Aix Marseille Université

NNT/NL : 2025AIXM0102

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 6 mai 2025

Coline MARCHÉ

Context effects in sound intensity encoding and perception

Discipline
Sciences pour l'ingénieur

Spécialité
Acoustique

École doctorale
Sciences pour l'Ingénieur : Mécanique, Physique, Micro et Nanoélectronique / ED353

Laboratoire
Laboratoire de Mécanique et d'Acoustique, UMR7031

Composition du jury

Laurel H. CARNEY Prof., University of Rochester	Rapporteure
Etienne GAUDRAIN Chargé de Recherche, CNRS, Centre de Recherche en Neurosciences	Rapporteur
Emmanuel PONSOT Chargé de Recherche, CNRS, Laboratoire STMS	Examinateur
Christian LORENZI Prof., ENS Paris	Président du jury Examinateur
Sabine MEUNIER Directrice de Recherche, CNRS, LMA	Directrice de thèse
Olivier MACHEREY Directeur de Recherche, CNRS, LMA	Directeur de thèse

Laboratoire de Mécanique et d'Acoustique

Affidavit

I, undersigned, Coline Marché, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific direction of Sabine Meunier and Olivier Macherey, in accordance with the principles of honesty, integrity, and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the French national charter for Research Integrity and the Aix-Marseille University charter on the fight against plagiarism.

This work has not been submitted previously either in France or in another country in the same or in a similar version as part of another examination or defense process.

In Marseille, March 10, 2025



This work is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License

List of publications and/or patents and participation in conferences

1) List of publications and/or patents realized in the framework of the thesis project: None

2) Participation in conferences and summer schools during the thesis period:

1. C. Marché, O. Macherey and S. Meunier. In Search of a Psychophysical Correlate of Dynamic Range Adaptation in Human Listeners: Intensity Discrimination within a Context. In *47th Annual ARO MidWinter Meeting*, Anaheim, United States of America, Feb. 2024.
   Poster presentation

2. C. Marché, S. Meunier and O. Macherey. Discrimination d'intensité et adaptation de la gamme dynamique. In *Journées Perception Sonore 2023*, Paris, France, Nov. 2023.
   Poster presentation

3. C. Marché, S. Meunier and O. Macherey. (Pure tone) intensity discrimination and dynamic range adaptation. In *10th Convention of the European Acoustics Association*, Turin, Italy, Sep. 2023.
   Oral presentation and proceedings paper

4. C. Marché, S. Meunier and O. Macherey. Discrimination d'intensité et adaptation de la gamme dynamique. In *Journées Jeunes Chercheur · se · s en Audition, Acoustique musicale et Signal audio 2023*, Paris, France, Apr. 2023.
   Poster presentation

5. C. Marché, O. Macherey, J. Grange and S. Meunier. Discrimination d'intensité et adaptation de la gamme dynamique des neurones du système auditif. In *16ème Congrès Français d'Acoustique*, Marseille, France, Apr. 2022.
   Oral presentation

Résumé

Des études animales ont montré que les neurones auditifs adaptent leur taux de décharge selon les statistiques du niveau sonore. Cette adaptation de la gamme dynamique (DRA) a été mesurée en présentant de courts segments de bruit ou sons purs avec certains niveaux présentés plus que les autres. La DRA induit un décalage des fonctions de réponse vers les niveaux les plus présentés. Elle produit un changement de sensibilité améliorant le codage des niveaux les plus probables. La DRA a été observée dans le nerf auditif (AN), impliquant que les cellules ciliées (HC), les synapses entre HC et AN et/ou l'AN y contribuent.

Le premier objectif était de trouver un corrélat perceptif à la DRA chez le normo-entendant. Nous avons mesuré la discrimination d'intensité pour des sons purs et large bande dans différents contextes avec l'hypothèse que les seuils de discrimination devraient être plus faibles quand le niveau moyen du contexte est identique à celui du signal test. Pour les signaux large bande, les résultats sont en accord avec l'hypothèse mais pour les sons purs la discrimination d'intensité se dégrade avec l'augmentation du niveau du contexte.

Le second objectif était de tester si l'AN contribue directement à la DRA. Nous avons mesuré des potentiels évoqués électriquement en réponse à des séquences de stimulus avec des distributions de niveau spécifiques chez des implantés cochléaires, pour qui les HC et leur synapse avec l'AN sont court-circuités. Les fibres de l'AN stimulées électriquement ont collectivement un comportement en accord avec la DRA.

Mots-clés : psychoacoustique, discrimination d'intensité, adaptation de la gamme dynamique, nerf auditif, implant cochléaire

Abstract

Animal studies have shown that auditory neurons adapt their firing rate depending on the sound level statistics. This Dynamic Range Adaptation (DRA) has been measured by presenting short segments of noise or tones with some levels being presented more than others. DRA manifests as a shift of the rate-level function towards the most presented levels. It produces a change in sensitivity enhancing the coding of the most probable levels. DRA has been observed in the auditory nerve (AN), implying that the hair cells, the hair cell-AN synapses and/or the AN contribute to this phenomenon.

The first aim of this thesis was to find a perceptual correlate of DRA in normal-hearing participants. We measured intensity discrimination for pure-tone and broadband-noise test signals embedded in different contexts with the hypothesis that discrimination thresholds should be lower when the context mean level matches the test signal level. For broadband signals, the results were fully consistent with this hypothesis while for pure tones, intensity discrimination always degraded with increases in context level, independently from the test signals.

The second aim was to test whether the AN directly contributes to DRA. We measured electrically-evoked compound action potential responses to electrical stimulus sequences with specific level distributions in cochlear implant listeners, for whom the hair cells and hair cell-AN synapses are bypassed. Electrically stimulated AN fibers collectively exhibit a behavior consistent with DRA, showing the AN contributes to DRA.

Keywords: psychoacoustics, intensity discrimination, dynamic range adaptation, auditory nerve, cochlear implants

Remerciements

" Je vais aller rajouter de l'entropie par là-bas. "

Tom Colinot, *JJCAAS 2024*

I would like to start by thanking the jury, for their interest in my work, for their time, their feedback, and their benevolence.

Je ne peux terminer l'écriture de ce manuscrit sans remercier le laboratoire qui m'a accueilli, pendant quatre ans déjà. Merci à toutes les personnes que j'ai croisé et avec qui j'ai échangé. En particulier, merci à toute l'équipe Sons pour son cadre chaleureux, son café aléatoire, et ses réunions d'équipe avec gâteau ; surtout, merci aux jeunes (et un peu moins). Merci Matthieu, Tom, Nathan, Martin, Romain, Alice, pour la musique, entre autres (et Joris et Anaïs pour les JJCAAS). Et encore plus petite échelle, merci aux psycho(pathes). Merci Gaston. Merci Guy pour l'auto-cyclo-mobile. Merci Ossen pour la bonne humeur, l'engagement, et les pâtisseries. Merci Manu pour les randonnées, les calembours, l'invitation et la culture (pas la confiture). Merci Marine pour les plantes, le thé et le chocolat quotidien. Merci à mes grand.e.s frère et soeur de thèse : Ana, Tim, pour la résilience ; merci Clément aussi. Merci infiniment Jacques pour l'expérience et la bienveillance ; les meilleurs partent les premiers.

Une pensée pour SATIS, en particulier Florence, Antoine, Rosine et Rémi, qui m'ont fait confiance pour gérer des cours entiers toute seule dès mon début de thèse.

Merci évidemment à l'EMS (Ensemble Musical du Sud) sans qui j'aurais dépéri (et ma clarinette pris la poussière) : pour tous les concerts et tous les bons moments, merci, j'ai de la peine de vous quitter.

Aux amis de Marseille et d'avant, merci ; je ne peux pas tous vous citer ici... Merci Lorène et Théo qui m'ont suivie (ou presque), Sophie pour s'être rapprochée en partant, Loann pour le cumul, Thomas pour les débats, et tous pour les sushis et les sorties. Merci aux grimpeurs, parce que quand ça ne va pas, tu montes. Merci Pierrot, Pascal, Valoo, Marco, Flo, et tous les autres.

Merci à ma famille, plus ou moins lointaine : Papy, Mamy, Mémé et tout le monde. Merci Papa et Maman (pour tout et plus encore : la montagne, les oiseaux, les plantes, les étoiles...), merci Cécile pour les conversations sans queue ni tête, merci Loïc pour la spontanéité. Et à bientôt dans un nouveau pays ?

Non, je ne vous oublie pas, et j'aurais du mal à vous oublier : merci Sabine et merci Olivier. Merci de m'avoir accordé votre confiance et votre soutien tout au long de ce projet. Merci de m'avoir formée, guidée, conseillée, et merci de vous être essayé à cette gymnastique consistant à essayer de penser comme moi, pour rendre tout cela compréhensible par d'autres. Cette fois, vous n'avez pas le droit de râler si c'est trop court.

Last but not least, Solène de loin (ou de Lyon) ; Amaury, Baptiste et Julie, Romain et Alex, merci de tout mon cœur d'être des amis si précieux.

Et pour ces moments de vie partagée, merci Chris.

Contents

Affidavit iii

Publications and communications v

Résumé vii

Abstract ix

Remerciements xi

List of Figures xvii

List of Tables xxv

Introduction 1

I The auditory system viewed by a psychoacoustician 3
I.1 Physiology . . . 3
I.1.1 The ear and auditory pathways . . . 3
I.1.1.1 Periphery . . . 3
I.1.1.2 Higher centers . . . 5
I.1.2 Hearing impairment and cochlear implants . . . 6
I.1.3 Intensity encoding . . . 8
I.1.3.1 Rate-place coding . . . 8
I.1.3.2 Temporal coding . . . 10
I.1.3.3 Other theories . . . 11
I.2 Perception of sound intensity . . . 12
I.2.1 Masking . . . 12
I.2.2 Loudness . . . 13
I.2.3 Intensity discrimination . . . 15
I.3 Dynamic range adaptation . . . 18
I.3.1 Presentation . . . 18
I.3.2 Possible origins . . . 21
I.3.3 Other phenomena . . . 21
I.3.4 Human studies . . . 23
I.3.5 Aims of the thesis . . . 24

II Intensity discrimination using pure tones 27
II.1 Introduction . . . 27
II.2 Experiment 1 : Intensity discrimination in a narrow-band context . . . 27
II.2.1 Material and methods . . . 27
II.2.1.1 Participants . . . 27

II.2.1.2 Hardware and software . . . 28
II.2.1.3 Stimuli . . . 28
II.2.1.4 Procedure . . . 29
II.2.1.5 Preliminary experiment . . . 30
II.2.1.6 Auditory nerve model . . . 31
II.2.2 Results . . . 33
II.2.3 Discussion . . . 35
II.2.3.1 Comparison with simulations . . . 36
II.2.3.2 Comparison with the literature . . . 37
II.2.3.3 Partial masking . . . 40
II.2.3.4 Mid-level hump and confusion . . . 41
II.2.3.5 Off-frequency listening . . . 42
II.2.3.6 Overview of the control experiments . . . 43
II.3 Control experiments . . . 43
II.3.1 Experiment 2 : Discrimination at lower levels . . . 43
II.3.1.1 Material and methods . . . 44
II.3.1.2 Results . . . 44
II.3.2 Experiment 3 : Discrimination with notched noise . . . 48
II.3.2.1 Material and methods . . . 48
II.3.2.2 Results . . . 48
II.3.3 Experiments 4 & 5 : Discrimination with gaps or using a diotic context . . . 50
II.3.3.1 Material and methods . . . 50
II.3.3.2 Results . . . 50
II.3.4 Discussion . . . 55
II.3.4.1 Interleaved . . . 55
II.3.4.2 Lower levels . . . 55
II.3.4.3 Added notched noise . . . 57
II.3.4.4 Masking by the context . . . 59
II.4 General discussion . . . 60

III Intensity discrimination using broadband signals 63
III.1 Introduction . . . 63
III.1.1 Pulse-Spreading Harmonic Complexes . . . 63
III.2 Experiment 6 : Discrimination with broadband signals . . . 65
III.2.1 Material and methods . . . 65
III.2.1.1 Participants . . . 65
III.2.1.2 Material . . . 65
III.2.1.3 Stimuli . . . 65
III.2.1.4 Procedure . . . 68
III.2.1.5 Preliminary experiment . . . 68
III.2.1.6 Auditory nerve model . . . 69
III.2.2 Results . . . 69
III.2.2.1 Modelling . . . 70
III.2.3 Discussion . . . 72
III.2.3.1 Comparison with the model . . . 73

III.2.3.2 Intensity discrimination thresholds of PSHC in silence 74
III.2.3.3 Maskers, mid-level hump and bandwidth 75
III.3 Experiment 7 : Discrimination in different frequency bands 76
III.3.1 Material and methods 76
III.3.2 Results 77
III.3.3 Discussion 78

IV Physiological response of the human auditory nerve 81
IV.1 Introduction 81
IV.2 Material and methods 82
IV.2.1 Participants 82
IV.2.2 Experimental hardware and software 82
IV.2.3 Stimulation and recording parameters 84
IV.2.4 Setting up current levels 85
IV.2.5 Experimental paradigms 86
IV.2.6 Session 88
IV.2.7 Analyses 88
IV.2.7.1 Very-long-term eCAP evolution 89
IV.2.7.2 Investigating the effect of randomizing current level presentation 89
IV.2.7.3 Quantifying dynamic range adaptation as a function of the HPR 89
IV.2.7.4 Comparing the slopes of the eCAP growth functions . 89
IV.2.7.5 Evaluating the temporal changes in different conditions 90
IV.2.7.6 Control analyses 91
IV.3 Results 92
IV.3.1 Very-long-term eCAP changes 92
IV.3.2 Differences between AGFs for CLASSIC and RANDOM paradigms 92
IV.3.3 Horizontal shift of the AGF as a function of HPR 94
IV.3.4 Change in response function shape 94
IV.3.4.1 Comparison at fixed current level 97
IV.3.4.2 Comparison at fixed amplitude 98
IV.3.4.3 Do the steeper slopes in the HPR reflect improved intensity encoding ? 98
IV.3.5 Temporal effects 100
IV.3.5.1 Evolution at 30 pps 100
IV.3.5.2 Changes at 300 pps 102
IV.3.6 Effect of pulse order 104
IV.3.7 Effect of preceding level 105
IV.4 Discussion 107
IV.4.1 Dynamic range adaptation at the level of the AN 107
IV.4.1.1 AGFs shift depending on the most probable levels . . . 107
IV.4.1.2 Differences with animal studies 108
IV.4.2 Effects of level variations 109
IV.4.3 Time constants 110
IV.4.4 Implications for CI coding 112

IV.5 Summary and conclusion 113

General conclusion **115**

References **119**

Appendix **133**

A Results of the preliminary experiment with pure tones for the selected participants 133

B Sensitivity indexes and DLs obtained with the model 134

B.1 Methods 134

B.2 Experiment 1 134

B.3 Experiment 2 136

B.4 Experiment 4 138

B.5 Experiment 6 138

C Additional data on eCAPs 141

C.1 Evolution of 30 pps classic control for L12 141

C.2 AGF for HPRs and RANDOM paradigms, for all participants . 141

C.3 Temporal evolution of normalized amplitude at 300 pps 147

List of Figures

I.1 The different components of the ear or auditory periphery, divided into the outer (external), middle and inner ear ; credits : OpenStax College. 3

I.2 Main centers of auditory processing throughout the auditory pathway ; the auditory nerve connects the cochlea to the brainstem which contains the cochlear nuclei and the olivocochlear systems. The information is then fed to the inferior colliculus in the brainstem and the medial geniculate body in the thalamus before it reaches the auditory cortex. From Butler and Lomber (2013). 5

I.3 Diagram of a cochlear implant in the human auditory system. The processor and its microphone are external and generally placed around the ear, while the transmitter is magnetically bonded to the receiver, under the skin. The signal then travels to the electrode array, surgically inserted inside the cochlea. Source : NIH/NIDCD. 7

I.4 Filtering process of a sound by a cochlear implant : the signal is first filtered as the outer and middle ear would ; it is then split through different bandpass filters ; the envelope of the output of each filter is then convoluted with electric pulse trains, and each electrode receives amplitude-modulated pulse trains matching the envelope of the filtered signal. 7

I.5 Theoretical examples of rate-level functions for two types of fibers in the cat auditory nerve : the low-spontaneous rate (SR) at the bottom, and high-SR fibers on top. From Plack (2013). 9

I.6 Envelope fluctuation profiles : the AN response of several frequency regions to a vowel over time is presented, along with the spectrum of the corresponding sound and the resulting amplitude fluctuation profile. The regions centered around the peaks in the spectrum have low amplitude fluctuations, while the in-between ones have higher fluctuations. From figure 2 in Carney (2018). 12

I.7 Equal-loudness curves : all the pure tones on a line have the same loudness level in phones. The lowest curve represents the minimum audible field (MAF) ; each curve above gives the level at which a pure tone of the selected frequency will sound as loud as the 1-kHz reference at this level. From Plack (2013). 14

I.8 Example of a stimulus used in Dean et al. (2005). Panel a) shows the distribution of levels over time for a broadband noise and a 12-dB high-probability region (HPR) centered around 63 dB SPL ; a 200-ms extract of the signal is shown in panel b) ; panel c) shows the probability distribution of the levels (replotted from Figure 1 of Dean et al. 2005). 19

I.9 Rate-level functions for four individual neurons measured by Dean et al. (2005). The grey line is the rate-level function in quiet, and each colored line represents the results for a specific HPR, represented in the x-axis (mean levels ; green : 39 dB SPL, blue : 51 dB SPL, red : 63 dB SPL and cyan : 75 dB SPL). The shift of the response is placing the steepest part of the rate-level function closer to the mean level of the HPR (replotted from Figure 1 of Dean et al. (2005)). 19

I.10 Different features of the dual-adaptation model of Wen et al. (2009). Panel A presents dynamic range adaptation, panel B classical adaptation, panel C a mix of the two and panel D what is observed in the measurement results. 20

II.1 Example trials from the main experiment for each condition, sorted by standard level and context condition. The soft test signals (standard and target) are represented in dark blue while the soft context is in light blue, the loud context is in orange and the loud test signals are in dark red. 29

II.2 Rate-level function of different neuron types to different stimuli, taken from the MAP auditory nerve model run by Grange et al. (2022) and replotted from their figure 6. Tracks of 8 s of fluctuating-level stimuli were averaged over ten randomization of 20 repetitions of each tested level (200 in total), with a 12-dB wide HPR of different mean levels (indicated in dB SPL in the legend). BBN : broadband noise. 32

II.3 Results of experiment 1 for both standard levels. Discrimination thresholds for all participants are plotted against context condition : the first and third quartile are limiting the blue box, and the median is presented as a blue bar within it. Whiskers include all non-outlier values that do not fall into the box ; blue circles are outliers (between 1.5 and 3 times the interquartile value away from the first or third quartiles) and red circles extreme values (more than 3 times the interquartile value away from the first or third quartiles). Pairwise t-tests on the discrimination thresholds revealed significant differences (represented as stars on the graph) between the loud context and the other conditions for both standards. 33

II.4 Rate-level functions with shift between contexts (a) and slopes (b) obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 1, averaged over ten simulations. The 55- and 75-dB SPL standard are represented respectively as black x and +. 35

II.5 Results of experiment 2 for six participants. Discrimination thresholds are plotted against context type (including silence), as in figure II.3. They are significantly higher in the loud context than in the other conditions for both standards, and in the soft context than in silence for the soft standard. 44

II.6 Discrimination thresholds for the six participants of experiment 2 for the 55-dB SPL standard in silence and in the 55-dB SPL context - the common conditions of experiment 1 and 2. 46
II.7 Rate-level functions with shift between contexts (a) and slopes (b) obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 2, averaged over ten simulations. The 43- and 55-dB SPL standard are represented respectively as black x and +. 46
II.8 Results of experiment 3 for the loud standard (75 dB SPL). Discrimination thresholds are plotted like in experiment 1 (see figure II.3). There is no change in the discrimination thresholds with the addition of different levels of notched noise : the results are still higher in the loud context than in the soft context, and the level of the notched noise does not seem to have an effect. 49
II.9 Results of experiment 4, with silent gaps between context and standard. The discrimination thresholds are plotted for both standard levels in both context levels, as in figure II.3. There is a significant increase in discrimination thresholds in the loud context compared to the soft context for both standard levels. 51
II.10 Rate-level functions with shift between contexts (a) and slopes (b) obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 4, averaged over ten simulations. The 55- and 75-dB SPL standard are represented respectively as black x and +. 52
II.11 Results of experiment 5. The discrimination thresholds are plotted for both standard levels in each context mean level, as in figure II.3. Only the soft standard displayed a significant increase in discrimination thresholds with the diotic context. 53
II.12 Difference in discrimination thresholds between the 75- and 55-dB SPL contexts for both standard levels in experiments 1, 4 and 5. 54
II.13 Review of discrimination thresholds obtained at difference frequencies. From Florentine et al. (1987). 57
II.14 Excitation patterns of the 75-dB SPL standard, 35- and 65-dB SPL notched noise used in experiment 3 ; in red, the ERB at 2 kHz. Made with the model from Glasberg and Moore (1990) ; Moore et al. (1997) ; ANSI (2007). 58

III.1 Example of harmonic distribution for a 7th-order PSHC. Each panel on the left displays one subgroup of harmonics and on the right the resulting waveform for this subgroup, with a certain phase shift applied (in grey), and the full signal at the bottom. Replotted from figure 1 from Hilkhuysen and Macherey (2014). 64
III.2 Example waveforms for two 50-ms excerpts of the signals used in this chapter. 66

III.3 Ouputs of seven auditory filters for two 50-ms excerpts of the signals used in this chapter. These are the output of gammatone filters from Slaney (1998). 67

III.4 Results of experiment 6 for all participants. Discrimination thresholds are plotted against context condition : the first and third quartile are limiting the blue box, and the median is presented as a blue bar within it. Whiskers include all non-outlier values that do not fall into the box ; blue circles are outliers (between 1.5 and 3 times the interquartile value away from the first or third quartiles) and red circles extreme values (more than 3 times the interquartile value away from the first or third quartiles). * : significant difference between conditions, see table III.4. . 70

III.5 Rate-level functions with shift between contexts (a) and slopes (b) obtained with the MAP model as configured by Grange et al. (2022) for HSR, MSR and LSR fibers in the ERB tuned around 1 kHz, for the PSHC and broadband context conditions used in experiment 6, averaged over ten simulations. 71

III.6 Results of experiment 7 for all participants and the three passbands used in the task. Discrimination thresholds are plotted against context level as in figure III.4. The thresholds are very variable for all conditions. 77

IV.1 Schematic temporal representation of two electric pulses, one cathodic-leading and one anodic-leading, with the durations used in the present study. Both phases of both pulses have equal amplitudes and duration. 84

IV.2 Example of recorded temporal response of the auditory nerve to a certain current level, averaged over 100 repetitions, for one participant. The negative peak is called N1 and the positive peak P2 : the amplitude of the eCAP response is the voltage difference between P2 and N1. 85

IV.3 Schematic representation of the stimuli used in the three paradigms. For illustration purposes, only three levels are represented, but twelve were effectively tested. a) CLASSIC, with pairs of pulses of both polarities presented one after the other at a fixed level ; b) RANDOM : all tested levels and polarities are mixed ; c) HPRs, with two particular levels with a probability of presentation of 80 % ; three different pairs of most probable levels were tested (for the purpose of the illustration, only one most probable level is presented for each of the three conditions). . . . 87

IV.4 Evolution of L11 amplitude in dB for the short CLASSIC paradigm, between the start and the end of the procedure. Individual data (one marker per participant) are completed by mean and 95 % confidence interval (in black). Test and retest were averaged. 92

IV.5 AGF for the CLASSIC and RANDOM paradigms for participant C70 at 300 pps (a), and C26 at 300 pps (b). 93

IV.6 Amplitude growth function, relative shifts and shit slopes. The AGF and relative shift relative to the RANDOM condition are plotted for one participant only (C5) : a) AGF 30 pps, b) 100 pps, c) 300 pps ; d) average relative shift of the corresponding HPR paradigm compared to RANDOM AGF : in this case, the mean current level difference between two consecutive HPRs was 3.14 dB ; e) slopes of the average relative shift for all rates and participants (one marker per participant) ; the across-subject mean and 95 % confidence intervals are shown in black next to the individual results. 95

IV.7 Examples of slopes calculated between two subsequent points for the HPRs and the RANDOM measurements for two participants at 300 pps. a) Slope maxima are observable for each paradigm and shift towards higher or lower levels depending on the most probable levels in the measurements ; b) in most cases the maximum was the first available slope and no evident shift appears. 96

IV.8 Slopes of interest for the HPR3 condition at 300 pps : zoomed in from figure IV.6c (HPR1 and HPR2 removed). The slope for the RANDOM curve at fixed amplitude is calculated with a linear regression on the two interpolated points on the RANDOM curve (purple X). 98

IV.9 Measures of response function changes in shape at 300 pps : a) slopes between the two HPR levels for the HPR paradigm and the RANDOM paradigm as well as the interpolated slope at fixed amplitude ; b) sensitivity index (slope over standard deviation) for the same conditions. 99

IV.10 Evolution of the response over groups of 100 pulses in the CLASSIC measurement at 30 pps : a) & b) examples of temporal evolutions of amplitude over the duration of the measurement for two participants, one with changes in amplitudes (a) and one without any evolution (b) ; c) amplitude evolution over the duration of the measurement for the first, second and last groups for all participants (one marker per participant, see figure IV.4), with the mean and 95 % confidence interval presented in black next to the individual results. Pulses in group 1 were recorded between 0 and 3.3 s after the start of the measurement, group 2 was 3.3 to 6.6 s after the start, and group 100 330 to 333.3 s (5 minutes 33 s, last group). 101

IV.11 Temporal evolution of the normalized amplitude for L3, L7 and L11 at 300 pps : a), b), & c) examples of temporal evolutions of the normalized amplitude over the duration of the measurement for different participants in both the CLASSIC and HPR measurements : one with an initial decrease only (a), one with a second decreasing slope (b), and one with an increase after the first decrease (c) ; d) adaptation index (normalized amplitude of the last pulse group) for all participants (see figure IV.4) and paradigms, with mean and 95 % confidence interval in black. . . . 103

IV.12 Time constant τ values at 300 pps for all participants with mean and 95 % confidence interval in black (one symbol per participant, see figure IV.4). τ was estimated as the time needed for a measurement to reach 36.8 % of its final value. 104

IV.13 Average difference in response amplitude between even and odd pulses in the HPR2 paradigm at 300 pps. All participants are represented, along with mean differences and 95 % confidence interval (black). . . . 105
IV.14 Average difference in response amplitude for HPR2 (L6 and L7) between preceding pulses with higher amplitude and with lower amplitude, at 30, 100 and 300 pps. All participants are represented, along with mean differences and 95 % confidence interval (black). 105
IV.15 Difference between response after a higher- and a lower-level pulse for the HPR2 paradigm at 300 pps, for the five participants who had values for both (C15, C18, C19, C35 and C70). 106
IV.16 As IV.14, but for the RANDOM paradigm and levels L2 to L11. 107
A.1 Detection thresholds in the three context conditions of experiment 1, for the ten selected participants. 133
B.1 Sensitivity index obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 1, averaged over ten simulations. The 55- and 75-dB SPL standard are represented respectively as black x and +. 135
B.2 Sensitivity index obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 2, averaged over ten simulations. The 43- and 55-dB SPL standard are represented respectively as black x and +. 137
B.3 Sensitivity index obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 4, averaged over ten simulations. The 55- and 75-dB SPL standard are represented respectively as black x and +. 138
B.4 Sensitivity index obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 1 kHz, for the broadband signals and context conditions used in experiment 6, averaged over ten simulations. The 50- and 70-dB SPL standard are represented respectively as black x and +. 140
C.1 Evolution of L12 amplitude in dB for the short CLASSIC paradigm, between the start and the end of the procedure, as in figure IV.4. . . . 141
C.2 Amplitude growth function for participants C2 and C15 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps. 142
C.3 Amplitude growth function for participants C16 and C17 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps. 143
C.4 Amplitude growth function for participants C18 and C19 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps. 144
C.5 Amplitude growth function for participants C35 and C70 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps. 145
C.6 Amplitude growth function for participants C46 and C26 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps. 146

C.7 Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C5 and C2. 148
C.8 Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C15 and C16. 149
C.9 Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C17 and C18. 150
C.10 Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C19 and C35. 151
C.11 Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C70 and C46. 152
C.12 Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participant C26. 153

List of Tables

II.1 Averaged results for experiment 1, with the corresponding condition letter in figure II.1. The average discrimination thresholds tend to increase in the 75-dB SPL context and they do not change much between the silence and 55-dB SPL context. 33
II.2 Pairwise t-tests for each standard and context mean level. Significant differences appear between the loud context and the other conditions for both standard levels : the loud context gives higher thresholds than the soft context and the silence (ns : non significant). 34
II.3 Slopes around both standard levels for each context and fiber types in figure II.4a. They are much lower for the 75-dB SPL standard as most fibers are already saturated. 36
II.4 Comparison of DLs (in dB) measured in experiment 1 and in the literature. DL is defined as $10log(1 + \Delta I/I)$ (rf : rise/fall time). 39
II.5 Details on the averaged results for experiment 2, with the corresponding condition letter in figure II.1. The average discrimination thresholds tend to increase with the context mean level for the 43-dB SPL standard. For the 55-dB SPL standard, the thresholds decrease in the 43-dB SPL context compared to silence, while they increase in the 55-dB SPL context. 45
II.6 Pairwise t-test for each standard and context mean level. The loud context gives significantly higher thresholds than the soft context for both standard levels, and than the silence for the soft standard. 45
II.7 Slopes around both standard levels for each context and fiber types in figure II.7b. 47
II.8 Details on the averaged results for experiment 3. The average discrimination thresholds and standard deviations tend to increase in the 75-dB SPL context but they do not change between the 35- and 65-dB SPL notched noise. 48
II.9 Repeated measures ANOVA (dof : degrees of freedom) on the data of experiment 3. The factors were context mean level and notched noise level. Only the level of the context had a significant effect on the discrimination thresholds. 49
II.10 Pairwise t-test between context mean levels in experiment 3. The discrimination thresholds are significantly different for both notched noise levels. 49
II.11 Averaged results for experiment 4, with 50-ms gaps between the presentation of the context and the standards. The discrimination thresholds increase with the context mean level. 50
II.12 Slopes around both standard levels for each context and fiber types in figure II.10b. They are much lower for the 75-dB SPL standard as most fibers are already saturated. 51

II.13 Averaged results for experiment 5, with the context presented to both ears. Average thresholds increase with context mean level. 53
II.14 Pairwise t-test comparisons between the thresholds with the same context mean level and standard level in experiments 1 and 4. Significant differences appear for the loud context for both standard levels : experiment 4 gives significantly lower discrimination thresholds in the 75-dB SPL context. 54
II.15 Pairwise t-test comparisons between thresholds of experiments 1 and 5 for each standard and context mean level. Significant differences appear for the loud context for both standard levels : experiment 5 gives significantly lower discrimination thresholds in the 75-dB SPL context. 54

III.1 Mean detection thresholds for the 16 selected participants for this study. They increased with the mean level of the context. 69
III.2 Averaged results for experiment 6. The mean discrimination thresholds are lowest when the context mean level is identical to the standard level. 69
III.3 Repeated measures ANOVA on standard level and context mean level. The investigated parameters and their interaction all have significant influences on the discrimination thresholds. 70
III.4 Pairwise t-test for each standard and context level. The context matching the standard level yields lower discrimination thresholds than the other context conditions for both standard levels ; only the silence and soft context for the loud standard are not significantly different. 71
III.5 Slopes around both standard levels for each context and fiber types in figure III.5a : three-points linear regression around the standard level. . 72
III.6 Mean detection thresholds and standard error for experiments 6 and 1. 75
III.7 Details on the averaged results for experiment 7. The mean discrimination thresholds tend to increase with the context mean level for the lowpass band, but tend to decrease between soft and loud contexts for the medium and highpass bands. 78

IV.1 Information on the CI users who took part. The electrode number is the stimulating electrode ; the recording electrode was 2 electrodes more basal. Age and CI use are indicated at time of testing. 83
IV.2 Average slopes and standard deviation for each condition, for the analysis of slope changes depending on rate and paradigm. All 11 participants are considered ; the number of values indicates the number of slope calculation that were possible. 97
IV.3 Pairwise comparisons between responses after a higher-level pulse and a lower-level one, for levels L3 to L11 in the RANDOM paradigm at 300 pps. 107
IV.4 Time constants evaluated in the present study. 111
IV.5 Time constants of classical adaptation and DRA in the literature. . . . 111
A.1 Mean detection thresholds for the 10 selected participants for experiment 1. They increased with the mean level of the context. 133

B.1 Sensitivity index $\delta'(L)_{type,CFi}$ in the ERB centered around 2 kHz and the three fiber types in figure B.1 for experiment 1, for both standard levels in each context. 135
B.2 Sensitivity index $\delta'(L)_{type,CFi}$ in the ERB centered around 2 kHz and the three fiber types in figure B.2 for experiment 2, for both standard levels in each context. 137
B.3 Sensitivity index $\delta'(L)_{type,CFi}$ in the ERB centered around 2 kHz and the three fiber types in figure B.3 for experiment 4, for both standard levels in each context. 139
B.4 Sensitivity index for both standard levels for each context and fiber types in figure B.4 for experiment 6 : three-points linear regression around the standard level. 139
C.1 Classification of the different encountered profiles for all paradigms, 300 pps temporal evolution. 147

Introduction

Hearing is one of the most developed and used senses in human ; most of us use it every day for communication, entertainment, and even unconsciously for spatial awareness and threat identification.

Our ears are constantly solicited ; yet, they can encounter very different sound environments whilst maintaining great performance. Our perception of sounds is defined by many factors, that allow us to perceive source type, distance, direction, meaning, emotion... and decide how to react to it. To many, and in cultures all around the world, speech and music are and have been used as ways to connect and interact with nature, deities, ancestors, peers... and self. The understanding of a sense as important and complex as hearing is essential to provide better solutions for hearing protections and prosthetics.

The perception of sound intensity is fundamental in auditory perception. It is highly linked to loudness and speech perception, that are very important for one's well-being and social life. In many languages, part of speech is dependent on fast (10s to 100s of ms) and modulated changes in sound intensity to convey meaning. Changes in sound intensity are also very important in carrying emotions and intent : in music, it gives life and character to a piece ; in speech, it allows us to go from a soft, caring whisper, to a loud and angry cry.

Although there are countless studies on all these topics and more, the mechanisms underlying the preservation of intensity information throughout the auditory system are not completely understood as of yet, considering how wide the perceived level range can be, and how good the resolution is. These mechanisms are of first importance when it comes to finding encoding strategies for hearing prosthetics, to reconstruct interpretable signals as accurately as possible.

Sound intensity perception is affected by context : this has been shown directly in detection tasks, intensity discrimination experiments, or amplitude modulation detection tasks, but also in loudness or speech perception studies. Listening history can lead to enhancement or reduction of the perceived intensity of sounds.

This is also true at the physiological level : the auditory system itself adapts its response depending on the sound statistics. One of these adaptation mechanisms concerns the dynamic range of the auditory neurons, that shifts to best encode the most probable levels. This phenomenon appears to be important to expand the range of accurately encoded sound levels in the auditory system, and the preservation of small intensity changes ; and as such, it is important for sound perception in general.

Gaining fundamental knowledge on how intensity is processed and represented in the auditory system under continuous stimulation nevertheless implies tight control of the environment and stimuli used for the experiments. This is an important limitation, as it precludes the use of natural sounds and everyday contexts. Observations in such conditions allow us to conclude on specific aspects of the results, stimuli and hypotheses, but can only reflect part of the processes that take place in more ecological environments.

This thesis focuses on the effects of continuous stimulation and level variation on the perception and encoding of intensity in the human auditory system. Both psychophysical and physiological aspects are investigated, in carefully monitored environments.

The first chapter establishes the important notions and provides an attempt at reviewing the literature of interest to understand the motivation and hypotheses behind this work. It starts with a broad presentation of the auditory system before focusing on the encoding of intensity. Different context effects on loudness and intensity discrimination are then discussed, followed by a presentation of dynamic range adaptation phenomenon, which is the mechanism that modifies the encoded level range of individual neurons depending on the previous short-term stimulation.

The second chapter describes the experiments we performed with pure tone and narrowband noise stimuli, in an attempt to find a perceptual correlate of this dynamic range adaptation using an intensity discrimination task in human. Different levels and stimuli were tested.

The third chapter provides another attempt at finding a psychophysical correlate of dynamic range adaptation, but this time using broadband signals.

The forth chapter tackles the second aim of this thesis : recordings of auditory nerve activity were performed with cochlear implants. Different stimulation paradigms with various level ranges are compared ; moreover, data on temporal evolution of the response was gathered.

Chapter I

The auditory system viewed by a psychoacoustician

I.1 Physiology

This section aims at reminding the reader of the physiology of the auditory structures of interest in this thesis, the potential mechanisms encoding sound intensity, and presenting cochlear implant as a research tool and peripheral auditory prosthesis.

I.1.1 The ear and auditory pathways

I.1.1.1 Periphery

The auditory periphery is considered as going from the large cartilaginous appendix commonly referred to as the ear - actually named pinna - to the cochlea, encompassing the outer, middle, inner ear, and sometimes the auditory nerve.

The "ear" is commonly divided into three parts, as shown in figure I.1.

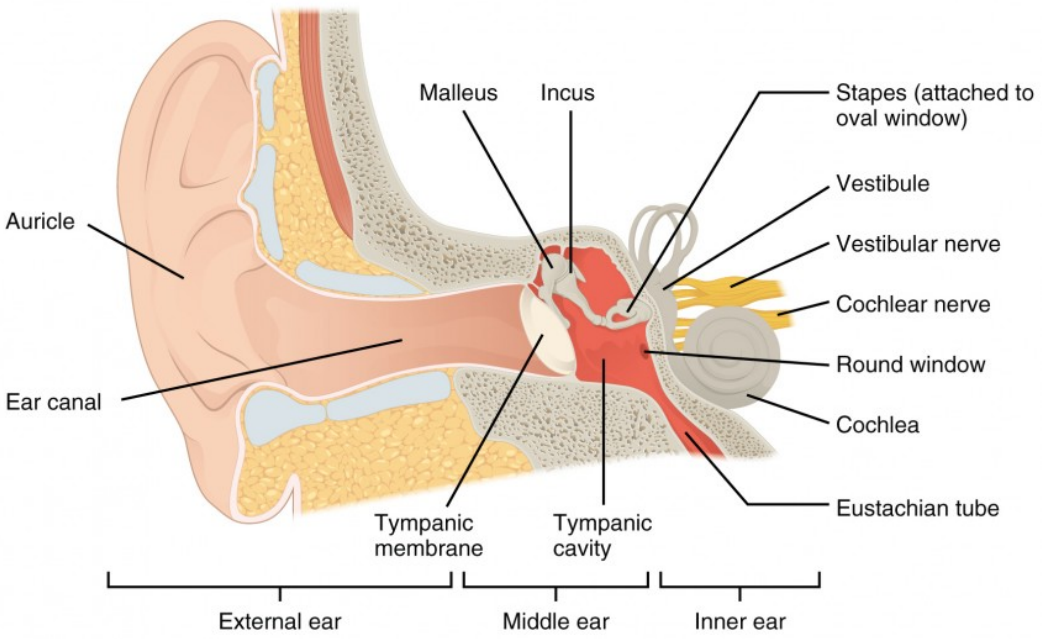


Figure I.1 – The different components of the ear or auditory periphery, divided into the outer (external), middle and inner ear ; credits : OpenStax College.

– The outer ear comprises the pinna, or auricle, and the auditory canal, funneling and filtering sounds from the outside world to the tympanic membrane (also called ear drum), which separates the outer ear from the middle ear. It transforms the vibrations of the air into a mechanical movement.

– The middle ear contains the three smallest bones in the human body, called the ossicles : malleus, incus and stapes. They are connected to the tympanic membrane, and move according to its vibrations to push on the oval window of the cochlea. Their role is impedance matching, between the air and the liquids inside the cochlea (perilymph and endolymph).

– The inner ear is essentially the cochlea itself, along with the vestibule and semi-circular canals : the latter are indicators of our position in space and give a reference for balance. Inside the cochlea's curled structure are three canals filled with liquid ; two of them are connected at their top end, the apex. In the third canal is the basilar membrane, on which is resting the organ of Corti. It is full of small hair cells. These cells are mechanotransducers, triggered by the movement of the basilar membrane caused by the traveling waves created by the ossicles ; they turn this mechanical force into electrical signals, which are then sent through the auditory nerve to higher centers.

Most of the hair cells are outer hair cells (OHC). Their role is to control - either enhance or suppress - the movement of the basilar membrane, and thus the gain of the signal, while the inner hair cells (IHC) are transmitting the signal itself.

An important property of the basilar membrane, which is also found in many other sites of the auditory system, is tonotopy : the hair cells are ordered in the cochlea in descending frequency response range. This means that the cells near the oval window, where the stapes pushes on the cochlea, are responding to high frequencies, while those at the apex respond to low frequencies. This can be explained by the mechanical properties of the basilar membrane : being stiffer and narrower at the base, it vibrates at high frequencies, while the apex is larger and more flexible, and captures the low frequencies.

The auditory nerve is the limit of the auditory periphery. It gathers all the nerve fibers (named spiral ganglion neurons in their full form) getting outputs from the cochlea, and directs the information to higher stages of the auditory system. In the absence of stimulation, auditory nerve fibers have some spontaneous activity to maintain function and connectivity. Nerve fibers have preferred frequency ranges to which they respond more, depending on the IHC they are connected to : the most sensitive part of this range is called their best, center, or characteristic frequency (Wever 1938 ; Greenwood 1990 ; Olson and Mountain 1991).

There are also efferent fibers connecting a feedback system to the cochlea (described later). The efferent fibers convey the information going from higher centers to the periphery, to modify its behavior.

The acoustic reflex is a mechanism affecting the response to sounds in the periphery. If a loud sound is heard, small muscles will stiffen the tympanic membrane and ossicle chain in the middle ear, and after a few milliseconds their movement in response to the

sound will decrease, so less power will be transmitted to the cochlea. This mechanism protects the inner ear from loud damaging sounds.

I.1.1.2 Higher centers

The rest of the auditory system is more intricate, as it involves many sites that interact and communicate both bottom-up and top-down. Some of the most important for this manuscript are represented in figure I.2.

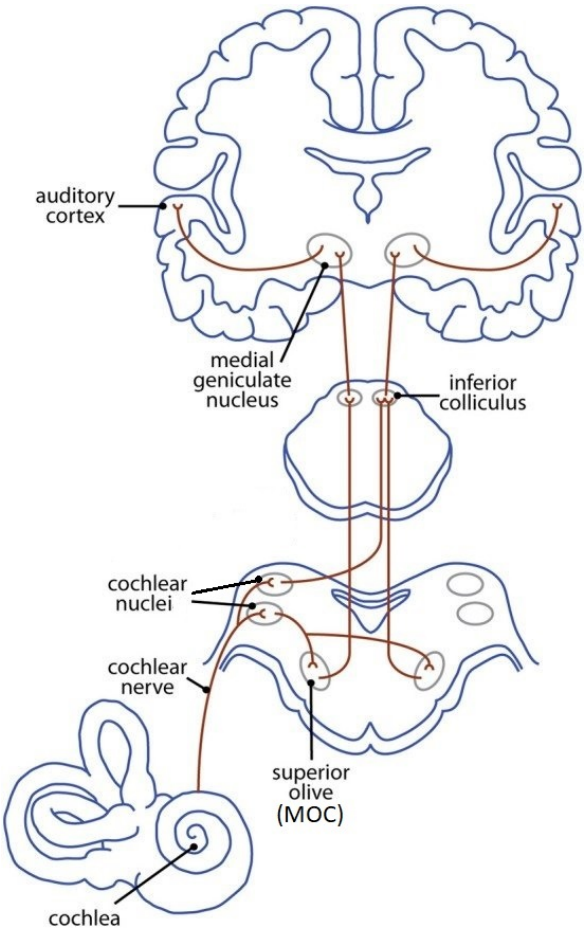


Figure I.2 – Main centers of auditory processing throughout the auditory pathway ; the auditory nerve connects the cochlea to the brainstem which contains the cochlear nuclei and the olivocochlear systems. The information is then fed to the inferior colliculus in the brainstem and the medial geniculate body in the thalamus before it reaches the auditory cortex. From Butler and Lomber (2013).

The cochlear nuclei - ventral and dorsal - are the first relays of the auditory signal from the auditory nerve to the brain. They separate the information into its temporal and spectral characteristics and send it to parallel processing pathways (Pannese et al. 2015).

The auditory system is often described starting from the ear, where sounds enter the system, and ascending to the brain, where they are transformed into a percept and given meaning. However this ascending path is completed by a descending one, and many crossings between neural centers. The bottom-up part of the auditory system is

called the afferent system, while the efferent part is a top-down one originating in the olivary complex, which contains the superior olive (Ciuman 2010).

The medial olivocochlear (MOC) bundle is an important part of the efferent system, located in the brainstem, next to the superior olive : efferent nerve fibers link it to the cochlea. It regulates the OHC activation and thus the basilar membrane mobility, hence controlling the gain of the transiting information from the IHC to the brain (Guinan et al. 2003 ; Guinan 2006 ; Backus and Guinan 2006).

The inferior colliculus is part of the midbrain. It is a collector : it gathers activity from the numerous sites processing auditory information in the brainstem and below, and sends them to the thalamus, which then connects the auditory cortex. It is an important site for sound localization. It is also involved in the efferent system : it receives and feeds back information from higher centers (Liu et al. 2024).

Finally, the auditory cortex is the part of the brain where the auditory inputs are converted into percepts. It is connected to other areas of the brain, among which the visual cortex and centers controlling movement.

I.1.2 Hearing impairment and cochlear implants

If there is a malfunction somewhere in the auditory system, it can lead to hearing impairment. There can be many causes for hearing loss : if the outer or middle ear are affected, the hearing loss is conductive. From the inner ear to the auditory nerve, it is sensorineural ; but if the origin is in higher centers, the loss can be termed as central or functional. Conductive hearing losses can be diagnosed by comparing hearing thresholds from acoustic stimulation with those obtained with bone conduction, with a vibrator placed on the skull (behind the ear). A sensorineural hearing loss can need other techniques for diagnosis because the transmission of sound by the outer and middle ear is normal.

The encoding of sound intensity in ears with different degrees of hearing loss and pathologies has always been a concern for hearing aid companies. For severe and profound hearing losses with damage to the auditory periphery (sensorineural hearing loss), a now rather common solution is the cochlear implant (CI), presented in figure I.3.

This device is composed of two main parts : the sound processor is placed behind the ear like a lot of other hearing aids, with an emitter placed on the skull, slightly higher back from the ear. The other part is implanted : a receiver is placed under the skin to face the emitter. It receives information from the emitter and stimulates several electrodes placed inside the cochlea, which deliver electric pulses to the auditory nerve (Macherey and Carlyon 2014).

These pulses can differ from one processor (or brand) to another, as do the encoding strategies and the stimulation paradigms, but in general, modern devices use monopolar, biphasic, charge-balanced pulse trains : the electrodes in the array are used for the stimulation, and external, ground electrodes are used as reference electrodes ; each pulse has a positive/anodic and a negative/cathodic phase with equal energy and duration, sometimes separated by a short interphase gap.

The overall principle behind signal processing by the implant is presented in figure I.4. The acoustic signal is filtered first to simulate the effects of the outer and middle

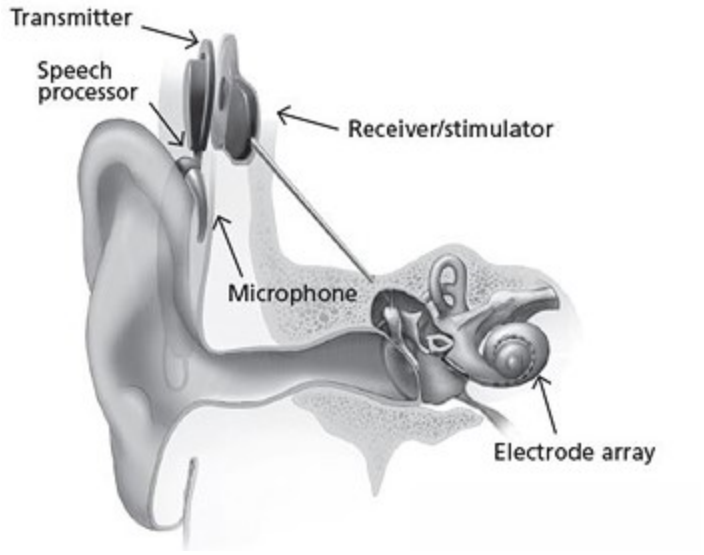


Figure I.3 – Diagram of a cochlear implant in the human auditory system. The processor and its microphone are external and generally placed around the ear, while the transmitter is magnetically bonded to the receiver, under the skin. The signal then travels to the electrode array, surgically inserted inside the cochlea. Source : NIH/NIDCD.

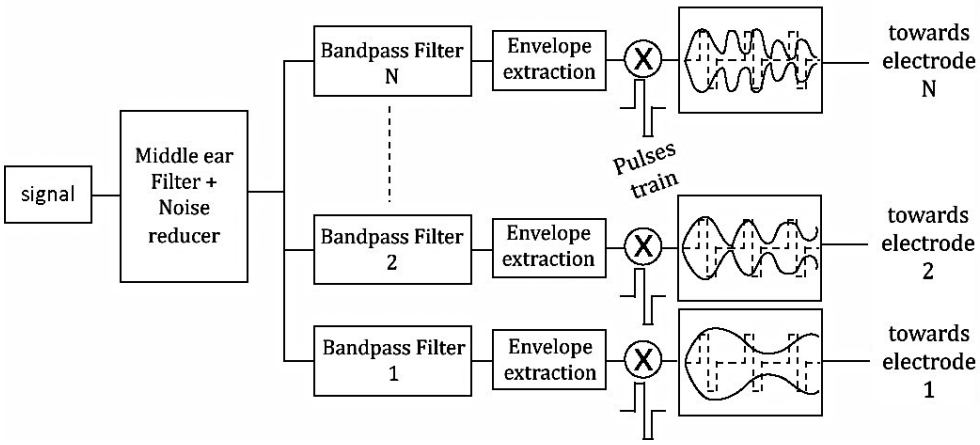


Figure I.4 – Filtering process of a sound by a cochlear implant : the signal is first filtered as the outer and middle ear would ; it is then split through different bandpass filters ; the envelope of the output of each filter is then convoluted with electric pulse trains, and each electrode receives amplitude-modulated pulse trains matching the envelope of the filtered signal.

ear (with more or less gain depending on the frequencies and levels) ; noise reduction algorithms are also applied. The signal is then divided in several frequency bands and filtered again using a bank of bandpass filters that extract the temporal envelope of the signal, simulating the effects of the cochlea. Finally, the envelope of each filter is convoluted with electrical pulse trains : the output of each filter is translated into amplitude-modulated electrical pulses, then sent to the corresponding electrode. Since the electrodes are placed at different depths in the cochlea, they each face a different frequency region, and the filtering of the signal is done in order to make

use of tonotopy. Some processors have more filters than electrodes, and use channel interactions to create intermediate, virtual channels. The number of channels defines the frequency resolution of each filter. However, some cochlear implant users cannot distinguish between neighboring electrodes. This can be attributed, at least partially, to the distance of the electrodes to the auditory nerve fibers, or to differences in neural survival (Hughes 2013 ; Macherey and Carlyon 2014).

One challenge in using this device is to know how to best encode the auditory information as a series of pulses in terms of amplitude, dynamic range, electrode choice, pulse rate...

And here comes the treat. An interesting feature of the electrode array is that it is not limited to stimulation : it is also able to record the neural response from the excited auditory nerve population (Lai 2004). This also makes cochlear implants an interesting tool for auditory research.

Measures of neural activity with cochlear implant is performed by using electrically evoked compound action potentials (eCAPs) ; details on the methods are given in Chapter IV.

I.1.3 Intensity encoding

Intensity, as defined in physical acoustics, is the measure linked to the pressure and velocity of the particles displaced by a soundwave, expressed in watts per square meter. In this work and in auditory research in general, the term is - and will be - used to refer to sound pressure levels, expressed in dB SPL (decibel sound pressure level) and measurable with a microphone. Physically, sound intensity is the sound pressure multiplied by the velocity of the particles : the speed and direction of the sound is taken into account in intensity, but not in sound pressure. Differences in sound pressure level are expressed in dB. Intensity is encoded by changes in neural activity in response to changes in sound pressure level, through various mechanisms (Viemeister 1988), described thereafter.

At the periphery, sound intensity is primarily reflected in the pressure at the tympanic membrane and in the basilar membrane's displacement and velocity. For both measurements, the higher the intensity, the greater the velocity. Ruggero et al. (1997) tested basilar membrane movement in healthy chinchilla cochleae. They recorded responses in different places to different levels and frequencies. They found an increase in the amplitude of the vibration with increasing sound level ; first linear at low levels, then with a much shallower slope. This change in behavior disappears in dead cochleae, with only a linear evolution of the vibration amplitude. This is due to the absence of compression induced by OHC electromotility.

The central question in this section is which neural information present at the level of the AN is used by the brain to perceive intensity ; several possibilities are discussed.

I.1.3.1 Rate-place coding

The next mechanism encoding intensity is the discharge rate of the nerve fibers ; the louder the signal, the higher the discharge rate, and it can be directly calculated from the pressure at the tympanic membrane as well as the basilar membrane displacement

(Sachs and Abbas 1974). However, the activity of the auditory neurons is inherently stochastic (Siebert 1965) : no neuron will respond twice in the same way to identical and independent stimuli, and no pair of neuron will respond the same.

The types of responsive nerve fibers are also indicative of the sound level, depending on their threshold and saturation levels. Indeed, there are differences in spontaneous firing rate and encoded level range (dynamic range) in the neural population. The rate-level function of a neuron is classically measured by playing one stimulus level at a time and recording the firing rate of the neuron in response to this stimulus, with several repetitions for averaging. The dynamic range of a neuron is the level range between its activation threshold (the level at which a significant increase in discharge rate compared to the spontaneous rate is observed), and the level at which it saturates (with a large enough slope reduction to appear much flatter than the steeper part). These two points are often taken respectively as 10 and 90 % of the difference between minimum and maximum values in the rate-level function. The neural population is generally divided in two groups : the low-spontaneous rate fibers (1 or less up to around 20 spikes/s), and the high-spontaneous rate fibers (from 20 to 100 or more spikes/s, Jackson and Carney 2005).

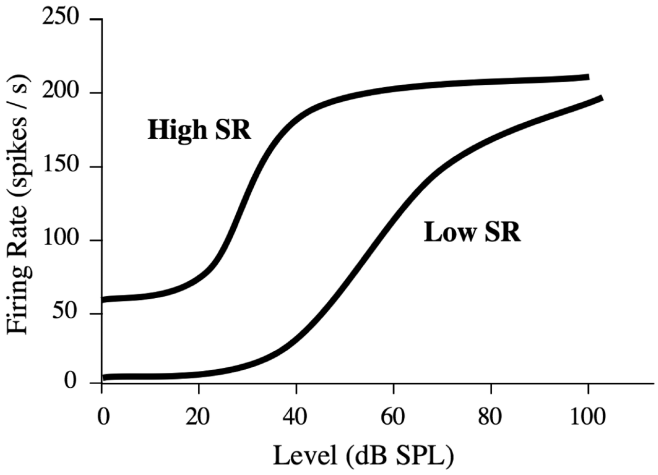


Figure I.5 – Theoretical examples of rate-level functions for two types of fibers in the cat auditory nerve : the low-spontaneous rate (SR) at the bottom, and high-SR fibers on top. From Plack (2013).

