de considérer chaque cycle de modulation comme une opportunité de détection, Sheft et Yost (1990) suggèrent de définir les observations à partir d'une fenêtre temporelle relativement courte. Viemeister et Wakefield (1989) utilisent cette approche sur des données neurales pour rendre compte de l'intégration temporelle dans une tâche de détection d'un signal.

Viemeister et Wakefield (1991) étudient dans quelle mesure le processus d'intégration temporelle résulterait d'une constante de temps courte ou d'une constante de temps longue. Pour ce faire, ils réalisent deux expérimentations différentes:

i) la première expérience mesure le seuil de détection d'une paire de brèves impulsions en fonction de l'écart temporel séparant les deux impulsions dans le silence. Leurs résultats suggèrent qu'une intégration de puissance n'est réalisée que pour des écarts temporels inférieurs à 5 ms. Pour des écarts temporels plus conséquents, les seuils de détection restent constants et les impulsions semblent être traitées séparément.

ii) dans la seconde expérience, de brèves impulsions espacées de 100 ms sont présentées dans des échancrures d'un bruit large bande. Les seuils de détection obtenus lorsque deux impulsions

sont présentées simultanément sont meilleurs que les seuils de détection obtenus lorsque les impulsions sont présentées séparément. Ce résultat suggère l'existence d'un processus d'intégration temporelle. Toutefois, les seuils de détection d'une paire d'impulsions ne sont pas affectés par une modification de l'intensité du bruit présent entre les deux impulsions. Cette insensibilité à l'intensité du bruit masquant présenté entre les deux impulsions est incompatible avec l'hypothèse d'une constante de temps longue et invalide de nombreux modèles. Par exemple, le modèle de Penner (1978) prédit une variation du seuil de détection lorsque l'intensité du bruit présenté entre les deux impulsions est modifiée. En effet, pour rendre compte du fait que les seuils de détection d'une paire d'impulsions sont plus faibles que lorsqu'une seule impulsion est présentée, il est nécessaire d'intégrer le son sur une durée au moins égale à l'écart temporel entre les deux impulsions.

Viemeister et Wakefield (1991) développent un modèle alternatif basé sur une hypothèse statistique proposée initialement dans le domaine de la vision (e.g., Pirenne, 1943; Lamar et al., 1947). Ils suggèrent d'échantillonner le signal à une cadence relativement rapide (constante de temps courte). Ces échantillons, ou observations, sont ensuite stockés en mémoire, et peuvent être traités séparément. La figure 25 représente une schématisation du modèle "multiple look" proposé par Viemeister & Wakefield (1991).

Le processus d'intégration temporelle résulterait d'une augmentation du nombre d'observations disponibles. A chaque observation est associée une probabilité que l'échantillon du signal dépasse une valeur seuil. Si la durée de stimulation augmente, le nombre d'observations disponibles augmente. La probabilité que le seuil soit dépassé pour au moins l'une des observations augmente, expliquant alors le processus d'intégration temporelle.

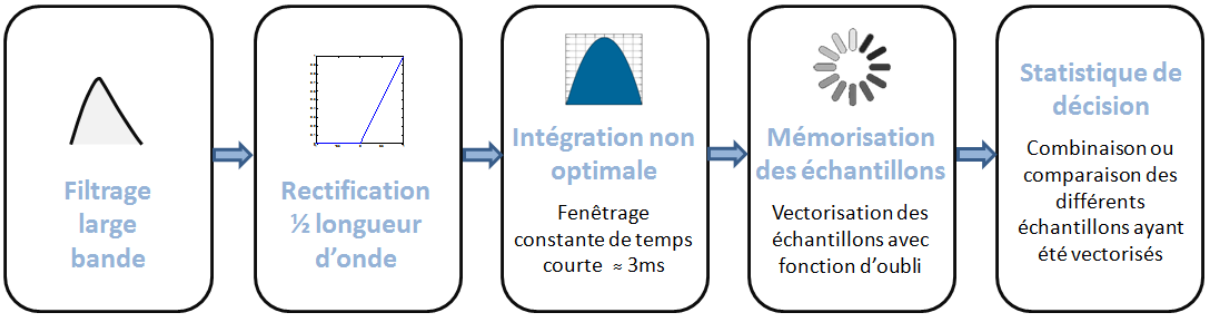


Figure 25. *Schématisation du modèle "multiple-look" de Viemeister & Wakefield (1991). Le signal est filtré sur une bande de fréquence, appelée bande critique, avant d'être rectifié. Il est ensuite intégré de façon non optimale (fenêtrage du signal avec une constante de temps courte , ie. environ 3 ms). Chaque observation est ensuite sauvegardée dans une mémoire à court terme (ayant une caractéristique de décroissance propre et une constante de temps plus longue, de l'ordre de quelques centaines de ms). Cette représentation en mémoire du signal peut être considérée comme un vecteur d'échantillons ou d'observations du signal d'entrée. Les observations sont utilisables de différentes façons: i) pour une tâche de sommation temporelle, le sujet peut utiliser l'intégralité des observations disponibles; ii) pour une tâche de détection d'une modulation, le sujet peut scruter le vecteur d'observation pour trouver la présence de fluctuations.*

Cette extension du modèle "multiple-looks" permet de rendre compte des données d'intégration temporelle de Viemeister et Wakefield (1991). Il permet également de modéliser les fonctions d'intégration temporelle observées pour des sons purs et pour du bruit. Le modèle fournit un cadre permettant de décrire les phénomènes liés à la résolution temporelle (tel que la détection d'intervalle de silence) sans pour autant avoir besoin de faire appel à une constante de temps variable.

Toutefois, le modèle "multiple looks" présente quelques limitations et imprécisions dans sa formulation:

La première limitation du modèle repose sur la fenêtre temporelle utilisée pour estimer les différentes observations successives. Dans leur modèle, les auteurs adoptent des fenêtres temporelles rectangulaires (3 ms), contigües et sans recouvrement. Un tel fenêtrage semble peu réaliste.

La seconde limitation repose sur la définition même de ce qui constitue une "observation". Une observation est décrite comme étant un "échantillon". Dans une approche aussi générale, un échantillon pourrait donc tout aussi bien être une quantité proportionnelle à l'indice de sensibilité d', un rapport de vraisemblance ou un échantillon de l'enveloppe issue de la forme d'onde du signal pour la fenêtre temporelle considérée.

La troisième limitation repose sur l'existence d'une probable corrélation entre les observations successives, ainsi que la façon dont les observations sont combinées avant l'étape de décision. Le modèle d'intégration temporelle proposé par Viemeister et Wakefield (1991) semble dépendre fortement de la façon dont les différentes observations sont combinées et pondérées.

La dernière limitation repose sur le fait que les caractéristiques mnésiques doivent être précisées, de façon à déterminer dans quelle mesure une observation décroît avec le temps et comment ce processus doit être modélisé.

Dau et al. (1997) proposent une extension de leur modèle PEMO (cf. section 3.1.3) de façon à rendre compte, entre autres, des effets d'intégration temporelle pour la détection d'une AM appliquée à un bruit. Leur modèle se base sur une version multi-bandes du banc de filtre de modulation, au sein duquel l'analyse des modulations est réalisée en parallèle en sortie de chacun des filtres auditifs simulés. Cette analyse multi-bande se justifie par le fait que l'étude de la représentation de l'AM au niveau du colliculus inférieur chez le chat (Schreiner et Langner, 1988) montre que les fréquences de modulation sont représentées orthogonalement à l'organisation tonotopique du colliculus inférieur. L'ensemble des autres étapes du modèle PEMO ont été décrites dans la section 3.1.3. Dans le modèle multi-bandes, la représentation interne du stimuli présente ici quatre dimensions: l'amplitude, le temps, la fréquence centrale de modulation et la fréquence audio. De façon à intégrer l'information au cours du temps, le modèle utilise en quelque sorte l'hypothèse "multiple looks", implémenté au niveau de la statistique de décision du modèle.

Le modèle PEMO utilise un détecteur optimal comme statistique de décision. La statistique de décision peut être vue comme un processus d'appariement de gabarit. Elle se base sur le calcul d'une inter-corrélation entre la représentation interne du stimulus (en 4 dimensions) et la représentation interne du gabarit (en 4 dimensions), extrait à un niveau supraliminaire.

Cela implique que la constante de temps est ajustée à la durée du signal de façon continue. En d'autres termes, le modèle dispose d'un continuum de constante de temps. Ainsi, aucune constante de temps fixe n'est nécessaire pour rendre compte de l'intégration temporelle à long terme. La pondération du stimuli au cours du temps est définie automatiquement par l'étendue temporelle du gabarit.

Le modèle de Dau et al. (1997) rend compte des données de Sheft et Yost (1990) avec une grande précision. Grâce à l'incorporation d'un mécanisme d'adaptation (Püschel, 1988; Dau et al., 1996a), le modèle rend compte de la dégradation des AMDLs à basse cadence de modulation pour des durées de stimulation courtes. L'adaptation simulée produit une forte augmentation du niveau de sortie du modèle lors du démarrage de la porteuse, et affecte la détection de la modulation du signal pour des cadences de modulation très faibles.

Le modèle PEMO permet également de rendre compte des longues constantes de temps observées en intégration temporelle (Sheft & Yost, 1990; Viemeister, 1979 ; Viemeister et Wakefield, 1991).

La structure du modèle PEMO fixe la limite de l'acuité temporelle. Cette constante de temps courte est dépendante du traitement de signal réalisé par le modèle mais totalement indépendante des paramètres expérimentaux. Le banc de filtre de modulation possède des constantes de temps très courtes, qui ne permettent toutefois pas de rendre compte de l'intégration temporelle à plus long terme.

La longue constante de temps prédite en intégration temporelle est réalisée au niveau de l'étape décisionnelle du modèle et reflète principalement la dépendance de l'inter-corrélation avec le nombre d'échantillons des vecteurs comparés. Ainsi, l'étape de décision est responsable de la

sensibilité globale à l'AM (le décalage vertical de la MTF) lorsque la durée de stimulation augmente (Viemeister, 1979). A l'inverse, les caractéristiques du banc de filtres de modulation déterminent la fréquence de coupure de la MTF, à partir de laquelle est déterminée la constante de temps pour la résolution temporelle.

Le modèle PEMO semble compatible avec la vision de Viemeister et Wakefield (1991), qui supposent que les tâches d'intégration temporelle et de résolution temporelle font appel à des stratégies différentes. Une tâche de résolution temporelle est probablement contrainte par des mécanismes relativement périphériques, tandis que l'intégration temporelle fait appel à des processus de plus haut niveau comme la sommation de probabilité ou l'échantillonnage multiple.

En conclusion, les travaux de Viemeister (1979) et Sheft et Yost (1990) montrent que les différents modèles proposés jusqu'alors, s'ils permettent d'expliquer une grande part des données d'intégration temporelle en AM, ne permettent pas d'en rendre compte parfaitement. Le modèle sous estime l'intégration temporelle en détection d'AM et surestime l'intégration temporelle en discrimination de profondeur d'AM. Par ailleurs, pour la détection d'AM, l'intégration temporelle peut être modélisée par une sommation de l'énergie, mais il faut faire l'hypothèse supplémentaire d'une sur-intégration pour des stimuli brefs.

Viemeister et Wakefield (1991) étendent le modèle "multiple looks" en proposant d'échantillonner le signal à une cadence relativement rapide. Ces échantillons indépendants, sont mémorisés, pondérés puis combinés. Le modèle "multiple looks" permet de rendre compte des données d'intégration temporelle. Toutefois, la formulation de ce modèle est encore imprécise et soulève de nombreuses interrogations: Qu'est-ce qu'une observation? Quel type de fenêtre utiliser? Comment combiner les différentes observations?

Dau et al. (1997) proposent probablement le modèle le plus abouti pour rendre compte de la détection d'AM et de son intégration temporelle. De par son processus d'appariement de gabarit reposant sur un calcul de corrélation, ce modèle rend compte tant des longues constantes de temps observées en intégration temporelle que des constantes de temps beaucoup plus courtes mesurées dans les tâches de résolution temporelle.

Ainsi, l'intégration temporelle reflète à la fois l'opération de mécanismes périphériques (adaptation nerveuse à court terme) aux cadences d'AM les plus lentes et l'opération de processus centraux, impliqués dans le stockage en mémoire et la prise de décision.

6.2 Intégration temporelle de la FM

Hartmann et Klein (1980) réalisent l'unique étude ayant à la fois évalué empiriquement et modélisé le processus d'intégration temporelle en détection de FM. Il est toutefois important de noter qu'antérieurement à leurs travaux, différents modèles ont été proposés de façon à rendre compte de la détection d'une FM (cf. section 3.2.2).

Par exemple, en 1937, Koch présentent une théorie de la détection d'une FM basée sur la détection des bandes spectrales latérales. Koch démontre que l'enveloppe des composantes spectrales devient bimodale lorsque l'index de modulation β ($\beta=\Delta f/F_m$) est supérieur à 1.5 ou 2. Koch suppose que la détection de la FM est possible lorsque Δf est suffisamment large pour permettre l'apparition de cette structure bimodale. Il propose donc que la détection d'une FM est contrainte par cette limite théorique (il serait donc impossible de détecter une FM en deçà de cette limite, i.e. β=1.5 - 2.0). Toutefois, cette prédiction se heurte aux données expérimentales, indiquant que l'index de modulation dans des tâches de détection de FM est généralement compris entre 0.14 et 0.4 (Kay & Matthews, 1972), soit des performances de détection 10 fois supérieures aux prédictions de Koch (1937).

Hartmann et Klein (1980) supposent que la détection d'une FM repose sur deux régimes distincts: un régime pour les cadences de modulation élevées, pour lequel la cadence de modulation dépasse la largeur du filtre auditif (c) et un régime à basse cadence de modulation (F_m < 6 Hz). Pour des cadences de modulation élevées, la représentation spectrale du signal permet de caractériser correctement le signal (Hartmann, 1978, 1979). Pour des cadences de modulation basses, la détection fait appel à des mécanismes purement temporels. Dans ce régime, le sujet peut détecter les variations temporelles engendrées par la modulation. Le système auditif intègrerait le signal sur un intervalle de temps suffisamment long pour pouvoir déterminer la hauteur du signal, mais suffisamment court pour détecter les variations de hauteur instantanée produite par la FM.

Hartmann et Klein (1980) proposent un modèle basé sur la théorie de la détection du signal, permettant de rendre compte de la détection d'une FM à basse cadence de modulation. Ce modèle permet de: i) rendre compte des différences observées entre seuils différentiels (DLF) mesurés par une tâche de discrimination de fréquence et seuils de détection d'une FM; ii) prédire la détectabilité d'une FM pour différents types d'onde de modulation, en fonction de la durée de stimulation.

A basse cadence de modulation, un signal de FM engendre la perception d'une hauteur qui varie au cours du temps. Les auteurs suggèrent que les sujets échantillonnent le signal pour réaliser différentes estimations de la hauteur du signal, et les comparent de façon à détecter une variation de hauteur.

Le modèle de Hartmann et Klein représente le signal de FM d'un point de vue interne sur une échelle tonotopique (donc, spatiale), soit : $x(t) = f_o + \Delta f_M (t) + N (t)$, avec f_o la fréquence centrale, $\Delta f_M (t)$ la modulante et N (t) le bruit interne.

Dans la mesure où l'information à détecter est contenue dans la variation de la fréquence du signal (i.e., la FM), les auteurs considèrent que la fréquence centrale du signal n'apporte aucune

information et décalent donc le signal de sorte que Fc = 0 Hz par souci de simplification mathématique. L'excursion de fréquence maximale est alors définie par x_m.

Un signal de FM peut être représenté d'un point de vue statistique par sa densité de probabilité fréquentielle p(x), qui indique la probabilité qu'un échantillon ait une valeur comprise entre x et x + dx. La densité de probabilité fréquentielle joue un rôle crucial dans le modèle de Hartmann et Klein. Il diffère du spectre du signal de FM dans le sens où la densité de probabilité fréquentielle est obtenue lorsque l'échantillonnage est court comparativement à la période de modulation, tandis que le spectre du signal nécessite un échantillonnage sur une échelle temporelle plus longue que la période de la modulation. En l'absence de bruit, la densité de probabilité fréquentielle est simplement proportionnelle au temps que le signal modulé passe à une coordonnée tonotopique donnée.

Un auditeur peut utiliser les échantillons de densité de probabilité fréquentielle successifs de différentes façons pour détecter la FM. En comparant deux échantillons successifs, un auditeur peut décider de la présence de la FM, si la différence entre les 2 échantillons est suffisamment importante.

Le modèle propose donc de calculer la densité de probabilité fréquentielle de deux échantillons successifs, puis de réaliser la différence entre ces deux densités de probabilités. On obtient ainsi une densité de probabilité pour la différence. Les auteurs font l'hypothèse que le bruit interne associé à chaque échantillon à la même variance.

Dans le cas où aucune modulation n'est présente, la différence de densité de probabilité est : $p_{\Delta 0N}(z) = r(z)$, avec r la convolution entre la densité de probabilité du bruit interne, supposé gaussien de moyenne nulle et de variance σ^2 et lui même et z la différence entre les deux échantillons x2 et x1.

Pour choisir l'intervalle contenant la modulation, l'observateur idéal choisira l'intervalle ayant la valeur absolue de z la plus grande. La densité de probabilité pour la valeur absolue de z est égale à deux fois la densité de probabilité pour la différence lorsque z est positif ou nul, et 0 sinon. Le

pourcentage de réponse correct, basé sur une simple différence est ensuite calculé à partir de la densité de probabilité obtenue précédemment.

Les opérations décrites ci-dessus constituent un modèle général selon lequel deux échantillons peuvent être utilisés pour obtenir une variable de décision afin de détecter une FM. A chaque échantillon correspond un bruit interne. Le modèle précédent relie donc bien la détection de FM à la discrimination de fréquence si l'on fait l'hypothèse que le bruit interne associé à l'échantillon dans la détection de FM présente des statistiques identiques au bruit associé à la perception de hauteur dans une tâche de discrimination de fréquence.

Hartmann et Klein (1980) suggèrent que la réalisation d'inter-corrélations entre les échantillons du signal et ceux issus d'un gabarit de la modulation, stocké en mémoire, pourrait permettre de détecter la présence d'une modulation plus efficacement. Mathématiquement, ce processus s'apparente à une somme de corrélations, chaque échantillon contribuant à cette somme.

Le modèle de Hartmann et Klein, basé sur des processus de soustractions et de corrélations, n'est valide que sous certaines hypothèses:

Un sujet ne peut réaliser ce processus d'inter-corrélation que si le signal est défini (connu) correctement. Dans le cas de la détection d'une FM, cela correspond à la situation où la modulante, la fréquence et la phase de départ du signal restent stables au cours de l'expérimentation. Toutefois, lorsque la profondeur de modulation est très faible de sorte qu'elle soit tout juste détectable, il est impossible de discerner la nature de la modulante pour un sujet (Klein & Hartmann, 1979). Pour cette raison, le processus de corrélation entre le stimulus et le gabarit de la modulation stocké en mémoire ne peut être aussi précis. Dans ce cas, le sujet utilise plus probablement la périodicité de la modulation: il recherche une augmentation de la hauteur, suivie d'une diminution de celle-ci un demi-cycle plus tard. Ainsi, Hartmann et Klein suggèrent que seule l'information de périodicité est utilisée pour corréler les différents échantillons.

Le modèle décrit précédemment estime le pourcentage de bonnes réponses obtenu à partir d'une seule soustraction entre deux échantillons. Toutefois, dans une tâche de détection de FM,

les stimuli sont suffisamment longs pour permettre au sujet de réaliser des soustractions multiples. A mesure que le sujet réalise des soustractions supplémentaires, sa performance en détection de FM augmente. Les auteurs suggèrent qu'il existe toutefois une limite au-delà de laquelle le sujet ne peut toutefois plus réaliser de soustractions supplémentaires.

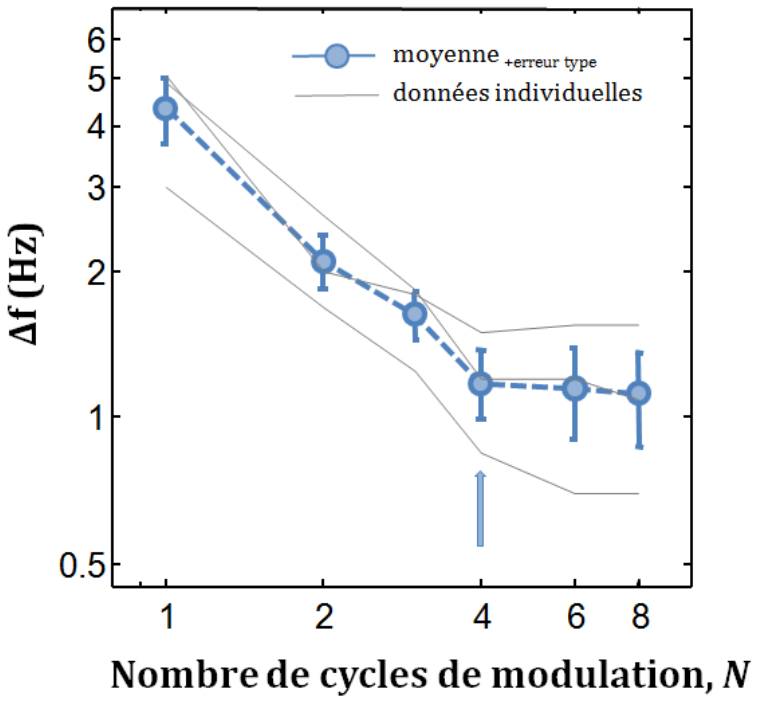


Figure 26. *Evolution des FMDLs en fonction du nombre de cycles de modulation. Les données individuelles des 3 sujets sont représentées par les lignes grises fines. La moyenne et l'erreur type des FMDL sont représentés par la ligne bleue pointillée. La flèche symbolise le nombre de cycles à partir duquel l'intégration temporelle n'est plus significative. Retracé à partir des données de Hartmann et Klein (1980)*

De façon à définir cette limite, Hartmann et Klein (1980) réalisent une mesure d'intégration temporelle en détection de FM. Hartmann et Klein cherchent à déterminer le nombre maximum de cycles de modulation (et donc le nombre maximal de soustractions) qu'un sujet peut utiliser pour réaliser une tâche de détection de FM. Ils supposent que la performance va augmenter lorsque le nombre de cycles de modulation N augmente, et ce de façon proportionnelle à $\sqrt{N}$, jusqu'à l'obtention d'un maxima au delà duquel la performance n'augmente plus.

Ils mesurent les FMDLs chez trois sujets normo-entendants (Fc= 0.8 kHz; Fm = 4 Hz; niveau de présentation de 75 dB SPL, procédure 2 bas, 1 haut, incrément fixe = 0.16 Hz, aucune AM interférente). Le nombre de cycle de modulation N varie entre 1 et 8 (cf. Figure 26).

Les FMDLs s'améliorent lorsque le nombre de cycles de modulation N augmente jusqu'à N=4. Au delà de cette limite, aucune des différences entre les FMDLs n'est statistiquement significative. Par ailleurs, l'intégration temporelle en FM observée par Hartmann et Klein (1980) est plus forte qu'attendue. Les FMDLs n'augmentent pas de façon proportionnelle à $\sqrt{N}$, même lorsque l'on fait abstraction du premier cycle de modulation, pour lequel un effet négatif de l'adaptation à court terme sur le FMDL ne peut être exclu.

Les auteurs en concluent que les cycles successifs de modulation ne peuvent être considérés comme étant statistiquement équivalents et indépendants. Dans ce cas, leur modèle, se basant sur l'hypothèse de soustractions indépendantes, ne peut être justifié que par souci de simplification mathématique.

Hartmann et Klein (1980) démontrent un effet d'intégration temporelle en détection de FM pour une cadence de modulation et une fréquence porteuse basse: les FMDLs mesurés chez des sujets normo-entendants s'améliorent lorsque le nombre de cycle de modulation augmente, jusqu'à atteindre une valeur asymptotique aux alentours de 4 cycles de modulation.

Hartmann et Klein proposent l'unique modèle permettant de rendre compte de ces données d'intégration temporelle de FM. Valable uniquement pour des basses cadences de modulation, il se rapproche qualitativement du modèle multiple looks de Viemeister et Wakefield (1991): le signal de FM est échantillonné, puis les densités de probabilités fréquentielles de chaque échantillon sont comparées deux à deux (par soustraction). Plusieurs soustractions successives sont réalisées, de façon à rendre compte de l'intégration temporelle. Toutefois, le modèle d'Hartmann et Klein diverge des modèles proposés pour la détection d'AM au sens où il considère que la détection de FM croît jusqu'à atteindre sa valeur asymptotique.

7. Objectifs principaux et plan de thèse

7.1 Problématique

L'objectif de cette thèse est d'améliorer notre compréhension du traitement d'AM et de FM et de clarifier les effets du vieillissement et de dommages cochléaires sur la sensibilité auditive à l'AM et à la FM.

Sur un plan fondamental, ce programme de recherche vise à tester l'existence de deux codes sensoriels distincts pour le traitement d'AM et de FM. Il s'agit ici de séparer la contribution relative des indices d'enveloppe temporelle et de structure temporelle fine à la détection d'AM et de FM lorsque la fréquence porteuse est basse. Ceci est réalisé à l'aide de plusieurs études psychophysiques évaluant l'effet de paramètres acoustiques tels que la cadence et le nombre de cycles de modulation et l'effet du vieillissement et d'une lésion cochléaire sur la détection d'AM et de FM. L'étude de l'intégration temporelle (à savoir, l'effet du nombre de cycles de modulation) permet également de séparer la contribution de facteurs de bas niveau (sensoriels) de la contribution de facteurs de plus haut niveau (mnésiques et décisionnels) à la détection d'AM et de FM. Enfin, la simulation de modèles computationnels permet d'apprécier la contribution d'informations d'enveloppe temporelle dans les différentes tâches de détection d'AM et de FM.

Sur un plan plus clinique, ce travail doctoral permet de préciser les conséquences audiologiques - sensorielles et cognitives - de la presbyacousie (une perte auditive liée au vieillissement) sur les traitements auditifs. Il s'agit de caractériser finement les troubles auditifs supraliminaires temporels associés à la presbyacousie, et de déterminer dans quelle mesure vieillissement et dommages cochléaires affectent le stockage en mémoire auditive à court terme ainsi que les stratégies d'écoute.

Ce programme de recherche sur le traitement d'AM et de FM est mis en œuvre de quatre manières:

1) Des mesures des capacités de détection d'AM et de FM sont réalisées chez *trois populations différentes*: des sujets normo-entendants jeunes, des sujets normo-entendants âgés et des sujets malentendants âgés. Ces mesures visent à séparer les effets du vieillissement et de la perte cochléaire sur le traitement d'AM et de FM, des effets souvent confondus dans les études antérieures.
2) L'étude *conjointe* des capacités de détection d'AM et de FM en basse et haute cadence de modulation pour une porteuse basse fréquence (< 1 kHz) vise à estimer les effets du vieillissement et de la perte cochléaire sur le traitement des informations d'enveloppe et de structure fine supposées être impliquées dans la détection d'AM et de FM (la sensibilité à l'AM et la sensibilité à la FM étant souvent étudiées séparément dans les études antérieures).
3) L'étude de l'*intégration temporelle* en détection d'AM et de FM pour chacun de ces groupes de sujets vise à estimer l'effet du vieillissement et de la perte cochléaire sur l'efficacité de traitement d'AM et de FM- à savoir, la capacité du système auditif central à utiliser de façon optimale l'information sensorielle disponible. En effet, les études antérieures ne permettent pas de déterminer dans quelle mesure les déficits de traitement d'AM et de FM observés chez des sujets âgés et/ou malentendants résultent de troubles du codage sensoriel des informations d'enveloppe et de structure fine ou de troubles de l'efficacité de traitement des modulations. Or, le vieillissement et/ou la perte auditive semblent affecter la mémoire auditive à court terme (Gallun et al., 2012) et/ou pourrait affecter les stratégies d'écoute. L'intégration temporelle est pour partie contrainte par des mécanismes cognitifs correspondant aux capacités de stockage en mémoire à court terme et aux processus de prise de décision pour la détection et discrimination d'AM et de FM. L'étude de l'intégration temporelle en détection d'AM et

de FM pour chacun de nos groupes de sujets permet ainsi de séparer les contributions de facteurs sensoriels et cognitifs aux déficits temporels observés.

4) Un modèle computationnel de traitement d'enveloppe temporelle basé sur le concept du banc de filtres de modulation et sur un mécanisme d'appariement de gabarit est développé puis testé dans les différentes conditions expérimentales. Les simulations permettent d'apprécier la contribution des informations d'enveloppe temporelle à la sensibilité et à l'intégration des indices d'AM et de FM pour les trois populations de sujets.

7.2 Plan de thèse

Le premier chapitre présente les résultats d'une étude psychophysique, dans laquelle les AMDLs et FMDLs sont évalués à deux cadences de modulation pour différentes durées de stimulation chez des sujets normo-entendants jeunes et âgés.

Ce premier projet de recherche psychophysique vise à: i) tester l'existence de deux codes sensoriels distincts pour le traitement d'AM et de FM ; ii) évaluer spécifiquement l'impact de l'âge sur la sensibilité à l'AM et la FM; iii) apprécier la contribution respective des facteurs sensoriels et de l'efficacité de traitement dans les déficits de détection d'AM et de FM causés par le vieilissement.

Les résultats montrent que: i) pour les deux groupes des sujets, les AMDLs et FMDLs s'améliorent de façon similaire lorsque la durée de stimulation augmente: le vieillissement préserve donc les processus centraux impliqués dans l'intégration temporelle; ii) le vieillissement altère les AMDLs et FMDLs, mais ce principalement en FM à basse cadence de modulation; iii) pour l'ensemble des groupes, l'intégration temporelle est différente pour l'AM et la FM (l'intégration étant plus forte en AM qu'en FM). Ces résultats corroborent l'hypothèse d'un double code pour la détection de FM et suggèrent que le vieillissement, bien qu'il dégrade tant

les indices d'enveloppe temporelle que les indices de TFS, affecterait plus fortement les indices de TFS.

Le deuxième chapitre prolonge l'étude précédente, et porte sur la sensibilité et l'intégration temporelle des indices d'AM chez des sujets âgés présentant une perte auditive neurosensorielle légère à modérée. Les données des sujets malentendants sont comparées à celles obtenues sur des sujets normo-entendants de l'étude précédente. Un modèle informatique de traitement d'enveloppe en sortie d'un filtre cochléaire est développé afin de rendre compte des données expérimentales. Cette étude combinant mesures psychoacoustiques et modélisation informatique permet de clarifier la contribution respective des facteurs sensoriels et non sensoriels à la détection d'AM chez les sujets malentendants.

Les résultats montrent que: i) les AMDLs sont meilleurs chez les sujets malentendants que chez les sujets normo-entendants; ii) l'intégration temporelle des indices d'AM est plus importante chez les sujets malentendants que chez les sujets normo-entendants. Les données psychophysiques et les simulations informatiques suggèrent que l'amélioration de la sensibilité aux indices d'AM et la plus forte intégration temporelle sont principalement attribuables à la perte de la compression cochléaire. Les simulations du modèle suggèrent également que la perte auditive préserve la mémoire à court terme et les mécanismes décisionnels impliqués dans la détection d'AM.

Le dernier chapitre présente les résultats d'une étude évaluant les FMDLs avec ou sans AM interférente chez des sujets malentendants âgés. Les données des sujets malentendants sont comparées aux FMDLs et AMDLs obtenus dans les études précédentes. Une nouvelle version du modèle de traitement d'enveloppe utilisant plusieurs canaux fréquentiels est développée afin de rendre compte des données psychophysiques. Cette dernière étude vise à clarifier la contribution des facteurs sensoriels (e.g., la dégradation de la conversion FM/AM) et non sensoriels dans la détection de FM chez les sujets malentendants.

Les résultats montrent que la perte cochléaire: i) dégrade la sensibilité à la FM, mais n'affecte pas l'intégration temporelle des indices de FM; ii) améliore la sensibilité à l'AM, et renforce le processus d'intégration temporelle. Pris ensemble, ces résultats suggèrent que: i) la détection d'AM et de FM font appel à des mécanismes distincts; ii) les lésions cochléaires dégradent les mécanismes impliqués dans la détection de FM, mais préservent les mécanismes de rétention en mémoire à court terme et le processus de prise de décision. Le modèle computationnel de traitement d'enveloppe permet de rendre compte: i) des AMDLs mesurés aux deux cadences de modulation; i) des FMDLs, mais ce uniquement pour la cadence la plus rapide. Cette incapacité du modèle à rendre compte des FMDLs à basse cadence de modulation suggère que les seuls indices d'enveloppe temporelle ne permettent pas d'expliquer les performances de sujets, et corrobore l'hypothèse de la contribution d'indices autres que ceux issus du pattern d'excitation pour la détection de FM à basse cadence.

Dans le dernier chapitre de cette thèse, les implications cliniques de ces études sont discutées du point de vue de la définition, du diagnostic et de la réhabilitation de la presbyacousie.

CHAPITRE 2

EFFETS DU VIELLISSEMENT SUR LA DETECTION D'AM ET DE FM

8. Comparing the effects of age on amplitude modulation and frequency modulation detection

8.1 Résumé:

Frequency modulation (FM) and amplitude modulation (AM) detection thresholds were measured at 40 dB sensation level for young (22-28 years) and older (44-66 years) listeners with normal audiograms for a carrier frequency of 500 Hz and modulation rates of 2 and 20 Hz. The number of modulation cycles, *N*, varied between 2 and 9. For FM detection, uninformative AM at the same rate as the FM was superimposed to disrupt excitation-pattern cues. For both groups, AM and FM detection thresholds were lower for the 2-Hz than for the 20-Hz rate, and AM and FM detection thresholds decreased with increasing *N*. Thresholds were higher for older than for younger listeners, especially for FM detection at 2 Hz, possibly reflecting the effect of age on the use of temporal-fine-structure cues for 2-Hz FM detection. The effect of increasing *N* was similar across groups for both AM and FM. However, at 20 Hz, older listeners showed a greater effect of increasing *N* than younger listeners for both AM and FM. The results suggest that ageing reduces sensitivity to both excitation-pattern and temporal-fine-structure cues for modulation detection, but more so for the latter, while sparing temporal integration of these cues at low modulation rates.

8.2 Article:

L'article suivant a été publié dans *The Journal of the Acoustical Society of America* en mai 2016.

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. DOI: 10.1121/1.4953019.

I. INTRODUCTION

Information about spectro-temporal changes in sounds, including speech, is conveyed in the auditory system by two mains types of cues: (1) excitation-pattern (place) cues, related to the distribution of excitation along the basilar membrane within the cochlea; (2) temporal-fine-structure (TFS) cues related to the waveform of the stimulus at each point on the basilar membrane. Place cues are conveyed in the auditory nerve by differences in firing rate of neurons with different characteristic frequencies whereas TFS cues are conveyed by the detailed timing of the action potentials in neurons with different characteristic frequencies.

There is evidence that sensitivity to TFS declines with increasing age (for a review, see Moore, 2014). However, most of the data supporting this idea were obtained using psychophysical tests whose outcome is affected both by the availability of TFS information and by "processing efficiency," *i.e.* the ability of the central auditory system to make use of available sensory information. One such test involves measurement of the threshold for detecting low-rate frequency modulation (FM); the rationale for this is described later. This paper presents a study in which thresholds for detecting both FM and amplitude modulation (AM) were measured for younger and older listeners with normal audiograms. It is generally assumed that AM detection does not depend on the use of TFS information, whereas detection of low-rate FM probably does (e.g., Moore and Sek, 1995). Hence, by comparing the effect of age on AM and FM detection in the same listeners, it should be possible to tease out the effects of changes in the use of TFS information and changes in processing efficiency. Modulation rates of 2 and 20 Hz were used, since it has been argued that FM detection for a 2-Hz rate depends on the use of TFS information (Moore and Sek, 1995; Sek and Moore; 1995; Ernst and Moore, 2010, 2012), while FM detection for a 20-Hz rate probably does not. Again, comparison of results at the two FM rates should allow us to tease out the effects of changes in the ability to use TFS information and changes in processing efficiency.

As a way of checking the role of TFS as opposed to excitation-pattern cues in the detection of FM,

modulation detection was also assessed as a function of the number of modulation cycles, *N* (between 2 and 9). AM and FM detection both improve (i.e., thresholds decrease) with increasing *N* (Viemeister, 1979; Hartmann and Klein, 1980; Sheft and Yost, 1990; Dau *et al*., 1997), an effect that probably reflects relatively central processes involving "multiple looks" (Viemeister and Wakefield, 1991) or a template-matching process (Hartmann and Klein, 1980; Dau *et al*., 1997). Here, we refer to this improvement as "temporal integration", not in the sense that energy is integrated over time, but in the sense that information is integrated over time. If AM and FM detection were both based solely on the detection of changes in excitation level, then one would expect that detection thresholds for AM and FM would improve in a similar way with increasing *N*. However, if low-rate FM detection is based on the use of TFS cues, then temporal integration might differ from that obtained for FM detection at a higher rate, and for AM detection. If such a difference in temporal integration were found, it would support the idea of two different mechanisms for FM detection. However, the lack of a difference in temporal integration would not disprove the idea that there are two mechanisms. We are not aware of any previous comparisons of temporal integration for AM and FM detection for a very low rate (2 Hz) and an intermediate rate (20 Hz).

A. Role of TFS cues in the detection of low-rate FM

We consider next evidence for the role of TFS in the detection of low-rate FM. Several studies have investigated whether the ability to detect, discriminate and recognize FM patterns depends on the use of excitation-pattern (place) or TFS information. Changes in frequency may be perceived by monitoring changes in excitation level (that is, temporal-envelope cues) at one place or at multiple places on the excitation pattern (Zwicker, 1952, 1956; Moore and Sek, 1994; Zwislocki and Nguyen, 1999). TFS information about FM is conveyed by changes in the pattern of phase locking over time. For most mammals, the precision of phase locking is constant for frequencies up to about 600-2000 Hz and then declines, becoming very weak at 3500-6000 Hz (Kiang *et al*., 1965; Rose *et al*., 1967; Johnson, 1980; Palmer *et al*., 1986), although the exact upper limit in humans is still debated (Heinz *et al*., 2001; Moore and Sek, 2009b; Verschooten

and Joris, 2014). Several researchers have suggested that changes in the pattern of phase locking over time may be used to perceive FM, at least for low FM rates and for carrier frequencies of 4000 Hz and below (Demany and Semal, 1989; Moore and Glasberg, 1989; Moore and Sek, 1995, 1996; Sek and Moore, 1995; Moore and Skrodzka, 2002). It has been proposed that TFS cues are not used to detect FM with rates above about 10 Hz because the mechanism for "decoding" the TFS information is "sluggish" and cannot track rapid changes in frequency (Moore and Sek, 1995, 1996; Sek and Moore, 1995).

B. Role of excitation pattern cues in FM detection.

To assess the role of excitation-pattern cues in FM detection, Moore and Sek (1996) measured FM detection thresholds for a wide range of combinations of carrier frequency, f_c, and modulation rate, f_m, with and without sinusoidal AM with a 6-dB peak-to-valley ratio applied to all stimuli in a forced-choice trial. The AM had the same modulation rate as the FM, and the starting phase of the AM was chosen at random for each stimulus. The AM was intended to disrupt excitation-pattern cues for FM detection by introducing large fluctuations in excitation level that were uninformative about the FM. The added AM adversely affected performance and, for f_c below 4000 Hz, the adverse effect increased with increasing f_m, consistent with the idea that excitation-pattern cues play a greater role for higher f_m. For f_c = 6000 Hz, the adverse effect of the added AM was similar for all f_m, consistent with the idea that, for very high f_c, excitation-pattern cues dominate for all f_m.

C. Effects of age on FM detection.

As noted earlier, the results of several studies suggest that sensitivity to TFS cues declines with increasing age (Ross *et al*., 2007; Grose and Mamo, 2010; Hopkins and Moore, 2011; Moore *et al*., 2012a,b; Füllgrabe, 2013; King *et al*., 2014; Füllgrabe *et al*., 2015). On the other hand, several studies indicate that frequency selectivity does not change with age provided that audiometric thresholds remain normal (Lutman *et al*., 1991; Peters *et al*., 1992; Hopkins and Moore, 2011),

which suggests that the excitation patterns evoked by modulated signals should be comparable for younger and older listeners. Hence, if FM is detected solely using excitation-pattern cues, the change in FM detection thresholds with age should be similar for low and high rates. In contrast, if low-rate FM detection depends on the use of TFS cues, then FM detection thresholds for a low rate should change with age more than FM detection thresholds for a high rate. However, there is no reason to expect that the effect of age would be different for low- and high-rate AM detection.

We are not aware of any previous studies that have assessed both AM and FM detection as a function of age, using both low and high modulation rates. Takahashi and Bacon (1992) measured AM detection thresholds as a function of AM rate (i.e. temporal modulation transfer functions, TMTFs) using a noise carrier. The data showed a progressive increase of thresholds with increasing age, especially for the lowest rate used (f_m = 2 Hz), but an analysis of variance (ANOVA) showed that the effect of age was not significant. He *et al*. (2007) measured TMTFs for f_c = 500 Hz. Listeners were given extensive training for f_m = 5 Hz. They found higher AM detection thresholds for older than for younger listeners for all AM rates except the lowest used (5 Hz). He *et al*. (2008) showed that sensitivity to FM for f_c = 500 Hz declined with increasing age, but they used only a single modulation rate, 5 Hz. Schoof and Rosen (2015) measured TMTFs using a noise carrier and measured FM detection thresholds for f_c=1000 Hz and a modulation rate of 2 Hz. Inconsistent with the above studies, they did not observe any significant effect of age on either AM or FM detection thresholds. Füllgrabe *et al*. (2015) measured TMTFs using a 4000-Hz sinusoidal carrier and found that AM detection thresholds were higher for older than for young listeners with matched normal audiograms. However, the shapes of the TMTFs were similar for the two groups, suggesting that increasing age is associated with reduced processing efficiency but not reduced temporal resolution for envelope changes.