The low-spontaneous rate fibers have a high-level activation threshold (lower curve in figure I.5), and a wide dynamic range with a shallow slope, sometimes even non-saturating. They mostly code medium-high levels. The majority of fibers are high-spontaneous rate fibers (upper curve), encoding lower levels, with a short and steep response function, and flat saturation profiles. They represent approximately 61 % of the neural population in cat (Liberman 1978). Sometimes a third group is distinguished from the low-spontaneous rate fibers : it is called the medium spontaneous rate fibers, which have a behavior somewhat in between the two groups (Sachs and Abbas 1974 ; Jackson and Carney 2005).

Because the different types of fiber have different dynamic ranges, one theory is that they are combined to get a broader dynamic range, at the population level : this is

called dynamic range stitching (Viemeister 1988). The HSR fibers would be encoding low to medium levels, and the LSR taking over for loud levels. There are however very little LSR fibers compared to the HSR ones : low and medium levels could be expected to be better represented in intensity encoding.

Sound intensity is also translated into the spread of excitation (Sachs and Abbas 1974) : because of the frequency tuning of the cochlea, only certain neurons react to certain frequencies, and when the sound level increases, a broader part of the basilar membrane gets progressively excited, inducing activity in fibers that are not specifically tuned to that frequency. The spread is more important on the high frequency side than on the low-frequency side (upward spread), even if it is controlled by basilar membrane compression, which limits the activity in neighboring regions while enhancing it in the region of interest. This controls the cochlea's very fine frequency tuning. The spread of excitation is however limited to pure tones or narrowband signals ; a wide signal could be exciting a very broad part of the auditory system at relatively low levels.

I.1.3.2 Temporal coding

Another key factor is the discharge synchrony (or phase locking), observed in auditory nerve fibers excited by pure tones (Johnson 1980 ; Colburn et al. 2003). Individual fibers tend to respond to specific phases of the signal, but not necessarily the same for all fibers, and they do not respond to each period of the signal. Synchrony is translated at low levels as a change in the timing of the spontaneous activity. For increases in level at very low sound levels, each fiber tends to progressively respond more during one half-period of the signal, before any increase in the average firing rate in response to an increase in the level of a sound. The fibers synchronize their activity to the shape of the signal before increasing their activity. Then, the neurons' individual firing rates only start to increase slightly below the synchrony saturation level (maximum observed synchrony) : synchrony thus provides information at levels at which they do not display increases in discharge rate (Johnson 1980). For some fibers with low characteristic frequencies, stimulation at lower frequencies sometimes results in a division of the population : responses to the start of each half period of the signal can be observed, and despite being greatly attenuated compared to the first half, the second half period is also represented in the temporal discharge pattern. Finally at medium-high levels, the relative phase at which the neurons discharge changes with the level : further increases in level cause the neurons to respond more to a specific phase of the signal, and to synchronize at the population level.

The phase at which the neurons will discharge is relative to the absolute frequency of the signal, not to the characteristic frequency of the excited fibers (Johnson 1980). At the population level, it is even possible to observe a crude representation of the rectified soundwaves when looking at discharge patterns over time (period histograms). However, these phase cues are believed to disappear around 5 kHz (as is the case in cat), because the fibers do not have enough resting time between periods and are unable to follow the temporal fine structure of the signal.

I.1.3.3 Other theories

The mechanisms previously described, even combined, cannot encompass the whole range of human hearing, and there are several reasons for that. The main argument against the effectiveness of combining the information given by the different fiber types, called dynamic range stitching, is the non-uniform distribution of neurons between types. A majority of neurons are HSR fibers ; they mostly code low levels. An overview of the possible intensity coding mechanisms by Barbour (2011) states that the diversity of thresholds, dynamic ranges, saturation levels and behaviors, can be used over most of the low-medium range of hearing, with numerous fibers for low levels, but that few fibers code for high levels.

More regions of the basilar membrane are activated with a level growth because of the spread of excitation, transmitting information to non-saturated cells. If the auditory system effectively combines the information gathered in different frequency regions, it is fair to think that spread of excitation extends the dynamic range of hearing. This effect would nevertheless be limited for broadband noises because there is already energy in different channels. The combination of temporal cues (synchrony and other phase cues) and rate-place cues (discharge rate, fiber types, spread of excitation) can account for a more important part of the hearing range, but the high frequencies and high levels are not as well represented as the rest of the hearing range.

Other theories on intensity encoding have surfaced. Heinz et al. (2001b) suggested that, by comparing the relative discharge phase of two neighboring frequency regions, it is possible to have a coincidence counter : discharges arriving at the same time from fibers with different characteristic frequencies are likely to be caused by the same signal. As phase locking is dependent on signal frequency and on level, the louder the sound, the more coincidence would happen. The number of simultaneous discharges may therefore provide information on the level of the sound and thus provide a cue for intensity discrimination. However, it cannot be very effective at high frequencies because of the roll-off of phase locking, nor at high levels because of the saturation of fibers. Note that these phase cues have also been hypothesized to encode the pitch of pure tones (Carlyon et al. 2012).

The response of neurons to sounds involves changes at different time scales : the one of interest in this paragraph is the low frequency fluctuation of neural response (Carney 2018 ; Brennan et al. 2023 ; Carney 2024). In a frequency region that consistently receives a lot of information (at high stimulation levels), fibers will be saturated, and the temporal envelope of their response will have little fluctuation. Fibers with characteristic frequencies further away from excitation frequencies will receive less steady information and display greater fluctuations. This reasoning is developed in figure I.6. That is where mid-level saturation becomes interesting : if two neighboring frequency regions have very different fluctuations profiles, the sound is rather soft, but if the sound is loud, close frequency regions will have similar (weak) fluctuation profiles because of the spread of excitation.

The role of the cochlear amplifier, efferent systems and adaptation mechanisms is consistent with this theory : they tend to enhance the differences in fluctuation profiles. Their role is to reduce the response to loud sounds and increase the response to soft sounds, thus increasing the fluctuations in both cases. The relative fluctuations of

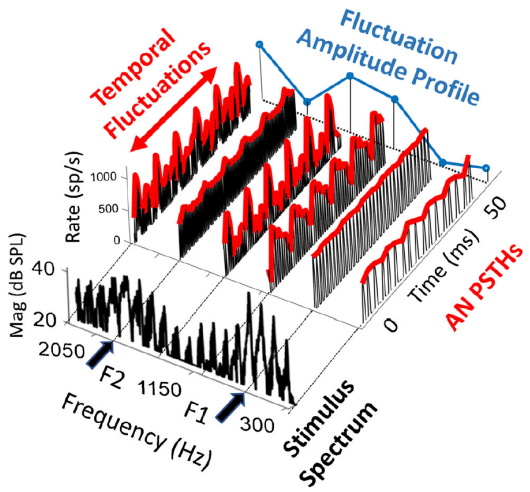


Figure I.6 – Envelope fluctuation profiles : the AN response of several frequency regions to a vowel over time is presented, along with the spectrum of the corresponding sound and the resulting amplitude fluctuation profile. The regions centered around the peaks in the spectrum have low amplitude fluctuations, while the in-between ones have higher fluctuations. From figure 2 in Carney (2018).

different frequency bands could thus encode intensity : a fluctuation profile with low values on most frequency bands will appear for a loud sound, while high fluctuations will indicate for a soft sound, and important differences throughout the frequency range will correspond to medium levels.

I.2 Perception of sound intensity

Psychoacoustics being the study of the relation between the acoustical properties of a sound and its perception by humans, it encompasses very diverse high- and low-level percepts and features. Our main interests for this work are the most related percept of intensity, namely loudness, and the measurement of intensity discrimination performance. This literature review also focuses on the various effects of an acoustic context on loudness and intensity perception, and thus requires the introduction of masking.

I.2.1 Masking

An important notion throughout this thesis is masking. One definition is given by Durlach (2006) :

> In general, masking can be said to occur whenever the reception of a specified set of acoustic stimuli (“targets”) is degraded by the presence of other stimuli (“maskers”).

With a definition that broad, any kind of interaction worsening performance in any psychophysical task can be classified as masking - it is not limited to detection tasks. Buss (2008) develops :

> Masking is said to be energetic when the peripheral response to the masker interferes with response to the signal; this occurs when the auditory channel or channels best representing the signal are also excited or suppressed by concurrent masker energy. [...] Threshold elevation that cannot be attributed to energetic masking is sometimes described as informational masking. Informational masking is typically assumed to occur central to the cochlea and is thought to be due to stimulus uncertainty and/or perceptual similarities between masker and signal.

This nomenclature is however highly limited, in that the definition depends on the reference for energetic masking we take in the auditory system, and that informational masking is a waste-basket category. It encompasses everything that is not energetic : effects of stimulus uncertainty, similarity, listener attention and memory... Watson (2005) warns against the use of such broad terms.

Moreover, there are no simple rules for masking additivity : the effect of several maskers on a target can be (and often are, see Durlach 2006) different from the sum of their respective effects on that same target.

In this manuscript, we will refer to energetic masking (peripheral, increasing detection thresholds) as simply "masking", while the different types of informational masking will be referred to with their most-common denomination in the literature we base it on.

I.2.2 Loudness

Loudness is the perception of the strength of a sound : it ranges from just audible to painful ; but it is mostly used with the terms "soft" or "quiet", and "loud". It is mainly linked to the intensity of a sound, but not exclusively.

The loudness unit is the sone : a sound twice as loud as a reference will have double the sone value ; a sound half as loud will have half the rating.

The loudness level is expressed in phones, relative to a 1-kHz pure tone ; a sound as loud as a X-dB SPL, 1-kHz pure tone will be at X phones. It is often used to plot equal-loudness curves, as shown in figure I.7. They represent the physical levels perceived as loud as a 1-kHz pure tone at different frequencies, from the absolute threshold up to the pain threshold - which is roughly around 120 to 130 dB SPL for most frequencies -, or the discomfort threshold, slightly lower. These graphs illustrate rather well the perceptual dynamic range of the different frequencies in human hearing. While normal ears are indeed very sensitive between 1 and 4 kHz, loudness takes notable changes in intensity to build up ; but at lower and higher frequencies, with much higher absolute thresholds and a similar pain threshold, small intensity variations translate into greater loudness differences (Relkin and Doucet 1997).

Loudness is affected by the duration of a sound : longer sounds are louder, up to a few hundred milliseconds for medium frequency pure tones. This is called the temporal integration of loudness. Loudness increases when the bandwidth of a signal increases above an auditory filter bandwidth, while the intensity remains the same. This is the spectral integration of loudness (Zwicker et al. 1957). The attack and decay time of the signal can also influence its loudness (Buus et al. 1997 ; Boullet et al. 2004).

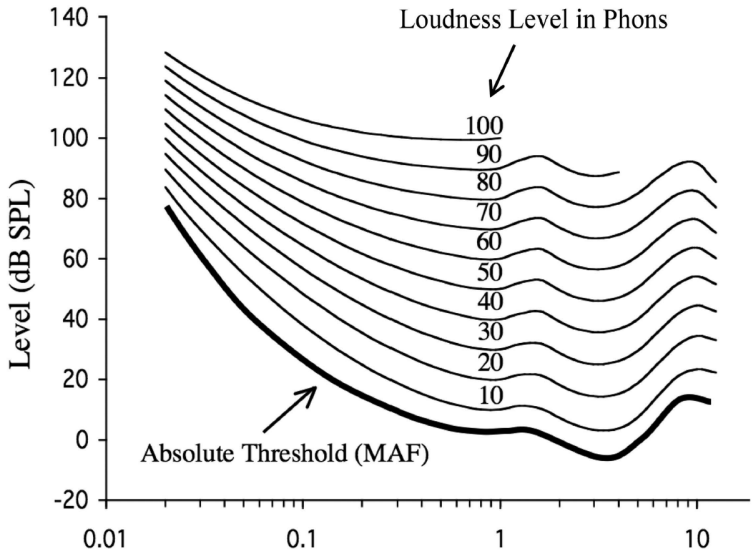


Figure I.7 – Equal-loudness curves : all the pure tones on a line have the same loudness level in phones. The lowest curve represents the minimum audible field (MAF) ; each curve above gives the level at which a pure tone of the selected frequency will sound as loud as the 1-kHz reference at this level. From Plack (2013).

The evolution of the loudness L of a sound with its level is often represented as a loudness function ; generally, Steven's Power Law is used as a reference for pure tones : at 1 kHz, $L = kI^{0.3}$ (Stevens 1966). In later work, loudness has been proved to behave differently in noise (Stevens 1966 ; Stevens and Guirao 1967 ; Moore and Glasberg 1993) : the loudness functions are steeper at lower levels with an elevated threshold compared to Steven's Power Law, then fold back to the classic loudness function at medium-high levels.

Loudness is difficult to relate directly to neural activity (Relkin and Doucet 1997) : the discharge rate of neurons is not a sufficient indicator for loudness. However, taking synchrony and spread of excitation into account gives better estimates of loudness. There has also been some research on loudness encoding in higher processing centers : temporal effects (integration) seem to take place at sub-cortical levels (Thwaites et al. 2016), while spectral integration is happening higher up, in the primary auditory cortex (Röhl et al. 2011 ; Moore et al. 2016, 2018 ; Fischenich et al. 2021).

It is also important to consider that the loudness of a sound is more important if the latter is presented to both ears (binaural) than only to one (monaural). However, loudness summation is not perfect : a diotic sound is less than twice louder than the same sound presented monaurally (Marozeau et al. 2006 ; Sivonen and Ellermeier 2011 ; Epstein and Florentine 2012).

The loudness of a sound can also be altered by the presentation of other sounds. The simplest example is masking : if one sound is presented simultaneously with another, it will sound weaker than if it is presented in isolation (Epstein 2007). Moreover, it is important to consider the temporal arrangement of the sounds when studying context effects on loudness. One pure tone presented after a first high-level sound can be perceived as quieter than when presented in isolation, and that same tone presented

before that same sound can also sound softer (Stevens and Guirao 1967) : this is called induced loudness reduction. However a first sound can also increase the loudness of a second sound (Plack 1996 ; Oberfeld 2007 ; Oxenham 2018) : this is loudness enhancement, or recalibration. It is generally observed for medium level differences between the loud masker and the tone.

There is also an influence of the nature of sounds. While Carlyon and Datta (1997) found that more peaky signals sound softer, Gockel et al. (2002) describe harmonic complexes with amplitude-modulated, peaky envelopes as louder than noise, with an impact of the relative phase of the different components. The reduction in loudness of peaky signals is consistent with basilar membrane compression, which tends to flatten the excitation profiles. The results from Gockel et al. (2002) however indicate that other factors such as annoyance, pitch strength and crest factor (which was much higher than for Carlyon and Datta 1997) could directly influence the loudness of their harmonic complexes.

Hence, loudness depends on the duration, time intervals, frequency content, and presentation order of the signals.

Another important aspect of loudness is that it is very different in normal hearing people and in the hearing-impaired population, hence the numerous studies attempting to compare different situations. For example, recruitment refers to the increase in absolute threshold with a steeper loudness growth function, similar to that observed for loudness in noise. It is a reduction in dynamic range in that thresholds get higher and comfortably loud levels remain the same, effectively yielding a steeper loudness growth. Recruitment is commonly observed in participants with hearing loss, and it leads to the degradation of speech perception (Moore and Glasberg 1993).

The end goal of comparing the normal hearing and hearing-impaired populations is often to try to improve our understanding of the physiological mechanisms behind loudness, and to provide better signal processing solutions for hearing aids and cochlear implants. This is especially true because noise influences performance sometimes differently in normal hearing and hearing impaired or cochlear implant users (Moore et al. 1985 ; Pienkowski and Hagerman 2009 ; Wang et al. 2015 ; Oxenham 2018 ; Marrufo-Pérez and Lopez-Poveda 2022) ; it informs us on the mechanisms behind certain phenomena (Stainsby and Moore 2006 ; Adel et al. 2017 ; Brennan et al. 2023 ; Carney 2024).

I.2.3 Intensity discrimination

Loudness is often related to another important measure in auditory perception : intensity discrimination (Schneider and Parker 1987 ; Plack 1996). Intensity discrimination seeks the smallest perceivable difference in sound level between two sounds, called the intensity discrimination threshold, just noticeable difference, or intensity difference limen. In the paradigm used in this thesis, one of these two sounds is the standard : it has a fixed level. The level of the second sound, called signal or target, is varied, but always stays above the level of the standard. The listening participant has to decide which of these two sounds has the higher level, as they are presented in random order. This measure is used to formulate and test hypotheses on the mechanisms behind intensity encoding. The participant can also be asked to find

the lower of the pair, and the process is inverted.

Siebert (1965) has been the first, using an analytical model with many individual auditory nerve fibers, to compare limits in simulated neural activity differences to limits in human performance in intensity discrimination. They combined fibers with different characteristics to try to define a limit discrimination performance, while taking into account the statistical variability in the response of each neuron. Using signal detection theory (Siebert 1968), they set the ground for a huge research tool with various signals : tones, noise, or tones in noise. The response of the simulated neurons were generally able to predict the order of magnitude of human performance in intensity discrimination but with discrepancies in its evolution when changing one parameter.

Intensity discrimination is known to be influenced by many factors, first of which are level and frequency. In humans, performance is about as good at all standard levels for broadband noise ; Weber's law states that at the discrimination threshold, the difference in intensity divided by the intensity of standard is a constant (Miller 1947). It appears that for pure tone stimuli, the results diverge slightly : the thresholds decrease when the standard level increases. This is the near-miss to Weber's law (McGill and Goldberg 1968 ; Moore and Raab 1974 ; Florentine et al. 1987). It is thought to arise from the spread of excitation : at high levels, more non-saturated fibers are available to encode intensity, and the combined information from all these fibers is more effective than that of less fibers tuned to the signal's frequency at lower levels. Consistently with this theory, it has been shown that presenting a notched noise in an intensity discrimination task with a pure tone increases the thresholds at medium-high levels, bringing back Weber's law (Schneider and Parker 1987). Intensity discrimination of noise bursts has been found not to depend on the bandwidth of the noise (Moore and Raab 1975 ; Schacknow and Raab 1976). Micheyl et al. (1997) studied intensity discrimination in broadband noise (0-15 kHz) ; they concluded that the MOC reflex, which can be monitored through otoacoustic emissions, was correlated with intensity discrimination performance.

At very high frequencies, performances in intensity discrimination start to worsen. This is supposedly linked to the spread of excitation that is limited when reaching the edges of the basilar membrane (Jesteadt et al. 1977 ; Florentine 1983 ; Florentine et al. 1987). Concerning duration, the longer the tone, the lower the discrimination threshold, up to 500 ms or more for most frequencies (Florentine 1986). This may be linked to temporal integration mechanisms : it has been proposed that stimulus energy is stored and summed over certain durations. Another theory is the multiple looks model (Viemeister and Wakefield 1991) : the auditory system is viewed as a detector that has limited observation capacity, and the longer the stimulus, the more looks it can take to detect it.

There is another deviation to Weber's law in pure tone discrimination. The discrimination thresholds at medium levels get much worse than at lower or higher levels, sometimes by up to 10 dB. This has been shown at high frequencies for different signal lengths (Carlyon and Moore 1984), and was seemingly unaffected by the addition of masking noise (Roverud and Strickland 2015b).

Forward and backward masking effects are present in intensity discrimination tasks : a noise presented before or after the stimulus will generally decrease the loudness of

the latter.

The mid-level hump is the increase of thresholds at medium level, generally observed in forward and backward masking (Plack and Viemeister 1992a,b). Zeng (1998) found it with a 1-kHz, 25-ms tone in the presence of a 1-kHz, 100-ms pure tone forward masker. Adding a notched noise did suppress their hump but not the increase in thresholds compared to the absence of masker.

Carlyon and Moore (1984) found the mid-level hump for tones at 4 and 6.5 kHz. They hypothesized that at medium levels and high frequencies, phase locking is not available, and the spread of excitation is not important enough to reduce the discrimination thresholds while it is at higher levels ; with less available information, thresholds increase. Another hypothesis for the mid-level hump is that the medium levels are too far from the two "perceptual anchors" that are the detection threshold and the discomfort level which make their loudness difficult to estimate accurately. Zeng et al. (1991) and Zeng (1998) proposed that an "intensity coding gap" (Relkin and Doucet 1991) created by the difference in recovery time between fiber types could prevent the medium levels from being properly encoded : the LSR fibers under recovery would be unable to encode them, while the fully recovered HSR fibers would be saturated. Basilar membrane compression has been identified as another possible cause by Baer et al. (2001), but as the mid-level hump has also been observed for short stimulus, the time course of basilar membrane compression would be too long to affect the stimuli ; it is unlikely to be a common cause. Zeng and Shannon (1995) also found the hump in CI users, suggesting it arises higher than in the cochlea.

Based on the previously defined notions, we can conclude that the mid-level hump might be a central phenomenon. It is probably sometimes paired with other sources of masking when in the presence of simultaneous or sequential maskers. It is related to the distance to an intensity reference, either internalized or created by a masker (Oberfeld et al. 2014).

Oberfeld (2008) used 1-kHz tones and maskers of 25 ms, and found a "mid-difference" hump, because it appears for medium masker-tone level differences. They discard the hypotheses previously described that are based on physiological lack of information at medium level, because their "mid-difference" hump is also present at low- and high levels, as long as the masker and probe had a medium (30-40 dB) difference in level.

Another phenomenon, referred to as “confusion”, is observed in signal detection tasks under forward masking (Neff 1985, 1986). Confusion refers to the inability to distinguish the signal from the masker : the signal can be detected but is not discriminated. This effect seems to be greatest for a pure tone around 40-ms long, forward masked by a narrow-band noise. Moreover, the similarity between the standard and masker as well as the fluctuations of the masker have a great influence (Neff 1986).

Confusion and mid-level hump might be related, but confusion seems more general and not limited to medium levels, or medium masker-standard level differences.

I.3 Dynamic range adaptation

I.3.1 Presentation

The more or less recent stimulation history can have various effects on the response of the auditory system itself, from the periphery upwards (Pérez-González and Malmierca 2014 ; Benda 2021 ; Willmore and King 2023), and it can also affect perception (Canevet and Meunier 1996 ; Marrufo-Pérez and Lopez-Poveda 2022).

For example, there are several forms and time scales of what is called classical spike-rate adaptation (Javel 1996 ; Meunier and Marchioni 1998 ; Ulanovsky 2004 ; Hughes et al. 2012) : the discharge rate of neurons evolves when a stimulation is long enough, decreasing over time to save resources from redundant information. It operates over a few milliseconds, up to several tens of seconds, and is present from the periphery upwards.

Costalupes et al. (1984) characterized the rate-level functions of cat auditory nerve fibers in the presence of noise. They played pure tones at the neurons best frequency, in a broadband noise that was either continuous, interrupted during the presentation of the tone, or presented simultaneously with the tone. They found a global shift of the rate-level functions of the neurons towards higher levels with increases in the level of the noise, as well as the classical reduction of the maximum rates in all three noise conditions. They suggested that the simultaneous part of the masker was producing the shift in dynamic range, while the non-simultaneous part was responsible mainly for the rate decrement. Rees and Palmer (1988) replicated these results in the inferior colliculus of guinea pigs.

Dean et al. (2005) used a different stimulus paradigm and found that the dynamic range of neurons in the inferior colliculus of guinea pigs could shift towards the most probable levels, as described below.

A broadband noise with level fluctuations was played for around seven minutes to awake, anesthetized guinea pigs, while recording the activity of several neurons in their inferior colliculus (Dean et al. 2005). The level of the noise changed every 50 ms (figure I.8 a and b) and followed a non-uniform distribution : over the whole level range, a band of 12 levels (width : 12 dB) called the high probability region (HPR, figure I.8 c) was chosen to have an 80 % chance overall of being played, while the rest of the tested level range shared the remaining 20 %. The HPR defined what they called the context, with the mean of the band being the mean presentation level. The response function of the neurons was computed by measuring the firing-rate in response to each 50-ms epoch, and was compared to that obtained with isolated 50-ms bursts.

The rate-level function of individual neurons was found to undergo a horizontal shift, compared to the classical paradigm for rate-level function measurement (grey line in figure I.9). This shift was dependent on the context mean level, and was thus called dynamic range adaptation (DRA). It is presented for several neurons and HPRs in figure I.9.

The width of the HPR mostly modified the slope of the response function, but the shift remained linked to the mean level. Indeed, Dean et al. (2005) tested a 24-dB large HPR (double the size of their original HPR), and found that some neurons changed

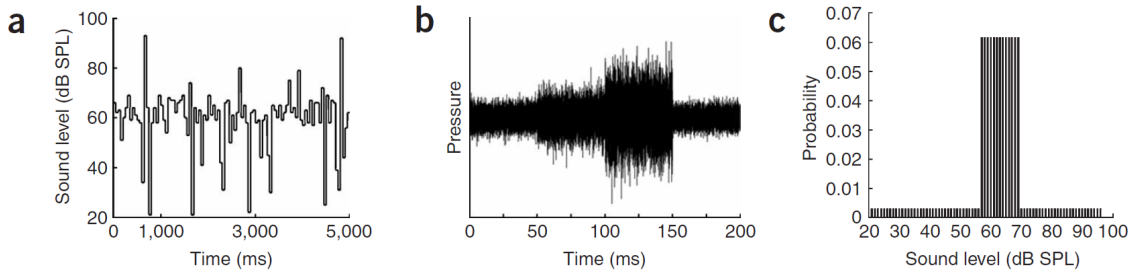


Figure I.8 – Example of a stimulus used in Dean et al. (2005). Panel a) shows the distribution of levels over time for a broadband noise and a 12-dB high-probability region (HPR) centered around 63 dB SPL ; a 200-ms extract of the signal is shown in panel b) ; panel c) shows the probability distribution of the levels (replotted from Figure 1 of Dean et al. 2005).

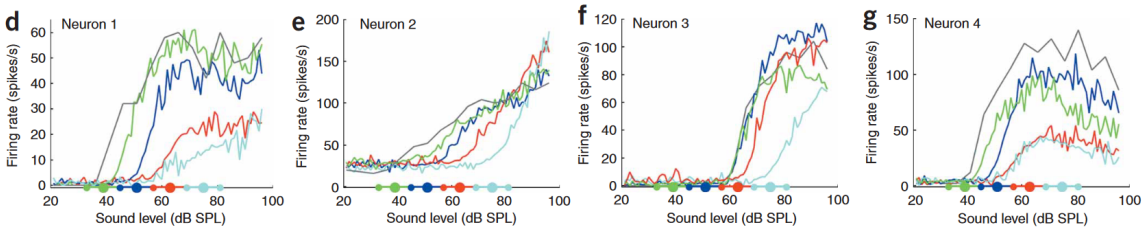


Figure I.9 – Rate-level functions for four individual neurons measured by Dean et al. (2005). The grey line is the rate-level function in quiet, and each colored line represents the results for a specific HPR, represented in the x-axis (mean levels ; green : 39 dB SPL, blue : 51 dB SPL, red : 63 dB SPL and cyan : 75 dB SPL). The shift of the response is placing the steepest part of the rate-level function closer to the mean level of the HPR (replotted from Figure 1 of Dean et al. (2005)).

their response, leading to a better representation of the edges of the HPR range. For bimodal distributions (with two separate 12-dB HPRs), each neuron seemed to adapt to one of the two HPRs and the overall population was approximately equally divided between the two HPRs. The response functions of the neuron population thus look collectively driven towards a better representation of the most probable levels.

They also evaluated the time constants of DRA by playing successive 5-s contexts and alternating between two HPR mean levels. The change in firing rate over each 5-s measurements was fitted with an exponential function. The time constant was measured to be around 160 ms on average for an increase in mean HPR level while it was twice as long for a decrease in level : 332 ms. Some neurons also exhibited a longer adaptation, with time constants in the order of tens of seconds (Dean et al. 2008). Other interesting factors were found in this study : neurons with higher characteristic frequencies tend to have slightly lower time constants, and neurons that adapt faster undergo greater dynamic range changes.

Wen et al. (2009) also found evidence for DRA in the auditory nerve of cats. In the auditory nerve, the high-spontaneous rate fibers seem more affected by DRA than the low-spontaneous rate fibers. They found a generally smaller magnitude of the shift (0.25 dB/dB) than in the inferior colliculus of guinea pigs (0.37 dB/dB), but it was

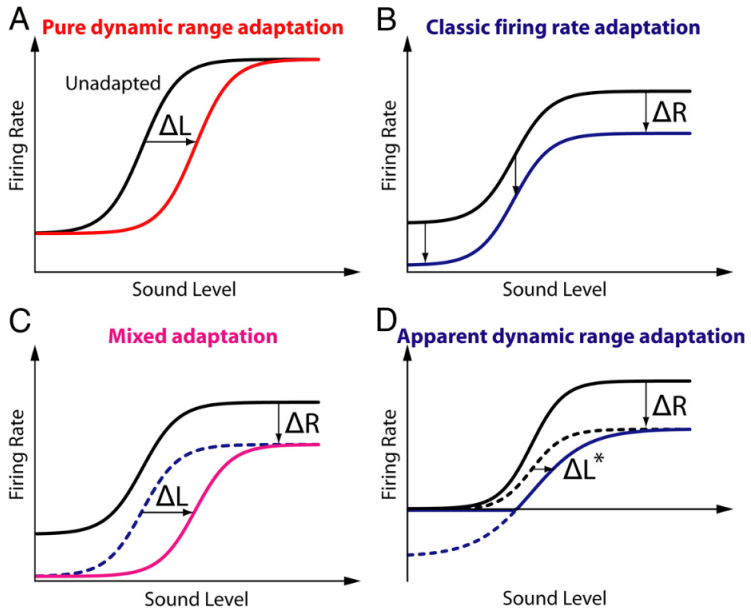


Figure I.10 – Different features of the dual-adaptation model of Wen et al. (2009). Panel A presents dynamic range adaptation, panel B classical adaptation, panel C a mix of the two and panel D what is observed in the measurement results.

definitely present, and while the adaptation time constant was frequency dependent in the inferior colliculus, no effect of the characteristic frequency was found in the auditory nerve. They later constructed a phenomenological model, described in figure I.10, to separate DRA from classical spikre-rate adaptation in response to continuous stimulation (Wen et al. 2012). They used the same signals as Dean et al. (2005), with the context mean level changing every 5 s. They found a mean of 211 ms for an increase in context mean level, and 262 ms for a decrease, i.e. similar to the time constants observed by Dean et al. (2008) in the inferior colliculus. As for classical adaptation, the time constants were 131 ms for an upward switch in mean level, and 272 ms for a downward switch.

Watkins and Barbour (2008) measured DRA in the auditory cortex of marmosets for monotonic and non-monotonic neurons. Non-monotonic neurons have a double dynamic range : their firing rate increases until a "best level" for which the firing rate is maximal, then decreases again. There are more and more non-monotonic neurons when going more central, and they probably convey specific information as the signal and the system get more complex. Compared to the lower HPR and to monotonic neurons, the higher HPR level was often not well represented by the non-monotonic neurons, because it fell in the second dynamic range for which firing rate decreases again. Contrary to their best coded level, the threshold of the non-monotonic neurons did not change with the mean level of the context, meaning they were still sensitive to low levels while the thresholds of the monotonic neurons shifted upwards. The two population are complementary to provide a good encoding of a great variety of levels even in a highly noisy background.

I.3.2 Possible origins

Grange et al. (2022) developed a numerical model based on the Matlab Auditory Periphery (MAP) model that simulates auditory nerve activity in response to sounds, together with a decoder to turn the auditory nerve activity back into sounds. More details about the model are given in the next chapter. They compared their model to the responses from Wen et al. (2009) and concluded that DRA could partly be due to peripheral efferent mechanisms, specifically a combination of the acoustic reflex and the MOC reflex. Indeed, both work to reduce the input gain of outer hair cells, modifying their behavior in response to sounds (especially the loud ones), and have time courses close to those found for DRA (Backus and Guinan 2006).

The MOC reflex was discarded by Wen et al. (2009) and Zilany and Carney (2010) as a primary cause for DRA for several reasons. First, the absence of frequency dependence of DRA in the auditory nerve, as the MOC reflex is frequency-specific. Second, their evolutions in terms of time course are opposite : the MOC reacts faster to a decrease in level than to an increase. Finally, in anesthetized animals the MOC is presumed to be absent, or at least importantly reduced. It could however participate in everyday hearing, especially because it also works on several time scales (Backus and Guinan 2006).

According to Zilany and Carney (2010), DRA is more likely the result of a process involving neurotransmitter release in the ribbon synapses between the inner hair cells and the auditory nerve fibers, as suggested by Wen et al. (2012). They used a peripheral model of the auditory system, and power-law dynamics for these synapses provided convincing adaptation and time courses for DRA.

Two-tone suppression was another hypothetical origin for DRA because of the wideband noise used (Costalupes et al. 1984 ; Rees and Palmer 1988) : adding a suppressor tone decreases the firing rate response and synchrony to the test tone as long as the suppressor is within the tuning curve of the neuron. The characteristic frequency of the neurons affects the growth of suppression, as does the relative frequency of the suppressor if the tone is at the characteristic frequency (Delgutte 1990). If the tone is elsewhere in the neuron's tuning curve, maximal suppression occurs for a suppressor at the characteristic frequency (Javel 1981 ; Javel et al. 1983). It was however also discarded by Dean et al. (2005) because of the absence of frequency dependence, and by Wen et al. (2009) because the frequency components all had the same relative power in the different HPR conditions, while two-tone suppression is observed with loud suppressors.

The other possible origins of - or mechanisms contributing to - DRA referenced by Wen et al. (2012) include some specific pre- or postsynaptic ion channels in the hair cells and/or auditory nerve fibers (namely, transduction channels and postsynaptic intrinsic membrane channels).

I.3.3 Other phenomena

Using the same broadband, fluctuating, unequally distributed stimuli as Dean et al. (2005), Robinson et al. (2016) played a different HPR every 5 s, and looked at the evolution of the time constant for each HPR change over time. They found a "meta-adaptation" operating over tens of seconds : in the inferior colliculus of guinea

pigs, the shifts adapted to changes in environments after less than two presentations of each context, almost to the point of dividing by two the time constant of adaptation to a change in HPR. In other words, the mechanisms behind DRA were adapting faster to HPR they had already encountered than to new ones.

Other studies looked at the influence of the level variance : Rocchi and Ramachandran (2018, 2020) found in the auditory cortex of awake macaques that the system also adapts to the overall variance of the presented levels, rather than to the mean. The slopes tend to increase for small ranges of test levels compared to wider ranges, for both the psychometric function of a tone detection task in noise and the rate-level function in response to the same sounds, but the slopes are unaffected by the levels themselves. In their second study, they found that attention also had an effect on DRA and adaptation to the variance : macaques involved in a psychophysical detection task showed larger shifts and slope changes than when listening passively.

The results from Rocchi and Ramachandran (2018) were very similar to the contrast gain control described by Rabinowitz et al. (2011) : the gain of the auditory cortex of anesthetized ferrets modifies the slopes of the rate-level functions to optimize the response to different contrasts environments, with level ranges differing in width. A high contrast (broad range) involved a more diverse range of levels and yielded shallower rate-level functions ; on the other hand, a low contrast (narrow range) leads to steeper rate-level functions, as less levels needed to be accurately represented.

They found this effect using harmonic complex stimuli, and observed less adaptation (both DRA and contrast gain control) to stimuli with high level contrasts than with low contrasts. They also investigated the effect of noise on the representation of natural sounds or speech at different sites of the auditory system of ferrets (Rabinowitz et al. 2013) : the lower areas were more sensitive to noise, and different adaptation mechanisms compensated the effects of the noise by i) suppressing the baseline changes through adaptation to the mean levels, and ii) optimizing the dynamic range through adaptation to contrast. Lohse et al. (2020) supported this by showing some adaptation to the level contrast in an intensity discrimination task in humans. They found lower thresholds for series with low contrasts (20-dB versus 40-dB wide level range), as well as in physiological measurements at several sites of the auditory system of mice. They were able to predict the human behavioral results from their measured mice physiological gains with a phenomenological model fitted to data from the inferior colliculus, thalamus and auditory cortex. This is consistent with the observation of the impact of the HPR width by Dean et al. (2005). DRA and contrast gain control would thus be complementary, one shifting the rate-level function (change in dynamic range mean level) and the other changing the overall slope of the rate-level function to match the range of probable levels, narrowing or broadening the range of encoded levels (change in dynamic range width).

Nagel and Doupe (2006) also found non-linear, level-dependent changes in the frequency tuning and gains of neurons in the forebrain of songbirds. These changes are consistent with combined forms of adaptation - including dynamic range adaptation. At low mean levels, the neurons acted as low-pass filters and time integrators ; while at higher mean levels, they behaved more like band-pass filters, with reductions in latency and decreases in gain with increases in level variance. This change in gain was compared to DRA, and it might also be related to auditory gain contrast.

I.3.4 Human studies

Many studies on adaptation to stimulus statistics mentioning DRA are questioning the involvement of the medial olivocochlear (MOC) reflex. In three back-to-back studies, Marrufo-Pérez et al. (2018, 2019, 2020) and later López-Ramos et al. (2024) conducted tasks of tone detection, amplitude/frequency modulation detection, and speech recognition in noise in normal hearing participants and CI users. They monitored the activation of the MOC reflex using otoacoustic-emissions and used vocoded signals to compare the results of normal hearing participants to those of cochlear implant users, for whom the MOC and the acoustic reflex are absent. They found that both normal-hearing and implanted listeners performed better in the different tasks when the noise was switched on 300 ms before each test trial compared when it was switched on at the same time. They concluded that the MOC reflex was apparently not necessary for an adaptation to noise, and that DRA, along with other mechanisms and interactions between stimuli, could well be a good explanation for their results.

Wojtczak et al. (2019) conducted amplitude modulation detection tasks for 40-ms pure tones (at 1 or 6 kHz) presented in 100-ms noise with different 400-ms precursors, along with otoacoustic emission measurements to monitor the activation of the MOC. They compared both measurements in search for a correlation between MOC and improved performance in behavioral tasks in noise. The noise and complex tone precursors decreased modulation detection thresholds, while the MOC did not seem linked with this reduction in thresholds. Moreover, they concluded that DRA could play a role, because the thresholds were lower when the precursor was identical to the masker (it had the same level within the auditory filter tuned to the test tone frequency).

In a follow-up study, Mesik and Wojtczak (2020) wanted to compare the mechanisms for amplitude modulation detection and frequency modulation detection of pure tones carriers in noise, arguing that the latter could just be a byproduct of phase locking and amplitude modulation. The idea is that frequency modulations at low frequencies rely on phase-locking cues, while at higher frequencies, frequency modulations are inducing amplitude modulations at the frequencies of the stimulus (FM-to-AM conversion). Their first experiment compared the effect of a noise precursor on amplitude and frequency modulation detection, and their second experiment focused on amplitude modulation with a roving-level precursor. They found similar effects of noise precursors on both tasks in noise in the first experiment, and an adaptation to the precursor level in the second : results were better for a precursor of the same level as the noise masker. They concluded that the perceptual similarity between the precursor and the masker was important, and that among the mechanisms at play, DRA could have had an effect on the second experiment - amplitude modulation with roving precursor level-, as the precursor with the same level as the masker yielded lower thresholds than other levels.

These studies do not mean that the efferent system cannot help tuning the peripheral response to the noise statistics, only that it is probably not the only process involved in the observed changes of the DRA studies cited in the previous subsection. The MOC is surely helping in everyday, normal hearing, but there are other causes to DRA, so it would be interesting to get physiological data from the human periphery.

Herrmann et al. (2020) investigated the presence of DRA in humans and its effect on perception. They compared the results of electroencephalographic (EEG) recording to those of a modulation detection experiment using the same stimuli : 100-ms bursts of white noise were played every 500 ms, using two different level ranges for the high probability region. In the psychophysical task the participants had to detect an amplitude modulation inserted in certain bursts, while the EEG recordings recorded activity for the different level configurations. For the EEG, the stimuli were not modulated. The amplitudes of the cortical responses evolved with the mean level of the stimulation : the first part of the neural response was larger for the less probable levels, and the activity in the second part increased along the mean context level. In the psychophysical task, the less probable levels had more chance to get detected when modulated. The EEG measurements showed results in agreement with both DRA and an error prediction mechanism, according to which the less probable levels are allocated more resources ; however, only the error prediction mechanism appeared in the amplitude modulation detection experiment.

In the same spirit, Simpson et al. (2014) designed an intensity discrimination task with using the method of constant stimuli. They had three standard levels and played one of them 80 % of the time. The difference in level between the two intervals was fixed at the discrimination threshold of each participant for each standard level, which was measured in a preliminary adaptive three down-one up procedure - the percentage correct was thus theoretically around 79.4 % for all standard levels. They monitored the correct response percentage (performance) over time depending on how likely the different standards were to appear considering the recent stimulation history. Performance was at first good for the most frequently presented level. But after the first minute performance improved for the less probable stimuli, and even more for the level most different to the most frequent, while performance decreased for the most frequent one. They also looked at the impact of changes in interstimulus intervals, and observed the same evolution of performance with the most frequent better at first then worse, and the less frequent improving over time.

Critically, the timescales used in these studies make them difficult to compare to the findings of Dean et al. (2005) and others, because of the long silent gaps (at least 350 ms) and short stimuli (50- or 100-ms probes). Time constants of DRA are shorter than the silent gaps in these studies : they were thus probably looking at different mechanisms. These studies indicate that in addition to a shorter-term adaptation (hundreds of milliseconds) to the most common stimulus characteristics, there would be another mechanism sensitive to long-term statistics (10s of seconds), as demonstrated by Robinson et al. (2016). This second time scale would in turn also enhance the response to less probable stimuli (error prediction). In a psychophysical task with measurements over several minutes, the protocols need to be built to focus on the short-term effects, as we are not really interested in the long-term ones. Long physiological measurements might catch both phenomena.

I.3.5 Aims of the thesis

Dynamic range adaptation might be very important in the perception of speech in noise, because of the need to hear small amplitude modulations to be able to

understand consonants and thus, words. However, it is not yet clear what are the direct consequences of DRA on sound intensity perception, and where it originates. And because it is not possible to simply replicate the animal studies with electrodes in several parts of the auditory system in humans, the challenge is to select the experimental setups that are more likely to deliver some evidence.

In terms of perceptual consequences of DRA, what we deemed to be the best lead was an intensity discrimination task. Interestingly, it was also one of the least explored tasks in the literature, compared to amplitude modulation detection for example (even if both tasks rely on the same cues for low modulation rates, see Gallun and Hafter 2006). The reasoning behind the choice of this task was the following : we hypothesize that a change in the intensity of a stimulus should be easier to detect if it produces a large change in neuron firing rate. Thus, if the stimulus falls into the steep part of the rate-level function of the neurons, a small intensity difference with the signal will lead to an important change in activity, and the intensity discrimination threshold will be low. But if the standard falls away from that steep part, the same change in activity will require a much larger intensity difference, so the discrimination threshold should be higher. By inscribing the discrimination task in a context similar to the animal studies, with a mean context level identical to the standard level, and another mean either higher or lower in level, we expect the best discrimination results to appear in the condition in which the mean level of the context is the same as the standard level.

Chapters II and III present the results of several intensity discrimination experiments that used narrow-band and broadband signals, respectively.

The second experimental path we wanted to explore was the measure of human physiological data in lower centers than the auditory cortex. The origin of DRA is uncertain, but its observation in the auditory nerve of cat (Wen et al. 2009) indicates than it could originate at least partially in the auditory nerve itself, or lower.

This study was focused on testing whether the auditory nerve can be a direct contributor to DRA. In humans, a way of measuring auditory nerve activity is with cochlear implants. A research software for cochlear implants allowed the recording of electrically evoked compound action potentials from the auditory nerve with different measurement paradigms, without any influence of the outer, middle and inner ear (as the medial olivocochlear reflex, the acoustic reflex and the IHC-auditory nerve synapse are bypassed with the cochlear implant).

Chapter IV addresses the experiments we were able to carry with cochlear implant users.

Chapter II

Intensity discrimination using pure tones

II.1 Introduction

The experiments presented in this chapter provide an attempt to find a perceptual correlate of DRA in humans. The aim is to measure intensity discrimination at different levels and in several contexts. We hypothesize that a shift in dynamic range should induce a change in intensity discrimination performance : specifically, for equal context and signal levels, the discrimination thresholds should be lower than when the context mean level and the standard level are different.

To test this hypothesis, we designed a protocol for measuring intensity discrimination of pure tones embedded in different contexts. The test signals consisted of a fixed-level standard and a target signal. We used pure tones as test signals, and the context was a concatenation of 50-ms segments of narrowband noise with different levels (see figure II.1). The context was continuous and only stopped during the presentation of the test signals. The main experiment explored the differences in intensity discrimination thresholds for different standard and context levels. Because the results were not entirely consistent with our hypothesis, several additional experiments were performed to achieve a better understanding of the effects of the context on intensity discrimination. Experiment 2 was a replication of experiment 1 performed at lower levels, experiment 3 tested the influence of off-frequency listening, and experiments 4 and 5 were focused on reducing the masking effects induced by the context. Moreover, auditory nerve activity in response to these stimuli in different experiments was simulated in an attempt to link the behavioral tasks results to physiological data.

II.2 Experiment 1 : Intensity discrimination in a narrow-band context

II.2.1 Material and methods

II.2.1.1 Participants

Ten normal-hearing (thresholds $\leqslant$ 20 dB HL between 125 and 8000 Hz) participants took part in the main experiment. They were part of a larger group of 43 participants, who went through a selection process - a preliminary experiment, described later. This preliminary experiment aimed at making sure that the test signals were clearly audible in the loud context. Three of the participants were experienced in psychophysical

tasks ; five of them were males. Their ages ranged from 22 to 57 years old (mean : 29.5, s.d. : 9.6).

The ten selected participants received at least one training session (i.e. a shorter version of the test conditions) before each experiment.

The research in this chapter was approved by ethics committee "Comité de protection des personnes Sud-Ouest et Outre Mer III" (2021-A01920-41)". Participants were paid according to the number of hours of testing.

II.2.1.2 Hardware and software

All signals were presented monaurally to the left ear using a Sennheiser HDA 200 headphone earpiece and a RME Fireface UCX sound card while participants were seated in a double-walled sound-proof booth. The test was entirely automated and programmed using Matlab (2022b, The MathWorks). Statistical analyses were run with the R software (4.4.2). Two red LED lights placed under the screen were synchronized with the two presentation of the two intervals, and feedback was provided onscreen after each response.

The system was calibrated using a B&K 4153 artificial ear, a B&K 4192 1/2" microphone, a GRAS 26AK preamplifier, and a GRAS 12AR conditioner. The levels were controlled by monitoring the gain applied to the numerical signal.

II.2.1.3 Stimuli

In physiological studies of DRA (e.g. Dean et al. 2005), the neural response to each segment of the stimulus context is recorded. This means there is no need for a "test signal". In contrast, for an intensity discrimination task, the participants need to know when to listen for the test signals to be able to make judgments. Previous research showed that the time constant of DRA at the level of the auditory nerve or of the inferior colliculus can be short (about 100-200 ms). This implies that introducing a gap between the context and the test signals to provide a cue to the participants would potentially reduce this effect. Therefore, we instead chose to have no gap between context and test signals but to use different sounds for context and test signals (standard and target), in addition to a visual cue indicating the intervals.

In this experiment, the context consisted of a series of 50-ms noise segments only interrupted for the presentation of the test signals (as shown in figure II.1). The noise segments all had the same spectral characteristics : they were synthesized by summing 1-Hz spaced sinusoids with random starting phases and frequencies between 1880 and 2120 Hz. The bandwidth was thus 240-Hz wide, equal to one equivalent rectangular bandwidth (ERB) centered around 2 kHz [1]. Each segment was shaped by 5-ms raised-cosine ramps at the onset and offset. The choice of a 1-ERB band was made to control the levels in a single auditory filter centered on 2 kHz. The context level was updated every 50 ms (i.e., every segment), and followed a Gaussian distribution centered around a specific mean level, with a standard deviation of 3 dB. The levels

1. The frequency limits of the bandwidth of the narrowband noise were defined considering 2000 Hz as the arithmetic mean of the two frequency boundaries. A geometrical mean with these limits gives 1996 Hz ; we consider the 4-Hz difference to be negligible.

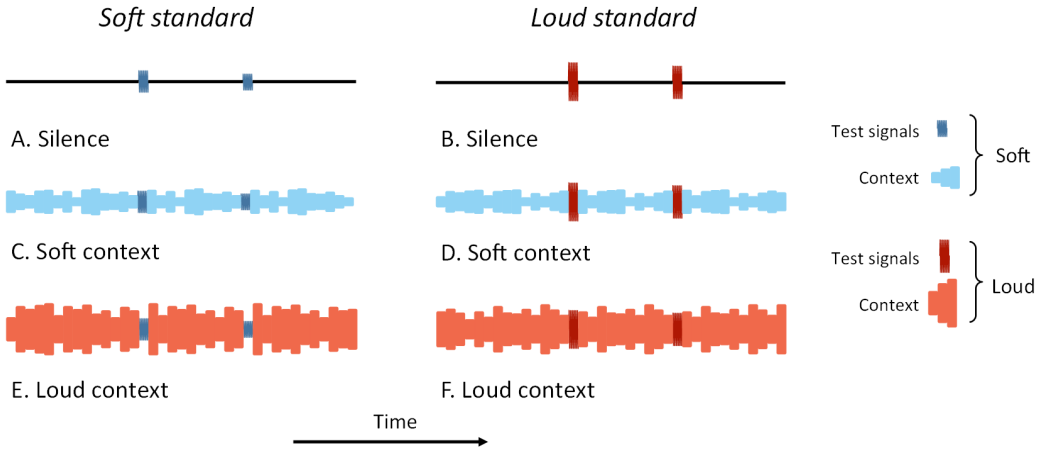


Figure II.1 – Example trials from the main experiment for each condition, sorted by standard level and context condition. The soft test signals (standard and target) are represented in dark blue while the soft context is in light blue, the loud context is in orange and the loud test signals are in dark red.

were however bounded to ± 10 dB around the mean : a random level was drawn for each segment and ditched if out of these limits. This was done to prevent the level from being much too loud or much too soft.

The test signals were pure tones, presented during the interruptions of the context, similarly to the "negatively gated noise" condition of Costalupes et al. (1984) which was gated off before the standard and back in after its presentation. Our test signals had a frequency of 2 kHz and a duration of 50 ms, with 5-ms onset and offset raised-cosine ramps.

The standard had a fixed level, and the target signal was constructed by adding two pure tones of the same frequency : one at the level of the standard, and the second ('added signal') at a chosen level. This ensured that there was always a strictly positive difference in level between the target signal and the standard. The level difference in dB was thus calculated as follow :

$$DL = 10 \times \log_{10}(1 + \frac{\Delta I}{I}) \tag{II.1}$$

where I is the standard intensity and ΔI is the difference in intensity between the standard + added signal and the standard, which corresponds to the intensity of the added signal.

Intensity discrimination thresholds were thus measured in six experimental conditions (2 standard levels x 3 context conditions, see figure II.1). The soft standard level was 55 dB SPL and the loud level was 75 dB SPL. The three context conditions were silence, and mean levels of 55 dB SPL (soft) and 75 dB SPL (loud) respectively.

II.2.1.4 Procedure

The procedure was a 2I-2AFC adaptive 2 down - 1 up interleaved method (Levitt 1971), with the standard presented in one (random) interval and the target signal in the other. Participants had to select the loudest sound of the two intervals.

The gap between the test signals was 500-ms long and was filled by the context. The exact same 500 ms of context were presented before, in between, and after the two test intervals containing the standard and target to limit the influence of short-term differences in context level on the judgments ; they however differed between trials. A trial was considered over after the participant answered, and the context was played throughout. The inter-trial interval, also filled with the context, was randomly varied across the experiment, ranging from 700 to 1500 ms (uniform distribution) to avoid anticipation from the participant.

The level of the added signal changed in 4-dB steps for the first two reversals, then in 2-dB steps for the rest of the procedure. The added signal at the start of the procedure had levels of 69 and 87 dB SPL for standard level of 55 and 75 dB SPL, respectively.

The intensity discrimination thresholds for the two standard levels were obtained simultaneously in each context condition using an interleaved procedure. For each trial, the standard level was randomly chosen among the two possibilities (soft or loud). Each discrimination threshold was calculated as the mean of the last six reversals points of each adaptive track ; the two tracks of the interleaved procedure continued until both tracks reached 12 reversals. There could thus be more than 12 reversals for one of the tracks, since both had an equal probability of being presented until the end of the procedure. The obtained thresholds were averaged over four repetitions for each measurement ; if the standard deviation of the four repetitions exceeded 3 dB, one or more repetitions were added until the standard deviation was below 3 dB (which happened only twice in total).

II.2.1.5 Preliminary experiment

The exact values of the soft and loud levels used in the experiment (55 and 75 dB) were initially chosen arbitrarily. We wanted this difference to be large to maximize our chances of seeing an effect of the context on intensity DLs. After a few pilot experiments, however, it appeared that the 55-dB SPL (soft) standard level could sometimes be totally masked by the 75-dB SPL (loud) context, making it impossible to perform the intensity discrimination task. The same problem was encountered by Moore et al. (1985) with noise detection in notched noise for narrow notches. As a result of the great variability in observed detection thresholds, we performed a subject selection task as a preliminary experiment.

The task was a signal detection task in the 75-dB SPL context. The context was the same narrowband noise (bandwidth 1880-2120 Hz) as in experiment 1, presented continuously except during the test intervals. Detection thresholds in the soft context (55 dB SPL mean level) and absolute thresholds (in silence) were also tested for every participant, for subsequent experiments and to compare our results with the literature.

The procedure was the same as in the main experiment : 2I-2AFC, 2 down - 1 up adaptive staircase, but with the standard in one of the intervals and silence in the other, separated by 500 ms. There were five repetitions of the three test conditions, all done during the same test session.

The selection criterion was that the detection threshold in the 75-dB SPL context had to be below 45 dB SPL, so that in the main intensity discrimination task, the

55-dB SPL standard would be presented *at least* at 10 dB sensation level (SL) in the loud context.

A total of 43 participants took part. The results showed a large variability between participants in detection thresholds : in the loud context the thresholds ranged between 35.6 and 72.5 dB SPL (between 27.0 and 55.1 dB SPL in the soft context, and from -2.52 to 24.9 dB SPL in silence). The correlations between detection threshold in the presence of the context and detection threshold in silence were small but significant (correlation coefficient $\rho = 0.31$, $p = 0.039$ for the soft context and $\rho = 0.46$, $p = 0.002$ for the loud context). This suggests that the context provides a source of masking that is not related to audibility.

Detection thresholds for the ten selected participants can be found in appendix A.

II.2.1.6 Auditory nerve model

Auditory nerve activity in response to the stimuli used in this experiment was simulated with the Matlab Auditory Periphery (MAP) model as configured by Grange et al. (2022). The parameters of the model were identical to those used in their study.