In summary, the goal of this study was to assess whether the increase in low-rate FM detection thresholds with increasing age reflects a decrease in the availability of TFS information or

whether it reflects a change in processing efficiency with age. To achieve this, both AM and FM detection thresholds were measured for two modulation rates and as a function of the number of modulation cycles, using younger and older listeners with normal audiograms. As a means of encouraging the use of TFS cues for FM detection, the thresholds for detecting FM were measured in the presence of added AM in all intervals of the forced-choice task (Moore and Sek, 1996). This was intended to disrupt the use of excitation-pattern cues. We hypothesized that age would have a greater effect on low-rate FM detection than on high-rate FM detection, but that the effect of age would be similar for low- and high-rate AM detection. We also hypothesized that temporal integration might differ for low-rate FM detection and for high-rate FM detection, or AM detection

II. METHOD

A. Listeners

There were 15 young listeners (Young), aged between 22 and 28 years (mean = 24 years; SD = 2 years) and 14 older listeners (Older), aged between 47 and 66 years (mean = 57 years; SD = 8 years). All listeners had audiometric thresholds less than 20 dB HL for the tested (right) ear for octave frequencies between 0.25 and 3 kHz (see Fig. 1). Some of the Older listeners (n=7) had elevated audiometric thresholds above 3 kHz. The mean audiometric threshold at 0.5 kHz was 8.1 dB HL (SD = 3.1 dB) for the Young group and 8.9 dB HL (SD = 4.9 dB) for the Older group. A *t*-test conducted on audiometric thresholds at 0.5 kHz showed no significant difference between groups [$t(27)= -0.61$; $p=0.54$]. 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).

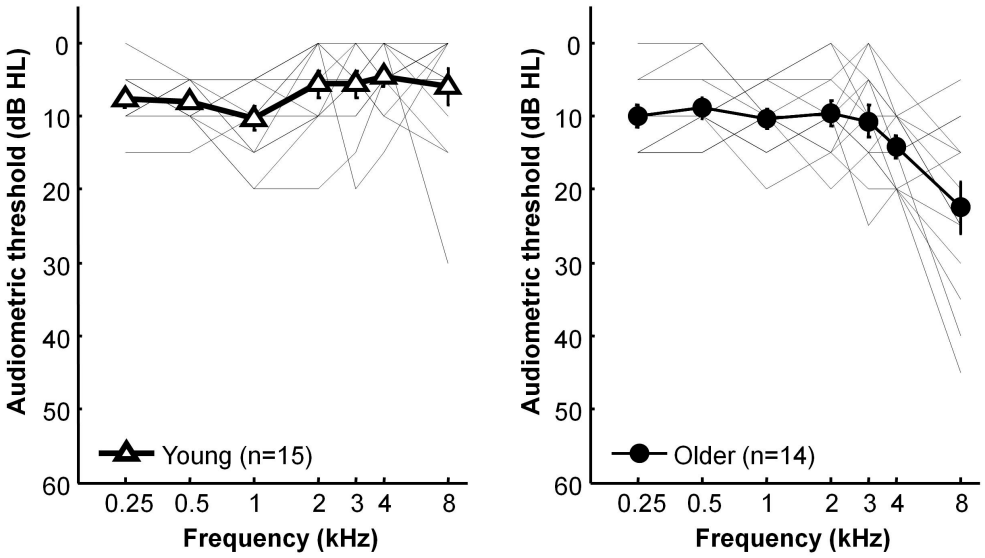


Figure 1. *Individual and mean audiometric thresholds for the younger (Young, left) and older (Older, right) listeners for the right ear. Error bars indicate ±1 standard error of the mean. They are omitted when they are smaller than the symbol used to plot the data.*

B. Stimuli

All stimuli were generated digitally at a sampling rate of 44.1 kHz. They were presented at a mean sensation level (SL) of 40 dB using Sennheiser HD 280 pro headphones (Old Lyme, CT) and an external soundcard (ECHO Indigo DJ 1-2, Santa Barbara, California; 16-bit resolution). Stimuli were presented only to the right ear. At 0.5 kHz, audiometric thresholds ranged between 5 and 15 dB HL for the Young listeners and between 0 and 15 dB HL for the Older listeners. Thus, presentation levels covered the same range for the two groups, i.e. 40-55 dB SPL.

A two-interval forced-choice task was used. Each trial contained a target (modulated) and a standard (unmodulated) stimulus, in random order, with a silent interval of 600 ms between them. The root-mean-square level of the two stimuli was, on average, the same. However, to discourage the use of level cues, the levels of the standard and target were roved independently within the range ±1.5 dB. For each AM and FM stimulus, raised-cosine ramps were applied at the onset and at the offset. The ramp durations were 250 and 25 ms for the modulation rates of 2

and 20 Hz, respectively. The duration of each signal was determined by the number of modulation cycles, *N*, and the modulation rate, f_m. *N* was set to 2, 3, 4, 5, or 9 cycles.

1. AM detection

The carrier was a sinusoid with f_c= 500 Hz. This was either unmodulated to produce the standard stimulus, or modulated sinusoidally in amplitude to produce the target stimulus. Eq. 1 describes the target stimulus T(t):

$$T(t) = [1 + m \sin(2\pi f_m t + 3\pi/2)]\sin(2\pi f_c t + \varphi) \quad [1]$$

where f_m is the modulation rate (2 or 20 Hz), φ is the starting phase of the carrier, and t is time (expressed in s). The value of φ was randomly chosen for each stimulus. The starting phase of the modulator, 3π/2 radians, resulted in a modulation minimum at the onset and offset of the stimulus. The modulation depth, *m*, was adaptively varied to determine the threshold.

2. FM detection

The carrier was a sinusoid with f_c = 500 Hz. The standard stimulus contained AM but not FM. The target stimulus contained both AM and FM. Eq. 2 describes the target stimulus T(t):

$$T(t) = AM(t) \sin[(2\pi f_c t + \Phi) + \beta (\sin(2\pi f_m t + \varphi))] \quad [2]$$

with: $AM(t) = [1 + 0.33\sin(2\pi f_m t + 3\pi/2)]$

and: $\beta = \Delta f / f_m$

where Δf is the frequency excursion in Hz, f_m is the modulation rate (2 or 20 Hz), φ is the starting phase of the FM, Φ is the starting phase of the carrier, and t is time in s. The values of φ, and Φ were randomly and independently chosen for each stimulus. Although the starting phase of the AM was fixed, its phase relative to that of the FM was random (because φ was chosen randomly), so the AM should still have been effective in disrupting the use of excitation-pattern cues. For the standard stimulus, Δf was set to 0. The value of Δf for the target was adaptively varied to determine the threshold.

C. Procedure

Detection thresholds were measured using a two-interval forced-choice adaptive procedure estimating the 70.7% point on the psychometric function (Levitt, 1971). Each interval was marked by a light on the computer screen. Listeners were asked to indicate which interval contained the fluctuating sound for the AM detection task or the "extra" pitch fluctuation for the FM detection task. Visual feedback as to the correct response was provided after each response. The tracking variable was m for the AM detection task and Δf for the FM detection task. A run started with the tracking variable well above the estimated detection threshold. The step size was a factor of 1.58 until two reversals had occurred, and 1.26 thereafter. Fourteen reversals were obtained for each run, and the threshold estimate for that run was taken as the geometric mean of the tracking variable at the last six reversals. Two threshold estimates were obtained for each condition. The final estimate of the threshold was taken as the geometric mean of the two threshold estimates.

Thresholds were measured first for AM detection and then for FM detection. Within a block (AM or FM), the order of conditions (5 values of N x 2 values of fm) was chosen using a latin-square design. This reduced group effects of learning and fatigue. A test session was terminated when the listener reported fatigue or when the experimenter judged that the listener was becoming fatigued. No training was given before the beginning of the experiment. However, each listener was presented with some practice trials at the start of each block

III. RESULTS

A. Effects of N and f_m

Figure 2 shows the mean modulation detection thresholds for each group plotted as a function of N, for f_m = 2 Hz (circles) and f_m = 20 Hz (triangles). The top and bottom panels show AM and FM detection thresholds, respectively. The thresholds for Young and Older listeners are shown by dashed and continuous lines, respectively. AM detection thresholds are expressed as $20\log_{10}m$.

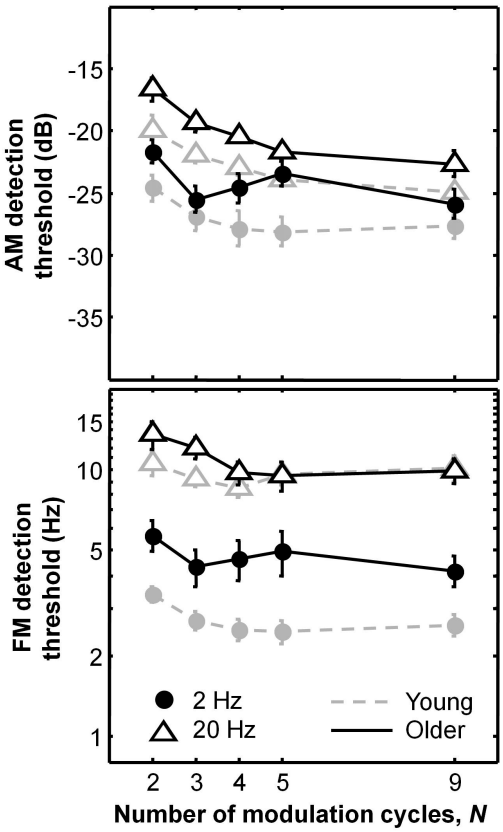


Figure 2. *Mean AM and FM detection thresholds for the Young listeners (dashed lines) and Older listeners (continuous lines), as a function of the number of modulation cycles, N, with modulation rate, f_m, as parameter (circles: 2 Hz; triangles: 20 Hz). The AM detection thresholds are expressed in dB as $20log_{10}m$. The FM detection thresholds are expressed in Hz (log scale). Error bars indicate ±1 standard error of the mean.*

A repeated-measures ANOVA was conducted on the AM detection thresholds (expressed as $20\log_{10}m$) with between-subjects factor group (2 levels) and within-subjects factors *N* (5 levels) and modulation rate (2 levels). The effect of group is discussed later. The effect of modulation rate was significant [$F(1, 27) = 37.8$, $p < 0.001$]. For both groups, AM detection thresholds were lower for f_m = 2 Hz than for f_m = 20 Hz for each value of *N*. For the Young listeners, the difference was about 5 dB for $N < 5$, and about 3 dB for $N = 9$. The thresholds for the Young listeners are broadly consistent with those found in previous studies of AM detection for normal-hearing listeners (Zwicker, 1952; Kohlrausch *et al*., 2000; Moore and Glasberg, 2001), although information is lacking about AM detection for very low modulation rates (2 Hz) and low carrier frequencies (500 Hz).

The effect of *N* was significant [$F(4, 108) = 26.5$, $p < 0.001$]. For each modulation rate, AM detection thresholds for Young listeners decreased with increasing *N*. For Older listeners, at f_m =

20 Hz, AM detection thresholds also decreased with increasing *N;* at f_m = 2 Hz, the change in AM detection thresholds as a function of *N* was somewhat irregular, but thresholds were lower when *N* = 9 cycles than when *N* = 2 cycles. For the Young listeners, the decrease was 3-5 dB when *N* was increased from 2 to 9 cycles, but most of the effect of *N* occurred for $N \leq 5$. We are not aware of any previous study of temporal integration for AM detection using a sinusoidal carrier. The effect of *N* on AM detection is broadly consistent with the results of Sheft and Yost (1990) obtained with gated wideband noise carriers and normal-hearing listeners.

There was a significant interaction between *N* and f_m [$F(4, 108) = 3.8$, $p = 0.006$]. This may reflect the fact that the decrease in AM detection thresholds as *N* increased from 2 to 9 was greater for f_m = 20 Hz (5.4 dB) than for f_m = 2 Hz (3.3 dB). None of the other interactions were significant.

A linear regression analysis was conducted on the individual AM detection thresholds (log-transformed AM detection thresholds versus log(*N*)). A repeated-measures ANOVA was conducted on the resulting slopes with between-subject factor group (2 levels) and within-subjects factor f_m (2 levels). The ANOVA showed a significant effect of f_m [$F(1, 27) = 9.0$, $p = 0.005$] but no significant effect of group [$F(1, 27) < 1$] and no significant interaction [$F(1,27) < 1$]. These analyses confirm that temporal integration for AM detection is stronger at 20 Hz than at 2 Hz for both groups.

A repeated-measures ANOVA was conducted on the (log-transformed) FM detection thresholds with the same factors as described above. The effect of group is discussed later. The effect of modulation rate was significant [$F(1, 27) = 234.1$, $p < 0.001$], thresholds being lower for f_m = 2 Hz than for f_m = 20 Hz. The thresholds for the Young listeners are comparable to those reported by Moore and Sek (1996) and Moore and Skrodzka (2002). The effect of *N* was significant [$F(4, 108) = 10.0$, $p < 0.001$]. For each modulation rate and each group, FM detection thresholds decreased with increasing *N* up to 3 or 4, with no consistent change for greater *N*. The decrease

for the Young listeners, corresponding to a factor of about 1.4 for f_m = 2 Hz and 1.1 for f_m = 20 Hz, is smaller than the factor of 1.8 found by Hartmann and Klein (1980) for f_m = 4 Hz and f_c = 800 Hz, possibly because we used interfering AM and they did not. The interaction between N and f_m was not significant [$F(4, 108) < 1$, $p > 0.05$], indicating that FM detection thresholds decreased similarly with increasing N for f_m = 2 Hz and f_m = 20 Hz.

B. Effects of group

For AM detection, the effect of group was significant [$F(1, 27) = 9.3$, $p = 0.005$]. For most conditions, the AM detection thresholds for the Older listeners were slightly higher (by about 2 dB) than those for the Young listeners. There were no interactions involving group, indicating that: i) the pattern of results for AM detection did not differ significantly for the two groups; ii) the Older group performed more poorly overall.

For FM detection, the effect of group was significant [$F(1, 27) = 8.6$, $p = 0.007$]. For most conditions, FM detection thresholds were higher for the Older than for the Young listeners, although this was not the case for N = 5 or 9 and f_m = 20 Hz. The interaction between group and f_m was significant [$F(1, 27) = 6.8$, $p = 0.014$], indicating that the effect of age was different for f_m = 2 Hz and f_m = 20 Hz. FM detection thresholds were higher for the Older than for the Young listeners by an average factor of about 1.7 for f_m = 2 Hz and 1.14 for f_m = 20 Hz. There was no significant interaction between group and N [$F(4, 108) < 1$, $p > 0.05$], indicating that the decrease in FM detection thresholds with increasing N was broadly similar for the Young and Older listeners. There was no significant interaction between group, f_m and N.

C. Further analysis of the effect of N

To compare temporal integration effects across groups and modulation type (AM vs FM), the data were normalized for each listener and each fm, by dividing each modulation threshold by the geometric mean detection threshold (expressed as m for AM detection, and in Hertz for FM

detection) across N for that listener. Figure 3 shows the mean normalized detection thresholds plotted as a function of N.

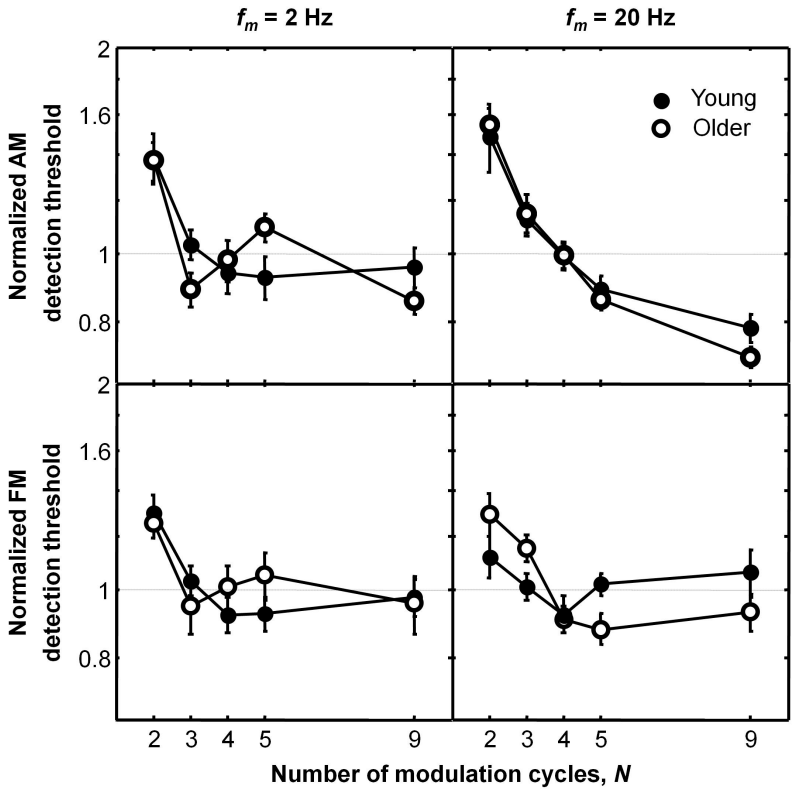


Figure 3. *Data replotted from Fig. 2, highlighting the effects of N. For each listener and each modulation rate, each modulation detection threshold was divided by the geometric mean detection threshold across N (expressed as m for AM detection, and in Hertz for FM detection). The arithmetic averages of the resulting normalized thresholds across listeners are plotted with error bars showing ±1 standard error.*

A repeated-measures ANOVA was conducted on the normalized detection thresholds with between-subjects factor group (2 levels) and within-subjects factors modulation type (2 levels), *N* (5 levels) and modulation rate (2 levels). The main effect of *N* was significant [$F(4, 108) = 43.3$, $p < 0.001$]. None of the other main effects were significant. The interaction between *N* and modulation type was significant [$F(4, 108) = 4.9$, $p = 0.001$]. For each modulation rate and each group, thresholds decreased more with increasing *N* for AM than FM. The interaction between *N* and modulation rate was significant [$F(4, 108) = 3.0$, $p = 0.022$]. Thresholds decreased more with increasing *N* for f_m = 20 Hz than for f_m = 2 Hz. No other two-way interactions were significant.

The three-way interaction between *N*, modulation type, and modulation rate approached but did not reach significance [$F(4, 108) = 2.3$, $p = 0.068$]. AM and FM detection thresholds tended to decrease similarly with increasing *N* for the two modulation rates, although there was a trend for AM detection thresholds to decrease more with increasing *N* for f_m = 20 Hz than for f_m = 2 Hz. There was no significant interaction between *N* and group [$F(4, 108) < 1$, $p > 0.05$] or between *N*, group and modulation type [$F(4, 108) < 1$, $p > 0.05$]. There was a three-way interaction between *N*, group and modulation rate [$F(4, 108) = 2.9$, $p = 0.025$], reflecting the observation that, for f_m = 20 Hz, AM and FM detection thresholds decreased slightly more with increasing *N* for the Older than for the Young listeners.

IV. DISCUSSION

A. Sluggishness and temporal integration for AM and FM detection

For both types of modulation (AM and FM), the modulation sensitivity of Young and Older listeners was poorer for the 20-Hz than for the 2-Hz modulation rate (see Fig. 2). However, this comparison was made with *N* equated across modulation rates, and, for a given *N*, the stimuli were a factor of 10 shorter in duration for the 20-Hz rate than for the 2-Hz rate. For stimuli with a fixed duration of a few hundred milliseconds, AM detection for rates close to 2 Hz tends to be *worse* than for rates close to 20 Hz (Sheft and Yost, 1990; Kohlrausch *et al*., 2000; Ernst and Moore, 2012). In contrast, FM detection for similar durations and for low and medium carrier frequencies tends to be better for rates close to 2 Hz than for rates close to 20 Hz (Moore and Sek, 1995, 1996; Sek and Moore, 1995), especially when uninformative AM is present in all intervals of a forced-choice task (Moore and Sek, 1996; Ernst and Moore, 2010, 2012), as in the present experiment. This pattern of results is consistent with the idea that FM detection does not depend solely on the use of excitation-pattern cues (transformation of FM to AM), but rather that an extra cue is used for FM detection at low rates, probably based on the use of TFS information.

Temporal integration (defined here by the improvement in modulation detection thresholds with increasing *N*) was found for both types of modulation (see Fig. 3). For FM detection, temporal integration did not differ markedly for the two FM rates. This suggests that temporal integration is distinct from the "sluggishness" that has been proposed to apply to the use of TFS information to detect low-rate FM (Moore and Sek, 1995, 1996; Sek and Moore, 1995). According to the "sluggishness" hypothesis, the rate of change of instantaneous frequency within a single FM cycle needs to be sufficiently slow for the TFS information to be extracted, but this appears to be separate from the process of combining information across successive FM cycles (temporal integration).

For AM detection, temporal integration was greater for the 20-Hz than for the 2-Hz rate. This may reflect limitations in short-term auditory memory since for the 2-Hz rate, the stimuli were a factor of 10 longer in duration, and the duration reached 4.5 sec when *N*=9. This duration exceeds the assumed duration of echoic memory (Darwin *et al.*, 1972; Ardoint *et al.*, 2008).

Temporal integration was greater for AM than FM, mainly for the 20-Hz rate (see Fig. 3). This seems puzzling at first sight, since it is usually assumed that, for a 20-Hz rate, FM is detected via FM-to-AM conversion. The difference in temporal integration for AM and FM may be linked to the fact that the FM detection thresholds were obtained in the presence of uninformative AM. The presence of the AM meant that the FM could be detected via changes in excitation level only if the fluctuations in excitation level were compared for the lower and upper sides of the excitation pattern, since the AM made the fluctuations at any single point on the pattern an unreliable cue. This comparison process may have different temporal integration properties from the process involved in simple AM detection.

As noted earlier, the improvement in FM detection threshold with increasing *N* was smaller than that reported by Hartmann and Klein (1980) for normal-hearing listeners using f_m = 4 Hz and f_c = 800 Hz. This discrepancy may be linked to the presence of uninformative AM in the present study but not in the study of Hartmann and Klein (1980). To assess this possibility, FM detection

thresholds were measured as a function of *N* (from 2 to 4) without interfering AM for 8 of the 15 Young listeners. The methods and procedures were identical to those described above. The data are shown in Fig. 4. FM detection thresholds collected with interfering AM for the same 8 Young listeners are plotted for comparison. A within-subjects ANOVA was conducted on the (log-transformed) FM detection thresholds with factors presence/absence of interfering AM (2 levels), *N* (3 levels) and modulation rate (2 levels). The main effect of interfering AM was significant [$F(1, 7) = 205.6$; $p < 0.001$], FM detection thresholds being better in the absence of interfering AM. The main effect of *N* was significant [$F(2, 14) = 30.5$; $p < 0.001$], FM detection thresholds decreased with increasing *N*. The main effect of modulation rate was significant [$F(1, 7) = 66.6$; $p < 0.001$], FM detection thresholds were better at f_m = 2 Hz than f_m = 20 Hz. The interaction between interfering AM and modulation rate was significant [$F(1, 7) = 22.4$; $p = 0.002$], the detrimental effect of the interfering AM being greater for f_m = 20 Hz than for f_m = 2 Hz. These findings are consistent with previous work (Moore and Sek, 1996; Ernest and Moore, 2010, 2012) and with the notion that excitation pattern cues play a greater role for f_m = 20 Hz than for f_m = 2 Hz.

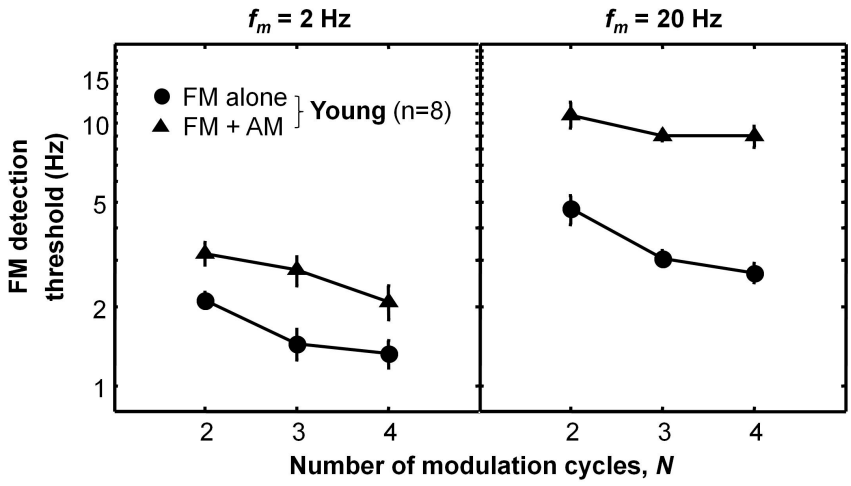


Figure. 4. *Mean FM detection thresholds for 8 Young listeners, as a function of N. The FM detection thresholds were measured with (filled triangles) and without (filled circles) interfering AM at the same rate as the FM.*

The two-way interaction between interfering AM and *N* was not significant at the 0.05 level

[$F(2,14) = 3$; $p = 0.089$] but the three-way interaction between interfering AM, N and modulation rate was marginally significant [$F(2,14) = 3.5$; $p = 0.057$]. For the 2-Hz rate, the decrease in threshold with increasing N was the same in the absence (factor of 1.5) and in the presence (factor of 1.5) of interfering AM. For the 20-Hz rate, the decrease in threshold with increasing N was greater in the absence (factor of 1.8) than in the presence (factor of 1.2) of interfering AM. The decrease factor observed here in the absence of AM at both modulation rates is only slightly less than the factor of about 2 found by Hartmann and Klein (1980) for an FM rate of 4 Hz, consistent with the idea that the smaller temporal integration found in the main experiment was at least partly due to the presence of the added AM.

B. Effects of age on AM and FM sensitivity

The modest but significant effect of age on AM detection observed in the present study (see Fig. 2, top panel) is consistent with most previous results obtained with older listeners than those used here (He *et al*., 2008: mean age = 71 years; Füllgrabe *et al*., 2015: mean age = 67 years; present study: Older group mean age = 57 years), although Schoof and Rosen (2015) reported no effect of age for detection of 20-Hz AM applied to a noise carrier.

The effect of age on FM detection for f_m = 2 Hz (mean thresholds a factor of 1.7 higher for the Older than for the Young listeners) is broadly consistent with the detrimental effects of age for a carrier frequency of 500 Hz reported by He *et al*. (2008) for f_m = 5 Hz (increase of FM detection threshold by a factor of about 2.5) and Grose and Mamo (2012) for f_m = 2 Hz (increase of FM detection threshold by a factor of about 1.8). However, Schoof and Rosen (2015) found no significant effect of age for detection of 2-Hz FM applied to a 1000-Hz carrier.

The small differences across studies that found an age effect are probably due to differences in the ages of the Older listeners, which were 47-66 years (mean = 57 years) for the present study, 65-77 years for the study of Grose and Mamo (2012), and a mean of 71 years for the study of He *et al*. (2008). The studies also varied in the extent to which they used stimulus manipulations to reduce the salience of excitation-pattern cues. The current study used interfering AM. Grose and

Mamo (2012) roved the carrier frequency over a small range but this might not have been very effective at reducing the use of excitation-pattern cues. He *et al*. (2008) did not attempt to reduce the role of excitation-pattern cues. The added AM in our study would be expected to produce a greater reliance on TFS cues for the 2-Hz rate, thereby increasing the effect of any age-related decline in the ability to use TFS cues. Therefore, the smaller effect of age found here compared to the studies of Grose and Mamo (2012) and He *et al*. (2008) probably reflects the fact that the older listeners were not as old as for the earlier studies.

The detrimental effect of age on AM detection was similar for the two modulation rates (see Fig. 2, top panel); the interaction between group and modulation rate was not significant. This is consistent with the results of Füllgrabe *et al*. (2015). In contrast, the detrimental effect of age on FM detection thresholds was greater for f_m = 2 Hz than for f_m = 20 Hz (see Fig. 2, lower panel); the interaction between group and modulation rate was significant. This is consistent with the idea that low-rate FM detection depends on the use of TFS cues, and that sensitivity to TFS declines with age. However, the detrimental effect of age on FM detection for the 2-Hz rate was modest, perhaps because the mean age of the Older group was only 57 yr. Also, there were large individual differences within the Older group. Figure 5 shows individual AM and FM detection thresholds averaged across *N* for each modulation rate (geometric mean). A detrimental effect of age on FM detection at f_m = 2 Hz occurred for 5 of the 14 Older listeners; the remainder had mean thresholds within the range found for the Young listeners. There was no significant correlation between age and (log-transformed) FM detection thresholds for f_m = 2 Hz for the Older group only (Pearson r = –0.42; p = 0.13). Hence, the poorest performers on the FM-detection task were not the oldest ones within the Older group. For the Older listeners only, there was no significant correlation between (log-transformed) FM detection thresholds at f_m = 2 Hz and absolute thresholds at 0.5 kHz (Pearson r = –0.12; p = 0.68) or mean absolute thresholds at 4 and 8 kHz (Pearson r = –0.07; p = 0.82). Further work is needed to understand the factors other than age that influence sensitivity to low-rate FM.

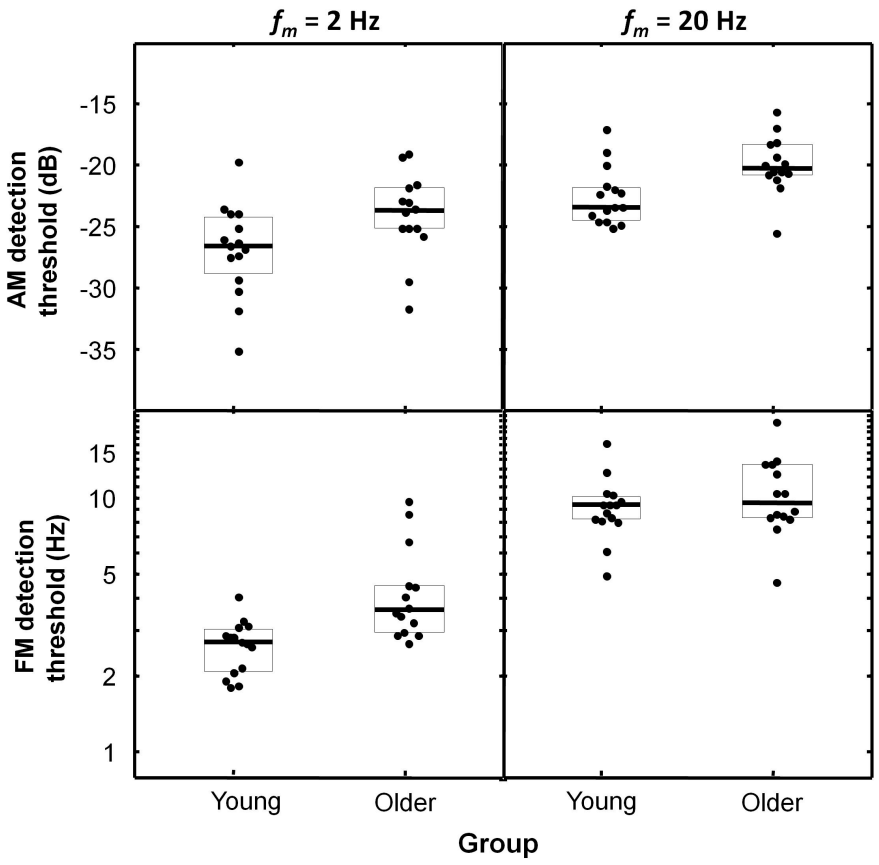


Figure 5. *The filled symbols show individual AM and FM detection thresholds averaged across N for each modulation rate. The bold lines show median values, and the lower and upper boundaries of the boxes show the first and third quartiles.*

Age did not affect temporal integration for AM and FM detection for the 2-Hz modulation rate (see Fig. 3, left panels) and temporal integration for the 20-Hz rate was actually slightly greater for the older than for the younger listeners (see Fig. 3, right panels). For FM at least, the greater temporal integration for the Older listeners occurred because they performed more poorly when the number of modulation cycles was small (N=2 and 3). These findings indicate that ageing spares the processes underlying temporal integration such as "multiple looks" (Viemeister and Wakefield, 1991) or a template-matching process (Hartmann and Klein, 1980; Dau *et al*., 1997) and support the notion that at least some aspects of "processing efficiency" do not decline with age. This is consistent with the finding that temporal integration for simple detection of pure tones in quiet does not vary significantly with age (Gehr and Sommers, 1999).

V. SUMMARY AND CONCLUSIONS

AM and FM detection thresholds were measured for a carrier frequency of 500 Hz and modulation rates of 2 and 20 Hz for Young and Older listeners with normal absolute thresholds below 3 kHz. FM detection thresholds were measured in the presence of uninformative AM in both intervals of a forced-choice trial, to disrupt the use of excitation-pattern cues. The number of modulation cycles, *N*, ranged from 2 to 9. The results show:

(1) For both groups and for each *N*, AM and FM detection thresholds were lower for the 2-Hz than for the 20-Hz rate.

(2) For both groups, AM and FM detection thresholds decreased with increasing *N*, this effect being greater for AM than FM.

(3) Thresholds were higher for older than for younger listeners, especially for FM detection at 2 Hz. This is interpreted as reflecting a detrimental effect of age on the use of temporal-fine-structure cues for low-rate FM detection.

(4) The effect of increasing *N* was similar across groups for both AM and FM for the 2-Hz rate. For the 20-Hz rate, the older listeners showed a slightly greater effect of increasing *N* than the younger listeners for both AM and FM. These findings suggest that ageing spares temporal integration of the cues used to detect AM and FM.

References

Ardoint, M., Lorenzi, C., Pressnitzer, D. and Gorea, A. (**2008**). "Perceptual constancy in the temporal envelope domain," J. Acoust. Soc. Am., **123**, 1591-1601.

Darwin, C. J., Turvey, M. T., and Crowder, R. G. (**1972**). "An auditory analogue of the Sperling partial report procedure : evidence for brief auditory storage," Cog. Psychol. **3**, 255-267.

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.

Demany, L., and Semal, C. (**1989**). "Detection thresholds for sinusoidal frequency modulation," J. Acoust. Soc. Am. **85**, 1295-1301.

Ernst, S. M. A., and Moore, B. C. J. (**2010**). "Mechanisms underlying the detection of frequency modulation," J. Acoust. Soc. Am. **128**, 3642-3648.

Ernst, S. M. A., and Moore, B. C. J. (**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.
Füllgrabe, C. (**2013**). "Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss," Am. J. Audiol. **22**, 313-315.
Füllgrabe, C., Moore, B. C. J, 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**, article 347, 1-25.
Gehr, S.E., and Sommers, M.S (**1999**). "The effects of age on temporal integration," J. Acoust. Soc. Am. **106**, 2208.
Grose, J. H., and Mamo, S. K. (**2010**). "Processing of temporal fine structure as a function of age," Ear Hear. **31**, 755-760..
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.
Hartmann, W. M., and Klein, M. A. (**1980**). "Theory of frequency modulation detection for low modulation frequencies," J. Acoust. Soc. Am. **67**, 935-946.
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.
Heinz M.G., Colburn H.S., and Carney L.H. (**2001**). "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve," Neur. Comput. **13**, 2273-2316.
Hopkins, K., and 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," J. Acoust. Soc. Am. **130**, 334-349.
Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," J. Acoust. Soc. Am. **68**, 1115-1122.
Kiang, N. Y. S. (**1965**). "Stimulus coding in the auditory nerve and cochlear nucleus," Acta Otolaryngol. **59**, 186-200.
King, A., Hopkins, K., and Plack, C. J. (**2014**). "The effects of age and hearing loss on interaural phase difference discrimination," J. Acoust. Soc. Am. **135**, 342-351.
Kohlrausch, A., Fassel, R., and Dau, T. (**2000**). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," J. Acoust. Soc. Am. **108**, 723-734.
Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467-477.
Lutman, M. E. (**1991**). "Degradations in frequency and temporal resolution with age and their impact on speech identification," Acta Otolaryngol. **111**, 120-126.
Moore, B.C.J., and Glasberg, B. R. (**1989**). "Mechanisms underlying the frequency discrimination of pulsed tones and the detection of frequency modulation," J. Acoust. Soc. Am. **86**, 1722-1732.
Moore, B.C.J., and Sek, A. (**1994)**. "Effects of carrier frequency and background noise on the detection of mixed modulation," J. Acoust. Soc. Am. **96**, 741–751.
Moore, B.C.J., and Sek, A. (**1995**). "Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation)," J. Acoust. Soc. Am. **97**, 2468-2478.
Moore, B.C.J., and Sek, A. (**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., 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 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.J., and Sęk, A. (**2009**). "Sensitivity of the human auditory system to temporal fine structure at high frequencies," J. Acoust. Soc. Am. **125**, 3186-3193.
Moore B.C.J, Vickers D.A., Mehta A. (**2012**). "The effects of age on temporal fine structure sensitivity in monaural and binaural conditions," Int. J Audiol. **51**, 715-721
Moore B.C.J, Glasberg B.R., Stoev M., Füllgrabe C., Hopkins K. (**2012**). 'The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies (L)," J. Acoust. Soc. Am. **131**, 1003-1006.
Moore, B.C.J. (**2014**). "Auditory Processing of Temporal Fine Structure: Effects of Age and Hearing Loss," (World Scientific, Singapore), pp. 1-182.
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.
Peters, R. W., and Moore, B.C.J. (**1992**). "Auditory filter shapes at low center frequencies in young and elderly hearing-impaired subjects," J. Acoust. Soc. Am. **91**, 256-266.
Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. **30**, 769-793.
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**, article 307, 1-14.
Sek, A., and Moore, B. C. J. (**1995**). "Frequency discrimination as a function of frequency, measured in several ways," J. Acoust. Soc. Am. **97**, 2479-2486.
Sheft, S., and Yost, W. A. (**1990**). "Temporal integration in amplitude modulation detection," J. Acoust. Soc. Am., **88**, 796-805.
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.
Verschooten, E., and Joris, P. X. (**2014**). "Estimation of neural phase locking from stimulus-evoked potentials," J. Assoc. Res. Otolaryngol., **15**, 767-787.
Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am., **66**, 1364-1380.
Viemeister, N. F., and Wakefield, G. H. (**1991**). "Temporal integration and multiple looks," J. Acoust. Soc. Am., **90**, 858-865.
Zwicker, E. (**1952**). "Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenz-modulation eines Tones (The limits of audibility of amplitude modulation and frequency modulation of a pure tone)," Acustica 2, 125-133.
Zwicker, E. (**1956**). ''Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs (The foundations for determining the information capacity of the auditory system),'' Acustica **6**, 356–381.
Zwislocki, J. J., and Nguyen, M. (**1999**). "Place code for pitch: A necessary revision," Acta Oto-laryngol., **119**, 140-145.

CHAPITRE 3

EFFETS DE LA PERTE AUDITIVE SUR LA SENSIBILITE ET L'INTEGRATION TEMPORELLE EN AM

9. Sensorineural hearing loss enhances auditory sensitivity and temporal integration for amplitude modulation

9.1 Résumé:

Amplitude-modulation (AM) detection thresholds (AMDTs) were measured at 40 dB sensation level for listeners with mild-to-moderate sensorineural hearing loss (age: 50-64 years) for a carrier frequency of 500 Hz and rates of 2 and 20 Hz. The number of modulation cycles, N, varied between 2 and 9. The data were compared with AMDTs measured for young and older normal-hearing listeners [Wallaert et al. (2016). J. Acoust. Soc. Am.]. As for normal-hearing listeners, AMDTs were lower for the 2-Hz than for the 20-Hz rate, and AMDTs decreased with increasing N. AMDTs were lower for hearing-impaired listeners than for normal-hearing listeners, and the effect of increasing N was greater for hearing-impaired listeners. A computational model based on the modulation-filterbank concept and a template-matching decision strategy was developed to account for the data. The psychophysical and simulation data suggest that the loss of amplitude compression in the impaired cochlea is mainly responsible for the enhanced sensitivity and temporal integration of temporal envelope cues found for hearing-impaired listeners. They also suggest that, for AM detection, cochlear damage is associated with increased internal noise, but preserved short-term memory and decision mechanisms.

9.2 Article:

L'article suivant a été publié dans *The Journal of the Acoustical Society of America* en février 2017.

Wallaert, N., Moore, B. C. J., 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. DOI: 10.1121/1.4976080

.

I. INTRODUCTION

Natural sounds, including speech and environmental sounds, convey salient amplitude-modulation (AM) cues that play a crucial role in their identification (*e.g.*, Shannon *et al.*, 1995; Gygi *et al.*, 2004). Many studies suggest that age and cochlear damage differentially affect the auditory processing of these AM cues (see Moore, 2007; Füllgrabe *et al.*, 2015). These effects are reviewed below. However, further work is needed to tease out the respective contributions of sensory and non-sensory (*e.g.*, memory, decision) factors to the effects of age and cochlear damage. For instance, recent work suggested that the decline in auditory sensitivity to AM associated with aging may result from reduced "processing efficiency", that is a reduced ability of the central auditory system to make use of available sensory information (Füllgrabe *et al.*, 2015). Changes in AM sensitivity associated with hearing loss may also result - at least partly - from reduced processing efficiency, especially when they are associated with aging, as in the case of presbycusis. The present study was designed to clarify the relative contributions of these sensory and non-sensory factors to AM detection for older hearing-impaired listeners by combining a psychophysical and modeling approach.

In a recent study (Wallaert *et al.*, 2016), AM detection thresholds (AMDTs) were measured at 40 dB sensation level (SL) for young (22-28 years) and older (44-66 years) listeners with normal audiograms, using a carrier frequency of 500 Hz and modulation rates of 2 and 20 Hz. The number of modulation cycles, *N*, varied between 2 and 9. For both groups, AMDTs were lower for the 2-Hz than for the 20-Hz rate, and AMDTs decreased with increasing *N*. AMDTs were higher for the older than for the younger normal-hearing (NH) listeners for both AM rates. A similar effect was found by Füllgrabe *et al.* (2015), but some studies have found no clear effect of age (Schoof and Rosen, 2014; Paraouty *et al.*, 2016; Paraouty and Lorenzi, 2016). The origin of the discrepancy across studies is not clear. It may partly reflect the fact that in several of these studies, the older listeners had higher absolute thresholds than the young listeners, even though both young and older listeners had thresholds within the normal range. Higher absolute thresholds may indicate a loss of cochlear compression, which can lead to better

AM detection and offset the possible detrimental effect of age (see below for further discussion of the effects of cochlear compression). When the absolute thresholds of the older and younger groups are careful matched, AM detection appears to be slightly worse for older than for young listeners (Füllgrabe *et al*., 2015). In the study of Wallaert *et al*. (2016), the effect of increasing *N* on AMDTs was similar across groups. This suggest that aging reduces sensitivity to AM cues, while sparing temporal integration of these cues.

The goal of the present study was to extend this investigation of AM sensitivity and temporal integration of AM cues to older listeners with sensorineural hearing loss. There are many studies of the effects of cochlear damage on AM sensitivity. These studies indicate that at comparable low SLs, hearing-impaired (HI) listeners show better AMDTs than NH listeners (Bacon and Gleitman, 1992; Moore *et al*., 1992, 1996; Füllgrabe *et al*., 2003). However, when compared at the same SPL, AMDTs are not consistently better for HI than for NH listeners (Grose *et al*., 2016; Schlittenlacher and Moore, 2016; Sek *et al*., 2015). The beneficial effect of hearing loss for stimuli presented at equal SLs was as large as 8 dB for low AM rates (Füllgrabe *et al*., 2003). These findings have 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) (Moore *et al*., 1992, 1996) 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). To the best of our knowledge, the effects of cochlear damage on the temporal integration of AM cues have not been investigated. Knowledge of these effects should help to clarify whether there are any effects of peripheral dysfunction on more central processes, since temporal integration of AM cues probably reflects central (*i.e.*, decision) processes involving "multiple looks" (Viemeister and Wakefield, 1991; Sheft and Yost, 1990) or a template-matching process (Dau *et al*., 1997). However, peripheral dysfunction may result in increased neural noise which could have consequences for many levels of processing. This point is addressed in our modeling.

In the present study, AMDTs were measured at 40 dB SL for listeners with mild-to-

moderate sensorineural hearing loss, aged between 50 and 64 years, using a 500-Hz sinusoidal carrier and modulation rates of 2 and 20 Hz. The number of modulation cycles, *N*, varied between 2 and 9. The data were compared with AMDTs measured for young and older NH listeners by Wallaert *et al*. (2016). We chose to compare performance across groups using a fixed SL in order to equate the audibility of the stimuli. However, this necessarily meant using a higher sound pressure level (SPL) for the HI than for the NH listeners. Since AM detection for NH listeners improves with increasing SPL (e.g., Kohlrausch et al., 2000), differences across groups might be partly a consequence of differences in SPL. This should be borne in mind when interpreting the results.

A popular model of AM processing is based on the assumption that AM detection is achieved by cross-correlating the outputs of AM channels with memory-stored templates according to an "optimal detector" scheme (Dau *et al.*, 1997a,b). Central limitations to AM processing are implemented as additive internal noise. As a consequence of nearly logarithmic compression of steady components by the adaptation stage of the model, the additive noise acts as multiplicative noise in the intensity domain, enabling the model to predict Weber's law for intensity discrimination. In a more recent version of this model (Ewert and Dau, 2004), the outputs of the AM channels were also logarithmically transformed prior to the addition of internal noise, to account for the Weber's law behaviour found for AM-depth discrimination (Ewert and Dau, 2004; Schlittenlacher and Moore, 2016). Effectively, this model version mimicked the effects of multiplicative internal noise in the AM domain. This model accounted successfully for a variety of AM detection data obtained using young NH listeners, including temporal integration of AM (Dau *et al.*, 1997b). Importantly, the temporal integration effects found for AM detection were interpreted as reflecting the properties of the correlation operation at the decision stage of the model (Dau *et al.*, 1997b). However, some form of *memory* process might be required to account for the effects of AM rate and stimulus duration on temporal integration for NH and HI listeners. This is suggested by a previous attempt to model auditory discrimination of complex AM patterns (Ardoint *et al*., 2008). In this work, model performance

was limited by decay of the memory trace of temporal-envelope cues. In the modeling part of the present study, we used a simplified front-end to the optimal detector approach to investigate whether performance could be predicted using a model that reproduces the main features of temporal-envelope processing implemented by low-level sensory mechanisms (cochlear filtering, instantaneous amplitude compression, adaptation), mid-level processes (bandpass AM filtering), and higher-level non-sensory processes (internal noises, memory decay in the AM domain, and template matching).

II. METHOD

A. Listeners

Eleven older hearing-impaired (HIo) listeners aged between 50 and 64 years (mean = 59 years; standard deviation, SD = 4 years) participated. Audiograms for the tested (right) ears of all HIo listeners are shown as black circles in Fig. 1. Sensorineural hearing loss was diagnosed based on air-conduction, bone-conduction and impedance audiometry. HIo listeners had audiometric thresholds ≤60 dB HL between 0.25 and 4 kHz. The audiometric threshold at the tested frequency (0.5 kHz) ranged between 30 and 50 dB HL (mean = 40 dB HL; SD = 6 dB), indicating that the hearing loss was mild to moderate (Goodman, 1965).

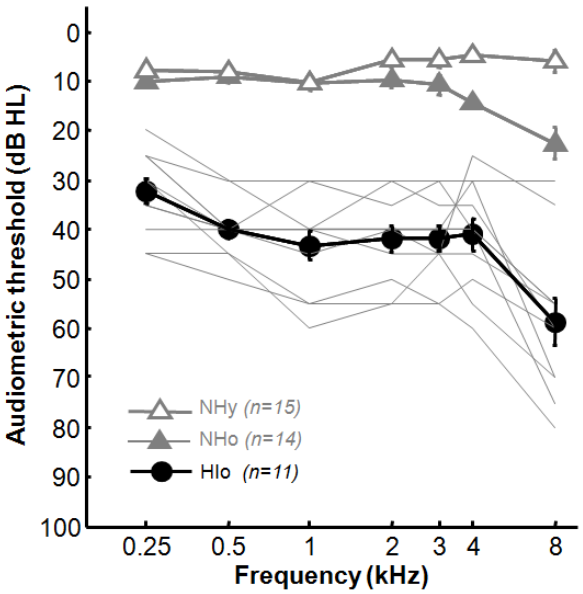


Figure 1. *Thin lines show individual audiometric thresholds and filled circles show mean audiometric thresholds for the older hearing-impaired (HIo) listeners for the right ear. The mean data for the young normal-hearing (NHy, open triangles) and older normal-hearing (NHo, filled gray triangles) listeners are replotted from Wallaert et al. (2016). Error bars indicate ±1 standard error of the mean. They are omitted when they would be smaller than the size of the symbol used to plot the data.*

The data for 15 young NH listeners (NHy), aged between 22 and 28 years (mean = 24 years; SD = 2 years), and 14 older listeners (NHo), aged between 47 and 66 years (mean = 57 years; SD = 8 years) collected by Wallaert *et al*. (2016) were compared with the present data. All NHy and NHo listeners had audiometric thresholds less than 20 dB HL for the tested (right) ear for octave frequencies between 0.25 and 3 kHz, but some of the NHo listeners had elevated audiometric thresholds for frequencies above 3 kHz. The mean audiometric threshold at 0.5 kHz was 9 dB HL (SD = 5 dB) for the NHo group and 8 dB HL (SD=3 dB) for the NHy group.

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 and N° IRB : 20153600001072).

B. Procedure

AMDTs were measured using a two-interval, two-alternative, forced-choice, adaptive procedure to estimate the 70.7% point on the psychometric function (Levitt, 1971). Each trial contained a target (modulated) and a standard (unmodulated) stimulus, in random order, with a silent interval of 600 ms between them. The root-mean-square level of the two stimuli was, on average, the same. However, to discourage the use of level cues, the levels of the standard and target stimuli were roved independently over a range of ±1.5 dB. The duration of each signal was determined by the number of modulation cycles, N, and the modulation rate, f_m. N was 2, 3, 4, 5, or 9 cycles.

Each interval was marked by a light on the computer screen. Listeners were asked to indicate which interval contained the fluctuating sound. Correct answer feedback was provided.

The tracking variable was the AM depth, m (see Eq. 1, below). A run started with the tracking variable well above the estimated detection threshold. The step size was a factor of 1.58 until two reversals had occurred, and 1.26 thereafter. Fourteen reversals were obtained for each run, and the threshold estimate for that run was taken as the geometric mean of the value of the

tracking variable at the last six reversals. Two threshold estimates were obtained for each condition. The final estimate of the threshold was taken as the geometric mean of the two threshold estimates.

C. Stimuli

The stimuli were identical to those used by Wallaert *et al*. (2016). All stimuli were generated digitally at a sampling rate of 44.1 kHz. They were presented using Sennheiser HD 280 pro headphones (Old Lyme, CT) and an external soundcard (ECHO Indigo DJ 1-2, Santa Barbara, California; 16-bit resolution). The presentation level was 40 dB SL except for HIo listeners with absolute thresholds of 40 dB HL or more, for whom the level was limited to 80 dB SPL. Levels were calibrated using a Bruel & Kjaer (Nærum, Denmark) type 2250 sound level meter and type 4153 ear simulator. Stimuli were presented only to the right ear.

The carrier was a sinusoid with frequency f_c= 500 Hz. This carrier was either unmodulated to produce the standard stimulus, or modulated sinusoidally in amplitude to produce the target stimulus. Eq. 1 describes the target stimulus T(t):

$$T(t) = [1 + m \sin(2\pi f_m t + 3\pi/2)] \sin(2\pi f_c t + \varphi), \quad [1]$$

where f_m is the modulation rate (2 or 20 Hz), φ is the starting phase of the carrier, and t is time (expressed in s). The value of φ was randomly chosen for each stimulus. The starting phase of the modulator, 3π/2 radians, resulted in a modulation minimum at the onset and offset of the stimulus. Raised cosine ramps lasting 250 and 25 ms were applied after the AM for the modulation rates of 2 and 20 Hz, respectively.

III. RESULTS

The right-most panel in Fig. 2 shows the mean AMDTs for the HIo listeners, plotted as a function of N (black lines and symbols). The mean thresholds for the NHy listeners (left panel) and NHo listeners (middle panel), as measured by Wallaert *et al.* (2016), are plotted for comparison. In each panel, the filled circles and open triangles show the detection thresholds for

f_m = 2 Hz and for f_m = 20 Hz, respectively. AMDTs are expressed as 20 $\log_{10} m$.

The mean AMDT for f_m = 2 Hz and N = 5 for the HIo listeners was identical to that reported by Füllgrabe *et al*. (2003) for a similar condition (f_m = 4 Hz; *fc* = 2000 Hz; duration = 2 s; presentation level = 40 dB SL) for listeners aged between 53 and 77 years with mild-to-moderate sensorineural hearing loss.

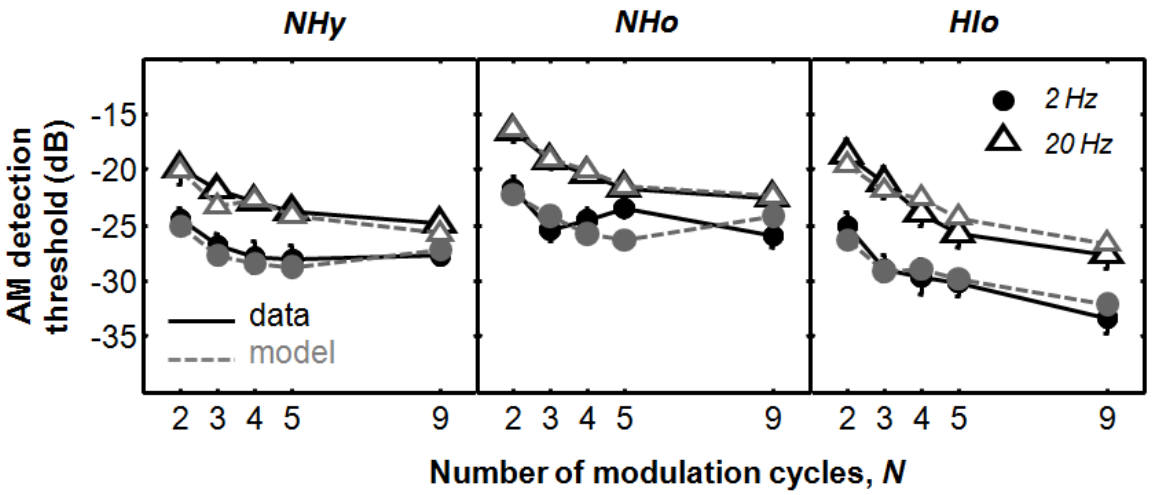


Figure 2. *Mean AM detection thresholds (AMDTs, black symbols and continuous lines) for the HIo group plotted as a function of N (right panel). The mean AMDTs for the NHy listeners (left panel) and NHo listeners (middle panel) are taken from Wallaert et al. (2016). Simulated AMDTs (grey symbols and dashed lines) are shown for the models for NHy, NHo and HIo listeners. The filled circles and open triangles show the thresholds for fm = 2 Hz and for fm = 20 Hz, respectively.*

Statistical analyses were restricted to comparison of the age-matched listeners: NHo and HIo. A repeated-measures analysis of variance (ANOVA) was conducted on the AMDTs with between-subjects factor group (2 levels: NHo and HIo) and within-subjects factors N (5 levels) and modulation rate (2 levels). The effects of each factor are described below.

A. Effects of f_m

The effect of modulation rate was significant [$F(1, 23) = 42.5$, $p < 0.001$]. For both groups, AMDTs were lower for f_m = 2 Hz than for f_m = 20 Hz. For the NHo listeners, the difference between AMDTs at f_m = 2 and 20 Hz was 5-6 dB for $N < 5$, and 2-3 dB for N = 5 or 9. For the HIo

listeners, the corresponding difference ranged between 4 and 7 dB, and was roughly independent of *N*.

B. Effects of *N*

The effect of *N* was significant [$F(4, 92) = 52.9$, $p < 0.001$]. For each modulation rate and each group, AMDTs decreased with increasing *N.* The decrease was greater for the HIo group for both modulation rates. For the NHo group, the decrease was 3-6 dB when *N* was increased from 2 to 9 cycles, but most of the effect of *N* occurred for $N \leq 5$. The HIo group showed a more monotonic and larger decrease of 8-9 dB.

There was a significant interaction between *N* and f_m [$F(4, 92) = 4.3$, $p = 0.003$]. This reflects the fact that AMDTs decreased more with increasing *N* for f_m = 20 Hz than for f_m = 2 Hz. For f_m = 2 Hz, AMDTs decreased by 6.2 dB (mean for the two groups; 3.8 dB for the NHo group; 8.6 dB for the HIo group) when *N* increased from 2 to 9. In comparison, for f_m = 20 Hz, AMDTs decreased by 7.4 dB (5.6 dB for the NHo group; 9.1 dB for the HIo group) when *N* increased from 2 to 9.

A linear regression analysis was conducted on the individual AMDTs (20log*m* versus log(*N*)). A repeated-measures ANOVA was conducted on the resulting slopes with between-subject factor group and within-subjects factor f_m. There was a significant effect of f_m [$F(1, 23) = 7.4$, $p = 0.012$], a significant effect of group [$F(1, 23) = 13.4$, $p = 0.001$] and no significant interaction [$F(1,23) = 1.1$, $p = 0.304$]. These analyses indicate that temporal integration for AM detection was: i) greater for the HIo than for the NHo group; ii) greater at 20 Hz than at 2 Hz for both groups.

C. Effects of group

The effect of group was significant [$F(1, 23) = 16.5$, $p < 0.001$]. For all conditions, the AMDTs were lower for the HIo than for the NHo group. In addition, there was a significant interaction between *N* and group [$F(4, 92) = 5.4$, $p < 0.001$]. The higher the value of *N*, the larger

was the difference in AMDTs between the NHo and HIo groups, suggesting that temporal integration was greater for the HIo group. When N = 9, the difference in AMDTs between the NHo and HIo groups reached 8.2 and 5.6 dB for f_m = 2 and f_m =20 Hz, respectively. None of the other interactions involving group was significant, indicating that the pattern of results for AM detection for the two modulation rates did not differ significantly for the two groups, although the NHo group performed more poorly overall. None of the other interactions was significant.

IV. MODELING AM PROCESSING

A. Rationale

The present data combined with those reported by Wallaert *et al*. (2016) show that aging *per se* degrades AM sensitivity while sparing temporal integration of AM cues, whereas hearing loss enhances both AM sensitivity and temporal integration of AM cues. It is unclear, however, whether the observed deficit in AM sensitivity associated with aging reflects a pure sensory deficit or, as suggested by Füllgrabe *et al*. (2015), a decrease of efficiency in processing temporal-envelope information. Assuming that such a reduction in processing efficiency does occur, it is unclear whether it results from an increase in internal noise (*i.e.*, noise occurring after envelope extraction), degraded memory and decision capacities, or some combination of these factors. It is also unclear whether the enhanced AM sensitivity and temporal integration of AM cues associated with hearing loss reflect only peripheral changes such as reduced frequency selectivity and loss of fast-acting compression, or whether hearing loss is also associated with reduced processing efficiency. Again, if this is the case, it is unclear whether reduced processing efficiency for HIo listeners is a consequence of an increase in internal noise, degraded memory and decision capacities, or some combination of these factors. It is also unclear whether aging and hearing loss affect the same sources of internal noise (additive and/or multiplicative noises). To address these issues, performance was compared with the predictions of an envelope-processing model that included central limitations, such as different sources of internal noise, and a memory decay. The model is described in the next section.

B. The model

A schematic diagram of the model is shown in Fig. 3. The first stage was a linear Gammatone filter that simulates the bandpass filtering of the basilar membrane in the cochlea (Patterson *et al.*, 1995). The center frequency was set to 583 Hz. This corresponds to the center frequency of an auditory filter centered 1 ERB_N (Glasberg and Moore, 1990) above 500 Hz. This filter was chosen to simulate off-frequency listening, consistent with the view that AM is perceived by monitoring fluctuations in excitation level on the high-frequency side of the excitation pattern (Zwicker, 1952, 1956; Moore and Sek, 1994).

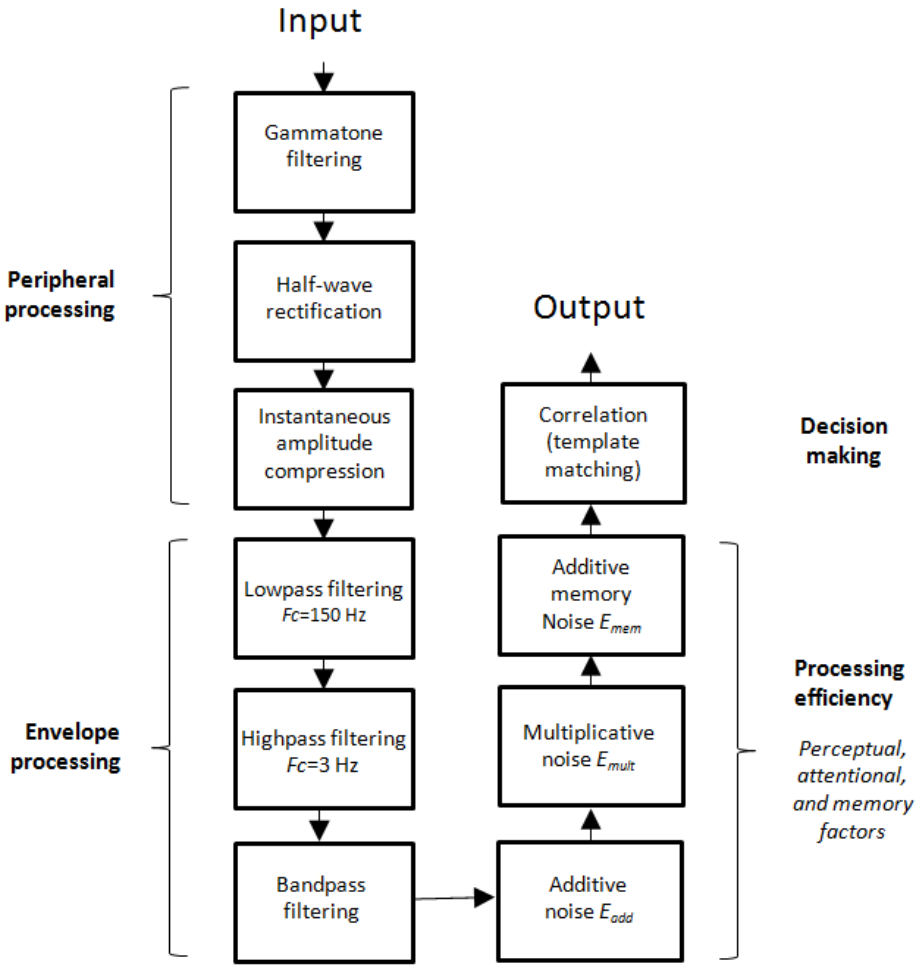


Figure 3. *Schematic of the model of AM processing.*

However, because the model employs a linear Gammatone filter (with compression applied in a subsequent stage), the exact choice of center frequency is somewhat arbitrary and is

not critical for determining the outcome. In the second stage, the output of the gammatone filter was half-wave rectified, subjected to instantaneous amplitude compression by raising the amplitude to the power C ($C \leq 1$), and lowpass filtered (cutoff = 150 Hz, rolloff = 6 dB/oct). The compressed envelope obtained in this way was then highpass filtered (rolloff = 6 dB/oct) to simulate the effects of adaptation in the AM domain, as done by Tchorz and Kollmeier (1999; see their Fig 2). The cutoff frequency of the highpass filter was an adjustable parameter. The resulting signal was passed through a bandpass filter to simulate frequency-selective processing in the AM domain (Dau *et al*., 1997a,b). Although Dau *et al*. (1997a,b) assumed that the modulation filters had a Q value of about 2, more recent work suggests that the Q value is closer to 1 (Ewert and Dau, 2000; Ewert *et al*., 2002; Joosten *et al.*, 2016; Lorenzi *et al*., 2001; Sek and Moore, 2002), and a value of 1 was assumed here (rolloff = 6 dB/oct). For each simulation, the center frequency of this modulation bandpass filter was set to the target AM rate (that is, 2 or 20 Hz).

The envelope at the output of the modulation filter was then "degraded" by three sources of internal noise. These internal noises were introduced to limit the performance of the model when processing AM. They were applied in the following order: i) firstly, an additive noise, E_{add}, as in Dau *et al*. (1997a,b), ii) then, a multiplicative noise, E_{mult}, as in Ewert and Dau (2004), and finally iii) an additive time-varying memory noise, characterized by parameter E_{mem}, which has not to our knowledge been used previously to model AM detection. All internal noises were modeled as Gaussian noises with zero mean, and added independently to each sample point of the output of the bandpass AM filter. For each type of internal noise, the magnitude was specified by its SD (expressed in arbitrary units (a.u.) for additive noises, and in dB for the multiplicative noise). E_{add} was introduced to limit intensity discrimination and AM sensitivity (Dau *et al*., 1997a,b). E_{mult} was introduced to simulate the Weber's law behaviour for AM-depth discrimination (Ewert and Dau, 2004; Schlittenlacher and Moore, 2016). The SD of the multiplicative noise was proportional to the root-mean-square value of the output of the AM filter (once combined with E_{add}). E_{mem} was introduced to simulate imperfect retention of

temporal-envelope information in each observation interval, especially for very long stimuli (*e.g.*, for 4.5-s long stimuli, as used when N = 9 and f_m = 2 Hz). Note that the approach used here to model memory noise implicitly assumes that short-term memory processes are independent from sensory processes (*e.g.*, temporal-envelope extraction). This is consistent with the view (common in vision research) that "memory processes are distinct from on-line perceptual analysis" (Magnussen, 2000; see also, Demany *et al*., 2003). A comparable approach was taken by Ardoint *et al*. (2008) to model discrimination of complex AM patterns. In the present work, the SD of the memory noise increased backward in time according to an exponential function with a 1.2-s half life. The parameter E_{mem} corresponds to maximum value of the SD of the memory noise. The resulting (noisy) output of the modulation filter was used as input to the decision stage of the model.

The decision stage of the model was realized as a simplified version of the optimal detector described by Dau *et al*. (1997a,b). In each trial, the output of the model was computed for the two stimuli in that trial. A “template” was derived at the beginning of each simulated threshold measurement (that is, at the beginning of each adaptive staircase) when the target AM depth, m, was 1. The template was generated as the difference between the model output in response to the target and standard stimuli. This was done with all three internal noise sources set to zero. The template was then correlated (using the Pearson correlation coefficient) with the model output for each interval in a trial. The interval leading to the largest correlation was taken as the target interval identified by the model for that trial. Stimuli were generated as for the behavioural experiment, except that the presentation level was fixed for all simulations. Thirty thresholds were simulated and averaged for each condition.

The values of some of the parameters were chosen based on previous data. The bandwidth of the Gammatone filter was set to 1 ERB_N (Glasberg and Moore, 1990) to model the data for the NHy and NHo groups, and to 2 ERB_N to model the data for the HIo group, simulating the reduced frequency selectivity associated with moderate sensorineural hearing loss (see Moore, 2007). The compression parameter C was set to 0.6 to model the data for the NHy and

NHo groups, simulating "normal" cochlear compression, and to 1.0 to model the data for the HIo group, simulating the loss of compression associated with cochlear hearing loss. The value of E_{mult} was set to 1 dB, as in Ewert and Dau (2004), so as to predict the threshold for detecting an increase in AM depth. The values of the other two noise parameters were chosen so as to fit the data for each group as well as possible. The values chosen are described below. The cutoff frequency of the highpass filter used to model adaptation was chosen so as to give a correct prediction of the relative values of the AMDTs for the two modulation rates. The effects of varying each of the adjustable parameters are described in more detail below.

C. Simulation results

Figure 2 compares the best-fitting model simulations (grey symbols and dashed lines) with the empirical data (black symbols and continuous lines) for the NHy, NHo and HIo groups. For each set of simulations the prediction error (averaged across *N*) was less than 1 dB. The cutoff frequency of the highpass filter used to simulate the effects of adaptation was 3 Hz for all three groups. E_{add} was set to 1.3, 1.7, and 6 arbitrary units (a.u.) to fit the data for the NHy, NHo and HIo groups, respectively, and E_{mem} was 0.9 a.u. for all three groups. Note that the absolute values of E_{add} and E_{mem} are not meaningful since the values required to fit the data would depend on the scaling of the signal magnitudes and the sampling rate used. However, the relative magnitudes for the different groups are meaningful. The effects of varying each adjustable parameter are illustrated in Fig. 4, and are discussed later in this section. However, first, we discuss the main aspects of the data and simulations shown in Fig. 2.

The model simulations in the left panel nicely captured the effects of AM rate and *N* on the AMDTs for the NHy listeners. The simulated thresholds improved by 2.2 and 5.7 dB at f_m = 2 and 20 Hz, respectively, when *N* was increased from 2 to 9. In comparison, the measured thresholds for the NHy listeners improved by 2.4 and 4.8 dB for f_m = 2 and 20 Hz, respectively, when *N* was increased from 2 to 9.

The model simulations in the middle panel of Fig. 2 capture the detrimental effect of age reasonably well, although the fit to the data was better for f_m = 20 than for f_m = 2 Hz. The increase in AMDTs observed for the NHo listeners when *N* was increased from 3 to 5 was not predicted by the model and may reflect errors of measurement; generally, one would expect the AMDTs to improve or remain constant with increasing *N*, rather than to worsen. When *N* was increased from 2 to 9, the simulated AMDTs improved by 2 and 6 dB for f_m = 2 and 20 Hz, respectively. In comparison, the measured AMDTs improved by 4.2 and 6 dB for f_m = 2 and 20 Hz, respectively.

The model simulations shown in the right panel of Fig. 2 capture the main trends in the data for the HIo listeners, namely the beneficial effects of hearing loss on AM sensitivity and temporal integration of AM information, although the simulated improvement in temporal integration was somewhat lower than observed. The simulated AMDTs improved by 6 and 7 dB for f_m = 2 and 20 Hz, respectively, when *N* was increased from 2 to 9. In comparison, the obtained AMDTs improved by 8.6 and 9 dB for f_m = 2 and 20 Hz, respectively.

Figure 4 illustrates the effect of varying the parameters of various stages of the model on predictions of the data for the NHy listeners. The top-left panel of Fig. 4 shows the effect of doubling the bandwidth of the simulated auditory filter from 1 ERB_N to 2 ERB_N, simulating the reduced frequency selectivity associated with cochlear hearing loss. This led to very slightly improved AMDTs, probably because increasing the bandwidth of the filter resulted in a slightly increased level of the signal relative to the internal noises used in the model. Recall that the filter was centered slightly above the carrier frequency of the signal.

The top-middle panel of Fig. 4 shows the effect of changing the adaptation stage of the model. When the cutoff frequency of the highpass filter was increased from 3 to 6 Hz, this had the effect of markedly elevating AMDTs for the 2-Hz AM rate and *N* = 9, giving predictions that were clearly inconsistent with the data. When the adaptation stage was removed ("off" condition), this led to predicted AMDTs for the 2-Hz AM rate that were markedly better than the obtained values for all values of *N*. Manipulating the characteristics of the adaptation stage had no effect on the predicted AMDTs for the 20-Hz AM rate. Overall, it seems clear that the

characteristics of the adaptation stage are important for predicting the relative values of the AMDTs for the 2- and 20-Hz AM rates, and for predicting the variation of AMDTs with *N* for the 2-Hz rate. The highpass filter cutoff frequency of 3 Hz gave good predictions of these aspects of the data for all three groups, suggesting that adaptation did not differ markedly across the groups.

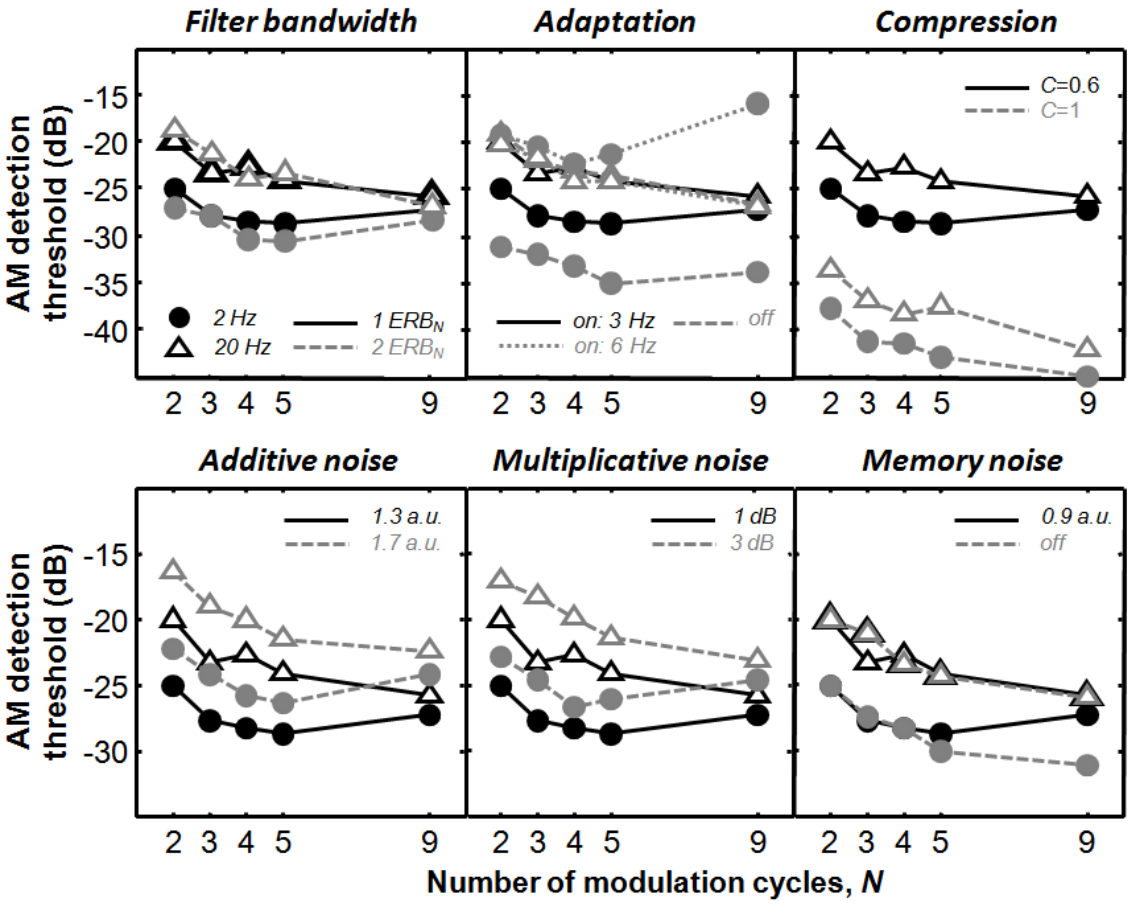


Figure 4. *Effects of modifying parameters of the model used to fit the data in Fig. 2 (for NHy listeners). Black continuous lines indicate predictions using the same parameters values as for Fig. 2. Grey dashed lines indicate predictions using the modified parameters values. The modifications were : (i) increasing the peripheral auditory filter bandwidth by a factor 2 (top-left panel); (ii) increasing the cutoff frequency of the highpass filter simulating adaptation from 3 to 6 Hz, or removing the highpass filter ("off") (top-middle panel); (iii) removing the compression (C = 1; top-right panel); (iv) changing the amount of additive noise from 1.3 to 1.7 a.u. (E_{add}; bottom-left panel); (v) changing the amount of multiplicative noise from 1 to 3 dB (E_{mult}; bottom-middle panel); and (vi) removing the memory noise ("off"; E_{mem}; bottom-right panel). Different y-axis ranges were used for the top and bottom panels, to improve visibility. See text for details.*

The top-right panel of Fig. 4 shows the effects of removing amplitude compression (*C* = 1.0) on the predicted AMDTs, simulating the loss of compression associated with cochlear

hearing loss. As expected, removing compression dramatically improved the predicted thresholds for both AM rates. This was a consequence of the expansion of temporal-envelope cues at the output of the Gammatone filter and the subsequent improvement in the internal signal-to-noise ratio; the representation of the envelope became larger relative to the two sources of additive noise, E_{add} and E_{mem}. The effect of removing compression was greatest for f_m = 2 Hz and N = 9, the condition for which the additive memory noise E_{mem} had its most important limiting effect. The predicted beneficial effect of loss of compression was actually much larger than observed in the data for the HIo group. That is why the value of E_{add} for that group was increased to 6 a.u., in order to predict overall values of the AMDTs that were consistent with the data.

The bottom-left panel of Fig. 4 shows the effects of increasing the additive noise (E_{add}) from 1.3 to 1.7 a.u. This led to an increase in predicted AMDTs for both AM rates and all N, but the increase was especially marked for the AM rate of 2 Hz and N = 9. The predicted AMDT for the latter condition was higher than the predicted AMDT for the AM rate of 2 Hz and N = 5, which is not consistent with the experimental data. Hence, the magnitude of the additive noise E_{add} is important for characterizing the variation of AMDTs with N for the AM rate of 2 Hz.

The bottom-middle panel of Fig. 4 shows the effects of increasing the multiplicative noise (E_{mult}) from 1 to 3 dB. To a first approximation, the effects resemble those produced by increasing the additive noise E_{add}, making it difficult to disentangle the effects of the two types of noise. Thus, the model does not make it possible to determine whether the general decrease in AM sensitivity associated with aging is caused by an increase in additive or multiplicative noise. However, the two noises do have distinct effects when the compression parameter is modified, since decreasing compression results in an increase of the effect of the multiplicative noise but not the additive noise. Hence, the relative magnitudes of the additive and multiplicative noises are important for predicting the results for the HIo listeners.

The bottom-right panel of Fig. 4 shows the effects of removing the memory noise (E_{mem}). This had little effect for the 20-Hz AM rate, which is not surprising given the short duration of

the stimuli even for N = 9. Removing the memory noise led to a marked improvement in predicted AMDTs for the 2-Hz AM rate and N = 9, which was not found in the experimental data. Hence, the memory noise appears important in accounting for the finding that for the 2-Hz AM rate performance did not improve as N was increased from 5 to 9.

Returning to the effects of compression, the additive internal noise E_{add} had to be increased substantially (by a factor of ~3) in the model for the HIo group (*i.e.*, to 6 a.u., compared to 1.7 a.u. in the model for the NHo group) to counterbalance the beneficial effect of loss of cochlear compression and achieve predicted AMDTs close to the obtained values. Accurate predictions could not be obtained by increasing the multiplicative noise E_{mult}. Hence it appears that hearing loss combined with age is associated with an increase in additive noise.

IV. DISCUSSION

A. Psychophysical results

For each value of N and for both AM rates, AMDTs for the HIo listeners were significantly lower (*i.e.*, better) than those for the NHo listeners tested at the same SL. The interaction between group and modulation rate was not significant. This beneficial effect of hearing loss on AM sensitivity is consistent with previous results obtained for older listeners with sensorineural hearing loss (*e.g.*, Bacon and Gleitman, 1992; Füllgrabe *et al*., 2003; Moore *et al*., 1992; Moore and Glasberg, 2001; Schlittenlacher and Moore, 2016) when the NH and HI groups were compared at equal SLs.

The beneficial effect of sensorineural hearing loss on AM sensitivity is often interpreted as resulting from the loss of cochlear compression and the associated loudness recruitment (*e.g.*, Moore *et al*., 1996), although animal studies suggest that abnormal transduction and short-term adaptation may also play a role (Kale and Heinz, 2010; Scheidt *et al*., 2010). Cochlear damage, and more specifically damage to outer hair cells, leads to loss of the fast-acting compression operating in the normal cochlea, yielding an abnormally rapid growth in loudness as sound level is increased (Moore, 2007). Because the amplitude fluctuations of incoming sounds are no

longer compressed, loudness recruitment has the effect of perceptually "magnifying" these amplitude fluctuations (Moore *et al*., 1996) and this may explain why improved AM sensitivity is often observed for hearing-impaired listeners when compared at the same SL.

In the present study, the interaction between *N* and group was significant, indicating that the ability to integrate AM information increases with sensorineural hearing loss. Indeed, for both rates, the higher the value of *N*, the larger was the difference in AMDTs between the NHo and HIo listeners. This difference reached 8-9 dB for $N = 9$. This indicates that sensorineural hearing loss does not impair and may even improve the central processes underlying temporal integration, such as the combination of "multiple looks" (Viemeister and Wakefield, 1991) or a template-matching process (Dau *et al*., 1997). This is consistent with the notion that at least some aspects of "processing efficiency" related to the short-term storage of AM information and decision making do not decline with sensorineural damage. This appears to contrast with the results for temporal integration for pure-tone detection which show reduced temporal integration for HI listeners (e.g., Florentine *et al*., 1988). Oxenham *et al*. (1997) proposed that this reduced temporal integration could be a consequence of reduced peripheral compression, although this interpretation has been questioned by Plack and Skeels (2007). At present, the origin of the reduced temporal integration for pure-tone detection for HI listeners remains unclear.

B. . Modeling results

The modeling study showed that it was possible to account for the data for NHy listeners using assumptions that formed the basis of previous models of envelope processing (Dau *et al.*, 1997a,b; Ardoint *et al*., 2008). Firstly, some form of high-pass envelope filtering, simulating short-term adaptation, is required to account for the effect of AM rate. Secondly, template matching can account for some aspects of the temporal integration of AM information. Thirdly, short-term memory limitations are required to account for the limited temporal integration that occurs for low (2-Hz) AM rates, when the stimulus duration exceeds 1-2 s.

The modeling study also showed that the data for NHo listeners could be accounted for without any change to the peripheral stages of the model (cochlear filtering, amplitude compression, and adaptation) or to the temporal-envelope analysis stage (AM bandpass filtering). This is consistent with the view that processing of low-rate AM is affected only slightly if at all by aging (Takahashi and Bacon, 1992; Füllgrabe *et al*., 2015). Finally, the data for NHo listeners could be accounted for without any change to the template-matching mechanism and memory limitations. The psychophysical and modeling data therefore suggest that aging spares short-term memory and decision processes. This is consistent with previous work showing that short-term memory (the temporary storage of information) is almost unaffected by age (*e.g*., Füllgrabe *et al*., 2015). However, the additive and/or multiplicative internal noise had to be slightly increased to account for the general decline in AM sensitivity for NHo listeners. This suggests that aging degrades at least some aspects of processing efficiency. The exact nature of the reduced processing efficiency needs to be clarified in future work. The increase in internal noise may be caused by envelope under-sampling resulting from a reduction in the number of auditory-nerve fibers (Kujawa and Liberman, 2009; Sergeyenko *et al*., 2013) and/or from greater neural noise at stages following the auditory nerve. It was not possible to determine from our data whether the general decrease in AM sensitivity associated with aging is caused by an increase in additive or multiplicative noise. However, if aging results in increased multiplicative noise, then poorer-than-normal AM-depth discrimination performance should be observed for NHo listeners. Schlittenlacher and Moore (2016) did not find such an effect, suggesting that multiplicative noise does not increase with increasing age.

Finally, the modeling study showed that the main trends in the data for HIo listeners could be reproduced if amplitude compression in off-frequency channels was reduced or abolished. There was no need to modify the adaptation, envelope extraction, envelope processing (AM bandpass filtering), short-term memory and template-matching stages of the model to account for the data for HIo listeners. This suggests that cochlear hearing loss spares modulation filtering (consistent with the data of Sek and Moore, 2006), short-term memory, and decision factors.

However, the additive internal noise had to be increased substantially to counterbalance the beneficial effect of loss of cochlear compression on AM sensitivity and predict AMDTs close to those obtained. This suggests that hearing loss degrades at least some aspects of (central) processing efficiency.

V. CONCLUSIONS

The effects of age and sensorineural hearing loss on AM sensitivity and temporal integration of AM cues were investigated. This was achieved by measuring AM detection thresholds (AMDTs) at 40 dB SL for a group of listeners with mild-to-moderate sensorineural hearing loss, aged between 50 and 64 years, using a carrier frequency of 500 Hz and modulation rates of 2 and 20 Hz. The number of modulation cycles, *N*, varied between 2 and 9. The data were compared with AMDTs measured for young and older normal-hearing listeners (Wallaert *et al.*, 2016).

The data for both AM rates show that aging *per se* degrades AM sensitivity while sparing temporal integration of AM cues, whereas hearing loss enhances both AM sensitivity and temporal integration of AM cues. When compared with the data for older normal-hearing listeners, the improvement in AM sensitivity associated with sensorineural hearing loss was as large as 8 dB, and temporal integration of AM cues was increased by 3-4 dB over the range of *N* tested here.

A computational model based on peripheral filtering and envelope extraction followed by a modulation filterbank and a template-based optimal detector was developed to account for the data. AM detection was limited by three internal noises (an additive and a multiplicative noise, and an additive memory noise), applied after modulation filtering.

Taken together, the psychophysical and simulation data suggest that the loss of amplitude compression in the impaired cochlea is mainly responsible for the enhanced sensitivity and temporal integration of temporal-envelope cues found for hearing-impaired listeners. They also suggest that, for AM detection, sensorineural hearing loss is associated with

a substantial increase in internal noise, but with preserved short-term memory and decision mechanisms.

References

Ardoint, M., Lorenzi, C., Pressnitzer, D. and Gorea, A. (**2008**). "Perceptual constancy in the temporal envelope domain," J. Acoust. Soc. Am., **123**, 1591-1601.

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.

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.

Demany, L., Montandon, G., and Semal, C. (**2003**). "Pitch discrimination as a function of the inter-stimulus interval: Evidence against a simple model of perceptual memory," J. Acoust. Soc. Am., **113**, 2290-2290.