The MAP model (Meddis 2006) simulates the response of several parts of the peripheral system to auditory stimuli. The outer and middle ear are simulated first as filters ; cochlear filtering is then modeled with the dual-resonance non-linear filter model of Lopez-Poveda and Meddis (2001), decomposed in several frequency bands (equivalent rectangular bands, ERBs). The response of the latter goes into the brainstem that controls efferent mechanisms like the MOC (which feeds back to the basilar membrane under the form of an attenuation coefficient for each ERB), and the acoustic reflex (which attenuates the middle ear response). 29 970 auditory nerve fibers (equally arranged over three spontaneous rates and 30 ERBs) are simulated.

The output of the model consists of i) the number of firings per fiber type and ERB sampled at 24 kHz ; ii) the temporal evolution of the attenuation applied on the basilar membrane movement by the MOC for each ERB, sampled at 100 Hz ; and iii) the attenuation of the middle ear response by the acoustic reflex over the duration of the stimulus, sampled at 24 000 Hz.

To obtain the results replotted in figure II.2, Grange et al. (2022) used 8-seconds long trains of 50-ms pure tones (or broadband noises) of different levels (from 0 to 80 dB SPL by 4-dB steps, 20 to 100 dB SPL for noise), randomly presented, with an 80 % probability of getting one of the four levels around the mean level of their context (36, 48, 60 or 72 dB SPL for pure tones ; 48, 60, 72 or 84 dB SPL for broadband signals).

Our simulations used stimuli identical to those used in the psychophysical task. The rate-level functions of the three types of auditory nerve fibers (HSR, MSR and LSR fibers) were plotted by computing the average firing rate of the 333 fibers for the ERB centered at 2 kHz (ERB 21), in response to the standard in each context condition used in the task. The stimuli consisted of 1.35-s long wav files sampled at 44 100 Hz. The file started with 50 ms of silence ; 700 ms of the desired context condition were then presented (silence, soft or loud context), followed by a 50-ms standard, then 500 ms of the same context, and finally 50 ms of silence. The standard level ranged from 1 to 89 dB SPL in 2-dB steps, and the context was identical for all standard

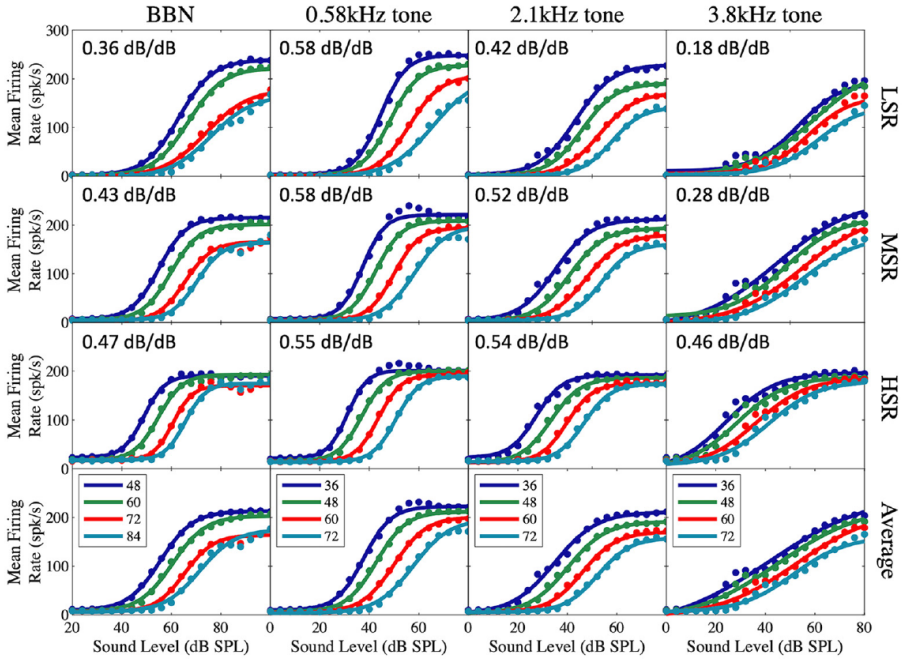
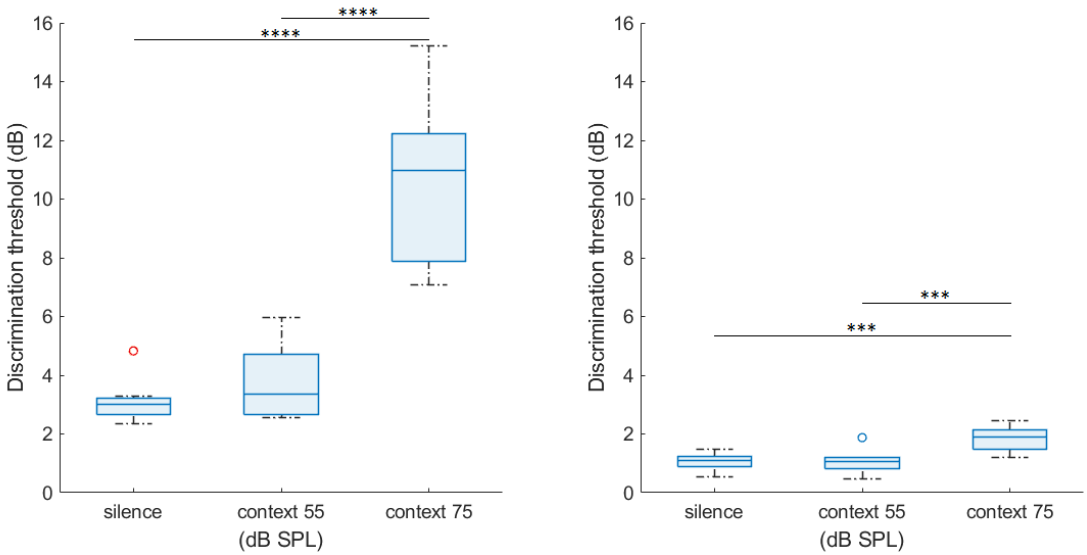


Figure II.2 – Rate-level function of different neuron types to different stimuli, taken from the MAP auditory nerve model run by Grange et al. (2022) and replotted from their figure 6. Tracks of 8 s of fluctuating-level stimuli were averaged over ten randomization of 20 repetitions of each tested level (200 in total), with a 12-dB wide HPR of different mean levels (indicated in dB SPL in the legend). BBN : broadband noise.

levels. Each condition (45 standard levels x 3 contexts) was simulated ten times with a different context (10 repetitions).

Sigmoid functions were fitted to the modeled rate-level functions, averaged over the ten repetitions, using a custom least-square method. The slopes of the fitted functions were obtained with linear regressions of three consecutive points. The horizontal shift between two simulated rate-level functions, obtained for two different HPR levels, was calculated as the shift of the midpoint level L_{50}, which is the level at which the point halfway between the maximum and minimum of the fitted sigmoid is found. That way, we could compute the slopes of the shift in rate-level functions, and compare them quantitatively to the results in the third column of figure II.2, and qualitatively to our hypotheses.

It is important to consider that the slopes of the rate-level functions are not representative of the whole picture : because the variability in fiber discharge tends to increase with firing rate, the sensitivity index is often used to evaluate whether a change in slope is translated into an improvement of the encoding, as is the case in Wen et al. (2009) and Grange et al. (2022). We thus used the same definition (Colburn et al. 2003) : we quantified the sensitivity index from the simulations. The methods and results obtained can be found in appendix B.2 for this experiment, and B.3 and B.4 for the control experiments.



(a) 55 dB SPL standard **(b)** 75 dB SPL standard

Figure II.3 – Results of experiment 1 for both standard levels. Discrimination thresholds for all participants are plotted against context condition : the first and third quartile are limiting the blue box, and the median is presented as a blue bar within it. Whiskers include all non-outlier values that do not fall into the box ; blue circles are outliers (between 1.5 and 3 times the interquartile value away from the first or third quartiles) and red circles extreme values (more than 3 times the interquartile value away from the first or third quartiles). Pairwise t-tests on the discrimination thresholds revealed significant differences (represented as stars on the graph) between the loud context and the other conditions for both standards.

Standard level	Context level	Condition	Mean (dB)	Standard error (dB)
55 dB SPL	silence	A	3.08	0.20
	55 dB SPL	C	3.68	0.33
	75 dB SPL	E	10.78	0.81
75 dB SPL	silence	B	1.06	0.08
	55 dB SPL	D	1.03	0.11
	75 dB SPL	F	1.85	0.12

Table II.1 – Averaged results for experiment 1, with the corresponding condition letter in figure II.1. The average discrimination thresholds tend to increase in the 75-dB SPL context and they do not change much between the silence and 55-dB SPL context.

II.2.2 Results

Intensity discrimination thresholds (DLs - difference limens) are presented as a function of context condition in figure II.3, and the mean and standard error of the group for each condition are presented in table II.1. The data were analyzed using a two-way repeated-measures ANOVA with treatment factors standard level (2 levels : 55- and 75 dB SPL), and context mean level (3 levels : silence, 55 dB SPL and 75 dB SPL). The Greenhouse-Geisser correction was used to correct p-values when the sphericity assumption was violated (as for all ANOVAs in this chapter). The analysis

revealed significant effects of standard level ($F_{1,9}$ = 133.3, p = 1.07e-6) and context mean level ($F_{2,18}$ = 100.4, p = 3.50e-7) on intensity discrimination performance, as well as a significant interaction between these factors ($F_{2,16}$ = 48.9, p = 9.63e-6). This means that the context condition affects intensity discrimination performance for the two standards differently.

Pairwise comparisons using the Bonferroni correction for multiple comparisons (as in the rest of the chapter) showed significant differences between context conditions for both standards. The results are presented in table II.2 : the thresholds in silence and in the soft context are very similar, whereas there is an increase in discrimination threshold in the loud context.

Standard level	Context 1	Context 2	t	p adjust	significance
55 dB SPL	silence	55 dB SPL	-1.73	0.354	ns
	silence	75 dB SPL	-8.66	3.51e-5	****
	55 dB SPL	75 dB SPL	-9.02	2.52e-5	****
75 dB SPL	silence	55 dB SPL	0.326	1	ns
	silence	75 dB SPL	-6.59	3.00e-4	***
	55 dB SPL	75 dB SPL	-7.15	1.62e-4	***

Table II.2 – Pairwise t-tests for each standard and context mean level. Significant differences appear between the loud context and the other conditions for both standard levels : the loud context gives higher thresholds than the soft context and the silence (ns : non significant).

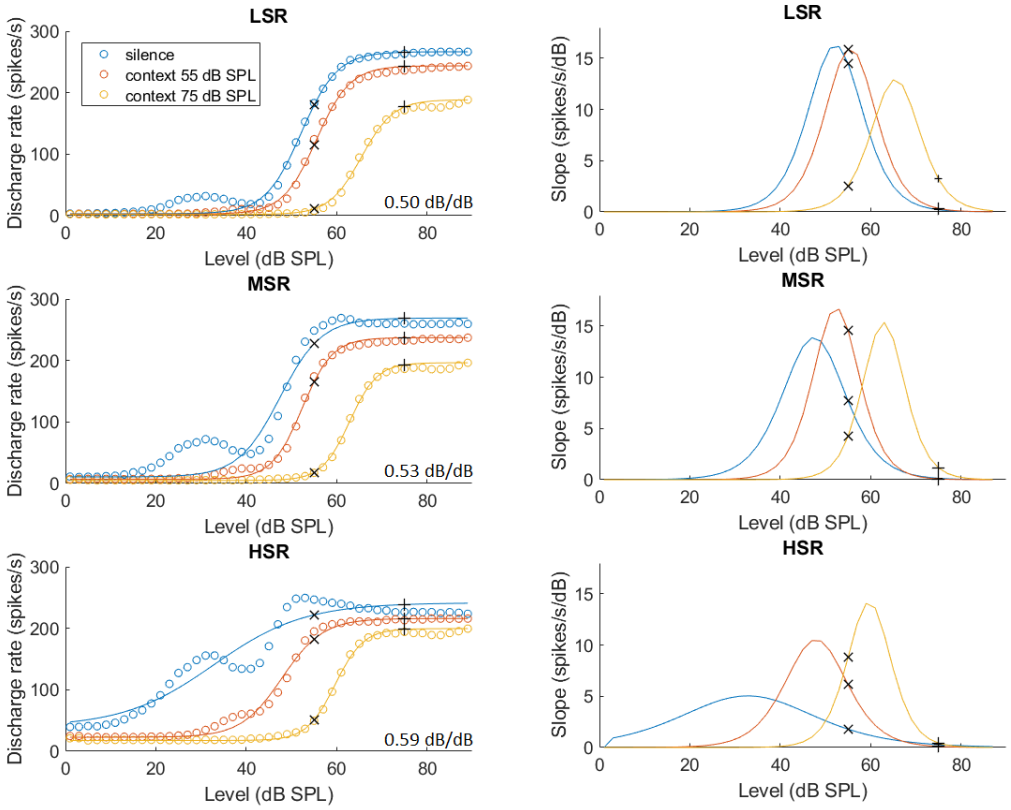
Figure II.3a shows the discrimination thresholds for a 55-dB SPL standard in different contexts. Consistently with our hypothesis, thresholds were higher when context and standard levels were different, with the loud context yielding the highest thresholds.

The discrimination thresholds for the 75-dB SPL standard are plotted in figure II.3b. In this case, discrimination thresholds were not lower for the matching-level condition, they were even slightly higher : this is not consistent with our hypothesis. Overall, the thresholds remained low and appeared to be less affected by context mean level than for the soft standard.

Figure II.4a shows the rate-level functions predicted by the model. They clearly indicate that there is a shift of the rate-level function in the presence of the context, as well as an overall reduction of the maximum firing rate when the level of the context increases. These rate-level functions are similar to those presented by Grange et al. (2022) in their figure 6 and replotted in figure II.2.

The responses to the pure tones in silence display one or two humps, around 30 and 50 dB SPL. These two disruptions in the sigmoid-like shape are also present at the lowest context mean levels, as well as in the results of Grange et al. (2022) ; they are less visible in their results, maybe due to the greater step size they have between levels (4 dB against 2 dB here). However there is no discussion of these humps in their paper, and they do not seem to be caused by the MOC or the acoustic reflex as they are still present in their simulations without efferents.

Figure II.4b shows the slopes around each standard level, that can be extracted from the rate-level functions (three-point linear interpolation). The values for each standard level are presented in table II.3. At 55 dB SPL (level of the soft standard),



(a) Rate-level functions. **(b)** Slopes.

Figure II.4 – Rate-level functions with shift between contexts (a) and slopes (b) obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 1, averaged over ten simulations. The 55- and 75-dB SPL standard are represented respectively as black x and +.

the slope is steeper in the soft context than in silence and than in the loud context for the LSR and the MSR fibers, but not for the HSR fibers. For the 75-dB SPL standard, the rate-level functions are mostly saturated (resulting in slopes below 1), except in the loud context for the LSR fibers.

The shift in rate-level function between the two context mean levels is computed as the horizontal shift of the level at which the function reaches 50 % of its firing rate range (difference between maximum and minimum firing rate L_{50}). It gives an idea of the strength of DRA, and will be discussed later. The values are indicated in figure II.4a ; the shift is largest for the HSR fibers (0.59 dB/dB), and slightly smaller for the MSR and LSR fibers (respectively 0.53 and 0.50 dB/dB). This is consistent with what was observed by Grange et al. (2022).

II.2.3 Discussion

The data showed a significant influence of the context on the intensity discrimination thresholds : they were higher in the loud context than in silence and than in the soft context for both standard levels. There was, however, no difference between the soft

Standard level	Fiber type	Context type	Slope (spikes/s/dB)
55 dB SPL	LSR	silence	14.57
		55 dB SPL	15.59
		75 dB SPL	2.93
	MSR	silence	7.82
		55 dB SPL	13.74
		75 dB SPL	4.99
	HSR	silence	1.80
		55 dB SPL	5.10
		75 dB SPL	9.92
75 dB SPL	LSR	silence	0.223
		55 dB SPL	0.394
		75 dB SPL	2.66
	MSR	silence	0.139
		55 dB SPL	0.089
		75 dB SPL	0.974
	HSR	silence	0.295
		55 dB SPL	0.200
		75 dB SPL	0.384

Table II.3 – Slopes around both standard levels for each context and fiber types in figure II.4a. They are much lower for the 75-dB SPL standard as most fibers are already saturated.

context and silence. The results therefore only agree with our initial hypothesis for the soft standard. For the loud standard, the discrimination thresholds were larger in the loud context.

The near-miss to Weber's law is clearly visible in the results of the discrimination task : the discrimination thresholds are lower for the loudest standard in all context conditions. Moreover, the difference in DL between standard levels increases with context level (2, 2.65, and 9 dB in average, respectively for silence, 55- and 75-dB SPL contexts). This is consistent with the significant interaction between the standard and context mean level on thresholds in the ANOVA.

The simulations clearly display a shift of the rate-level function and maximum slope with the context mean level. The slopes are consistent with the hypothesis, with steeper rate-level functions in the context matching the standard level in most cases.

The following paragraphs discuss different aspects of these results and describe the hypotheses on which the subsequent control experiments are based.

II.2.3.1 Comparison with simulations

If we compare the results of our simulations (figure II.4a) to those obtained by Grange et al. (2022) with 2.1 kHz pure tones and a slightly different paradigm, replotted in figure II.2, the rate-level functions appear to be very similar.

The rate-level functions for HSR and MSR are saturated for the high-level standard, so the following slope comparisons are mostly based on LSR responses and should be considered with some caution.

The changes in slope observed at the standard level when changing the context are in agreement with our hypothesis but not with the results of the intensity discrimination task. The slope of the rate-level functions is steeper for a context of the same mean level as the standard in most cases. The only deviation is observed for the 55-dB SPL standard and the HSR fibers : in this case the slope is steeper for the 75-dB SPL context. Therefore, the results obtained in our intensity discrimination task do not reflect these changes in slopes.

The fact that the predicted horizontal shift values of the midpoint L_{50} of the rate-level functions are close to those from Grange et al. (2022) is reassuring. They might however be under- or over-estimated because of the small number of functions considered (only two context mean levels) compared to the four functions for Grange et al. (2022) : the silence was also not taken into account in the measurement of these horizontal shifts. The estimation of shift is our study and in Grange et al. (2022) is higher than what was found by Wen et al. (2009) for anesthetized animals (average : 0.25 dB/dB). The value of the shift in the AN measured by Wen et al. (2009) is closer to those obtained by Grange et al. (2022) in their simulations without efferents. This would be consistent with the idea that the MOC and acoustic reflex contribute to the effect of DRA : the MOC was supposedly absent in the anesthetized animals of Wen et al. (2009), and its presence could enhance the shift.

One of the differences between the physiological studies (Wen et al. 2009) and our first experiment is the use of a much narrower bandwidth : in the present study, the standard was a pure tone and the context a narrowband noise, while Dean et al. (2005) and following studies used broadband noise. The level used in one particular frequency band was higher in experiment 1 than the level of a broadband noise of the same overall level would be.

II.2.3.2 Comparison with the literature

The DLs presented in silence in table II.4 show the discrimination thresholds measured in silence in experiment 1 compared to the thresholds measured by Florentine (1986) and Carlyon and Moore (1984) in comparable conditions. It can be observed that our thresholds are higher than those reported in both of these studies. Florentine (1986) used a shorter rise/fall time than us, which may be an explanation for their lower thresholds (due to spectral splatter). However, it is unlikely to be the sole explanation as Carlyon and Moore (1984) used a rise/fall time identical to ours (5 ms), and the duration of their signal was 26 ms in most conditions ; only at 55 dB SPL, a duration of 50 ms was tested. They found DLs of 0.97 dB at 500 Hz and 1.45 dB at 4 kHz, whereas we found a threshold of 3.08 dB at 2 kHz, i.e. higher than a factor 2. It is unlikely that the difference across studies is due to the use of different frequencies, as it has been shown that DL does not change much with frequency for frequencies under 8 kHz (Florentine et al. 1987). It has to be noted that the signals used in Florentine et al. (1987) were longer than ours (500 vs. 50 ms), but they also tested shorter durations and did not find any interaction between stimulus duration and frequency on DL (Florentine 1986).

A more likely explanation for this difference comes from differences in training. We had 10 participants, while Carlyon and Moore (1984) had only 3, including the

1st author, and all their participants were trained for at least 15 hours, which was not our case ; Florentine (1986) also had three participants with training "until no systematic change in performance was apparent". Florentine et al. (1987) have shown that training and motivation can lower discrimination thresholds by a factor of 1.5 to 2, which is in the same order of magnitude as the difference in DLs between the present study and that of Carlyon and Moore (1984).

The protocol itself could also have an impact on the thresholds : in experiment 1, the two standard levels were interleaved, while the other studies tested the levels one by one. Studies with roving-level versus fixed-level intensity discrimination (Berliner and Durlach 1973 ; Berliner et al. 1977) found higher DLs for larger level-roving ranges for 1-kHz pure tones having durations ranging from 200 ms to 1.25 s. Mori and Ward (1992) compared results for fixed and roving level intensity discrimination of 50-ms, 1-kHz pure tones, with or without maskers : the transmission of information (calculated from the percent of correct responses, the mean level and standard deviation of the stimuli, and confusion matrix) was always better for fixed-level discrimination than for roving level, and was affected by masking. The lowest values of information transmission were obtained for roving-level with a 100-ms, broadband noise forward masker with respectively 50-ms, 100-ms, or no delay between masker and standard. Backward and simultaneous maskers also degraded the transmission of information, compared to the condition without masker. The fixed-level conditions with masker were however similar to the conditions with roving level and masker, especially with no delay between masker and standard. We could thus consider that with a continuous context like ours, that could be compared to forward and backward maskers without delay, the interleaved protocol does not have any effect on the discrimination thresholds. The discrimination thresholds in silence in the present study are thus probably overestimated compared to those in context.

To add on that, Pienkowski and Hagerman (2009) studied the effect of the width of the level-roving range on performance in an intensity discrimination task with 4-kHz pure tones with durations of 4 or 300 ms. The standard level was either fixed or roved over a 15 dB range centered around a mean ; four levels were tested. An interleaved version was also tested, in which all four levels were tested randomly, resulting in a total roving range of 60 dB. They concluded that performance was worse with wide roving of the tested levels, with a larger effect for longer stimuli and medium levels.

Similarly, Oxenham and Buus (2000) looked at the combined effects of stimulus level, duration (5 to 500 ms), and roving over 20 dB for 1- or 2-kHz tones. They also found an increase in the thresholds when roving the standard level, with increases in the difference between fixed and roving levels with stimulus duration.

The last lines of table II.4 show the DLs measured in context in experiment 1 of the present study. In the 55-dB SPL context, DLs were not significantly different from those in silence, showing that the context had less effect on DLs than the 75-dB SPL context. For the latter however, thresholds were significantly higher than in silence and than in the 55-dB SPL context. This increase was, in particular, very strong for the standard at 55 dB SPL.

	Present study		Florentine (1986)		Carlyon and Moore (1984)				
Context	Standard	2 kHz 50 ms rf = 5 ms	Standard	1 kHz 50 ms rf = 1 ms	Standard	500 Hz 26 ms rf = 5 ms	4 kHz 26 ms rf = 5 ms	500 Hz 50 ms rf = 5 ms	4 kHz 50 ms rf = 5 ms
Silence	55 dB SPL	3.08	45 dB SPL 65 dB SPL 85 dB SPL	∼ 1.5 ∼ 0.88 ∼ 0.7	25 dB SPL 55 dB SPL 85 dB SPL	∼ 2.08 ∼ 1.7 ∼ 1.4	∼ 2.08 ∼ 1.7 ∼ 0.5	∼ 0.97	∼ 1.45
	75 dB SPL	1.06							
55 dB SPL	55 dB SPL	3.68							
	75 dB SPL	1.03							
75 dB SPL	55 dB SPL	10.78							
	75 dB SPL	1.85							

Table II.4 – Comparison of DLs (in dB) measured in experiment 1 and in the literature. DL is defined as $10log(1+\Delta I/I)$ (rf : rise/fall time).

II.2.3.3 Partial masking

One possible reason for the large DL increase in the loud context is that the context partially masks the standard (and target). It has been shown that discrimination thresholds increase when the sensation level of the sound decreases (Florentine et al. 1987). Partial masking resulting in a decrease of the sensation level of the signal, we may expect it to induce an increase in DL. Florentine et al. (1987) measured, in silence, for a pure tone of 2 kHz, a DL increase of about 1 dB when the level of the pure tone was decreased from 55 down to 10 dB SL and an increase of about 1.3 dB when the level of the pure tone was decreased from 75 down to 30 dB SL.

In our experiment, the participants were selected so that the standard detection threshold in the loud context was lower than 45 dB SPL. Thus, the sensation level of the soft standard in the loud context was always greater than 10 dB SL (actually ranging between 10 and 16.5 dB SL across participants) and greater than 30 dB SL for the loud standard in the loud context (between 30 and 36.5 dB SL). At 2 kHz, the average detection threshold of the group in silence was 3.0 dB ; it can therefore be considered that dB SL $\sim$ dB SPL.

From 55 dB SL (silence condition in our experiment for the soft standard) to 10 dB SL (the minimum sensation level of the soft standard in the loud context), we observe an increase of DL of 7.7 dB in average, which is considerably larger than the 1 dB found by Florentine et al. (1987). In contrast, from 75 dB SL (silence condition in our experiment for the loud standard) to 30 dB SL (the minimum sensation level of the loud standard in the loud context), we observe an increase of DL of 0.8 dB, which is more in line with the data from Florentine et al. (1987).

Data from Florentine et al. (1987) were, however, obtained for 500-ms long signals whereas we used 50-ms signals. Florentine (1986) measured the effect of duration on DL for several frequencies and levels. She found an interaction between level and duration, showing that DL increased by factors of 1.5 at 40 dB SPL and 1.75 at 85 dB SPL when stimulus duration decreased from 500 to 50 ms. This means that for long durations, DLs decrease faster with increasing level than for shorter stimuli. These data, therefore, demonstrate that the short duration of our signals is unlikely to be the explanation of the difference in DL variation when the sensation level increases between the present study and Florentine et al. (1987).

It is however important to note that this comparison is made between sensation level relative to the absolute threshold for Florentine (1986) ; Florentine et al. (1987) (without noise), and the sensation level within a context is relative to the detection threshold with a non-simultaneous noise context in the present study. Some of these comparisons might thus be taken with caution.

It is interesting to note that Widin et al. (1986) measured intensity discrimination of pure tones following a forward masker with a 5-ms delay between the masker and the test signal and reported that in some cases, the resulting increase in DL was greater than that expected on the basis of a shift in the sensation level of the pure tone alone ; this is in line with our results.

II.2.3.4 Mid-level hump and confusion

Another explanation for the increase of DLs in the loud context comes from studies on intensity discrimination in forward masking. It has been shown that forward-masked intensity discrimination thresholds are larger than in silence, even for masker and standard delays as long as 100 ms. In an intensity discrimination task with forward masking, each signal (standard or target) is preceded by a masker. Typically, the signals are presented 100 ms after the maskers, which usually have a finite duration of about 100 ms (e.g. Zeng et al. 1991 ; Oberfeld 2008). This paradigm is different from ours, where the "masker" (context) is continuous throughout the experiment except during the presentation of the standard and target, so there is no silence between masker and standard. Zeng and Turner (1992) measured the intensity discrimination threshold of a 1-kHz pure tone standard preceded by a 90 dB SPL masking narrow-band noise (900-1100 Hz) as a function of the level of the standard (from 20 to 100 dB SPL) for several delays between the end of the maskers and the beginning of the standard and target. They showed that DL increased with increasing standard level up to about 50 dB SPL, then decreased for standard levels up to 90-100 dB SPL to eventually reach threshold in silence. As introduced in Chapter I, this non-monotonic function is called the mid-level hump and has been shown in several studies (Zeng et al. 1991 ; Zeng and Turner 1992 ; Plack et al. 1995 ; Zeng 1998 ; Oberfeld 2008).

For a mid-level standard of 40-60 dB SPL (i.e., the level at which the DL was highest for a 90-dB SPL masker), Zeng and Turner (1992) measured the effect of masker level on DL (for a delay between masker and standard of 100 ms). They found a DL equal to the DL in silence for masker levels below 40 dB SPL, followed by a significant increase in DL when the masker level increased from 40 up to 90 dB SPL, to reach a DL 6 dB higher than in silence for a 90-dB SPL masker.

It should be noted that the mid-level hump is also observed when the standard is not directly masked by the masking noise. Zeng et al. (1991) measured DLs with a masker presented 100 ms before the standard. With this delay, the detection threshold of the standard was equal to the detection threshold in silence, yet the DL was degraded by the presence of the masker, and showed the mid-level hump. In Oberfeld (2008), detection thresholds with the same delay increased by only 3 dB on average between silence and a 55-dB SPL masker, and by 5 dB between silence and a 85-dB SPL masker.

The mid-level hump is also larger for shorter delays (Zeng and Turner 1992). We may expect it to be particularly large in our experiment where we used no delay between context and test signals. This may partly explain the almost 8-dB difference in threshold between the DL of the soft standard in the loud context and the DL of the soft standard in silence. Part of this difference in DL may also be attributed to partial masking of the context on the standard.

Oberfeld et al. (Oberfeld and Stahn 2012 ; Oberfeld et al. 2014) hypothesized that the mid-level hump could result from poor segregation of the target from the masker. Their hypothesis was that masker and target are perceptually grouped together, and that the level difference between masker and target needs to be increased a lot for the two sounds to be perceived as different objects. They found that this "object-based selective attention" explanation could, at least partially, account for the mid-level hump.

Related to this failure of perceptually segregating context and test signals is the effect of confusion (Neff 1985, 1986). If context and test signals were perceptually too similar, participants may have had difficulties to identify when the test signals were being presented. Ironically, we would expect this effect to be stronger when context mean level and standard level are the same (which are the conditions for which DRA is expected to improve performance). In our experiment in particular, the context and standard are similar (in duration and frequency content) ; the context level is fluctuating, it is narrowband, and presented both before and after the standard without delay : as presented in Chapter I, these factors are all parameters expected to increase confusion. The standard is a 50-ms pure tone, which makes it the ideal candidate for the effects of confusion. Several factors can help reduce the effect of confusion : the addition of a delay between masker and standard, a difference in presentation mode of the masker relative to the standard (for example lateralization, contralateral masker, or binaural masker ; see Oberfeld et al. 2012 ; Roverud and Strickland 2015a).

To conclude, the increase of DL observed for the standards in the loud context in the present study could be due to the mid-level hump and/or confusion, which may therefore potentially obscure the effects of DRA.

II.2.3.5 Off-frequency listening

As the context has a rather narrow bandwidth, it is possible that some spread of excitation helps intensity discrimination when the standard is louder than the context : that could explain why the discrimination thresholds were not higher in the soft context than in the loud context for the loud standard, but the other way round. During the presentation of the context, the targeted neurons may adapt to the statistics of the context, which is soft and thus unfavorable for the loud standard. When the latter is played, the spread of excitation would elicit activity in neurons outside the frequency regions excited by the context, which would have received no prior information and thus would not have adapted to any stimulus. These unadapted regions of the auditory system would respond only to the standard, and could in turn provide additional information for intensity discrimination, thereby lowering the thresholds.

For a context mean level equal to the standard level or higher, the spread of excitation of the standard would be the same or smaller than that of the context, so off-frequency may not help. If the context is louder than the standard, the thresholds should therefore increase. In the case of matching context and standard levels, dynamic range adaptation is expected to improve the encoding of intensity differences, so the DLs should be lower than in the (non-matching) conditions without taking spread of excitation into account.

The discrimination thresholds for the loud standard could thus be caused by a mixture of several effects, going in potentially opposite directions : i) the context masks the standard and increases DLs (forward + backward masking & confusion), with a stronger impact for the loud context mean level ; ii) DRA lowers the DLs in the loud context and increases them in the soft context ; and iii) off-frequency listening lowers the DLs in the soft context. This hypothetic combination does not display results in agreement with DRA in the present experiment : the effects of masking and off-frequency listening are thus possibly more important. Suppressing off-frequency

listening would leave only the effects of DRA and of masking, and should therefore increase the discrimination thresholds in the soft context but not in the loud context.

II.2.3.6 Overview of the control experiments

In order to test whether the different phenomena described above affected our data and obscured the effect of DRA on intensity discrimination, four additional control experiments were designed.

Experiment 2 tested the same conditions as experiment 1, but using lower levels. This experiment had several aims. First, using lower standard levels may reduce the effect of the mid-level hump (Zeng and Turner 1992). Second, as shown by the model simulations, lower levels may avoid being in the saturated portion of the rate-level function for HSR fibers. Third, the DLs measured for the loud standard were extremely low (less than 2 dB on average) and it is possible that we reached the limits of discrimination performance (i.e., that participants cannot do any better than what they already do), thereby preventing the effect of DRA to pop out. Using lower levels should degrade performance and may leave headroom to measure an effect of DRA.

In experiment 3, a notched noise was added to the stimuli in order to limit off-frequency listening.

Since the mid-level hump, confusion, forward- and backward-masking effects are reduced with increases in delay, silent gaps between the context and the test signals were introduced in experiment 4.

In experiment 5, the context was played diotically, while the standard remained monaural. This aimed to reduce confusion effects as well as masking.

The auditory nerve model was also run in experiments 2 and 4 for comparison with the predictions of experiment 1.

II.3 Control experiments

II.3.1 Experiment 2 : Discrimination at lower levels

This second experiment was the same as the first one, except for a change in tested levels. As mentioned above, one reason for running this experiment is to try to reduce the mid-level hump which is often seen with loud maskers and medium-level standards. Another reason comes from comparison with physiological studies. One of the differences between the latter and experiment 1 is the use of a much narrower bandwidth : in the present study, the standard was a pure tone and the context a narrowband noise, while Dean et al. (2005) and following studies generally used broadband noise. The level used in the 2-kHz frequency band was higher in experiment 1 than for the broadband noise used in physiological studies.

Moreover, the simulated rate-level functions from the first experiment suggests that the 75-dB SPL standard was always in the saturated part of the neurons' response, even in the presence of a context : lower levels might allow to observe other points in the rate-level function that might not be saturated. Furthermore, the near-miss to Weber's law predicts that the discrimination thresholds should be higher for lower

standard levels ; it could thus be easier to observe an effect of DRA than with the high-level, 75-dB SPL standard.

II.3.1.1 Material and methods

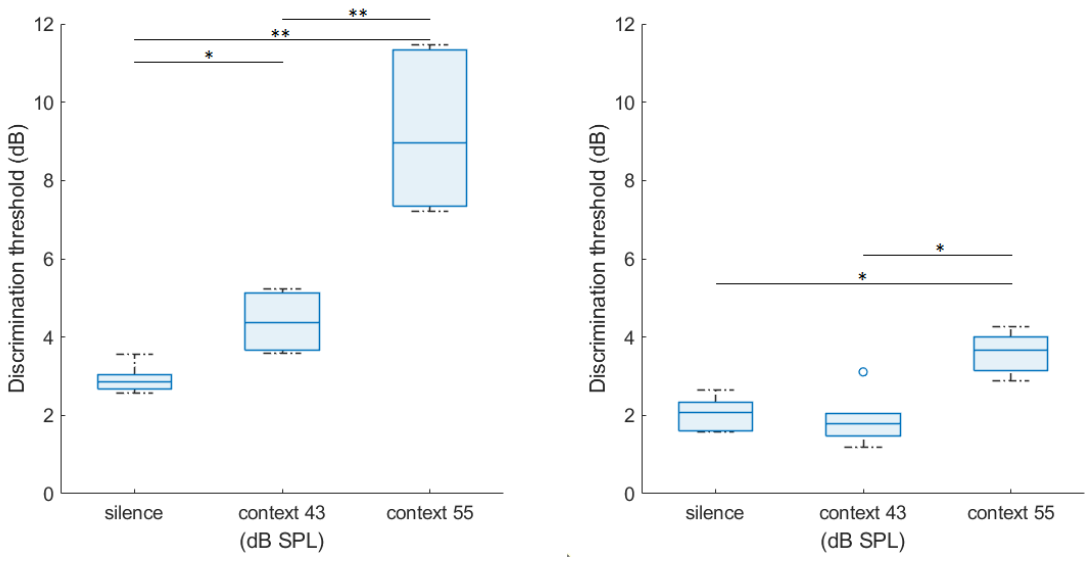
The tested levels for both context and standard were 43- and 55 dB SPL. The loud level was the same as the soft level of experiment 1. The lower level (43 dB SPL) was chosen based on the results of the preliminary experiment described in the Methods subsection of experiment 1. As we wanted to keep the difference in levels between the two contexts above 10 dB and identical for standards and contexts, the soft level was chosen 10 dB above the mean detection level of all ten participants in the 55-dB SPL context, which was 33 dB SPL (see table A.1). Six participants were included ; the four others had detection thresholds above 36 dB SPL in the 55-dB SPL context.

The exact same protocol and material was used as in experiment 1.

We also used the MAP model to run some simulations with the used levels and compare them with those obtained for experiment 1 and in Grange et al. (2022).

II.3.1.2 Results

The discrimination thresholds are presented in figure II.5a for the soft, 43-dB SPL standard, and figure II.5b for the loud, 55-dB SPL standard.



(a) 43 dB SPL standard **(b)** 55 dB SPL standard

Figure II.5 – Results of experiment 2 for six participants. Discrimination thresholds are plotted against context type (including silence), as in figure II.3. They are significantly higher in the loud context than in the other conditions for both standards, and in the soft context than in silence for the soft standard.

Compared to experiment 1 (figure II.3), there was an overall increase in intensity discrimination thresholds. However, the same general pattern was observed : the discrimination thresholds were significantly higher in the loud context than in the soft one both for the soft and loud standard (table II.5).

Standard level	Context level	Condition	Mean (dB)	Standard error (dB)
43 dB SPL	silence	A	2.93	0.14
	43 dB SPL	C	4.39	0.32
	55 dB SPL	E	9.21	0.79
55 dB SPL	silence	B	2.05	0.17
	43 dB SPL	D	1.90	0.28
	55 dB SPL	F	3.60	0.22

Table II.5 – Details on the averaged results for experiment 2, with the corresponding condition letter in figure II.1. The average discrimination thresholds tend to increase with the context mean level for the 43-dB SPL standard. For the 55-dB SPL standard, the thresholds decrease in the 43-dB SPL context compared to silence, while they increase in the 55-dB SPL context.

As in experiment 1, a two-way repeated-measures ANOVA (with factors standard level : 43 or 55 dB SPL, and context mean level : silence, 43 or 55 dB SPL) revealed significant effects of the standard level ($F_{1,5} = 69.4$, p = 4.07e-4), of the context mean level ($F_{2,10} = 49.2$, p = 4.10e-4), and of the interaction ($F_{2,10} = 31.2$, p = 0.002). The mean level of the context thus affects the different standard levels differently, as suggested by the average thresholds in table II.5. This is confirmed by the pairwise comparisons between context conditions for each standard level, presented in table II.6. The only pairwise comparison that was not significant is between silence and soft context for the loud standard.

Standard level	Context 1	Context 2	t	p adjust	significance
43 dB SPL	silence	43 dB SPL	-4.73	0.016	*
	silence	55 dB SPL	-7.11	0.003	**
	43 dB SPL	55 dB SPL	-6.77	0.003	**
55 dB SPL	silence	43 dB SPL	1.17	0.879	ns
	silence	55 dB SPL	-4.71	0.016	*
	43 dB SPL	55 dB SPL	-4.87	0.014	*

Table II.6 – Pairwise t-test for each standard and context mean level. The loud context gives significantly higher thresholds than the soft context for both standard levels, and than the silence for the soft standard.

Interestingly, there is an overlap between the conditions performed here and those performed in experiment 1 : the 55-dB SPL standard in silence and in the 55-dB SPL context were common to these two experiments. The results from experiment 1 of the same six participants are shown in figure II.6 along with those from experiment 2. The discrimination thresholds in the 55-dB SPL context were very similar between the two experiments ; in silence, they were lower in experiment 2 (t = 4.24, p = 0.008). This might be due to the interleaved protocol, as will be discussed later.

Again, simulations were run using the stimuli of the psychophysical task ; the rate-level functions and slopes are respectively presented in figures II.7a and II.7b. The same protocol, fitting algorithm, analysis methods and number of averaging as simulations for experiment 1 were used.

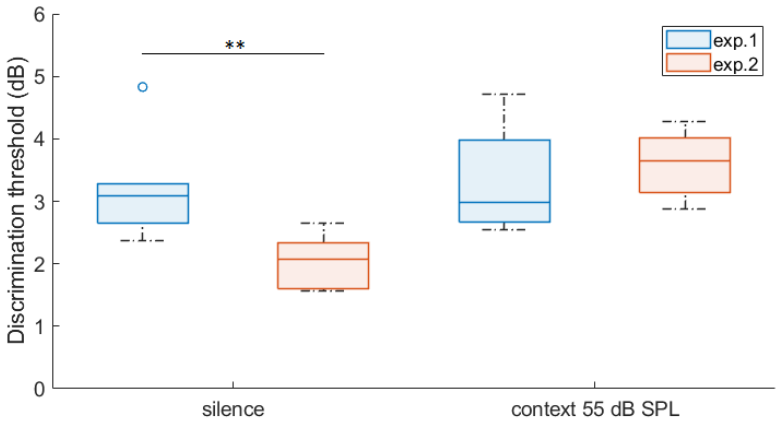
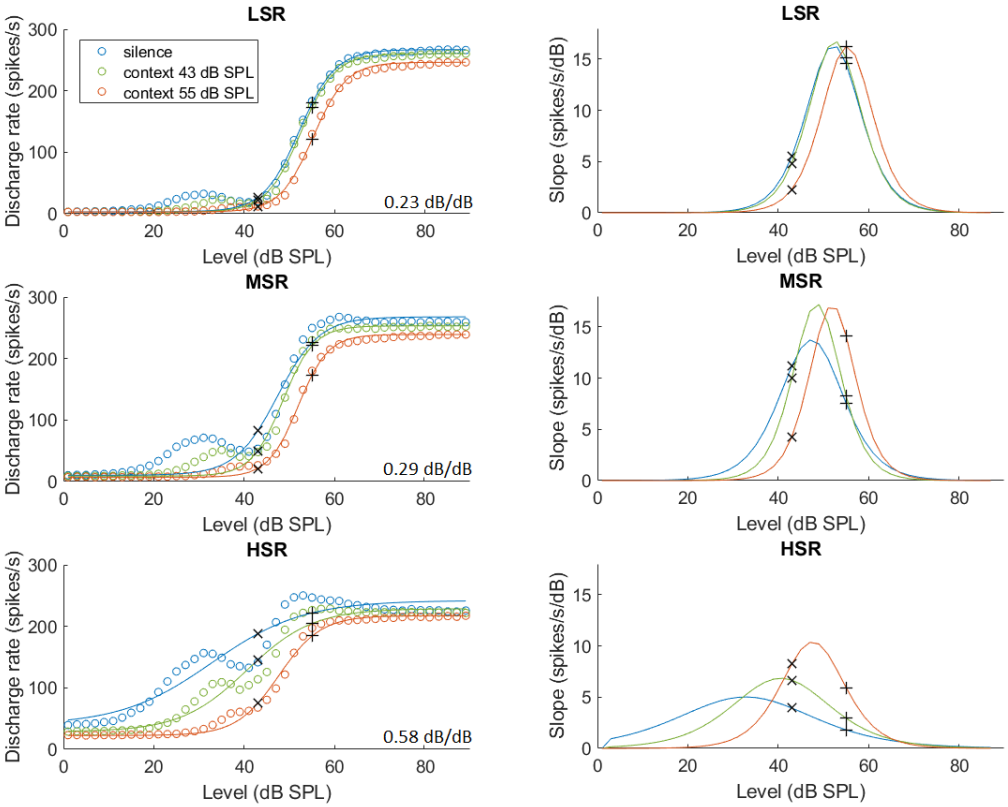


Figure II.6 – Discrimination thresholds for the six participants of experiment 2 for the 55-dB SPL standard in silence and in the 55-dB SPL context - the common conditions of experiment 1 and 2.



(a) Rate-level functions. **(b)** Slopes.

Figure II.7 – Rate-level functions with shift between contexts (a) and slopes (b) obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 2, averaged over ten simulations. The 43- and 55-dB SPL standard are represented respectively as black x and +.

Standard level	Fiber type	Context type	Slope (spikes/s/dB)
43 dB SPL	LSR	silence	5.52
		43 dB SPL	4.83
		55 dB SPL	2.28
	MSR	silence	11.18
		43 dB SPL	10.00
		55 dB SPL	4.25
	HSR	silence	3.93
		43 dB SPL	6.66
		55 dB SPL	8.24
55 dB SPL	LSR	silence	14.55
		43 dB SPL	15.12
		55 dB SPL	16.24
	MSR	silence	7.55
		43 dB SPL	8.26
		55 dB SPL	14.14
	HSR	silence	1.81
		43 dB SPL	2.94
		55 dB SPL	5.94

Table II.7 – Slopes around both standard levels for each context and fiber types in figure II.7b.

As expected, the rate-level functions for silence and 55-dB SPL context (respectively blue and red) are almost identical to those of experiment 1 (figure II.4a). One of the main differences is that the rate-level functions are not saturated for the loud standard (figure II.7b).

The slope values measured at the standard levels are presented in table II.7, and are steeper than those of experiment 1 (table II.3). The rate-level functions are not saturated for the loud standard - as expected. For the 43-dB SPL standard, the slopes are steeper in the 43-dB SPL context than in the 55-dB SPL one, except for the HSR fibers ; for both the LSR and MSR however, silence brings the steepest slopes. For the 55-dB SPL standard, the slopes are always steeper in the 55-dB SPL context, and shallowest in silence. These simulations globally support our hypothesis, but disagree with the results of the behavioral task for the 55-dB SPL standard.

The horizontal shift in rate-level function, indicated in figure II.7a, is much larger for the HSR fibers (0.58 dB/dB) than for the MSR and LSR fibers (respectively 0.29 and 0.23 dB/dB). For the HSR fibers, the shift value is similar to that obtained in experiment 1 (figure II.4a), but the shifts for the LSR and MSR fibers are much smaller in this experiment. Grange et al. (2022) found that, for frequencies of 2.1 and 3.8 kHz, HSR fibers adapted more than the other types ; but we have no explanation on why the shift is smaller for the LSR and MSR fibers than in experiment 1.

II.3.2 Experiment 3 : Discrimination with notched noise

In an attempt to limit off-frequency listening, a notched noise was added to the stimuli. The thresholds were expected to increase in every condition because of the added noise compared to experiment 1, but our hypothesis was that the effect should be more dramatic for the loud standard in the soft context, hypothetically great enough to change the tendency in figure II.3b, and get the discrimination thresholds of the loud standard in the soft context mean higher than in the loud context (i.e. revealing an effect consistent with DRA).

II.3.2.1 Material and methods

The material and protocol of experiment 3 were the same as in experiment 1. The stimuli were based on experiment 1. We only tested the 75-dB SPL standard, in both 55- and 75-dB SPL context mean levels (still respectively named soft and loud), but not in silence. The notched noise was a pink noise, bandpass filtered between 1000 and 4000 Hz (8th order Butterworth filters, 48 dB/octave), then bandstop filtered in the ERB around 2 kHz to match the context bandwidth (1880 - 2120 Hz - IIR filter, 280 dB/octave). Two different notched noise levels were tested, 35 dB SPL and 65 dB SPL. The notched noise was played continuously throughout the task, from a 2-h looped wave file. All the participants of experiment 1 took part.

II.3.2.2 Results

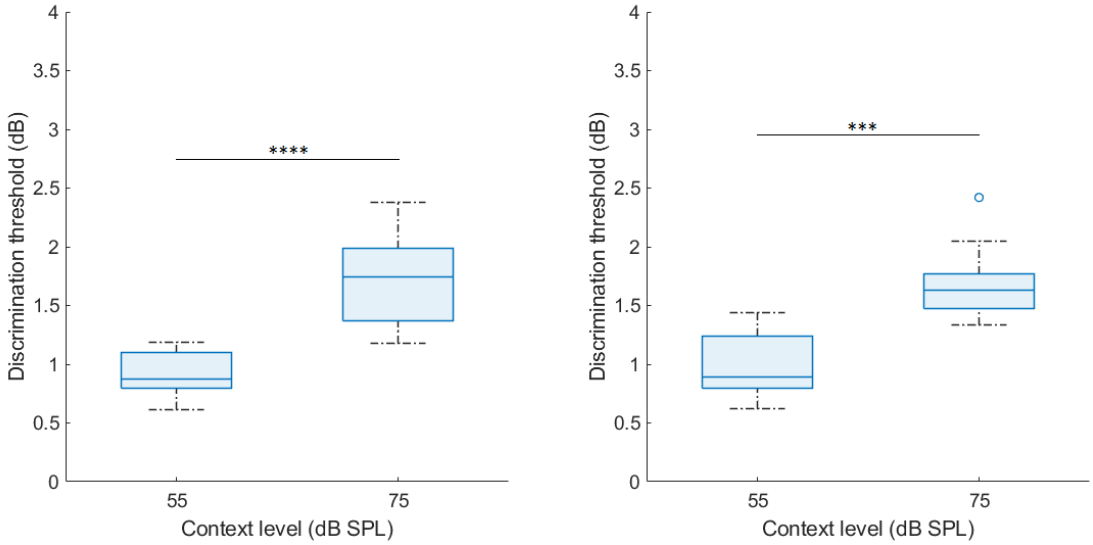
Figure II.8 shows the discrimination thresholds depending on the context mean level for the different notched noise levels, and table II.8 gives the average discrimination thresholds and standard error in each case.

Notched noise level	Context level	Mean (dB)	Standard error (dB)
35 dB SPL	55 dB SPL	0.93	0.06
	75 dB SPL	1.73	0.13
65 dB SPL	55 dB SPL	1.00	0.09
	75 dB SPL	1.70	0.10

Table II.8 – Details on the averaged results for experiment 3. The average discrimination thresholds and standard deviations tend to increase in the 75-dB SPL context but they do not change between the 35- and 65-dB SPL notched noise.

Figure II.8a shows the thresholds obtained with the 35-dB SPL notched noise, and figure II.8b with the 65-dB SPL notched noise. As in experiment 1, the discrimination thresholds increased with an increase in context mean level. However, contrary to our hypothesis, no difference was observed with the addition of a notched noise of any level.

The data were analyzed in a two-way repeated-measures ANOVA with factors context mean level (2 levels : 55 and 75 dB SPL), and notched noise level (2 levels : 35 and 65 dB SPL). The analysis (table II.9) revealed a significant effect of the context mean level, but not of the notched noise level. In fact, the average discrimination thresholds are very close to those obtained in experiment 1 (table II.1). The pairwise



(a) 35-dB SPL notched noise **(b)** 65-dB SPL notched noise

Figure II.8 – Results of experiment 3 for the loud standard (75 dB SPL). Discrimination thresholds are plotted like in experiment 1 (see figure II.3). There is no change in the discrimination thresholds with the addition of different levels of notched noise : the results are still higher in the loud context than in the soft context, and the level of the notched noise does not seem to have an effect.

Effect	dof	dof err.	F	p	significance
notched noise level (nn level)	1	9	0.053	0.824	ns
context mean level	1	9	188.284	2.44e-7	*
nn level×context mean level	1	9	0.250	0.629	ns

Table II.9 – Repeated measures ANOVA (dof : degrees of freedom) on the data of experiment 3. The factors were context mean level and notched noise level. Only the level of the context had a significant effect on the discrimination thresholds.

comparisons (table II.10) reveal the same differences between contexts as those obtained in experiment 1.

Notched noise	Context 1	Context 2	t	p value	significance
35 dB SPL	55 dB SPL	75 dB SPL	-7.73	2.91e-5	****
65 dB SPL	55 dB SPL	75 dB SPL	-6.09	1.82e-4	***

Table II.10 – Pairwise t-test between context mean levels in experiment 3. The discrimination thresholds are significantly different for both notched noise levels.

The discrimination thresholds are not influenced by the notched noise : pairwise comparisons (not shown) confirm that there is no significant change in discrimination thresholds between experiment 1 and 3. The average threshold and standard error are very similar. The notched noise did not change the direction in which the discrimination thresholds evolved compared to experiments 1 and 2 for the loud standard : thresholds were still higher in the loud, 75-dB SPL context, than in the soft context.

II.3.3 Experiments 4 & 5 : Discrimination with gaps or using a diotic context

II.3.3.1 Material and methods

The last two experiments (4 and 5) used the same protocol and material as experiment 1. These experiments were designed to reduce the masking effects that may have prevented the observation of DRA : the mid-level hump and the confusion effect. According to several studies, these effects should be reduced with an increase in delay between masker and signal, or with the addition of cues that help separate perceptually the masker and standard (Moore 1981 ; Neff 1985, 1986 ; Zeng et al. 1991 ; Zeng and Turner 1992).

All participants of experiment 1 took part in both experiments. The condition in silence was not tested. Discrimination thresholds for both standard levels were still measured with an interleaved procedure.

For experiment 4, 50-ms gaps were added before and after the presentation of the test signals. Care was taken to keep this gap shorter than the time constants of DRA observed at the level of the auditory nerve and of the inferior colliculus, so we do not expect the gap to cancel potential effects of adaptation.

The MAP model was used to extract the rate-level functions of the auditory nerve in response to the stimuli used in experiment 4, to compare to the rate-level functions obtained for experiment 1. It was also run on the (not-tested) condition in silence.

For experiment 5, the context was presented diotically, but the standard was kept monaural to reduce confusion effects.

II.3.3.2 Results

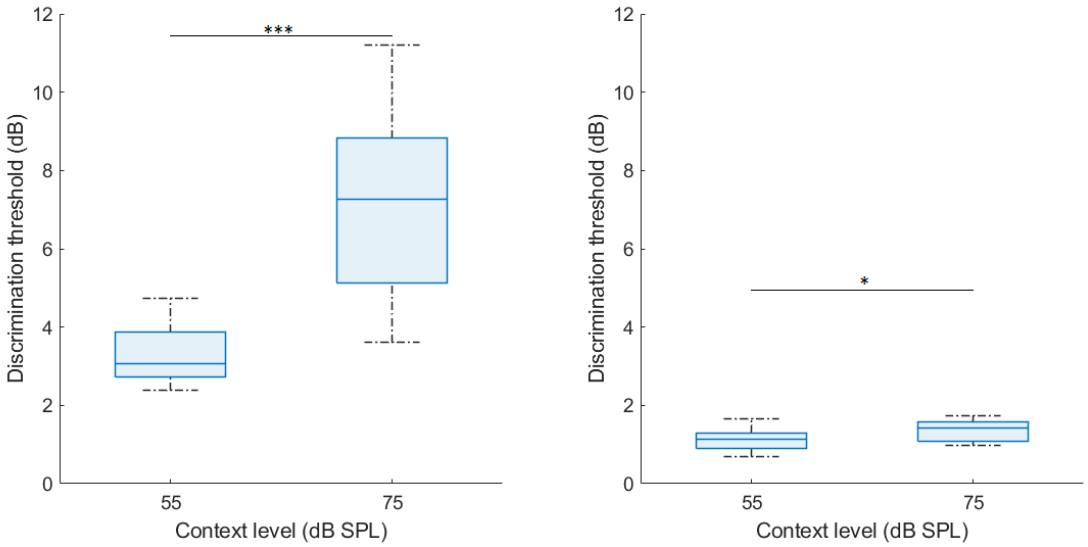
Experiment 4

For experiment 4, the discrimination thresholds with added gaps presented in figure II.9 were lower in the soft context than in the loud context for both standard levels, as in experiment 1. The data were analyzed in a two-way repeated-measures ANOVA, with factors standard level (2 levels : 55 and 75 dB SPL), and context mean level (2 levels : 55 and 75 dB SPL).

Standard level	Context level	Mean (dB)	Standard error (dB)
55 dB SPL	55 dB SPL	3.31	0.26
	75 dB SPL	7.22	0.72
75 dB SPL	55 dB SPL	1.13	0.10
	75 dB SPL	1.37	0.09

Table II.11 – Averaged results for experiment 4, with 50-ms gaps between the presentation of the context and the standards. The discrimination thresholds increase with the context mean level.

The analysis showed significant effects of the standard level ($F_{1,5}$ = 79.7, p = 9.13e-6), the context mean level ($F_{2,10}$ = 48.9, p = 6.39e-5) and their interaction ($F_{2,10}$ = 33.3 p = 2.69e-4). The discrimination thresholds increased with the mean level



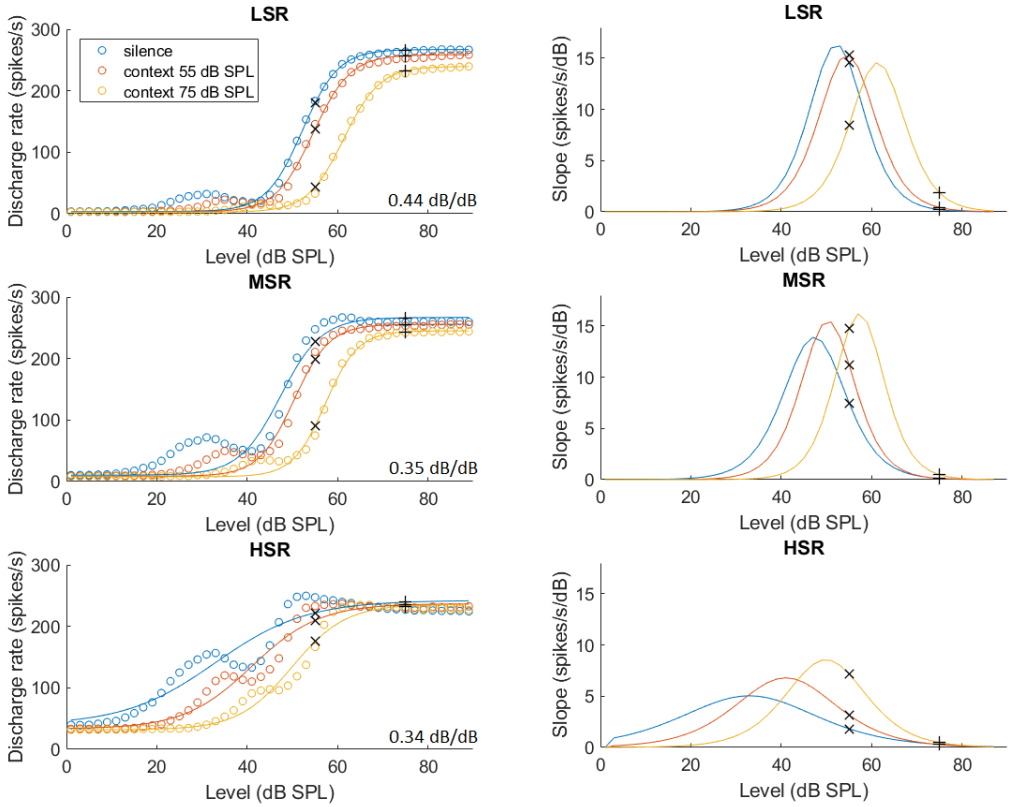
(a) 55 dB SPL standard **(b)** 75 dB SPL standard

Figure II.9 – Results of experiment 4, with silent gaps between context and standard. The discrimination thresholds are plotted for both standard levels in both context levels, as in figure II.3. There is a significant increase in discrimination thresholds in the loud context compared to the soft context for both standard levels.

Standard level	Fiber type	Context type	Slope (spikes/s/dB)
55 dB SPL	LSR	silence	14.56
		55 dB SPL	15.31
		75 dB SPL	8.48
	MSR	silence	7.49
		55 dB SPL	11.23
		75 dB SPL	14.78
	HSR	silence	1.82
		55 dB SPL	3.16
		75 dB SPL	7.17
75 dB SPL	LSR	silence	0.214
		55 dB SPL	0.429
		75 dB SPL	1.85
	MSR	silence	0.130
		55 dB SPL	0.118
		75 dB SPL	0.477
	HSR	silence	0.298
		55 dB SPL	0.282
		75 dB SPL	0.485

Table II.12 – Slopes around both standard levels for each context and fiber types in figure II.10b. They are much lower for the 75-dB SPL standard as most fibers are already saturated.

of the context (see table II.11), as in the previous experiments, but were lower in the loud context than in experiment 1.



(a) Rate-level functions. **(b)** Slopes.

Figure II.10 – Rate-level functions with shift between contexts (a) and slopes (b) obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 4, averaged over ten simulations. The 55- and 75-dB SPL standard are represented respectively as black x and +.

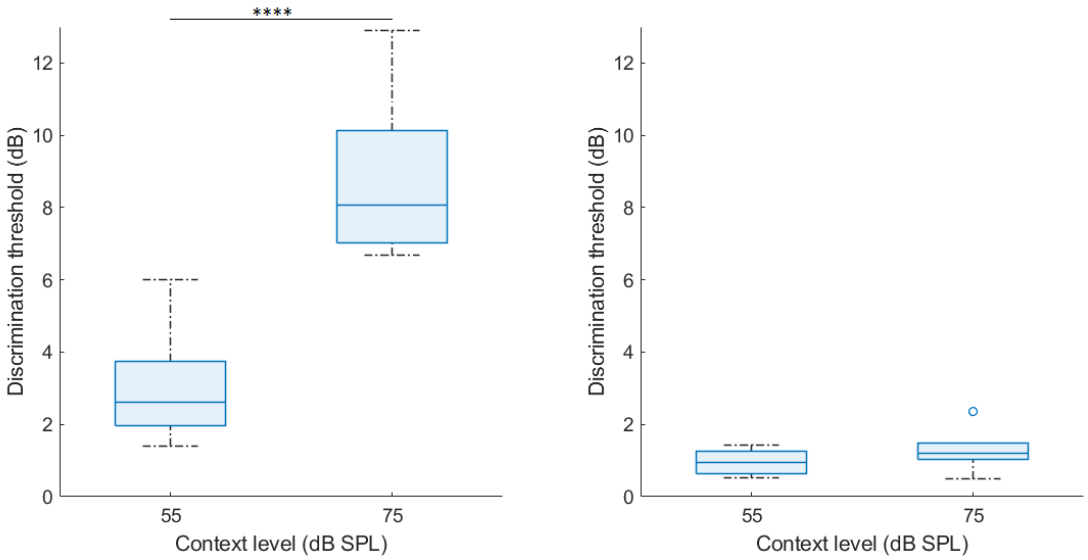
The difference between contexts for both standards was also significant (soft standard : t = -6.45, p = 1.19e-4, loud standard : t = -2.39, p = 0.04).

Figure II.10a presents the rate-level functions obtained by simulations run with the model from Grange et al. (2022). The evolution of the firing rate is plotted as a function of excitation level for each context condition and fiber type, in the ERB around the 2-kHz pure tone standard. Figure II.10b presents the slopes in the different context conditions of the corresponding rate-level functions.

The slopes measured at the standard levels are presented in table II.12. The LSR and MSR fibers have steeper slopes than in experiment 1 for the 55-dB SPL standard in the 75-dB SPL context, but shallower ones for the 75-dB SPL standard in the same context (table II.3) ; the slope for the HSR fibers in both contexts is greater in experiment 4 than in experiment 1 for the 75-dB SPL standard while it is lower for the 55-dB SPL standard.

With gaps, the shift is similar for all fiber types ; it is indicated in figure II.10a. The shift is slightly higher for LSR fibers than for MSR and HSR (respectively 0.44, 0.35 and 0.34 dB/dB) : compared to experiment 1 (figure II.4a), the shifts are slightly reduced.

Experiment 5



(a) 55 dB SPL standard **(b)** 75 dB SPL standard

Figure II.11 – Results of experiment 5. The discrimination thresholds are plotted for both standard levels in each context mean level, as in figure II.3. Only the soft standard displayed a significant increase in discrimination thresholds with the diotic context.

Standard level	Context level	Mean (dB)	Standard error (dB)
55 dB SPL	55 dB SPL	3.03	0.44
	75 dB SPL	8.76	0.69
75 dB SPL	55 dB SPL	0.95	0.10
	75 dB SPL	1.25	0.15

Table II.13 – Averaged results for experiment 5, with the context presented to both ears. Average thresholds increase with context mean level.

The discrimination thresholds for the diotic presentation of the context are presented in table II.13 and figure II.11 : for the loud standard, the average threshold is just higher in the loud context than in the soft context, while it is much higher in the loud context for the soft standard. As in the first experiment, a two-way repeated-measures ANOVA (with the same factors as experiment 4) showed significant effects of standard level ($F_{1,9}$ = 103.8, p = 3.06e-6), context mean level ($F_{1,9}$ = 107.2, p = 2.68e-6) and their interaction ($F_{1,9}$ = 211.7, p = 1.47e-7).

The pairwise t-tests revealed a significant increase in threshold for the diotic presentation of the context for the soft standard only (t = -12.4, p = 5.90e-7). A small increase was also visible for the loud standard, but not significant (t = -1.92, p = 0.087).

Figure II.12 shows the average difference in discrimination thresholds between the 75-dB SPL context and the 55-dB SPL one, for both standard levels, for experiments 1, 4 and 5. This difference in discrimination thresholds is always much lower for the

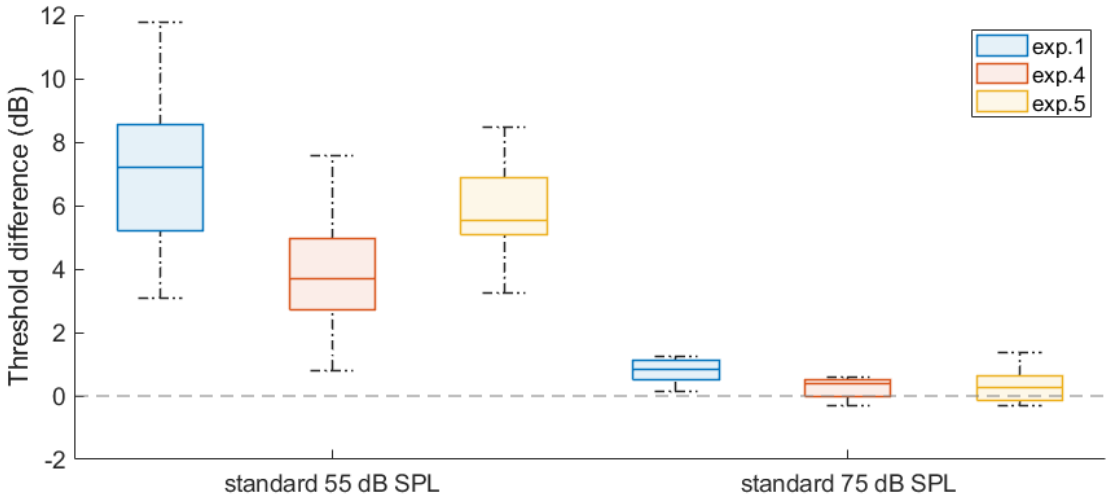


Figure II.12 – Difference in discrimination thresholds between the 75- and 55-dB SPL contexts for both standard levels in experiments 1, 4 and 5.

Standard level	Context	t	p value	significance
55 dB SPL	55 dB SPL	1.56	0.154	ns
	75 dB SPL	4.05	0.003	**
75 dB SPL	55 dB SPL	-0.866	0.409	ns
	75 dB SPL	3.09	0.013	*

Table II.14 – Pairwise t-test comparisons between the thresholds with the same context mean level and standard level in experiments 1 and 4. Significant differences appear for the loud context for both standard levels : experiment 4 gives significantly lower discrimination thresholds in the 75-dB SPL context.

Standard level	Context	t	p value	significance
55 dB SPL	55 dB SPL	2.01	0.076	ns
	75 dB SPL	4.00	0.003	**
75 dB SPL	55 dB SPL	0.634	0.542	ns
	75 dB SPL	3.84	0.004	**

Table II.15 – Pairwise t-test comparisons between thresholds of experiments 1 and 5 for each standard and context mean level. Significant differences appear for the loud context for both standard levels : experiment 5 gives significantly lower discrimination thresholds in the 75-dB SPL context.

loud than for the soft standard. For both standard levels, the difference is greatest in experiment 1 ; for the soft standard, it is lowest with the addition of gaps (experiment 4), and with the binaural context (experiment 5) for the loud standard. Moreover, pairwise comparisons presented in tables II.14 and II.15 confirm that the discrimination thresholds in experiments 4 and 5 are lower in the loud context for both standard levels than in experiment 1. These comparisons confirm that both the gaps and the diotic context have an effect on the discrimination thresholds.

II.3.4 Discussion

Overall, the data obtained both in the main and control experiments were only partially consistent with the initial hypothesis, at least for the soft standard : the discrimination thresholds were never lower in the loud context than in the soft context. Below we discuss the results and implications of the control experiments ; first, the possible impact of the procedure on the thresholdds ; second, the absence of change in tendencies for lower standard levels ; third, the absence of impact of the notched noise on the loud standard ; and last, the possible causes underlying the effects of introducing gaps and presenting the context diotically.

II.3.4.1 Interleaved

The discrimination threshold for the 55-dB SPL standard in silence was lower in experiment 2 than in experiment 1, for the same participants. It has previously been shown that the random alternation of standard levels in the procedure increases the thresholds compared to using one-level track measurements (Oxenham and Buus 2000 ; Pienkowski and Hagerman 2009), and that a large level difference between tracks produce a larger deterioration in threshold (Mori and Ward 1992). Because we reduce the level difference between standards in experiment 2, the interleaved protocol might have less impact on the discrimination thresholds.

It is also possible that the interleaving has less effect on louder levels in the protocol. Indeed, Berliner and Durlach (1973) ; Berliner et al. (1977) described an "edge effect" : participants perform better when tested at the edges of the roved level range.

Interleaving trials at different levels can also induce long-term changes in discrimination performance. Simpson et al. (2014) first measured individual discrimination thresholds at 79.4 % with a 2IFC, three down-one up adaptive procedure at 35, 55 and 75 dB SPL, for 50-ms bursts of broadband noise. They then tracked the intensity discrimination performance over time for the three interleaved levels with different probabilities of appearance. They found an increase in performance after a minute for the level that were furthest from the most probable level, over groups of 40 trials, consistently with Berliner et al. (1977). But even more interesting, performance for the 75-dB SPL stimulus never got below 65 %, regardless of whether it was the most or the least probable level in the task, while the other two levels frequently dropped below that value (between 60 and 75 % for the 35-dB SPL standard, and between 55 and 70 % for the 55-dB SPL one - max for the 75-dB SPL : 80 %).

The loudest level might thus have an advantage over other levels. The discrimination thresholds could increase in interleaved protocol compared to separate measurements, but with less impact on the loudest level, and more for larger differences between standard levels. That may explain the observed difference in discrimination thresholds for the 55-dB SPL standard in silence between experiments 1 and 2 (figure II.6).

II.3.4.2 Lower levels

The use of lower levels was expected to bring a change in the observed effect of the context compared to experiment 1. In experiment 2, the difference in discrimination thresholds between silence and loud context for the soft standard was smaller than

in experiment 1 (6.3 dB instead of 7.7), probably due to a reduction in the mid-level hump.

Because the step-size between levels was reduced, the shift in rate-level functions predicted by the model was smaller and both standards were in unsaturated parts of the response functions. The saturation of the neurons is thus, according to the model, not the cause of the observed tendencies. Other mechanisms might overshadow DRA, with the mid-level hump being a good candidate.

The tendencies observed in experiment 1 are confirmed at lower levels (figures II.3 and II.5) : the louder the context, the higher the discrimination thresholds. The pattern of results was still not consistent with DRA despite a reduction of the energy in the stimuli.

The fact that the threshold for the 55-dB SPL standard in the 55-dB SPL context was higher than in the 43-dB SPL context could suggest a flooring effect : discrimination thresholds cannot go under a certain value. The lower levels were supposed to yield higher DLs and thus eliminate the possible flooring effect that would limit performance, but they did not. The increase in threshold could be linked to a reduction of sensation level in the 55-dB SPL context compared to the 43-dB SPL context. It would be possible to consider a limit to human performance, and that there is a limit to intensity discrimination thresholds.

However, no such limit has been characterized. The lowest thresholds for the 55-dB SPL standard in experiment 2 are lower than those observed in experiment 1. Moreover, figure 5 in Florentine et al. (1987) reviews the discrimination thresholds obtained in seven different studies for pure tones at several levels and different frequencies. This figure, replotted in figure II.13, suggests that the thresholds can go below 1 dB at high levels, for various frequencies, including 2 kHz. Finally, five participants had thresholds below 1 dB in experiment 1 for the 75-dB SPL standard, in silence and/or in the soft context.

The smaller difference between the tested levels might have contributed to this absence of change, and no firm conclusion can be drawn. Unfortunately, it is only possible to test lower levels if the mean level of the loud context is also decreased, otherwise the soft standard would be inaudible in the loud context. Some participants from experiment 1 were already excluded of experiment 2 because their detection thresholds in the 55-dB SPL context were too high ; the choice of levels of all the experiments is constrained by the level of the loud context, and the desired level difference between soft and loud stimuli.

Another way to increase the DLs (worsen performance) could be to use a shorter standard duration. However, pilot studies suggested that the participant selection would have been even more difficult ; having the same number of participant would have required a smaller level difference between contexts.

The results of this second experiment thus confirm the previously observed tendencies : the inconsistencies between our hypotheses based on DRA and the results of Experiment 1 are neither due to a flooring effect nor to a the level of the loud standard driving too many fibers in saturation.

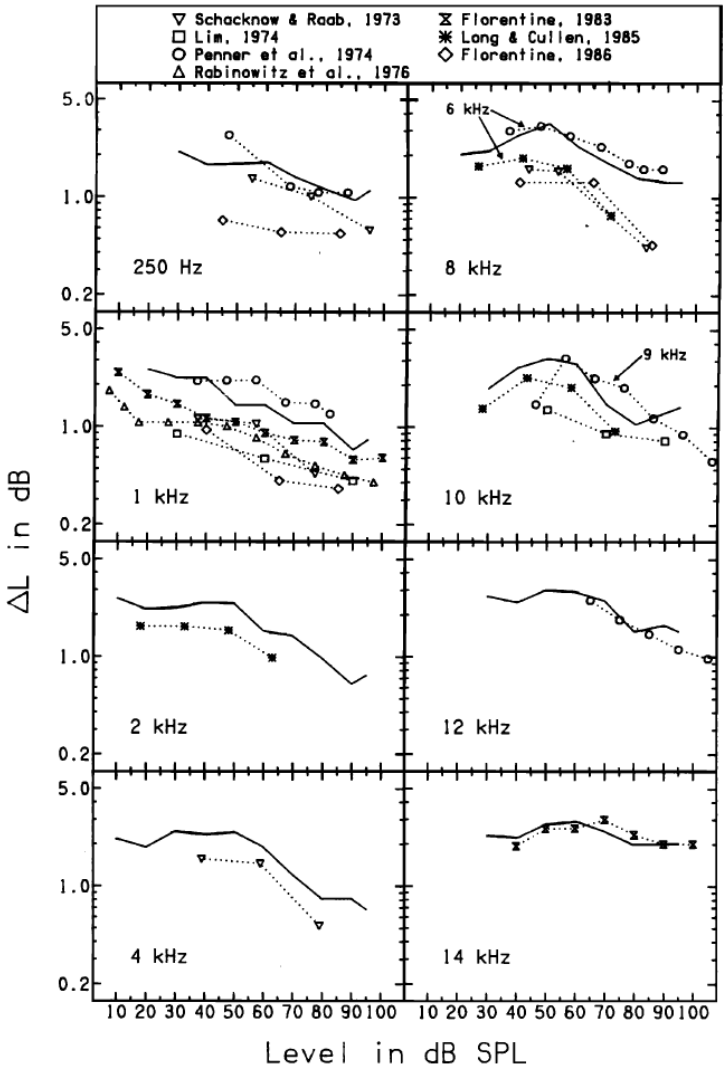


Figure II.13 – Review of discrimination thresholds obtained at difference frequencies. From Florentine et al. (1987).

II.3.4.3 Added notched noise

If the results do not seem to be a consequence of the high levels used with narrowband stimuli, they are apparently not caused either by off-frequency listening.

The results of experiment 3 showed no effect of the notched noise on the discrimination thresholds for the 75-dB SPL standard. The discrimination thresholds were higher in the 75-dB SPL than in the 55-dB SPL context, independently of the notched noise level, and they were not different from those obtained without notched noise. This would suggest either that the low thresholds observed in the soft context compared to the loud context are not due to off-frequency listening, or that the notched noise failed to mask it. This hypothesis is tested in the following.

Moore et al. (1997) proposed a loudness model, with one of the first step being the calculation of the excitation pattern. We used it to compare the excitation patterns of the 75-dB SPL standard and of the 35- and 65-dB SPL notched noise ; the results are plotted in figure II.14.

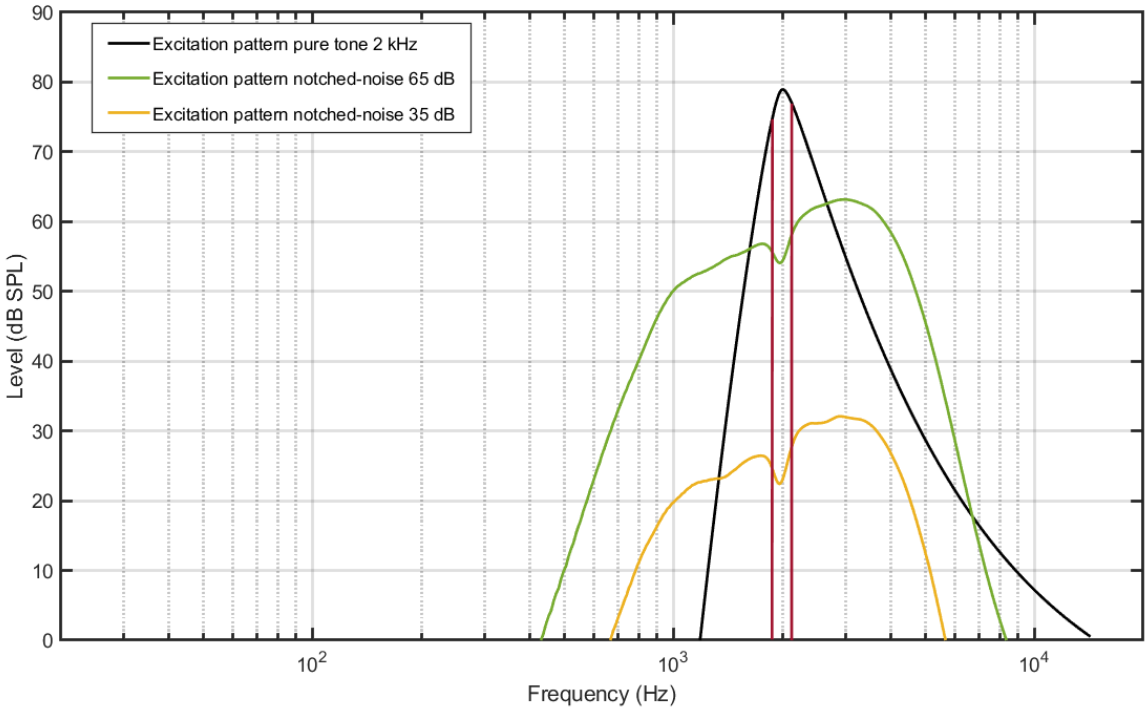


Figure II.14 – Excitation patterns of the 75-dB SPL standard, 35- and 65-dB SPL notched noise used in experiment 3 ; in red, the ERB at 2 kHz. Made with the model from Glasberg and Moore (1990) ; Moore et al. (1997) ; ANSI (2007).

We can infer that every frequency region under the excitation pattern of the notched noise should be masked by the latter and would, therefore, not be used in intensity discrimination. From these patterns, little effect of the 35-dB SPL notched noise on the standard would be expected, but the 65-dB SPL masker should mask enough of the excitation pattern of the standard to increase the discrimination thresholds - even if the whole pattern outside the ERB around 2 kHz is not completely masked. Indded, components above 6 kHz are not masked by the notched-noise. From Killion (1978), the minimum audible pressure from 4 to 10 kHz is around 15 dB, thus components above 7 kHz may be audible or very weak.

In several studies, off-frequency was successfully reduced using notched noise (e.g. Carlyon and Moore 1984 ; Schneider and Parker 1987.

Our intensity discrimination task nevertheless showed no effect of the notched noise, even compared to experiment 1. Plack and Viemeister (1992a,b) suggested that the notched noise had more effect on a standard with a forward or backward masker than just limiting off-frequency listening. They found that adding a notched noise to a discrimination task of pure tones with forward or backward masking reduced the mid-level hump, effectively decreasing the thresholds. Studies by Carlyon and Moore (1984) and Zeng (1998) had similar results. This reduction in discrimination thresholds appeared for standard levels up to 70-80 dB SPL, at different frequencies (from 500 to 6500 Hz). The maskers were either narrowband noise or pure tones, played with 100-ms delays before the standard, with levels ranging from 75 to 94 dB SPL.

We would thus expect a decrease in discrimination thresholds with the addition

of the notched noise for medium (30-40 dB) level-masker level differences, and maybe even slightly at high levels.

Moreover, studies of tone detection in noise (Carlyon 1987) and increment detection in continuous stimuli or intensity discrimination (Carlyon and Moore 1986) showed that the notched noise could reduce thresholds if the signal was preceded by a narrowband noise masker centered on the standard frequency. They proposed that the notched noise induced adaptation to the masker.

The addition of the continuous notched noise could also be reducing confusion by masking part of the level fluctuations of the context and broadening its bandwidth (Neff 1986).

These two contradicting effects suggest that the effect of the notched noise could be opposite to expectations because of the context.

Furthermore, any effect of the notched noise starts to disappear around 70-80 dB SPL (Carlyon and Moore 1984 ; Plack and Viemeister 1992a,b ; Zeng 1998), which is exactly the level range our loud standard is in.

It is also possible that the off-frequency components are suppressed by the context itself, because of time-frequency masking (Necciari et al. 2016).

II.3.4.4 Masking by the context

The addition of gaps (experiment 4) was effective in reducing the discrimination thresholds in the loud context. As in experiment 1, the discrimination thresholds are lower in the soft context than in the loud context for both standard levels, but the discrimination thresholds were lower for both standards in the loud context in experiment 4 than in experiment 1 (figure II.9).

This reduction of thresholds in the loud context is consistent with previous work with gaps. For example, Carlyon (1989) showed that an increase in the delay between a narrowband masker and a pure tone standard caused a decrease of the discrimination thresholds. They are thus less affected by masking (confusion/mid-level hump), as suggested by Neff (1985, 1986) and Zeng and Turner (1992), but not enough to observe DRA.

According to the time constants measured by Backus and Guinan (2006) in human (increase in level : 272 ms, decrease : 159 ms), the MOC could have started to decay in the silent gap before the start of the standard in experiment 4 ; the firing rate of the neurons could thus have increased slightly and the discrimination thresholds would be lower than without the gap, if that is the case. This is reflected in the modeled rate-level functions (figure II.10a) : the shift is smaller with than without the gap (experiment 1).

Moreover, the effects of the modification of the context from monaural in experiment 1 to diotic in experiment 5 are also noticeable. With the diotic context, an increase in discrimination thresholds between soft and loud contexts is still observed, but for the soft standard only ; whereas for the loud standard no significant difference is found between contexts. Compared to experiment 1, a significant decrease of the discrimination thresholds in experiment 5 appears in the loud context for both standard levels (table II.15).

This is consistent with the effects of a binaural or lateralized masker on a monaural standard, as reported by Oberfeld et al. (2012, 2014) and Roverud and Strickland (2015a) in intensity discrimination, and Hirsh (1948) ; Punch and Carhart (1973) ; Moore and Glasberg (1982) and Neff (1985) for detection thresholds.

These last experiments indicate that there is some masking and confusion caused by the context. The modifications (silent gaps and diotic context) brought small improvements in the discrimination thresholds in the loud context, without modifying those in the soft context for both standard levels and experiments. These changes are interpreted as releases from masking ; greater releases (if possible) could even change the trends, if they could be obtained without canceling the potential effect of DRA on the task.

The increase in discrimination threshold obtained for the 55-dB SPL standard between silence (experiment 1) and loud context for both experiments 4 and 5, or between contexts for these two experiments, is still greater than that expected based on differences in sensation level. Maybe DRA is the main cause of this greater difference : since the rate-level function is supposedly shifted towards higher levels in the 75-dB SPL context, the DL might be larger with a nearly-continuous context than with short precursors with gaps before the standard.

II.4 General discussion

To conclude, our results were not fully consistent with our predictions based on DRA. More specifically, the results for the loud standard did not match our initial hypothesis. The context level and temporal proximity with the standard generate some masking and confusion. The discrimination thresholds thus increase with the mean level of the context, and DRA is not visible in the results.

Several studies, including Viemeister (1988) and Colburn et al. (2003), have warned against the direct derivation of intensity discrimination performance from discharge rate only from animals, and Weber's law itself disagrees with the response function of auditory neurons ; there are still missing links between intensity encoding and perception.

Other encodings of intensity could be considered : for example, spread of excitation and frequency channel combinations are probably important here. Temporal aspects of the signals could also play a role. It is however difficult to know how to incorporate these aspects to the physiological studies of DRA in anesthetized animals, that give information for separate individual fibers.

It is also possible that the choice of the model we used to simulate the rate-level functions of auditory nerve fibers (Grange et al. 2022) was not the best ; its MOC and AR effects have great magnitudes. It would be necessary to combine the different fiber types and frequency bands to see the greater picture, but choosing how to do that is tricky. Other numerical models of the periphery could be used and maybe predict the discrimination thresholds better.

We could have tried to compare different standard levels in one context level, with the hypothesis that the standard of the same level as the mean of the context would have lower discrimination thresholds than other standard levels, because it fell into

the steepest part of the rate-level function. However, the near-miss to Weber's law suggests that higher standard levels will have lower intensity discrimination thresholds. The use of pure tone standards thus prevented this comparison ; we had to compare the same standard in different context mean levels.

It is possible to conclude from the results of the control experiments that the context partially masks the standards, and that there is some confusion between the context and the standard for both standard levels. Gallun and Hafter (2006) looked at amplitude modulation as similar to intensity increment detection and hypothesized that both could use the same mechanisms ; it is then possible that the fluctuations of the context are perceived as increments and thus confuse the listeners. Colburn et al. (2003) found that approaching saturation of auditory nerve fiber discharge rate itself has less effect on the information than the fluctuations in their responses. Indeed, the fluctuations in response tend to increase with discharge rate ; moreover, since the level of the context is fluctuating, the fluctuations in discharge rate in response to the standard must be even more important depending on the directly preceding level.

It is also possible that observing DRA in intensity discrimination tasks is not possible. Maybe DRA just makes us slightly better at perceiving intensity differences in noise, and performances would be terrible without it, but it is too tied to other phenomena to be distinguished in the present psychophysical tasks.

The effects of the continuous narrowband noise context on the discrimination thresholds of pure tones were more complex than anticipated, and the control experiments did not answer all of our questions.

One feature that changed in the present stimuli compared to the physiological studies on DRA is the definition of the most probable levels. Instead of the high probability region of Dean et al. (2005), we used a Gaussian distribution for the levels of the noise segments. This difference could lead to a behavioral change of the neurons. For example, Rabinowitz et al. (2011) described an adaptation of the neurons to the variety of encountered levels, that they called contrast gain control; a high contrast meant that many different levels were played while a low contrast indicated fewer differences. This adaptation either expanded or compressed the dynamic range depending on the contrast context, resulting in an optimization of the representation of stimuli level changes.

This idea was investigated by Lohse et al. (2020) who observed changes in the rate-level function's slope from the inferior colliculus up with broadband noise and harmonic complexes. They compared their auditory gain contrast, an adaptation to the level fluctuation range in mice auditory system, to human psychophysical data with the same stimuli, and found lower intensity discrimination thresholds for stimuli with little contrast - the condition that displayed the steepest slope of the rate-level function. Together with dynamic range adaptation, this change in slope could lead to a kind of stimuli normalization dependent on the context. However, our study with pure tones and narrowband noise failed to demonstrate any effect of DRA, and could not give any information on contrast gain control. Since we use a Gaussian distribution for the level of the context segments, we have a relatively low contrast which should induce a steepening of the response function.

Finding a way to study intensity discrimination with a similar paradigm to ours

but with broadband signals seems important. The studies who first found DRA at several sites of the auditory system of animals used broadband signals while we used pure tones and narrowband noise. Targeting a specific region of the auditory system and looking at behavioral results might be too tricky, while using wide-band signals could take advantage of the entire neuron population response.

Chapter III

Intensity discrimination using broadband signals

III.1 Introduction

In the previous chapter, we tried to uncover a perceptual correlate of DRA using pure tones in an intensity discrimination task. This chapter aims to extend these experiments using broadband signals. The hypotheses, material and protocol were very similar.

Most of the physiological studies described in the first chapter used broadband signals and measured the evolution of the responses of neurons of different parts of the auditory system of animals, in response to sounds with different level distributions. Only sometimes did they use pure tones at the characteristic frequency of the observed neuron. In their first study in the auditory nerve, Wen et al. (2009) found that, for the same SPL mean HPR levels, the shift in rate-level functions attributed to DRA was less important when using broadband noise than when using pure tones at the characteristic frequencies of the tested neurons. They however argued that the signals, which had the same overall level, had different mean levels relative to the neurons' threshold. Hence, the difference was probably not due to the properties of DRA, but more likely to the differences in overall tested level. The broadband signals used in this chapter had lower levels per ERB than the pure tones and narrowband noise context used in the previous chapter. The rate-level functions should therefore be less saturated, and the excitation broader, getting the experiment closer to the configuration of the animal studies. Moreover, there should not be any influence of off-frequency listening.

Similar to Chapter II, we wanted to use no gaps between context and test signals to maximize our chances of observing an effect due to DRA. Therefore, once again, we had to select different stimuli for the context and the test signals. Our constraint was that both context and test signals should have the same frequency content but should differ in timbre. For the context, we used broadband white noise, as in physiological studies. For the test signals, we decided to use pulse-spreading harmonic complexes (PSHC), as described in Hilkhuysen and Macherey (2014).

III.1.1 Pulse-Spreading Harmonic Complexes

A PSHC is a harmonic complex sound with a specific phase relationship between harmonics that provides a way to decouple the fundamental frequency and the envelope repetition rate of the waveform. It is, for example, possible to generate a 1-Hz f_0 complex with an envelope repeating at 100 pulses per second.

The harmonics are divided into harmonic subgroups. An illustrated example of this

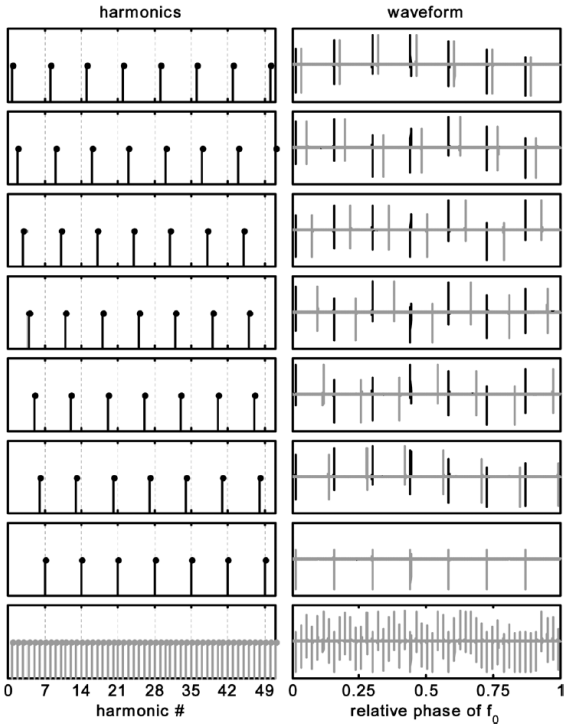


Figure III.1 – Example of harmonic distribution for a 7th-order PSHC. Each panel on the left displays one subgroup of harmonics and on the right the resulting waveform for this subgroup, with a certain phase shift applied (in grey), and the full signal at the bottom. Replotted from figure 1 from Hilkhuysen and Macherey (2014).

division is presented in figure III.1 for a 7th-order PSHC ($k = 7$) : the first harmonic in the selected band falls into the first subgroup, the second in the second subgroup, the third and fourth respectively in the third and fourth subgroups, and so on until the 7th group ; the eighth harmonic is then assigned to the first subgroup, the ninth to the second subgroup, and so on. The harmonics of each subgroup are summed together with the same starting phase, creating a pulse train that repeats at $k \times f_0$ Hz. In this specific example, this yields 7 pulse trains. These seven signals are then added with a temporal delay so that the pulses are evenly spread across the period of the stimulus, effectively yielding an envelope rate of $k^2 \times f_0$ pulses per second.

A PSHC is defined by its fundamental frequency f_0 (generally low, e.g. 1 or 2 Hz), its center frequency $f_c = \sqrt{f_{min} \times f_{max}}$ (with f_{min} and f_{max} respectively the lowest and highest frequencies of the bandwidth), its order k, and its pulse rate. k defines the number of harmonic subgroups within the stimulus.

A specificity of the PSHC is that the pulse rate (which depends on the PSHC order and the f_0) can be optimized to minimize the fluctuations at the output of a given auditory filter (Hilkhuysen and Macherey 2014). Studies by Hilkhuysen and Macherey (2014) and Mesnildrey et al. (2016) provide the necessary equations for this "optimal" pulse rate as a function of the center frequency f_c of the stimulus: $rate = 37 + 151 \times \frac{f_c}{1000} + 0.17 \times (\frac{f_c}{1000})^2$, and order $k = \sqrt{\frac{rate}{f_0}}$.

III.2 Experiment 6 : Discrimination with broadband signals

Experiment 6 consisted in an intensity discrimination task using broadband PSHC as signals, and a nearly continuous broadband noise context presented between the test intervals.

III.2.1 Material and methods

III.2.1.1 Participants

Sixteen participants completed the experiment. Seven were female ; their ages ranged from 19 to 67 years old (mean : 33.9, s.d. : 13.5). They all had absolute detection thresholds below or equal to 20 dB HL for pure tone frequencies ranging from 125 to 8000 Hz on both ears, and seven of them had no previous experience in psychophysical tasks.

This research was approved by ethics committee "Comité de protection des personnes Sud-Ouest et Outre Mer III" (2021-A01920-41)". Participants were paid according to the number of hours of testing.

As in Chapter II (experiment 1), they completed a selection task described below to ensure that the presentation levels of all test signals were above 10 dB SL in the discrimination task. A training session was also performed, to get the participants accustomed to the stimuli and to the different conditions of the discrimination task. The training consisted of short versions of all experimental conditions.

III.2.1.2 Material

The experiments were completed in a double-walled soundproof booth with a Sennheiser HDA 200 earpiece connected to a RME Fireface UCX sound card. Calibration was performed in the same way as for the pure tones and narrowband signals (Chapter II), with a B&K 4153 artificial ear, a GRAS 26AK preamplifier, and a GRAS 12AR conditioner. All sounds were presented to the left ear.

The procedure was programmed using custom Matlab code (2022b). Statistical analysis was done with the R software (4.4.2).

Test intervals were indicated by two red LED lights placed under the screen, and feedback was provided by a button color change on screen after answering.

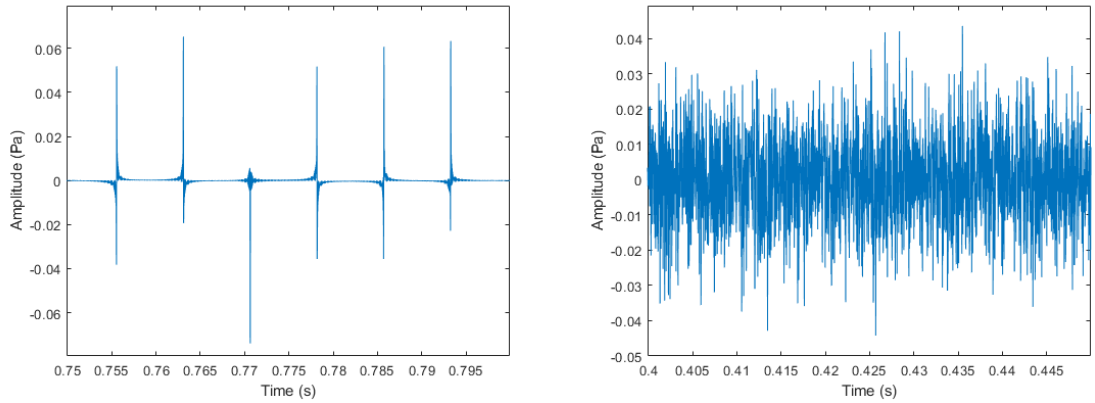
III.2.1.3 Stimuli

Intensity discrimination thresholds were measured for pulse-spreading harmonic complexes presented in a broadband noise context. The bandwidth of the test signals ranged from 20 to 20 000 Hz. This bandwidth was chosen to be as wide as possible to match the physiological studies and the classically admitted human frequency range.

The standard used for the task was a 50-ms PSHC (see III.1.1) with a fundamental frequency f_0 of 2 Hz, and a central frequency $f_c = \sqrt{20 \times 20000} = 632.4$ Hz ; one waveform example is presented in figure III.2a. The pulse rate of 132.6 pps was calculated to minimize the intrinsic modulations at the center frequency of the band

(Mesnildrey et al. 2016). The PSHC order was thus $k = 8$. The output of seven simulated auditory filters between 20 and 10 500 Hz is displayed in figure III.3a.

The context was a broadband noise. It had the same bandwidth as the signal and was constructed similarly to the PSHC (with all the harmonics of 2 Hz between 20 and 20000 Hz), but all the sinusoids were summed in random phase. That way, both signals had the same long-term spectral content. As in the physiological studies, the context level changed every 50 ms - every segment -, and followed a Gaussian distribution centered on the chosen mean level (3-dB standard deviation, with a limit in the possible levels defined as the mean ± 10 dB). A 50-ms example is presented in figure III.2b. For comparison with the auditory filter's outputs of the PSHC, the corresponding outputs are plotted for an exemplar of the noise in figure III.3b.



(a) PSHC exemple **(b)** Broadband noise example

Figure III.2 – Example waveforms for two 50-ms excerpts of the signals used in this chapter.

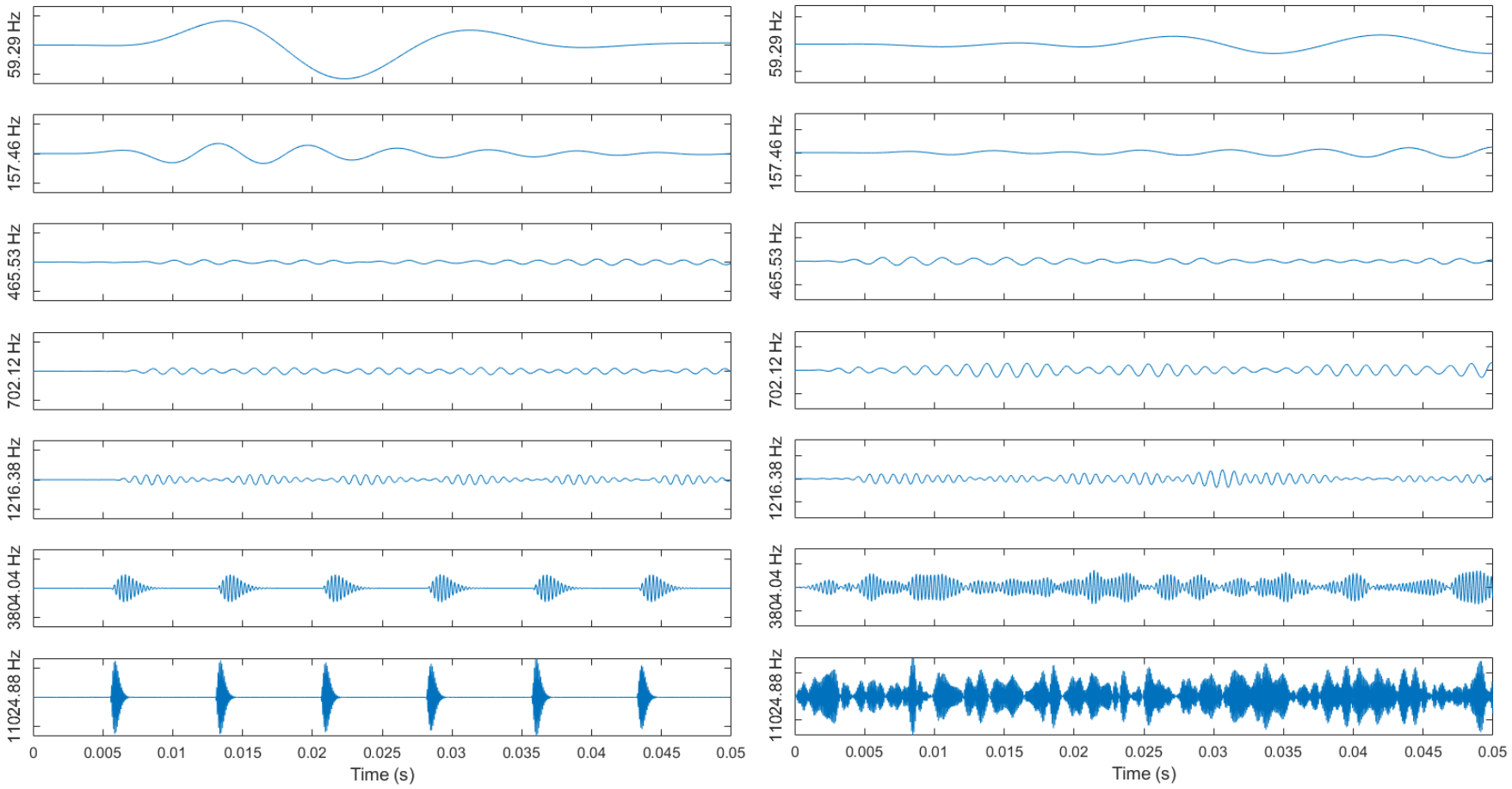
The test level referred to as loud was chosen to be 70 dB SPL. We initially programmed the experiment with a loud level of 75 dB SPL context (to match the levels of Chapter II) but we informally found the levels were uncomfortably loud. The selected level for the soft level was 50 dB SPL (both were used for context and standard). Hence, the level difference between the two tested levels was still 20 dB, as in experiment 1.

The noise context was pre-generated as one hundred 1-s sound files. For each noise segment (i.e. each 50-ms time interval), a portion of 50 ms of one of these files was randomly extracted and played.

The target signal was created by adding an identical PSHC with a variable level to the standard. The difference between target and standard level was calculated with the following formula :

$$DL = 10 \times \log_{10}\left(1 + \frac{\Delta I}{I}\right) \tag{III.1}$$

where I is the standard intensity and ΔI is the difference in intensity between the standard + added signal and the standard.



(a) PSHC **(b)** Broadband noise

Figure III.3 – Ouputs of seven auditory filters for two 50-ms excerpts of the signals used in this chapter. These are the output of gammatone filters from Slaney (1998).

III.2.1.4 Procedure

Intensity discrimination thresholds were measured in a 2 intervals, 2 alternatives forced choice, two down - one up adaptive staircase (Levitt 1971). Participants had to select the louder sound (target). The standard and target were presented in random order during the two test intervals, separated by a gap of 500 ms. For the first trial, the added signal was 14 dB above the standard level, so the target level was 15.6 dB above the level of the standard. Inter-trial interval could vary from 700 to 1500 ms after the participant answer to avoid anticipation.

The added signal level decreased in 4-dB steps for the first two reversals of the task, then in 2-dB steps for the rest of the experiment. Thresholds for the two standard levels were obtained simultaneously for each context level, using an interleaved procedure. Both tracks went on until at least 12 reversals were obtained for both (6 for the training sessions). One of the tracks could thus total more than 12, as the probability of getting either was always 50 %. Each threshold was the average of the last six reversals of the adaptive procedure, regardless of the total number of reversals.

There were six tested conditions in total (2 standard levels and 3 context conditions). The discrimination threshold for each participant and condition was the average of four repetitions of the measurement (or more, until the standard deviation of the thresholds was below 3 dB). The context was presented continuously and only stopped during the presentation of the standards, with 5-ms rise and fall times around the test intervals. To limit possible short-term fluctuation effects that could be caused by different preceding levels between intervals within a trial, the context before and after both test intervals was the same, so 500 ms of context were identical before the first interval, after the second one and in-between. Thresholds were also acquired without context.

III.2.1.5 Preliminary experiment

To select the participants for the discrimination task, a detection task of the test signal was performed with the same signals as in the main experiment. The signal was thus a 50-ms, 2-Hz f0, 8th-order PSHC with a bandwidth ranging from 20 Hz to 20 kHz. The detection thresholds were obtained in silence, in the soft context (50 dB SPL), and in the loud context (70 dB SPL). The context was a broadband noise (20-20 000 Hz), presented continuously except during the test intervals.

The protocol was a 2I-2AFC, 2 down - 1 up adaptive staircase with 12 reversals, with the standard in one of the intervals and silence in the other, separated by 500 ms : the participants had to select the interval in which they heard the signal. There were five repetitions of the three context conditions, all done during the same test session.

The selection criterion was to have a detection threshold in the loud context lower than 40 dB SPL, so that the soft standard (50 dB SPL) in the discrimination task would be at least 10 dB above threshold in the loud context. This was a way to ensure that this configuration would actually be feasible as a discrimination task ; if some signals were not perceived in the loud context, then the nature of the task would change, and the measured threshold would not reflect intensity discrimination performance.

All participants were selected [1] ; their thresholds are presented in table III.1. The

1. They were preselected from other psychophysical experiments

Context	Mean (dB SPL)	Standard error (dB)
silence	8.78	0.79
50 dB SPL	24.91	0.73
70 dB SPL	33.96	0.67

Table III.1 – Mean detection thresholds for the 16 selected participants for this study. They increased with the mean level of the context.

lowest detection thresholds were obtained in silence. Thresholds for the soft context were higher than in silence, and the loud context yielded the highest threshold.

III.2.1.6 Auditory nerve model

The auditory nerve model previously described in the Material and methods section of experiment 1 in Chapter II was also used. It is described in Grange et al. (2022) and based on the MAP model (Meddis 2006). It simulates the response of the auditory nerve fiber population, and provides output from the acoustic reflex and medial olivocochlear (MOC) efferents.

The stimuli were similar to those used for pure tones : 700 ms of context, 50-ms standard, 500 ms of the same context, with standard levels ranging from 20 to 100 dB SPL. The rate-level functions depending on the context were computed. The slopes of the fitted sigmoid functions were also computed with linear regressions of three consecutive points.

As in the previous chapter, the sensitivity index was computed from the model in an attempt to correct the slopes with the variability of the firing rates. The obtained values can be found in appendix B.5.

III.2.2 Results

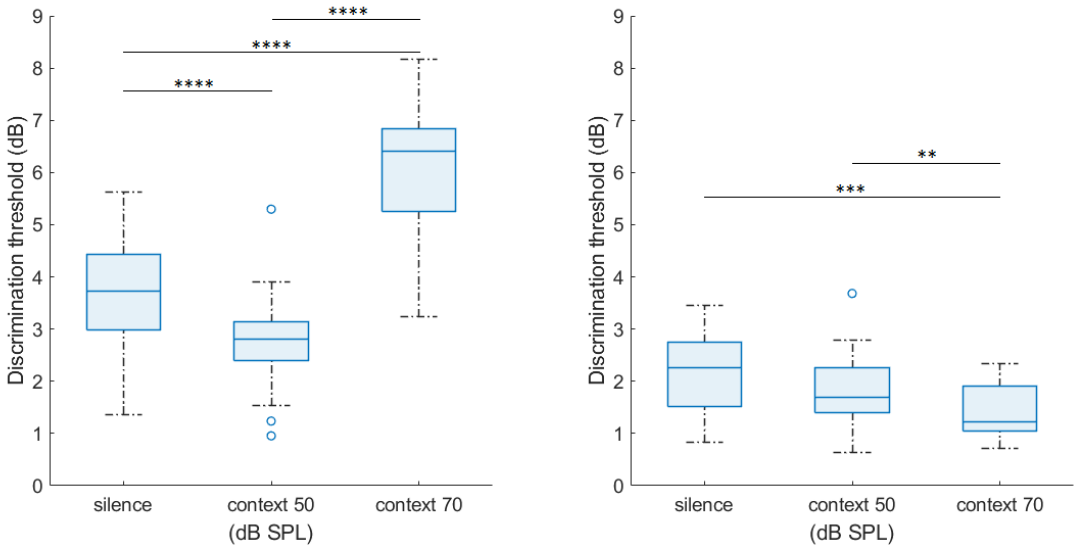
Figure III.4 presents the obtained thresholds, along with table III.2.

Standard level	Context level	Mean (dB)	Standard error (dB)
50 dB SPL	silence	3.73	0.29
	50 dB SPL	2.77	0.26
	70 dB SPL	5.99	0.32
70 dB SPL	silence	2.17	0.20
	50 dB SPL	1.88	0.19
	70 dB SPL	1.38	0.13

Table III.2 – Averaged results for experiment 6. The mean discrimination thresholds are lowest when the context mean level is identical to the standard level.

A two-way repeated-measures ANOVA was run on the data, with factors standard level (2 levels : 50 and 70 dB SPL), and context (3 levels : silence, 50 dB SPL and 70 dB SPL). The analysis showed highly significant effects of both factors and of their interaction (table III.3).

Furthermore, pairwise comparisons (table III.4) indicate that the discrimination threshold in the loud context was significantly different than for the other contexts.



(a) 50 dB SPL standard **(b)** 70 dB SPL standard

Figure III.4 – Results of experiment 6 for all participants. Discrimination thresholds are plotted against context condition : the first and third quartile are limiting the blue box, and the median is presented as a blue bar within it. Whiskers include all non-outlier values that do not fall into the box ; blue circles are outliers (between 1.5 and 3 times the interquartile value away from the first or third quartiles) and red circles extreme values (more than 3 times the interquartile value away from the first or third quartiles). * : significant difference between conditions, see table III.4.

Effect	dof	dof err.	F	p	significance
standard level	1	15	296.399	2.73e-11	****
context level	2	30	51.833	1.85e-10	****
standard level × context level	2	30	79.971	9.49e-13	****

Table III.3 – Repeated measures ANOVA on standard level and context mean level. The investigated parameters and their interaction all have significant influences on the discrimination thresholds.

There was a clear decrease in threshold when the context mean level was the same as the standard for both standard levels (threshold difference compared to the other context level : -3.22 dB for the soft standard, -0.50 dB for the loud standard ; compared to silence : -0.96 dB for the soft standard and -0.79 dB for the loud standard). Moreover the thresholds in the loud context and in silence for the soft standard were also significantly different.