Demany, L., Montandon, G., and Semal, C. (**2005**). "Internal noise and memory for pitch," in Pressnitzer, D., de Cheveigné, A., McAdams, S., and Collet, L. (Eds) *.Auditory Signal Processing: Physiology, Psychoacoustics, and Models* (Springer, New York), pp 136-144.

Ewert, S. D., and Dau, T. (**2000**). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am., 108, 1181-1196.

Ewert, S. D., Verhey, J. L., and Dau, T. (**2002**). "Spectro-temporal processing in the envelope-frequency domain," J. Acoust. Soc. Am., **112**, 2921-2931.

Ewert, S. D., and Dau, T. (**2004**). "External and internal limitations in amplitude-modulation processing," J. Acoust. Soc. Am. **116**, 478-490.

Florentine, M., Fastl, H., and Buus, S. (**1988**). "Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking," J. Acoust. Soc. Am. **84**, 195-203.

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. (**2013**). "Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss," Am. J. Audiol. **22**, 313-315.

Füllgrabe, C., Moore, B. C. J, 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, 1-25.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear Res., **47**, 103-138.

Goodman, A. (1965). "Reference zero levels for pure-tone audiometer," ASHA **7**, 262-263.

Gygi, B., Kidd, G. R., and Watson, C. S. (**2004**). "Spectral-temporal factors in the identification of environmental sounds," J. Acoust. Soc. Am., **115**, 1252-1265.

Joosten, E.R.M., Shamma, S.A., Lorenzi, C., and Neri, P. (**2016**). "Dynamic reweighting of auditory modulation filters," PLOS Comput. Biol., **12**(7). doi: 10.1371/journal.pcbi.1005019.

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.

Kohlrausch, A., Fassel, R., and Dau, T. (**2000**). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," J. Acoust. Soc. Am. **108**, 723-734.

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.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am., **49**, 467-477.
Lorenzi, C., Soares, C., and Vonner, T. (**2001**). "Second-order temporal modulation transfer functions," J. Acoust. Soc. Am., **110**, 1030-1038.
Magnussen, S. (**2000**). "Low-level memory processes in vision," Trends Neurosci., **23**, 247-251.
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," Brit. J. Audiol., **26**, 229-237.
Moore, B. C. J., and Sek, A. (**1994**). "Effects of carrier frequency and background noise on the detection of mixed modulation," J. Acoust. Soc. Am., **96**, 741-751.
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.J., Shailer, M.J., Schooneveldt, G.P. (**1992**). "Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss," Brit. J. Audiol., **26**, 229–237.
Moore, B. C. J., 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. (**2007**). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues, 2nd Ed.* (Wiley, Chichester), pp. 1-332.
Oxenham, A.J., Moore, B.C.J, and Vickers, D.A. (**1997**). "Short-term temporal integration: evidence for the influence of peripheral compression," J. Acoust. Soc. Am. **101**, 3676-3687.
Paraouty, N., and Lorenzi, C. (**2017**). "Using individual differences to assess modulation-processing mechanisms and age effects, " Hear. Res. 344, 38–49.
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.
Patterson, R. D., Allerhand, M. H., and Giguère, C. (**1995**). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," J. Acoust. Soc. Am. **98**, 1890-1894.
Plack, C.J., and Skeels, V. (**2007**). "Temporal integration and compression near absolute threshold in normal and impaired ears," J. Acoust. Soc. Am. **122**, 2236-2244.
Schlittenlacher, J. and Moore, B.C.J. (**2016**). "Discrimination of amplitude-modulation depth by subjects with normal and impaired hearing," J. Acoust. Soc. Am. **140**, 3487-3495.
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.
Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science, **270**, 303-304.
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.
Schoof, T., and Rosen, S.M. (**2014**). "The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners," Front. Aging Neurosci., **6**, 1-14.
Sek, A., and Moore, B.C.J. (**2002**). "Mechanisms of modulation gap detection," J. Acoust. Soc. Am. **111**, 2783-2792.
Sek, A., and Moore, B. C. J. (**2006**). "Perception of amplitude modulation by hearing-impaired listeners: The audibility of component modulation and detection of phase change in three-component modulators," J. Acoust. Soc. Am. **119**, 507-514.
Sheft, S., and Yost, W. A. (**1990**). "Temporal integration in amplitude modulation detection," J. Acoust. Soc. Am., **88**, 796-805.
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.
Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am., **66**, 1364-1380.

Viemeister, N. F., and Wakefield, G. H. (**1991**). "Temporal integration and multiple looks," J. Acoust. Soc. Am., **90**, 858-865.
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.
Zwicker, E. (**1952**). "Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones (The limits of audibility of amplitude modulation and frequency modulation of a pure tone)," Acustica **2**, 125-133.
Zwicker, E. (**1956**). "Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs (The foundations for determining the information capacity of the auditory system)," Acustica **6**, 356-381.

CHAPITRE 4

EFFETS DE LA PERTE AUDITIVE SUR L'INTEGRATION TEMPORELLE EN AM ET EN FM

10. Differential effects of sensorineural hearing loss on temporal integration of amplitude- and frequency-modulation cues

10.1 Résumé:

The effect of the number of modulation cycles (N) on frequency-modulation (FM) detection thresholds (FMDTs) was measured in the absence and presence of interfering amplitude modulation (AM) for hearing-impaired (HI) listeners, using a 500-Hz sinusoidal carrier and FM rates of 2 and 20 Hz. The data were compared with FMDTs for normal-hearing (NH) listeners and AM detection thresholds (AMDTs) for NH and HI listeners [Wallaert, Moore and Lorenzi (2016). J. Acoust. Soc., **139**, 3088-3096; Wallaert, Moore, Ewert and Lorenzi (2017). J. Acoust. Soc., **141**, 971-980]. FMDTs were higher for HI than for NH listeners, but the effect of increasing N was similar across groups. In contrast, AMDTs were lower and the effect of increasing N was greater for HI listeners than for NH listeners. A model of temporal-envelope processing based on the modulation-filterbank concept and a template-matching decision strategy accounted better for the FMDTs at 20 Hz than at 2 Hz and failed to predict the temporal integration of FM. These results suggest that different mechanisms underlie AM and FM detection at very low rates and that hearing loss impairs the mechanisms responsible for FM detection, but preserves the memory and decision processes responsible for temporal integration of FM.

10.2 Article:

Wallaert, N., Varnet L., Moore, B. C. J., and Lorenzi, C. (2017). "Differential effects of sensorineural hearing loss on detection and temporal integration of amplitude- and frequency-modulation cues,"

I. INTRODUCTION

Communication sounds, including speech and animal vocalizations, contain relatively slow (<5-10 Hz) patterns of frequency modulation (FM) that play an important role in the discrimination and identification of the sounds (e.g., Rees and Malmierca, 2005; Zeng et al., 2005). Consistent with this idea, several psychophysical studies have shown that auditory sensitivity to slow FM is associated with speech identification for both normal-hearing (NH) and hearing-impaired (HI) individuals when speech reception is limited by acoustic degradation or competing speech sounds (Buss et al., 2004; Strelcyk and Dau, 2009; Johannesen et al., 2016). To understand the nature of the link between FM detection and speech perception, it is important to clarify the mechanisms involved in FM detection and the effects of cochlear hearing loss on FM perception (e.g., Paraouty et al., 2016; Whiteford et al., 2017. That is the goal of the present study.

Many studies indicate that sensorineural hearing loss impairs sensitivity to FM, sometimes substantially (for a review, see Moore, 2014). This has been reported for a wide range of audiometric configurations, stimulus levels, carrier frequencies, modulation rates and modulation patterns (e.g., Moore and Skrodzka, 2002; Lacher-Fougère and Demany, 1998; Buss et al., 2004; Strelcyk and Dau, 2009; Ernst and Moore, 2012; Johannesen et al., 2016; Kortlang et al., 2016; Paraouty et al., 2016). The deficits for hearing-impaired (HI) listeners are generally interpreted as resulting from degradation of the sensory encoding of FM. In the normal auditory system, information about FM is potentially conveyed by two cues: (1) temporal-envelope cues, related to the dynamic changes in the distribution of excitation along the basilar membrane within the cochlea (resulting in fluctuations in firing rate of neurons with characteristic frequencies below and above the carrier frequency); (2) temporal fine-structure (TFS) cues related to the waveform of the stimulus at each point on the basilar membrane (and the corresponding fine timing of the action potentials in neurons with different characteristic frequencies) (Rees and Malmierca, 2005; Paraouty et al., 2016). It has been proposed that for normal-hearing (NH) listeners FM is encoded via TFS cues when the FM rate is low (<5 Hz) and

when the carrier frequency is below about 4 kHz (Demany and Semal, 1986, 1989; Sek and Moore, 1995; Moore and Sek, 1996; Moore and Skrodzka, 2002; Ernst and Moore, 2010; Paraouty et al., 2016; Wallaert et al., 2016), and via temporal-envelope cues when the FM is fast or when the carrier frequency is higher than 4 kHz (Zwicker, 1956; Maiwald, 1967; Saberi and Hafter, 1995; Moore and Sek, 1996; Ernst and Moore, 2010).

The ability to use envelope cues may be reduced in the impaired auditory system because of reduced conversion of FM into temporal-envelope cues at the outputs of the cochlear filters, resulting from the loss of outer hair cells in the cochlea and the consequent broadening of the filters (Glasberg and Moore, 1986). A reduced ability to use TFS cues may result from: (1) Reduced precision of neural phase locking or a reduced number of neurons conveying phase locking (Kujawa and Liberman, 2009); (2) Changes in the relative phase of response at different points along the basilar membrane in the cochlea (Ruggero, 1994), which would disrupt central mechanisms decoding TFS information based on correlation of the outputs of adjacent places (Shamma, 1985); (3) More complex and more rapidly varying neural TFS resulting from broader cochlear filters (Moore, 2008); (4) A mismatch between neural TFS information and the place on the basilar membrane that would "normally" respond to that information (Henry et al., 2016); (5) Central changes following cochlear hearing loss, such as loss of neural inhibition (cf. Moore, 2008).

For HI listeners, the poorer-than-normal sensitivity to high-rate FM may result mainly from reduced conversion of FM into temporal-envelope cues, whereas the poorer-than-normal sensitivity to low-rate FM for low carrier frequencies may result from abnormal encoding and/or use of TFS cues. However, the origin of the observed deficits in FM detection remains a matter of debate (Ernst and Moore, 2012; Sheft et al., 2012; Whiteford et al., 2017; Kortlang et al., 2016; Paraouty et al., 2016; Wallaert et al., 2016), and it is possible that the deficits result partly from reduced "processing efficiency", that is a reduced ability of the central auditory system to make use of available sensory information, perhaps caused by a decline in short-term memory and decision-making capacities. Consistent with this idea, recent work suggests that

aging and sensorineural hearing loss are associated with reduced processing efficiency (modeled as increased "internal noise") for amplitude modulation (AM) detection (Füllgrabe et al., 2015; Paraouty et al., 2016; Wallaert et al., 2017).

The present study was designed to clarify the relative contributions of sensory and non-sensory factors to FM detection for older HI listeners by combining a psychophysical and modeling approach. To do this, temporal integration was compared for AM and FM detection. Temporal integration here refers to the improvement in modulation detection thresholds with increasing number of modulation cycles, N (Viemeister, 1979; Hartmann and Klein, 1980). The temporal integration of AM and FM is assumed to reflect central (i.e., decision) processes involving "multiple looks" (Viemeister and Wakefield, 1991; Sheft and Yost, 1990) or a template-matching process (Dau et al., 1997). In a recent study (Wallaert et al., 2017), AM detection thresholds (AMDTs) were measured for NH listeners and listeners with mild-to-moderate sensorineural hearing loss using a carrier frequency of 500 Hz and modulation rates of 2 and 20 Hz. AMDTs decreased (i.e., improved) with increasing N for both groups and AMDTs were smaller for the HI than for the NH group. Also, temporal integration of AM (the beneficial effect of increasing N) was greater for the HI group.

A computational model based on the modulation-filterbank concept and a template-matching decision strategy (Dau et al., 1997) was developed to account for the data. The data and the simulations suggested that the loss of amplitude compression in the impaired cochlea was mainly responsible for the enhanced sensitivity and temporal integration of AM found for the HI listeners. They also suggested that, for AM detection, cochlear damage is associated with increased internal noise, but preserved short-term memory and decision mechanisms. We thus reasoned that, if FM detection relies on the use of temporal-envelope cues resulting from FM-to-AM conversion at the outputs of the cochlear filters, sensorineural hearing loss should have similar effects on the temporal integration of AM and FM cues. In other words, temporal integration of FM should be greater for HI listeners than for NH listeners.

In the present study, FM detection thresholds (FMDTs) were measured at 40 dB

sensation level (SL) for listeners with mild-to-moderate sensorineural hearing loss (aged between 50 and 64 years) using a 500-Hz sinusoidal carrier and modulation rates of 2 and 20 Hz. FMDTs were measured with and without interfering AM at the same rate. The AM was present in both intervals of the forced-choice task, and was therefore uninformative, whereas the FM was present in only one interval. The task was to identify the interval with the FM. The interfering AM was used to make it more difficult for the listeners to use temporal-envelope cues. In particular, if FM is detected by monitoring changes in excitation at a single point on the excitation pattern (Zwicker, 1952, 1956), the interfering AM should greatly impair performance (Moore and Sek, 1996). FMDTs were also measured with and without interfering AM for young NH listeners and older NH listeners matched in age with the hearing-impaired listeners. The data were compared with FMDTs measured by Wallaert et al. (2016) for young and older NH listeners and AMDTs measured by Wallaert et al. (2016, 2017) for young and older NH listeners and older HI listeners under identical conditions. Finally, the data were compared to the predictions of a model of temporal-envelope processing based on the assumption that modulation detection is achieved by cross-correlating the outputs of AM channels with memory-stored templates according to an "optimal detector" scheme (Dau et al., 1997a,b; see also Ardoint et al., 2008). The model incorporated stages simulating temporal-envelope processing by low-level sensory mechanisms (cochlear filtering, instantaneous amplitude compression, adaptation), mid-level processes (bandpass AM filtering), and higher-level non-sensory processes (internal noise, memory decay in the temporal-envelope domain, and template matching).

II. METHOD

A. Listeners

1. Hearing-impaired listeners

Twelve older hearing-impaired (HIo) listeners aged between 50 and 64 years (mean = 59 years; standard deviation, SD = 4 years) were tested for FM detection with interfering AM. Nine

of these were also tested for FM detection without interfering AM. The audiograms for the tested (right) ears of all 12 listeners are shown in Fig. 1. All had audiometric thresholds ≤60 dB hearing level (HL) between 0.25 and 4 kHz. The severity of hearing loss was classified as mild to moderate based on audiometric thresholds. The audiometric thresholds at the tested frequency (0.5 kHz) ranged between 30 and 50 dB HL (mean = 40 dB HL; SD = 5.6 dB). For the 9 HIo listeners who were tested for FM detection without interfering AM, the audiometric threshold at the tested frequency ranged between 30 and 45 dB HL (mean = 38 dB HL; SD = 5 dB). The cochlear origin of hearing loss was diagnosed based on air-conduction, bone-conduction and impedance audiometry. Eleven out of the 12 HIo listeners participated in a previous study that investigated temporal integration for AM detection (Wallaert et al., 2017).

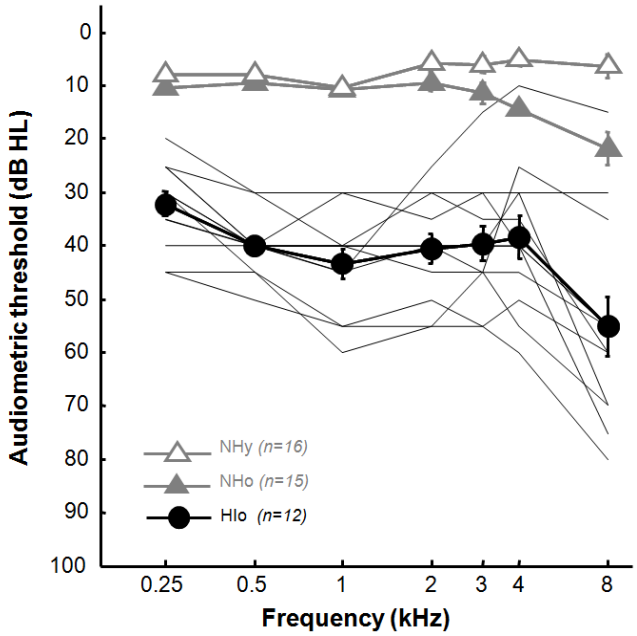


Figure 1. *Individual (thin lines) and mean (filled circles and black thick lines) audiometric thresholds for the older hearing-impaired (HIo) listeners for the right (tested) ear. Mean thresholds for the young normal-hearing (NHy) group are shown by open gray triangles and mean thresholds for the older normal-hearing (NHo) group are shown by filled gray triangles. Error bars indicate ±1 standard error of the mean. They are omitted when they would be smaller than the size of the symbol used to plot the data.*

2. Normal-hearing listeners

FMDTs were measured with and without interfering AM for 11 young normal-hearing (NHy) listeners aged between 22 and 28 years (mean = 24 years; SD = 2 years) and 11 older normal-hearing (NHo) listeners aged between 47 and 66 years (mean = 55 years; SD = 8 years). The mean audiometric threshold at 0.5 kHz was 7.7 dB HL (SD = 3.4 dB) for the NHy group, and

9.5 dB HL (SD = 5.7 dB) for the NHo group. A t-test conducted on audiometric thresholds at 0.5 kHz showed no significant difference between NH groups [t(20)=0.91; p=0.375]. The data for 8 of the 11 NHy listeners were reported by Wallaert et al. (2016).

FMDTs with interfering AM and AMDTs were measured for 16 NHy listeners aged between 22 and 28 years (mean = 24 years; SD = 2 years) and 15 NHo listeners aged between 45 and 66 years (mean = 56 years; SD = 8 years). The data for 15 of the NHy listeners and 14 of the NHo listeners were reported by Wallaert et al. (2016, 2017). All of these NHy and NHo listeners had audiometric thresholds less than 20 dB HL for the tested (right) ear for octave frequencies between 0.25 and 3 kHz, but some of the NHo listeners had elevated audiometric thresholds for frequencies above 3 kHz. The mean audiometric threshold at 0.5 kHz was 7.8 dB HL (SD = 3.1 dB) for the NHy group and 9.3 dB HL (SD = 5 dB) for the NHo group. Audiometric thresholds at 0.5 kHz were not significantly different between the NHy and NHo groups.

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 and N° IRB : 20153600001072).

B. Procedure

FMDTs were measured using a two-interval forced-choice adaptive procedure estimating the 70.7% point on the psychometric function. Each trial contained a target and a standard stimulus, in random order, with a silent interval of 600 ms between them. The root-mean-square level of the two stimuli was, on average, the same. However, to discourage the use of level cues, the levels of the standard and target stimuli were roved independently over a range of ±1.5 dB. The onset and offset ramp durations were 250 and 25 ms for the modulation rates of 2 and 20 Hz, respectively. The FM was present in only one interval of a trial, selected at random, while in conditions with interfering AM, the AM was present in both intervals of a trial, and was therefore uninformative. The duration of each signal was determined by the number of modulation cycles, N, and the modulation rate, fm. N was 2, 3, 4, 5, or 9 cycles when FM was presented with

interfering AM and 2, 3 or 4 cycles when FM was presented without any interfering AM. Each interval was marked by a light on the computer screen. Listeners were asked to indicate which interval contained the "extra" pitch fluctuation. Visual feedback as to the correct response was provided after each response.

The tracking variable was the frequency excursion in hertz, Δf (see section C for details). A run started with the tracking variable well above the estimated detection threshold. The step size was a factor of 1.58 until two reversals had occurred, and 1.26 thereafter. Fourteen reversals were obtained for each run, and the threshold estimate for that run was taken as the geometric mean of the value Δf at the last six reversals. Two threshold estimates were obtained for each condition. The final estimate of the threshold was taken as the geometric mean of the two threshold estimates.

C. Stimuli

The stimuli were identical to those used by Wallaert et al. (2016). All stimuli were generated digitally at a sampling rate of 44.1 kHz. They were presented using Sennheiser HD 280 pro headphones (Old Lyme, CT) and an external soundcard (ECHO Indigo DJ 1-2, Santa Barbara, California; 16-bit resolution) at a mean SL of 40 dB, except for HIo listeners with absolute thresholds of 40 dB HL or more, when the level was set to 80 dB SPL. Levels were calibrated using a Bruel & Kjaer 2250 sound level meter and ear simulator type 4153 (Naerum, Denmark) complying with IEC 60318-1. Stimuli were presented only to the right ear.

For the FMDTs measured with interfering AM, the standard stimulus contained AM but not FM. The target stimulus contained both AM and FM. Eq. 1 describes the target stimulus T(t):

$$T(t) = AM(t)\sin[(2\pi fct + \Phi) + \beta(\sin(2\pi fmt + \varphi))] \quad (1)$$

with: $AM(t) = [1 + m\sin(2\pi fmt + 3\pi/2)]$ with $m = 0.33$

and: $\beta = \Delta f / fm$

where Δf is the frequency excursion in Hz, fm is the modulation rate (2 or 20 Hz), φ is the starting phase of the FM, fc is the carrier frequency (500 Hz), Φ is the starting phase of the

carrier, and t is time in s. The values of φ, and Φ were randomly and independently chosen for each stimulus. Although the starting phase of the AM was fixed, its phase relative to that of the FM was random (because φ was chosen randomly), so the AM should still have been effective in disrupting the use of temporal-envelope cues. For the standard stimulus, Δf was set to 0. For the FMDTs measured without interfering AM, the value of m was set to 0.

III. RESULTS

A. AMDTs compared to FMDTs with and without AM (*N*=2-4 cycles)

Figure 2 shows mean AMDTs (top panels, taken from Wallaert et al., 2016, 2017) and FMDTs (bottom panels) as a function of N, for the NHy (left-most panels), NHo (middle panels) and HIo (right-most panels) listeners. FMDTs (lower panels) were measured with (grey lines) and without (black lines) interfering AM. In each panel, the filled circles and open triangles show the thresholds for fm = 2 Hz and fm = 20 Hz, respectively. AMDTs are expressed as 20log10m (dB) and FMDTs are expressed as Δf (Hz).

The reader is referred to Wallaert et al. (2016) for analysis of the AMDTs and FMDTs for the NHy and NHo listeners and for discussion of the effects of age on AM and FM detection. The FMDTs for the HIo listeners are comparable to those reported previously (Moore and Skrodzka, 2002; Lacher-Fougère and Demany, 1998; Buss et al., 2004; Strelcyk and Dau, 2009; Ernst and Moore, 2012; Johannesen et al., 2016; Paraouty et al., 2016) for FM rates below 20 Hz and carrier frequencies below 1 kHz. In addition, the results are consistent with those of Kortlang et al. (2016), for which random slow FM (1-4 Hz) was used with and without superimposed AM.

A mixed-model analysis of variance (ANOVA) was conducted on the AMDTs with between-subjects factor group (3 levels: NHy, NHo and HIo) and within-subjects factors N (3 levels) and modulation rate fm (2 levels). The effect of modulation rate was significant [$F(1, 28) = 47.53$, $p < 0.001$]. For all three groups (NH and HI listeners), AMDTs were lower (i.e., better) for fm = 2 Hz than for fm = 20 Hz. The effect of N was significant [$F(2, 56) = 26.01$, $p < 0.001$]. For each modulation rate and each group, AMDTs decreased (i.e., improved) with increasing N.

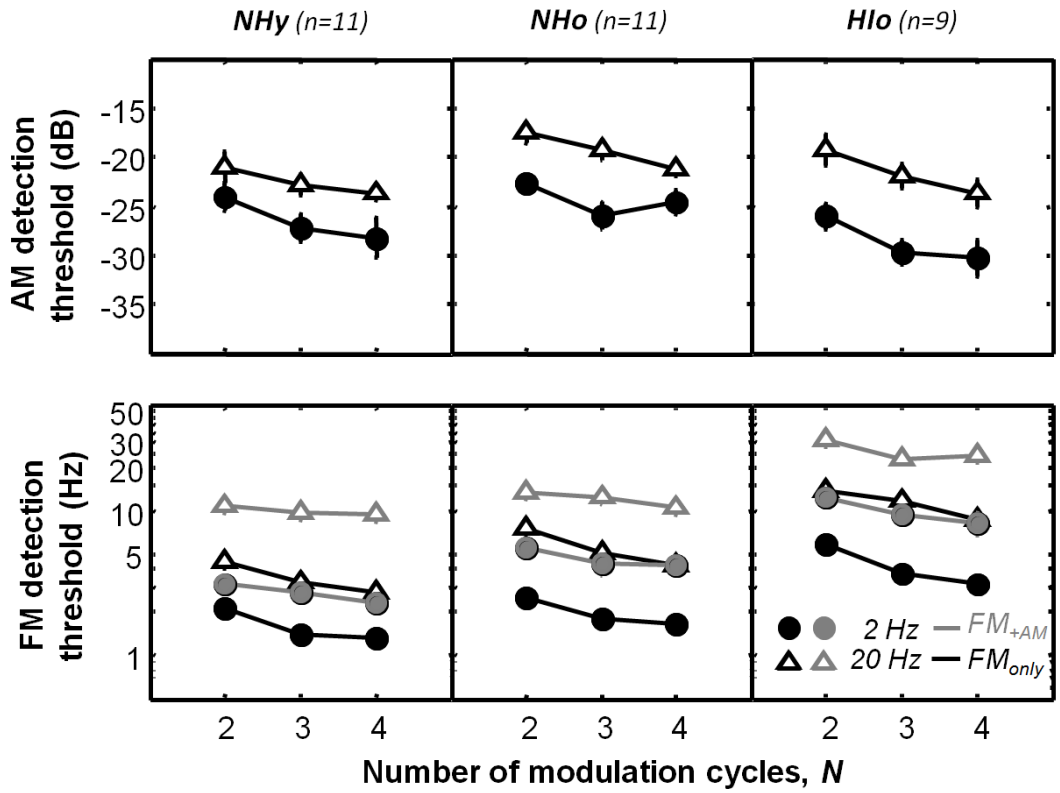


Figure 2. *Mean AM detection thresholds (AMDTs, top panels) and FM detection thresholds (FMDTs, bottom panels) as a function of N for the NHy (left panels), NHo (middle panels) and HIo (right panels) groups. In each panel, the filled circles and open triangles show the thresholds for fm = 2 and 20 Hz, respectively. FMDTs measured without interfering AM ("FM-only" condition) are shown as black symbols and black lines. FMDTs measured with interfering AM ("FM_{+AM}" condition) are shown as gray symbols and gray lines.*

The interaction between N and fm was not significant [F(2, 56) = 1.43, p = 0.24], indicating that AMDTs decreased similarly with increasing N for the two modulation rates. The interaction between N and group was not significant [F(4, 56) < 1]. This indicates that the temporal integration of AM cues was similar for all groups. This appears to be inconsistent with the results of Wallaert et al. (2017), who found greater temporal integration of AM for HIo listeners than for NHy and NHo listeners. However, the study of Wallaert et al. (2017) included higher values of N (5 and 9), making it easier to detect differences across groups. The effect of group was significant [F(2, 28) = 3.46, p = 0.045]. Post hoc comparisons were based on the Fisher least-significant difference (LSD) test. For all conditions, AMDTs were higher for the NHo than for the NHy group (p=0.049), and were lower for the HIo group than for the NHo group (p=0.022), indicating a deleterious effect of greater age and a beneficial effect of hearing loss.

The interaction between modulation rate, fm, and group [$F(2, 28) = 1.28$, $p = 0.29$] was not significant, indicating that the effect of modulation rate was similar for the three groups. None of the other interactions involving group was significant (N x fm x group: $F(4, 56) < 1$).

A mixed-model ANOVA was conducted on the (log-transformed) FMDTs with within-subject factors presence/absence of interfering AM (2 levels), N (3 levels) and modulation rate fm (2 levels), and between-subjects factor group (3 levels). The effect of group was significant [$F(2,28) = 96.5$; $p < 0.001$], FM detection thresholds being higher for the older groups than for the younger group, and being higher for the HIo group than for the NHo group. The effect of interfering AM was significant [$F(1,28) = 350.4$; $p < 0.001$], FM detection being worse in the presence of interfering AM as show by post hoc comparison based on the Tukey test ($p<0.001$).

The effect of N was significant [$F(2,56) = 94.3$; $p < 0.001$], FM detection thresholds decreasing with increasing N. The effect of modulation rate was significant [$F(1,28) = 308.9$; $p < 0.001$], FM detection being better at fm = 2 Hz than fm = 20 Hz. The interaction between interfering AM and modulation rate was significant [$F(1,28) = 5.8$; $p = 0.022$], the detrimental effect of the interfering AM being greater for fm = 20 Hz than for fm = 2 Hz (Tukey test ; $p<0.001$). This finding is consistent with previous work (Moore and Sek, 1996; Ernst and Moore, 2010, 2012) and with the notion that temporal-envelope cues play a greater role for fm = 20 Hz than for fm = 2 Hz. The interaction between interfering AM and N was significant [$F(2,56) = 5.6$; $p = 0.006$], the decrease in threshold with increasing N being greater in the absence than in the presence of interfering AM. The interaction between interfering AM, group and modulation rate was significant [$F(2,28) = 5.8$; $p = 0.008$]. To analyse further this triple interaction, an additional mixed-model ANOVA was conducted on the ratio between the FMDTs measured with and without interfering AM (an estimate of the interference effect caused by AM), with within-subject factors N (3 levels) and modulation rate fm (2 levels), and between-subjects factor group (3 levels). The interaction between group and modulation rate was significant [$F(2,28) = 4.895$; $p = 0.015$]. Post hoc analysis (Tukey test) indicated that the deleterious effect of interfering AM was greater at fm = 20 Hz than at fm = 2 Hz for the NHy group only ($p=0.005$). This pattern of

results is consistent with the idea that NHy listeners can detect FM for low carrier frequencies and very low FM rates using a cue other than that resulting from FM-to-AM conversion (a cue presumably based on the use of TFS) and that the ability to use this cue is reduced with increasing age and with hearing loss.

B. AMDTs compared to FMDTs with AM (*N*=2-9 cycles)

Figure 3 shows mean AMDTs (top panels) and FMDTs measured with interfering AM (bottom panels) as a function of N for the NHy (left-most panels), NHo (middle panels) and HIo (right-most panels) listeners. The FMDTs were always measured with interfering AM. N ranged from 2 to 9 cycles.

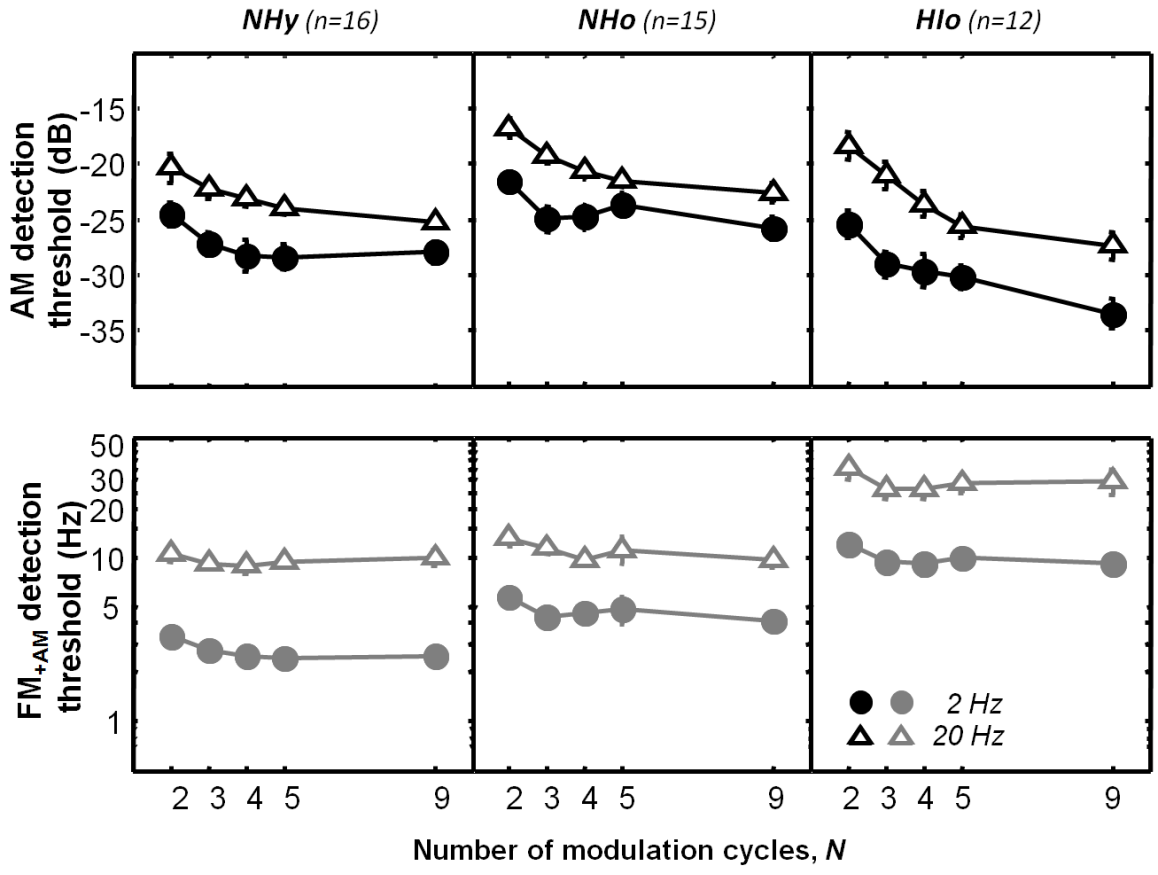


Figure 3. *As Fig. 2, except for the greater range of N, and the fact that FMDTs were measured only with interfering AM.*

A mixed-model ANOVA was conducted on the AMDTs with between-subjects factor group (3 levels: NHy, NHo and HIo) and within-subjects factors N (5 levels) and modulation rate

(2 levels). The effect of modulation rate was significant [$F(1, 40) = 73.79$, $p < 0.001$]. For all groups, AMDTs were lower (i.e., better) for fm = 2 Hz than for fm = 20 Hz. The effect of N was significant [$F(4, 160) = 63.63$, $p < 0.001$]. For each modulation rate and each group, AMDTs decreased (improved) with increasing N. The interaction between N and fm was significant [$F(4, 160) = 3.70$, $p = 0.007$]. Consistent with the results of Wallaert et al. (2017), the interaction between N and group was significant [$F(8, 160) = 3.68$, $p < 0.001$]. The effect of group was significant [$F(2, 40) = 9.95$, $p< 0.001$]. Post hoc comparisons were based on the Fisher LSD test. For all N, AMDTs were higher for the NHo group than for the NHy group ($p=0.002$), but were lower for the HIo group than for the NHo group ($p<0.001$), indicating a deleterious effect of greater age but a beneficial effect of hearing loss. The interaction between fm and group [$F(2, 40) =1.65$, $p = 0.20$] was not significant: the effect of fm was similar for the three groups. None of the other interactions involving group was significant (N x fm x group: [$F(8, 160) < 1$]) .

A linear regression analysis was conducted on the individual AMDTs (logm versus log(N)). A repeated-measures ANOVA was conducted on the resulting slopes with between-subject factor group (3 levels) and within-subjects factor fm (2 levels). There was a significant effect of fm [$F(1, 40) = 7.64$, $p = 0.008$], a significant effect of group [$F(2, 40) = 9.1$, $p < 0.001$] and no significant interaction [$F(2,40)<1$, $p=0.611$]. Post hoc analysis (LSD test) analyses indicate that temporal integration for AM detection was: i) greater at 20 Hz than at 2 Hz ($p = 0.008$); ii) greater for the HIo than for the two NH groups ($p<0.002$).

A mixed-model ANOVA was conducted on the FMDTs with between-subjects factor group (3 levels) and within-subjects factors N (5 levels) and modulation rate (2 levels). The effect of modulation rate was significant [$F(1, 40) = 22.76$, $p < 0.001$]. For all groups, FMDTs were lower for fm = 2 Hz than for fm = 20 Hz ($p<0.001$; Tukey test). The effect of N was significant [$F(4, 160) = 14.5$, $p < 0.001$]. The interaction between N and fm was not significant [$F(4, 160) < 1$], suggesting that FMDTs decreased similarly with increasing N for the two modulation rates. The interaction between N and group was not significant [$F(8, 160) < 1$], indicating that FMDTs decreased similarly with increasing N for all groups. The effect of group

was significant [$F(2, 40) = 66.2$, $p < 0.001$]. For all conditions, FMDTs were greater for the NHo than for the NHy group ($p=0.01$; Tukey test) and were greater for the HIo than for the NHo group ($p<0.001$; Tukey test), indicating deleterious effects of both age and hearing loss. There was a significant interaction between fm and group [$F(2, 40) =3.6$, $p = 0.03$]. The ratio between FMDTs for fm = 20 and 2 Hz was 3.6 for the NHy group, 2.4 for the NHo group, and 2.9 for the HIo group. The difference between the NHY and NHo groups is consistent with the finding of Wallaert et al. (2016) that age had a greater effect on FMDTs at 2 Hz than at 20 Hz. This might reflect the effects of age on the use of TFS cues for 2-Hz FM detection. Hearing loss was associated with an increase in FMDTs for both modulation rates, consistent with the results of Lacher-Fougère and Demany (1998), Moore and Skrodzka (2002) and Strelcyk and Dau (2009). This may happen because hearing loss reduces the ability to use both temporal envelope cues (reduced FM-to-AM conversion because of broader auditory filters) and TFS cues. None of the other interactions involving group was significant.

IV. MODELING AM AND FM PROCESSING

A. Rationale

The present data combined with those reported by Wallaert et al. (2016, 2017) show a differential effect of hearing loss on AM and FM detection and integration: FM detection was poorer for the HIo than for the NHo group but temporal integration of FM cues was similar for the two groups, whereas AM detection was better for the HIo than for the NHo group and temporal integration of AM cues was greater for the HIo than for the NHo group. At first sight, this dissociation suggests different mechanisms for AM and FM detection and integration when fc is low (here, 500 Hz) and suggests that hearing loss preserves the memory and decision processes responsible for temporal integration of FM cues. To assess whether this interpretation is correct, the FM detection thresholds were compared with the predictions of an envelope-processing model whose parameters were initially adjusted to fit the AMDTs. The model is described in section IV.B.

B. The model

The general structure of the model is similar to that described by Wallaert et al. (2017). The first stage was a set of five linear gammatone filters (Patterson et al., 1995) that simulate the bandpass filtering of the basilar membrane in the cochlea. One filter (called the "on-frequency" filter) was centered on the carrier frequency of the signal (500 Hz). The remaining four filters (the "off-frequency" filters) were centered one and two Cams (units of the ERBN-number scale; Glasberg and Moore, 1990) above and below the carrier frequency. In the second stage, the outputs of the five Gammatone filters were half-wave rectified. For simulating normal hearing, the output of the on-frequency filter was processed by a broken-stick input-output function, which applied compression for levels above a given threshold level, chosen so that compression occurred only for stimuli with levels above 40 dB SPL. The compression was implemented by raising the amplitude to the power 0.3. The rectified signals at the outputs of the four off-frequency auditory filters were not compressed. This crudely simulates the physiological finding that compression on the basilar membrane occurs mainly for input frequencies close to the characteristic frequency of the place whose response is being measured (Robles and Ruggero, 2001).

The signals derived from the outputs of the five auditory filters were highpass filtered (cutoff = 2.5 Hz; rolloff = 20 dB/decade) to simulate the effects of adaptation in the AM domain (Tchorz and Kollmeier, 1999). The resulting signals were passed through a set of 10 bandpass (modulation) filters to simulate frequency-selective processing in the AM domain (Dau et al., 1997a,b). The modulation filters had a Q value of 1 and a rolloff of ±20 dB/decade (Ewert and Dau, 2000; Ewert et al., 2002; Lorenzi et al., 2001; Sek and Moore, 2002). The center frequencies of the modulation bandpass filters ranged between 2 and 120 Hz (Moore et al., 2009) and the center frequencies of the modulation filters were spaced logarithmically. 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. This was done to simulate the loss of sensitivity to envelope phase for

rates above 10 Hz (Dau et al., 1997a,b). Each envelope of the envelope (the so-called "venelope"; Ewert et al., 2002) was scaled to ensure that the root-mean-square value at the output of modulation filters was the same before and after the Hilbert transformation.

The outputs of the modulation filters (the envelope for fm < 10 Hz; the venelope for fm≥ 10 Hz) were down-sampled by a factor 10 and then "degraded" by two sources of internal noise, which were introduced to limit the performance of the model. These noises were: (i) an additive noise, Eadd, and (ii) an additive time-varying memory noise, characterized by parameter Emem. Both internal noises were modeled as Gaussian noises with zero mean and with magnitudes specified by their standard deviation. Randomly drawn samples of each noise were added independently to each sample point of the outputs of the bandpass AM filters. Eadd was introduced to limit intensity discrimination and AM sensitivity (Dau et al., 1997a,b). Emem was introduced to simulate imperfect retention of temporal-envelope information (Ardoint et al., 2008; Wallaert et al., 2017). This noise primarily played a role for very long stimuli (e.g., for 4.5-s stimuli, as used when N = 9 and fm = 2 Hz). The standard deviation of the memory noise increased backward in time according to an exponential function with a 1.4-s half life. The parameter Emem corresponds to the maximum value of the standard deviation of the memory noise. The resulting (noisy) signals were used as input to the decision stage of the model.

The decision stage was realized as a simplified version of the optimal detector described by Dau et al. (1997a,b). The output of the model was computed for the two stimuli in each trial. A "template" was generated at the beginning of each simulated threshold measurement (that is, for the first trial of each adaptive staircase) as the difference between the model output in response to the target and standard stimuli. This was done with the two internal noise sources set to zero. Note that this largely cancels out the effects of the added AM on the internal representation, and therefore it represents an idealized version of what actual subjects do. For each subsequent trial, the template was cross-correlated with the model output for each interval in that trial. The lags used by the cross-correlation device were restricted to ±1 modulation cycle. The interval leading to the largest cross-correlation was taken as the target interval identified by the model for that

trial. Stimuli were generated as for the behavioral experiment, except that the presentation level was fixed for all simulations. Ninety thresholds were simulated and averaged for each condition.