To summarize, these results are completely in line with the expected effect of DRA.

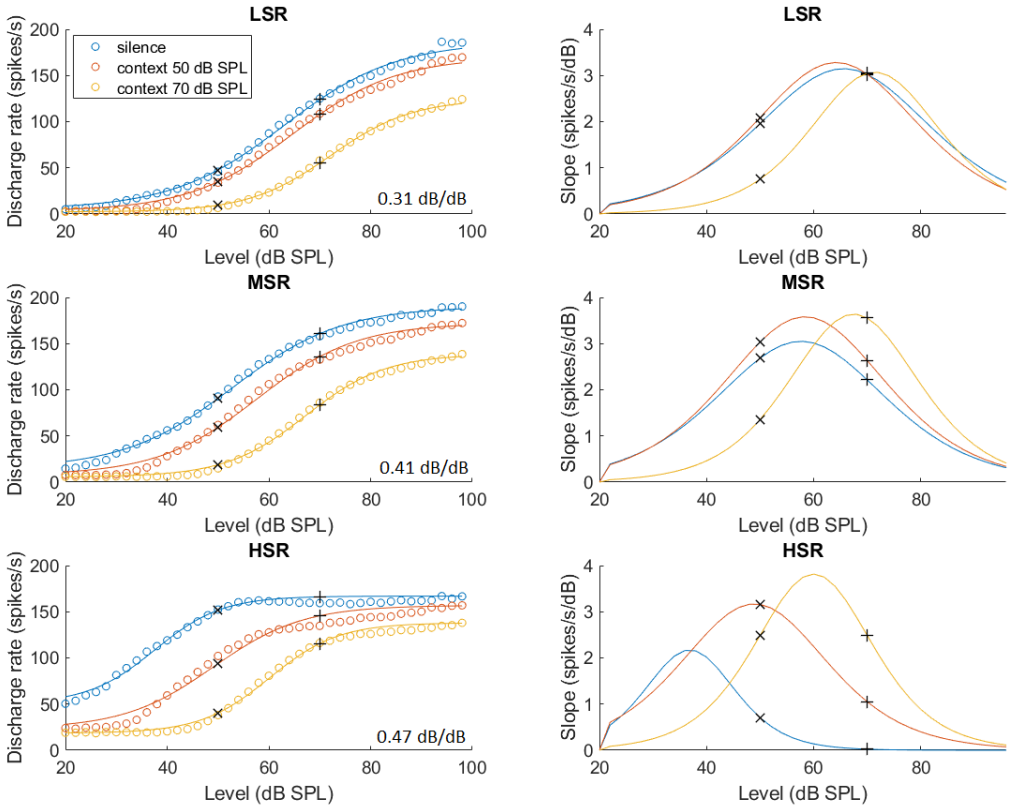
III.2.2.1 Modelling

The rate-level functions obtained with the model are plotted in figure III.5a, along with the slopes for the corresponding functions in figure III.5b. The rate-level functions are shown for the ERB centered on 1 kHz (as Grange et al. (2022) did for their

broadband signal).

Standard level	Context 1	Context 2	t	p adjust	significance
50 dB SPL	silence	50 dB SPL	6.72	2.06e-5	****
	silence	70 dB SPL	-7.28	8.04e-6	****
	50 dB SPL	70 dB SPL	-12.5	7.11e-9	****
70 dB SPL	silence	50 dB SPL	1.71	0.324	ns
	silence	70 dB SPL	4.76	7.59e-4	***
	50 dB SPL	70 dB SPL	3.99	0.004	**

Table III.4 – Pairwise t-test for each standard and context level. The context matching the standard level yields lower discrimination thresholds than the other context conditions for both standard levels ; only the silence and soft context for the loud standard are not significantly different.



(a) Rate-level functions. **(b)** Slopes

Figure III.5 – Rate-level functions with shift between contexts (a) and slopes (b) obtained with the MAP model as configured by Grange et al. (2022) for HSR, MSR and LSR fibers in the ERB tuned around 1 kHz, for the PSHC and broadband context conditions used in experiment 6, averaged over ten simulations.

The change in average discharge rate ($L50$, point at 50 % between the minimum and maximum of the rate-level function) between the soft and loud context, divided by the level difference between contexts, represents the horizontal shift of the rate-level

Standard level	Fiber type	Context type	Slope (spikes/s/dB)
50 dB SPL	LSR	silence	2.73
		50 dB SPL	2.58
		70 dB SPL	1.13
	MSR	silence	4.11
		50 dB SPL	3.69
		70 dB SPL	1.97
	HSR	silence	2.38
		50 dB SPL	3.63
		70 dB SPL	3.08
70 dB SPL	LSR	silence	3.80
		50 dB SPL	3.80
		70 dB SPL	3.88
	MSR	silence	2.16
		50 dB SPL	2.91
		70 dB SPL	4.01
	HSR	silence	0.242
		50 dB SPL	1.25
		70 dB SPL	2.42

Table III.5 – Slopes around both standard levels for each context and fiber types in figure III.5a : three-points linear regression around the standard level.

function, indicated in figure III.5a. The shift is lowest for the LSR fibers, and highest for the HSR (respectively LSR = 0.31, MSR = 0.41, HSR = 0.47 dB/dB). The shift values obtained for all fiber types are close to those found for broadband noise in Grange et al. (2022) and displayed in figure II.2 in the previous chapter.

The slopes are shown in figure III.5b. Their values computed at the tested standard levels, presented in table III.5, are steeper in silence for the LSR and MSR fibers for the 50-dB SPL standard, and in the 50-dB SPL context for the HSR fibers. For the 70-dB SPL standard, they are always steeper in the loud context. The differences in slopes between contexts are smaller for the loud standard than for the soft standard for LSR and MSR fibers, consistently with the differences in discrimination thresholds between contexts. For the 50-dB SPL standard, the slope for the HSR fibers is consistent with DRA, while it is the case for all fibers for the 70-dB SPL standard.

III.2.3 Discussion

First of all, the results are in agreement with our initial hypothesis on the effect of DRA. For both standard levels, the discrimination thresholds were lower when the context mean level matched the standard level than when it did not.

The discrimination thresholds for the soft standard were qualitatively similar to those observed in Chapter II, showing significantly higher thresholds in the loud context than in the soft context or than in silence.

More surprisingly, for these broadband signals, the discrimination threshold in the soft context was even lower than in silence, whereas there was no difference between

these conditions for pure tones.

For the loud standard, the intensity discrimination threshold was lower in the loud context than in silence and than in the soft context. This result is opposite to that found with pure tones. This shows that it is possible to go *below* the discrimination threshold performance in silence, with soft or loud standards and the corresponding context.

Therefore, we show that embedding the test signals in a context can improve intensity discrimination performance compared to when presenting them in silence. This is somewhat surprising given we know from the selection task that the context increases detection threshold. So this means that the test signals were presented at a lower sensation level when embedded in the context compared to when presented in silence. We may have expected this decreased sensation level to increase threshold rather than decreasing them Jesteadt et al. (1977) ; Florentine and Buus (1981).

Overall, our results are consistent with DRA.

III.2.3.1 Comparison with the model

The MAP model as configured and run by Grange et al. (2022) allowed us to simulate the rate-level function of auditory neurons in response to audio stimuli. Our hypothesis on DRA states that, if the slope around the level of the standard is steep, the discrimination thresholds will be lower than if the slope is shallower. The model enables the identification of the steepest rate level functions for both standard levels, depending on the context condition.

We ran simulations with the model for our stimuli ; the obtained rate-level functions for the ERB at 1 kHz, presented in figure III.5a, resemble those obtained by Grange et al. (2022) (first column of their figure 6, reproduced in figure II.2 in the previous chapter). The slopes of the rate-level functions for the different fiber types predict the 50-dB SPL standard to show better results in the soft context (HSR) or in silence (MSR-LSR), but worse in the loud context. For the 70-dB SPL standard, all fibers have steeper slopes in the loud context, even if they are almost saturated for the HSR. These results are thus in agreement with our hypothesis and roughly with the intensity discrimination thresholds, which was not the case for pure tones in the previous chapter.

Values for the shift in dynamic range between context mean levels are smaller for the broadband signals than for the narrowband signals (for LSR, MSR and HSR, respectively : for pure tones, 0.50, 0.53 and 0.59 dB/dB ; 0.31, 0.41 and 0.47 dB/dB for PSHC). They are still consistent with the values of rate-level function shifts obtained by Grange et al. (2022). Wen et al. (2009) found a generally larger shift for pure tones than broadband noise ; they suggested that it was not due to DRA itself, but rather to differences in level relative to the individual neuron threshold, which was higher for noise than for tones.

If we compare the slope of the different simulations obtained with the stimuli used in this chapter (figure III.5a), to those obtained with the 2-kHz pure tones and narrowband noise stimuli used in the previous chapter (figure II.4a), discrepancies appear. For the pure tones, the rate-level function for each context was mostly saturated at 75 dB SPL (loud context) in the response of neurons tuned at 2 kHz, except for the LSR fibers in the 75-dB SPL context (table II.3) : however our hypothesis on DRA implies that the

discrimination thresholds should be lower in this context level, but they were not. For the broadband PSHC, the slopes were steeper for the 50-dB SPL standard in the 50-dB SPL context for the HSR fibers only (table III.5), and the discrimination thresholds were lower in this condition. It is then possible that the HSR fibers matter more than the LSR fibers in predicting discrimination thresholds ; this would be consistent with their much greater number in the auditory system.

III.2.3.2 Intensity discrimination thresholds of PSHC in silence

The discrimination thresholds for PSHC in silence obtained in this experiment are interesting. For both standard levels, they were higher than those we found for pure tones in experiment 1 (see II.1).

Moore and Raab (1975) tested discrimination thresholds for 10- and 250-ms bursts of broadband noise ; their results, presented as $10 * log_{10}(\Delta I/I)$, are translated into $10 * log_{10}(1 + \Delta I/I)$ for comparison here. They found values of DL for two participants at 31 and 71 dB SPL of respectively 1.35 and 1.83 dB for the 10-ms bursts, and 0.75 and 0.77 dB for the 250-ms burst. Viemeister (1974) obtained intensity discrimination thresholds for broadband and narrowband noises in quiet. They found relatively stable DLs for a broadband noise (100-15000 Hz) for levels ranging from 37 to 87 dB SPL, with values between 1 and 0.8 dB. The DLs for their narrowband noise (650-2000 Hz) however slightly decreased with increasing level : from 1.5 dB at 46 dB SPL to 0.8 at 76 dB SPL. For the closest values to 50-55 dB SPL, they thus had a lower discrimination threshold in the broadband noise, while the values were similar at 70-75 dB SPL. The DLs for our broadband PSHC were thus not expected to show a decrease with an increase in standard level, contrary to narrowband signals.

The discrimination thresholds values of these two studies are overall much lower than what we found. This is not due to a change in duration of the standard, since our DLs were higher than for a 10-ms noise burst (Moore and Raab 1975). In the present study however, we had much more participants with little training, and an interleaved procedure, which have been shown to increase discrimination thresholds. These differences could also explain the higher discrimination thresholds in silence in this experiment than in the literature.

We have no explanation as to why the discrimination thresholds in silence decreased with increases in standard level in the present experiment, except the possible bias towards high levels of the interleaved procedure (Oxenham and Buus 2000 ; Pienkowski and Hagerman 2009). The near-miss to Weber's law is supposedly absent in broadband signal because there is no spread of excitation (McGill and Goldberg 1968), and it is eliminated for pure tones by presenting broadband noise maskers (Viemeister 1974 ; Widin et al. 1986).

Moreover, the PSHC standard is a broadband signal, so its discrimination thresholds in silence should not be affected by off-frequency listening. However near-miss implies that discrimination thresholds for noise are supposed to be higher than for pure tones at high levels (McGill and Goldberg 1968 ; Moore and Raab 1974), which was the case for the loud standard in our study.

III.2.3.3 Maskers, mid-level hump and bandwidth

The discrimination thresholds in context agree with our hypothesis based on DRA for both standard levels.

The difference in discrimination thresholds between context conditions were lower in the present experiment than those measured with pure tones. Indeed, for the soft standard in both context mean levels, the difference in discrimination threshold was lower for PSHC (3.22 dB) than for a 2-kHz pure tone (7.10 dB), and similarly for the loud standard (-0.80 dB for PSHC, +0.80 dB for pure tones). This decrease in the difference could be caused by the reduction of the mid-level hump and/or confusion.

Experiment 6 PSHC in broadband noise 16 participants			Experiment 1 2-kHz pure tones in narrow-band noise 10 participants		
Context	Mean (dB SPL)	Standard error (dB)	Context	Mean (dB SPL)	Standard error (dB)
silence	8.78	0.79	silence	3.78	1.23
50 dB SPL	24.91	0.73	55 dB SPL	32.96	1.46
70 dB SPL	33.96	0.67	75 dB SPL	40.76	0.87

Table III.6 – Mean detection thresholds and standard error for experiments 6 and 1.

Florentine et al. (1987) found decreases in thresholds for pure tones of 1 dB between standard levels of 10 and 55 dB SL. Between the loud context and silence, the sensation level of the soft standard increased from 16 to 41 dB SL, and the discrimination threshold decreased by 2.3 dB : this is much lower than the difference between silence and loud context for the pure tones reported in Chapter II for a similar change in sensation level (7.7 dB).

Since the standard had a very specific timbre in experiment 6, the confusion effects found with pure tones may be greatly reduced. Neff (1985) defined confusion as the inability to distinguish the signal from the masker, and these effects seem particularly strong for pure tones in narrowband noises (Neff 1986).

It is thus possible that along with confusion, the mid-level hump was reduced, especially for the 50-dB SPL standard in the 70-dB SPL standard for which it should be strongest (Zeng and Turner 1992) ; the remaining difference with Florentine et al. (1987) might be due to partial masking, the interleaved protocol we used, or the type of signal.

To our knowledge, very few studies reported lower discrimination thresholds in the presence of a masker compared to when presented in silence. Oberfeld (2008) found lower discrimination thresholds for a 1-kHz, 30-ms standard at 25 dB SPL when playing a forward masker identical to the standard at 10 dB SPL with a 100-ms delay. For standards at 55 and 85 dB SPL however, they only found higher discrimination thresholds with maskers than in silence. Their results for the 55- and 85-dB SPL standards are more in line with the results from Chapter II that used pure tones. For the 25-dB SPL standard, they put forward a "cueing effect" as presented by Moore and Glasberg (1982) as a possible explanation for the decrease in threshold, but the link is unclear : the study by Moore and Glasberg (1982) was on the effect of

masker lateralization and the addition of an auditory cue on tone detection (that would reduce confusion), and they did not compare their results to a condition in silence.

Carlyon and Beveridge (1993) tested intensity discrimination of 1-kHz, 25-ms pure tones at 30 and 80 dB SPL and at an individually chosen medium level (40, 50 or 60 dB SPL) with a 85-dB SPL, 100-ms narrowband noise masker (800-1200 Hz) played 170 ms before the interval. They found lower discrimination thresholds than without maskers when only the target (standard + increment) was preceded by a masker (the standard was presented without masker). They explained that with loudness enhancement ; but in the present experiment, the context is present before (and after) both intervals, conditions for which they found much higher thresholds. Their discrimination thresholds were actually even higher if the masker was presented before the standard only. However, they found no discrimination thresholds lower than without masker when the masker-probe interval was reduced to 15 ms : this is supported by Arieh and Marks (2003), who showed that there is little to no loudness enhancement for inter-stimulus delays below 150 ms. Since we have no delay between context and standard, loudness enhancement is unlikely to affect our results in the same way.

The relative absence of trends similar to ours in the literature would support the idea that the choice of signals is important in the observation of a decrease in discrimination thresholds for sounds presented within a context. The reductions in masking and perceptual similarity between context and standard compared to the experiments in the previous chapter are certainly part of the reasons why we observed results consistent with DRA.

There is another question to raise here : the test signals used for this experiment were optimized to get the flattest response profile at the output of the auditory filters for a characteristic frequency $f_c = 632$ Hz (Hilkhuysen and Macherey 2014). Given our stimulus bandwidth extended from 20 to 20 000 Hz, this means that the PSHC rate was not optimized in the low and high parts of the spectrum. As shown in figure III.3, the outputs of auditory filters with a low characteristic frequency are very similar to those obtained with random noise. Therefore, we would expect the center part and the high-frequency part of the PSHC to be most important for intensity discrimination.

Finally, restricting the overall bandwidth of the stimuli might improve the discrimination thresholds even further : by selecting only the optimized and high frequency regions, which are quite different from the noise, and thus suppressing the low frequencies that are very similar to noise, it is possible that the effect of DRA would be even stronger. Acquisition of the discrimination thresholds in the different frequency regions should thus display different trends.

III.3 Experiment 7 : Discrimination in different frequency bands

III.3.1 Material and methods

The participants, material, methods and protocol were identical to the first experiment in all context conditions ; only the bandwidth of the stimuli differed.

Only the loud standard was tested in this experiment.

The experiment was performed for the same PSHC signal filtered in three different ways. For the medium condition, the PSHC was bandpass filtered so that its passband was 2- octaves wide and the geometric center frequency remained the same as that of the broadband signal (632 Hz). The cutoff frequencies were 316 and 1264 Hz. For the lowpass and highpass conditions, the broadband PSHC was filtered by a lowpass or a highpass filter with cutoff 316 and 1264 Hz, respectively. The broadband level of the PSHC before filtering was 50 dB SPL or 70 dB SPL for the soft and loud conditions, respectively. This means that the amount of energy in each band remained the same as for the broadband stimulus but that the overall level of the filtered stimuli was lower than for the broadband. The filters were sixth order Butterworth low- and high-pass filters, applied to both the standards and the context. For the highpass band, the level of the standard was 69 dB SPL ; for the medium band, 57 dB SPL, and for the lowpass band, 51 dB SPL. The loud context had the same level as the standard and therefore depended on the bandwidth, and the soft level context was 20 dB lower.

In total, discrimination thresholds were measured in 9 conditions : three different frequency bands for the loud standard tested in silence, in the soft context, and in the loud context.

III.3.2 Results

Figure III.6 presents the results for each tested band : the low frequencies on the left, the medium band in the middle, and the high frequencies on the right. Table III.7 presents the mean and standard error for each condition : the discrimination thresholds in quiet were similar for all frequency bands and standard levels.

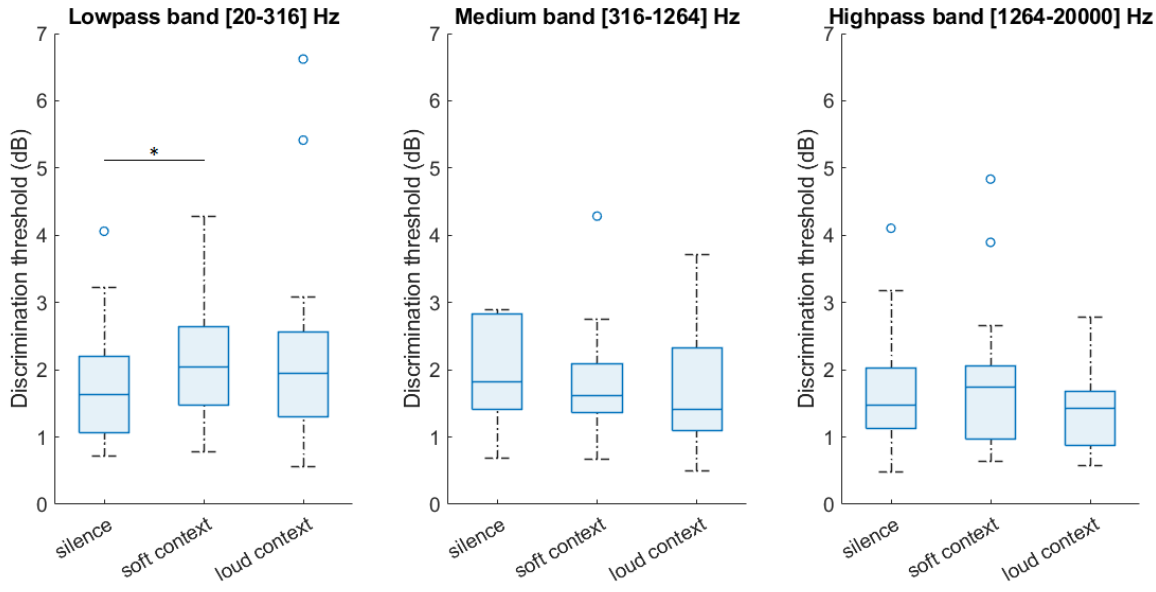


Figure III.6 – Results of experiment 7 for all participants and the three passbands used in the task. Discrimination thresholds are plotted against context level as in figure III.4. The thresholds are very variable for all conditions.

First, a two-way repeated-measures ANOVA was used to analyze the data, with factors context (3 levels : silence, soft and loud) and bandwidth (3 levels : low, medium, high). The analysis showed significant effects of the bandwidth ($F_{(2,30)} = 5.423$, $p = 0.010$), and of the interaction between the context and the bandwidth of the signal ($F_{(4,60)} = 3.968$, $p = 0.026$), but no effect of the context ($F_{(2,30)} = 1.929$, $p = 0.166$).

Additional one-way ANOVAs were run on the different bandwidths to investigate the effect of the context (three levels : silence, soft and loud). The context had a significant effect on the low-frequency band ($F_{(2,30)} = 3.46$, $p = 0.045$) and on the high-frequency band ($F_{(2,30)} = 5.25$, $p = 0.011$), but not on the medium frequency band ($F_{(2,30)} = 1.56$, $p = 0.228$).

Signal bandwidth	Context level (dB SPL)	Mean (dB)	Standard error (dB)
Lowpass	silence	1.80	0.22
	soft context (31)	2.15	0.25
	loud context (51)	2.31	0.40
Medium	silence	1.97	0.18
	soft context (37)	1.83	0.21
	loud context (57)	1.70	0.22
Highpass	silence	1.67	0.24
	soft context (49)	1.88	0.29
	loud context (69)	1.40	0.15

Table III.7 – Details on the averaged results for experiment 7. The mean discrimination thresholds tend to increase with the context mean level for the lowpass band, but tend to decrease between soft and loud contexts for the medium and highpass bands.

The only significant difference in pairwise comparisons between the context levels for each filtered band appears for the lowpass-band, between the silence and soft context ($t = -2.73$, $p = 0.046$). It is consistent with the effect of the context level in the second ANOVA (one-way).

III.3.3 Discussion

The results display no difference in discrimination thresholds between silence and loud context, or between soft and loud context.

This lack of differences in discrimination thresholds in either direction for the medium band is puzzling. The wholeband signal displayed a clear decrease of thresholds in the loud context (figure III.4b) and no difference between silence and soft context.

It is possible that filtering the signal affects it perceptually. A confusion effect as defined by Neff (1985, 1986) could be influencing the results, especially in the lowpass band in which the context and standard at the output of the auditory filters are very similar. However the output of the different auditory filters (figure III.3) suggest that the standard and noise are still very different in the medium and highpass bands.

The MOC has been proposed as a possible cause for DRA by Grange et al. (2022) : they found smaller shifts of rate-level functions with broadband noise than pure tones, even if they tested higher levels for noise than tones (see figure II.2),

but they do not discuss this. This would be consistent with a smaller effect of DRA on the discrimination thresholds for the broadband signals compared to the pure tone/narrowband combination (which is opposite to our results). However, Guinan et al. (2003) found that for the same SPL level, a broadband noise elicited a greater MOC response than a pure tone, a click, or a tone pip. Since the MOC is supposed to enhance DRA, both studies have contradictory findings.

The only cue we have on the effect of the bandwidth on DRA is the comment from Wen et al. (2009) discussed earlier: " the greater rates of dynamic range shifts observed for CF tones than for broadband noise may reflect a bias in our measurement procedure rather than a genuine difference in AN fibers' response properties between these two stimuli." It is thus not clear how DRA should react to a change in bandwidth.

Since there is no clear difference between the thresholds obtained in the different conditions, and that the decrease in discrimination thresholds in the loud context disappears, it is possible that all frequency regions are important to obtain the results of experiment 6. The high, rapid envelope fluctuations of the highpass-filtered standard (composed of pulses, visible at the output of the corresponding auditory filters) are apparently not much of an advantage nor disadvantage for intensity discrimination performance, nor is the flat output of the auditory filters for the medium band.

For the lowpass band, we expected the results to reflect the perceptual fusion of the context and standard (confusion), given the high similarity between auditory filter outputs. There was an increase in threshold compared to silence for the soft context, which would support the theory, but not in the loud context, in which the context and standard have the same level.

There is some evidence that a masker's similarity to the standard can affect the discrimination thresholds (Neff 1985 ; Oberfeld et al. 2014). Schlauch et al. (1997) carried an intensity discrimination study with a forward masker, looking at duration and spectral effects. They tested discrimination thresholds for pure tones and broadband noise with identical maskers, as well as pure tones with broadband maskers. They concluded that the effect of the masker was linked to the perceptual similarity between masker and standard, but that adding a cue to the masker to aid segregation did not suppress its effect. The discrimination thresholds with cues only decreased for low-level signals compared to the same stimuli without cue.

Some participants indeed reported that, on some trials, the standard was completely indistinguishable from the loud context, and two outliers are visible in figure III.6. However the similarity between context and standard affected the results way less than we anticipated. There might be enough differences in the higher frequency bands to enable segregation of standard and context.

To conclude, the intensity discrimination thresholds were consistent with DRA only when using broadband signals.

Chapter IV

Physiological response of the human auditory nerve

IV.1 Introduction

DRA has been observed at multiple levels of the auditory system including the auditory nerve (AN), the inferior colliculus and the auditory cortex of several non-human animal species (Dean et al. 2005 ; Watkins and Barbour 2008 ; Wen et al. 2009). Furthermore, these studies showed that the magnitude of the dynamic range shift is larger at central levels than at the AN suggesting DRA is initiated at the periphery and somehow "amplified" more centrally (Wen et al. 2009).

The mechanisms underlying DRA are still under debate. The fact that it is already present at the level of the auditory nerve suggests either that the hair-cells, the hair cell-auditory nerve synapse, and/or the auditory nerve are involved. Using a dual-adaptation phenomenological model, Wen et al. (2012) showed that the time constants of both DRA and classic firing-rate adaptation were correlated, suggesting these two types of adaptation may share a common basis. Given firing-rate adaptation is believed to arise, at least partly, from neurotransmitter depletion, the ribbon synapse appears as a likely site of initiation for DRA. A modeling study by Zilany and Carney (2010) showed that the inclusion of power-law dynamics in a phenomenological model of the inner hair cell-AN synapse could account for physiological data of DRA. More recently, using the computational model of Meddis et al. (2013), Grange et al. (2022) showed that both the acoustic and medial-olivocochlear reflexes may also contribute to DRA. However, efferent reflexes are unlikely to be the sole contributor because the various studies that observed DRA were performed in anesthetized animals for which the efferent reflexes are reduced.

Although there is no evidence that the AN directly contributes to DRA, it is already known that the auditory nerve shows firing-rate adaptation in response to direct electrical stimulation. This has been shown by single-fiber and electrically evoked compound action potential measurements (Zhang et al. 2007 ; McKay et al. 2013). This suggests that the firing-rate adaptation measured in response to acoustic stimulation may, at least partially, arise from the AN. In a recent study, Marrufo-Pérez et al. (2018) proposed that DRA may also arise at this level of the system. They showed that normal hearing listeners were better at recognizing speech in a noisy background when the noise started 300 ms before the word than when it started simultaneously. Interestingly, they also found a similar effect in CI listeners. This implies either that their paradigm does not measure DRA or that DRA is already present in CI despite the fact that CIs bypass the hair-cell AN synapse. Answering this question has both theoretical and clinical implications. First, if DRA is present

in direct electrical stimulation of the AN, this would strongly suggest that the AN also directly contributes to DRA in acoustic hearing. Second, if DRA is not present in electrical hearing, this would urge for the development of signal processing strategies that can mimic it and adapt the stimulation levels in real time depending on the context.

Measures of auditory nerve activity can be recorded in CI users through electrically evoked compound action potentials (eCAPs) (Hughes 2013). Usual measures of eCAP responses in humans with CI involve recording and averaging the responses to individual pulses presented at the same level. Amplitude growth functions (AGFs) are obtained by repeating this procedure at different levels. Here we include additional test protocols inspired from physiological DRA studies to investigate the effect of specific current level distributions on the AGFs. We systematically manipulate this current level distribution, and do so at different pulse rates assuming that DRA should be stronger at higher stimulation rates, as is the case for firing rate adaptation (Hu et al. 2010 ; Boulet et al. 2016 ; Adel et al. 2017).

IV.2 Material and methods

IV.2.1 Participants

Eleven CI listeners took part (c.f. table IV.1). All participants used devices manufactured by Cochlear Corporation and had been using their implant for more than 18 months at the time of testing. These participants were pre-selected from the hospital and laboratories databases based on their large eCAP responses (amplitude > 80 µV). This research was approved by the ethics comittees “Comité de protection des personnes Sud-Est I” (2022-A01600-43) for testing in Marseille (France) and by the "National Research Ethics Committee for the East of England" (ref. number 00/327) for testing in Cambridge (UK). All listeners were compensated for their time and travel expenses.

IV.2.2 Experimental hardware and software

eCAPs were measured using a CP910 speech processor connected to a programming pod (enabling communication between the computer and the CI processor). At the beginning of the session, Custom Sound EP 6.0 (Cochlear Limited) was used to identify an electrode with a large eCAP response amplitude (> 80 µV at a comfortably loud level) and to set the range of current levels to be tested. All levels were kept within the compliance range of the device and impedances were measured both at the beginning and at the end of the experiment.

Code	Gender	Age	CI array type (electrode type)	CI use (years)	Deafness before implant (years)	Etiology	Tested ear	Electrode
C2	F	66	CI24RE (CA)	17	4	Genetic, progressive	Right	11
C5	F	68	CI24RE (CA)	16	5	Genetic	Right	11
C15	F	45	CI622	4	5	Unknown	Right	14
C16	M	53	CI622	4	4	Genetic	Right	7
C17	F	72	CI622	3	4	Cochlear otosclerosis	Left	15
C18	M	52	CI622	3	15	Genetic Wolfram syndrome	Right	17
C19	F	25	CI622	2	3	Albright syndrome	Left	12
C26	F	61	CI522	7	3	Genetic	Right	6
C35	F	77	CI622	3	4	Meniere	Right	13
C46	F	46	CI622	1	39	Unknown	Right	12
C70	M	52	CI622	1	1	Unknown	Left	17

Table IV.1 – Information on the CI users who took part. The electrode number is the stimulating electrode ; the recording electrode was 2 electrodes more basal. Age and CI use are indicated at time of testing.

The main experiment used a custom Python script that served as an interface to NIC4.3 (Nucleus Implant Communicator 4.3) software library functions to precisely control stimulation and recording. The NIC4 software library provides low-latency streaming capabilities and supports eCAP recording, thereby allowing us to use minutes-long stimuli and retrieve the response to each individual pulse. Data processing and analysis was performed in Python and Matlab (The Mathworks, 2022b). The R software (4.4.2) was used for statistical analyses.

IV.2.3 Stimulation and recording parameters

The stimuli were delivered to a single electrode located roughly in the middle of the array [1], as indicated in table IV.1 for each participant. The stimuli were symmetric biphasic pulse trains (see figure IV.1) presented in monopolar mode using MP2 (located on the implant casing) as return electrode. The two phases of each pulse had durations of 37 µs and were separated by a gap of 8 µs. All measures were performed at three pulse rates (30, 100 and 300 pps), as detailed in section IV.2.5. These stimuli were presented in several experimental conditions which differed in their current level distribution, as described in subsection IV.2.5.

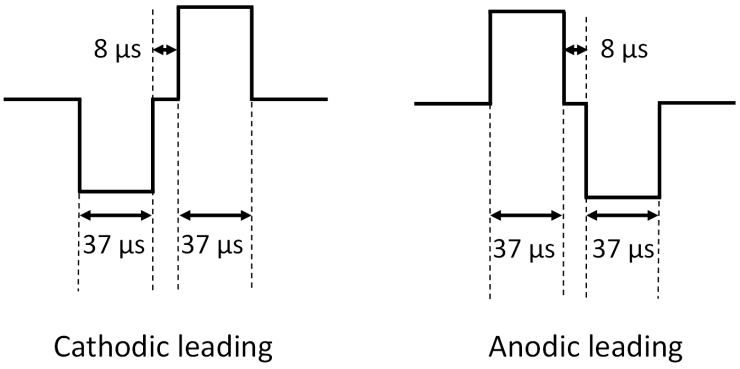


Figure IV.1 – Schematic temporal representation of two electric pulses, one cathodic-leading and one anodic-leading, with the durations used in the present study. Both phases of both pulses have equal amplitudes and duration.

The eCAP response to each individual pulse was recorded using an electrode contact 1.5 mm basal to the stimulating electrode, with reference to the MP1 ground electrode located under the temporalis muscle. Each eCAP measure started 73 µs after the electrical pulse offset and lasted 1500 µs. The recording was sampled at a rate of 21.3 kHz, and the gain of the amplifier was 40 dB.

The electrical artefact was reduced using the alternating polarity paradigm which averages the eCAP responses to anodic- and cathodic-first pulses. One drawback of this paradigm is that anodic- and cathodic-first responses usually show slightly different latencies. A consequence is that the eCAP amplitude responses are typically lower than when measured using the more common masker-probe paradigm (Macherey

1. The CI manufactured by Cochlear Corporation has 22 intracochlear electrodes.

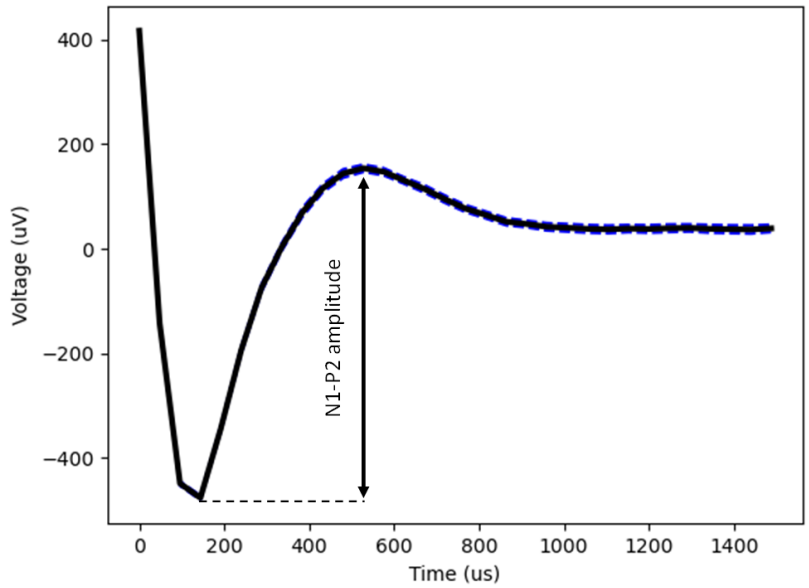


Figure IV.2 – Example of recorded temporal response of the auditory nerve to a certain current level, averaged over 100 repetitions, for one participant. The negative peak is called N1 and the positive peak P2 : the amplitude of the eCAP response is the voltage difference between P2 and N1.

et al. 2008 ; Baudhuin et al. 2016).[2] However, we do not expect this to affect the responses differently in the different experimental conditions. Most importantly, the alternating polarity paradigm allowed us to continuously monitor the AN response to constant-rate pulse trains which would have not been possible using the masker-probe paradigm. The highest rate tested (300 pps) was relatively low with respect to usual clinical rates in order to record the response to each pulse without contamination from the artefact of the following pulse. Nevertheless, previous studies have shown that this rate is high enough to show eCAP adaptation (Clay and Brown 2007) and this will be confirmed by the results of the present experiment.

An example of an eCAP response averaged over 100 repetitions of both polarities is shown in figure IV.2 for participant C70. The amplitude of the eCAP response was defined as the difference in amplitude between the second positive peak (P2) and the preceding negative peak (N1), averaged across all repetitions and pulse-leading polarities. The AGF represents the evolution of the eCAP response amplitude as a function of current level.

IV.2.4 Setting up current levels

After the electrode selection, comfortably loud levels were measured for each of the three pulse rates using stimulus durations of 3, 1, and 0.3 s for rates of 30, 100 and 300 pps, respectively. This was done by progressively increasing current level starting at 10 clinical units (CU). The step size of the current-level increase was 10 CU until reaching 70 CU, then 5 CU until detection threshold. It decreased to 2 CU until “soft but

2. This paradigm adds the response to the masker and to the probe in isolation, and subtracts the response to the combination, to get the response to the probe only.

comfortable", and finally 1 CU until the stimulus was perceived as comfortably loud. Each time a stimulus was presented, the participant was asked to provide feedback on their perceived loudness by pointing to a loudness chart ranging from 0 ("inaudible") to 10 ("too loud"). eCAPs were recorded in response to all stimuli used to determine comfortably loud levels. The eCAP threshold was selected as the lowest level at which a consistent response could be observed with the clinical software in all measurements averaged over 100 pulses each (at least two measurements were performed).

Next, twelve current levels (labeled L1 to L12), log-spaced in amperes, were selected so that L12 was the lowest MCL measured across the three pulse rates (which was for all participants that obtained at 300 pps) and L1 was just below eCAP threshold at 30 pps. Due to the limited current level resolution of the device and the choice of fixed steps between levels, however, for several participants more than one current level could be below eCAP threshold.

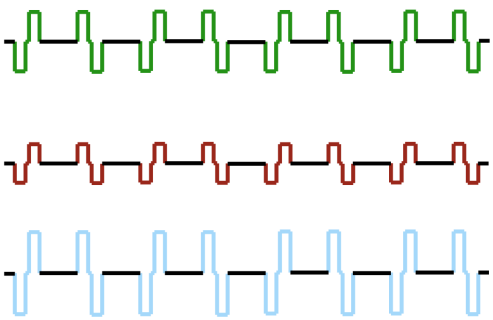
IV.2.5 Experimental paradigms

eCAP growth functions were measured in three different experimental paradigms, referred to as CLASSIC, RANDOM and HPR, as described below. In the animal studies, all 50-ms noise segments were analyzed ; similarly, all pulses are analyzed in the present study. For each of these paradigms, AGFs were collected at the three pulse rates (30, 100 and 300 pps).

For the CLASSIC paradigm, the responses to each of the twelve levels were measured separately by presenting a block of hundred pairs of constant-current level pulses with alternating polarity (c.f. figure IV.3a). Twelve blocks (corresponding to the 12 levels) were presented in random order for the three different pulse rates with gaps of 1 s between blocks. The CLASSIC paradigm was also repeated for 3 levels (L3, L7 and L11) using longer stimuli consisting of 5000 pulse pairs. This was done at the three tested rates to compare the responses to the other paradigms for stimuli of equal durations.

In the RANDOM paradigm, the responses to the twelve levels were collected within the same block by presenting all levels in a pseudo-random order (c.f. figure IV.3b) ; responses to pulses of both polarities at each level were extracted from the total measurement and averaged separately. More specifically, 417 sequences of 24 randomly ordered pulses (12 levels x 2 leading polarities) were constructed. The 417 sequences were played successively with no interruption. This specific number of sequences was chosen so that approximately 5000 pulses of each polarity were presented, yielding measurement durations of 5 minutes 33 seconds at 30 pps, 1 minute 51 seconds at 100 pps and 33 seconds at 300 pps.

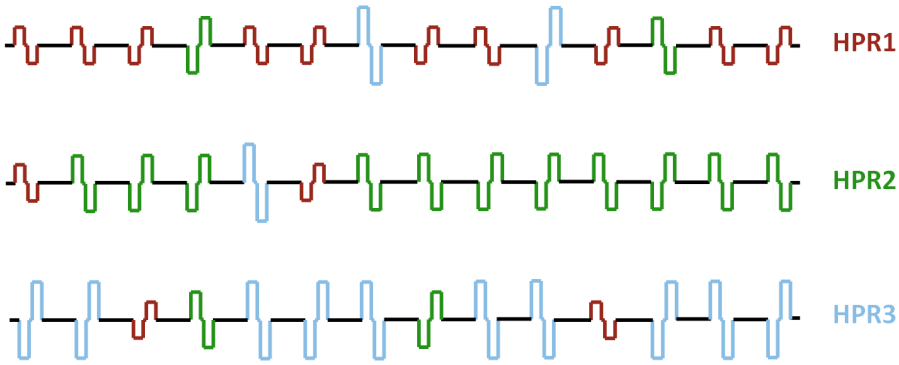
The HPR paradigm was inspired by the dynamic range adaptation measurement protocol introduced by Dean et al. (2005) and is illustrated in figure IV.3c. The 12 current levels were separated in two groups or "regions" : the high- and the low-probability regions (HPR and LPR, respectively). The HPR pulses consisted of 2 adjacent current levels chosen among the 12 levels and had a probability of presentation of 80 % (40 % for each level). The LPR pulses encompassed the remaining 10 levels and had a probability of presentation of 20 %. In practical terms, a total of 100 sequences



(a) CLASSIC paradigm with alternating polarity. Each level is presented separately.



(b) RANDOM paradigm with alternating polarity. Each pulse is presented an equal number of times.



(c) HPR paradigms with alternating polarity, for three different most probable levels. Each consecutive pulse group contains one pulse of each polarity for the LPR.

Figure IV.3 – Schematic representation of the stimuli used in the three paradigms. For illustration purposes, only three levels are represented, but twelve were effectively tested. a) CLASSIC, with pairs of pulses of both polarities presented one after the other at a fixed level ; b) RANDOM : all tested levels and polarities are mixed ; c) HPRs, with two particular levels with a probability of presentation of 80 % ; three different pairs of most probable levels were tested (for the purpose of the illustration, only one most probable level is presented for each of the three conditions).

of 100 pulses were presented and measured. Each sequence consisted of 100 pulses (2 levels x 2 polarities x 20 = 80 pulses for the HPR, and 10 levels x 2 polarities = 20 pulses for the LPR), presented in random order. This means that at least 200 measures (including both polarities) were obtained at each level (4000 pulses per level in the HPR, 200 in the LPR). Note that here again, a total of 5000 pulses of each polarity were presented, yielding stimulus durations equal to those of the RANDOM paradigm. The HPR paradigm was repeated three times using different HPR levels (HPR1 : L2-L3, HPR2 : L6-L7 and HPR3 : L10-L11).

IV.2.6 Session

Before to start the main experiment measurement, short versions of all paradigms were presented to the participants to verify that they could comfortably listen to those for extended periods of time (respectively 12, 5.25 and 3 s for each rate). For participants C19 and C26, the current level ranges had to be shifted down by 2 and 4 CU, respectively. The participants were then invited to relax and read or watch a subtitled movie for the rest of the experiment.

eCAPs were collected for all participants for five paradigms presented in the following order : CLASSIC, RANDOM, HPR1, HPR2 and HPR3. For each paradigm, the responses were collected at three pulse rates (30, 100 and 300 pps). To ensure that the responses obtained in the different conditions were comparable, we verified that there was no long-term eCAP changes within the testing session by repeating the CLASSIC paradigm at 30 pps at the end of the CLASSIC and RANDOM paradigms, and after each of the three HPR paradigms.

For five participants (C5, C15, C16, C17 and C19), the measurements were repeated a second time on the same day. Likewise, retests were performed on the same day for C2, C18, C35 and C46, except at 100 pps due to time constraints. For participants C26 and C70, these retests at 100 pps were only obtained for the CLASSIC and RANDOM paradigms.

IV.2.7 Analyses

For each participant, and each experimental condition, eCAP recordings were averaged separately at each level. The resulting mean eCAP traces were visually checked by two experimenters (Coline Marché and Olivier Macherey) in order to exclude traces with no visible eCAP responses. N1 and P2 amplitudes were automatically extracted from the averaged eCAP waveform by finding the lowest-amplitude sample within the first 375 µs and the highest-amplitude sample occurring within 500 µs after N1. The eCAP amplitude was then calculated by subtracting N1 from P2 (see figure IV.2). When available, the eCAP response amplitudes for test and retest were averaged.

The noise level associated with each eCAP was computed as in Undurraga et al. (2012). Briefly, four equally-spaced time points (M) were chosen from 375 µs to 1200 µs after the pulse presentation, separated by interval $\Delta t = 280$ µs. For each time point m and each individual eCAP recording i, the variance of the noise was calculated as the difference between the amplitude of the sample for each eCAP recording ($s_i(m\Delta t)$) and the amplitude of this same sample averaged over all the recordings ($s_{moy}(m\Delta t)$), squared, summed over the total number of recordings N and divided by $N-1$. The total variance of the noise (σ^2_{noise}) was defined as the averaged variance over the four time points :

$$\sigma^2_{noise} = \frac{1}{M}\sum_{m=1}^{M}\left[\frac{1}{N-1}\sum_{i=1}^{N}(s_i(m\Delta t) - s_{moy}(m\Delta t))^2\right] \tag{IV.1}$$

The SNR was further defined as the variance of the averaged eCAP amplitude (σ^2_{eCAP}) divided by the variance of the noise (σ^2_{noise}) :

$$SNR = \frac{\sigma^2_{eCAP}}{\sigma^2_{noise}} \tag{IV.2}$$

eCAPs with an amplitude lower than 2 standard deviations of the noise ($2\sigma_{noise}$) were excluded from the analysis.

IV.2.7.1 Very-long-term eCAP evolution

Results in section IV.3.1.

The control 30-pps CLASSIC amplitude responses measured at the end of each paradigm were analyzed to inform us on any long-term change in eCAP amplitude over the duration of the experiment. Analyses were performed on L11 and L12.

IV.2.7.2 Investigating the effect of randomizing current level presentation

Results in section IV.3.2.

The eCAP amplitude growth functions (AGFs) obtained in the CLASSIC and RANDOM paradigms were first compared. We hypothesized that the shape of the functions should be different at high rates due to increased adaptation in the CLASSIC paradigm, more specifically that the amplitudes for the CLASSIC paradigm should be lower than those for the RANDOM paradigm.

IV.2.7.3 Quantifying dynamic range adaptation as a function of the HPR

Results in section IV.3.3.

If dynamic range adaptation is produced at the level of the auditory nerve, we would expect a HPR-dependent horizontal shift of the AGF. For each participant and each pulse rate, the shift of each HPR AGF relative to the RANDOM AGF was calculated using the same method as Brochier et al. (2021) to quantify the effect of inter-phase gap on eCAP : the two portions of the HPR and RANDOM AGFs showing overlapping values of eCAP amplitudes were first extracted. The shift was then calculated as the mean difference in current level between these two functions (see figure IV.6c). Nine values of shift were obtained this way (3 rates x 3 HPRs) for each participant.

Given the magnitude of the current level difference between HPRs differed across participants (due to differences in their dynamic ranges), for each rate, we fitted linear regression functions to the three shift values as a function of the difference in mean current level of the HPR (see figure IV.6d) compared to the RANDOM paradigm. The slope of this function represents the shift in dynamic range expressed in dB/dB for each rate, and can be seen as the eCAP version of the DRA shifts reported in animal studies that used single-fiber recordings (Dean et al. 2005 ; Wen et al. 2009).

IV.2.7.4 Comparing the slopes of the eCAP growth functions

Results in section IV.3.4.

The second hypothesis we wanted to test was whether this shift provides an advantage in terms of intensity coding. To do that, the slopes of the eCAP growth

functions obtained in each HPR condition were compared to those obtained in the RANDOM condition in two different ways (see figure IV.8).

For the first comparison, the slope of the HPR eCAP growth function (expressed in dB/dB) was computed at the mean current level of the HPR pulses (L2-L3, L6-L7 or L10-L11), and compared to the slope of the RANDOM AGF at the *same current level*. If a difference is observed, this would mean that the change in eCAP amplitude does not simply reflect a vertical shift of the eCAP AGF.

For the second comparison, the slope of the HPR AGF measured at the mean current level of the HPR pulses was compared to the slope of the RANDOM function measured at a current level that gives the *same eCAP amplitude*. The same-amplitude points of the RANDOM paradigm were obtained by interpolation from the closest neighboring levels, and the slope computed between these two interpolated points. In this case, a difference in slopes would imply that the change in eCAP amplitude does not solely reflect a horizontal shift of the eCAP AGF.

Finally, even if the HPR conditions show an increase in slope, it does not necessary imply that this provides an advantage in terms of intensity coding. More specifically, if the trial-to-trial variability of the eCAP response amplitude is higher in the HPR case, this might mitigate any advantage due to the increased slope. In addition to these slope measurements, we therefore also measured the standard deviation σ of the eCAP amplitude response A at these specific levels for the HPR3 paradigm. We computed the coefficient of variation :

$$V = \frac{\sigma_{L10} + \sigma_{L11}}{A_{L10} + A_{L11}} \tag{IV.3}$$

and the sensitivity index as given by Wen et al. (2009) :

$$\delta' = \frac{A_{L11} - A_{L10}}{(L11 - L10) \times \sigma} \tag{IV.4}$$

with $\sigma = \sqrt{\frac{\sigma_{L11}^2 + \sigma_{L10}^2}{2}}$ and $\sigma_{Lx}^2 = \sigma_{Lx}(N1)^2 + \sigma_{Lx}(P2)^2$

This sensitivity index represents the slope of the AGF divided by the standard deviation of the amplitude.

IV.2.7.5 Evaluating the temporal changes in different conditions

Results in sections IV.3.5.1 and IV.3.5.2.

The last aspect of the data we wanted to investigate was to quantify the time constant of adaptation for each experimental paradigm. The decay of the eCAP response was measured at three levels (L3, L7 and L11) for the CLASSIC and RANDOM paradigms and at L3 for HPR1, L7 for HPR2 and L11 for HPR3.

For each paradigm and each level analyzed, the eCAP amplitude was averaged in groups of 100 consecutive pulses (representing 3 seconds and 333 ms for rates of 30 and 300 pps, respectively). Each eCAP amplitude was normalized by the mean eCAP response to the first group of 100 pulses in the CLASSIC paradigm at 30 pps (collected at the same current level). The reason is that this condition shows minimal adaptation

and should therefore be representative of the response to the very first pulse of all our stimuli (see section IV.3.5.1).

The strength of the adaptation was quantified using the adaptation index which represents the normalized eCAP amplitude for the last group of 100 pulses in the recording (Clay and Brown 2007).

Previous physiological studies used exponential fits to evaluate de time constant of DRA. Here however, the diversity in shape of the adaptation functions did not allow us to proceed the same way. So to be able to compare our data to the physiological studies, we considered the time constant τ to be the time taken by the response to reach 36.8 % of its final value (100th group). This particular value was chosen because it corresponds to the time constant of an exponential function.

IV.2.7.6 Control analyses

Results in sections IV.3.6 and IV.3.7.

Lastly, two control analyses were performed to evaluate the possible influence of local effects on our measures.

First, it has been shown that within this range of rates, the eCAP response to a constant-level pulse train sometimes shows alternating amplitudes with the amplitude of odd-numbered pulses being larger than that of even-numbered pulses (Wilson 1997 ; Hughes et al. 2012 ; Carlyon and Deeks 2015). Although we would expect this effect to be reduced in the RANDOM and HPR paradigm due to the randomized level presentation, we computed the AGFs for even pulses and odd pulses separately. This analysis was restricted to levels L6 and L7 obtained in the HPR2 measurements at 300 pps. Note that this analysis could not be performed for the CLASSIC paradigm because all odd-numbered pulses were cathodic-leading and all even-numbered pulses were anodic-leading. After artefact cancellation, therefore, any amplitude difference between odd and even pulses was lost due to the averaging.

Second, we may expect short-term refractoriness due to the pulse immediately preceding the pulse we record from to have an influence on the measured eCAP amplitude (Boulet et al. 2016). To investigate the contribution of this local effect on the shift of the HPR AGF, the responses to L6 and L7 presented in HPR2 were separated in two groups : the responses to pulses which were preceded by a pulse with a low level (between L1 and L5), and those preceded by a pulse with a high level (between L8 and L12), taking care that the same number of positive and negative pulses were averaged. The averaged responses for these two groups of responses were further compared. For the RANDOM paradigm, responses for L2 to L11 were computed depending on if the directly preceding level was lower or higher than the pulse level at which the response was measured. The cases where the preceding level was identical were excluded.

IV.3 Results

IV.3.1 Very-long-term eCAP changes

The "very long term" evolution of the eCAP amplitude was monitored through CLASSIC measurements at 30 pps (see analyses in section IV.2.7.1), repeated multiple times during the session (i.e., at the start of the experiment and then after each CLASSIC/RANDOM/HPR measurement), giving a total of 4 repetitions of this measurement (8 for participants who did test and retest). The response amplitude for the ith measurement was computed for each participant, and the amplitude values were averaged over test and retest when available. An example of this evolution is shown for level L11 for all participants in figure IV.4.

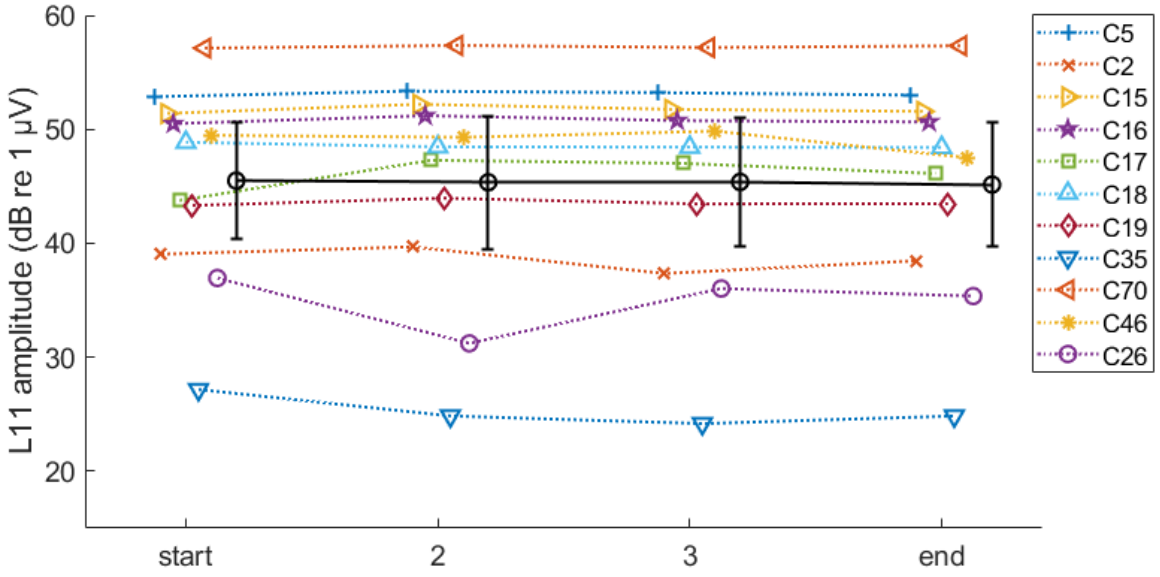


Figure IV.4 – Evolution of L11 amplitude in dB for the short CLASSIC paradigm, between the start and the end of the procedure. Individual data (one marker per participant) are completed by mean and 95 % confidence interval (in black). Test and retest were averaged.

A one-way repeated-measures ANOVA with measurement time as a treatment factor (4 levels) showed no consistent change in amplitude across the session ($F_{(3,30)}$) = 0.740, p = 0.489). This was also true for L12 (see appendix C.1).

We thus conclude that there was no systematic changes in eCAP amplitudes within the test session. The eCAP responses obtained in the different paradigms can, therefore, be compared with one another.

IV.3.2 Differences between AGFs for CLASSIC and RANDOM paradigms

Figures IV.5a and IV.5b show examples of AGFs obtained in the RANDOM and CLASSIC paradigms for participants C70 and C26 at 300 pps (see analyses in section IV.2.7.2). For C70, eCAP responses were obtained from L2 to L12 ; for C26, more levels

turned out to be below eCAP threshold, and the figure shows responses from L4 to L12. In all cases, the amplitude at each level was calculated based on 100 measurements. At 30 and 100 pps, the two paradigms produced similar AGFs (data not shown).

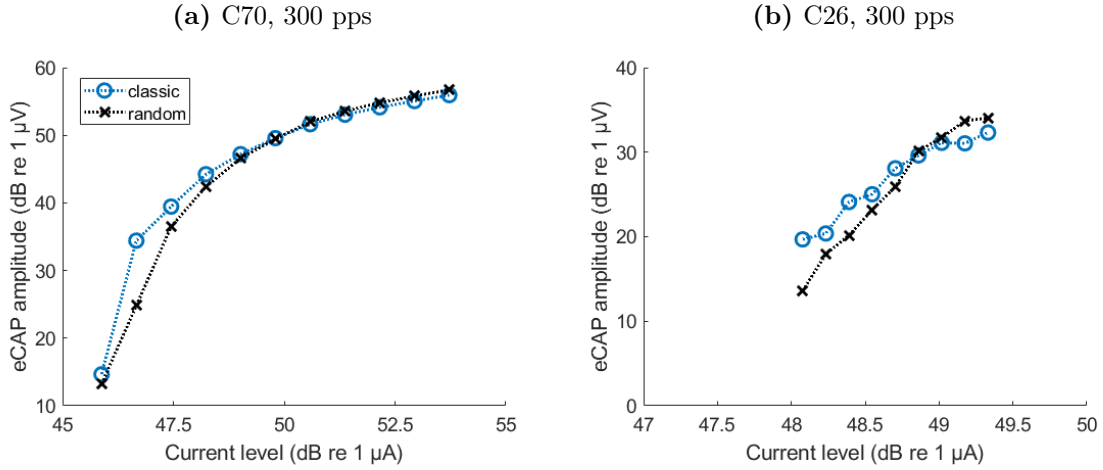


Figure IV.5 – AGF for the CLASSIC and RANDOM paradigms for participant C70 at 300 pps (a), and C26 at 300 pps (b).

However at 300 pps, the CLASSIC function appears to have a greater response amplitude at low levels, but the RANDOM AGF is steeper and shows higher amplitudes at high levels. The two AGFs cross around L7 for participant C70, and L9 for C26. This trend was actually observed in most participants, but the level at which the AGFs crossed at 300 pps varied across listeners. For 4 participants, the CLASSIC AGF was below the RANDOM AGF at 300 pps at all levels.

Because all participants did not have the same range of tested levels, and because the individual eCAP thresholds were different, linear-mixed models were fitted to the whole dataset. The dependent variable was the difference in eCAP amplitude between the RANDOM and the CLASSIC paradigms. Participants were considered as a random effect by including a random intercept in the model ; rate was a fixed factor, and level was continuous. For each model (incorporating fixed factors progressively), the -2 log-likelihood was used as indicators of the fit. These values were further compared using chi-squared tests.

There was a significant improvement of the fit with the addition of the level ($\chi^2(1) = 3.973$, p = 0.046) compared to the random intercept model ; moreover, the slope estimate (0.251, p = 0.044) of the model indicated that the difference between RANDOM and CLASSIC increases with increases in current level, consistently with our observations. The factor rate also improved the model ($\chi^2(2) = 23.60$, p = 7.52e-6), as did the interaction between both factors ($\chi^2(2) = 9.528$, p = 0.009) : the greatest (or more positive) differences were observed at the highest rates and levels, and the smallest (or more negative) differences at the lowest rates and levels.

Our initial hypothesis was that the RANDOM paradigm should yield larger amplitudes than the CLASSIC because of reduced adaptation. While this was consistently true at high stimulation levels, this was not always the case at low levels, as shown in

Figure IV.5. It is possible that the adaptation remains weak for the CLASSIC at low levels. For RANDOM, at low levels, the reduction in amplitude sometimes observed may be due to stronger refractory effects of one pulse over another (the levels are all randomized for RANDOM, so the presentation of a low-level pulse can be preceded by a very high-level pulse).

IV.3.3 Horizontal shift of the AGF as a function of HPR

Figures IV.6a to IV.6c shows examples of AGFs for the RANDOM, HPR1, HPR2 and HPR3 paradigms at the three tested pulse rates, for participant C5 (see analyses in section IV.2.7.3). The AGFs of all the other participants can be found in appendix C.2.

For each participant, each HPR paradigm and each rate, the horizontal shift of the AGF was measured relative to the AGF obtained in the RANDOM paradigm, thereby giving a total of nine shift values per participant. These values are shown in figure IV.6d, for participant C5. A positive shift indicates that the AGF is shifted towards higher current levels compared to RANDOM.

Figure IV.6e shows the slope of the shift as a function of the rate, for all participants (one symbol per participant). This slope represents the amount by which the HPR AGFs shifts as a function of the HPR level. The mean and 95 % confidence intervals were computed for the slopes at each pulse rate and plotted as the black circles and bars in figure IV.6e.

For all rates, there is a shift consistent with DRA : as the HPR level increases, the AGF shifts horizontally towards higher current levels.

Furthermore, a one-way repeated-measures ANOVA showed a significant effect of pulse rate on the slope ($F_{2,20} = 9.189$, $p = 0.008$), demonstrating that the shift of the AGF is larger at high than at low rates, consistent with our hypothesis. Pairwise comparisons also revealed a significantly steeper slope for 300 pps than 30 pps ($t = -4.95$, $p = 0.002$) and than 100 pps ($t = -4.57$, $p = 0.003$). There was no difference between the slopes obtained at 30 and 100 pps ($t = -0.380$, $p = 1$).

To conclude, these results show that the eCAP AGF is influenced by the experimental paradigm and the pulse rate, and exhibits a behavior consistent with DRA.

IV.3.4 Change in response function shape

The previous section revealed a horizontal shift of the AGF in the HPR paradigms compared to the RANDOM paradigm. In this section we investigate how the shape of the AGF changes with the paradigm and stimulation rate, and whether this global shift produces an improved encoding of the HPR levels. This is done by comparing the slopes of the AGFs (see analyses in section IV.2.7.4).

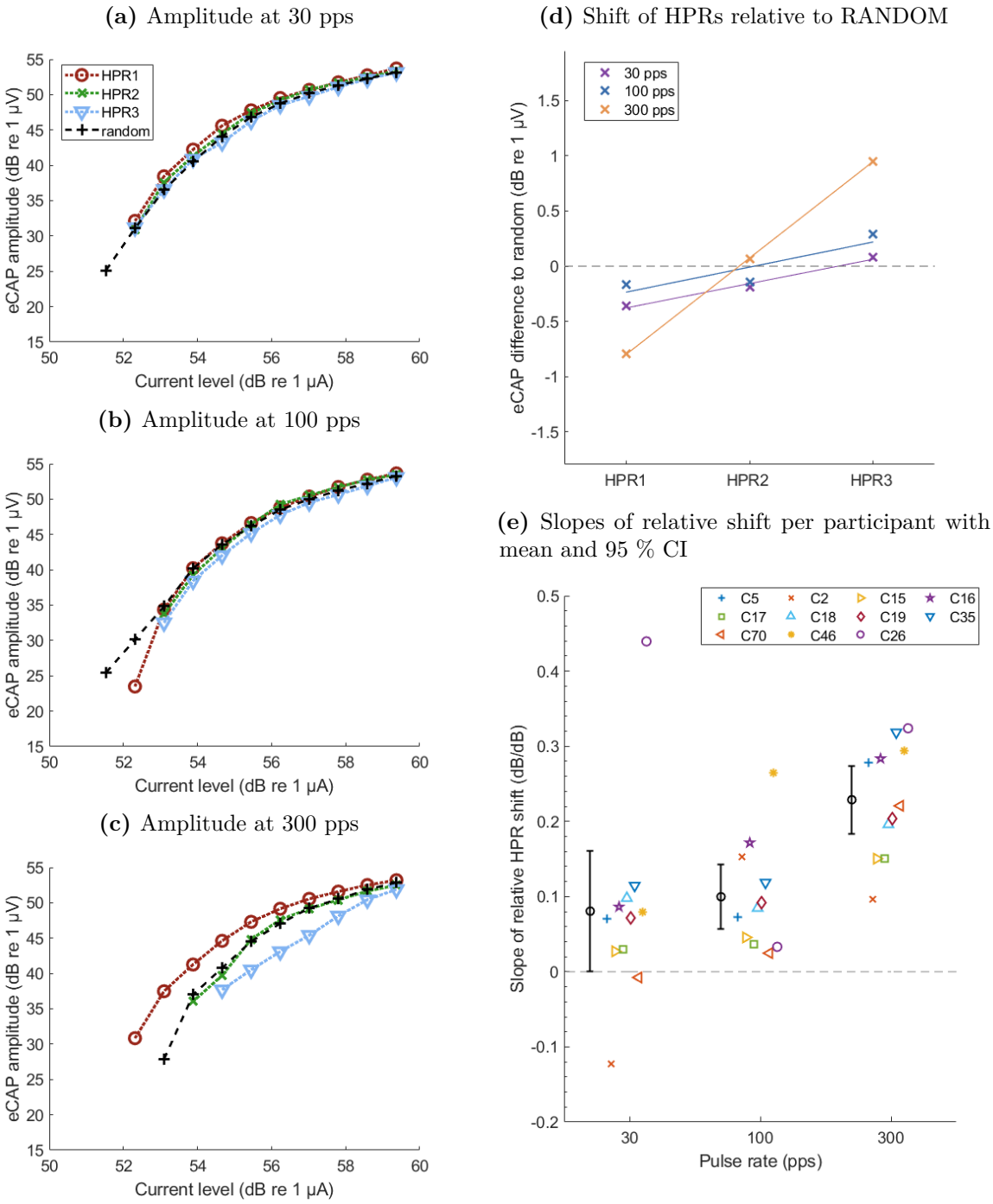
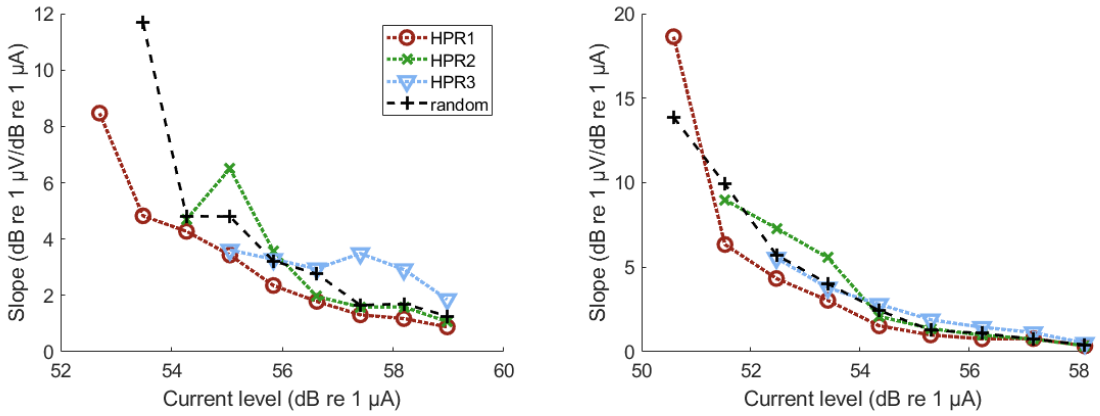


Figure IV.6 – Amplitude growth function, relative shifts and shit slopes. The AGF and relative shift relative to the RANDOM condition are plotted for one participant only (C5) : a) AGF 30 pps, b) 100 pps, c) 300 pps ; d) average relative shift of the corresponding HPR paradigm compared to RANDOM AGF : in this case, the mean current level difference between two consecutive HPRs was 3.14 dB ; e) slopes of the average relative shift for all rates and participants (one marker per participant) ; the across-subject mean and 95 % confidence intervals are shown in black next to the individual results.



(a) Subject C5 **(b)** Subject C18

Figure IV.7 – Examples of slopes calculated between two subsequent points for the HPRs and the RANDOM measurements for two participants at 300 pps. a) Slope maxima are observable for each paradigm and shift towards higher or lower levels depending on the most probable levels in the measurements ; b) in most cases the maximum was the first available slope and no evident shift appears.

Examples of AGF slope vs. current level functions obtained in the RANDOM and HPR paradigms at 300 pps are shown in figure IV.7 for participants C5 and C18. Figure IV.7a shows an effect of HPR on the slopes consistent with the hypothesis of improved intensity coding in the HPR region : the slope for HPR1 is highest at the lowest levels and then monotonically decreases, while the slopes of HPR2 and HPR3 show non-monotonic evolutions with local maxima at intermediate and at high levels, respectively. Note this behavior was, however, not observed in all participants. For example, figure IV.7b (C18) reveals small differences across HPR paradigms, with functions showing no local maxima.

The slopes of the AGFs obtained in the HPR and RANDOM paradigms were computed in two different ways. (i) for the "same-level" condition, they were compared at the mean level of the HPR pulses (in figure IV.8 : HPRX, light blue, and RANDOM same current, black). For the "same-amplitude" condition, the HPR slope at the mean level of the HPR pulses was compared to the RANDOM slope at a level producing the same eCAP amplitude (RANDOM same amplitude, purple in figure IV.8). The slope values obtained in each HPR and averaged across participants are presented in table IV.2 ; missing data were due to the responses at these particular levels being too noisy or below eCAP threshold to compute the slope.

The comparison of interest is between the slopes for the HPR and for the RANDOM (at fixed current level or fixed response amplitude) paradigms. To investigate the influence of rate on these slopes, the slope differences between HPR and RANDOM were fitted using two distinct linear mixed-effects models (one for fixed level and one for fixed amplitude comparison). Participants were considered a random factor (modeled as a random intercept), while rate and HPR paradigm were fixed factors.

Rate (pps)	Condition	Nb of values	Mean (dB/dB)	Std
30	HPR1	2	13.7	1.17
	RANDOM same current	4	11.4	4.22
	RANDOM same amplitude	2	14.3	2.62
	HPR2	11	14.2	13.7
	RANDOM same current	11	11.4	12.2
	RANDOM same amplitude	10	10.2	9.47
	HPR3	11	7.26	7.82
	RANDOM same current	11	6.32	6.54
	RANDOM same amplitude	10	7.03	6.51
100	HPR1	2	20.0	8.55
	RANDOM same current	3	12.6	6.74
	RANDOM same amplitude	1	12.1	-
	HPR2	10	13.0	10.7
	RANDOM same current	11	13.1	11.0
	RANDOM same amplitude	10	12.6	10.8
	HPR3	11	5.22	3.45
	RANDOM same current	11	5.44	7.12
	RANDOM same amplitude	9	5.11	3.45
300	HPR1	1	28.6	-
	RANDOM same current	1	18.4	-
	RANDOM same amplitude	0	-	-
	HPR2	10	21.2	15.6
	RANDOM same current	10	14.5	10.5
	RANDOM same amplitude	7	20.3	23.0
	HPR3	11	10.6	9.19
	RANDOM same current	11	6.72	6.96
	RANDOM same amplitude	11	7.97	7.05

Table IV.2 – Average slopes and standard deviation for each condition, for the analysis of slope changes depending on rate and paradigm. All 11 participants are considered ; the number of values indicates the number of slope calculation that were possible.

IV.3.4.1 Comparison at fixed current level

The comparisons at fixed current level was intended to establish whether the AGF shift observed in section IV.3.3 for the HPR paradigm could be explained by a purely vertical shift of the AGF compared to the RANDOM paradigm.

Adding the factor rate in the analysis produced a significantly better fit ($\chi^2(2) = 6.268$, $p = 0.044$). There was however no interaction between rate and HPR paradigm ($\chi^2(4) = 2.019$, $p = 0.732$). Therefore, this analysis demonstrates that the difference in slope between HPR and RANDOM depends on the pulse rate.

Additional pairwise comparisons showed the slopes to be steeper for the HPR paradigms than for the RANDOM paradigm at 300 pps for both HPR2 ($t = 2.39$, $p = 0.044$) and HPR3 ($t = 3.31$, $p = 0.008$). This analysis shows that the AGF shift is not purely vertical.

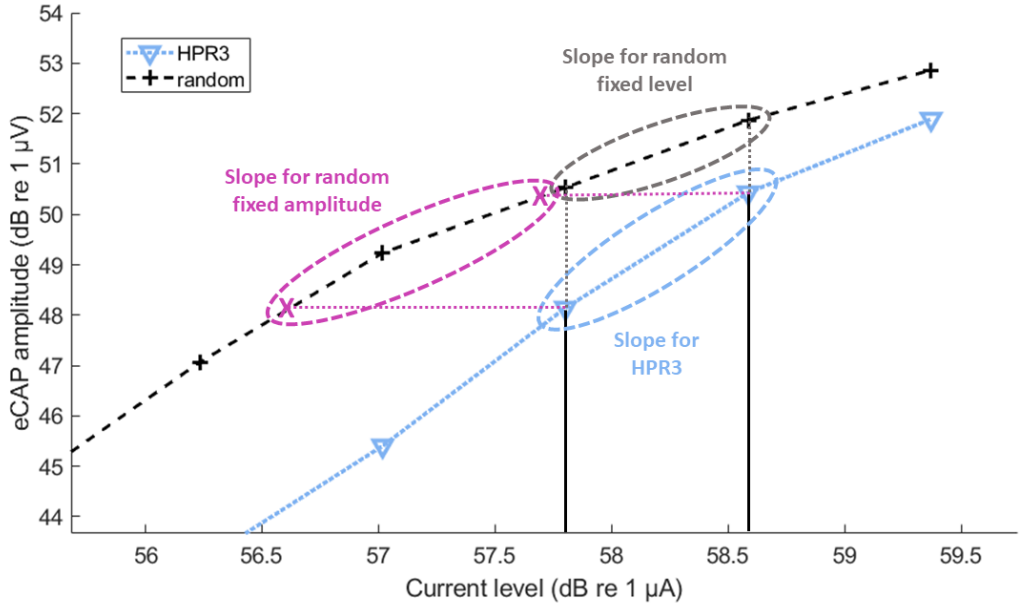


Figure IV.8 – Slopes of interest for the HPR3 condition at 300 pps : zoomed in from figure IV.6c (HPR1 and HPR2 removed). The slope for the RANDOM curve at fixed amplitude is calculated with a linear regression on the two interpolated points on the RANDOM curve (purple X).

IV.3.4.2 Comparison at fixed amplitude

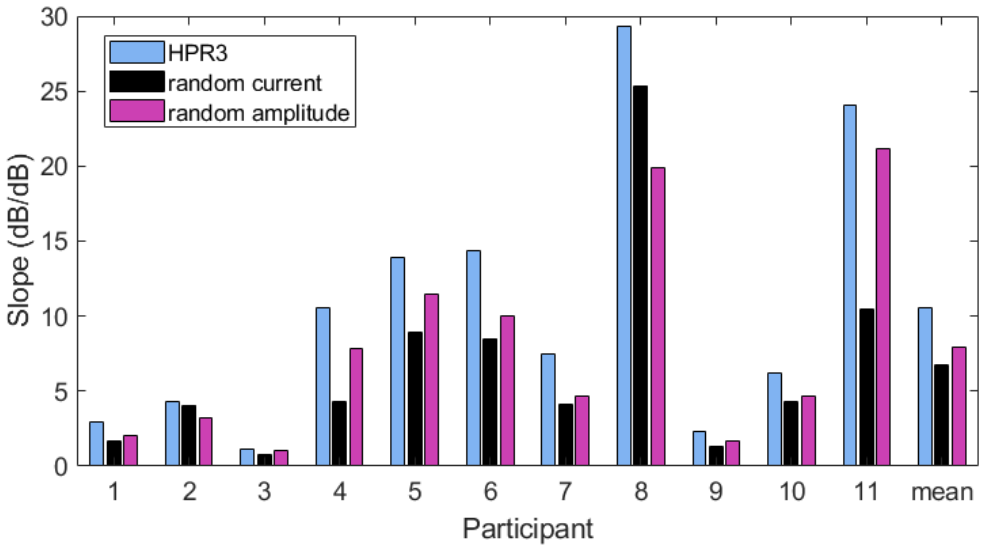
When compared at a fixed eCAP amplitude (blue vs purple in figure IV.8), the same linear mixed models were used. The fit of the model was not improved by the addition of the rate ($\chi^2(2) = 0.078$, $p = 0.962$). However, there was a significant interaction between rate and HPR paradigm ($\chi^2(3) = 10.58$, $p = 0.014$). This indicates that only certain rate-HPR combinations produce changes in slope. Additional pairwise comparisons showed a significant difference between RANDOM and HPR at 300 pps for the HPR3 paradigm ($t = 3.35$, $p = 0.007$) but not in the other conditions.

To conclude, it appears that there is a change in the shape of the response function for the HPR3 measurements at 300 pps. The AGF shift is thus not purely horizontal either ; the shape of the response function of the neurons changes at the population level.

IV.3.4.3 Do the steeper slopes in the HPR reflect improved intensity encoding ?

The individual values for the slopes for HPR3, RANDOM fixed current and RANDOM fixed amplitude at 300 pps are plotted in figure IV.9a. As seen in the previous section, the AGF slope is steeper at the HPR levels of the HPR3 than for the RANDOM paradigm. This was true both when comparing these slopes at a fixed current level and at a fixed eCAP response amplitude. However, to demonstrate that this translates into an improved intensity encoding, we still need to show that the response in the HPR3 paradigm is not more variable than for the RANDOM paradigm. To test that, we computed the coefficients of variation of the eCAP amplitude at the levels and

(a) Slopes per participant with mean



(b) Sensitivity index per participant with mean

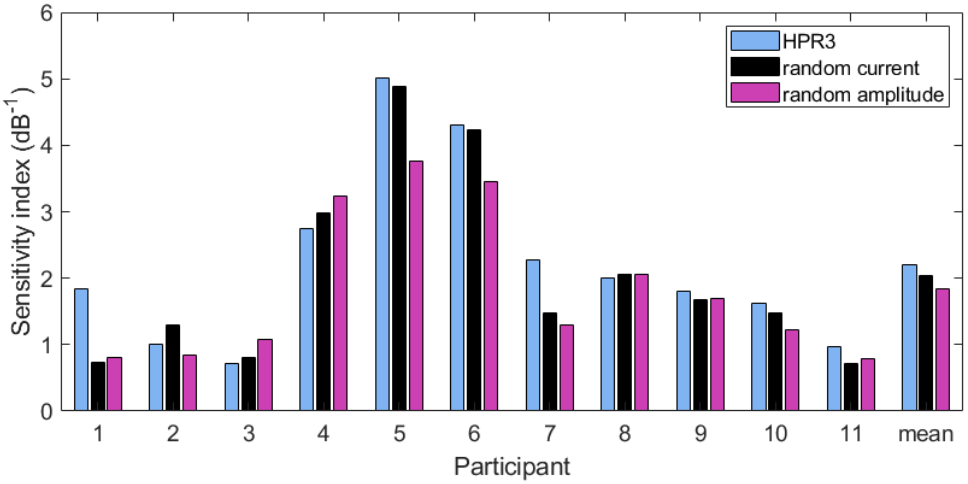


Figure IV.9 – Measures of response function changes in shape at 300 pps : a) slopes between the two HPR levels for the HPR paradigm and the RANDOM paradigm as well as the interpolated slope at fixed amplitude ; b) sensitivity index (slope over standard deviation) for the same conditions.

paradigms at which each of these three slopes was computed. This coefficient is defined as the standard deviation of the difference between P2 and N1, divided by the mean N1-P2 eCAP amplitude.

We found the coefficient of variation to be significantly smaller in the RANDOM paradigm than in the HPR3 when compared at fixed current levels ($t = -3.62$, $p = 0.014$). There was no difference when compared at fixed response amplitude ($t = -0.504$, $p = 0.625$). This analysis therefore shows that the increased slope observed for the HPR3 compared to RANDOM at the HPR current level is concomitant with an

increase in the variability of the response.

We then followed the approach of Wen et al. (2009) and computed the sensitivity index associated with the three slopes. This sensitivity index represents the ratio between the AGF slope and the standard deviation of the amplitude (see section IV.2.7.4) and is shown in figure IV.9b. Pairwise comparisons revealed no significant difference in sensitivity index between HPR3 and RANDOM fixed current (t = -1.42, p = 0.187), nor between HPR3 and RANDOM fixed amplitude (t = -2.10, p = 0.062).

To conclude, despite a significant increase in the slope of the AGF observed in the HPR3 paradigm, the data do not conclusively show an improved encoding of intensity/current level in the eCAP responses.

IV.3.5 Temporal effects

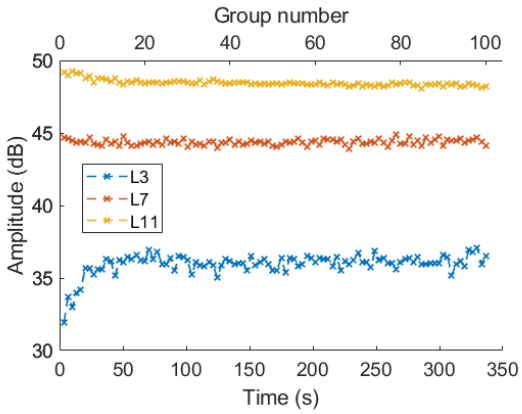
The long duration of our measurements allows us to investigate adaptation over several time scales (see analyses in section IV.2.7.5). The very long term changes were already addressed in section IV.3.1 by analyzing the short CLASSIC measurements repeated throughout the protocol. In the following paragraphs, we first study the long term evolution of the response for the long CLASSIC paradigm at 30 pps, for which responses to L3, L7 and L11 were recorded during 5 minutes 30 seconds each (10 000 pulses). The other rates were recorded for shorter durations (1 minute 30 s for 100 pps, 33 s for 300 pps).

IV.3.5.1 Evolution at 30 pps

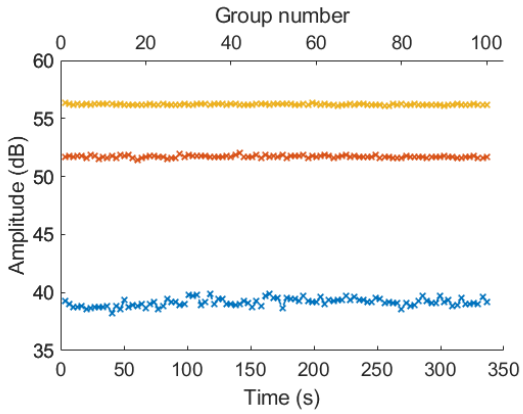
Figures IV.10a and IV.10b show two examples of the temporal evolution of the eCAP amplitude in response to the CLASSIC paradigm for participants C46 and C70 at 30 pps. For C46, the changes in amplitude depend on level : the amplitude decreases for L7 and L11, but it increases for L3 ; this pattern was robust, as it was observed both for the test and retest measurements. For C70, however, the amplitudes did not change across the measurement. Overall, of all participants, five (including C70) displayed no changes in amplitude over time.

The eCAP amplitude (in dB re 1 µV) in response to the first (0 to 3.33 s), second (3.33 to 6.66 s) and last groups of 100 pulses (330 to 333.3 s) of the 30 pps CLASSIC measurements at all three levels is plotted in figure IV.10c.

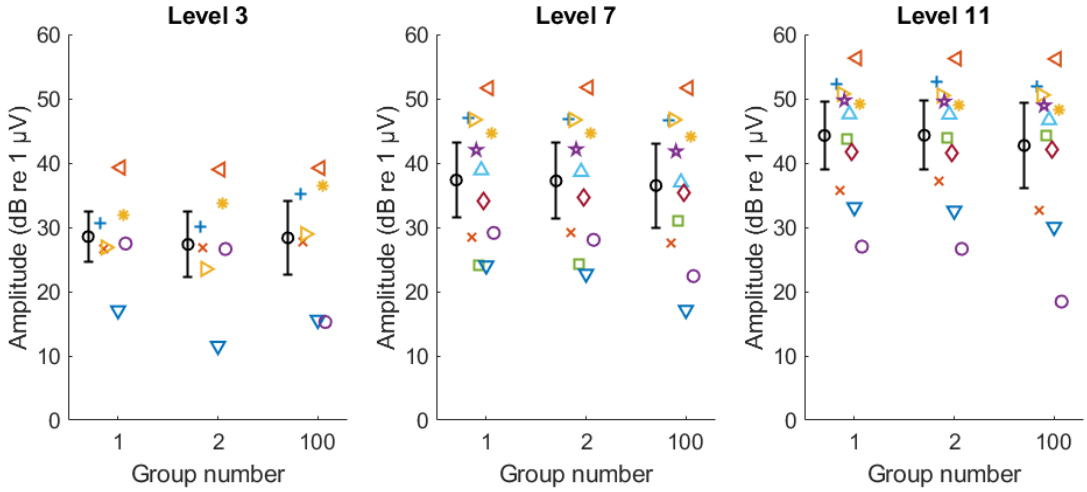
Linear mixed models were fitted to these amplitude data, with current level and time point (group) as fixed factors and participant as random factor. As expected, including current level in the model significantly improved the fit compared to the random intercept model ($\chi^2(2) = 3782$, p < 2e-16). The inclusion of time point also had an effect ($\chi^2(2) = 15.51$, p = 4.28e-4) showing there is a significant change in amplitude as a function of time. However, interestingly, pairwise comparisons showed no change between the first and second time points, for the three tested levels (L3 : t = 1.31, p = 0.240 ; L7 : t = 0.708, p = 0.495 ; L11 : t = -0.199, p = 0.846). This means that the eCAP amplitude remains relatively constant at the very beginning of the measurement. The decrease in response between the beginning and very end of the recording involves longer (order of minutes) time constants.



(a) Evolution for C46 at 30 pps



(b) Evolution for C70 at 30 pps



(c) Response amplitudes to each level for the CLASSIC paradigm at 30 pps.

Figure IV.10 – Evolution of the response over groups of 100 pulses in the CLASSIC measurement at 30 pps : a) & b) examples of temporal evolutions of amplitude over the duration of the measurement for two participants, one with changes in amplitudes (a) and one without any evolution (b) ; c) amplitude evolution over the duration of the measurement for the first, second and last groups for all participants (one marker per participant, see figure IV.4), with the mean and 95 % confidence interval presented in black next to the individual results. Pulses in group 1 were recorded between 0 and 3.3 s after the start of the measurement, group 2 was 3.3 to 6.6 s after the start, and group 100 330 to 333.3 s (5 minutes 33 s, last group).

Hence, there is a change in amplitude over time but it seems to be slow. Therefore, we assume in the following that the mean response to the first 100 pulses is a good approximation of the response to the very first pulse of any stimulus presented in the various paradigms. The eCAP amplitude of the first group of the response to the 30-pps CLASSIC paradigm was thus defined as the reference for normalization of other eCAP responses in the rest of the temporal analysis.

IV.3.5.2 Changes at 300 pps

Because the AGF shift obtained for the HPR paradigms were largest at 300 pps, here we focus on the temporal evolution of the response at this rate. The time course of adaptation was investigated for all paradigms at three different levels (L3, L7 and L11), over 33 s. Each point consisted of the eCAP amplitude measured over a group of 100 pulses. To compare evolutions between levels and subjects, all values were normalized by the reference defined earlier : the eCAP amplitude of the first group of 100 pulses in the CLASSIC paradigm at 30 pps.

Figures IV.11a to IV.11c show three examples of temporal evolutions at 300 pps, for L7 or L11 in the CLASSIC and HPR paradigms. For the majority of the recordings, there was an initial decrease in the normalized amplitude, followed either by a plateau (figure IV.11a) or by a secondary, longer-term decrease (figure IV.11b). However some variability was also observed across participants, with some conditions showing less common and more surprising behaviors (e.g. figure IV.11c). The complete classification of the observed profiles and figures can be found in appendix C.3 (table C.1).

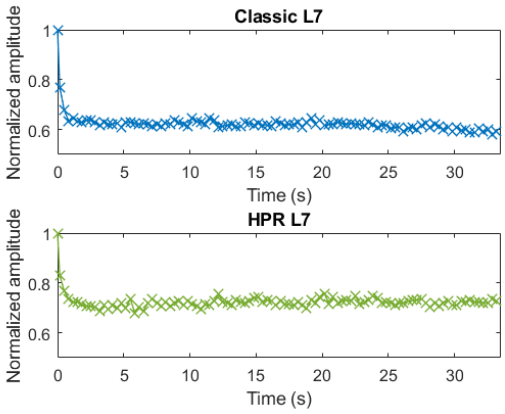
The adaptation index shown in figure IV.11d corresponds to the value of the normalized eCAP amplitude of the last (100th) group in the measurement. These adaptation indices were analyzed using a two-way repeated measures ANOVA with factors paradigm (3 levels : CLASSIC, HPR and RANDOM) and pulse level (2 levels : L7 and L11). The statistical analysis was only run on L7 (HPR2) and L11 (HPR3) levels, because of the lack of points for L3 (HPR1).

There was an effect of the paradigm on the adaptation index ($F_{(2,20)} = 11.085$, $p = 0.006$), but no effect of the level ($F_{(1,10)} = 0.300$, $p = 0.596$) or interaction ($F_{(2,20)} = 0.559$, $p = 0.580$). Note the absence of an effect of level was expected because the amplitudes were normalized.

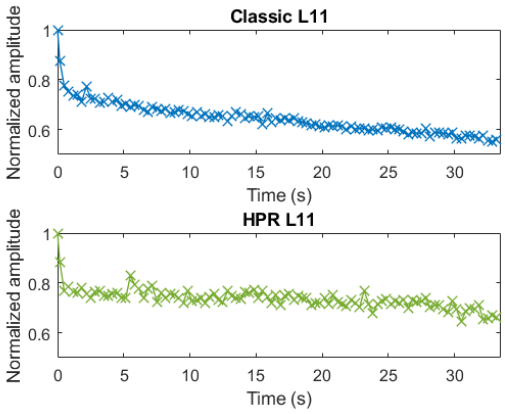
Pairwise comparisons showed adaptation indices to be lower (more adaptation) for CLASSIC than for HPR ($t = -3.82$, $p = 0.003$) and for CLASSIC than for RANDOM ($t = -3.91$, $p = 0.002$), as well as lower for HPR than for RANDOM ($t = -2.80$, $p = 0.032$). Of the tested paradigms, the CLASSIC paradigm thus induced the strongest adaptation, and the RANDOM paradigm the weakest, as hypothesized.

The values of adaptation index higher than 1 mean that the response amplitude was sometimes larger than the reference. This was caused either by the response increasing over time, or by it being very noisy.

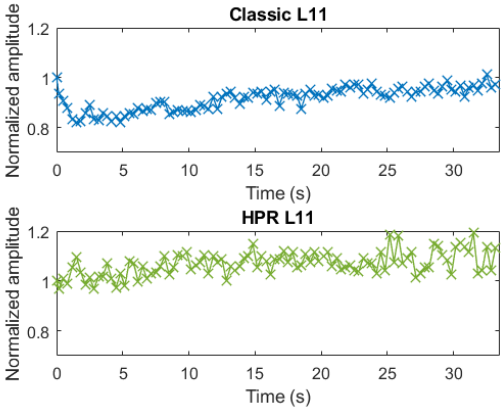
Figure IV.12 presents the individual values of the estimated time constant τ at 300 pps with mean and standard error. No obvious difference between levels or paradigms was observed. The τ values obtained for the CLASSIC and HPR paradigms were mostly of a few hundreds of ms (e.g., for L7 : CLASSIC = 0.222 s, HPR = 0.165 s ; for L11 : CLASSIC = 0.227 s, HPR = 0.160 s).



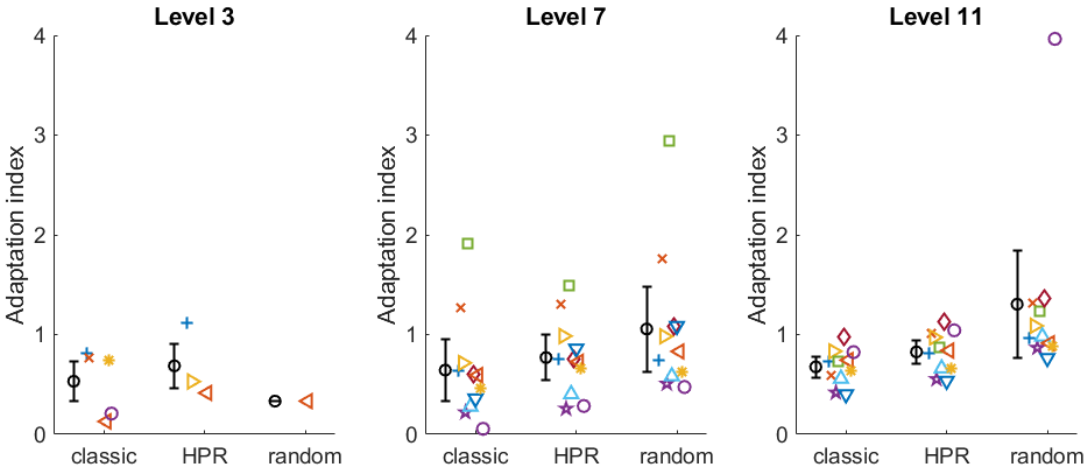
(a) Evolution for C70



(b) Evolution for C15



(c) Evolution for C17 (shifted scale)



(d) Adaptation index for all participants.

Figure IV.11 – Temporal evolution of the normalized amplitude for L3, L7 and L11 at 300 pps : a), b), & c) examples of temporal evolutions of the normalized amplitude over the duration of the measurement for different participants in both the CLASSIC and HPR measurements : one with an initial decrease only (a), one with a second decreasing slope (b), and one with an increase after the first decrease (c) ; d) adaptation index (normalized amplitude of the last pulse group) for all participants (see figure IV.4) and paradigms, with mean and 95 % confidence interval in black.

Time constants for the RANDOM paradigm were not computed because there was little to no decrease in most cases, the values of adaptation index differed from zeros mostly because of the variability of the amplitude response and the possible misevaluation of the reference (see appendix C.3).

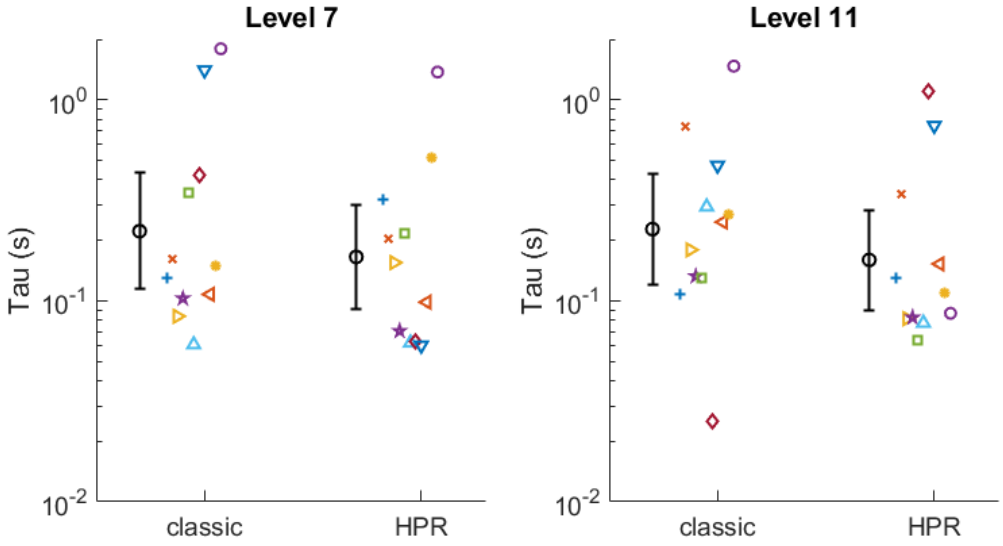


Figure IV.12 – Time constant τ values at 300 pps for all participants with mean and 95 % confidence interval in black (one symbol per participant, see figure IV.4). τ was estimated as the time needed for a measurement to reach 36.8 % of its final value.

There were large variations between participants and paradigms in the obtained τ values. We classified informally the different trends influencing the values of τ. For the participants with a secondary decrease in amplitude (figure IV.11b) or small, longer-term changes only, the value of τ could sometimes take values of more than 1 s (for L7 : 2 participants in CLASSIC, 1 in HPR ; for L11 : 1 participant in CLASSIC, 1 in HPR). The opposite trend was also present : for the participants with a secondary increase (figure IV.11c), or little to no adaptation, the value of τ was typically lower than 70 ms (for L7 : 1 participant in CLASSIC, 3 in HPR ; for L11 : 1 in CLASSIC, 1 in HPR).

IV.3.6 Effect of pulse order

The average difference between even and odd pulses is plotted in figure IV.13 for the HPR2 paradigm at 300 pps (see analyses in section IV.2.7.6).

For most participants, the difference remained small. Furthermore, there was no consistent difference across participants. One exception was participant C26. This was due to its eCAP amplitudes being very noisy with small amplitudes. To conclude, the shift in AGF we observed do not seem to be influenced by eCAP alternation.

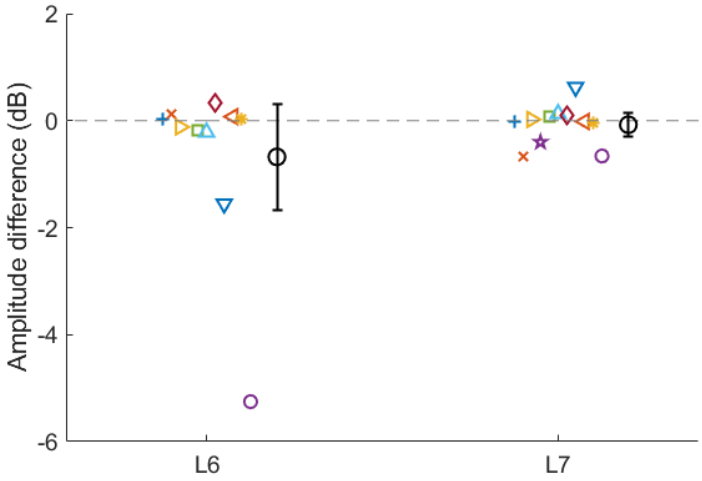


Figure IV.13 – Average difference in response amplitude between even and odd pulses in the HPR2 paradigm at 300 pps. All participants are represented, along with mean differences and 95 % confidence interval (black).

IV.3.7 Effect of preceding level

A last analysis was performed to evaluate the potential influence of the level of the pulse immediately preceding (n-1) the pulse for which we measure the eCAP amplitude (n). As our RANDOM and HPR paradigms involve random presentation of the levels, and because we see evolutions of the response that depend on the paradigm and HPR levels, we wanted to know if local effects had an influence on our results (see analyses in section IV.2.7.6). The following analysis was initially restricted to the HPR2 results at 300 pps because the pulses in this HPR are in the middle of the level range that was measured, so that there are as many preceding pulses with a lower level as preceding pulses with a higher level.

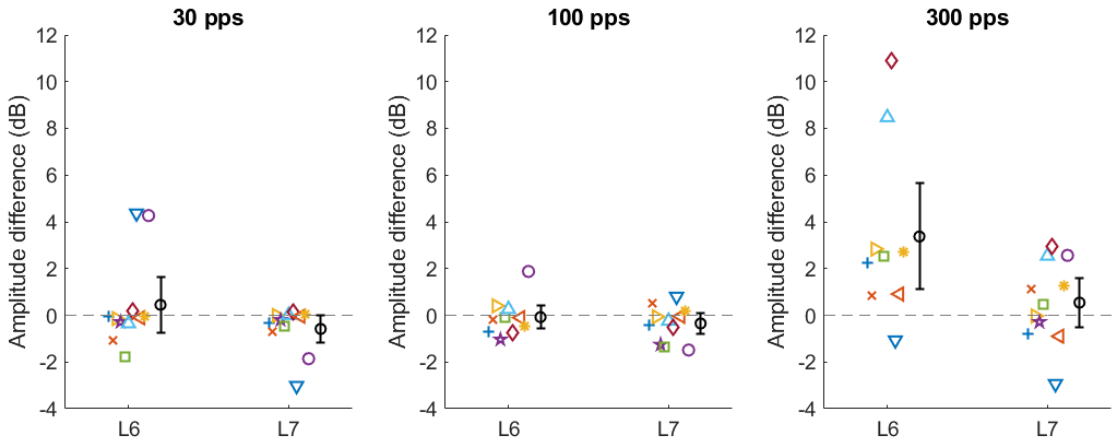


Figure IV.14 – Average difference in response amplitude for HPR2 (L6 and L7) between preceding pulses with higher amplitude and with lower amplitude, at 30, 100 and 300 pps. All participants are represented, along with mean differences and 95 % confidence interval (black).

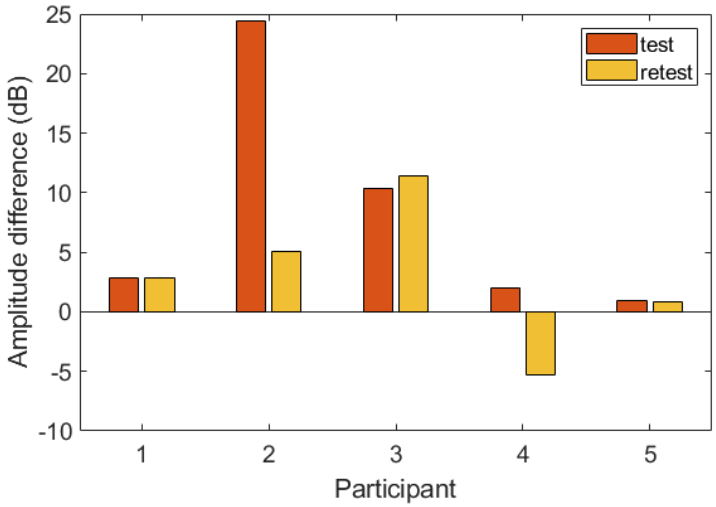


Figure IV.15 – Difference between response after a higher- and a lower-level pulse for the HPR2 paradigm at 300 pps, for the five participants who had values for both (C15, C18, C19, C35 and C70).

The difference between the mean eCAP response obtained after high-level pulses and that obtained after low-level pulses is plotted on the right of figure IV.14 for both levels at 300 pps.

The response amplitude was significantly higher for a higher level preceding pulse if the probe was at L6 ($t = 2.64$, $p = 0.030$), but not if it was at L7 ($t = 1.01$, $p = 0.337$). This is surprising because we were initially expecting the opposite effect : that a high-level preceding pulse would place more fibers in a refractory state and that these fibers would not be able to fire in response to the probe pulse. Inspection of the data at 30 and 100 pps revealed that this effect was not present at lower rates. Could it be a statistical fluke ? Figure IV.15 shows the data at 300 pps for test and retest for five participants for whom test and retest data were available for these conditions : two of them show relatively large differences, but the general trend remained. To conclude, at 300 pps, there was an unexpected effect of the level of the preceding pulse on the eCAP amplitude that we will further discuss in section IV.4.2.

To investigate if this local effect was specific to the HPR paradigm, the eCAP amplitudes obtained in the RANDOM paradigm were also analyzed in a similar way. The results are presented in figure IV.16.

Amplitudes were fitted with a linear-mixed model, with the participant as random factor, and fixed factors pulse level (from L2 to L11), pulse rate (30, 100 and 300 pps) and preceding level type (higher or lower). The analysis revealed a significantly better fit when including the pulse level ($\chi^2(1) = 744.2$, $p < 2.2\text{e-}16$), the rate ($\chi^2(2) = 20.39$, $p = 3.74\text{e-}5$), and the interaction between rate and level ($\chi^2(2) = 25.70$, $p = 2.62\text{e-}6$), but not the preceding level or the interactions between preceding level and rate and/or probe level. Pairwise comparisons at 300 pps (table IV.3) indicate that the difference between higher- and lower-level preceding pulse is significant for L3, L10 and L11. At low levels, the effect is the same as that observed for L6 in the HPR2 ; however, no effect was observed at medium levels in the RANDOM paradigm.

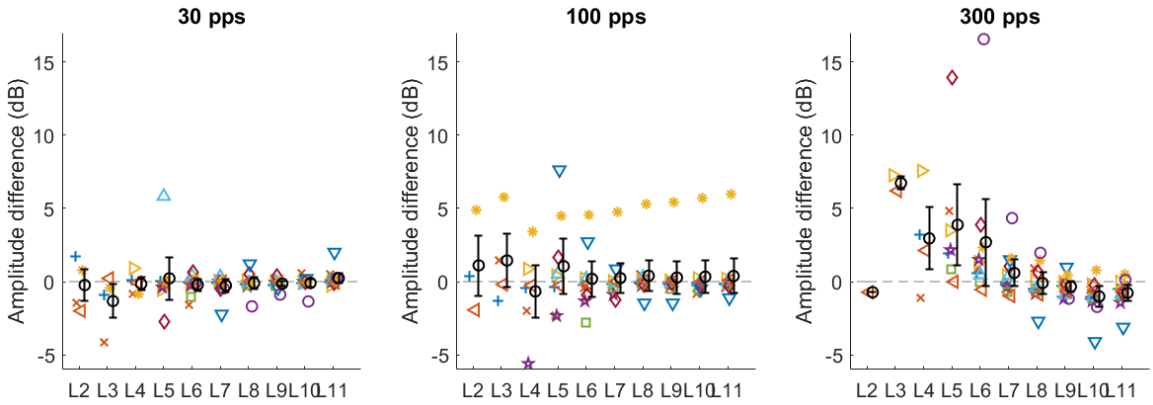


Figure IV.16 – As IV.14, but for the RANDOM paradigm and levels L2 to L11.

Level	t	p	significance
3	12.7	0.050	*
4	1.65	0.198	ns
5	2.18	0.072	ns
6	1.69	0.125	ns
7	1.31	0.221	ns
8	-0.246	0.811	ns
9	-1.67	0.125	ns
10	-2.74	0.021	*
11	-2.68	0.023	*

Table IV.3 – Pairwise comparisons between responses after a higher-level pulse and a lower-level one, for levels L3 to L11 in the RANDOM paradigm at 300 pps.

To summarize, the response to L6 in the HPR2 at 300 pps was lower when preceded by a lower-level pulse. For the RANDOM paradigm at 300 pps, the response to low-level probes was lower after lower-level preceding pulses than after higher-level pulses ; for high level probes, the response is higher after lower-level preceding pulses. Finally, the preceding levels had no effect at the other rates.

IV.4 Discussion

IV.4.1 Dynamic range adaptation at the level of the AN

IV.4.1.1 AGFs shift depending on the most probable levels

The main finding of this work is that the shape of the eCAP AGF depends on the current level distribution of the pulses (see section IV.3.3). In response to the HPR paradigm, for which a narrow range of levels is presented 80 % of the time, the AGF shifts towards the mean level of the HPR. This shift is larger at high pulse rates and amounts to 0.23 dB/dB at 300 pps (figure IV.6e).

We also found that the AGFs were steeper in the HPR2 and HPR3 paradigms than in the RANDOM paradigm when compared at the same current levels (corresponding

to the high-probability level of the HPR paradigms), and for the HPR3 when compared at the same eCAP amplitude (figure IV.9a).

This last observation indicates that the shift is neither purely horizontal nor purely vertical, at least for the HPR3 AGF compared to the RANDOM AGF. This observation bears some similarity with data obtained in AN fibers in response to acoustic stimulation. Wen et al. (2012) showed these data to be consistent with a dual-adaptation model, combining a vertical shift caused by classical spike-rate adaptation, and a horizontal shift caused by DRA. It is however difficult to compare directly individual neural responses to eCAP recordings, as the latter provide no information on the activity of single fibers. The differences between these two types of measurements will be further discussed in section IV.4.1.2.

The fact that the slopes were steeper for the HPR than for the RANDOM paradigms when compared at the same current level is consistent with an improved representation of the most probable levels. More specifically, the pulses presented in the HPR2 and RANDOM paradigms have the same mean level, but the range of most presented level is much narrower for the HPR2. Hence, a narrower level range induces a steepening of the response function. This would be consistent with observations from Dean et al. (2005) in the inferior colliculus : they found wider level ranges of enhanced sensitivity for wider HPR ranges. This result also echoes the work of Rabinowitz et al. (2011) ; Rocchi and Ramachandran (2020) and Lohse et al. (2020) on stimulus level contrast in the inferior colliculus and auditory cortex. Both studies found a dependence of the sensitivity (gain) of the neurons on the total range of stimulation levels : narrower level ranges yielded steeper response functions.

The estimation of time constants in the present study (figure IV.12) allows for a qualitative comparisons with physiological experiments : they are comparable with the time constants of DRA and classical firing-rate adaptation from Dean et al. (2005) and Wen et al. (2012), that range on average between 160 and 330 ms in the inferior colliculus (for the two sorts of adaptation), between 210 and 260 ms for DRA and between 130 and 270 ms for firing-rate adaptation in the AN.

This finding suggests that the AN may be a contributor to DRA in acoustic hearing.

IV.4.1.2 Differences with animal studies

However, eCAPs do not directly reflect single-fiber activity.

There is evidence from computational models that eCAP amplitude relates to the number of firings of individual fibers (Dong et al. 2020). Nevertheless, the eCAP remains a compound response, which is affected by synchrony (the response will be smaller if fibers do not fire at the same time), and which may in some cases not only reflect orthodromic spike propagation but also antidromic activity (Miller et al. 2004). Moreover, there might be additional differences caused by degradation of neural health, which may saturate the neurons and prevent DRA, and distance from the electrodes to the neurons that would reduce the eCAP amplitude, but we don't expect these parameters to affect our results.

An important difference between our data and physiological studies of DRA is that we did not conclusively show that the shift in the AGF translated into an improvement in intensity coding. Despite the increased AGF slope in the HPR region, sensitivity

indices were not different between the RANDOM and HPR paradigms (figure IV.9b, see section IV.3.4). Wen et al. (2009) computed the same sensitivity index, and found greater sensitivity for the HPR levels than for the rest of the tested level range.

It is nevertheless possible that the contribution of the AN to DRA is limited. In contrast, as mentioned in the introduction, other mechanisms may contribute to DRA in acoustic stimulation, such as the hair cell-auditory nerve synapse and the efferent system. The results from the model in Rabinowitz et al. (2013) indeed suggests that this steepening of the response function of neurons remains weak in the AN, compared to the inferior colliculus or auditory cortex : the amount of adaptation increases when going up the auditory system. Wen et al. (2009) also found smaller amounts of DRA in the AN than Dean et al. (2005) in the inferior colliculus.

IV.4.2 Effects of level variations

Since the development of reverse telemetry in cochlear implants, there have been many attempts at correlating various aspects of sound perception in cochlear implant users with objective measures of auditory nerve activity : absolute thresholds, amplitude modulation detection, speech perception, neural survival (McKay et al. 2013 ; De Vos et al. 2018 ; Garcia et al. 2021 ; Skidmore et al. 2022)... Mostly with regular pulse trains of equal current levels, or simple masker-probe pairs (see footnote 2), and rates ranging from a few pulses per second to several thousands.