C. Choice of model parameters

To model the data for the NHy and NHo groups, the bandwidths of the Gammatone filters were set to either 0.7 or 1 ERBN (Glasberg and Moore, 1990). The same values were used for these two groups because auditory filters do not broaden with increasing age when the audiometric threshold remains normal (Lutman *et al*., 1991; Peters and Moore, 1992). The values of ERBN represent average values of the bandwidths of the auditory filters estimated using data from simultaneous masking experiments (mostly with notched noise). The smaller value of 0.7 ERBN was used based on the suggestion that human cochlear tuning may be sharper than estimated using simultaneous notched-noise masking (Oxenham and Shera, 2003) and also based on previous attempts to model AM and FM detection for NH listeners, which showed that filters with bandwidths less than 1 ERBN were required to give a good fit to the data (Moore and Sek, 1994). To model the data for the HIo group, the bandwidths of the gammatone filters were set to either 1.4 or 2 ERBN, simulating the reduced frequency selectivity associated with moderate sensorineural hearing loss (Glasberg and Moore, 1986; Moore, 2007). The compression parameter was set to 0.3 to model the data for the NHy and NHo groups, simulating "normal" cochlear compression for the on-frequency filter, and to 1.0 to model the data for the HIo group, simulating the loss of compression associated with cochlear hearing loss. The cutoff frequency of the highpass filter used to model adaptation was chosen so as to predict the relative values of the AMDTs for the two modulation rates. The values of the two noise parameters were chosen by first adjusting parameter E_{add} and then parameter E_{mem} so as to fit the AMDTs for the NHy listeners. Subsequently, parameter E_{add} was adjusted to fit the AMDTs for the NHo and HIo listeners, while keeping E_{mem} constant. Consistent with the modeling study of Wallaert *et al.*, (2017), the values of E_{add} obtained in this way were higher than for the NHy listeners by a factor 1.3 for the NHo listeners and by a factor 5.3 for the HIo listeners.

D. Simulation results

Figures 4 to 7 show the predictions of the model as symbols connected by dashed lines for AM detection (top panels) and FM detection (bottom panels). Black lines and symbols denote AMDTs and FMDTs measured without interfering AM. Gray lines and symbols denote FMDTs measured with interfering AM. Symbols connected by solid lines (when present) show the obtained thresholds.

1. Effect of simulated auditory filter width on predicted AMDTs

Consider first the effect of the bandwidths of the simulated auditory filters for the two NH groups on AM detection (top-left and top-middle panels in Figs. 4-7). Generally, the data were fitted more accurately with the bandwidth of 0.7 ERBN than with the bandwidth of 1 ERBN, especially for the 20-Hz rate. This is consistent with the proposal that the cochlear filters in NH ears are somewhat sharper than estimated using simultaneous notched-noise masking (Oxenham and Shera, 2003). The AM-detection data for the HIo listeners were fitted roughly equally well for the simulated auditory filters with bandwidths of 1.4 and 2 ERBN.

2. Effects of age and hearing loss on predicted AMDTs

The model predictions shown in the top panels of Figs. 4-7 captured the effects of AM rate, *N* and group (*i.e.*, the effects of age and hearing loss) on the AMDTs reasonably well. However, the model predicted greater temporal integration for AM than observed for the NHy listeners when f_m = 20 Hz (see the top-left panels in Figs. 6 and 7), and lower AM thresholds than observed for the HIo listeners when f_m = 2 Hz (top-right panels of Figs. 4-7), although the discrepancy in the latter case was small when the widths of the simulated filters were 2 ERBN (see Figs. 5 and 7). The model made slightly less accurate predictions than the model described by Wallaert *et al*. (2017). However, average absolute deviation between the obtained and predicted thresholds (across *N*) was reasonably small, ranging between 0.2 and 2 dB across conditions and groups.

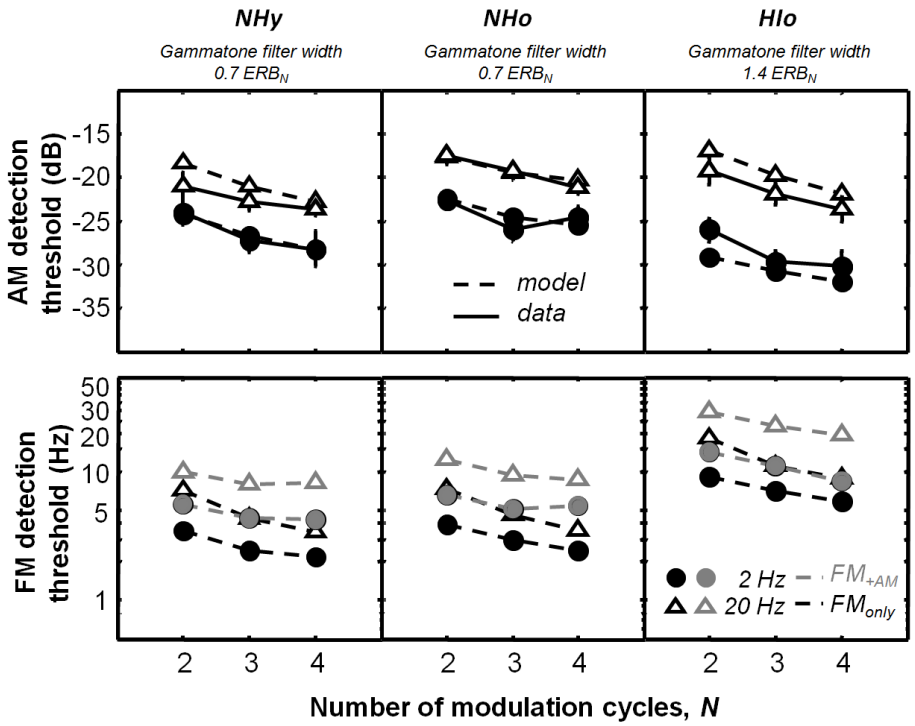


Figure 4. *Predicted AMDTs and FMDTs (dashed lines) for the models simulating the NHy, NHo and HIo listeners. The widths of the gammatone filters were set to 0.7 ERBN for the NHy and NHo models, and 1.4 ERBN for the HIo model. Obtained AMDTs are shown by black symbols and continuous black lines in the top panels. Otherwise as Fig. 2.*

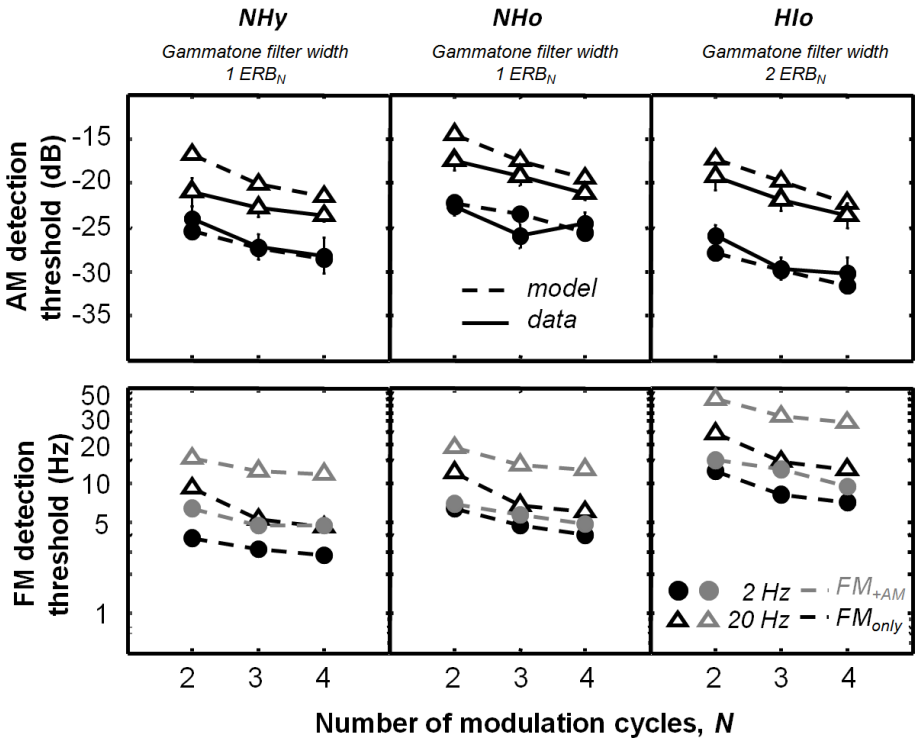


Figure 5. *As Fig. 4, but with the widths of the gammatone filters set to 1 ERBN for the NHy and NHo models, and 2 ERBN for the HIo model.*

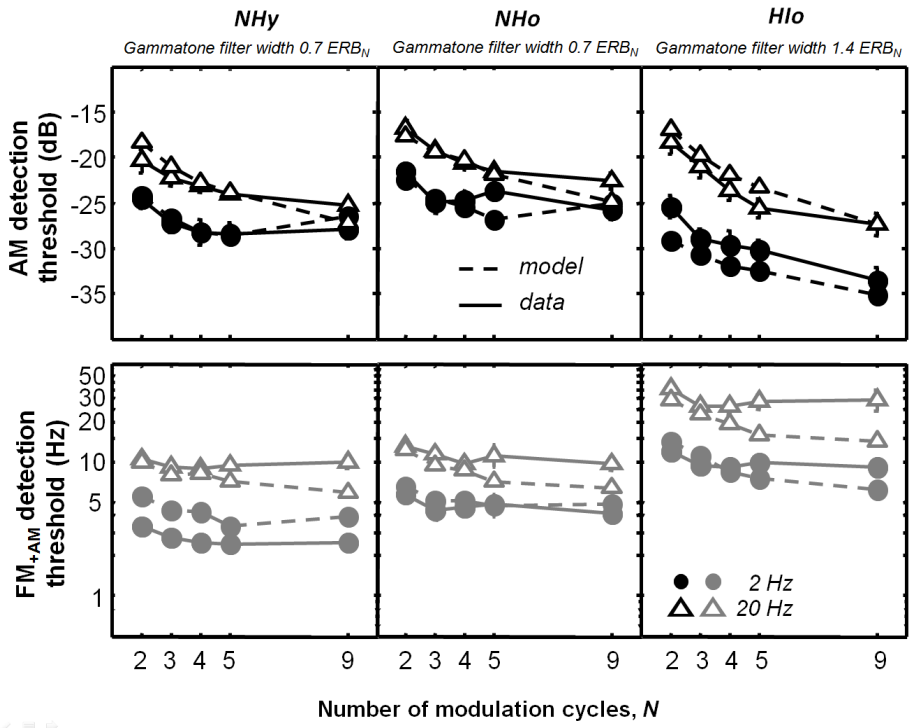


Figure 6. *Obtained (solid lines) and predicted (dashed lines) AMDTs (top) and FMDTs (bottom) for each group. The widths of the gammatone filters were set to 0.7 ERBN for the NHy and NHo models, and 1.4 ERBN for the HIo model. FMDTs were measured only with interfering AM.*

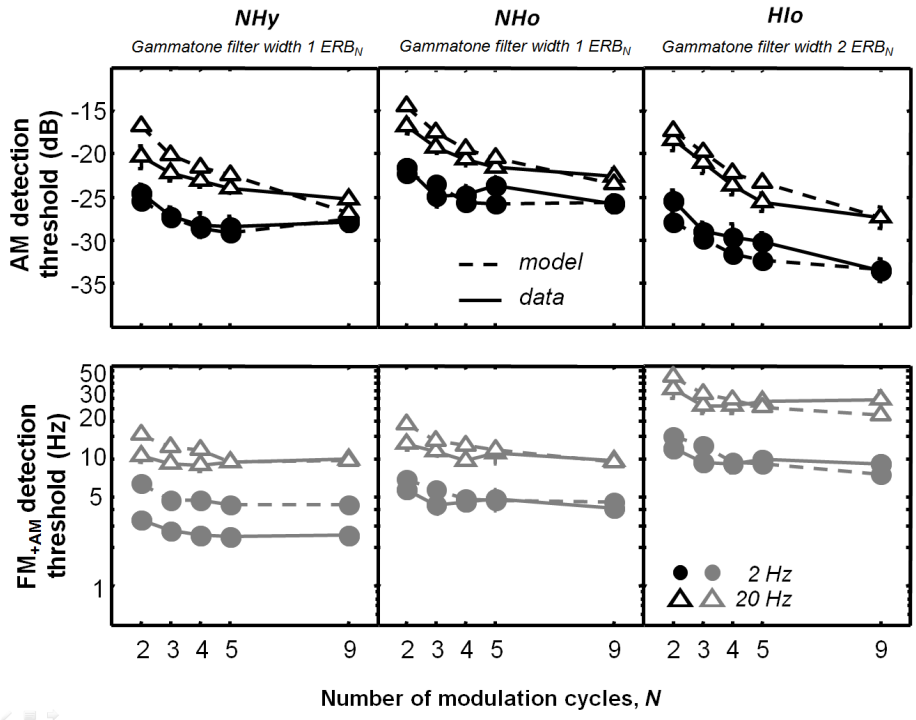


Figure 7. *As Fig. 6, but with the widths of the gammatone filters set to 1 ERBN for the NHy and NHo models, and 2 ERBN for the HIo model.*

3. Effects of age and hearing loss on predicted FMDTs

The model predictions shown in the bottom panels of Figs. 4-7 captured some of the main trends in the data for FM detection. Specifically, for both values of the widths of the gammatone filters, the predictions captured the effects of modulation rate (the predicted FMDTs were always lower for f_m = 2 than for f_m = 20 Hz), the effects of N (the predicted FMDTs improved gradually when N increased from 2 to 9), and the effects of interfering AM (the predicted FMDTs were always poorer in the presence of interfering AM). Consistent with the general form of the results, the model predicted that FMDTs would be higher with greater age and with hearing loss. However, the predicted FMDTs were often higher than the obtained FMDTs, especially for the NHy group in the presence of interfering AM and for f_m = 2 Hz (see, for example, the bottom-left panels in Figs. 6 and 7). The discrepancy between the predictions and the data was somewhat less when the width of the simulated gammatone filters was 0.7 ERBN than when it was 1 ERBN, (compare Figs. 6 and 7), but it nevertheless remained quite large for the 0.7 ERBN bandwidth.

As can be seen by comparing Figs. 2, 4 and 5, FMDTs without and with interfering AM with N = 2, 3 and 4 were predicted more accurately when the widths of the gammatone filters were set to 0.7 ERB_N for the NH model and 1.4 ERB_N for the HI model. The goodness of fit was assessed by taking the absolute value of the logarithm of the *ratio* between predicted and obtained FMDTs and averaging this across values of N and the presence/absence of AM; the smaller this average, called log-AV, the better is the fit. For the NH model, log-AV was 0.12 when the filter width was 0.7 ERB_N and 0.21 when it was 1 ERB_N. For the HI model, log-AV was 0.12 when the filter width was 1.4 ERB_N and 0.17 when it was 2 ERB_N.

The use of sharper gammatone filters resulted in larger temporal-envelope fluctuations generated by FM-to-AM conversion and this allowed the model to predict lower FMDTs, that were more consistent with the data. However, for FMDTs obtained with interfering AM using a wider range of N (Fig. 3), the predictions for the two older groups and f_m = 20 Hz were more accurate for the wider simulated auditory filters (Fig. 7) than for the narrower simulated auditory filters (Fig. 6).

The model predicted FMDTs with and without interfering AM (data in Fig. 2) more accurately for f_m = 20 Hz than for f_m = 2 Hz. For the NH model with filter width 0.7 ERB_N, log-AV was 0.17 when f_m = 2 Hz and 0.07 when f_m = 20 Hz. For the HI model with filter width 1.4 ERB_N, log-AV was 0.16 when f_m = 2 Hz and 0.07 when f_m = 20 Hz. This is consistent with the idea that FMDTs for f_m = 20 Hz are based entirely on the use of temporal envelope cues, while some other cue or cues play a role for f_m = 2 Hz.

4. Effect of interfering AM on predicted FMDTs

Interestingly, the model did predict the greater detrimental effect of interfering AM on FMDTs for f_m = 20 Hz than for f_m = 2 Hz. However, inconsistent with the data (Fig. 2, bottom panels), the predicted effect of interfering AM was of similar magnitude for the NHy, NHo and HIo models for both modulation rates (Figs. 4 and 5, bottom panels). For both values of the width of the gammatone filters, the ratio of predicted FMDTs with and without interfering AM (averaged across *N*) ranged between 1.3 and 1.6 for f_m = 2 Hz and was about 2 for f_m = 20 Hz for the NHy, NHo and HIo models. Corresponding obtained ratios were 1.8 for the NHy listeners and 2.6 for the NHo and HIo listeners for f_m = 2 Hz, and 2.8 for the NHy listeners, 2.1 for the NHo listeners and 2.5 for the HIo listeners for f_m = 20 Hz. This inconsistency between the data and predictions occurred because the predicted FMDTs obtained without interfering AM for f_m = 2 Hz were systematically higher (*i.e.*, poorer) than the obtained FMDTs for all three groups. The predicted FMDTs obtained with interfering AM for f_m = 2 Hz were close to the obtained FMDTs. This suggests that the NHy, NHo and HIo listeners used a cue distinct from temporal-envelope cues when detecting FM, and that this extra cue (presumably, TFS) was still available - but to a lesser extent – for the NHo and HIo listeners. It appears that this additional cue can be used more effectively in the absence than in the presence of interfering AM.

5. Temporal integration of FM

Whatever the width of the gammatone filters, the predicted temporal integration for FM in the presence of interfering AM, when measured using a wide range of *N* (2-9 cycles), was much higher than observed for each group of listeners for f_m= 20 Hz. The discrepancy was smaller for f_m= 2 Hz, but was still apparent for the HIo group. As shown in Fig. 6, when the assumed bandwidths of the auditory filters were small (0.7 ERBN for the two NH groups and 1.4 ERBN for the HIo group), this "over-integration" effect led to predicted FMDTs that were below obtained FMDTs for large *N*. As shown in Fig. 7, when the assumed bandwidths of the auditory filters were larger (1 ERBN for the two NH groups and 2 ERBN for the HIo group), the predicted FMDTs were above the obtained FMDTs for small *N*.

IV. DISCUSSION

C. Effects of age and hearing loss on AM detection and temporal integration

For each value of *N* and for both AM rates, AMDTs for the HIo listeners were significantly lower (*i.e.*, better) than those for the NHo listeners tested at the same SL. In previous studies, this beneficial effect of sensorineural hearing loss on AM sensitivity was interpreted as resulting from the loss of cochlear compression and the associated loudness recruitment (Moore *et al*., 1996; Wallaert *et al.*, 2017). The interaction between *N* and group was significant, temporal integration of AM being greater for the HIo group than for the two NH groups. This effect was predicted by the model (see the top panels of Figs. 6 and 7), consistent with the idea that it arises from loss of amplitude compression in the cochlea. Overall the results are consistent with the notion that at least some aspects of "processing efficiency" related to the short-term storage of temporal-envelope information and decision making are spared by cochlear damage (Wallaert *et al.*, 2017).

Overall, the model presented here made slightly less accurate predictions of AMDTs than the version of the model described by Wallaert *et al*. (2017), which used a single off-frequency gammatone filter centered 1 Cam above the carrier frequency). More precisely, temporal

integration of AM at f_m=20 Hz for NHy listeners was better predicted by Wallaert *et al*. (2017). This difference resulted partly from differences in the number and center frequencies of the simulated auditory filters used to convert FM into AM and partly from differences in the implementation of the template-matching decision device (a simple correlation process in Wallaert *et al*. (2017) versus cross-correlation in the current model). It appears that the cross-correlation device is more efficient than a simple correlation device and real listeners. The cross-correlation device is more robust to (internal) noise than a simple correlation device. The current simulations also suggest that real listeners are less optimal than assumed here. It may be the case that human listeners cannot make perfect comparisons between their internal template and the noisy output of the modulation filters. In future studies, this central inefficiency could be modeled by adding a temporal jitter to the lags used by the cross-correlation process, hence limiting more strongly the comparison between the internal template and the output of the modulation filters at high modulation rates.

B. Effects of age and hearing loss on FM detection

The FMDTs for the HIo listeners were significantly higher (poorer) than those for the NHo listeners. This detrimental effect of hearing loss on FM detection is consistent with previous results obtained for older listeners with sensorineural hearing loss (Moore and Skrodzka, 2002; Lacher-Fougère and Demany, 1998; Buss *et al.*, 2004; Strelcyk and Dau, 2009; Ernst and Moore, 2012; Paraouty *et al*., 2016) when the NH and HI groups were compared at equal SLs.

The detrimental effect of age on FM detection predicted by the model depended on a small increase in the additive internal noise used to degrade the internal representation of temporal-envelope cues at the outputs of the modulation filters. The detrimental effect of hearing loss on FM detection predicted by the model depended on the combined effects of (i) the broadening of the gammatone filters and the consequent reduction of FM-to-AM conversion and (ii) a substantial increase in the additive internal noise.

Although the model reproduced some trends in the data, the model did not provide a

unified account of AM and FM detection. The modeling showed that whatever the bandwidth of the simulated auditory filters, the predicted FMDTs for f_m = 2 Hz were higher than obtained (especially for the NHy listeners) and temporal integration of FM was over-estimated for both modulation rates for all three groups. Although it is conceivable that more accurate predictions of FMDTs and temporal integration for f_m = 20 Hz may be obtained by using a more accurate simulation of cochlear frequency selectivity and by making the cross-correlation process less optimal (see above), it seems unlikely that such a model would succeed at predicting the low FMDTs found for the NHy listeners for f_m = 2 Hz. The general failure of this multi-band model of temporal-envelope processing argues for separate mechanisms for AM and FM detection for very low modulation rates and low carrier frequencies.

C. Interference effects of AM on FM

The interaction between the presence/absence of interfering AM and modulation rate was significant, the detrimental effect of the AM being greater for f_m = 20 Hz than for f_m = 2 Hz (Fig. 2), but this effect was mainly driven by the results for the NHy group, as supported by the significant three-way interaction between group, presence/absence of AM, and modulation rate. The results for the NHy group replicate previous findings (Moore and Sek, 1996; Ernst and Moore, 2010, 2012). The results are consistent with the idea that, for the NHy listeners, FM detection depended partly on a cue or cues other than temporal envelope, presumably TFS cues, and TFS cues only slightly disrupted by interfering AM. The ability to use TFS cues worsens with increasing age (Moore *et al*., 2012a; Moore et al., 2012b) and with hearing loss (Hopkins and Moore, 2007; Moore, 2014) and this may account for the finding that the effect of the interfering AM for f_m = 2 Hz was greater for the NHo and HIo groups than for the NHy group.

Although this explanation is plausible, the model based on temporal-envelope cues alone did predict a greater effect of the interfering AM on FM detection for f_m = 20 Hz than for f_m = 2 Hz, when the auditory filters were assumed to have bandwidths of 1 ERB_N for the NH groups

and 2 ERB_N for the HIo group (see Fig. 5, bottom). This happened because, in the model, AM phase information was removed for modulation filters tuned above 10 Hz. This loss of phase information limited the effectiveness of the cross-correlation device (the template-matching process) in extracting the pattern of envelope cues resulting from FM-to-AM conversion. However, the interference effect predicted by the model for f_m = 2 Hz was much smaller than the obtained interference effect, and the model predicted similar interference effects for all three groups, which is inconsistent with the data. Therefore, it seems that the model does not correctly predict the pattern of interference effects produced by the AM.

Assuming that FM detection for f_m = 2 Hz was mediated by the use of TFS cues, at least for the NHy group, the obtained results indicate that interfering AM does have some adverse effect on the ability to use TFS cues. This may occur because the precision of phase locking and the phase at which nerve spikes occur both depend somewhat on stimulus level (Johnson, 1980) and on whether the amplitude is increasing or decreasing (Moore, 1973).

D. Temporal integration of FM cues

The temporal integration of FM information was similar for the three groups. Thus, sensorineural hearing loss does not impair the central processes underlying temporal integration of FM information, such as the combination of "multiple looks" (Viemeister and Wakefield, 1991) or template matching (Dau *et al*., 1997). For FM detection in the presence of interfering AM, temporal integration was very small for all three groups (Fig. 3, bottom panels); FMDTs changed by less than a factor of 2 as *N* was increased from 2 to 9. This may have happened because the disruptive effect of the AM depended primarily on the phase of the AM relative to that of the FM in the signal interval (Moore and Sek, 1992); some relative phases are more disruptive than others. For example when the instantaneous frequency is moving upwards, this leads to a decrease in excitation level on the low-frequency side of the excitation pattern, but this might be cancelled or changed into an increase if the amplitude is also

increasing. Although the relative phase of the AM and FM changed from one trial to the next, it was fixed within a given stimulus. Essentially, for a stimulus with *N* cycles, the same information was repeated *N* times. The "noise" associated with the random variation in the relative phase of the AM and FM may have limited the amount of temporal integration.

The small amount of temporal integration for FM in the presence of interfering AM was not predicted by the model (Figs. 6 and 7, bottom panels). It may be the case that a template based on a single sample of AM+FM with random relative phase would not represent adequately what the auditory system would do to perform the task. A more realistic simulation may be to use an internal template based on a sample of FM alone. This strategy was simulated in the FM+AM condition (see Figure 8) but the predictions of NHy data were no more accurate. Thus, as pointed out above, the failure of the model probably occurs because human listeners cannot make perfect comparison between their internal template and the output of modulation filters, or they cannot construct a perfect template based on the difference between the internal representations of each interval.

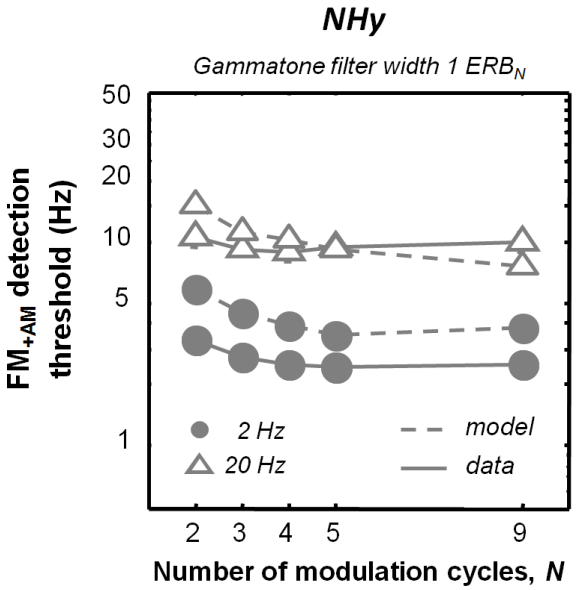


Figure 8. *As Fig. 7 for the NHy model, but the model uses an internal template based on a sample of FM alone.*

V. CONCLUSIONS

The effects of age and sensorineural hearing loss on the detection and temporal integration of AM and FM were investigated. This was achieved by measuring FM detection thresholds (FMDTs) at 40 dB SL for a group of listeners with mild-to-moderate sensorineural hearing loss, aged between 50 and 64 years, using a carrier frequency of 500 Hz and modulation rates of 2 and 20 Hz. The number of modulation cycles, *N*, varied between 2 and 9. The data were compared with FMDTs and AM detection thresholds (AMDTs) measured for young and older normal-hearing listeners (Wallaert *et al*., 2016, 2017).

FMDTs were higher for the hearing-impaired listeners than for the normal-hearing listeners, but the effect of increasing *N* was similar across groups. In contrast, AMDTs were lower (better) and the effect of increasing *N* was greater for the hearing-impaired than for the normal-hearing listeners.

A model based on peripheral filtering and envelope extraction followed by a modulation filterbank and a template-based optimal detector was developed to interpret the data. AM and FM detection was limited by two internal noises (an additive noise and an additive memory noise) applied after modulation filtering, whose standard deviation was adjusted to fit the AMDTs. Overall, the model accounted better for FMDTs for the 20-Hz modulation rate than for the 2-Hz rate, and predicted greater temporal integration for FM than observed for all groups.

Taken together, the psychophysical data and modeling results suggest that (i) FM detection for a modulation rate of 2 Hz and low carrier frequencies probably depends on a cue or cues other than those arising from FM-to-AM conversion, for example cues related to temporal fine structure; (ii) greater age is associated with increased FMDTs and this can be explained partly by increased internal noise and partly by reduced sensitivity to temporal fine structure; (iii) hearing loss impairs the mechanisms responsible for FM detection, but preserves the memory and decision processes responsible for temporal integration of FM.

References

Ardoint, M., Lorenzi, C., Pressnitzer, D. and Gorea, A. (**2008**). “Perceptual constancy in the temporal envelope domain,” J. Acoust. Soc. Am., **123**, 1591-1601.

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.

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 Acust. Acust. **88**, 334–346.

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.

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., and Semal, C. (**1989**). "Detection thresholds for sinusoidal frequency modulation," J. Acoust. Soc. Am., **85**, 1295–1301.

Ernst, S. M., and Moore, B. C. J. (**2010**). "Mechanisms underlying the detection of frequency modulation," J. Acoust. Soc. Am., **128**, 3642–3648.

Ernst, S. M., and Moore, B. C. J. (**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., and Dau, T. (**2000**). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am., **108**, 1181–1196.

Ewert, S. D., Verhey, J. L., and Dau, T. (**2002**). "Spectro-temporal processing in the envelope-frequency domain," J. Acoust. Soc. Am., **112**, 2921–2931.

Füllgrabe, C., Moore, B. C. J., 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**, 1-25..

Glasberg, B. R., and Moore, B. C. J. (**1986**). "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," J. Acoust. Soc. Am., **79**, 1020–1033.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103-138.

Hartmann, W. M., and Klein, M. A. (**1980**). "Theory of frequency modulation detection for low modulation frequencies," J. Acoust. Soc. Am., **67**, 935–946.

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.

Hopkins, K., and Moore, B. C. J. (**2007**). "Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information," J. Acoust. Soc. Am. **122**, 1055-1068.

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**, 1-14.

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," J. Acoust. Soc. Am. **68**, 1115-1122.

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.

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.

Lacher-Fougère, S., and Demany, L. (**1998**). "Modulation detection by normal and hearing-impaired listeners," Audiology, **37**, 109–121.

Lorenzi, C., Soares, C., and Vonner, T. (**2001**). "Second-order temporal modulation transfer functions," J. Acoust. Soc. Am., **110**, 1030–1038.
Lutman, M. E., Gatehouse, S., and Worthington, A. G. (**1991**). "Frequency resolution as a function of hearing threshold level and age," J. Acoust. Soc. Am. **89**, 320-328.
Maiwald, D. (**1967**). "Die berechnung von modulationsschwellen mit hilfe eines funktionsschemas". Acustica, **18**, 193-207.
Moore, B. C. J. (**1973**). "Frequency difference limens for narrow bands of noise," J. Acoust. Soc. Am. 54, 888-896.
Moore, B. C. J. (**2007**). "Cochlear hearing loss: physiological, psychological and technical issues". (Wiley, Chichester, UK), pp 332.
Moore, B. C. J. (**2008**). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," J. Assoc. Res. Otolaryngol., **9**, 399–406.
Moore, B. C. J. (**2014**). "Auditory processing of temporal fine structure: Effects of age and hearing loss," (World Scientific, Singapore), pp. 1-182.
Moore, B. C. J., and Sek, A. (**1994**). "Effects of carrier frequency and background noise on the detection of mixed modulation," J. Acoust. Soc. Am. **96**, 741-751.
Moore, B. C. J., and Sek, A. (**1992**). "Detection of combined frequency and amplitude modulation," J. Acoust. Soc. Am. **92**, 3119-3131.
Moore, B. C. J, and Sek, A. (**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., 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. J., Glasberg, B. R., Stoev, M., Füllgrabe, C., and Hopkins, K. (**2012a**). "The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies," J. Acoust. Soc. Am. **131**, 1003-1006.
Moore, B. C. J., Vickers, D. A., and Mehta, A. (**2012b**). "The effects of age on temporal fine structure sensitivity in monaural and binaural conditions," Int. J. Audiol. **51**, 715-721.
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. J., Füllgrabe, C., and Sek, A. (**2009**). "Estimation of the center frequency of the highest modulation filter," J. Acoust. Soc. Am. **125**, 1075-1081.
Oxenham, A. J., and Shera, C. A. (**2003**). "Estimates of human cochlear tuning at low levels using forward and simultaneous masking," J. Assoc. Res. Otolaryngol., **4**, 541–554.
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.
Peters, R. W., and Moore, B. C. J. (**1992**). "Auditory filters and aging: filters when auditory thresholds are normal," in Auditory Physiology and Perception, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 179-185.
Rees, A., and Malmierca, M. S. (**2005**). "Processing of dynamic spectral properties of sounds," Int. Rev. Neurobiol., **70**, 299–330.
Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305-1352.
Ruggero, M. A. (**1994**). "Cochlear delays and traveling waves: comments on 'experimental look at cochlear mechanics':[A. Dancer, Audiology 1992; 31: 301-312]," Audiology, **33**, 131–142.
Saberi, K., and Hafter, E. R. (**1995**). "A common neural code for frequency-and amplitude-modulated sounds," Nature, **374**, 537-539.
Sek, A., and Moore, B. C. J. (**1995**). "Frequency discrimination as a function of frequency, measured in several ways," J. Acoust. Soc. Am. **97**, 2479-2486.

Sek, A., and Moore, B. C. J. (**2002**). "Mechanisms of modulation gap detection," J. Acoust. Soc. Am., **111**, 2783–2792.

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.

Sheft, S., and Yost, W. A. (**1990**). "Temporal integration in amplitude modulation detection," J. Acoust. Soc. Am., **88**, 796–805.

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.

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.

Viemeister, N. F. (**1979**)." Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am., **66**, 1364–1380.

Viemeister, N. F., and Wakefield, G. H. (**1991**). "Temporal integration and multiple looks," J. Acoust. Soc. Am., **90**, 858–865.

Wallaert, N., Moore, B. C. J., 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., 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.

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.

Zeng, F.-G., Nie, K., Stickney, G. S., Kong, Y.-Y., Vongphoe, M., Bhargave, A., Cao, K. (**2005**)." Speech recognition with amplitude and frequency modulations," Proc. Natl. Acad. Sci. U.S.A., **102**, 2293–2298.

Zwicker, E. (**1952**). "Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones (The limits of audibility of amplitude modulation and frequency modulation of a pure tone)," Acustica 2, 125-133.

Zwicker, E. (**1956**). "Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs (The foundations for determining the information capacity of the auditory system)," Acustica 6, 356-381.

CHAPITRE 5

DISCUSSION ET CONCLUSIONS

11. Discussion et conclusion

Ce programme de recherche doctoral, combinant mesures psychoacoustiques et simulations informatiques, a permis de : 1) préciser la nature des mécanismes impliqués dans le traitement des indices d'AM et de FM, et 2) clarifier l'impact du vieillissement et de la perte auditive neurosensorielle sur la sensibilité aux composantes d'AM et de FM de signaux acoustiques complexes.

11.1. Synthèse sur le traitement de l'AM:

11.1.1. Effet de la cadence de modulation

Lorsque la durée de stimulation est ajustée de façon à égaliser le nombre de cycle, la sensibilité aux indices d'AM diminue pour l'ensemble des sujets lorsque la cadence de modulation augmente (AMDLs moins bons à 20 Hz qu'à 2 Hz), suggérant une certaine inertie des mécanismes impliqués dans l'encodage et/ou le décodage des indices d'AM. Cet effet contraste avec les précédentes mesures de sensibilité à l'AM réalisées pour des durées de stimulation fixes de quelques centaines de ms (Sheft & Yost, 1990; Kohlrausch et al., 2000; Ernst & Moore, 2012), indiquant que les seuils de détection d'AM sont meilleurs pour une cadence de modulation de 20 Hz que pour une cadence de modulation de 2 Hz. Un modèle computationnel de traitement d'enveloppe temporelle basé sur le concept du banc de filtres de modulation et sur un mécanisme d'appariement de gabarit rend compte de cet effet de cadence de modulation, résultant du filtrage passe haut simulant l'adaptation nerveuse à court terme et de la nature du mécanisme de décision (un mécanisme d'appariement de gabarit). Ces résultats suggèrent que l'inertie du traitement de l'AM résulte à la fois de la contribution de facteurs périphériques et de facteurs centraux au traitement temporel des informations d'enveloppe.

11.1.2. Effet du nombre de cycles

Les seuils de détection d'AM s'améliorent substantiellement pour l'ensemble des groupes de sujets à mesure que le nombre de cycles présentés à l'auditeur augmente, en accord avec les données de la littérature (e.g., Viemeister, 1979, Sheft & Yost, 1990, Dau et al., 1997). Ce phénomène d'intégration temporelle d'AM est plus important pour une cadence de modulation rapide (20 Hz versus 2 Hz), probablement en raison de limitations induites par la mémoire auditive à court terme lorsque la durée de stimulation dépasse les capacités de stockage de la mémoire échoïque et plus particulièrement, les capacités du registre mnésique spécialisé dans la rétention des informations de sonie (Darwin et al., 1972; Clement et al., 1999; Ardoint et al., 2008). Les travaux de modélisation montrent que l'intégration temporelle est sous tendue par deux mécanismes centraux: la capacité de stockage et le mécanisme décisionnel (i.e., appariement de gabarits). Ceci est conforme aux résultats de Dau et al. (1997). Toutefois, pour simuler l'intégration temporelle aux durées de stimulation importantes (plusieurs secondes), il a été nécessaire de prendre en compte la contribution de facteurs mnésiques. Nos simulations révèlent que les indices d'enveloppe seraient a priori stockés dans un registre de mémoire à court terme dont la constante de temps (ou la demi-vie) serait d'environ 1.2 à 1.4 secs.

11.1.3. Effet du vieillissement

Nos résultats corroborent l'existence d'un effet délétère et global du vieillissement sur la sensibilité à l'AM quelle que soit la cadence de modulation (en accord avec He et al., 2008; Füllgrabe et al., 2015; Kumar & Sangamanatha, 2011; Purcell et al., 2004). Toutefois, cet effet non sélectif, bien que significatif, est de taille relativement modeste (approximativement 2-3 dB de différence entre les seuils des auditeurs jeunes et ceux des auditeurs âgés). Nos travaux de simulation permettent de rendre compte des données des sujets âgés en augmentant le bruit interne (additif) dans le domaine de l'enveloppe par un facteur 1.7, suggérant une perte relativement modeste d'efficacité de traitement centrale, ce conformément à l'hypothèse formulée par Füllgrabe (2015).

Par ailleurs, le vieillissement ne modifie pas l'intégration temporelle observée en AM, ce qui suggère que la rétention d'information et les processus de décision impliqués dans l'intégration temporelle sont intacts. Si certains aspects de l'efficacité de traitement centrale se dégradent avec l'âge (comme le suggère l'augmentation du bruit interne), d'autres aspects de l'efficacité de traitement semblent donc être préservés. Nos travaux de modélisation permettent d'étudier séparéement les différents aspects de l'efficacité de traitement (e.g. fidélité de la représentation interne, stockage, et prise de décision) qui pourraient être affectés par le vieillissement. Toutefois, l'origine exacte de ces troubles d'efficacité de traitement demeure encore incertaine: dégradation des capacités attentionnelles, diminution du rapport signal-sur-bruit dans la représentation interne de l'enveloppe consécutive à une déafférentation périphérique voire centrale (Kujawa & Liberman, 2009; Sergeyenko et al., 2013), etc.

Notons qu'il eut été également possible de simuler les données des normo-entendants âgés en diminuant le gain des filtres de modulation. Cette diminution du gain engendrerait une diminution du rapport signal sur bruit interne, conduisant à une dégradation de la performance. Ceci pourrait permettre de rendre compte des effets du vieillissement sans augmentation du bruit interne additif. On ne peut donc pas statuer avec certitude quant à l'origine sensorielle ou centrale de la perte de sensibilité à l'AM avec l'âge. Seules les mesures de bruit interne dans le domaine de l'enveloppe temporelle réalisées à l'aide de la méthode de double-passe (Green, 1964) ou du bruit équivalent permettront de répondre à ces questions.

11.1.4. Effet de la perte auditive

Nos données psychophysiques corroborent l'idée selon laquelle la perte auditive améliore la sensibilité à l'AM quelle que soit la cadence de modulation lorsqu'on égalise le niveau de présentation en dB SL (en accord avec Füllgrabe et al., 2003; Moore et al., 1996; Bacon & Gleitman, 1992; Glasberg & Moore, 1990, Schlittenlacher & Moore, 2016). Nos travaux de modélisation confirment l'hypothèse selon laquelle ce rehaussement des indices d'enveloppe est attribuable au recrutement de sonie consécutif à la perte du mécanisme actif cochléaire.

Contrairement au vieillissement, la perte auditive améliore très fortement l'intégration temporelle observée en AM. Notre modèle permet de rendre compte de cette sur-intégration par une diminution de la compression cochléaire. Quelque soit l'étude de modélisation (modèle mono bande versus multi-bandes), cette perte de compression s'accompagne d'une forte amélioration des AMDLs simulés, qu'il convient de compenser par une augmentation substantielle (d'un facteur 3 à 5) du bruit interne additif dans le domaine de l'enveloppe. Ces résultats suggèrent à nouveau: i) que les mécanismes de stockage et de prise de décision impliqués dans l'intégration temporelle demeurent intacts malgré une atteinte neurosensorielle; ii) un déficit d'efficacité de traitement pour l'AM à la suite d'une lésion cochléaire.

En conclusion, l'étude de la détection et de l'intégration des indices d'AM corrobore notre approche théorique, telle qu'implémentée dans le modèle PEMO (Dau et al. 1997), à condition de contraindre les capacités de traitement d'enveloppe par des limitations mnésiques.

11.2. Synthèse sur le traitement de FM:

11.2.1. Effet de la cadence de modulation

La sensibilité aux indices de FM diminue lorsque la cadence de modulation augmente (FMDLs plus mauvais à 20 Hz qu'à 2 Hz), corroborant a priori l'hypothèse d'une inertie (ou "sluggishness") des mécanismes impliqués dans le traitement de la FM (Moore & Sek, 1996; Moore & Glasberg, 1987a; Gockel et al., 2001). Ces résultats sont conformes aux FMDLs obtenus dans les études précédentes (Moore & Sek, 1996; Ernst and Moore, 2010, 2012) lorsque la détection de FM est mesurée en présence d'une AM interférente, que ce soit en égalisant la durée de stimulation ou le nombre de cycles présentés. Ces résultats contrastent toutefois avec la meilleure sensibilité observée en AM pour une cadence de modulation de 20 Hz lorsque les signaux sont égalisés en durée de stimulation. Cette différence dans le pattern de résultat renforce l'hypothèse selon laquelle la détection de FM ne repose pas exclusivement sur l'utilisation des indices dynamiques de pattern d'excitation (des indices d'enveloppe) issus de la conversion FM/AM.