Our RANDOM paradigm is probably closer to the everyday experience of CI users than the CLASSIC paradigm, that is used in the operating room and in clinics. Since there are several differences between these two paradigms, we can raise some questions on the separation of adaptation and AN responsiveness in studies using a CLASSIC paradigm at high rates. In the RANDOM paradigm, there is no or little adaptation (figure IV.11, see section IV.3.5.2), so the eCAP amplitudes may more closely reflect AN responsiveness with limited neural adaptation. Furthermore, CLASSIC paradigms require low rates to avoid adaptation : the RANDOM paradigm, which is less affected by adaptation, could potentially be performed at higher pulse rates with the same number of repetitions, thereby shortening the time needed to collect AGFs in patients.

Few studies have used pulse trains of varying amplitude in electrical stimulation. Jeng et al. (2009) and Hu et al. (2010) used sinusoidally amplitude modulated pulse trains to measure the effects of the modulation on the eCAP in the cat AN. They played 400-ms pulse trains at 5000 pps, with an amplitude modulation of variable depth at 100 Hz. Their results displayed more adaptation for unmodulated, low-level pulse trains, and neurons with low firing rates at the start of the measurement. The AN fiber threshold increased over time, inducing a shift of the AGF : the N1-P2 amplitude evolved over repetitions of the same modulated pulse train. The high rate and short duration they used contrast with the present study, but the adaptation observed in their data is similar to ours. It also appears that we have more adaptation for the CLASSIC than for the other paradigms that are more modulated. Moreover both studies by Jeng et al. (2009) and Hu et al. (2010) showed a decay of the amplitudes over 400 ms, with a first rapid initial decay for the unmodulated signal.

In human, Tejani et al. (2017) reached similar conclusions. They played 4000-pps amplitude-modulated pulse trains lasting at most 15 ms to 8 CI users. The amplitude

modulations varied from 125 to 1000 Hz. They recorded the response to each pulse of the pulse train during at least two modulations cycles (or more at higher modulation frequencies) : they used the preceding pulses as masker in a masker-probe paradigm to remove the artifact (see footnote 2), and progressively increased the duration of the pulse train ; the masker duration thus depended on the pulse of interest. As in the animal studies (Jeng et al. 2009 ; Hu et al. 2010), they found greater responses at high modulation frequencies, and more adaptation for unmodulated pulse trains. They hypothesized that it was because the changes in level between pulses were larger for high modulation frequencies, thus inducing less adaptation.

These studies demonstrate some asymmetry in the modulated response, with higher amplitudes after a lower- than after a higher-level pulse. This is consistent with the effect of the level we observed at high levels in the RANDOM paradigm, and with the decrease in adaptation index with the increase in the diversity of current levels, between the three paradigms.

The effect of the level of the immediately preceding pulse on the eCAP amplitudes collected in the RANDOM paradigm (figure IV.16) shows that the RANDOM paradigm is not exempt from interactions between pulses.

In the RANDOM paradigm at 300 pps, the lower eCAP amplitude recorded in response to high-level pulses (e.g. L10 and L11) after a higher-level pulse compared to after a lower-level pulse might be due to refractoriness (Boulet et al. 2016) : the response to a preceding higher-level pulse puts more fibers in a refractory state, and less fibers are available to fire in response to the next high-level pulse.

However, the eCAP amplitude in response to a low-level pulse (e.g. L3) was found to be lower if the preceding pulse has a lower level than if it has a higher level. This observation is opposite to our predictions based on an effect of refractoriness. One possible explanation is that this effect is caused by accommodation (Miller et al. 2011 ; Boulet et al. 2016). The inter-pulse interval of 3.3 ms in the present study is consistent with the range of time intervals where accommodation is observed or predicted by computational models. The effect of the previous level observed at 300 pps in the HPR2 paradigm is similar (figures IV.14 and IV.15) : for L6, the eCAP amplitude was on average higher if the previous level had a higher level, but no change was observed for L7. It is possible that since L7 is further in level from the "lower-level" previous pulses than L6, this difference prevents accommodation to occur for L7.

We need to acknowledge that this accommodation explanation is purely speculative but we do not see other potential explanations to this result. Interestingly, accommodation has been observed in animal CI studies as well as predicted by computational models but, to our knowledge, never in human CI listeners. The eCAP measurement paradigms used here provide an interesting opportunity to deepen our understanding of these effects and of their prevalence in CI listening.

IV.4.3 Time constants

The estimation of time constants for the CLASSIC and HPR paradigms in the present study allows for a qualitative comparison with physiological experiments (figure IV.12, see table IV.4) : they are comparable with the time constants of DRA and classical adaptation from Dean et al. (2005) and Wen et al. (2012), summarized in

table IV.5.

The main differences in protocol are that they used acoustic stimuli in individual fibers in anesthetized animals, and that they looked at the time constants for changes between contexts, while we looked at the response at the population level to electrical pulses after a complete absence of stimulation in human, and cannot tease out the contributions of DRA and classical spike rate adaptation.

Present study	CLASSIC		HPR	
Level	L7	L11	L7	L11
τ (ms)	222	227	165	160

Table IV.4 – Time constants evaluated in the present study.

Study	Dean et al. (2005)		Wen et al. (2009)			
Mechanism			Classical adaptation		DRA	
Change in mean level	Increase	Decrease	Increase	Decrease	Increase	Decrease
τ (ms)	160	330	130	270	210	260

Table IV.5 – Time constants of classical adaptation and DRA in the literature.

The long-term change at 30 pps observed in some participants takes place over tens of seconds or even minutes (figure IV.10a) : it resembles the adaptation observed by Clay and Brown (2007) at 300 pps. They studied the evolution of eCAP amplitude for pulses trains at 15, 80 and 300 pps using the masker-probe artefact suppression paradigm, over pulse trains of 5 minutes. They found an increase of adaptation with pulse rate but very little adaptation at 15 pps. Huarte et al. (2014) also tested several pulse rates : they had a baseline rate of 10 pps to study the time scale of the recovery of the response after stimulation at higher rates, and pulse trains at 100, 250 or 495 pps with a masker-probe paradigm for artefact cancellation. They found evolutions over tens of seconds, similar to Clay and Brown (2007), and with different evolution profiles, similar to those we observed (figures IV.11a and IV.11b) - with or without a secondary decrease in amplitude, after a first important amplitude drop. Though both studies did not quantify the time constants of the adaptation they observed, the temporal evolution and adaptation index in the present study are comparable to their findings.

In the present study, the measures at 30 pps also spanned several minutes and showed some, albeit weak, adaptation (figure IV.10a). For 300 pps, the measurements only lasted 33 seconds, but the evolution profiles are very similar to the start of those in the studies from Clay and Brown (2007) and Huarte et al. (2014). On the shorter term changes, few studies have tried to quantify the time constants of short-term adaptation in human CI participants. Hay-McCutcheon et al. (2005) sent pulse trains at 1000 pps with a masker-probe paradigm ; they recorded the responses of trains from 2- to 100-pulses long to assess adaptation. They found a first decrease in amplitude within the first milliseconds of the pulse train, and a rather steady response afterwards, with alternation (changes in amplitude between odd and even pulses).

Animal studies are much more proficient. As an example, Hu et al. (2010) showed in the cat AN that most of the adaptation is taking place in the first 50 ms at 5000 pps, over a 400 ms, amplitude-modulated pulse train. This would be consistent with the first drop we observe in most CLASSIC and HPR measurements at 300 pps (see appendix C.3).

Most studies on alternation display a great drop in amplitude in the first few pulses of the stimulation : results from Hughes et al. (2012) showed a greater drop at high stimulation rates. Huarte et al. (2014) also showed this drop between their baseline at 10 pps and the first recording from their faster pulse trains at 100, 250 and 495 pps. There is thus a very fast adaptation with a time constant of a few ms in eCAP response.

Some of our participants clearly display rapid decreases in amplitude at 300 pps. Very-short-term scales are difficult to analyze with our paradigms and in alternating polarity : a very fast evolution at 300 pps takes place in the first 100 pulses (333 ms), but averaging over less pulses is much noisier and prevents from studying other paradigms than the CLASSIC one. However, the effect of directly preceding level at 300 pps indicates that local, short-term changes affect the results, consistently with the definition of adaptation by Boulet et al. (2016), in addition to longer time scales.

It is important to note that the temporal evolution of RANDOM measurements as implemented in this experiment is harder to track because of the lower number of averages per level in the same time frame as the CLASSIC paradigm, so the evolution of response amplitude over time is noisier. Actually, the RANDOM measurements for some participants display some amplitude changes at 300 pps (initial decrease and plateau), and can be found in appendix C.3.

IV.4.4 Implications for CI coding

It is interesting to note that the concept of dynamic range adaptation has been implemented in cochlear implants and other hearing aids (James et al. 2002) before the first published papers on DRA (Dean et al. 2005). Adaptive dynamic range optimization (ADRO) is a signal processing strategy that intervenes before the speech coding algorithms : it regulates the gain of the different passband filters (Blamey 2005). The end goal is to restore intensity ranges for hearing-impaired, through the selection of the most-informative available channels, and the amplification of the information that falls out of the users range of hearing.

Initially designed for CI-hearing aids combinations, ADRO has proved beneficial for both devices separately (James et al. 2002 ; Blamey 2005) as well as for their combination (Iwaki et al. 2008). It runs a continuous statistical analysis of each frequency channel, and selects the individual level range of interest. No compression is applied if the selected dynamic range is smaller than the level range between threshold and uncomfortable level of the participant. It thus takes individual characteristics and preferences into account.

Its efficiency in speech processing has been assessed with words, sentences from different corpa, in quiet, in multitalker babble and in noise, at different levels... And most studies show a significant improvement of speech recognition performance with ADRO (see review in Blamey 2005 and in the introduction of Iwaki et al. 2008).

Peripheral mechanisms like the IHC-AN synapses and the MOC, which could participate in DRA, are however bypassed in CI. There may be benefits in tuning a signal processing strategy to reflect the DRA observed in acoustic hearing. The analysis window for ADRO processing is generally quite short (2 ms in James et al. (2002) for example), while DRA takes a hundred milliseconds to fully develop. Maybe considering additional, slower time constants, and modifying the magnitude of the dynamic range shift could provide additional benefits on top of those already observed. Furthermore, encoding strategies like ADRO are more flexible than compression.

IV.5 Summary and conclusion

This study provides several new findings about AN responses of human CI users when stimulated with different current levels distributions.

- When using current level distributions with a narrow range of levels presented 80 % of the time, the eCAP amplitude growth function shifted towards the most presented level (mean shift : 0.23 dB/dB).
- The slopes of the amplitude growth functions were steeper in the most presented level range compared to our control condition. However, this increase in slope was concomitant with an increased variability in the eCAP response amplitude. This implies that our data do not conclusively show that the shift of the AGF improves intensity encoding.
- eCAP adaptation, defined as the decrease in response amplitude over time, was found to be strongest when the current level was constant and weakest when the current level showed pulse-to-pulse random variations.
- At 300 pps, the mean time constants of adaptation ranged from 160 to 222 ms depending on the paradigm. These values are comparable to the measured time constants for DRA and classical spike-rate adaptation in animals.
- Overall, these eCAP measures are consistent with the DRA phenomenon observed in previous physiological animal studies and suggests that the AN may be a direct contributor. Future studies will be needed to identify the other potential contributors in acoustic hearing in order to restore this mechanism in CIs.

General conclusion

Summary

The goal of this thesis was to investigate the presence of dynamic range adaptation (DRA) in the human auditory system, through psychophysical and physiological measurements.

In Chapters II and III, we tried to uncover a perceptual correlate of DRA using an intensity discrimination task. The hypothesis was that if the context increases the slope of the rate-level function around its mean level, the intensity discrimination thresholds should be lower. Thus, DRA should yield lower intensity discrimination thresholds if the level of the signal was the same as the context mean level than if the levels were mismatched.

In Chapter II, the test signal was a 50-ms, 2-kHz pure tone standard, played in a narrowband noise context interrupted during the presentation of the standard and probe. The level of the context changed every 50 ms, according to a Gaussian distribution around a mean that was either at 55 or 75 dB SPL ; two standard levels (55 and 75 dB SPL) were tested together in both context mean levels and in silence. Ten normal-hearing participants took part ; all signals were only presented to the left ear.

The results showed a significant interaction between tone level and context level but not consistent with our hypothesis on the perceptual consequences of DRA. The discrimination threshold for the soft standard was lower in the soft than in the loud context ; however, discrimination threshold for the loud standard was also lower for the soft context, which is inconsistent with our initial hypothesis. Several control experiments were run to try to explain why DRA was not highlighted in our experiment. Off-frequency listening was discarded with the addition of notched noise on the task. A task with lower tested levels was run to avoid levels at which the rate-level functions of auditory neurons would be saturated and reduce possible masking effects. It displayed the same tendencies for the soft standard, and a larger discrimination threshold in the loud context than in the soft context for the loud standard - still not in line with our hypothesis on DRA. Silent gaps were added between context and standard in an experiment, and the context was presented diotically in another, to reduce masking effects (e.g., mid-level hump and confusion) ; the results revealed lower discrimination thresholds in the loud context compared to the main experiment, for both standard levels (maybe due to decreasing of masking and/or confusion), but still no trace of DRA. Some energetic masking of the standards by the context, as well as confusion and other masking effects, were certainly preventing the observation of DRA for the loud standard in the main experiment.

The change from pure tone and narrowband to broadband signals provided different results. The intensity discrimination task with a pulse-spreading harmonic complex in a broadband noise got lower discrimination thresholds in the context with the same mean level than the standard level, compared to the other context mean level and

silence. This was true only if the signals bandwidth was 20 to 20 000 Hz : the filtered signals in the control experiment, while several octaves wide, did not show significant differences. The bandwidth and level used in the main experiment are certainly central to the observation of DRA in the intensity discrimination task.

The measurements of auditory nerve activity with cochlear implant in Chapter IV revealed changes in the response patterns depending on the stimulation paradigm, diversity and mean of the presented levels. The sensitivity was however not affected by the paradigm, preventing to conclude on an improvement of the encoding of certain levels. The horizontal shift of the amplitude-growth functions observed at 300 pps was in average of 0.23 dB/dB ; and the estimated time constants for the observed adaptation were in average around 200 ms. These values are consistent with the literature on dynamic range adaptation. Our data also points out that several time scales of adaptation are at play but the different mechanisms are difficult to separate due to the observed differences in response profile. It is possible that the AN contributes directly to dynamic range adaptation, but that it is not translated into an improvement in the encoding of the levels.

This study provides signs of DRA in both physiological and psychophysical results in human.

Perspectives

The results of the intensity discrimination tasks bring new questions.

The results with broadband signals question the importance of the nature of the signal to observe a decrease in discrimination thresholds in the context of mean level equal to that of the standard. Moreover, the control experiments with pure tones indicate that different forms of masking (energetic and informational) are at play. It could be interesting to investigate other ways to reduce the effects of masking, both with the narrowband signals and the filtered broadband signal (since we have no better explanation for these results either). For example, we could take inspiration in the work of Moore and Glasberg (1982) or Oberfeld (2008), and add cues to the contralateral ear or use lateralization.

It is also possible to try and combine changes in contrast in addition of mean level with broadband signals : this would complement the studies on perceptual consequences of adaptation to stimulus statistics, with both DRA and contrast studies (Dean et al. 2005 ; Lohse et al. 2020). While DRA is interested in the mean context level, contrast is about the variance of the levels of the stimuli : low contrasts (small variance) display steeper rate-level functions than high contrasts (with great variance). That would mean varying the width of the distribution of the levels in the context in addition to its mean level. An intensity discrimination task with broadband noise and harmonic complexes could be designed with different context mean levels for DRA, and standard deviations for contrast gain control. We would expect lower thresholds in the low-contrast context with the same mean level as the standard.

The implication of the MOC reflex in the results of the intensity discrimination task with broadband signals could also be evaluated, for example with contralateral and binaural context and broadband signals. This change in protocol should yield the same

trends with these conditions than the main experiment with PSHC (lower thresholds with matching context and test signal level), if the MOC is indeed responsible for the results with ipsilateral broadband noise context (Lilaonitkul and Guinan 2009). On the contrary, thresholds would be less affected by the context with contralateral and binaural stimulation if the MOC is not driving the differences in thresholds between context levels.

The results of the different psychophysical experiments of Chapters II and III bring the need to build models that can account for our observations : neither the rate-level functions and their slopes, nor the sensitivity index calculated with the model could account for the obtained discrimination thresholds.

The possibility to record eCAPs with individual pulse responses invites further research. Other paradigms could be tested : it could be interesting to investigate other level distributions within one track, including wider and narrower level ranges, to extend the results on preceding level effect and its coupling with the changes in shape of the AGFs. Reproducing the switching-HPR paradigm tested in the physiological studies, for which the HPR is changed every 5 seconds, could also provide complementary data on temporal evolution of the adaptation we observe with the eCAP. It could also be possible to optimize the range of tested levels to observe adaptation over a few pulses only.

Another interesting topic would be to look further into the adaptation observed for some participants in the RANDOM paradigm, with longer measurements at high rates for example.

An additional prospect could be to design an intensity discrimination task for CI users with different pulse rates and most probable level ranges to compare the effect of adaptation. We could use stimuli similar to those used in our study with CI users, and even measure eCAPs during the task. Kreft et al. (2004a,b) have shown that increases in pulse rate increase the dynamic range of the participants and the maximum acceptable level. Over the tested rates (from 200 to 6500 pps), the number of 'discriminable steps' does not change : intensity discrimination thresholds increase with pulse rate. It could be interesting to complete these studies by testing intensity discrimination using lower pulses rates, for which there is little to no adaptation in the auditory nerve. Moreover, these results could be extended further by adding a continuous context similar to the HPR paradigms in the same way as in the psychophysical experiments in this manuscript. Discrimination thresholds with a context could be measured using different rates (for example 30 and 300 pps, to monitor adaptation) and level distributions for the context, to explore the effect of DRA on intensity discrimination thresholds in CI users. Discrimination thresholds at (very) low rates should be less affected by the introduction of the context than higher rates, and might even have more 'discriminable steps' than found in the literature at higher rates. This could also comfort the electrophysiological results : observing the same trends as in Chapter III with cochlear implant users at high pulse rates or with other participants with sensorineural hearing loss for example would comfort the idea that DRA is implicated and originates in the AN.

Finally, the use of auditory brainstem response (ABR) to record more central activity with the stimuli used in the intensity discrimination tasks could provide insight on DRA at higher levels in the auditory system, between the AN response and intensity

discrimination performance. It could also offer insights with more central responses on the results of the task, because it has been established here that we could not rely on modeled auditory nerve activity only.

References

Adel Y., Hilkhuysen G., Noreña A., Cazals Y., Roman S. and Macherey O. 2017, Forward Masking in Cochlear Implant Users: Electrophysiological and Psychophysical Data Using Pulse Train Maskers, Journal of the Association for Research in Otolaryngology, 18(3), 495–512

ANSI. 2007, ANSI S3.4-2007 - Procedure for the Computation of Loudness of Steady Sounds

Arieh Y. and Marks L. E. 2003, Time course of loudness recalibration: Implications for loudness enhancement, The Journal of the Acoustical Society of America, 114(3), 1550–1556

Backus B. C. and Guinan J. J. 2006, Time-course of the human medial olivocochlear reflex, The Journal of the Acoustical Society of America, 119(5), 2889–2904

Baer T., Moore B. C. and Marriage J. 2001, Detection and intensity discrimination of brief tones as a function of duration by hearing-impaired listeners, Hearing Research, 159(1-2), 74–84

Barbour D. L. 2011, Intensity-invariant coding in the auditory system, Neuroscience & Biobehavioral Reviews, 35(10), 2064–2072

Baudhuin J. L., Hughes M. L. and Goehring J. L. 2016, A Comparison of Alternating Polarity and Forward Masking Artifact-Reduction Methods to Resolve the Electrically Evoked Compound Action Potential, Ear & Hearing, 37(4), e247–e255

Benda J. 2021, Neural adaptation, Current Biology, 31(3), R110–R116

Berliner J. E. and Durlach N. I. 1973, Intensity perception. IV. Resolution in roving-level discrimination, The Journal of the Acoustical Society of America, 53(5), 1270–1287

Berliner J. E., Durlach N. I. and Braida L. D. 1977, Intensity perception. VII. Further data on roving-level discrimination and the resolution and bias edge effects, The Journal of the Acoustical Society of America, 61(6), 1577–1585

Blamey P. J. 2005, Adaptive Dynamic Range Optimization (ADRO): A Digital Amplification Strategy for Hearing Aids and Cochlear Implants, Trends in Amplification, 9(2), 77–98

Boulet J., White M. and Bruce I. C. 2016, Temporal Considerations for Stimulating Spiral Ganglion Neurons with Cochlear Implants, Journal of the Association for Research in Otolaryngology, 17(1), 1–17

Boullet I., Marozeau J. and Meunier S. 2004, *Temporal summation of loudness as a function of frequency and temporal pattern*, in 33rd International Congress and Exposition on noise control engineering (Internoise), Prague, Czech Republic

Brennan M. A., Svec A., Farhadi A., Maxwell B. N. and Carney L. H. 2023, Inherent envelope fluctuations in forward masking: Effects of age and hearing loss, The Journal of the Acoustical Society of America, 153(4), 1994

Brochier T., McKay C. M. and Carlyon R. P. 2021, Interpreting the Effect of Stimulus Parameters on the Electrically Evoked Compound Action Potential and on Neural Health Estimates, Journal of the Association for Research in Otolaryngology, 22(1), 81–94

Buss E. 2008, The effect of masker level uncertainty on intensity discrimination, The Journal of the Acoustical Society of America, 123(1), 254–264

Butler B. E. and Lomber S. G. 2013, Functional and structural changes throughout the auditory system following congenital and early-onset deafness: implications for hearing restoration, Frontiers in Systems Neuroscience, 7

Buus S., Florentine M. and Poulsen T. 1997, Temporal integration of loudness, loudness discrimination, and the form of the loudness function, The Journal of the Acoustical Society of America, 101(2), 669–680

Canevet G. and Meunier S. 1996, Effect of Adaptation on Auditory Localization and Lateralization, Acta Acustica united with Acustica, 82

Carlyon R. P. 1987, A release from masking by continuous, random, notched noise, The Journal of the Acoustical Society of America, 81(2), 418–426

Carlyon R. P. 1989, Changes in the masked thresholds of brief tones produced by prior bursts of noise, Hearing Research, 41(2-3), 223–235

Carlyon R. P. and Beveridge H. A. 1993, Effects of forward masking on intensity discrimination, frequency discrimination, and the detection of tones in noise, The Journal of the Acoustical Society of America, 93(5), 2886–2895

Carlyon R. P. and Datta A. J. 1997, Excitation produced by Schroeder-phase complexes: Evidence for fast-acting compression in the auditory system, The Journal of the Acoustical Society of America, 101(6), 3636–3647

Carlyon R. P. and Deeks J. M. 2015, Combined neural and behavioural measures of temporal pitch perception in cochlear implant users, The Journal of the Acoustical Society of America, 138(5), 2885–2905

Carlyon R. P., Long C. J. and 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

Carlyon R. P. and Moore B. C. J. 1984, Intensity discrimination: A severe departure from Weber's law, The Journal of the Acoustical Society of America, 76(5), 1369–1376

Carlyon R. P. and Moore B. C. J. 1986, Continuous versus gated pedestals and the "severe departure" from Weber's law, The Journal of the Acoustical Society of America, 79(2), 453–460

Carney L. H. 2018, Supra-Threshold Hearing and Fluctuation Profiles: Implications for Sensorineural and Hidden Hearing Loss, Journal of the Association for Research in Otolaryngology, 19(4), 331–352

Carney L. H. 2024, Neural Fluctuation Contrast as a Code for Complex Sounds: The Role and Control of Peripheral Nonlinearities, Hearing Research, 443, 108966

Ciuman R. R. 2010, The efferent system or olivocochlear function bundle - fine regulator and protector of hearing perception, International journal of biomedical science: IJBS, 6(4), 276–288

Clay K. M. S. and Brown C. J. 2007, Adaptation of the Electrically Evoked Compound Action Potential (ECAP) Recorded from Nucleus CI24 Cochlear Implant Users, Ear & Hearing, 28(6), 850–861

Colburn S. H., Carney L. H. and Heinz M. G. 2003, Quantifying the Information in Auditory-Nerve Responses for Level Discrimination, JARO - Journal of the Association for Research in Otolaryngology, 4(3), 294–311

Costalupes J. A., Young E. D. and Gibson D. J. 1984, Effects of continuous noise backgrounds on rate response of auditory nerve fibers in cat, Journal of Neurophysiology, 51(6), 1326–1344

De Vos J. J., Biesheuvel J. D., Briaire J. J., Boot P. S., Van Gendt M. J., Dekkers O. M., Fiocco M. and Frijns J. H. M. 2018, Use of Electrically Evoked Compound Action Potentials for Cochlear Implant Fitting: A Systematic Review, Ear & Hearing, 39(3), 401–411

Dean I., Harper N. S. and McAlpine D. 2005, Neural population coding of sound level adapts to stimulus statistics, Nature Neuroscience, 8(12), 1684–1689

Dean I., Robinson B. L., Harper N. S. and McAlpine D. 2008, Rapid neural adaptation to sound level statistics, Journal of Neuroscience, 28(25), 6430–6438

Delgutte B. 1990, Two-tone rate suppression in auditory-nerve fibers: Dependence on suppressor frequency and level, Hearing Research, 49(1-3), 225–246

Dong Y., Briaire J. J., Biesheuvel J. D., Stronks H. C. and Frijns J. H. 2020, Unravelling the temporal properties of human eCAPs through an iterative deconvolution model, Hearing Research, 395, 108037

Durlach N. 2006, Auditory masking: Need for improved conceptual structure, The Journal of the Acoustical Society of America, 120(4), 1787–1790

Epstein M. 2007, An introduction to induced loudness reduction, The Journal of the Acoustical Society of America, 122(3), EL74–EL80

Epstein M. and Florentine M. 2012, Binaural loudness summation for speech presented via earphones and loudspeaker with and without visual cues, The Journal of the Acoustical Society of America, 131(5), 3981–3988

Fischenich A., Hots J., Verhey J. and Oberfeld D. 2021, Temporal Loudness Weights Are Frequency Specific, Frontiers in Psychology, 12, 588571

Florentine M. 1983, Intensity discrimination as a function of level and frequency and its relation to high-frequency hearing, The Journal of the Acoustical Society of America, 74(5), 1375–1379

Florentine M. 1986, Level discrimination of tones as a function of duration, The Journal of the Acoustical Society of America, 79(3), 792–798

Florentine M. and Buus S. 1981, An excitation-pattern model for intensity discrimination, The Journal of the Acoustical Society of America, 70(6), 1646–1654

Florentine M., Buus S. and Mason C. R. 1987, Level discrimination as a function of level for tones from 0.25 to 16 kHz, The Journal of the Acoustical Society of America, 81(5), 1528–1541

Gallun F. J. and Hafter E. R. 2006, Amplitude modulation sensitivity as a mechanism for increment detection, The Journal of the Acoustical Society of America, 119(6), 3919–3930

Garcia C., Goehring T., Cosentino S. and al. 2021, The Panoramic ECAP Method: Estimating Patient-Specific Patterns of Current Spread and Neural Health in Cochlear Implant Users, Journal of the Association for Research in Otolaryngology, 22(5), 567–589

Glasberg B. R. and Moore B. C. 1990, Derivation of auditory filter shapes from notched-noise data, Hearing Research, 47(1-2), 103–138

Gockel H., Moore B. C. J. and Patterson R. D. 2002, Influence of Component Phase on the Loudness of Complex Tones, Acta Acustica united with Acustica, 88

Grange J., Zhang M. and Culling J. 2022, The Role of Efferent Reflexes in the Efficient Encoding of Speech by the Auditory Nerve, Journal of Neuroscience, 42(36), 6907–6916

Greenwood D. D. 1990, A cochlear frequency-position function for several species—29 years later, The Journal of the Acoustical Society of America, 87(6), 2592–2605

Guinan J. J. 2006, Olivocochlear Efferents: Anatomy, Physiology, Function, and the Measurement of Efferent Effects in Humans:, Ear and Hearing, 27(6), 589–607

Guinan J. J., Backus B. C., Lilaonitkul W. and Aharonson V. 2003, Medial Olivocochlear Efferent Reflex in Humans: Otoacoustic Emission (OAE) Measurement Issues and the Advantages of Stimulus Frequency OAEs, Journal of the Association for Research in Otolaryngology, 4(4), 521–540

Hay-McCutcheon M. J., Brown C. J. and Abbas P. J. 2005, An analysis of the impact of auditory-nerve adaptation on behavioral measures of temporal integration in cochlear implant recipients, The Journal of the Acoustical Society of America, 118(4), 2444–2457

Heinz M. G., Colburn H. S. and Carney L. H. 2001a, Evaluating Auditory Performance Limits: II. One-Parameter Discrimination with Random-Level Variation, Neural Computation, 13(10), 2317–2338

Heinz M. G., Colburn S. H. and Carney L. H. 2001b, Rate and timing cues associated with the cochlear amplifier: Level discrimination based on monaural cross-frequency coincidence detection, The Journal of the Acoustical Society of America, 110(4), 2065–2084

Herrmann B., Augereau T. and Johnsrude I. S. 2020, Neural Responses and Perceptual Sensitivity to Sound Depend on Sound-Level Statistics, Scientific Reports, 10(1), 9571

Hilkhuysen G. and Macherey O. 2014, Optimizing pulse-spreading harmonic complexes to minimize intrinsic modulations after auditory filtering, The Journal of the Acoustical Society of America, 136(3), 1281–1294

Hirsh I. J. 1948, The Influence of Interaural Phase on Interaural Summation and Inhibition, The Journal of the Acoustical Society of America, 20(4), 536–544

Hu N., Miller C. A., Abbas P. J., Robinson B. K. and Woo J. 2010, Changes in Auditory Nerve Responses Across the Duration of Sinusoidally Amplitude-Modulated Electric Pulse-Train Stimuli, Journal of the Association for Research in Otolaryngology, 11(4), 641–656

Huarte A., Ramos A., Morera C. and al. 2014, Evaluation of neural response telemetry (NRT™) with focus on long-term rate adaptation over a wide range of stimulation rates, Cochlear Implants International, 15(3), 136–144

Hughes M. L. 2013, Objective measures in cochlear implants, Objective measures in cochlear implants (San Diego, CA: Plural Pub.), oCLC: 893575536

Hughes M. L., Castioni E. E., Goehring J. L. and Baudhuin J. L. 2012, Temporal response properties of the auditory nerve: Data from human cochlear-implant recipients, Hearing Research, 285(1-2), 46–57

Iwaki T., Blamey P. and Kubo T. 2008, Bimodal studies using adaptive dynamic range optimization (ADRO) technology, International Journal of Audiology, 47(6), 311–318

Jackson B. S. and Carney L. H. 2005, The Spontaneous-Rate Histogram of the Auditory Nerve Can Be Explained by Only Two or Three Spontaneous Rates and Long-Range Dependence, Journal of the Association for Research in Otolaryngology, 6(2), 148–159

James C. J., Blamey P. J., Martin L., Swanson B., Just Y. and Macfarlane D. 2002, Adaptive Dynamic Range Optimization for Cochlear Implants: A Preliminary Study:, Ear and Hearing, 23(Supplement), 49S–58S

Javel E. 1981, Suppression of auditory nerve responses I: Temporal analysis, intensity effects and suppression contours, The Journal of the Acoustical Society of America, 69(6), 1735–1745

Javel E. 1996, Long-term adaptation in cat auditory-nerve fiber responses, The Journal of the Acoustical Society of America, 99(2), 1040–1052

Javel E., McGee J., Walsh E. J., Farley G. R. and Gorga M. P. 1983, Suppression of auditory nerve responses. II. Suppression threshold and growth, iso-suppression contours, The Journal of the Acoustical Society of America, 74(3), 801–813

Jeng F.-C., Abbas P. J., Hu N., Miller C. A., Nourski K. V. and Robinson B. K. 2009, Effects of temporal properties on compound action potentials in response to amplitude-modulated electric pulse trains in guinea pigs, Hearing Research, 247(1), 47–59

Jesteadt W., Wier C. C. and Green D. M. 1977, Intensity discrimination as a function of frequency and sensation level, The Journal of the Acoustical Society of America, 61(1), 169–177

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

Killion M. C. 1978, Revised estimate of minimum audible pressure: Where is the "missing 6 dB"?, The Journal of the Acoustical Society of America, 63(5), 1501–1508

Kreft H. A., Donaldson G. S. and Nelson D. A. 2004a, Effects of pulse rate and electrode array design on intensity discrimination in cochlear implant users, The Journal of the Acoustical Society of America, 116(4), 2258–2268

Kreft H. A., Donaldson G. S. and Nelson D. A. 2004b, Effects of pulse rate on threshold and dynamic range in Clarion cochlear-implant users (L), The Journal of the Acoustical Society of America, 115(5), 1885–1888

Lai W. 2004, An NRT cookbook: guidelines for making NRT measurements, version 3.1, An NRT cookbook: guidelines for making NRT measurements, version 3.1 (Basel: Cochlear AG), oCLC: 59512915

Levitt H. 1971, Transformed Up-Down Methods in Psychoacoustics, The Journal of the Acoustical Society of America, 49(2B), 467–477

Liberman M. C. 1978, Auditory-nerve response from cats raised in a low-noise chamber, The Journal of the Acoustical Society of America, 63(2), 442–455

Lilaonitkul W. and Guinan J. J. 2009, Human Medial Olivocochlear Reflex: Effects as Functions of Contralateral, Ipsilateral, and Bilateral Elicitor Bandwidths, Journal of the Association for Research in Otolaryngology, 10(3), 459–470

Liu M., Wang Y., Jiang L., Zhang X., Wang C. and Zhang T. 2024, Research progress of the inferior colliculus: from Neuron, neural circuit to auditory disease, Brain Research, 1828, 148775

Lohse M., Bajo V. M., King A. J. and Willmore B. D. B. 2020, Neural circuits underlying auditory contrast gain control and their perceptual implications, Nature Communications, 11(1), 324

Lopez-Poveda E. A. and Meddis R. 2001, A human nonlinear cochlear filterbank, The Journal of the Acoustical Society of America, 110(6), 3107–3118

López-Ramos D., Marrufo-Pérez M. I., Eustaquio-Martín A., López-Bascuas L. E. and Lopez-Poveda E. A. 2024, Adaptation to Noise in Spectrotemporal Modulation Detection and Word Recognition, Trends in Hearing, 28, 23312165241266322

Macherey O. and Carlyon R. P. 2014, Cochlear implants, Current Biology, 24(18), R878–R884

Macherey O., Carlyon R. P., Van Wieringen A., Deeks J. M. and Wouters J. 2008, Higher Sensitivity of Human Auditory Nerve Fibers to Positive Electrical Currents, Journal of the Association for Research in Otolaryngology, 9(2), 241–251

Marozeau J., Epstein M., Florentine M. and Daley B. 2006, A test of the Binaural Equal-Loudness-Ratio hypothesis for tones, The Journal of the Acoustical Society of America, 120(6), 3870–3877

Marrufo-Pérez M. I., Eustaquio-Martín A., Fumero M. J., Gorospe J. M., Polo R., Gutiérrez Revilla A. and Lopez-Poveda E. A. 2019, Adaptation to noise in amplitude modulation detection without the medial olivocochlear reflex, Hearing Research, 377, 133–141

Marrufo-Pérez M. I., Eustaquio-Martín A. and Lopez-Poveda E. A. 2018, Adaptation to Noise in Human Speech Recognition Unrelated to the Medial Olivocochlear Reflex, Journal of Neuroscience, 38(17), 4138–4145

Marrufo-Pérez M. I. and Lopez-Poveda E. A. 2022, Adaptation to noise in normal and impaired hearing, The Journal of the Acoustical Society of America, 151(3), 1741–1753

Marrufo-Pérez M. I., Sturla-Carreto D. D. P., Eustaquio-Martín A. and Lopez-Poveda E. A. 2020, Adaptation to Noise in Human Speech Recognition Depends on Noise-Level Statistics and Fast Dynamic-Range Compression, Journal of Neuroscience, 40(34), 6613–6623

McGill W. J. and Goldberg J. P. 1968, A study of the near-miss involving Weber's law and pure-tone intensity discrimination, Perception & Psychophysics, 4(2), 105–109

McKay C. M., Chandan K., Akhoun I., Siciliano C. and Kluk K. 2013, Can ECAP Measures Be Used for Totally Objective Programming of Cochlear Implants?, Journal of the Association for Research in Otolaryngology, 14(6), 879–890

Meddis R. 2006, Auditory-nerve first-spike latency and auditory absolute threshold: A computer model, The Journal of the Acoustical Society of America, 119(1), 406–417

Meddis R., Lecluyse W., Clark N. R., Jürgens T., Tan C. M., Panda M. R. and Brown G. J. 2013, in *Basic Aspects of Hearing*, ed. B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon, & H. E. Gockel, Vol. 787 (New York, NY: Springer New York), 11–20, series Title: Advances in Experimental Medicine and Biology

Mesik J. and Wojtczak M. 2020, Effects of noise precursors on the detection of amplitude and frequency modulation for tones in noise, The Journal of the Acoustical Society of America, 148(6), 3581–3597

Mesnildrey Q., Hilkhuysen G. and Macherey O. 2016, Pulse-spreading harmonic complex as an alternative carrier for vocoder simulations of cochlear implants, The Journal of the Acoustical Society of America, 139(2), 986–991

Meunier S. and Marchioni A. 1998, Effect of Mid-Term Adaptation on Pure-Tone Detection, Acta Acustica united with Acustica, 84(3), 503–512

Micheyl C., Perrot X. and Collet L. 1997, Relationship between auditory intensity discrimination in noise and olivocochlear efferent system activity in humans., Behavioral Neuroscience, 111(4), 801–807

Miller C. A., Abbas P. J., Hay-McCutcheon M. J., Robinson B. K., Nourski K. V. and Jeng F.-C. 2004, Intracochlear and extracochlear ECAPs suggest antidromic action potentials, Hearing Research, 198(1-2), 75–86

Miller C. A., Woo J., Abbas P. J., Hu N. and Robinson B. K. 2011, Neural Masking by Sub-threshold Electric Stimuli: Animal and Computer Model Results, Journal of the Association for Research in Otolaryngology, 12(2), 219–232

Miller G. A. 1947, Sensitivity to Changes in the Intensity of White Noise and Its Relation to Masking and Loudness, The Journal of the Acoustical Society of America, 19(4), 609–619

Moore B. C. J. 1981, Interactions of masker bandwidth with signal duration and delay in forward masking, The Journal of the Acoustical Society of America, 70(1), 62–68

Moore B. C. J. and Glasberg B. R. 1982, Contralateral and ipsilateral cueing in forward masking, The Journal of the Acoustical Society of America, 71(4), 942–945

Moore B. C. J. and Glasberg B. R. 1993, Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech, The Journal of the Acoustical Society of America, 94(4), 2050–2062

Moore B. C. J., Glasberg B. R. and Baer T. 1997, A Model for the Prediction of Thresholds, Loudness, and Partial Loudness, Journal of the Audio Engineering Society, 45(4), 224–240

Moore B. C. J., Glasberg B. R., Hess R. F. and Birchall J. P. 1985, Effects of flanking noise bands on the rate of growth of loudness of tones in normal and recruiting ears, The Journal of the Acoustical Society of America, 77(4), 1505–1513

Moore B. C. J., Glasberg B. R., Varathanathan A. and Schlittenlacher J. 2016, A Loudness Model for Time-Varying Sounds Incorporating Binaural Inhibition, Trends in Hearing, 20, 233121651668269

Moore B. C. J., Jervis M., Harries L. and Schlittenlacher J. 2018, Testing and refining a loudness model for time-varying sounds incorporating binaural inhibition, The Journal of the Acoustical Society of America, 143(3), 1504–1513

Moore B. C. J. and Raab D. H. 1974, Pure-tone intensity discrimination: some experiments relating to the "near-miss" to Weber's law, The Journal of the Acoustical Society of America, 55(5), 1049–1054

Moore B. C. J. and Raab D. H. 1975, Intensity discrimination for noise bursts in the presence of a continuous, bandstop background: effects of level, width of the bandstop, and duration, The Journal of the Acoustical Society of America, 57(2), 400–405

Mori S. and Ward L. M. 1992, Intensity and frequency resolution: Masking of absolute identification and fixed and roving discrimination, The Journal of the Acoustical Society of America, 91(1), 246–255

Nagel K. I. and Doupe A. J. 2006, Temporal Processing and Adaptation in the Songbird Auditory Forebrain, Neuron, 51(6), 845–859

Necciari T., Laback B., Savel S., Ystad S., Balazs P., Meunier S. and Kronland-Martinet R. 2016, Auditory Time-Frequency Masking for Spectrally and Temporally Maximally-Compact Stimuli, PLoS ONE, 11(11), e0166937

Neff D. L. 1985, Stimulus parameters governing confusion effects in forward masking, The Journal of the Acoustical Society of America, 78(6), 1966–1976

Neff D. L. 1986, Confusion effects with sinusoidal and narrow-band noise forward maskers, The Journal of the Acoustical Society of America, 79(5), 1519–1529

Oberfeld D. 2007, Loudness changes induced by a proximal sound: Loudness enhancement, loudness recalibration, or both?, The Journal of the Acoustical Society of America, 121(4), 2137–2148

Oberfeld D. 2008, The mid-difference hump in forward-masked intensity discrimination, The Journal of the Acoustical Society of America, 123(3), 1571–1581

Oberfeld D. and Stahn P. 2012, Sequential Grouping Modulates the Effect of Non-Simultaneous Masking on Auditory Intensity Resolution, PLoS ONE, 7(10), e48054

Oberfeld D., Stahn P. and Kuta M. 2012, Binaural release from masking in forward-masked intensity discrimination: Evidence for effects of selective attention, Hearing Research, 294(1-2), 1–9

Oberfeld D., Stahn P. and Kuta M. 2014, Why Do Forward Maskers Affect Auditory Intensity Discrimination? Evidence from "Molecular Psychophysics", PLoS ONE, 9(6), e99745

Olson E. S. and Mountain D. C. 1991, *In vivo* measurement of basilar membrane stiffness, The Journal of the Acoustical Society of America, 89(3), 1262–1275

Oxenham A. J. 2018, Loudness Context Effects and Auditory Enhancement in Normal, Impaired, and Electric Hearing, Acta Acustica united with Acustica, 104(5), 839–843

Oxenham A. J. and Buus S. 2000, Level discrimination of sinusoids as a function of duration and level for fixed-level, roving-level, and across-frequency conditions, The Journal of the Acoustical Society of America, 107(3), 1605–1614

Pannese A., Grandjean D. and Frühholz S. 2015, Subcortical processing in auditory communication, Hearing Research, 328, 67–77

Pienkowski M. and Hagerman B. 2009, Auditory intensity discrimination as a function of level-rove and tone duration in normal-hearing and impaired subjects: The “mid-level hump” revisited, Hearing Research, 253(1-2), 107–115

Plack C. J. 1996, Loudness enhancement and intensity discrimination under forward and backward masking, The Journal of the Acoustical Society of America, 100(2), 1024–1030

Plack C. J. 2013, The Sense of Hearing, The Sense of Hearing, 3rd edn. (London: Routledge)

Plack C. J., Carlyon R. P. and Viemeister N. F. 1995, Intensity discrimination under forward and backward masking: Role of referential coding, The Journal of the Acoustical Society of America, 97(2), 1141–1149

Plack C. J. and Viemeister N. F. 1992a, The effects of notched noise on intensity discrimination under forward masking, The Journal of the Acoustical Society of America, 92(4), 1902–1910

Plack C. J. and Viemeister N. F. 1992b, Intensity discrimination under backward masking, The Journal of the Acoustical Society of America, 92(6), 3097–3101

Punch J. and Carhart R. 1973, Influence of interaural phase on forward masking, The Journal of the Acoustical Society of America, 54(4), 897–904

Pérez-González D. and Malmierca M. S. 2014, Adaptation in the auditory system: an overview, Frontiers in Integrative Neuroscience, 8

Rabinowitz N., Willmore B., Schnupp J. and King A. 2011, Contrast Gain Control in Auditory Cortex, Neuron, 70(6), 1178–1191

Rabinowitz N. C., Willmore B. D. B., King A. J. and Schnupp J. W. H. 2013, Constructing Noise-Invariant Representations of Sound in the Auditory Pathway, PLoS Biology, 11(11), e1001710

Rees A. and Palmer A. R. 1988, Rate-intensity functions and their modification by broadband noise for neurons in the guinea pig inferior colliculus, The Journal of the Acoustical Society of America, 83(4), 1488–1498

Relkin E. M. and Doucet J. R. 1991, Recovery from prior stimulation. I: Relationship to spontaneous firing rates of primary auditory neurons, Hearing Research, 55(2), 215–222

Relkin E. M. and Doucet J. R. 1997, Is loudness simply proportional to the auditory nerve spike count?, The Journal of the Acoustical Society of America, 101(5), 2735–2740

Robinson B. L., Harper N. S. and McAlpine D. 2016, Meta-adaptation in the auditory midbrain under cortical influence, Nature Communications, 7(1), 13442

Rocchi F. and Ramachandran R. 2018, Neuronal adaptation to sound statistics in the inferior colliculus of behaving macaques does not reduce the effectiveness of the masking noise, Journal of Neurophysiology, 120(6), 2819–2833

Rocchi F. and Ramachandran R. 2020, Foreground stimuli and task engagement enhance neuronal adaptation to background noise in the inferior colliculus of macaques, Journal of Neurophysiology, 124(5), 1315–1326

Roverud E. and Strickland E. A. 2015a, The effects of ipsilateral, contralateral, and bilateral broadband noise on the mid-level hump in intensity discrimination, The Journal of the Acoustical Society of America, 138(5), 3245–3261

Roverud E. and Strickland E. A. 2015b, Exploring the source of the mid-level hump for intensity discrimination in quiet and the effects of noise, The Journal of the Acoustical Society of America, 137(3), 1318–1335

Ruggero M. A., Rich N. C., Recio A., Narayan S. S. and Robles L. 1997, Basilar-membrane responses to tones at the base of the chinchilla cochlea, The Journal of the Acoustical Society of America, 101(4), 2151–2163

Röhl M., Kollmeier B. and Uppenkamp S. 2011, Spectral loudness summation takes place in the primary auditory cortex, Human Brain Mapping, 32(9), 1483–1496

Sachs M. B. and Abbas P. J. 1974, Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli, The Journal of the Acoustical Society of America, 56(6), 1835–1847

Schacknow P. N. and Raab D. H. 1976, Noise-intensity discrimination: Effects of bandwidth conditions and mode of masker presentation, The Journal of the Acoustical Society of America, 60(4), 893–905

Schlauch R. S., Lanthier N. and Neve J. 1997, Forward-masked intensity discrimination: Duration effects and spectral effects, The Journal of the Acoustical Society of America, 102(1), 461–467

Schneider B. A. and Parker S. 1987, Intensity discrimination and loudness for tones in notched noise, Perception & Psychophysics, 41(3), 253–261

Siebert W. M. 1965, Some implications of the stochastic behavior of primary auditory neurons, Kybernetik, 2(5), 206–215

Siebert W. M. 1968, in *Recognizing Patterns*, mit press, cambridge, ma edn., ed. P. Kolers & M. Eden, 104–133

Simpson A. J. R., Harper N. S., Reiss J. D. and McAlpine D. 2014, Selective Adaptation to "Oddball" Sounds by the Human Auditory System, Journal of Neuroscience, 34(5), 1963–1969

Sivonen V. P. and Ellermeier W. 2011, in *Loudness*, ed. M. Florentine, A. N. Popper, & R. R. Fay, Vol. 37 (New York, NY: Springer New York), 169–197, series Title: Springer Handbook of Auditory Research

Skidmore J., Ramekers D., Colesa D. J., Schvartz-Leyzac K. C., Pfingst B. E. and He S. 2022, A Broadly Applicable Method for Characterizing the Slope of the Electrically Evoked Compound Action Potential Amplitude Growth Function, Ear and Hearing, 43(1), 150–164

Slaney M. 1998, Auditory toolbox: A matlab toolbox for auditory modeling work, Interval Research Corproation, Technical Report, 10, 1194

Stainsby T. H. and Moore B. C. J. 2006, Temporal masking curves for hearing-impaired listeners, Hearing Research, 218(1), 98–111

Stevens S. S. 1966, Power-Group Transformations under Glare, Masking, and Recruitment, The Journal of the Acoustical Society of America, 39(4), 725–735

Stevens S. S. and Guirao M. 1967, Loudness functions under inhibition, Perception & Psychophysics, 2(10), 459–465

Tejani V. D., Abbas P. J. and Brown C. J. 2017, Relationship Between Peripheral and Psychophysical Measures of Amplitude Modulation Detection in Cochlear Implant Users, Ear & Hearing, 38(5), e268–e284

Thwaites A., Glasberg B. R., Nimmo-Smith I., Marslen-Wilson W. D. and Moore B. C. J. 2016, Representation of Instantaneous and Short-Term Loudness in the Human Cortex, Frontiers in Neuroscience, 10

Ulanovsky N. 2004, Multiple Time Scales of Adaptation in Auditory Cortex Neurons, Journal of Neuroscience, 24(46), 10440–10453

Undurraga J. A., Carlyon R. P., Wouters J. and Van Wieringen A. 2012, Evaluating the Noise in Electrically Evoked Compound Action Potential Measurements in Cochlear Implants, IEEE Transactions on Biomedical Engineering, 59(7), 1912–1923

Viemeister N. F. 1974, Intensity discrimination of noise in the presence of band-reject noise, The Journal of the Acoustical Society of America, 56(5), 1594–1600

Viemeister N. F. 1988, Intensity coding and the dynamic range problem, Hearing Research, 34(3), 267–274

Viemeister N. F. and Wakefield G. H. 1991, Temporal integration and multiple looks, The Journal of the Acoustical Society of America, 90(2), 858–865

Wang N., Kreft H. A. and Oxenham A. J. 2015, Loudness Context Effects in Normal-Hearing Listeners and Cochlear-Implant Users, Journal of the Association for Research in Otolaryngology, 16(4), 535–545

Watkins P. V. and Barbour D. L. 2008, Specialized neuronal adaptation for preserving input sensitivity, Nature Neuroscience, 11(11), 1259–1261

Watson C. S. 2005, Some Comments on Informational Masking, Speech and Hearing Sciences

Wen B., Wang G. I., Dean I. and Delgutte B. 2009, Dynamic range adaptation to sound level statistics in the auditory nerve, Journal of Neuroscience, 29(44), 13797–13808

Wen B., Wang G. I., Dean I. and Delgutte B. 2012, Time course of dynamic range adaptation in the auditory nerve, Journal of Neurophysiology, 108(1), 69–82

Wever E. G. 1938, II The Width of the Basilar Membrane in Man, Annals of Otology, Rhinology & Laryngology, 47(1), 37–47

Widin G. P., Viemeister N. F. and Bacon S. P. 1986, Effects of forward and simultaneous masking on intensity discrimination, The Journal of the Acoustical Society of America, 80(1), 108–111

Willmore B. D. B. and King A. J. 2023, Adaptation in auditory processing, Physiological Reviews, 103(2), 1025–1058

Wilson B. S. 1997, Thirty years of the *British Journal of Audiology* : Guest editorial: The future of cochlear implants, British Journal of Audiology, 31(4), 205–225

Wojtczak M., Klang A. M. and Torunsky N. T. 2019, Exploring the Role of Medial Olivocochlear Efferents on the Detection of Amplitude Modulation for Tones Presented in Noise, Journal of the Association for Research in Otolaryngology, 20(4), 395–413

Zeng F.-G. 1998, Interactions of forward and simultaneous masking in intensity discrimination, The Journal of the Acoustical Society of America, 103(4), 2021–2030

Zeng F.-G. and Shannon R. V. 1995, Possible origins of the non-monotonic intensity discrimination function in forward masking, Hearing Research, 82(2), 216–224

Zeng F.-G. and Turner C. W. 1992, Intensity discrimination in forward masking, The Journal of the Acoustical Society of America, 92(2), 782–787

Zeng F.-G., Turner C. W. and Relkin E. M. 1991, Recovery from prior stimulation II: Effects upon intensity discrimination, Hearing Research, 55(2), 223–230

Zhang F., Miller C. A., Robinson B. K., Abbas P. J. and Hu N. 2007, Changes Across Time in Spike Rate and Spike Amplitude of Auditory Nerve Fibers Stimulated by Electric Pulse Trains, Journal of the Association for Research in Otolaryngology, 8(3), 356–372

Zilany M. S. A. and Carney L. H. 2010, Power-Law Dynamics in an Auditory-Nerve Model Can Account for Neural Adaptation to Sound-Level Statistics, Journal of Neuroscience, 30(31), 10380–10390

Zwicker E., Flottorp G. and Stevens S. S. 1957, Critical Band Width in Loudness Summation, The Journal of the Acoustical Society of America, 29(5), 548–557

Appendix

A Results of the preliminary experiment with pure tones for the selected participants

Context	Mean (dB SPL)	Standard error (dB)
silence	3.78	1.24
50 dB SPL	32.96	1.46
70 dB SPL	40.76	0.87

Table A.1 – Mean detection thresholds for the 10 selected participants for experiment 1. They increased with the mean level of the context.

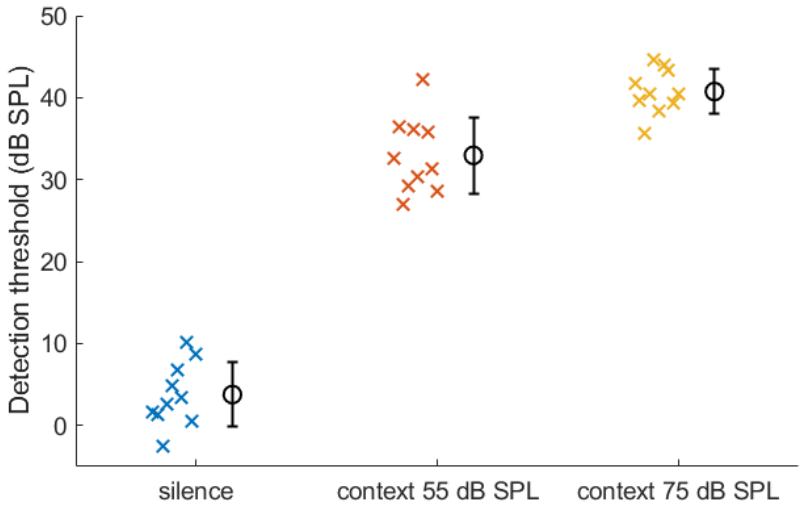


Figure A.1 – Detection thresholds in the three context conditions of experiment 1, for the ten selected participants.

B Sensitivity indexes and DLs obtained with the model

B.1 Methods

The sensitivity index δ' was defined as the ratio between the slope of the rate-level function and the variance of the signal (Colburn et al. 2003). Because the models differed and we only have one rate-level function per fiber type and ERB, the equations we used are the following :

$$(\delta'(L))^2 = \sum_{i=0}^{I}(N_{HSR} \times (\delta'(L)_{HSR,CFi})^2 + N_{MSR} \times (\delta'(L)_{MSR,CFi})^2 + N_{LSR} \times (\delta'(L)_{LSR,CFi})^2) \quad (5)$$

with N_{type} the respective approximate proportion of fibers of each type, I the total number of ERBs in the model (30), L the stimulus level,

$$\delta'(L)_{type,CFi} = \frac{dR_{type,CFi}(L)}{dL} \times \frac{1}{\sqrt{V(L)_{type,CFi}}} \quad (6)$$

the sensitivity index for one fiber type and ERB, with R the mean firing rate of one nerve fiber over the duration of the stimulus and the repetitions of the simulation, and $V(L)_{type,CFi}$ the variance of the firing rate over the duration of the stimulus averaged over repetitions.