11.2.2. Effet du nombre de cycles

Pour l'ensemble des groupes testés, les seuils de détection de FM s'améliorent significativement lorsque le nombre de cycles augmente, conformément aux résultats d'Hartmann et Klein (1980). L'intégration temporelle apparait surtout lorsque le nombre de cycles de modulation est relativement faible (N < 4). Les FMDLs atteignent ensuite une valeur asymptotique. Contrairement à la détection d'AM, la taille de l'effet d'intégration temporelle observée en FM est identique aux deux cadences de modulation. Cela suggère que le processus d'intégration temporelle est indépendant du processus de "sluggishness" (Moore & Sek, 1995, 1996; Sek & Moore, 1995) proposé de façon à expliquer les conditions pour lesquelles les indices de TFS seraient utilisables. Le mécanisme responsable du "slugishness" selon lequel les changements de fréquence instantanée au sein d'un cycle de modulation ne peuvent être détectables que s'ils sont suffisamment lents pour permettre l'extraction des indices de TFS est donc distinct du processus responsable de l'intégration temporelle résultant d'une combinaison des informations issues des différents cycles successifs de FM.

11.2.1. Effet du vieillissement

Ce programme de recherche constitue la seule étude ayant évalué l'effet de l'âge sur la sensibilité à la FM pour des cadences de modulation lentes et rapides, en essayant de contrôler l'utilisation des indices de pattern d'excitation (par la présence d'une AM interférente). Nos données empiriques montrent un effet délétère sélectif de l'âge sur la détection de FM, limité à la cadence de modulation la plus faible. Cet effet délétère du vieillissement est globalement conforme aux données de la littérature (Grose & Mamo, 2012; Het et al., 2008; Whiteford et al., 2017; Paraouty et al., 2016; Kortlang et al., 2016). Cette dégradation sélective des FMDLs renforce l'hypothèse selon laquelle: i) la détection de FM à basse cadence de modulation repose pour partie sur l'utilisation deuxième indice, probablement la TFS; ii) le vieillissement dégrade plus fortement la sensibilité à ce deuxième indice.

Comme pour l'AM, le vieillissement ne modifie pas substantiellement l'intégration temporelle en FM. L'intégration temporelle à 20 Hz est légèrement plus forte chez les sujets âgés que chez les sujets jeunes. Ces résultats suggèrent donc que le vieillissement préserve certains aspects de l'efficacité de traitement centrale (stockage et prise de décision) pour la détection de FM.

11.2.1. Effet de la perte auditive

Nos données empiriques démontrent qu'une lésion cochléaire dégrade significativement la sensibilité à la FM aux deux cadences de modulation (e.g., Moore & Skrodzka, 2002; Lacher-Fougère & Demany, 1998; Buss et al., 2004; Strelcyk & Dau, 2009; Ernst & Moore, 2012; Paraouty et al., 2016). Cette altération de la sensibilité à la FM associée à une perte neurosensorielle peut être attribuable à: i) une réduction de la conversion FM/AM consécutivement à la perte de sélectivité fréquentielle; ii) un appauvrissement des capacités d'encodage et/ou de décodage des indices de TFS.

Du point de vue de l'intégration temporelle, l'amélioration des FMDLs lorsque le nombre de cycles de modulation augmente est identique chez les sujets malentendants et chez les sujets normo-entendants, suggérant que les capacités d'intégration des indices de FM sont préservées malgré une atteinte cochléaire. Ainsi, comme pour l'AM, la perte auditive préserve les processus impliqués dans l'intégration temporelle (mémoire à court terme, prise de décision, etc.).

11.2.1. Ce que la modélisation nous a appris

Le modèle computationnel de traitement d'enveloppe temporelle (dont les paramètres sont fortement contraints par l'ajustement aux données d'AM) est utilisé pour prédire les FMDLs. Il permet de rendre compte de certains aspects des données empiriques:

i. l'effet de la cadence de modulation et l'effet du nombre de cycles de modulation résulte, comme pour l'AM: a) du filtrage passe haut introduit de façon à modéliser l'adaptation neurale à court terme et; b) des limitations mnésiques (bruit mnésique additif) et du processus de prise de décision (appariement de gabarit);

ii. l'effet différentiel de l'AM interférente sur les seuils de détection de FM aux deux cadences de modulation est correctement prédit par le modèle computationnel. Cet effet différentiel de l'AM interférente, initialement interprété comme résultant de l'existence d'un double code pour la détection de la FM et d'une inertie des mécanismes impliqués dans le décodage des indices de TFS, peut être simulé à partir d'un modèle utilisant uniquement les indices d'enveloppe. La modélisation suggère que l'effet d'interférence de l'AM sur la FM résulterait de la conversion FM/AM et du masquage d'AM (Houtgast, 1989). La plus forte interférence de l'AM à 20 Hz est observée et s'explique par la perte de sensibilité à la phase d'enveloppe pour les filtres de modulation accordés sur des cadences supérieures à 10 Hz. Cette perte de sensibilité à la phase d'enveloppe pourrait limiter la capacité du processus d'appariement de gabarit à extraire de façon optimale les indices d'AM issus de la conversion FM/AM.

iii. effet du vieillissement ou de la perte auditive: comme pour l'AM, le modèle suggère que: a) l'effet délétère du vieillissement sur les seuils de détection de FM à 20 Hz résulte d'une dégradation de l'efficacité de traitement centrale; b) l'altération globale des FMDLs à 20 Hz suite à une lésion cochléaire provient a minima d'un effet combiné de la perte de sélectivité fréquentielle et d'un déficit de l'efficacité de traitement centrale dans le domaine de l'enveloppe.

11.3. Comparaison des traitements d'AM et de FM

Sur le *plan psychophysique*, l'étude conjointe de l'AM et de la FM, et plus précisement l'étude de l'effet de la cadence de modulation et du nombre de cycles de modulation sur les AMDLs et les FMDLs chez différentes populations permet d'observer:

- ✓ Une dissociation entre les effets du vieillissement et de la perte auditive sur la sensibilité à l'AM et à la FM. En ce qui concerne l'AM, l'âge dégrade globalement la sensibilité aux indices d'enveloppe quelle que soit la cadence de modulation tandis que la perte auditive l'améliore substantiellement. En ce qui concerne la FM, le vieillissement affecte sélectivement le

traitement des cadences de modulation tandis que la perte auditive dégrade globalement la sensibilité à la FM.

- ✓ Un effet du nombre de cycle différent en AM et en FM. Les seuils AMDLs s'améliorent plus fortement que les seuils FMDLs lorsque le nombre de cycles de modulation augmente. De plus, le processus d'intégration temporelle est observé pour l'ensemble des nombres de cycles testés (*N*= 2-9) en AM, tandis que les FMDLs diminuent lorsque N augmente, avant de converger vers une valeur asymptotique au-delà de 4 cycles de modulation. Enfin, l'effet du nombre de cycles est similaire en FM, quelle que soit la cadence de modulation, alors que l'intégration temporelle observée en AM est plus importante à 20 Hz qu'à 2 Hz.
- ✓ Une dissociation entre l'effet de l'âge et de la perte auditive sur l'intégration temporelle. Le vieillissement n'affecte pas l'intégration temporelle en AM et en FM tandis que la perte cochléaire améliore l'intégration temporelle en AM et préserve l'intégration temporelle en FM.

Pris ensemble, ces éléments corroborent l'hypothèse de codes distincts intervenant dans la détection d'AM et de FM. En effet, les seuls indices issus du pattern d'excitation ne permettent pas d'expliquer: i) la dégradation sélective des FMDLs avec l'âge; ii) l'intégration temporelle plus faible en FM qu'en AM; iii) l'absence de sur-intégration chez les patients malentendants en FM; iv) l'amélioration de la sensibilité à l'AM et la dégradation de la sensibilité à la FM chez le sujet malentendant; v) l'absence d'effet de la cadence de modulation sur l'intégration temporelle en FM. La détection de FM, si elle repose pour partie sur les indices issus de la conversion FM/AM pour des cadences de modulation rapides, ne peut s'expliquer que par l'existence d'indices autres à basse cadence de modulation, à savoir des indices probablement basés sur l'utilisation de l'information de TFS.

Sur le plan *computationnel*, un modèle de traitement d'enveloppe temporelle basé sur le concept du banc de filtres de modulation et sur un mécanisme d'appariement de gabarit a été ajusté de façon à rendre compte des données d'AM. Ce modèle, qui est de fait très contraint, permet de

simuler schématiquement les données à 20 Hz, comme attendu si la conversion FM/AM a lieu. Par contre, le modèle ne permet pas de rendre compte des FMDLs à basse cadence de modulation (2 Hz) chez l'ensemble des sujets. De plus, les seuils simulés prédisent une sur-intégration en FM pour la plupart des conditions expérimentales, quelque soit la cadence de modulation. Cette sur-intégration conduit le modèle à dépasser les performances humaines dans de nombreuses conditions. Le modèle de traitement d'enveloppe temporelle échoue donc à rendre compte des données empiriques de FM.

S'il est concevable d'ajuster plus finement les prédictions du modèle à la détection et l'intégration en FM à 20 Hz (en modifiant la sélectivité fréquentielle et/ou en rendant le mécanisme d'appariement de gabarit moins efficace), il est pour autant fortement probable que l'on ne puisse rendre compte de la détection de FM à 2 Hz pour l'ensemble des groupes avec les mêmes paramètres. Pris ensemble, ces éléments suggèrent que la détection de FM repose sur deux mécanismes distincts.

De plus, notre modèle rend mieux compte des FMDLs obtenus à 2 Hz chez les sujets âgés et malentendants lorsque la détection de FM est réalisée en présence d'une AM interférente plutôt que sans AM interférente. Ce résultat est de prime abord surprenant dans la mesure où l'AM interférente a été introduite de façon à contraindre l'utilisation des indices de pattern d'excitation. Toutefois, l'introduction de l'AM ne permet pas d'abolir totalement l'utilisation des indices de pattern d'excitation, comme le démontre les relativement bonnes prédictions du modèle lorsque les FMDLs sont mesurés en présence d'une AM interférente. Dans la condition où la FM est présentée sans AM, les résultats suggèrent que les capacités d'encodage et/ou d'utilisation des indices de TFS, bien que dégradées, ne sont pas abolies chez les sujets âgés et malentendants (les seuils réels étant meilleurs que ceux prédits par simulation). Lorsque les FMDLs sont mesurés en présence d'une AM interférente, les simulations suggèrent que les capacités d'extraction des indices de TFS sont meilleures en l'absence d'AM interférente chez les sujets âges et malentendants (les seuils prédits étant meilleurs que les données réelles). Ces

éléments peuvent être interprétés comme: i) révélant une dépendance du phase locking sur la TFS vis-à-vis de l'intensité de stimulation, i.e. une diminution de la précision du phase locking à faible intensité (Palmer & Russell, 1986 ; Ernst & Moore, 2010 ; Paraouty et al., unpublished); ii) l'existence d'un mécanisme spatio-temporel, ie. les informations de phase locking sur la TFS ne seraient exploitables par le système auditif périphérique que si et seulement si le codage d'enveloppe est intact (e.g., Carney et al., 2002; Shamma, 1985).

Pour finir, les données empiriques montrent que: i) l'intégration temporelle en AM est plus importante que l'intégration temporelle mesurée en FM, quel que soit le groupe testé: ii) les pattern d'intégration temporelle diffèrent substantiellement en AM et en FM. Les simulations démontrent que le pattern d'intégration temporelle résulte: i) du processus d'appariement de gabarit ou du mécanisme d'observations multiples qui améliore la sensibilité à l'AM et la FM; ii) des limitations mnésiques qui dégradent la sensibilité aux modulations aux longues durées de stimulation.

Ces différences d'intégration temporelle suggérent l'existence de mécanismes différents pour le traitement d'AM et de FM: Existe t-il des mécanismes décisionnels distincts pour l'AM et la FM? Le mécanisme décisionnel impliqué dans la détection de FM bénéficie t-il moins d'une augmentation du nombre d'observations? Les structures mnésiques impliquées dans la détection des modulations sont -elles spécifiques à l'AM et la FM? Des mesures des capacités de stockage à court terme pour la perception d'AM et de FM (basées par exemple sur la méthode de discrimination retardée, développée par Clément et al., 1999), permettront de répondre à cette dernière question.

A ce jour, aucune étude n'a tenté de caractériser les structures mnésiques impliquées dans la détection d'AM et de FM, ainsi que les fonctions d'oubli de la mémoire auditive pour ces deux types de modulation. Différents travaux (Semal & Demany, 1991, 1993; Semal et al., 1996; Deutsch, 1972; Starr & Pitt, 1997) ont démontré que la hauteur et la sonie d'un son pourraient être traités par des modules séparés dans la mémoire auditive à court terme. En raison de la

probable différence des structures mnésiques impliquées dans la perception de la hauteur et de la sonie (Clement et al., 1999), on peut raisonnablement supposer que l'AM et la FM pourraient faire appel respectivement à des mémoires à court terme distinctes. Des travaux additionnels sont toutefois nécessaires pour confirmer cette hypothèse.

Concernant les processus centraux impliqués dans la détection de FM, les travaux sur l'adaptation à l'AM et la FM de Tansley et Suffield (1983) suggèrent que la FM est détectée par des canaux sélectifs. Bien que ceci ne change rien quant à la nature des mécanismes périphériques impliqués dans l'encodage d'une FM, une telle sélectivité dans le décodage de la FM, si elle était démontrée, pourrait contraindre le traitement central de FM et les capacités de stockage et de décision en FM, ainsi que les effets du vieillissement et de la perte auditive.

11.4. Implications en audiologie clinique

Ce programme de recherche précise les mécanismes impliqués dans le traitement de l'AM et de la FM ainsi que l'effet du vieillissement et de la perte auditive neurosensorielle sur la perception de ces modulations temporelles.

Quelles conséquences pour la compréhension des déficits d'intelligiblité de la parole chez le sujet malentendant? Nos résultats démontrent un effet délétère de la perte auditive sur la sensibilité à la FM, et ce pour les cadences de modulation lentes et rapides. L'impact d'une telle dégradation de la sensibilité à la FM sur l'intelligiblité de la parole a été évalué lors de précédents travaux (Lorenzi et al., 2012; cf. Annexe). Cette étude par vocodeur évalue les effets d'une perte auditive sur la capacité à utiliser des indices de parole de FM, et plus précisément la capacité à utiliser les indices de parole issus de la conversion de la FM en AM à la sortie des filtres cochléaires. Les résultats de cette étude montraient que, conformément à nos mesures psychophysiques et à nos simulations réalisées avec des stimuli de FM non-linguistiques, les capacités des sujets malentendants à utiliser des indices de FM (ainsi que ceux issus de la conversion FM/AM) en reconnaissance de la parole sont dégradées. Toutefois, ce déficit est

relativement modeste, et apparaît peu sensible à de fortes dégradations de la sélectivité fréquentielle. Cependant, cette étude ne permet pas de préciser la contribution des indices de FM autres que ceux issus de la conversion FM-AM à l'intelligiblité de la parole, ainsi que l'effet d'une perte auditive neurosensorielle sur la perception de ces indices.

Caractérisation de la presbyacousie: La presbyacousie (une perte auditive associée au vieillissement, diagnostiquée à l'audiométrie tonale, cf. ISO 7029), constitue la cause de surdité de perception la plus fréquente chez l'adulte. Nos données psychophysiques et de modélisation raffinent cette définition clinique. Elles ont en effet permis de dissocier l'effet du vieillissement (chez des sujets présentant des seuils audiométriques normaux) et les effets d'une perte auditive neurosensorielle sur la perception de l'AM et de la FM. Au final, nos travaux démontrent que la perte auditive: i) améliore la sensibilité à l'AM en préservant les capacités d'encodage des indices d'enveloppe temporelle; ii) dégrade substantiellement la sensibilité à la FM, peut-être en raison d'une altération des mécanismes impliqués dans l'encodage ou le décodage des indices de TFS. Les processus centraux impliqués dans l'intégration temporelle (capacités de stockage et mécanisme de prise de décision) semblent être préservés tant par le viellissement que par la perte auditive. Nos résultats de simulations informatiques suggèrent toutefois l'existence d'un déficit d'efficacité de traitement tant chez les sujets âgés que chez les sujets malentendants, modélisé par une réduction du rapport signal sur bruit dans le représentation interne du signal. L'origine de ce déficit d'efficacité de traitement centrale reste encore à préciser.

Diagnostic des troubles auditifs supraliminaires: La caractérisation des troubles auditifs induits par le vieillissement et la perte audditive neurosensorielle est incomplète si elle se limite à l'audiométrie tonale et/ou vocale. Différents travaux suggèrent que le vieillissement affecte précocément la sensibilité aux indices de TFS, tant sur des tâches monaurales que binaurales (e.g. Grose & Mamo, 2010, 2012). Hopkins et Moore (2010a) proposent le "TFS-LF" test de façon à évaluer l'intégrité des indices de TFS via une tâche de détection d'IPD. Nos données psychophysiques démontrent que le vieillissement affecte sélectivement la sensibilité à la FM à

basse cadence de modulation, avec toutefois une grande variabilité interindividuelle. Sur le plan clinique, la mesure du seuil de détection de FM à basse cadence de modulation chez le sujets âgés via un test monaural permettrait d'évaluer rapidement l'intégrité des mécanismes impliqués dans l'encodage/le décodage des indices de TFS.

Réhabilitation par prothèse auditive: Les aides auditives ne permettent pas de compenser directement les effets d'un déficit dans le traitement des indices de TFS. Cependant, elles pourraient apporter une aide indirecte, par exemple via l'utilisation de microphones directionnels ou d'algorithmes de réduction de bruit. Ces deux dispositifs visent à améliorer le rapport signal sur bruit, de façon à diminuer les effets du vieillissement et de la perte auditive. L'effet des réducteurs de bruit sur la sensibilité à l'AM et la FM a été évalué par différentes études (Ives et al., 2013; 2014). Ces travaux basés sur la réduction de bruit montrent que la détection d'AM et de FM complexes n'est pas améliorée par les réducteurs de bruit utilisant la soustraction spectrale chez les sujets malentendants. Ils corroborent les résultats obtenus par Alcantara et al. (2003), démontrant que la réduction de bruit ne permet pas d'améliorer significativement l'intelligibilité de la parole en environnement bruyant. D'autres algorithmes de réduction de bruit doivent donc être développés et testés.

12. Références

Alcántara, J. I., Moore, B. C., Kühnel, V., & Launer, S. (2003). Evaluation of the noise reduction system in a commercial digital hearing aid: *International Journal of Audiology*, *42*(1), 34–42.

Arai, T., & Greenberg, S. (1997). The temporal properties of spoken Japanese are similar to those of English. Proc. *EUROSPEECH*. '97, 1011-1014.

Arai, T., Pavel, M., Hermansky, H., & Avendano, C. (1996). Intelligibility of speech with filtered time trajectories of spectral envelopes. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on* (Vol. 4, p. 2490–2493). IEEE.

Ardoint, M., & Lorenzi, C. (2010). Effects of lowpass and highpass filtering on the intelligibility of speech based on temporal fine structure or envelope cues. *Hearing research*, *260*(1), 89–95.

Ardoint, M., Sheft, S., Fleuriot, P., Garnier, S., & Lorenzi, C. (2010). Perception of temporal fine-structure cues in speech with minimal envelope cues for listeners with mild-to-moderate hearing loss. *International Journal of Audiology*, *49*(11), 823-831.

Arlinger, S. D., Jerlvall, L. B., Ahren, T., & Holmgren, E. C. (1977). Thresholds for linear frequency ramps of a continuous pure tone. *Acta oto-laryngologica*, *83*(1-6), 317–327.

Aslin, R. N. (1989). Discrimination of frequency transitions by human infants. *The Journal of the Acoustical Society of America*, *86*(2), 582–590.

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.

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.

Bacon, S. P., & Viemeister, N. F. (1985). Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners. *Audiology*, *24*(2), 117–134.

Baker, R. J., & Rosen, S. (2006). Auditory filter nonlinearity across frequency using simultaneous notched-noise masking. *The Journal of the Acoustical Society of America*, *119*(1), 454–462.

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), 219–227.

Bartzokis, G. (2004). Age-related myelin breakdown: a developmental model of cognitive decline and Alzheimer's disease. *Neurobiology of aging*, *25*(1), 5–18.

Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., & Shinn-Cunningham, B. G. (2015). Individual differences reveal correlates of hidden hearing deficits. *Journal of Neuroscience*, *35*(5), 2161–2172.

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.

Boettcher, F. A., Madhotra, D., Poth, E. A., & Mills, J. H. (2002). The frequency-modulation following response in young and aged human subjects. *Hearing research*, *165*(1), 10–18.

Bracewell, R. N. (1986). *The Fourier transform and its applications* (Vol. 31999). McGraw-Hill New York.

Bruckert, L., Herrmann, M., & Lorenzi, C. (2006). No adaptation in the amplitude modulation domain in trained listeners. *The Journal of the Acoustical Society of America*, *119*(6), 3542–3545.

Burns, E. M., & Viemeister, N. F. (1981). Played-again SAM: Further observations on the pitch of amplitude-modulated noise. *The Journal of the Acoustical Society of America*, *70*(6), 1655–1660.

Buss, E., Hall III, J. W., & Grose, J. H. (2004). Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear and hearing*, *25*(3), 242–250.

Buss, E., Hall III, J. W., & Grose, J. H. (2006). Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. *The Journal of the Acoustical Society of America*, *120*(5), 2777–2788.

Buss, E., Hall, J. W., Grose, J. H., & Dev, M. B. (1999). Development of adult-like performance in backward, simultaneous, and forward masking. *Journal of Speech, Language, and Hearing Research*, *42*(4), 844–849.

Carlyon, R. P., Long, C. J., & Micheyl, C. (2012). Across-channel timing differences as a potential code for the frequency of pure tones. *Journal of the Association for Research in Otolaryngology*, *13*(2), 159–171.

Carney, L. H., Heinzy, M. G., Evilsizer, M. E., Gilkeyz, R. H., & Colburn, H. S. (2002). Auditory phase opponency: a temporal model for masked detection at low frequencies. *Acta Acustica United with Acustica*, *88*(3), 334–347.

Cazals, Y., Pelizzone, M., Saudan, O., & Boex, C. (1994). Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. *The Journal of the Acoustical Society of America*, *96*(4), 2048–2054.

Chen, H., & Zeng, F.-G. (2004). Frequency modulation detection in cochlear implant subjects. *The Journal of the Acoustical Society of America*, *116*(4), 2269–2277.

Colburn, H. S., Carney, L. H., & Heinz, M. G. (2003). Quantifying the information in auditory-nerve responses for level discrimination. *Journal of the Association for Research in Otolaryngology*, *4*(3), 294–311.

Colombo, J., & Horowitz, F. D. (1986). Infants' attentional responses to frequency modulated sweeps. *Child development*, 287–291.

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.

Dau, T., Püschel, D., & Kohlrausch, A. (1996a). A quantitative model of the ''effective''signal processing in the auditory system. I. Model structure. *The Journal of the Acoustical Society of America*, *99*(6), 3615–3622.

Dau, T., Püschel, D., & Kohlrausch, A. (1996b). 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.

Dau, T., Verhey, J., & Kohlrausch, A. (1999). Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers. *The Journal of the Acoustical Society of America*, *106*(5), 2752–2760.

Dawes, P., & Bishop, D. V. (2008). Maturation of visual and auditory temporal processing in school-aged children. *Journal of Speech, Language, and Hearing Research*, *51*(4), 1002–1015.

De Boer, E. (1985). Auditory time constants: A paradox? In *Time resolution in auditory systems* (p. 141–158). Springer. Consulté à l'adresse http://link.springer.com/chapter/10.1007/978-3-642-70622-6_9

Demany, L., & Semal, C. (1986). On the detection of amplitude modulation and frequency modulation at low modulation frequencies. *Acta Acustica united with Acustica*, *61*(4), 243–255.

Demany, L., & Semal, C. (1989). Detection thresholds for sinusoidal frequency modulation. *The Journal of the Acoustical Society of America*, *85*(3), 1295–1301.

Diedler, J., Pietz, J., Bast, T., & Rupp, A. (2007). Auditory temporal resolution in children assessed by magnetoencephalography. *NeuroReport*, *18*(16), 1691–1695.

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.

Dooley, G. J., & Moore, B. C. (1988). Detection of linear frequency glides as a function of frequency and duration. *The Journal of the Acoustical Society of America*, *84*(6), 2045–2057.

Dorman, M. F., Loizou, P. C., & Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *The Journal of the Acoustical Society of America*, *102*(4), 2403–2411.

Drullman, R. (1995). Temporal envelope and fine structure cues for speech intelligibility. *The Journal of the Acoustical Society of America*, *97*(1), 585–592.

Drullman, R., Festen, J. M., & Plomp, R. (1994a). Effect of reducing slow temporal modulations on speech reception. *The Journal of the Acoustical Society of America*, *95*(5), 2670–2680.

Drullman, R., Festen, J. M., & Plomp, R. (1994b). Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America*, *95*(2), 1053–1064.

Dubbelboer, F., & Houtgast, T. (2007). A detailed study on the effects of noise on speech intelligibility a. *The Journal of the Acoustical Society of America*, *122*(5), 2865–2871.

Dudley, H. (1939). Remaking speech. *The Journal of the Acoustical Society of America*, *11*(2), 169–177.

Dudley, H. W. (1938). *System for the artificial production of vocal or other sounds*.

Eddins, D. A. (1993). Amplitude modulation detection of narrow-band noise: Effects of absolute bandwidth and frequency region. *The Journal of the Acoustical Society of America*, *93*(1), 470–479.

Eddins, D. A. (1999). Amplitude-modulation detection at low-and high-audio frequencies. *The Journal of the Acoustical Society of America*, *105*(2), 829–837.

Edwards, B. W., & Viemeister, N. F. (1994a). Frequency modulation versus amplitude modulation discrimination: Evidence for a second frequency modulation encoding mechanism. *The Journal of the Acoustical Society of America*, *96*(2), 733-740.

Edwards, B. W., & Viemeister, N. F. (1994b). Psychoacoustic equivalence of frequency modulation and quasi-frequency modulation. *The Journal of the Acoustical Society of America*, *95*(3), 1510–1513.

Ernst, S. M., & Moore, B. C. (2010). Mechanisms underlying the detection of frequency modulation. *The Journal of the Acoustical Society of America*, *128*(6), 3642–3648.
Erulkar, S. D., Butler, R. A., & Gerstein, G. L. (1968). Excitation and inhibition in cochlear nucleus. II. Frequency-modulated tones. *Journal of Neurophysiology*, *31*(4), 537–548.
Evans, E. F. (1975). The sharpening of cochlear frequency selectivity in the normal and abnormal cochlea. *Audiology*, *14*(5-6), 419–442.
Ewert, S. D., & Dau, T. (2000). Characterizing frequency selectivity for envelope fluctuations. *The Journal of the Acoustical Society of America*, *108*(3), 1181–1196.
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.
Ewert, S. D., Verhey, J. L., & Dau, T. (2002). Spectro-temporal processing in the envelope-frequency domain. *The Journal of the Acoustical Society of America*, *112*(6), 2921–2931.
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.
Flanagan, J. L. (1972). Voices of men and machines. *The Journal of the Acoustical Society of America*, *51*(5A), 1375–1387.
Flanagan, J. L. (1980). Parametric coding of speech spectra. *The Journal of the Acoustical Society of America*, *68*(2), 412–419.
Florentine, M., & Buus, S. (1981). An excitation-pattern model for intensity discrimination. *The Journal of the Acoustical Society of America*, *70*(6), 1646–1654.
Formby, C. (1985). Differential sensitivity to tonal frequency and to the rate of amplitude modulation of broadband noise by normally hearing listeners. *The Journal of the Acoustical Society of America*, *78*(1), 70–77.
Formby, C. (1987). Modulation Threshold Functions for Chronically Impaired Menière Patients: Seuils de modulation et maladie de Menière chronique. *Audiology*, *26*(2), 89–102.
Formby, C. C. (1982). Differential sensitivity to tonal frequency and to the rate of modulation of broad-band noise by hearing-impaired listeners.
Formby, C., & Muir, K. (1988). Modulation and gap detection for broadband and filtered noise signals. *The Journal of the Acoustical Society of America*, *84*(2), 545–550.
Forrest, T. G., & Green, D. M. (1987). Detection of partially filled gaps in noise and the temporal modulation transfer function. *The Journal of the Acoustical Society of America*, *82*(6), 1933–1943.
Fu, Q.-J. (2002). Temporal processing and speech recognition in cochlear implant users. *Neuroreport*, *13*(13), 1635–1639.
Füllgrabe, C., Berthommier, F., & Lorenzi, C. (2006). Masking release for consonant features in temporally fluctuating background noise. *Hearing Research*, *211*(1), 74–84.
Füllgrabe, C., Meyer, B., & Lorenzi, C. (2003). Effect of cochlear damage on the detection of complex temporal envelopes. *Hearing research*, *178*(1), 35–43.
Füllgrabe, C., Moore, B. C., & Stone, M. A. (2014). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience*, *6*.

Gallun, F. J., Diedesch, A. C., & Beasley, R. (2012). Impacts of age on memory for auditory intensity a. *The Journal of the Acoustical Society of America*, *132*(2), 944–956.

Gates, G. A., Couropmitree, N. N., & Myers, R. H. (1999). Genetic associations in age-related hearing thresholds. *Archives of otolaryngology–head & neck surgery*, *125*(6), 654–659.

Gerken, G. M., Bhat, V. K., & Hutchison-Clutter, M. (1990). Auditory temporal integration and the power function model. *The Journal of the Acoustical Society of America*, *88*(2), 767–778.

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.

Gifford, R. H., & Bacon, S. P. (2005). Psychophysical estimates of nonlinear cochlear processing in younger and older listeners. *The Journal of the Acoustical Society of America*, *118*(6), 3823–3833.

Gifford, R. H., Bacon, S. P., & Williams, E. J. (2007). An examination of speech recognition in a modulated background and of forward masking in younger and older listeners. *Journal of Speech, Language, and Hearing Research*, *50*(4), 857–864.

Gilbert, G., Bergeras, I., Voillery, D., & Lorenzi, C. (2007). Effects of periodic interruptions on the intelligibility of speech based on temporal fine-structure or envelope cues. *The Journal of the Acoustical Society of America*, *122*(3), 1336–1339.

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.

Giraud, A.-L., Price, C. J., Graham, J. M., Truy, E., & Frackowiak, R. S. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, *30*(3), 657–664.

Glasberg, B. R., & Moore, B. C. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *The Journal of the Acoustical Society of America*, *79*(4), 1020–1033.

Glasberg, B. R., & Moore, B. C. (1988). Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech. *Scandinavian Audiology*.

Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing research*, *47*(1), 103–138.

Gleich, O., Hamann, I., Klump, G. M., Kittel, M., & Strutz, J. (2003). Boosting GABA improves impaired auditory temporal resolution in the gerbil. *Neuroreport*, *14*(14), 1877–1880.

Gnansia, D., Lazard, D. S., Léger, A. C., Fugain, C., Lancelin, D., Meyer, B., & Lorenzi, C. (2014). Role of slow temporal modulations in speech identification for cochlear implant users. *International journal of audiology*, *53*(1), 48–54.

Gockel, H., Moore, B. C., & Carlyon, R. P. (2001). Influence of rate of change of frequency on the overall pitch of frequency-modulated tones. *The Journal of the Acoustical Society of America*, *109*(2), 701–712.

Goldstein, J.L., & Srulovicz, P. (1977). Auditory-nerve spike intervals as an adequate basis for aural frequency measurement, In: Evans E.F, & Wilson, J.P., (Eds), *Psychophysics and Physiology of hearing*., London: Academic Press, pp. 337–347.

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.

Grant, K. W. (1987). Frequency modulation detection by normally hearing and profoundly hearing-impaired listeners. *J. Speech Hear. Res*, *30*, 558–563.
Grant, K. W., Summers, V., & Leek, M. R. (1998). Modulation rate detection and discrimination by normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *104*(2), 1051–1060.
Green, D. M. (1960). Auditory detection of a noise signal. *The Journal of the Acoustical Society of America*, *32*(1), 121–131.
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*, John Wiley, New York
Green, D. M., & Swets, J. A. (1974). *Signal Detection Theory and Psychophysics*, Krieger, Huntington, New York.
Green, D.M., & Forrest, T.G. (1988) Detection of amplitude modulation and gaps in noise. In: H. Duifhuis, H.P. Wit and J.P. Horst (Eds.), *Basic Issues in Hearing*, Academic Press, New York, USA.
Grose, J. H., & Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear and hearing*, *31*(6), 755.
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), 49–54.
Grose, J. H., Mamo, S. K., Buss, E., & Hall, J. W. (2015). Temporal processing deficits in middle age. *American journal of audiology*, *24*(2), 91–93.
Hall III, 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.
Hartmann, W. M. (1978). The effect of amplitude envelope on the pitch of sine wave tones. *The Journal of the Acoustical Society of America*, *63*(4), 1105–1113.
Hartmann, W. M. (1979). Detection of amplitude modulation. *The Journal of the Acoustical Society of America*, *65*(S1), S59–S59.
Hartmann, W. M., & Hnath, G. M. (1982). Detection of mixed modulation. *Acta Acustica united with Acustica*, *50*(5), 297–312.
Hartmann, W. M., & Klein, M. A. (1980). Theory of frequency modulation detection for low modulation frequencies. *The Journal of the Acoustical Society of America*, *67*(3), 935–946.
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.
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.
Heinz, M. G., Colburn, H. S., & Carney, L. H. (2001). Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. *Neural computation*, *13*(10), 2273–2316.
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.
Heinz, M. G., Swaminathan, J., Boley, J. D., & Kale, S. (2010). Across-fiber coding of temporal fine-structure: effects of noise-induced hearing loss on auditory-nerve

responses. In *The neurophysiological bases of auditory perception* (p. 621–630). Springer.
Heinz, M. G., & Young, E. D. (2004). Response growth with sound level in auditory-nerve fibers after noise-induced hearing loss. *Journal of neurophysiology*, *91*(2), 784–795.
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.
Helfer, K. S., & Vargo, M. (2009). Speech recognition and temporal processing in middle-aged women. *Journal of the American Academy of Audiology*, *20*(4), 264–271.
Henry, K. S., & Heinz, M. G. (2012). Diminished temporal coding with sensorineural hearing loss emerges in background noise. *Nature neuroscience*, *15*(10), 1362–1364.
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.
Hilbert, D. (1904). *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*. Leipzig: Teubner.
Hopkins, K., & Moore, B. C. (2010a). Development of a fast method for measuring sensitivity to temporal fine structure information at low frequencies. *International journal of audiology*, *49*(12), 940–946.
Hopkins, K., & Moore, B. C. (2010b). The importance of temporal fine structure information in speech at different spectral regions for normal-hearing and hearing-impaired subjects. *The Journal of the Acoustical Society of America*, *127*(3), 1595–1608.
Hopkins, K., & Moore, B. C. (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.
Hopkins, K., Moore, B. C., & Stone, M. A. (2008). Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *The Journal of the Acoustical Society of America*, *123*(2), 1140–1153.
Hopkins, K., Moore, B. C., & 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.
Houtgast, T. (1989). Frequency selectivity in amplitude-modulation detection. *The Journal of the Acoustical Society of America*, *85*(4), 1676–1680.
Houtgast, T., & Steeneken, H. J. M. (1984). A multi-language evaluation of the RASTI-method for estimating speech intelligibility in auditoria. *Acta Acustica united with Acustica*, *54*(4), 185–199.
Houtgast, T., & Steeneken, H. Jm. (1973). The modulation transfer function in room acoustics as a predictor of speech intelligibility. *Acta Acustica United With Acustica*, *28*(1), 66–73.
Humes, L. E., & Christopherson, L. (1991). Speech identification difficulties of hearing-impaired elderly personsthe contributions of auditory processing deficits. *Journal of Speech, Language, and Hearing Research*, *34*(3), 686–693.
Ives, D. T., Kalluri, S., Strelcyk, O., Sheft, S., Miermont, F., Coez, A., Bizaguet, E., 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.
Ives, D. T., Calcus, A., Kalluri, S., Strelcyk, O., Sheft, S., & Lorenzi, C. (2013). Effects of noise reduction on AM and FM perception. *Journal of the Association for Research in Otolaryngology*, 14(1), 149–157.
Jeffress, L. A. (1967). Stimulus-Oriented Approach to Detection Re-Examined. *The Journal of the Acoustical Society of America*, *41*(2), 480–488.
Jeffress, L. A. (1968). Mathematical and electrical models of auditory detection. *The Journal of the Acoustical Society of America*, *44*(1), 187–203.
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.
Johannesen, P. T., Pérez-González, P., Kalluri, S., Blanco, J. L., & 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 in hearing*, *20*, 2331216516641055.
Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *The Journal of the Acoustical Society of America*, *68*(4), 1115–1122.
Joosten, E. R., Shamma, S. A., Lorenzi, C., & Neri, P. (2016). Dynamic reweighting of auditory modulation filters. *PLoS computational biology*, *12*(7), e1005019.
Jørgensen, S., Ewert, S. D., & Dau, T. (2013). A multi-resolution envelope-power based model for speech intelligibility. *The Journal of the Acoustical Society of America*, *134*(1), 436–446.
Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological reviews*, *84*(2), 541–577.
Joris, P. X., & Verschooten, E. (2013). On the limit of neural phase locking to fine structure in humans. In *Basic Aspects of Hearing* (p. 101–108). Springer.
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.
Kale, S., & Heinz, M. G. (2012). Temporal modulation transfer functions measured from auditory-nerve responses following sensorineural hearing loss. *Hearing research*, *286*(1), 64–75.
Kay, R. H. (1982). Hearing of modulation in sounds. *Physiological Reviews*, *62*(3), 894–975.
Kay, R. H., & Matthews, D. R. (1972). On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency-modulated tones. *The Journal of Physiology*, *225*(3), 657–677.
Kettner, R. E., Feng, J. Z., & Brugge, J. F. (1985). Postnatal development of the phase-locked response to low frequency tones of auditory nerve fibers in the cat. *Journal of Neuroscience*, *5*(2), 275–283.
Kock, W. E. (1937). A new interpretation of the results of experiments on the differential pitch sensitivity of the ear. *The Journal of the Acoustical Society of America*, *9*(2), 129–134.
Koenig, W., Dunn, H. K., & Lacy, L. Y. (1946). The sound spectrograph. *The Journal of the Acoustical Society of America*, *18*(1), 19–49.

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.

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.

Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, *115*(1), 362–378.

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.

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.

Kumar, U., AV, S., & others. (2011). Temporal processing abilities across different age groups. *Journal of the American Academy of Audiology*, *22*(1), 5–12.

Lacher-Fougère, S., & Demany, L. (1998). Modulation detection by normal and hearing-impaired listeners. *Audiology*, *37*(2), 109–121.

Lamar, E. S., Hecht, S., Shlaer, S., & Hendley, C. D. (1947). Size, Shape, and Contrast in Detection of Targets by Daylight Vision. I. Data and Analytical Description1. *JOSA*, *37*(7), 531–545.

Lawson, J. L., & Uhlenbeck, G. E. (1950). Threshold signals. New York: Mc Graw-Hill

Lee, J. (1994). Amplitude modulation rate discrimination with sinusoidal carriers. *The Journal of the Acoustical Society of America*, *96*(4), 2140–2147.

Lee, J., & Bacon, S. P. (1997). Amplitude modulation depth discrimination of a sinusoidal carrier: Effect of stimulus duration. *The Journal of the Acoustical Society of America*, *101*(6), 3688–3693.

Léger, A. C., Desloge, J. G., Braida, L. D., & Swaminathan, J. (2015). The role of recovered envelope cues in the identification of temporal-fine-structure speech for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *137*(1), 505–508.

Levi, E. C., Folsom, R. C., & Dobie, R. A. (1995). Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults. *Hearing research*, *89*(1), 21–27.

Levi, E. C., & Werner, L. A. (1996). Amplitude modulation detection in infancy: Update on 3-month-olds. *Association for Research in Otolaryngology*, *19*, 142.

Liberman, M. C., & Dodds, L. W. (1984). Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. *Hearing research*, *16*(1), 55–74.

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.

Liégeois-Chauvel, C., Lorenzi, C., Trébuchon, A., Régis, J., & Chauvel, P. (2004). Temporal envelope processing in the human left and right auditory cortices. *Cerebral Cortex*, *14*(7), 731–740.

Lobarinas, E., Salvi, R., & Ding, D. (2013). Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas. *Hearing research*, *302*, 113–120.

Logan, B. F. (1977). Information in the zero crossings of bandpass signals. *Bell Labs Technical Journal*, *56*(4), 487–510.

Lorenzi, C., Debruille, L., Garnier, S., Fleuriot, P., & Moore, B. C. (2009). Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal. *The Journal of the Acoustical Society of America*, *125*(1), 27–30.

Lorenzi, C., Dumont, A., & Fullgrabe, 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., Gatehouse, S., & Lever, C. (1999). Sound localization in noise in normal-hearing listeners. *The Journal of the Acoustical Society of America*, *105*(3), 1810–1820.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. (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.

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.

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.

Lorenzi, C., Wallaert, N., Gnansia, D., Leger, A. C., Ives, D. T., Chays, A., Cazals, Y. (2012). Temporal-envelope reconstruction for hearing-impaired listeners. *Journal of the Association for Research in Otolaryngology*, *13*(6), 853–865.

Lunner, T., Hietkamp, R. K., Andersen, M. R., Hopkins, K., & Moore, B. C. (2012). Effect of speech material on the benefit of temporal fine structure information in speech for young normal-hearing and older hearing-impaired participants. *Ear and hearing*, *33*(3), 377–388.

Lutman, M. E., Gatehouse, S., & Worthington, A. G. (1991). Frequency resolution as a function of hearing threshold level and age. *The Journal of the Acoustical Society of America*, *89*(1), 320–328.

Madden, J. P., & Fire, K. M. (1996). Detection and discrimination of gliding tones as a function of frequency transition and center frequency. *The Journal of the Acoustical Society of America*, *100*(6), 3754–3760.

Maiwald, D. (1967). Die berechnung von modulations schwellen mit hilfe eines funktions schemas. *Acustica*, *18*(193).

Makary, C. A., Shin, J., Kujawa, S. G., Liberman, M. C., & Merchant, S. N. (2011). Age-related primary cochlear neuronal degeneration in human temporal bones. *Journal of the Association for Research in Otolaryngology*, *12*(6), 711–717.

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.

Meddis, R. (1986). Simulation of mechanical to neural transduction in the auditory receptor. *The Journal of the Acoustical Society of America*, *79*(3), 702–711.

Meddis, R. (1988). Simulation of auditory–neural transduction: Further studies. *The Journal of the Acoustical Society of America*, *83*(3), 1056–1063.

Meddis, R., O'Mard, L. P., & Lopez-Poveda, E. A. (2001). A computational algorithm for computing nonlinear auditory frequency selectivity. *The Journal of the Acoustical Society of America*, *109*(6), 2852–2861.

Mehraei, G., Gallun, F. J., Leek, M. R., & Bernstein, J. G. (2014). Spectrotemporal modulation sensitivity for hearing-impaired listeners: Dependence on carrier center frequency and the relationship to speech intelligibility. *The Journal of the Acoustical Society of America*, *136*(1), 301–316.

Miller, G. A., & Taylor, W. G. (1948). The perception of repeated bursts of noise. *The Journal of the Acoustical Society of America*, *20*(2), 171–182.
Møller, A. R. (1976). Dynamic properties of excitation and two-tone inhibition in the cochlear nucleus studied using amplitude-modulated tones. *Experimental brain research*, *25*(3), 307–321.
Moody, D. B., Cole, D., Davidson, L. M., & Stebbins, W. C. (1984). Evidence for a reappraisal of the psychophysical selective adaptation paradigm. *The Journal of the Acoustical Society of America*, *76*(4), 1076–1079.
Moore, B. C. (2007). *Cochlear hearing loss: physiological, psychological and technical issues*. John Wiley & Sons.
Moore, B. C. (2014). *Auditory processing of temporal fine structure: Effects of age and hearing loss*. World Scientific.
Moore, B. C., & Ernst, S. M. (2012). Frequency difference limens at high frequencies: Evidence for a transition from a temporal to a place code. *The Journal of the Acoustical Society of America*, *132*(3), 1542–1547.
Moore, B.C., & Glasberg, B. R. (1986). The role of frequency selectivity in the perception of loudness, pitch and time, in Moore, B.C., (Ed), *Frequency Selectivity in Hearing*,. Academic, London.
Moore, B. C., & Glasberg, B. R. (1987a). Factors affecting thresholds for sinusoidal signals in narrow-band maskers with fluctuating envelopes. *The Journal of the Acoustical Society of America*, *82*(1), 69–79.
Moore, B. C., & Glasberg, B. R. (1987b). Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. *Hearing research*, *28*(2), 209–225.
Moore, B. C., & Glasberg, B. R. (1989). Mechanisms underlying the frequency discrimination of pulsed tones and the detection of frequency modulation. *The Journal of the Acoustical Society of America*, *86*(5), 1722–1732.
Moore, B. C., Glasberg, B. R., Plack, C. J., & Biswas, A. K. (1988). The shape of the ear's temporal window. *The Journal of the Acoustical Society of America*, *83*(3), 1102–1116.
Moore, B. C. J. (1976). Comparison of frequency DL's for pulsed tones and modulated tones. *British Journal of Audiology*, *10*(1), 17–20.
Moore, B. C., & Oxenham, A. J. (1998). Psychoacoustic consequences of compression in the peripheral auditory system. *Psychological review*, *105*(1), 108.
Moore, B. C., & Sek, A. (1994). Effects of carrier frequency and background noise on the detection of mixed modulation. *The Journal of the Acoustical Society of America*, *96*(2), 741–751.
Moore, B. C., & Sek, A. (1995). Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation). *The Journal of the Acoustical Society of America*, *97*(4), 2468–2478.
Moore, B. C., & 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.
Moore, B. C., & Sek, A. (2009). Development of a fast method for determining sensitivity to temporal fine structure. *International journal of audiology*, *48*(4), 161–171.
Moore, B. C., & Sęk, A. (2009). Sensitivity of the human auditory system to temporal fine structure at high frequencies. *The Journal of the Acoustical Society of America*, *125*(5), 3186–3193.

Moore, B. C., Shailer, M. J., & Schooneveldt, G. P. (1992). Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss. *British journal of audiology*, *26*(4), 229–237.

Moore, B. C., & 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.

Moore, B. C., Vickers, D. A., & Mehta, A. (2012). The effects of age on temporal fine structure sensitivity in monaural and binaural conditions. *International journal of audiology*, *51*(10), 715–721.

Moore, B. C., 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.

Munson, W. A. (1947). The growth of auditory sensation. *The Journal of the Acoustical Society of America*, *19*(4), 584–591.

Nabelek, I., & Hirsh, I. J. (1969). On the discrimination of frequency transitions. *The Journal of the Acoustical Society of America*, *45*(6), 1510–1519.

Neher, T., Lunner, T., Hopkins, K., & Moore, B. C. (2012). Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners (L) a. *The Journal of the Acoustical Society of America*, *131*(4), 2561–2564.

Nelson, D. A., & Schroder, A. C. (1997). Linearized response growth inferred from growth-of-masking slopes in ears with cochlear hearing loss. *The Journal of the Acoustical Society of America*, *101*(4), 2186–2201.

Nelson, P. B., & Jin, S.-H. (2004). Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *115*(5), 2286–2294.

Nelson, P. B., Jin, S.-H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *113*(2), 961–968.

Noordhoek, I. M., & Drullman, R. (1997). Effect of reducing temporal intensity modulations on sentence intelligibility. *The Journal of the Acoustical Society of America*, *101*(1), 498–502.

Otsuka, S., Furukawa, S., Yamagishi, S., Hirota, K., & Kashino, M. (2016). Relation between cochlear mechanics and performance of temporal fine structure-based tasks. *Journal of the Association for Research in Otolaryngology*, *17*(6), 541–557.

Oxenham, A. J. (1996). Temporal integration as a function of masker bandwidth. *The Journal of the Acoustical Society of America*, *99*(4), 2565–2574.

Oxenham, A. J., & Simonson, A. M. (2006). Level dependence of auditory filters in nonsimultaneous masking as a function of frequency. *The Journal of the Acoustical Society of America*, *119*(1), 444–453.

Ozimek, E., & Sęk, A. (1988). AM difference limens for noise bands. *Acta Acustica united with Acustica*, *66*(3), 153–160.

Palmer, A. R., & 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. *Hearing research*, *24*(1), 1–15.

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.

Paraouty, N., & Lorenzi, C. (2017). Using individual differences to assess modulation-processing mechanisms and age effects. *Hearing research*, *344*, 38–49.
Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *The Journal of the Acoustical Society of America*, *59*(3), 640–654.
Patterson, R. D., Allerhand, M. H., & Giguere, C. (1995). Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform. *The Journal of the Acoustical Society of America*, *98*(4), 1890–1894.
Patterson, R. D., Nimmo-Smith, I., Weber, D. L., & Milroy, R. (1982). The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *The Journal of the Acoustical Society of America*, *72*(6), 1788–1803.
Penner, M. J. (1978). A power law transformation resulting in a class of short-term integrators that produce time-intensity trades for noise bursts. *The Journal of the Acoustical Society of America*, *63*(1), 195–202.
Peters, R. W., & Moore, B. C. (1992). Auditory filter shapes at low center frequencies in young and elderly hearing-impaired subjects. *The Journal of the Acoustical Society of America*, *91*(1), 256–266.
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.
Pichora-Fuller, M. K., & Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends in amplification*, *10*(1), 29–59.
Pirenne, M. H. (1943). Binocular and uniocular threshold of vision. *Nature*.
Plack, C. J., & Carlyon, R. P. (1995). Differences in frequency modulation detection and fundamental frequency discrimination between complex tones consisting of resolved and unresolved harmonics. *The Journal of the Acoustical Society of America*, *98*(3), 1355–1364.
Plomp, R. (1983). The role of modulation in hearing. In *Hearing—Physiological bases and psychophysics.* Springer. pp. 270–276.
Plomp, R., & Bouman, M. A. (1959). Relation between hearing threshold and duration for tone pulses. *The Journal of the Acoustical Society of America*, *31*(6), 749–758.
Pollack, I. (1968). Detection of rate of change of auditory frequency. *Journal of experimental psychology*, *77*(4), 535.
Pumplin, J. (1985). Low-noise noise. *The Journal of the Acoustical Society of America*, *78*(1), 100–104.
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.
Püschel, D. (1988). *Prinzipien der zeitlichen Analyse beim Hören*, Doctoral thesis, University of Göttingen.
Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *The Journal of the Acoustical Society of America*, *114*(1), 446–454.
Qin, M. K., & Oxenham, A. J. (2006). Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. *The Journal of the Acoustical Society of America*, *119*(4), 2417–2426.
Rees, A., & Palmer, A. R. (1989). Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by

broadband noise. *The Journal of the Acoustical Society of America*, *85*(5), 1978–1994.
Rice, S. O. (1973). Distortion produced by band limitation of an FM wave. *Bell Labs Technical Journal*, *52*(5), 605–626.
Richards, V. M., Buss, E., & Tian, L. (1997). Effects of modulator phase for comodulation masking release and modulation detection interference. *The Journal of the Acoustical Society of America*, *102*(1), 468–476.
Robles, L., & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. *Physiological reviews*, *81*(3), 1305–1352.
Robles, L., Ruggero, M. A., & Rich, N. C. (1986). Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input–output functions, tuning curves, and response phases. *The Journal of the Acoustical Society of America*, *80*(5), 1364–1374.
Rodenburg, M. (1972). *Sensitivity of the auditory system to differences in intensity*, Doctoral thesis, Medical Faculty of Rotterdam.
Rodenburg, M. (1977). Investigations of temporal effects with amplitude modulated signals, In: Evans E.F, & Wilson, J.P., (Eds), *Psychophysics and Physiology of hearing*., London: Academic Press, pp. 429–437.
Rose, J. E., Hind, J. E., Anderson, D. J., & Brugge, J. F. (1971). Some effects of stimulus intensity on response of auditory nerve fibers in the squirrel monkey. *Journal of Neurophysiology*, *34*(4), 685–699.
Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions: Biological Sciences*, 367–373.
Rosen, S., & Howell, P. (2010). Signals and Systems for Speech and Hearing, 2nd edition: Emerald Press, London.
Ruggero, M. A. (1994). Cochlear Delays and Traveling Waves: Comments on 'Experimental Look at Cochlear Mechanics':[A. Dancer, Audiology 1992; 31: 301-312] Ruggero. *Audiology*, *33*(3), 131–142.
Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2011). Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proceedings of the National Academy of Sciences*, *108*(37), 15516–15521.
Saberi, K., & Hafter, E. R. (1995). A common neural code for frequency-and amplitude-modulated sounds. *Nature*, *374*(6522), 537.
Salvi, R. J., Wang, J., & Ding, D. (2000). Auditory plasticity and hyperactivity following cochlear damage. *Hearing research*, *147*(1), 261–274.
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), 23–33.
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*.
Schouten, M. E. H. (1985). Identification and discrimination of sweep tones. *Perception & psychophysics*, *37*(4), 369–376.
Schreiner, C. E., & Langner, G. (1988). Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *Journal of neurophysiology*, *60*(6), 1823–1840.
Schreiner, C. E., & Urbas, J. V. (1988). Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hearing research*, *32*(1), 49–63.
Schroeder, M. R. (1966). Vocoders: Analysis and synthesis of speech. *Proceedings of the IEEE*, *54*(5), 720–734.

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. (1995). Frequency discrimination as a function of frequency, measured in several ways. *The Journal of the Acoustical Society of America*, *97*(4), 2479–2486.
Sek, A., & Moore, B. C. (2000). Detection of quasitrapezoidal frequency and amplitude modulation. *The Journal of the Acoustical Society of America*, *107*(3), 1598–1604.
Sek, A., & Moore, B. C. (2002). Mechanisms of modulation gap detection. *The Journal of the Acoustical Society of America*, *111*(6), 2783–2792.
Sellick, P. M., Patuzzi, R., & Johnstone, B. M. (1982). Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. *The journal of the acoustical society of America*, *72*(1), 131–141.
Sergeant, R. L., & Harris, J. D. (1962). Sensitivity to unidirectional frequency modulation. *The Journal of the Acoustical Society of America*, *34*(10), 1625–1628.
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.
Shamma, S. A. (1985a). Speech processing in the auditory system I: The representation of speech sounds in the responses of the auditory nerve. *The Journal of the Acoustical Society of America*, *78*(5), 1612–1621.
Shamma, S. A. (1985b). 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.
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.
Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303.
Sheft, S., Ardoint, M., & Lorenzi, C. (2008). Speech identification based on temporal fine structure cues. *The Journal of the Acoustical Society of America*, *124*(1), 562–575.
Sheft, S., Shafiro, V., Lorenzi, C., McMullen, R., & Farrell, C. (2012). Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech perception. *Ear and hearing*, *33*(6), 709.
Sheft, S., & Yost, W. A. (1989). Detection and recognition of amplitude modulation with tonal carriers. *The Journal of the Acoustical Society of America*, *85*(S1), S121–S121.
Sheft, S., & Yost, W. A. (1990). Temporal integration in amplitude modulation detection. *The Journal of the Acoustical Society of America*, *88*(2), 796–805.
Siebert, W. M. (1965). Some implications of the stochastic behavior of primary auditory neurons. *Kybernetik*, *2*(5), 206–215.
Siebert, W. M. (1970). Frequency discrimination in the auditory system: Place or periodicity mechanisms? *Proceedings of the IEEE*, *58*(5), 723–730.
Smith, R. L., & Brachman, M. L. (1980). Response modulation of auditory-nerve fibers by AM stimuli: effects of average intensity. *Hearing research*, *2*(2), 123–133.
Smith, R. L., Brachman, M. L., & Frisina, R. D. (1985). Sensitivity of auditory-nerve fibers to changes in intensity: A dichotomy between decrements and increments. *The Journal of the Acoustical Society of America*, *78*(4), 1310–1316.
Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*(6876), 87–90.

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. *The Journal of the Acoustical Society of America*, *112*(2), 720–727.

Søndergaard, P. L., Decorsiere, R., & Dau, T. (2011). On the relationship between multi-channel envelope and temporal fine structure. In *Proceedings of the International Symposium on Auditory and Audiological Research*, Vol. 3, pp. 363–370.

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.

Stickney, G. S., Nie, K., & Zeng, F.-G. (2005). Contribution of frequency modulation to speech recognition in noise a. *The Journal of the Acoustical Society of America*, *118*(4), 2412–2420.

Stone, M. A., Füllgrabe, C., Mackinnon, R. C., & Moore, B. C. (2011). The importance for speech intelligibility of random fluctuations in "steady" background noise. *The Journal of the Acoustical Society of America*, *130*(5), 2874–2881.

Stone, M. A., Füllgrabe, C., & Moore, B. C. (2008). Benefit of high-rate envelope cues in vocoder processing: Effect of number of channels and spectral region. *The Journal of the Acoustical Society of America*, *124*(4), 2272–2282.

Stone, M. A., & Moore, B. C. (2003). Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *The Journal of the Acoustical Society of America*, *114*(2), 1023–1034.

Stone, M. A., Moore, B. C., & Füllgrabe, C. (2011). The dynamic range of useful temporal fine structure cues for speech in the presence of a competing talker. *The Journal of the Acoustical Society of America*, *130*(4), 2162–2172.

Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing a. *The Journal of the Acoustical Society of America*, *125*(5), 3328–3345.

Strickland, E. A., & Viemeister, N. F. (1996). Cues for discrimination of envelopes. *The Journal of the Acoustical Society of America*, *99*(6), 3638–3646.

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.

Swaminathan, J., Reed, C. M., Desloge, J. G., Braida, L. D., & Delhorne, L. A. (2014). Consonant identification using temporal fine structure and recovered envelope cues. *The Journal of the Acoustical Society of America*, *135*(4), 2078–2090.

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.

Tansley, B. W., & Suffield, J. B. (1983). Time course of adaptation and recovery of channels selectively sensitive to frequency and amplitude modulation. *The Journal of the Acoustical Society of America*, *74*(3), 765–775.

Trainor, L. J., Samuel, S. S., Desjardins, R. N., & Sonnadara, R. R. (2001). Measuring temporal resolution in infants using mismatch negativity. *Neuroreport*, *12*(11), 2443–2448.

Trainor, L. J., & Trehub, S. E. (1989). Aging and auditory temporal sequencing: Ordering the elements of repeating tone patterns. *Attention, Perception, & Psychophysics*, *45*(5), 417–426.

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.

Tsumura, T., Sone, T., & Nimura, T. (1973). Auditory detection of frequency transition. *The Journal of the Acoustical Society of America, 53(1),* 17-25.

Turner, R. E., & Sahani, M. (2011). Demodulation as probabilistic inference. *IEEE Transactions on Audio, Speech, and Language Processing*, *19*(8), 2398–2411.

Tyler, R. S., Wood, E. J., & Fernandes, M. (1983). Frequency resolution and discrimination of constant and dynamic tones in normal and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *74(4),* 1190-1199.

Van Buuren, R. A., Festen, J. M., & Houtgast, T. (1999). Compression and expansion of the temporal envelope: Evaluation of speech intelligibility and sound quality. *The Journal of the Acoustical Society of America*, *105*(5), 2903–2913.

Van Eyken, E., Van Camp, G., & Van Laer, L. (2007). The complexity of age-related hearing impairment: contributing environmental and genetic factors. *Audiology and Neurotology*, *12*(6), 345–358.

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.

Van Zanten, G.A. (1980). Temporal modulation transfer functions for intensity modulated noise bands," In van den Brink, G., & Bilse, F.A. (Eds), *Psychophysical, Physiological, and Behavioral Studies in Hearing*, Delft University Press.

Van Zanten, G. A., & Senten, C. J. J. (1983). Spectro-Temporal Modulation Transfer Function (STMTF) for various types of temporal modulation and a peak distance of 200 Hz. *The Journal of the Acoustical Society of America*, *74*(1), 52–62.

Varnet, L., Ortiz-Barajas, M. C., Guevara Erra, R., Gervain, J., & Lorenzi, C. (2017). A cross-linguistic study of speech modulation spectra. *The Journal of the Acoustical Society of America*, *141*(5), 3701–3702.

Viemeister, N. F. (1970a). Discrimination between Beats and a Pure Tone. *The Journal of the Acoustical Society of America*, *47*(1A), 119–119.

Viemeister, N. F. (1970b). Intensity discrimination: Performance in three paradigms. *Perception & Psychophysics*, *8*(6), 417–419.

Viemeister, N. F. (1977). Temporal Factors in Audition: A Systems Analysis Approach, In: Evans E.F, & Wilson, J.P., (Eds), *Psychophysics and Physiology of hearing*., London: Academic Press, pp. 419-428.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *The Journal of the Acoustical Society of America*, *66*(5), 1364–1380.

Viemeister, N. F., & Wakefield, G. H. (1989). Multiple looks and temporal integration. *The Journal of the Acoustical Society of America*, *86*(S1), S23–S23.

Viemeister, N. F., & Wakefield, G. H. (1991). Temporal integration and multiple looks. *The Journal of the Acoustical Society of America*, *90*(2), 858–865.

Voelcker, H. B. (1966). Toward a unified theory of modulation part I: Phase-envelope relationships. *Proceedings of the IEEE*, *54*(3), 340–353.

Wakefield, G. H., & Nelson, D. A. (1985). Extension of a temporal model of frequency discrimination: Intensity effects in normal and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *77*(2), 613–619.

Wakefield, G. H., & Viemeister, N. F. (1984). Selective adaptation to linear frequency-modulated sweeps: Evidence for direction-specific FM channels? *The Journal of the Acoustical Society of America*, *75*(5), 1588–1592.
Wakefield, G. H., & Viemeister, N. F. (1990). Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise. *The Journal of the Acoustical Society of America*, *88*(3), 1367–1373.
Wang, S., Xu, L., & Mannell, R. (2011). Relative contributions of temporal envelope and fine structure cues to lexical tone recognition in hearing-impaired listeners. *Journal of the Association for Research in Otolaryngology*, *12*(6), 783–794.
Werner, L. A. (1999). Forward masking among infant and adult listeners. *The Journal of the Acoustical Society of America*, *105*(4), 2445–2453.
Werner, L. A. (2007). Issues in human auditory development. *Journal of communication disorders*, *40*(4), 275–283.
Werner, L. A., Folsom, R. C., Mancl, L. R., & Syapin, C. L. (2001). Human auditory brainstem response to temporal gaps in noise. *Journal of Speech, Language, and Hearing Research*, *44*(4), 737–750.
Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., & Gillenwater, J. M. (1992). Infant auditory temporal acuity: Gap detection. *Child development*, *63*(2), 260–272.
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*, 1–15.
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.
Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child development*, 611–624.
Willott, J., & Lister, J. (2003). The aging auditory system: anatomic and physiologic changes and implications for rehabilitation. *International Journal of Audiology*, *42*, 2S3–2S10.
Wojtczak, M., & Viemeister, N. F. (2003). Suprathreshold effects of adaptation produced by amplitude modulation. *The Journal of the Acoustical Society of America*, *114*(2), 991–997.
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.
Xu, L., & Pfingst, B. E. (2003). Relative importance of temporal envelope and fine structure in lexical-tone perception (L). *The Journal of the Acoustical Society of America*, *114*(6), 3024–3027.
Xu, L., & Zheng, Y. (2007). Spectral and temporal cues for phoneme recognition in noise. *The Journal of the Acoustical Society of America*, *122*(3), 1758–1764.
Yates, G. K. (1987). Dynamic effects in the input/output relationship of auditory nerve. *Hearing research*, *27*(3), 221–230.
Yost, W. A., & Moore, M. J. (1987). Temporal changes in a complex spectral profile. *The Journal of the Acoustical Society of America*, *81*(6), 1896–1905.
Yost, W. A., & Sheft, S. (1989). Across-critical-band processing of amplitude-modulated tones. *The Journal of the Acoustical Society of America*, *85*(2), 848–857.

Yost, W. A., Sheft, S., & Opie, J. (1989). Modulation interference in detection and discrimination of amplitude modulation. *The Journal of the Acoustical Society of America*, *86*(6), 2138–2147.
Zurek, P. M., & Formby, C. (1981). Frequency-discrimination ability of hearing-impaired listeners. *Journal of Speech and Hearing Research*, *24*(1), 108.
Zwicker, E. (1952). Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenz-modulation eines Tones. *Acustica*, *2*(3), 125–133.
Zwicker, E. (1953). Die Veränderung der Modulationsschwellen durch verdeckende Töne und Geräusche. *Acustica*, *3*(4), 274–278.
Zwicker, E. (1956). Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs. *Acustica*, *6*(4), 365–381.
Zwicker, E., (1970). Masking and psychological excitation as consequences of the ear's frequency analysis. In:Plomp, R., & Smoorenburg, G.F., (Eds.), *Frequency analysis and periodicity detection in hearing*, Sijthoff, Leiden, 376
Zwicker, E., & Graf, L. (1987). Modulationsschwellen bei Verdeckung. *Acustica*, *64*(3), 148–154.
Zwicker, E., & Kaiser, W. (1952). Der Verlauf der Modulationsschwellen in der Hörfläche. *Acustica*, *2*(6), 239–246.
Zwicker, E., & Fastl, H., (1990). *Psychacoustics - Facts and models*. Berlin: Springer Science & Business Media
Zwislocki, J. (1960). Theory of temporal auditory summation. *The Journal of the Acoustical Society of America*, *32*(8), 1046–1060.
Zwislocki, J. J. (1969). Temporal summation of loudness: An analysis. *The Journal of the Acoustical Society of America*, *46*(2B), 431–441.
Zwislocki, J. J., & Nguyen, M. (1999). Place code for pitch: A necessary revision. *Acta oto-laryngologica*, *119*(2), 140–145.

13. Annexe

JARO (2012)
DOI: 10.1007/s10162-012-0350-3
© 2012 Association for Research in Otolaryngology

JARO
Journal of the Association for Research in Otolaryngology

Research Article

Temporal-Envelope Reconstruction for Hearing-Impaired Listeners

Christian Lorenzi[1], Nicolas Wallaert[1], Dan Gnansia[2], Agnès Claire Leger[1], David Timothy Ives[1], André Chays[3], Stéphane Garnier[4], and Yves Cazals[5]

[1]*Equipe Audition (CNRS, Universite Paris Descartes, Ecole normale superieure), Institut d'Etude de la Cognition, Ecole normale superieure, Paris Sciences et Lettres, 29 rue d'Ulm, 75005 Paris, France*

[2]*Neurelec, 2720 Chemin Saint Bernard, 06224 Vallauris Cedex, France*

[3]*Hôpital Robert Debré, Reims, France*

[4]*Groupement Entendre SAS, Paris, France*

[5]*Université Paul Cézanne, UMR CNRS 6231, 13916 Marseille, France*

Received: 18 December 2011; Accepted: 9 September 2012

ABSTRACT

Recent studies suggest that normal-hearing listeners maintain robust speech intelligibility despite severe degradations of amplitude-modulation (AM) cues, by using temporal-envelope information recovered from broadband frequency-modulation (FM) speech cues at the output of cochlear filters. This study aimed to assess whether cochlear damage affects this capacity to reconstruct temporal-envelope information from FM. This was achieved by measuring the ability of 40 normal-hearing listeners and 41 listeners with mild-to-moderate hearing loss to identify syllables processed to degrade AM cues while leaving FM cues intact within three broad frequency bands spanning the range 65–3,645 Hz. Stimuli were presented at 65 dB SPL for both normal-hearing listeners and hearing-impaired listeners. They were presented as such or amplified using a modified half-gain rule for hearing-impaired listeners. Hearing-impaired listeners showed significantly poorer identification scores than normal-hearing listeners at both presentation levels. However, the deficit shown by hearing-impaired listeners for amplified stimuli was relatively modest. Overall, hearing-impaired data and the results of a simulation study were consistent with a poorer-than-normal ability to reconstruct temporal-envelope information resulting from a broadening of cochlear filters by a factor ranging from 2 to 4. These results suggest that mild-to-moderate cochlear hearing loss has only a modest detrimental effect on peripheral, temporal-envelope reconstruction mechanisms.

Keywords: speech, hearing loss, envelope reconstruction, amplitude modulation, frequency modulation

INTRODUCTION

A wide range of studies suggest that amplitude-modulation (AM) cues play an important role in speech intelligibility (e.g., Drullman 1995; Shannon et al. 1995). However, additional studies conducted with normal-hearing (NH) or hearing-impaired (HI) listeners have suggested that frequency-modulation (FM) cues may also be important for speech intelligibility when speech cues are severely degraded by background noise, interfering voices, filtering, infinite compression, peak clipping, or gating (e.g., Zeng et al. 2005; Gilbert and Lorenzi 2006, 2010; Lorenzi et al. 2006, 2009; Drennan et al. 2007; Ardoint and Lorenzi 2010; Gnansia et al. 2009, 2010; Hopkins et al. 2008, 2010; Ardoint et al. 2011; Eaves et al. 2011).

Correspondence to: Christian Lorenzi · Equipe Audition (CNRS, Universite Paris Descartes, Ecole normale superieure), Institut d'Etude de la Cognition · Ecole normale superieure, Paris Sciences et Lettres · 29 rue d'Ulm, 75005 Paris, France. Telephone: +33 144322674; fax: +33 144322680; email: lorenzi@ens.fr

Published online: 25 September 2012

Psychoacoustical investigations (e.g., Saberi and Hafter 1995; Moore and Sek 1996) proposed that the perception of FM cues is mediated by two early auditory mechanisms: (1) a purely temporal mechanism based on phase-locking in auditory-nerve fibers, extracting so-called temporal fine-structure cues from FM, and (2) a second mechanism based on an AM (or so-called temporal envelope) reconstruction process. In this case, the differential attenuation of cochlear filtering converts the frequency excursions of FM into AM fluctuations at the output of cochlear filters.

In agreement with the second mechanism, psychoacoustical, electrophysiological, and modeling studies showed that the narrow-band speech AM patterns can be reconstructed from the broadband FM components as a result of cochlear filtering (Ghitza 2001; Zeng et al. 2004; Gilbert and Lorenzi 2006; Sheft et al. 2008; Heinz and Swaminathan 2009; Ibrahim and Bruce 2010; Swaminathan and Heinz 2012; see also Apoux et al. 2011). Gilbert and Lorenzi (2006) demonstrated that recovered AM cues contribute substantially to the identification of vocoded syllables retaining FM cues only when a wide analysis filter bandwidth is used to generate the vocoded signals. More precisely, their results showed that recovered AM cues played a major role in the identification of FM-vocoded syllables when the bandwidth of analysis filters used to vocode speech was greater than about four times the bandwidth of a normal cochlear filter (i.e., >4 ERB_N, where ERB_N stands for the average equivalent rectangular bandwidth of the cochlear filter as determined using young normally hearing listeners tested at moderate sound levels; cf. Glasberg and Moore 1990; Moore 2007). Consistent with these behavioral findings, the correlation between the original and the recovered AM components computed at the output of a simulated cochlear filterbank was also found to increase with the bandwidth of analysis bands (Gilbert and Lorenzi 2006; Sheft et al. 2008).

Thus, when only FM speech cues are presented to NH listeners within broad frequency bands, the recovered temporal-envelope information may serve as a strong cue for speech identification and help maintain robust speech intelligibility. HI listeners typically show normal or near-normal ability to use AM speech cues (e.g., Turner et al. 1995; Baskent 2006; Lorenzi et al. 2006, 2009). However, they may not be able to recover temporal-envelope information from FM as well as NH listeners do because cochlear filters tend to broaden with the amount of hearing loss (see Moore 2007, for a review). This prediction is supported qualitatively by the results of a modeling study reporting simulations of a computational model of auditory-nerve fibers' responses to a vocoded speech signal retaining broadband FM cues only (Heinz and Swaminathan 2009). Simulations obtained with an auditory-nerve fiber model tuned to a low center frequency (550 Hz) showed that cochlear envelopes recovered from FM cues extracted from one, two, or four broad frequency bands spanning the range 80–8820 Hz may be reduced up to about 40 % following a 30-dB hearing loss due to selective outer-hair-cell damage.[1] The model tuning curves were based on those obtained for NH cats. Modeling work performed by Ibrahim and Bruce (2010) indicated that the model of cat AM tuning underestimates envelope recovery in humans, for which frequency tuning was estimated to be two to three times sharper than cats (Recio et al. 2002; Shera et al. 2002). For this reason, the simulation data of Heinz and Swaminathan (2009) cannot be taken as precise predictions of the effect of cochlear damage on envelope recovery in humans. Nevertheless, they suggest that human subjects with cochlear hearing loss and the associated reduction in frequency selectivity should show poorer-than-normal vocoded-speech identification performance if a wide analysis filter bandwidth is used to generate the vocoded speech signals retaining FM cues only. The present study aimed to test this hypothesis and quantify precisely the potential deficits in speech-envelope reconstruction by measuring the ability of human listeners with mild to moderate hearing loss to identify vocoded speech signals retaining broadband FM cues only.

METHODS

The first experiment assessed the capacity of NH and HI listeners to identify nonsense syllables processed to retain FM cues only within three, ~7–8-ERB_N-wide analysis bands spanning the range 65–3,645 Hz.

A second experiment was conducted on NH listeners to assess the capacity to identify the FM-vocoded syllables on the sole basis of the putative envelope cues recovered at the output of auditory filters with normal (1-ERB_N wide) or broader-than-normal (2- or 4-ERB_N wide) bandwidths.

[1] A neural cross-correlation coefficient (ρenv) was used to predict the similarity of temporal-envelope cues between single auditory-nerve (AN) fiber spike-train responses to an intact speech token and corresponding FM-vocoded version. Recovery of speech envelope cues was demonstrated by a large envelope correlation ρ_{env} between the neural responses to the original speech token and to the corresponding FM-vocoded version. Simulation of a 30-dB hearing loss due to selective outer-hair-cell damage in the cochlear model (which corresponds to a broadening in cochlear tuning by a factor 1.4–2.3) resulted in a 40 % reduction of the ρ_{env} metric quantifying the amount of envelope reconstruction (see Heinz and Swaminathan 2009 for more details).

Participants

Forty NH and 41 HI listeners participated. The study was approved by CPP Ile de France (07018—ID RCB, 2007-A00343-50). Audiograms were measured for each listener using a Madsen Aurical +3.08 audiometer and TDH39 headphones in a single-wall sound-attenuating booth for the tested ear (right ear for all listeners). Pure-tone average (PTA) estimates were calculated for audiometric frequencies between 0.125 and 4 kHz for each individual NH and HI listener.

The listeners with normal hearing showed audiometric thresholds not exceeding 20 dB hearing level (HL) between 0.125 and 4 kHz. Their mean audiogram is shown in Figure 1 (dashed line). The listeners with normal hearing were considered as a single NH group according to the current experimental design. However, they were divided into two groups—young [n=16; mean age=24 years; standard deviation (SD)= 3 years; range=19–29 years; mean PTA=4.4 dB HL] and older (n=24; mean age=46 years; SD=10 years; range=31–63 years; mean PTA=6.5 dB HL) listeners —for further statistical analysis because of differences in their audiometric configuration. Young NH listeners showed audiometric thresholds not exceeding 20 dB HL up to (and including) 8 kHz. Fifteen among 24 older NH listeners showed audiometric thresholds not exceeding 20 dB HL up to (and including) 8 kHz; the remaining nine older NH listeners showed higher audiometric thresholds at 6 and 8 kHz (range=–5 to 50 dB HL).

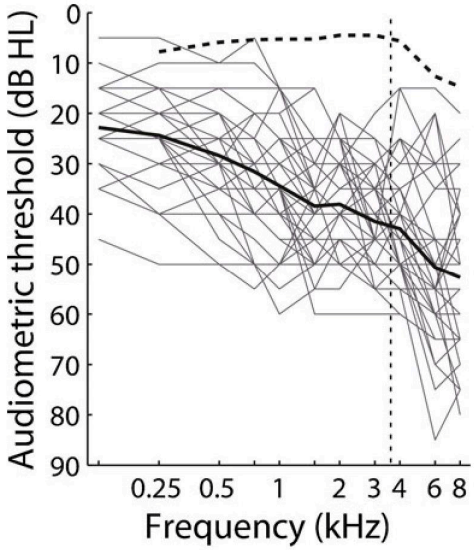


FIG. 1. Pure-tone audiograms for the tested (i.e., right) ear of the 41 HI listeners. The *bold continuous line* shows the mean audiogram of the 41 HI listeners. The *bold dashed line* shows the mean audiogram of the NH listeners. The *vertical thin dashed line* shows the high cutoff frequency (3.6 kHz) of the speech stimuli used in the present experiment.

The listeners with hearing loss were aged between 28 and 70 years (n=41; mean age=58 years). Their audiograms [individual (gray lines) and mean data (bold line)] are shown in Figure 1. Air-conduction, bone-conduction, and impedance audiometry for the HI listeners were consistent with sensorineural impairment. HI listeners showed audiometric thresholds not exceeding 60 dB HL between 0.125 and 4 kHz. Their PTA calculated for audiometric frequencies between 0.125 and 4 kHz ranged between 20 and 48 dB HL (mean PTA=35 dB HL; SD=6 dB HL), indicating that the hearing loss was mild to moderate (Goodman 1965).

All speech stimuli were delivered monaurally to their right ear via Sennheiser HD280 headphones. For NH listeners, the presentation level of the stimuli was 65 dB SPL. Overall levels were calibrated using a Bruel & Kjaer (2250) sound level meter and ear simulator (B&K 4153) complying with IEC 60318-1. For each HI listener, presentation level was adjusted individually according to the "Cambridge formula" (Moore and Glasberg 1998), which corresponds to a modified half-gain linear amplification rule aiming to limit the upward spread of masking produced by amplification. Maximum presentation level never exceeded 90 dB SPL.

Speech identification tasks

Speech material. Forty-eight nonsense vowel-consonant-vowel-consonant-vowel (VCVCV) stimuli were spoken in quiet by a French female talker (mean VCVCV duration= 1,272 ms; standard deviation=113 ms), recorded digitally via a 16-bit A/D converter at a 44.1-kHz sampling frequency and equalized in root-mean-square (rms) power. These VCVCV stimuli consisted of three recordings of 16 /aCaCa/ utterances (C=/p, t, k, b, d, g, f, s, ʃ, v, z, j, m, n, r, l/). The fundamental frequency of the female voice was estimated to be 219 Hz using the algorithm of de Cheveigne and Kawahara (2002).

Speech-processing conditions. The original speech signals were submitted to three different processing schemes. Stimuli processed using the first scheme (referred to as "3-band AM+FM") contained AM and FM information from the speech waveform. Stimuli processed using the second scheme (referred to as "3-band FM") contained speech information in their FM only, but temporal-envelope cues were potentially recoverable at the output of auditory filters. Stimuli were finally generated using a third scheme (referred to as "Recov") so as to force listeners to identify consonants primarily on the basis of the envelope cues recovered at the output of simulated auditory filters from the 3-band FM stimuli. Three-band AM +FM and 3-band FM stimuli were presented for identification to NH and HI listeners. Recov stimuli were presented for identification to NH listeners only.

Each /aCaCa/ signal was initially bandpass filtered using fourth-order gammatone filters (Patterson 1987) to generate the 3-band AM+FM and 3-band FM stimuli. For both processing conditions, the filterbank consisted of three, ~7–8-ERB_N-wide contiguous analysis bands spanning the range of 65 to 3,645 Hz.[2] The three analysis bands were centered on 263, 940, and 2,535 Hz, respectively. Their bandwidths were 395, 936, and 2,220 Hz, respectively. The three analysis bands encompassed the energy of the fundamental frequency (band 1), the first and second formant (band 2), and the third formant (band 3) of the French vowel /a/ used in this study, respectively. A speech-shaped noise (SSN) was generated whose spectrum matched the average spectrum of the whole set of /aCaCa/ stimuli. This SSN was passed through a highpass Butterworth filter (cutoff frequency=3,645 Hz, slope=108 dB/octave) and added to the processed speech stimuli (see below) to prevent the use of information from the transition band (Warren et al. 2004). The overall level of the highpass filtered noise was 12 dB below the overall level of the speech stimuli. The filtered noise was gated on and off with the target speech item, and shaped using a raised-cosine function with 10-ms rise/fall times. A different sample of the filtered noise was used for each presentation.

Three-band AM+FM condition. The Hilbert transform was applied to the bandpass-filtered speech signal in order to decompose the latter into its AM (modulus of the Hilbert analytic signal) and FM (cosine of the argument of the Hilbert analytic signal). For each band, the AM and FM components were multiplied. Impulse responses were peak-aligned for the AM (using a group delay of 16 ms) and the FM components across frequency channels (Hohmann 2002). The modulated signals were finally weighted and summed over the three frequency bands. The weighting compensated for imperfect superposition of the bands' impulse responses at the desired group delay. The weights were optimized numerically to achieve a flat frequency response (Hohmann 2002). The stimuli were equated in root mean square (rms) power.

Three-band FM condition. The Hilbert transform was applied to the bandpass-filtered speech signal in order to decompose the latter into its AM and FM. The AM component was modified by replacing all amplitudes above a level of 20 dB below the band rms level by +1 and all amplitudes below that level by 0. This was achieved to avoid the amplification of low-level recording noise to the same level as the target speech (a signal-processing artifact that might have reduced speech intelligibility in previous studies conducted with HI listeners; see Hopkins et al. 2010). The modified AM was then lowpass-filtered (cutoff frequency=128 Hz; 56 dB/octave rolloff) using a zero-phase Butterworth filter to smooth fast transitions in the AM, and multiplied with the FM component. As for the AM+FM condition, impulse responses were peak-aligned for the AM and the FM components across frequency channels. The modulated signals were finally weighted and summed over the three frequency bands. The resulting stimuli were equated in rms power.

The effects of speech processing and degraded auditory frequency selectivity are illustrated in Figures 2 and 3 for a single speech utterance (/ababa/). The highpass-filtered SSN was not added to the processed speech signals for clarity. The left and right panels of Figure 2 show the waveform of the /ababa/ utterance in the 3-band AM+FM and 3-band FM conditions, respectively. Figure 2 shows that speech processing for the 3-band FM condition behaves like infinite peak clipping or like a very fast, multi-channel compressor with an infinite compression ratio (cf. Licklider and Pollack 1948; Gilbert and Lorenzi 2006). Figure 3 displays the temporal envelopes computed by a model of early auditory processing in response to the 3-band AM+FM (top panel) and 3-band FM (bottom panels) /ababa/ stimuli. Temporal envelopes were computed at the output of 80 gammatone filters tuned between 80 and 3758 Hz, via half-wave rectification and lowpass filtering at 20 Hz (Butterworth filter, rolloff=12 dB/octave).

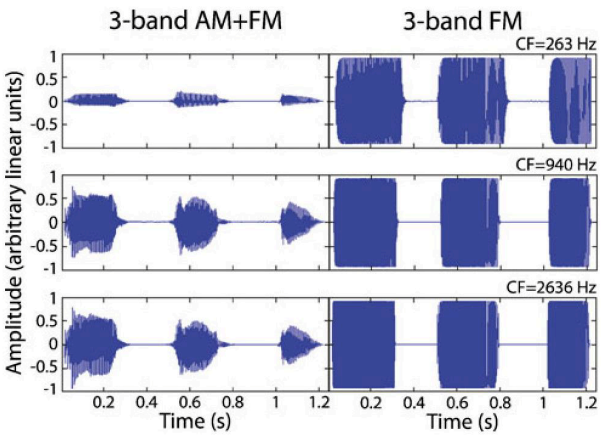


FIG. 2. Waveforms of a single /ababa/ utterance in the 3-band AM+FM (*left panel*) and 3-band FM (*right panel*) conditions. Waveforms are shown for each analysis band of center frequency CF.

[2] More precisely, speech sounds were initially split into four frequency bands spanning the range 65–8,958 Hz, and the fourth band centered on 6,325 Hz was omitted. Gilbert and Lorenzi (2006) and Heinz and Swaminathan (2009) used a comparable frequency range to generate four-band vocoded stimuli: 80–8,020 Hz in Gilbert and Lorenzi (2006) and 80–8,820 Hz in Heinz and Swaminathan (2009). As a consequence, the "4-band vocoder" conditions were somewhat comparable in Gilbert and Lorenzi (2006), Heinz and Swaminathan (2009), and in the present study.

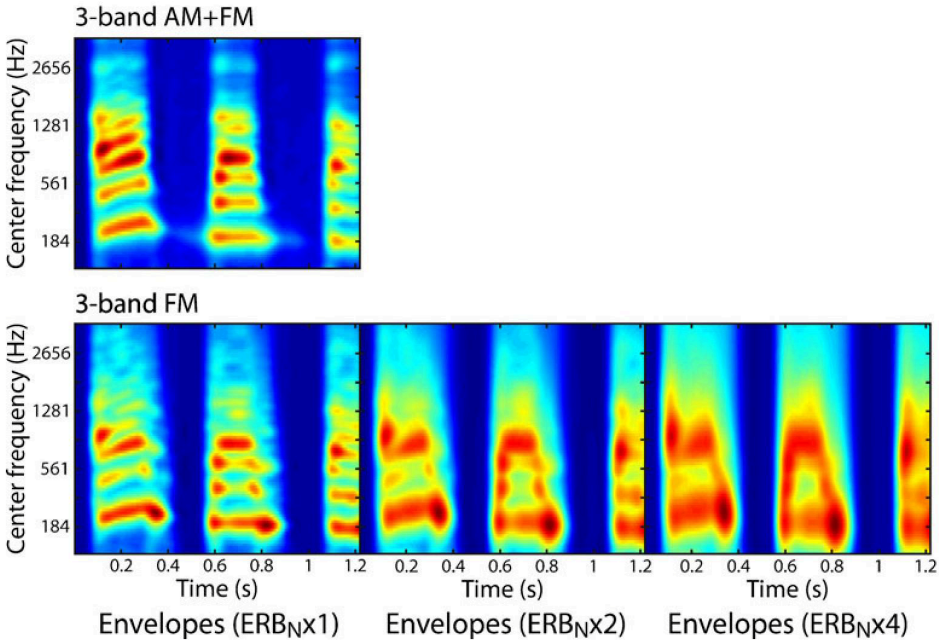


FIG. 3. Temporal envelopes of a single /ababa/ utterance in the 3-band AM+FM (*top panel*) and 3-band FM (*bottom panels*) conditions. Temporal envelopes are computed at the output of 80 1-ERB_N-wide (*top panel*, and *left panel, bottom row*), 2-ERB_N-wide (*middle panel, bottom row*), and 4-ERB_N-wide (*right panel, bottom row*) gammatone filters. In each panel, the center frequency (CF) of the gammatone filter is plotted on the ordinate, and time is plotted on the abscissa. A color scale ranging from blue to red is used, with a higher value for the red component.

Temporal envelopes were computed using 1-ERB_N-wide (top panel and left panel, bottom row) gammatone filters to simulate normal auditory filters. The model also simulated the effects of broadening auditory filters by factors of 2 (middle panel, bottom row) and 4 (right panel, bottom row), which is representative of mild/moderate and severe hearing loss, respectively (e.g., Tyler et al. 1984; Glasberg and Moore 1986; Dubno and Dirks 1989; Peters and Moore 1992; Stone et al. 1992; see also Moore 2007 for a review). Figure 3 shows that the original temporal-envelope speech cues (top panel) are recovered from FM cues at the output of simulated auditory filters when the latter are 1-ERB_N wide (left panel, bottom row). The recovered temporal-envelope pattern shown in the left panel of the bottom row displays the harmonic structure, the formants and their transitions, and the phonetic/syllabic segments of the /ababa/ signal. As expected, the middle and right panels of the bottom row show that the temporal-envelope cues recovered at the output of simulated auditory filters are substantially smeared when the bandwidth of auditory filters is broadened by a factor ranging between 2 and 4. In this case, the formants and the syllabic segments of the /ababa/ signal are only signaled by gross (recovered) envelope features encompassing broad frequency regions.

Recov condition. The 3-band FM signals were passed through a bank of 32 gammatone auditory filters, each 1-, 2-, or 4-ERB_N wide with center frequencies ranging from 80 to 8,583 Hz, and spaced along an ERB scale. In each band, the temporal envelopes were extracted using the Hilbert transform and lowpass-filtered using a zero-phase Butterworth filter. The envelopes were lowpass-filtered at $ERB_N/2$ for analysis filters whose bandwidths were greater than 128 Hz. Otherwise, the envelopes were lowpass-filtered at 64 Hz. These envelopes were then used to amplitude-modulate sine waves having the same frequencies as the original center frequencies of the auditory filters, but with random starting phase. In order to avoid delays in envelopes induced by filtering, impulse responses were peak-aligned for the envelope (using a group delay of 16 ms) and the carrier across frequency channels (Hohmann 2002). This envelope alignment was applied to the different stimuli. The modulated signals were finally weighted and summed over the 32 frequency bands. The weighting compensated for imperfect superposition of the bands' impulse responses at the desired group delay. The weights were optimized numerically to achieve a flat frequency response (Hohmann 2002; see also Gnansia et al. 2008, 2009, 2010 for further details).

This experimental condition was included to assess the capacity of NH listeners to identify the FM-vocoded speech sounds on the basis of the envelope

cues recovered at the output of auditory filters with normal (1-ERB_N wide) or broader-than-normal (2- or 4-ERB_N wide) bandwidths. As indicated above, the PTA calculated for audiometric frequencies between 0.125 and 4 kHz ranged between 20 and 48 dB HL, indicating that the hearing loss of the present HI listeners was globally mild to moderate (Goodman 1965). However, at each audiometric frequency, pure-tone sensitivity varied widely across individuals. At low frequencies (≤0.5 kHz) where envelope reconstruction was shown to be most prominent in previous modeling studies (e.g., Sheft et al. 2008; Heinz and Swaminathan 2009), audiometric thresholds ranged between 5 and 50 dB HL across HI listeners. At mid-frequencies (0.75–3 kHz), audiometric thresholds ranged between 5 and 60 dB HL across HI listeners. Broadening factors by 2 and 4 (i.e., 2- or 4-ERB_N-wide gammatone filters) were thus chosen to match the average auditory-filter broadening reported in the literature (e.g., Moore 2007) for HI listeners with the same range of audiometric thresholds as in here (5–60 dB HL). Note that a broadening of auditory filters by a factor of 2 should be considered even for mild levels of hearing loss. For the low-frequency region and for hearing losses between 5 and 30 dB HL, Moore (2007) reported a maximum broadening by a factor of 1.7 at 0.4 and 0.8 kHz; Santurette and Dau (2007) reported a maximum broadening by a factor of 1.4 at 0.5 kHz, and Strelcyk and Dau (2009) reported a maximum broadening by a factor of 1.4 at 0.75 kHz. For the mid-frequency region, Badri et al. (2011) reported a maximum broadening by a factor of 1.7 at 2 kHz.

Testing procedure

Speech identification was conducted in a sound-attenuating booth. All listeners were tested in a single session lasting between 1 and 2 h. This session was divided into several blocks, each corresponding to a given experimental condition. Before the actual testing session, the listeners were familiarized with speech material by listening to each original (i.e., unprocessed) logatome three times. During each block, the 48 processed /aCaCa/stimuli were presented in random order. On each trial, a single /aCaCa/ stimulus was presented to the listener with the task to identify the /aCaCa/ stimulus among 16 alternatives displayed on a computer screen. The alternatives were letter strings corresponding to the nonsense syllables (e.g., "ababa", "adada", etc.). Listeners responded by selecting a given letter string on the screen with a computer mouse. Response time was not limited, and no feedback was provided to listeners. Percent correct scores were corrected for chance level (6.25 %). The correction used was: $PC\ corrected\ for\ chance = [(PC\ raw - Chance\ level)/(100 - Chance\ level)] \times 100$, where PC stands for percent correct scores (Baskent 2006). In the corrected scores, a performance at 0 % represents a performance at chance level. The corrected scores were then transformed to rationalized arcsine units (RAU; Studebaker 1985). RAU scores are similar to percent correct scores for values of the latter from 15 % to 85 %, but RAU scores are "stretched" outside that region (reducing the effects of the bounded percent-correct scale). Speech identification was always assessed first for 3-band AM+FM speech stimuli in order to familiarize all listeners with the experimental procedure in the less distorted speech-processing condition. Then, speech identification was assessed for 3-band FM speech stimuli. Finally, speech identification was assessed for Recov speech stimuli. However, for each speech-processing condition, HI listeners were tested in random order for the two levels of stimulus presentation. In the Recov condition, NH listeners were tested in random order for the three bandwidths of analysis filters (1-, 2-, and 4-ERB_N-wide filters). For each HI listener, testing lasted about 15 min per experimental condition.

RESULTS

Experiment 1

The left and middle panels of Figure 4 show individual (filled symbols) and mean (horizontal bar in each box) identification scores expressed in RAU for the NH (circles) and HI (squares) listeners. The identification data are shown for the 3-band AM+FM (left panel) and 3-band FM (middle panel) conditions. NH data are shown for stimuli presented at a level of 65 dB SPL. HI data are shown for stimuli presented at a level of 65 dB SPL (condition "HI") or for stimuli amplified linearly using the Cambridge formula (condition "HI+").

Figure 4 shows that for most experimental conditions, HI data were more scattered than NH data. When stimuli were presented at 65 dB SPL, the identification scores of two HI listeners were extremely poor and slightly greater than chance level (−16 RAU) for both speech-processing conditions. At both presentation levels, the identification scores of 15 NH and five HI listeners were at ceiling level (116 RAU) in the 3-band AM+FM speech-processing condition. Only the 3-band FM condition appeared totally immune to ceiling *and* floor effects for both NH and HI listeners when stimulus audibility was controlled via linear amplification.

For each group of listeners and presentation level, average identification scores decreased by 9 to 30 RAU when the AM component was removed within each analysis band. Still, average identification scores remained above 50 RAU (and thus, well above chance level; here −16 RAU) across groups and presentation

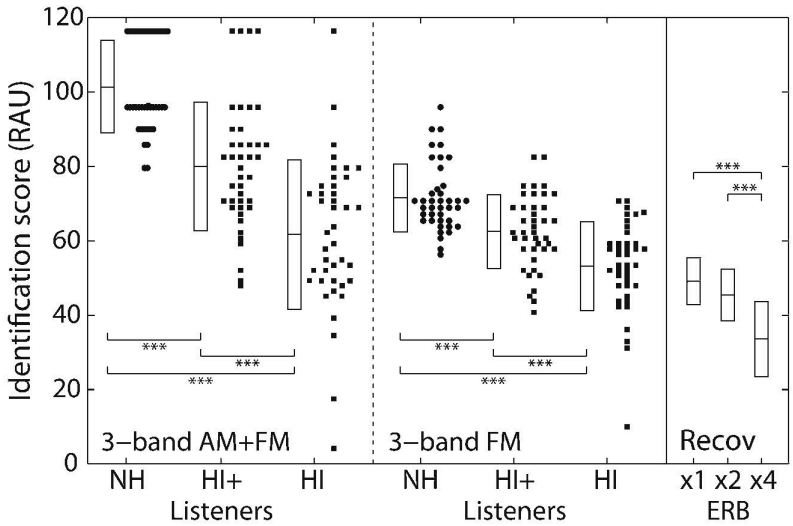


FIG. 4. Individual (*filled symbols*) and mean (*horizontal bar* in each box) identification scores (in RAU). *Open boxes* represent ±1 standard deviation about the mean for each condition. Percent-correct identification scores were corrected for chance level (6.25 % correct) and transformed to rationalized arcsine units (RAU). Floor and ceiling values correspond to −16 and 116 RAU, respectively. *Left* and *middle panels*: data are shown for the NH (*circles*) and HI (*squares*) listeners in the 3-band AM+FM (*left panel*) and 3-band FM (*middle panel*) conditions. NH data are shown for stimuli presented at a level of 65 dB SPL. HI data are shown for stimuli presented at a level of 65 dB SPL (*HI*) or amplified linearly (*HI+*). *Asterisks* (***) indicate that all comparisons of speech scores (NH/HI, NH/HI+, and HI/HI+) were significant at the 0.001 level. *Right panel*: mean identification scores obtained by NH listeners in the Recov condition, that is for 3-band FM speech passed through a tone-excited vocoder using 1-ERB_N-wide analysis filters ($ERB_N \times 1$), 2-ERB_N-wide analysis filters ($ERB_N \times 2$), or 4-ERB_N-wide analysis filters ($ERB_N \times 4$). Again, *asterisks* (***) indicate that comparisons of speech scores (ERB_N × $1/ERB_N$ × 4 and $ERB_N \times 2/ERB_N \times 4$) were significant at the 0.001 level.

levels after AM removal. For NH listeners, the removal of AM reduced average identification scores by 30 RAU. For HI listeners, the removal of AM had a greater deleterious effect when stimuli were amplified (17 RAU) than when they were presented at 65 dB SPL (9 RAU). The effect of removing AM was larger for NH listeners because of the stretching effect produced by the RAU transform on percent-correct identification scores (15 NH listeners being at ceiling performance for the 3-band AM+FM condition). For each speech-processing scheme, average scores measured for HI listeners increased when stimuli were submitted to linear amplification. This increase was two times greater for 3-band AM+FM stimuli (18 RAU) than for 3-band FM stimuli (9 RAU). However, for each presentation level, average scores of HI listeners remained poorer than normal (by 9 to 40 RAU) in each speech-processing condition. In the 3-band FM condition, the difference between NH and HI listeners' mean identification scores was two times smaller when stimuli were amplified (9 RAU) than when they were presented at 65 dB SPL (18 RAU).

Student t tests[3] (with Bonferroni correction) were conducted on the RAU scores obtained by NH and HI listeners using 3-band AM+FM stimuli. These analyses showed significant differences between RAU scores when comparing (1) NH and HI listeners' data at 65 dB SPL ($p<0.001$), (2) NH listeners' data measured at 65 dB SPL and HI listeners' data measured with amplified stimuli ($p<0.001$), and (iii) HI listeners' data measured with and without amplification ($p<0.001$).

Similar Student t tests (with Bonferroni correction) were then conducted on the RAU scores obtained by NH and HI listeners using 3-band FM stimuli (these two statistical analyses were conducted separately because listeners were always tested first with 3-band AM+FM stimuli). Again, statistical analyses showed significant differences between RAU scores when comparing (1) NH and HI listeners' data at 65 dB SPL ($p<0.001$), (2) NH listeners' data measured at 65 dB SPL and HI listeners' data measured with amplified stimuli ($p<0.001$), and (3) HI listeners' data measured with and without amplification ($p<0.001$).

In summary, both NH and HI listeners demonstrated on average a clear ability to identify consonants on the sole basis of the broadband FM speech cues. The ability of HI listeners to identify 3-band FM speech stimuli was significantly poorer than normal with or without linear amplification of speech stimuli. However, the overall deficit shown by HI listeners with amplified 3-band FM speech stimuli was modest,

[3] All statistical tests were performed using the R software (R Project).

amounting to 9 RAU. Interestingly, the effect of amplification was greater for 3-band AM+FM than for 3-band FM stimuli, suggesting that the ability to identify consonants processed to retain FM cues only was less affected by a change in stimulus audibility than the ability to identify consonants with intact modulation components. However, this result should be treated with caution because all listeners were tested first with 3-band AM+FM stimuli.

Experiment 2

The right panel of Figure 4 shows the mean scores (horizontal bar in each box) obtained by NH listeners in the Recov condition. For each Recov condition, the area around the mean identification score indicates ±1 standard deviation about the mean.

Consistent with previous work (e.g., Gilbert and Lorenzi 2006), NH listeners could identify 3-band FM speech stimuli on the basis of recovered envelope cues well above chance level. Mean identification scores were 49, 45, and 34 RAU for 1-, 2,- or 4-ERB_N-wide analysis filters, respectively. The mean identification score obtained with 1-ERB_N-wide analysis filters (49 RAU) was lower than the mean score obtained by NH listeners with 3-band FM stimuli (72 RAU), suggesting that factors other than peripheral envelope reconstruction may play a role in the identification of 3-band FM speech. Previous work suggests that the ability to use temporal fine-structure cues conveyed by patterns of phase locking in auditory-nerve fibers may have contributed to the identification of vocoded consonants (e.g., Lorenzi et al. 2006, 2009; Sheft et al. 2008; Heinz and Swaminathan 2009; Ardoint et al. 2010; however, see Swaminathan and Heinz 2012). On the other hand, it is likely that speech identification based on recovered envelope cues was underestimated with the present simulation approach because (1) human peripheral tuning may be two or more times sharper than previously estimated (Shera et al. 2002), such that envelope reconstruction may be stronger than originally assumed (Ibrahim and Bruce 2010), and (3) tone-vocoders degrade somewhat the transmission of envelope cues (Kates 2011). Thus, the *relative* pattern of performance obtained in the Recov condition (i.e., the effect of broadening analysis filters on identification scores) may be more relevant to the present investigation than absolute scores. Figure 4 shows that broadening analysis filters by a factor of 2 reduced identification scores by 4 RAU only, whereas broadening analysis filters by a factor of 4 reduced identification scores by 15 RAU. These data are in line with the modest deficits (9 RAU) observed for HI listeners with 3-band FM speech in the amplified condition (HI+) (as shown in the middle panel of Fig. 4).

A repeated measures analysis of variance (ANOVA) conducted on the RAU scores of the NH listeners in the three Recov conditions showed a significant effect of the broadening of the analysis filters [$F(2,78)=52$, $p<0.001$]. A post hoc analysis [Tukey HSD (honestly significant difference) test] showed that broadening analysis filters by a factor of 2 did not degrade significantly identification performance for NH listeners compared to the case where analysis filters were 1-ERB_N wide ($p=0.1$). However, broadening analysis filters by a factor of 4 reduced performance significantly (all $p<0.001$).

DISCUSSION

The ability to use recovered envelope cues

Mean identification scores measured for both NH and HI listeners using syllables processed to retain FM cues only were above 60 RAU—and thus, well above chance level (−16 RAU)—when stimulus audibility was controlled via linear amplification. This finding is consistent with the results of experiment 2 where mean identification scores based on the recovered envelope cues only were shown to range between about 35 and 50 RAU for NH listeners. These results suggest that, despite potentially important variations in auditory frequency selectivity (see Moore 2007), both NH listeners and listeners with mild-to-moderate hearing loss are able to maintain robust speech intelligibility in quiet by using temporal-envelope information recovered from the *broadband* FM speech cues. Note, however, that additional speech cues (such as temporal fine-structure cues) may have contributed to the identification of FM-vocoded speech identification for both NH and HI listeners.

The current study also showed that our listeners with mild-to-moderate cochlear hearing loss (mean PTA= 35 dB HL) were significantly worse than NH listeners in the 3-band FM speech condition. This suggested that their ability to reconstruct temporal-envelope cues was degraded by cochlear damage, consistent with the modeling work conducted by Heinz and Swaminathan (2009) with comparable vocoded-speech stimuli. The deficit shown by HI listeners was relatively small, amounting to 9 RAU for 3-band FM speech after stimulus amplification. Moreover, the effect of broadening analysis filters by a factor of 2 did not yield a substantial drop in 3-band FM speech identification performance for NH listeners (4 RAU). Altogether, these results indicate that HI listeners with mild-to-moderate hearing loss remain able to use temporal-envelope information recovered from broadband FM speech cues, although this capacity weakens because of degraded frequency selectivity.

A simulation study based on the correlational approach developed by Zeng et al. (2004), Gilbert and Lorenzi (2006), and Sheft et al. (2008) was conducted to better understand the limited effect of cochlear damage on the identification of 3-band FM speech stimuli. Quantitative estimates of the fidelity of temporal-envelope reconstruction were obtained by calculating the correlation between the temporal envelopes of the original speech stimuli and the 3-band FM speech stimuli at the output of 30 gammatone filters with center frequencies uniformly spaced along an ERB scale ranging from 80 to 3,500 Hz. In each band, the temporal envelopes were extracted using the Hilbert transform and lowpass filtering at $ERB_N/2$ Hz (using a 180th order finite impulse response filter). The correlation estimates were calculated for 1-, 2-, and 4-ERB_N-wide gammatone filters. For each condition, the correlation estimates were finally averaged across the entire set of 48 /aCaCa/ speech stimuli. A Fischer z transform was used before averaging correlations. The correlation estimates are reported in Figure 5 with results obtained with 1-, 2-, and 4-ERB_N-wide gammatone filters shown by bold, dashed, and dotted lines, respectively. The numbers in parentheses are the mean correlation across gammatone filters for each bandwidth of gammatone filters. Figure 5 shows that the mean correlation is relatively high (>0.5), and decreases only slightly as a function of gammatone filters' bandwidth. More precisely, the mean correlation is not affected by a broadening of gammatone filters by a factor of 2 (however, note that correlation estimates decrease slightly in the 200–500-Hz range). The mean correlation decreases only from 0.6 to 0.5 when the gammatone filters are broadened by a factor of 4. Most of the reduction in envelope reconstruction occurs for filters tuned at or close to the fundamental frequency of the female speaker (219 Hz). Similar results (not shown here) were obtained with other metrics assessing the fidelity of temporal-envelope transmission, such as the "depth-dependent" correlation index developed by Sheft et al. (2008) that takes into account changes in modulation depth at the output of gammatone filters (see Sheft et al. 2008 for further details).

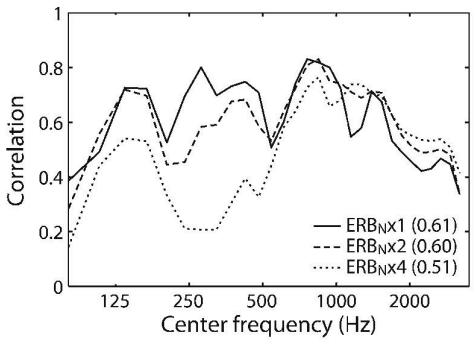


FIG. 5. Assessment of temporal-envelope reconstruction at the output of 1-ERB_N (*bold line*), 2-ERB_N (*dashed line*), and 4-ERB_N (*dotted line*)-wide gammatone filters. Mean correlation estimates between the original speech envelopes and the envelopes of the stimuli in the 3-band FM speech-processing conditions are shown as a function of gammatone filter center frequency. *Numbers* between *parentheses* correspond to the mean correlation estimate across channels in the respective gammatone-filter bandwidth condition.

These simulated data suggest that the loss of frequency selectivity typically associated with mild to moderate cochlear hearing loss should only have a modest deleterious effect on peripheral temporal-envelope reconstruction. This is in good agreement with the empirical data collected for most of our HI listeners. Thus, the differentiation process that converts FM excursions into temporal-envelope fluctuations at the output of cochlear filters appears hardly affected, if at all, by the reduction of the slope of auditory filters tuned to low/middle frequencies (when filters' bandwidth changes from 1 to 4 ERB_N, the low- and high-frequency slopes of gammatone filters decrease from about 80 to 10 dB/octave and from about 100 to 35 dB/octave, respectively). The global decrease in modulation depth of the reconstructed envelopes resulting from the broadening of auditory filters should have little impact on the reception of speech cues because auditory sensitivity to temporal-envelope cues is usually normal (or even better than normal) for HI listeners when audibility is carefully controlled for (Bacon and Gleitman 1992; Moore et al. 1992, 1996; Füllgrabe et al. 2003; Kale and Heinz 2010, 2012). The main deleterious effects of cochlear damage on the intelligibility of 3-band FM speech sounds would then correspond to (1) a reduction in the number of independent peripheral channels which transmit envelope information (see also Swaminathan and Heinz 2011) and (2) a reduction in the ability to discriminate and use envelope information for speech identification. Previous studies demonstrated that speech intelligibility in quiet is not affected by a reduction in frequency resolution by a factor of 2–4 (e.g., Baer and Moore 1993; Shannon et al. 1995; Baskent 2006). Other studies indicated that HI listeners may show poorer-than-normal AM discrimination abilities (Grant et al. 1998; Buss et al. 2004). This may partially explain why removing AM speech cues had a smaller detrimental effect on speech intelligibility for HI listeners compared to NH listeners in the present study. However, several studies showed that HI listeners are nearly as good as NH listeners at understanding AM-vocoded speech (e.g., Turner et al. 1995; Baskent 2006; Lorenzi et al. 2006). This suggests that HI listeners are as efficient as NH listeners at using complex temporal-envelope speech cues. Therefore, it is not surprising that the smearing of envelope cues illustrated in Figure 3 has only a small impact on consonant identification for the present HI listeners.

The small deficit shown by HI listeners contrasts with the dramatic deficits reported by Lorenzi et al. (2006) and Ardoint et al. (2010) for HI listeners with mild-to-moderate cochlear hearing loss. The speech material and speech processing in these studies differ in three aspects that, together, could account for the difference in HI listeners' performance between previous and current investigations. First, VCV stimuli were used by Lorenzi et al. (2006) and Ardoint et al. (2010), whereas VCVCV stimuli were used in the current study. Thus, all listeners were provided with more redundant consonantal cues in the present study, and this may have limited the detrimental effects of speech processing (i.e., FM vocoding) on HI listeners' performance. Secondly, analysis bands with no speech information at a particular time were filled with distracting background sound in Lorenzi et al. (2006) and Ardoint et al. (2010). As a result, most FM speech sounded harsh and noisy. This may have posed a particular problem to the HI listeners who would have suffered more from masking between channels because of their broadened cochlear filters (Moore 2008). In the present study, such masking effects were strongly limited because FM information was extracted in each analysis band only when the band signal was above −20 re: RMS (that is, when the band signal was assumed to correspond to speech instead of low-level recording noise). Finally, speech sounds were split into 16 2-ERB_N-wide analysis filters in Lorenzi et al. (2006) and Ardoint et al. (2010), versus three 8-ERB_N-wide analysis filters in the present investigation. Gilbert and Lorenzi (2006) showed that for NH listeners, the amount of envelope reconstruction expected at the output of cochlear filters (and the ability to use these reconstructed envelope speech cues) increased strongly with the bandwidth of the FM-vocoder analysis bands (and thus, with the bandwidth of FM speech cues). It is reasonable to assume that the negative impact of reduced frequency selectivity caused by cochlear damage on the amount of envelope reconstruction was smaller when envelope cues were reconstructed from broadband FM cues (as in the case of a 3-band FM vocoder) rather than from narrowband FM cues (as in the case of a 16-band FM vocoder). However, it is difficult to reconcile this explanation with the outcome of the simulation studies conducted by Heinz and Swaminathan (2009) and Ardoint et al. (2010), showing very limited effects of reduced frequency selectivity on envelope reconstruction for 16-band FM speech stimuli. Two explanations may be proposed: (1) the contribution of (neural) temporal fine-structure cues to the identification of FM vocoded speech sounds may vary inversely with the availability of reconstructed envelope cues (in that case, a limited capacity to encode and/or use temporal fine-structure cues may explain the dramatic deficits reported for HI listeners with 16-band FM-speech stimuli; Buss et al. 2004; Lorenzi et al. 2006; Moore 2008), or (2) the contribution of reconstructed envelope cues to FM-speech identification has been underestimated when FM cues are extracted from 16 narrow analysis bands (as suggested by Ibrahim and Bruce 2010).

Effects of age

The HI listeners who participated in the present study were globally older than the NH listeners: the age of NH listeners ranged between 19 and 63 years whereas the age of HI listeners ranged between 28 and 70 years. Correlation analyses performed between the individual RAU scores of the HI listeners in 3-band FM condition (in the amplified condition) and their age while factoring out the effect of the PTAs estimated below 4 kHz were not significant (r=−0.02, p=0.92), demonstrating that the identification deficits shown by HI listeners for 3-band FM stimuli could not be accounted for by aging per se. This result is consistent with the outcome of previous studies showing that neither auditory frequency selectivity (e.g., Peters and Moore 1992) nor the ability to discriminate complex temporal-envelope cues (e.g., Takahashi and Bacon 1992; Lorenzi et al. 2006) are strongly affected by aging.

Contribution of audibility

Correlation analyses performed between the individual RAU scores of the HI listeners in 3-band FM condition (in the amplified condition) and their PTAs estimated below 4 kHz while factoring out the effect of age were not significant (r=−0.18, p=0.27), demonstrating that the identification deficits shown by HI listeners for 3-band FM stimuli could not be accounted for by differences in PTA estimated below 4 kHz. Statistical analyses also showed that linear amplification had a much smaller effect on speech identification for 3-band FM stimuli than for 3-band AM+FM stimuli. Altogether, these results suggest that the reduced audibility shown by HI listeners was not responsible for their speech identification deficit in the amplified, 3-band FM condition.

Training effects

Previous work assessing the intelligibility of speech stimuli processed to retain only FM cues showed that the capacity to use the degraded speech cues was highly susceptible to training effects for NH listeners when a high frequency resolution (e.g., 16 or 32 bands) was used to generate vocoded stimuli (Lorenzi et al. 2006; Gilbert et al. 2007; Hopkins et al. 2010). In

the present study, all listeners were tested only once with a complete set of 3-band FM speech stimuli presented at a given sound level. Consequently, the capacity to use recovered envelope cues may have been underestimated in each group. An additional experiment was run on 10 among the 40 NH listeners to assess the extent to which the capacity to identify 3-band FM speech stimuli is affected by training. All NH listeners were tested again for ten successive blocks of 48, 3-band FM speech stimuli during a single session. An ANOVA conducted on the training data showed that consonant identification scores improved significantly by 9 RAU between the first and last training block [$F(10,90)=3$; $p<0.01$]. This small, although significant training effect suggests that the capacity of NH and HI listeners to use recovered envelope cues was not substantially underestimated when performing the current consonant identification task.

Clinical applications

The present speech identification test based on processed nonsense speech material (3-band FM syllables) may lead to interesting clinical applications. This test (1) is easy and relatively fast to administer, (2) does not require training, (3) yields high levels of performance in normal control listeners, (4) is relatively immune to floor and ceiling effects in normal control listeners and listeners with a mild-to-moderate hearing loss, and (5) is less influenced by reduced audibility than a comparable speech test using unprocessed speech material. The outcome of this study indicates that poorer-than-normal frequency selectivity impairs—although modestly—the capacity to recover important envelope information from broadband FM speech cues. This test may thus be of use for the fast diagnosis of supra-threshold auditory deficits responsible for impoverished speech perception in daily situations where AM cues are degraded by nonlinear distortions (e.g., clipping, compression; Houtgast and Steeneken 1985).

CONCLUSIONS

The present study compared the ability of a large cohort of normal-hearing and hearing-impaired listeners to identify nonsense syllables processed to retain FM cues only within three broad frequency bands.

1. Consistent with previous modeling work, hearing-impaired listeners showed significantly poorer identification scores than normal-hearing listeners irrespective of presentation level.
2. The detrimental effect of cochlear hearing loss was relatively modest, amounting to 9 RAU when speech stimuli were amplified to compensate for reduced audibility.
3. Overall, the hearing-impaired data were consistent with a poorer-than-normal ability to reconstruct temporal-envelope speech cues resulting from a broadening of cochlear filters by a factor ranging from 2 to 4.

These results indicate that temporal-envelope reconstruction from broadband FM is an important, early auditory mechanism contributing to the robust perception of speech sounds in degraded listening conditions, as also suggested by recent work by Swaminathan and Heinz (2012). These results also suggest that most people suffering from mild to moderate cochlear hearing loss can make efficient use of reconstructed envelope cues despite degradations in frequency selectivity. Still, our results suggest that poorer-than-normal frequency selectivity impairs somewhat temporal-envelope reconstruction mechanisms. For hearing-impaired listeners showing the worst capacity to use recovered envelope cues, poor speech perception is anticipated in listening situations where AM speech cues are attenuated (for instance, by peak clipping or amplitude compression). The current study shows directions to develop a clinical test that may prove useful for the fast diagnosis of supra-threshold auditory deficits.

ACKNOWLEDGMENTS

C. Lorenzi was supported by two grants (PRESBYCUSIS and HEARFIN projects) from the French ANR agency. D. T. Ives was supported by a postdoctoral grant from Starkey France. A. Léger was supported by a CIFRE grant from the ANRT agency and Neurelec. The authors thank Alain de Cheveigné, Barbara Shinn-Cunningham, Joseph Hall, and four anonymous reviewers for valuable input on a previous version of this manuscript.

REFERENCES

Apoux F, Millman RE, Viemeister NF, Brown CA, Bacon SP (2011) On the mechanism involved in the recovery of envelope information from temporal fine structure. J Acoust Soc Am 130:273–282

Ardoint M, Lorenzi C (2010) Effects of lowpass and highpass filtering on the intelligibility of speech based on temporal fine structure or envelope cues. Hear Res 260:89–95

Ardoint M, Sheft S, Fleuriot P, Garnier S, Lorenzi C (2010) Perception of temporal fine structure cues in speech with minimal envelope cues for listeners with mild-to-moderate hearing loss. Int J Audiol 49:823–831

Ardoint M, Agus T, Sheft S, Lorenzi C (2011) Importance of temporal-envelope speech cues in different spectral regions. J Acoust Soc Am Exp Lett 130:EL115–EL121

Bacon SP, Gleitman RM (1992) Modulation detection in subjects with relatively flat hearing losses. J Speech Hear Res 35:642–653

Badri R, Siegel JH, Wright BA (2011) Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiograms. J Acoust Soc Am 129:852–863

Baer T, Moore BCJ (1993) Effects of spectral smearing on the intelligibility of sentences in noise. J Acoust Soc Am 93:1229–1241

Baskent D (2006) Speech recognition in normal hearing and sensorineural hearing loss as a function of the number of spectral channels. J Acoust Soc Am 120:2908–2925

Buss E, Hall JW 3rd, Grose JH (2004) Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. Ear Hear 25:242–250

de Cheveigne A, Kawahara H (2002) YIN, a fundamental frequency estimator for speech and music. J Acoust Soc Am 111:1917–1930

Drennan WR, Won JH, Dasika VK, Rubinstein JT (2007) Effects of temporal fine structure on the lateralization of speech and on speech understanding in noise. J Assoc Res Otol 8:373–383

Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. J Acoust Soc Am 97:585–592

Dubno JR, Dirks DD (1989) Auditory filter characteristics and consonant recognition for hearing-impaired listeners. J Acoust Soc Am 85:1666–1675

Eaves JM, Kitterick PT, Summerfield AQ (2011) Benefit of temporal fine structure to speech perception in noise measured with controlled temporal envelopes. J Acoust Soc Am 130:501–507

Füllgrabe C, Meyer B, Lorenzi C (2003) Effect of cochlear damage on the detection of complex temporal envelopes. Hear Res 178:35–43

Ghitza O (2001) On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception. J Acoust Soc Am 110:1628–1640

Gilbert G, Lorenzi C (2006) The ability of listeners to use recovered envelope cues from speech fine structure. J Acoust Soc Am 119:2438–2444

Gilbert G, Lorenzi C (2010) Role of spectral and temporal cues in restoring missing speech information. J Acoust Soc Am 128: EL294–EL299

Gilbert G, Bergeras I, Voillery D, Lorenzi C (2007) Effects of periodic interruption on the intelligibility of speech based on temporal fine-structure or envelope cues. J Acoust Soc Am 122:1336–1339

Glasberg BR, Moore BCJ (1986) Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. J Acoust Soc Am 79:1020–1033

Glasberg BR, Moore BCJ (1990) Derivation of auditory filter shapes from notched-noise data. Hear Res 47:103–138

Gnansia D, Jourdes V, Lorenzi C (2008) Effect of masker modulation depth on speech masking release. Hear Res 239:60–68

Gnansia D, Pean V, Meyer B, Lorenzi C (2009) Effects of spectral smearing and temporal fine structure degradation on speech masking release. J Acoust Soc Am 125:4023–4033

Gnansia D, Pressnitzer D, Pean V, Meyer B, Lorenzi C (2010) Intelligibility of interrupted and interleaved speech in normal-hearing listeners and cochlear implantees. Hear Res 14:46–53

Goodman A (1965) Reference zero levels for pure-tone audiometer. ASHA 7:262–263

Grant KW, Summers V, Leek MR (1998) Modulation rate detection and discrimination by normal-hearing and hearing-impaired listeners. J Acoust Soc Am 104:1051–1060

Heinz MG, Swaminathan J (2009) Quantifying envelope and fine-structure coding in auditory nerve responses to chimaeric speech. J Assoc Res Otolaryngol 10:407–423

Hohmann V (2002) Frequency analysis and synthesis using a Gammatone filterbank. Acust Acta Acust 88:433–442

Hopkins K, Moore BCJ, Stone MA (2008) Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. J Acoust Soc Am 123:1140–1153

Hopkins K, Moore BCJ, Stone MA (2010) The effects of the addition of low-level, low-noise noise on the intelligibility of sentences processed to remove temporal envelope information. J Acoust Soc Am 128:2150–2161

Houtgast T, Steeneken HJM (1985) A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. J Acoust Soc Am 77:1069–1077

Ibrahim RA, Bruce IC (2010) Effects of peripheral tuning on the auditory nerve's representation of speech envelope and temporal fine structure cues. In: Lopez-Poveda EA, Palmer AR, Meddis R (eds) The neurophysiological bases of auditory perception, chapter 40. Springer, New York, pp 429–438. doi:10.1007/978-1-4419-5686-6_40

Kale S, Heinz M (2010) Envelope coding in auditory nerve fibers following noise-induced hearing loss. J Assoc Res Otolaryngol 11:657–673

Kale S, Heinz M (2012) Temporal modulation transfer functions measured from auditory-nerve responses following sensorineural hearing loss. Hear Res 286:64–75

Kates JM (2011) Spectro-temporal envelope changes caused by temporal fine-structure modification. J Acoust Soc Am 129:3981–3990

Licklider JCR, Pollack I (1948) Effects of differentiation, integration and infinite peak clipping upon the intelligibility of speech. J Acoust Soc Am 20:42–51

Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BCJ (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. Proc Natl Acad Sci USA 103:18866–18869

Lorenzi C, Debruille L, Garnier S, Fleuriot P, Moore BCJ (2009) Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal. J Acoust Soc Am 125:27–30

Moore BCJ (2007) Cochlear hearing loss: physiological, psychological and technical issues. Wiley, Chichester

Moore BCJ (2008) The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. J Assoc Res Otolaryngol 9:399–406

Moore BCJ, Glasberg BR (1998) Use of a loudness model for hearing-aid fitting. I. Linear hearing aids. Br J Audiol 32:317–335

Moore BCJ, Sek A (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 BCJ, Shailer MJ, Schooneveldt GP (1992) Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss. Br J Audiol 26:229–237

Moore BCJ, Wotczak M, Vickers DA (1996) Effect of loudness recruitment on the perception of amplitude modulation. J Acoust Soc Am 100:481–489

Patterson RD (1987) A pulse ribbon model of monaural phase perception. J Acoust Soc Am 82:1560–1586

Peters RW, Moore BCJ (1992) Auditory filter shapes at low center frequencies in young and elderly hearing-impaired subjects. J Acoust Soc Am 91:256–266

Recio A, Rhode WS, Kiefte M, Kluender KR (2002) Responses to cochlear normalized speech stimuli in the auditory nerve of cat. J Acoust Soc Am 111:2213–2218

Saberi K, Hafter ER (1995) A common neural code for frequency- and amplitude-modulated sounds. Nature 374:537–539

Santurette S, Dau T (2007) Binaural pitch perception in normal-hearing and hearing-impaired listeners. Hear Res 223:29–47

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. Science 270:303–304

Sheft S, Ardoint M, Lorenzi C (2008) Speech identification based on temporal fine structure cues. J Acoust Soc Am 124:562–575

Shera CA, Guinan JJ Jr, Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc Natl Acad Sci USA 99:3318–3323

Stone MA, Glasberg BR, Moore BCJ (1992) Simplified measurement of impaired auditory filter shapes using the notched-noise method. Br J Audiol 26:329–334

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

Studebaker GA (1985) A "rationalized" arcsine transform. J Speech Hear Res 28:455–462

Swaminathan J, Heinz MG (2011) Predicted effects of sensorineural hearing loss on across-fiber envelope coding in the auditory nerve. J Acoust Soc Am 129:4001–4013

Swaminathan J, Heinz MG (2012) Psychophysiological analyses demonstrate the importance of neural envelope coding for speech perception in noise. J Neurosci 32:1747–1756

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

Turner CW, Souza PE, Forget LN (1995) Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. J Acoust Soc Am 97:2568–2576

Tyler RS, Hall JW, Glasberg BR, Moore BCJ, Patterson RD (1984) Auditory filter asymmetry in the hearing impaired. J Acoust Soc Am 76:1363–1368.

Warren RM, Bashford JA Jr, Lenz PW (2004) Intelligibility of bandpass filtered speech: steepness of slopes required to eliminate transition band contributions. J Acoust Soc Am 115:1292–1295

Zeng FG, Nie K, Liu S, Stickney G, Del Rio E, Kong YY, Chen H (2004) On the dichotomy in auditory perception between temporal envelope and fine structure cues. J Acoust Soc Am 116:1351–1354

Zeng FG, Nie K, Stickney GS, Kong YY, Vongphoe M, Bhargave A, Wei C, Cao K (2005) Speech recognition with amplitude and frequency modulations. Proc Natl Acad Sci USA 102:2293–2298

Résumé

Les signaux de communication contiennent des modulations d'amplitude (AM) et de fréquence (FM) qui jouent un rôle dans l'identification des sons. Le but de ce programme de recherche est de préciser les mécanismes impliqués dans la perception d'AM et de FM, ainsi que de clarifier les effets du vieillissement et de la perte auditive neurosensorielle sur ceux-ci. Les résultats démontrent que: 1) pour des cadences de modulation rapides, la détection d'AM et de FM sont encodés par un mécanisme commun, probablement basé sur des indices d'enveloppe temporelle. Aux basses cadences de modulation, l'AM et la FM sont encodées par des mécanismes distincts, utilisant probablement et respectivement des indices d'enveloppe temporelle et de structure temporelle fine; 2) le vieillissement dégrade la sensibilité à l'AM et la FM (ie. les indices d'enveloppe temporelle et de structure temporelle fine), mais affecte plus fortement ces derniers; 3) la perte auditive n'affecte pas la sensibilité à l'AM, mais dégrade la sensibilité à la FM; 4) l'intégration temporelle d'AM est préservée chez les sujets âgés, mais s'améliore en présence d'une perte auditive. En FM, l'intégration temporelle est préservée avec l'âge et la perte auditive. Un modèle computationnel d'enveloppe temporelle a été développé et les résultats montrent que les processus décisionnels et mnésiques impliqués dans l'intégration temporelle sont préservés avec l'âge et la perte auditive. Toutefois, certains aspects de l'efficacité de traitement déclinent avec l'âge et encore plus fortement avec une perte auditive. Pris ensemble, ces travaux de recherche permettent de redéfinir le diagnostic de la presbyacousie.

Mots Clés

enveloppe temporelle, intégration temporelle, modulation d'amplitude, modulation de fréquence, perte auditive, structure temporelle fine, vieillissement

Abstract

Communication sounds contain amplitude (AM) and frequency (FM) modulations that play a role in sound identification. The aim of this research program is to clarify the mechanisms involved in AM and FM perception, as well as to assess the effects of age and sensorineural hearing loss. The results showed that: 1) for high modulation rates, AM and FM detection are encoded by a common mechanism, probably based on temporal envelope cues. For low modulation rates, AM and FM detection are encoded by distinct mechanisms, based respectively on temporal envelope and temporal fine structure cues; 2) ageing degrades the sensitivity to AM and FM (temporal envelope and fine structure cues), but affects more importantly the latter; 3) hearing loss does not affect AM sensitivity, but degrades FM sensitivity; 4) temporal integration of AM is spared with age, but improves with hearing loss. Temporal integration of FM is preserved with both age and hearing loss. A computational model based on temporal envelope information was developed and the results demonstrated that the decisional and memory factors involved in temporal integration are preserved with age and hearing loss. However, some aspects of processing efficiency decline with age and more importantly, with hearing loss. Taken together, this research program redefines the diagnosis of presbycusis.

Keywords

ageing, amplitude modulation, frequency modulation, hearing loss, temporal envelope, temporal fine structure, temporal integration.