It is questionable to study the variance over the average activity of hundreds (or thousands) of fibers, but the model does not provide individual fiber responses, so these results are presented for comparison purposes.

The values of $\delta'(L)_{type,CFi}$ for all three fiber types were compared to the simulations by Grange et al. (2022) for the ERB around the standard (2 kHz). This measure ensured that the changes in slopes were not caused by increases in variability of the firing rate with level, and can be used to compare our simulation quantitatively to those of Grange et al. (2022), and qualitatively to the physiological measurements of Wen et al. (2009). The slopes were computed with a linear regression between the two levels around the standard level.

According to Heinz et al. (2001a) and Colburn et al. (2003), the discrimination threshold can be estimated with $DL = 1/\delta'$ if Weber's law is accurate. DL estimates were also calculated with the model. Values of N_{type} for optimal combination were taken from Liberman (1978) : "*high-, medium-, and low-spont*[aneous rate] *units comprised 61 %, 23 %, and 16 %, respectively, of the units sampled*".

B.2 Experiment 1

The sensitivity index is represented in figure B.1. The curves with context look similar to those obtained by Grange et al. (2022), and the shift of the maximum sensitivity with context is evident ; it is however not exactly aligned with the mean level of the context.

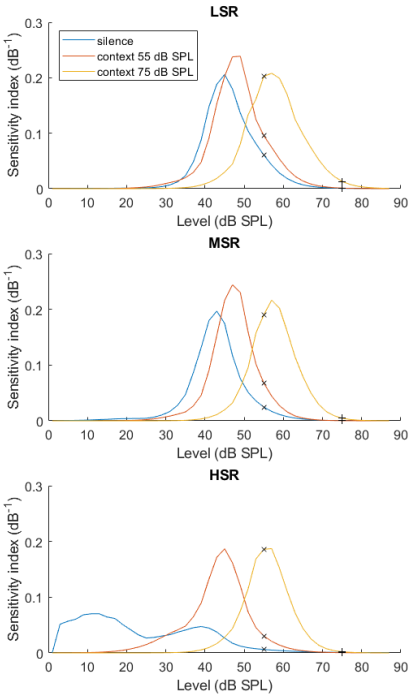


Figure B.1 – Sensitivity index obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 1, averaged over ten simulations. The 55- and 75-dB SPL standard are represented respectively as black x and +.

Standard level	Fiber type	Context type	δ' (dB^{-1})
55 dB SPL	LSR	silence	0.0606
		55 dB SPL	0.0954
		75 dB SPL	0.2023
	MSR	silence	0.0247
		55 dB SPL	0.0677
		75 dB SPL	0.1901
	HSR	silence	0.0062
		55 dB SPL	0.0304
		75 dB SPL	0.1859
75 dB SPL	LSR	silence	0.0006
		55 dB SPL	0.0010
		75 dB SPL	0.0131
	MSR	silence	0.0004
		55 dB SPL	0.0003
		75 dB SPL	0.0046
	HSR	silence	0.0009
		55 dB SPL	0.0004
		75 dB SPL	0.0021

Table B.1 – Sensitivity index $\delta'(L)_{type,CFi}$ in the ERB centered around 2 kHz and the three fiber types in figure B.1 for experiment 1, for both standard levels in each context.

The sensitivity index for each standard level can be extracted from these graphs (c.f. cross symbols in figure B.1) ; these values are also given in table B.1 for both standard levels in the three context conditions and for the three fiber types. For both standards, the sensitivity index is more important in the loud context than in silence or than in the soft context for all fiber types. For the 75-dB SPL standard, the rate-level functions are mostly saturated (resulting in sensitivity indexes below 0.01), except in the loud context for the LSR fibers.

The DLs obtained from the model ($DL = 1/\delta'$) decrease with the increase of the context mean level : in silence, soft context and loud context respectively, the estimates are : for the soft standard, 3.03, 2.53 and 2.18 dB ; for the loud standard, 7.36, 6.55 and 3.44 dB.

The sensitivity index is however not aligned with the changes in slopes (the maximum sensitivity is not obtained exactly at the HPR level), as the standard deviation increases with the discharge rate. For example, the sensitivity for the 55-dB SPL standard is higher in the loud context ; but the slopes for this level are steeper in the soft context for all fiber types, while they are really low for the loud context. The changes in sensitivity index around the standard levels are, therefore, not in agreement with our hypothesis, whereas the slopes are - except for the soft standard for which the HSR rate-level functions are steeper in the loud context, all slopes are steeper in the context mean level identical to the standard (see II.3). The index does not reflect the results of the intensity discrimination task either.

The DL estimate in silence for the soft standard (3.03 dB) is close to the actual measured threshold (3.08 dB), but neither Weber's law or the near-miss are predicted by the model, and the estimates for the loud standard are much higher than the real values (and that for the soft standard). To our knowledge, nobody has tried to derive DLs from the MAP model the way we did ; Heinz et al. (2001a,b) and Colburn et al. (2003) found the near-miss to Weber's law (and the mid-level hump at high frequencies) in their use of the optimal combination for δ'. Maybe the optimal combination of fibers is not well supported by the MAP model as configured in this study. It could be a consequence of the absence of differences between fibers of the same type in the same ERB or an overestimation of the increase in variability of the discharges with the increase in level.

B.3 Experiment 2

The sensitivity index (see table B.2) is much lower for the 55-dB SPL standard than for the 43-dB SPL one.

As in the previous experiment, estimates of the DL were calculated from the sensitivity index. The estimated thresholds were, this time, lower in the context of the same mean level as the standard for both standards : for the 43-dB SPL standard : 1.80 dB in silence, 1.60 dB in the 43-dB SPL context, and 1.76 dB in the 55-dB SPL context ; and for the 55-dB SPL standard, respectively 2.97 dB, 2.94 dB and 2.56 dB. For the 43-dB SPL standard, they are lower than the measured DLs. For the 55-dB SPL standard, they are closer to the measured DLs : a little too high in silence and in the 43-dB SPL context, but too low in the 55-dB SPL context.

The DLs estimated with the model were lower in the context of the same mean

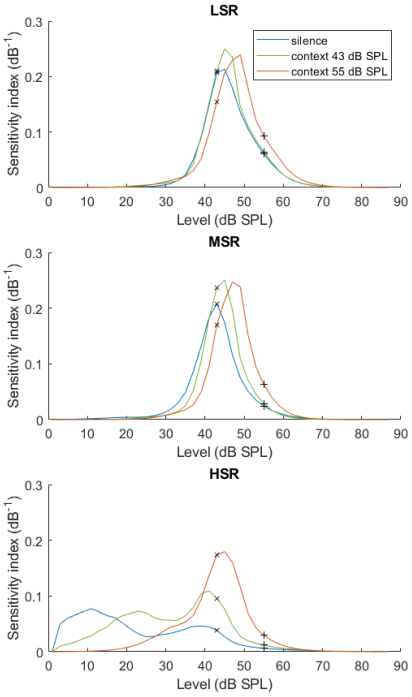


Figure B.2 – Sensitivity index obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 2, averaged over ten simulations. The 43- and 55-dB SPL standard are represented respectively as black x and +.

Standard level	Fiber type	Context type	δ' (dB^{-1})
43 dB SPL	LSR	silence	0.2065
		43 dB SPL	0.2105
		55 dB SPL	0.1546
	MSR	silence	0.2079
		43 dB SPL	0.2364
		55 dB SPL	0.1697
	HSR	silence	0.0382
		43 dB SPL	0.0959
		55 dB SPL	0.1735
55 dB SPL	LSR	silence	0.0606
		43 dB SPL	0.0642
		55 dB SPL	0.0928
	MSR	silence	0.0242
		43 dB SPL	0.0284
		55 dB SPL	0.0632
	HSR	silence	0.0062
		43 dB SPL	0.0117
		55 dB SPL	0.0295

Table B.2 – Sensitivity index $\delta'(L)_{type,CFi}$ in the ERB centered around 2 kHz and the three fiber types in figure B.2 for experiment 2, for both standard levels in each context.

level as the standard, consistently with our hypothesis, but the real DLs obtained in the discrimination task were not.

B.4 Experiment 4

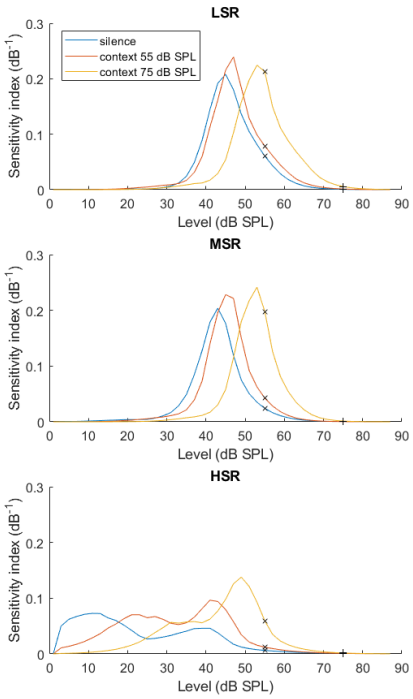


Figure B.3 – Sensitivity index obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 2 kHz, for the 2-kHz standard and narrowband context conditions used in experiment 4, averaged over ten simulations. The 55- and 75-dB SPL standard are represented respectively as black x and +.

A comparison of the sensitivity indices (presented in table B.3) with table B.1 highlights similar indices for the responses in silence, and various differences in the two context mean levels. In general, the sensitivity index for experiment 4 was similar to that obtained in experiment 1. As in experiment 1, the index is higher in the 75-dB SPL context, for both standards and all fiber types. It is however generally smaller or close to that in experiment 1 ; the 75-dB SPL standard is still in the saturated part of the response function. The addition of the gap did not change much in the model.

The estimated DLs for this experiment are of 2.79 dB in the 55-dB SPL context and 2.09 dB in the 75-dB SPL one for the soft standard ; 6.80 dB and 4.78 dB respectively for the 75-dB SPL standard. They are different from the measured thresholds (even more than those estimated for experiment 1), and still decrease with the increase of the context mean level.

B.5 Experiment 6

The sensitivity index for experiment 6 is presented in figure B.4. It can be found for both standard levels in table B.4, for all contexts and fiber types, in the ERB centered

Standard level	Fiber type	Context type	δ' (dB^{-1})
55 dB SPL	LSR	silence	0.0605
		55 dB SPL	0.0785
		75 dB SPL	0.2126
	MSR	silence	0.0241
		55 dB SPL	0.0424
		75 dB SPL	0.1979
	HSR	silence	0.0063
		55 dB SPL	0.0121
		75 dB SPL	0.0593
75 dB SPL	LSR	silence	0.0006
		55 dB SPL	0.0012
		75 dB SPL	0.0056
	MSR	silence	0.0003
		55 dB SPL	0.0003
		75 dB SPL	0.0014
	HSR	silence	0.0009
		55 dB SPL	0.0009
		75 dB SPL	0.0017

Table B.3 – Sensitivity index $\delta'(L)_{type,CFi}$ in the ERB centered around 2 kHz and the three fiber types in figure B.3 for experiment 4, for both standard levels in each context.

Standard level	Fiber type	Context type	Sensitivity index
50 dB SPL	LSR	silence	0.0322
		50 dB SPL	0.0351
		70 dB SPL	0.0499
	MSR	silence	0.0235
		50 dB SPL	0.0288
		70 dB SPL	0.0445
	HSR	silence	0.0082
		50 dB SPL	0.0204
		70 dB SPL	0.0370
70 dB SPL	LSR	silence	0.0155
		50 dB SPL	0.0166
		70 dB SPL	0.0274
	MSR	silence	0.0076
		50 dB SPL	0.0100
		70 dB SPL	0.0207
	HSR	silence	0.0004
		50 dB SPL	0.0040
		70 dB SPL	0.0107

Table B.4 – Sensitivity index for both standard levels for each context and fiber types in figure B.4 for experiment 6 : three-points linear regression around the standard level.

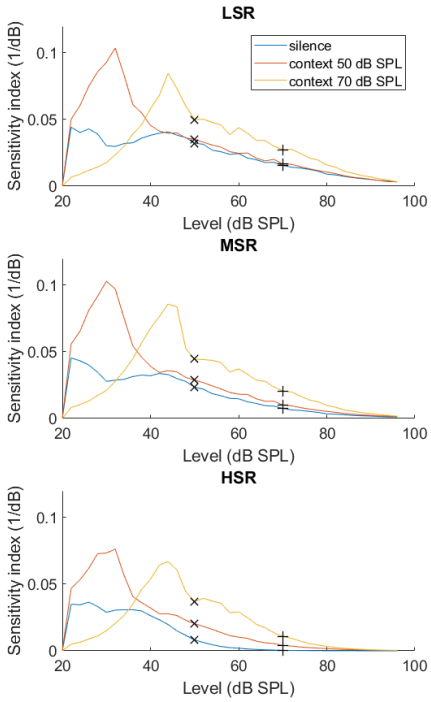


Figure B.4 – Sensitivity index obtained with the MAP model of Grange et al. (2022) for HSR, MSR and LSR fibers tuned around 1 kHz, for the broadband signals and context conditions used in experiment 6, averaged over ten simulations. The 50- and 70-dB SPL standard are represented respectively as black x and +.

around 1 kHz. For both standards, the index is highest in the loud context, and lowest in silence. This is not consistent with the slopes for the soft standard that are generally steeper in the soft context, nor with the results of the discrimination task. The 70-dB SPL standard falls into the saturated part of the rate-level functions for HSR : the values are really low in silence and in the soft context. However, the evolution of the index is consistent with the hypothesis and the discrimination thresholds.

Estimated DLs using an optimal combination of all simulated fibers were lower in the loud context for both standards (for the soft and loud standard respectively : 4.26 & 6.34 dB), but they are higher in silence for the soft standard (5.95 dB) and in the soft context for the loud standard (8.71 dB). The remaining values are the soft standard in the soft context (5.14 dB) and the loud standard in silence (8.05 dB).

The near-miss observed in the discrimination thresholds is not present in the estimated DLs, which are generally higher for the loud standard. They do not reflect the results of the discrimination task or the hypothesis either.

The sensitivity index for this same 1-kHz ERB gives very different results : it is always higher in the loud context. In the present case with broadband signals, it is however possible that looking at one ERB only does not reflect the results of the discrimination task. Yet the estimated DLs, obtained by inverting the combination of the sensitivity indexes for all ERBs and fiber types (Colburn et al. 2003), are also far from the behavioral results.

C Additional data on eCAPs

C.1 Evolution of 30 pps classic control for L12

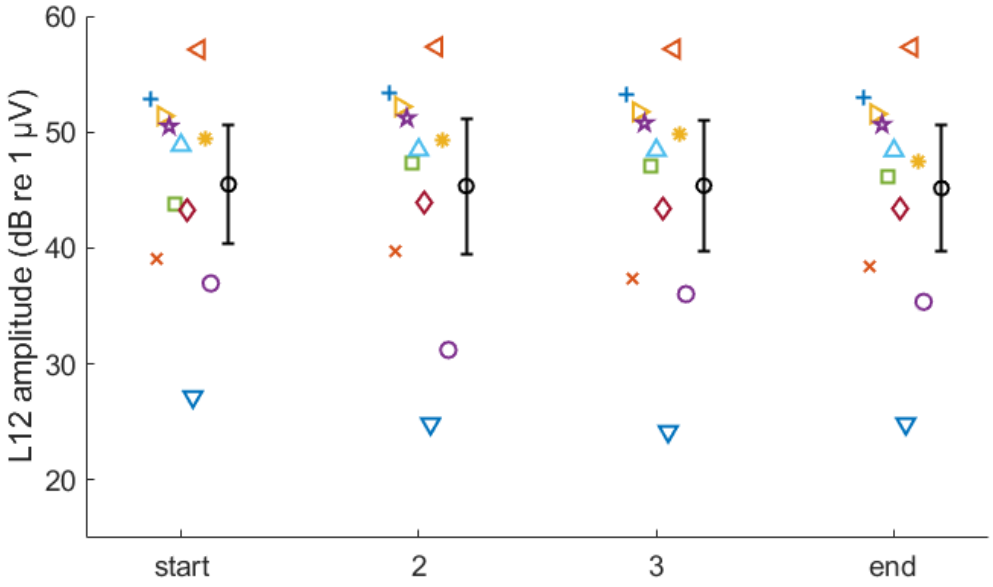
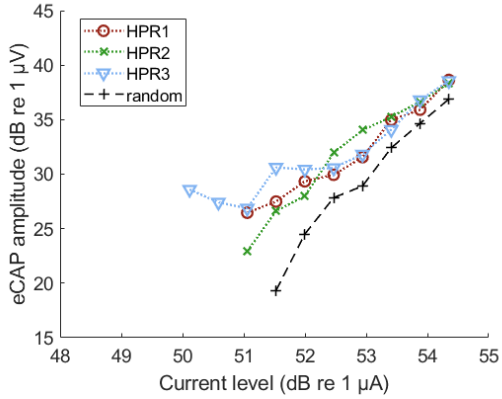


Figure C.1 – Evolution of L12 amplitude in dB for the short CLASSIC paradigm, between the start and the end of the procedure, as in figure IV.4.

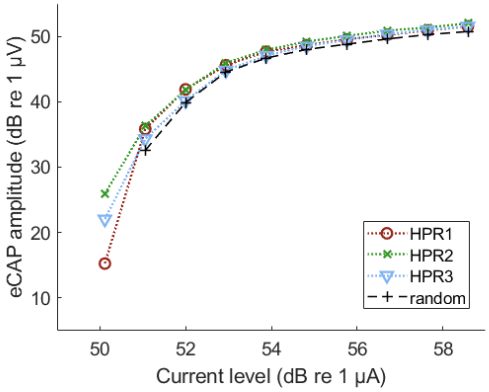
C.2 AGF for HPRs and RANDOM paradigms, for all participants

AGFs for C5 can be found in figures IV.6a to IV.6c.

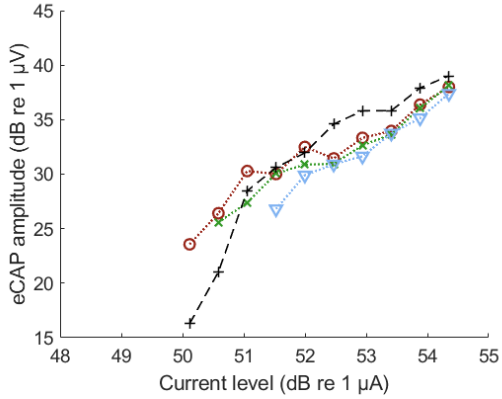
(a) Amplitude at 30 pps for C2



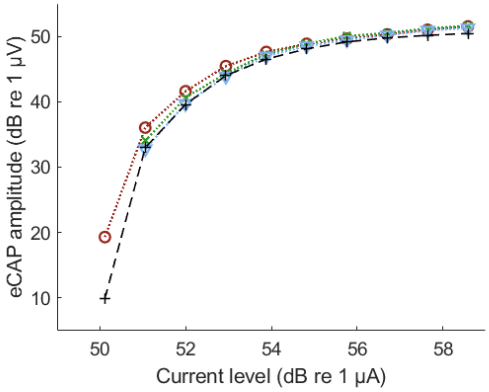
(d) Amplitude at 30 pps for C15



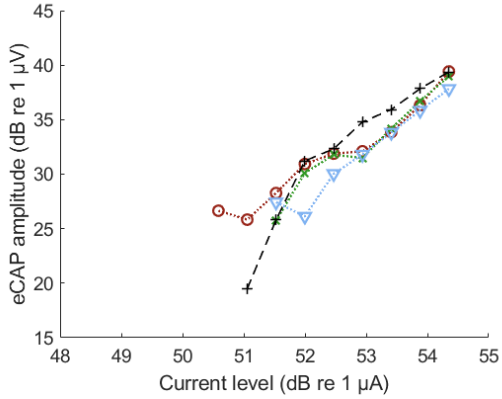
(b) Amplitude at 100 pps for C2



(e) Amplitude at 100 pps for C15



(c) Amplitude at 300 pps for C2



(f) Amplitude at 300 pps for C15

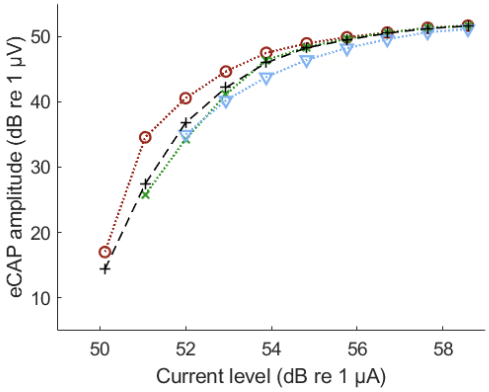
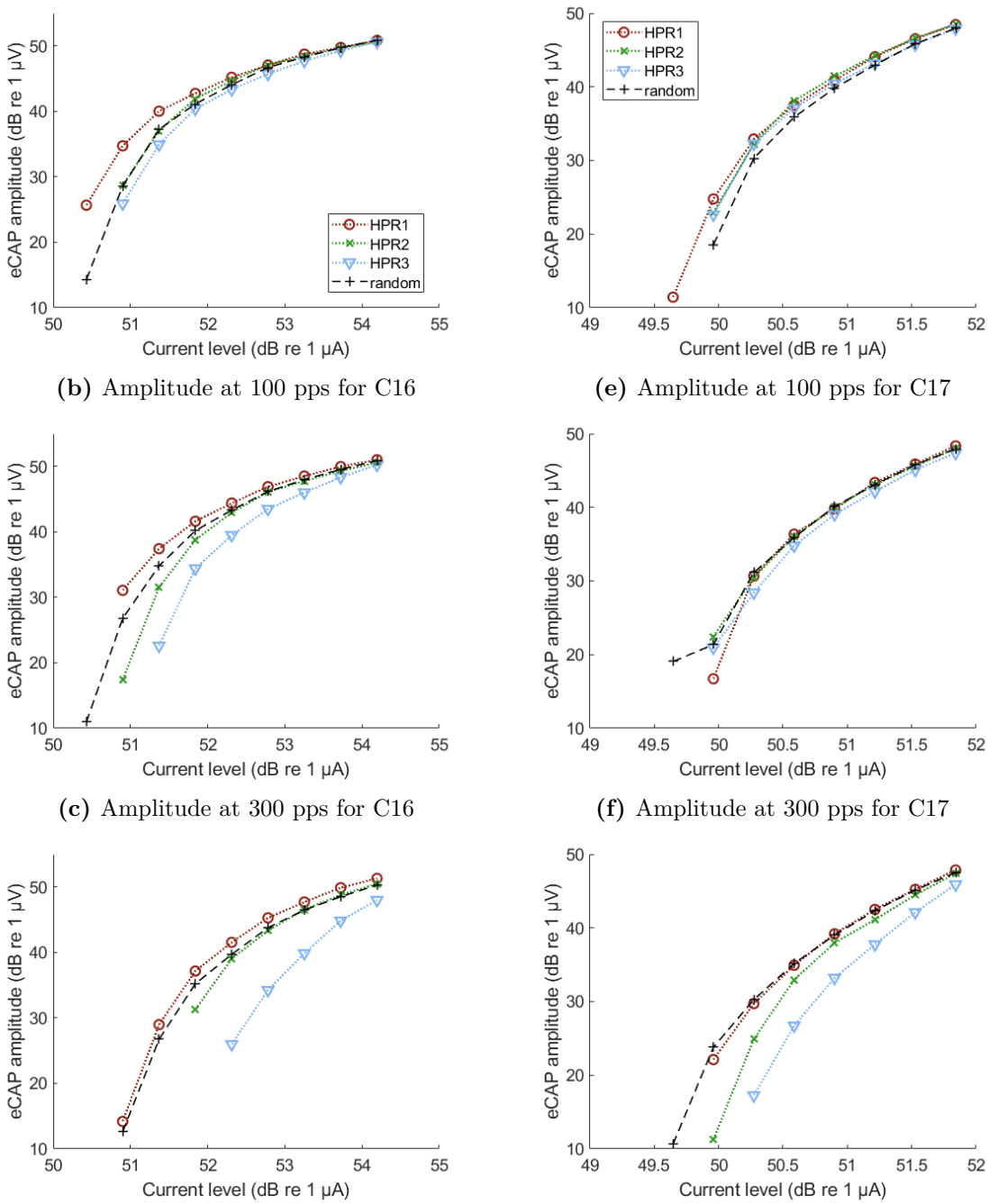


Figure C.2 – Amplitude growth function for participants C2 and C15 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps.



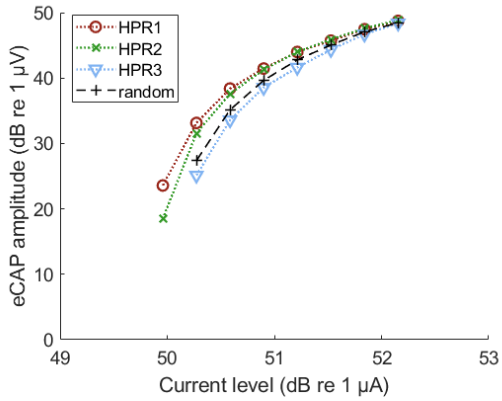
(a) Amplitude at 30 pps for C16 **(d)** Amplitude at 30 pps for C17

(b) Amplitude at 100 pps for C16 **(e)** Amplitude at 100 pps for C17

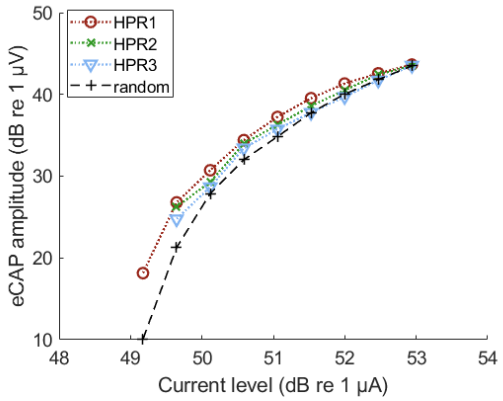
(c) Amplitude at 300 pps for C16 **(f)** Amplitude at 300 pps for C17

Figure C.3 – Amplitude growth function for participants C16 and C17 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps.

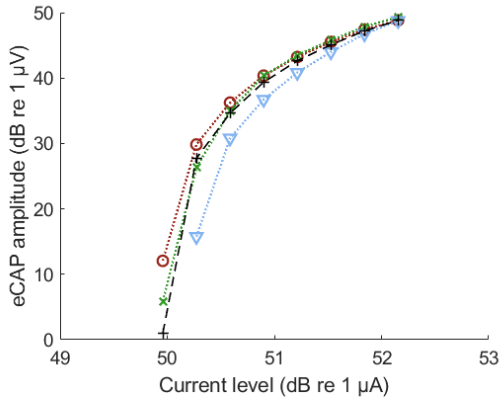
(a) Amplitude at 30 pps for C18



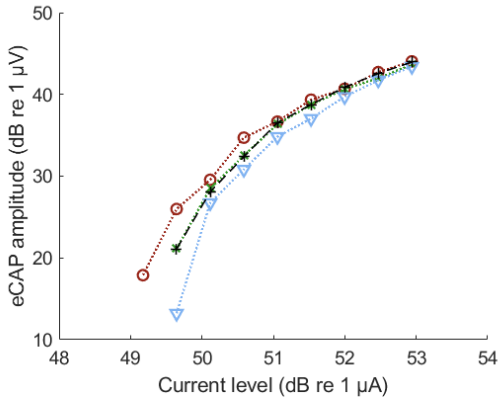
(d) Amplitude at 30 pps for C19



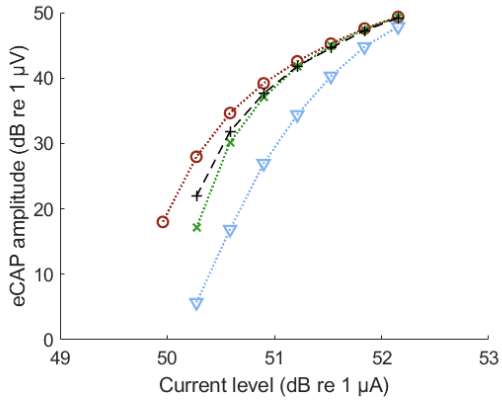
(b) Amplitude at 100 pps for C18



(e) Amplitude at 100 pps for C19



(c) Amplitude at 300 pps for C18



(f) Amplitude at 300 pps for C19

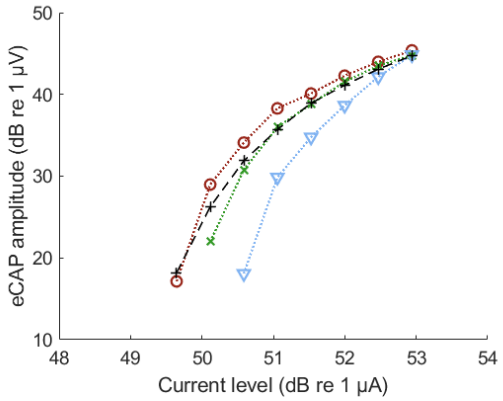
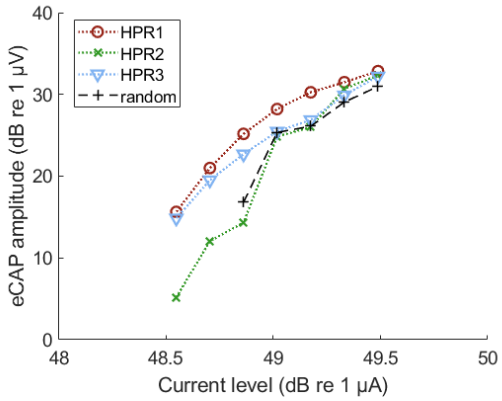
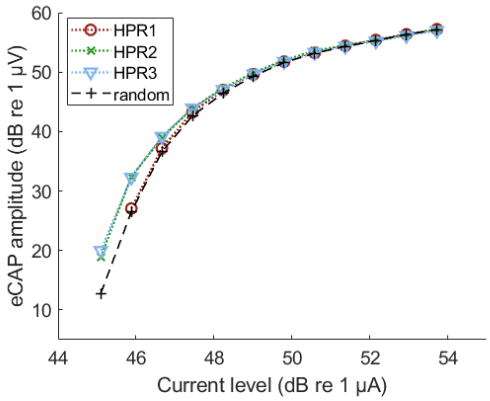


Figure C.4 – Amplitude growth function for participants C18 and C19 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps.

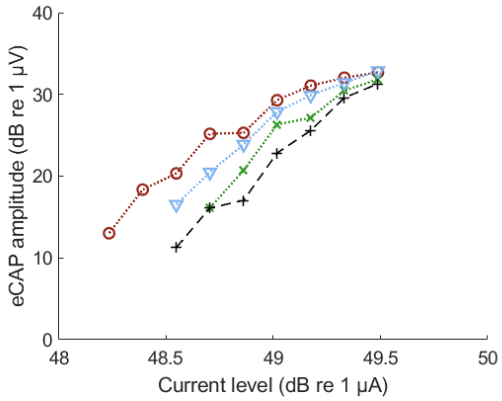
(a) Amplitude at 30 pps for C35



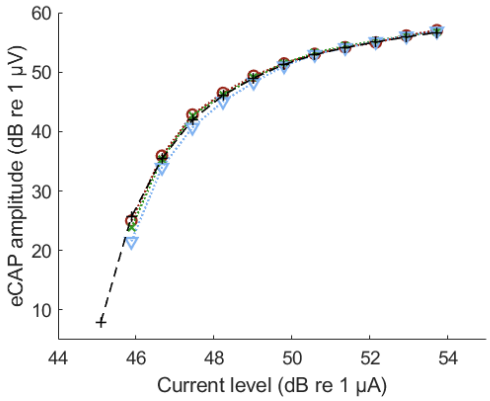
(d) Amplitude at 30 pps for C70



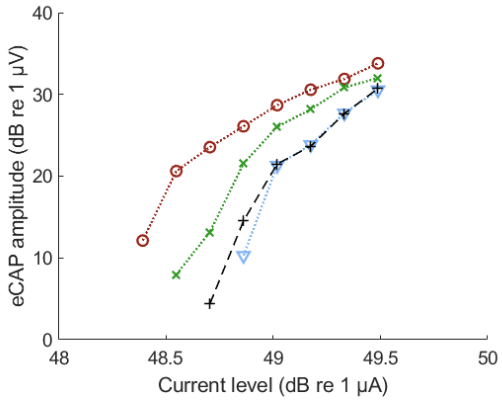
(b) Amplitude at 100 pps for C35



(e) Amplitude at 100 pps for C70



(c) Amplitude at 300 pps for C35



(f) Amplitude at 300 pps for C70

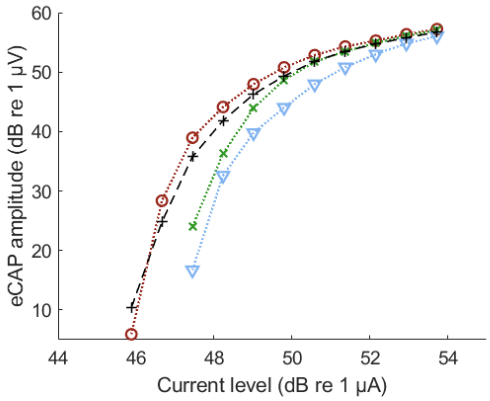
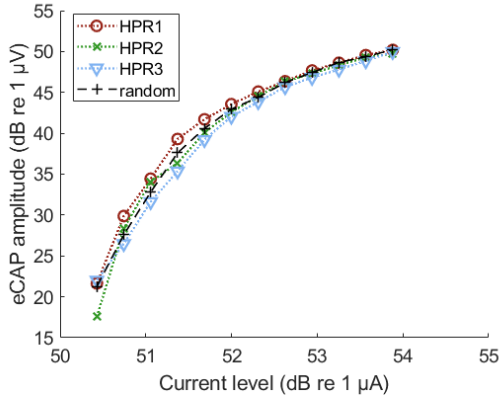
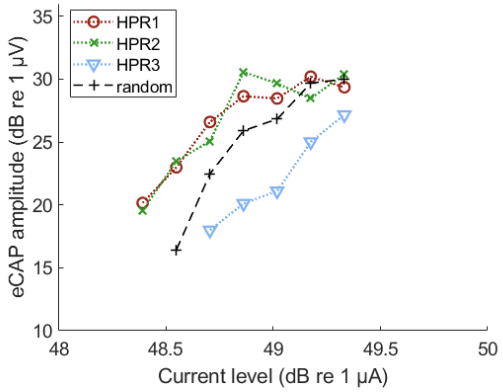


Figure C.5 – Amplitude growth function for participants C35 and C70 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps.

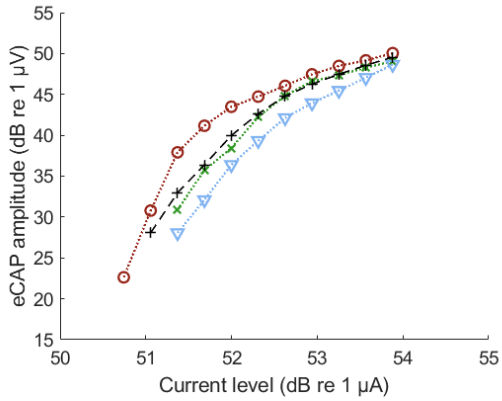
(a) Amplitude at 30 pps for C46



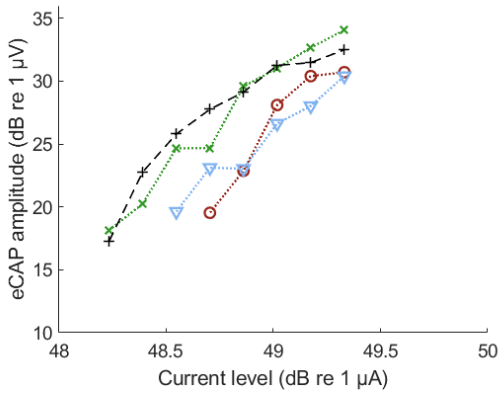
(d) Amplitude at 30 pps for C26



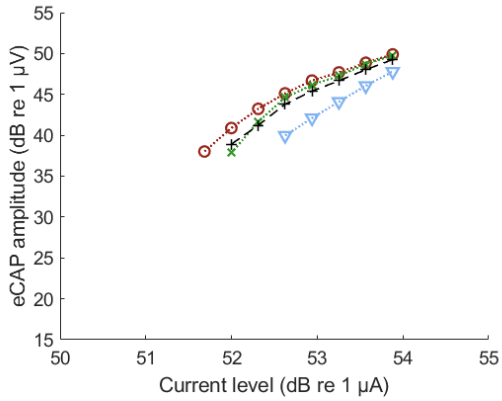
(b) Amplitude at 100 pps for C46



(e) Amplitude at 100 pps for C26



(c) Amplitude at 300 pps for C46



(f) Amplitude at 300 pps for C26

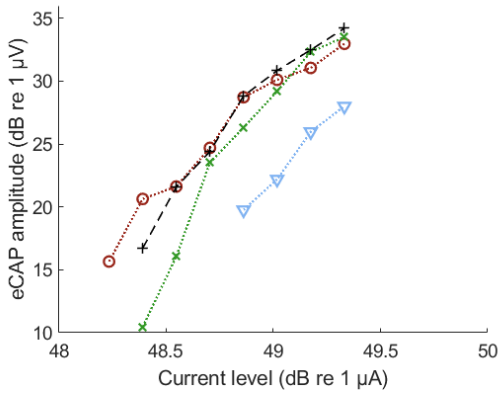


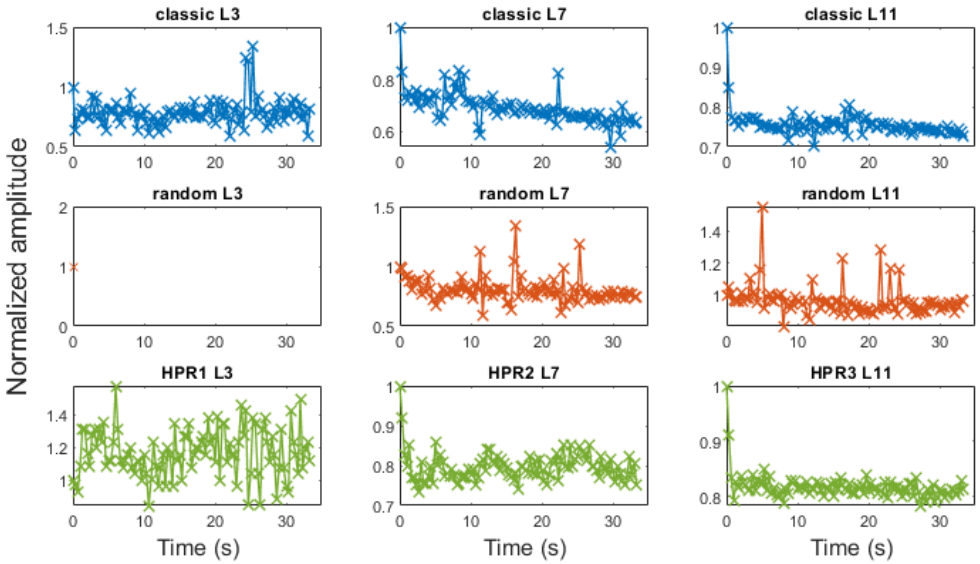
Figure C.6 – Amplitude growth function for participants C46 and C26 : a&d) AGF 30 pps, b&e) 100 pps, c&f) 300 pps.

C.3 Temporal evolution of normalized amplitude at 300 pps

Profile	Paradigm	Level	Number	Example figure
First decrease only	RANDOM	L7	5	C.9b
		L11	3	
	CLASSIC	L7	3	IV.11a
		L11	4	
	HPR	L7	6	
		L11	5	C.7a
Secondary decrease	RANDOM	L7	0	
		L11	0	
	CLASSIC	L7	5	
		L11	4	IV.11b
	HPR	L7	2	
		L11	2	
Secondary increase	RANDOM	L7	0	
		L11	0	
	CLASSIC	L7	1	
		L11	2	IV.11c
	HPR	L7	1	
		L11	1	
Increase (initial or longer)	RANDOM	L7	2	
		L11	5	C.9a
	CLASSIC	L7	2	
		L11	0	
	HPR	L7	1	
		L11	1	IV.11c
No evolution	RANDOM	L7	5	
		L11	3	
	CLASSIC	L7	0	
		L11	1	
	HPR	L7	1	
		L11	2	

Table C.1 – Classification of the different encountered profiles for all paradigms, 300 pps temporal evolution.

(a) C5



(b) C2

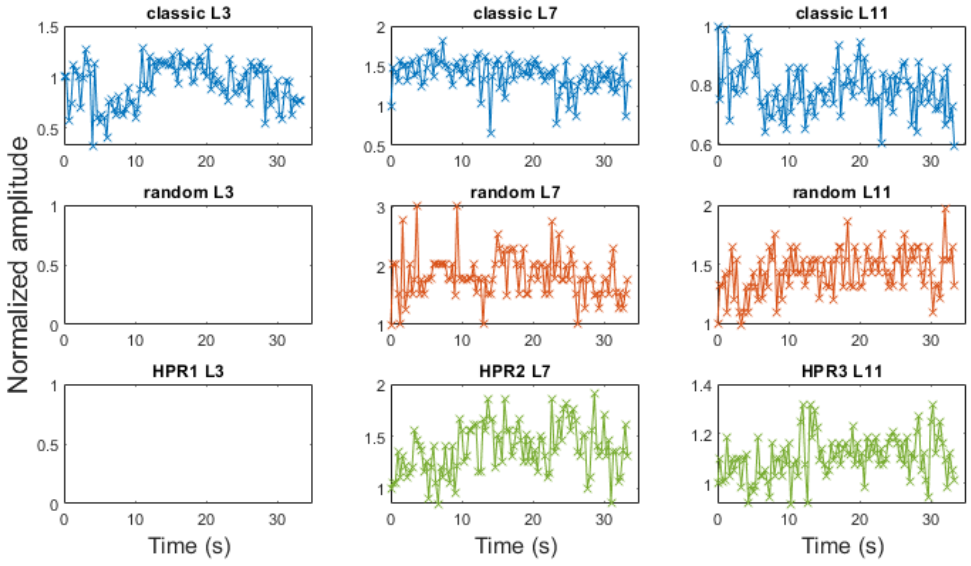
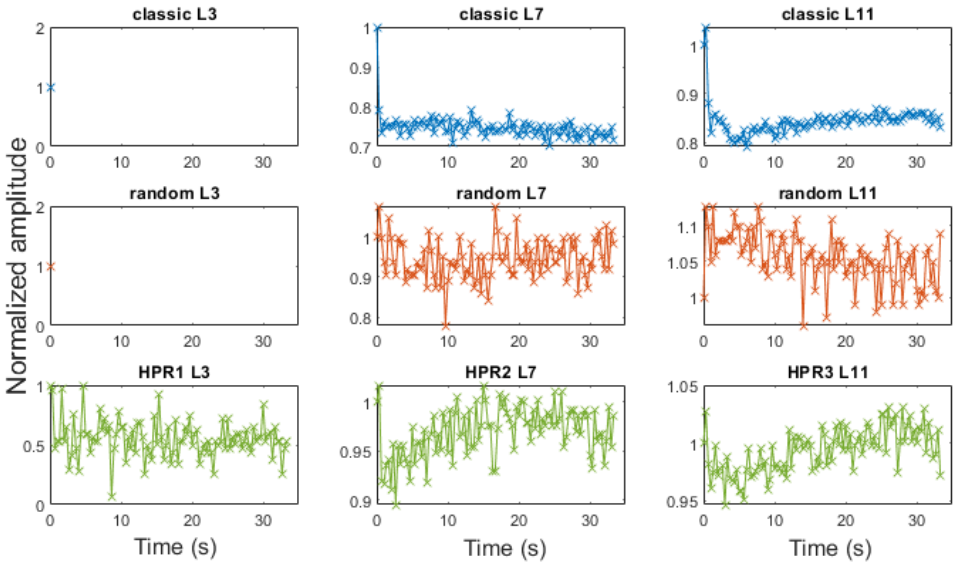


Figure C.7 – Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C5 and C2.

(a) C15



(b) C16

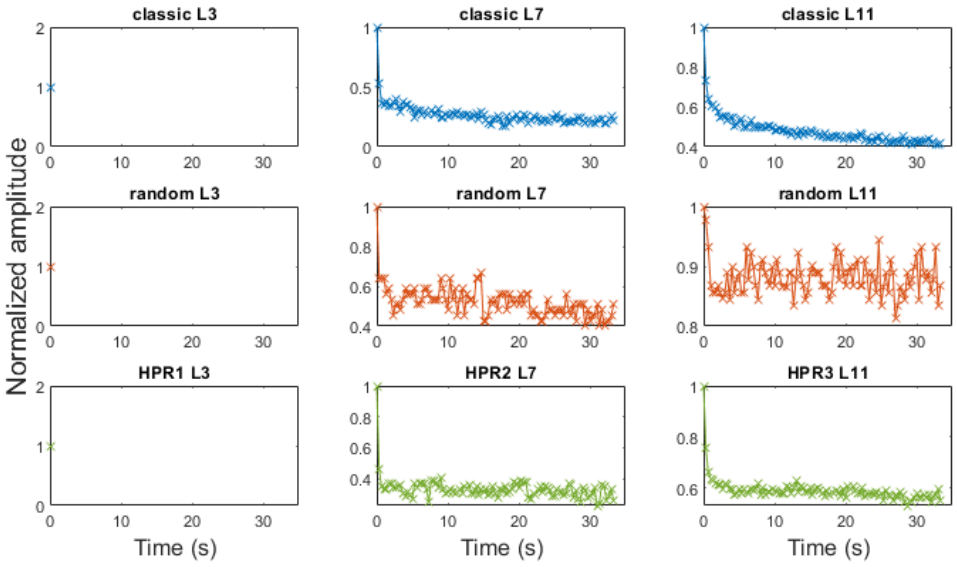
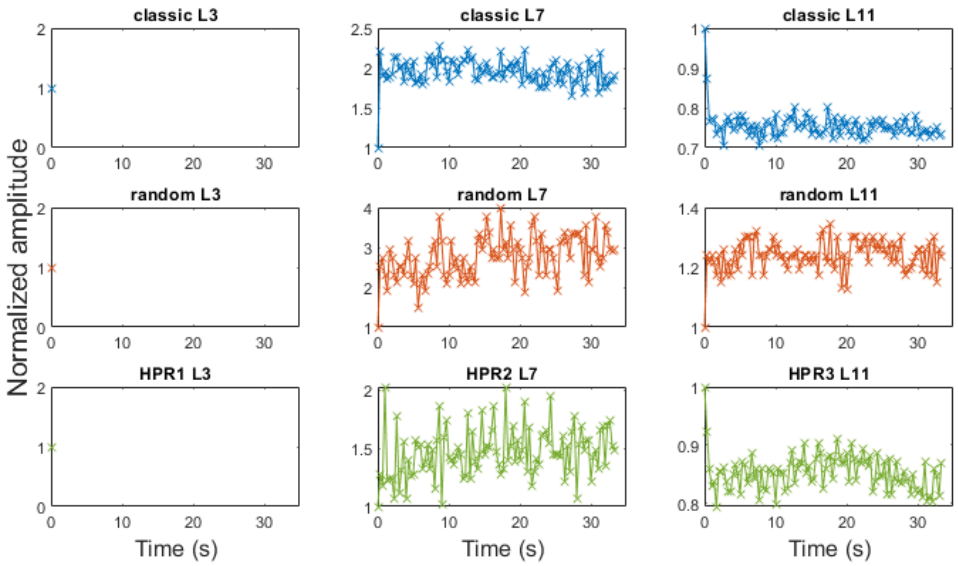


Figure C.8 – Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C15 and C16.

(a) C17



(b) C18

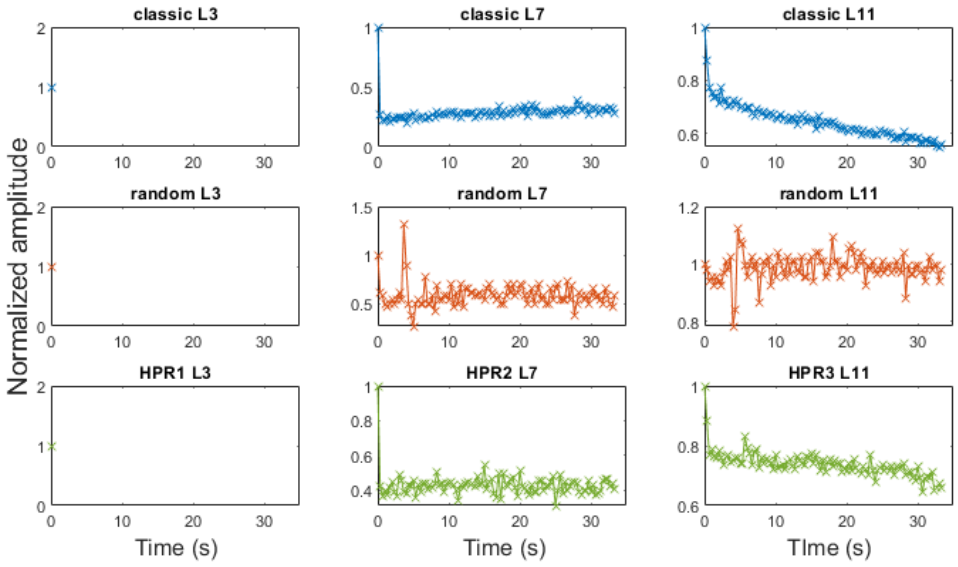
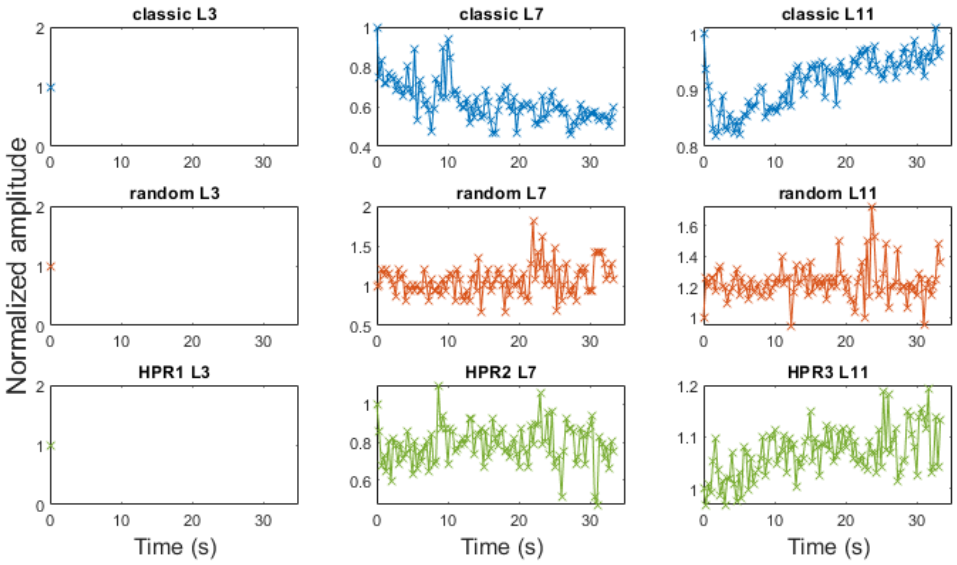


Figure C.9 – Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C17 and C18.

(a) C19



(b) C35

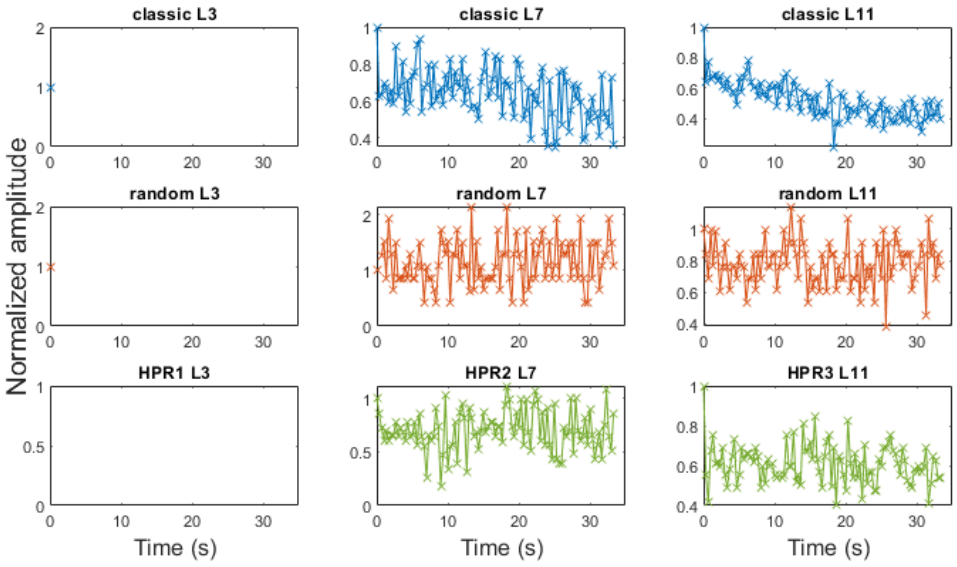
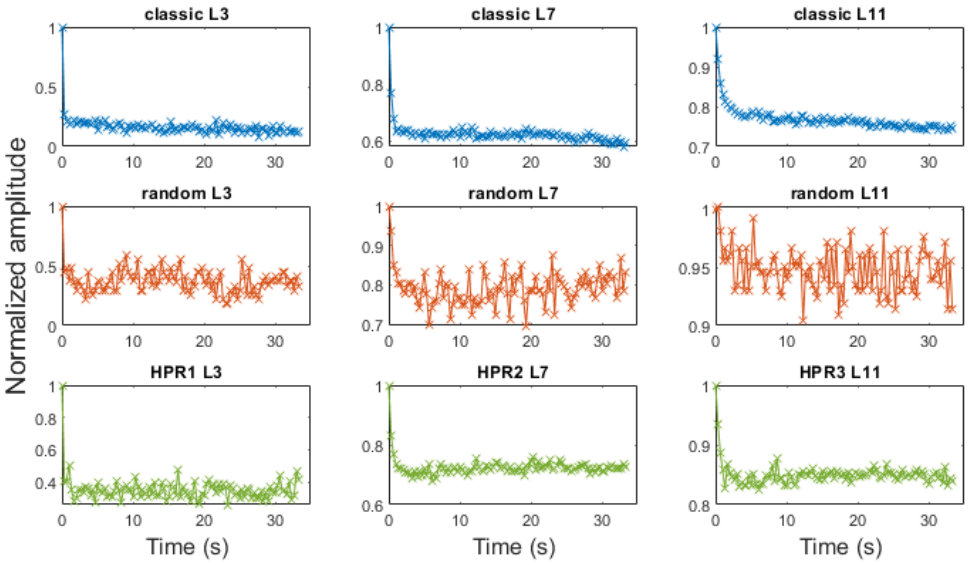


Figure C.10 – Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C19 and C35.

(a) C70



(b) C46

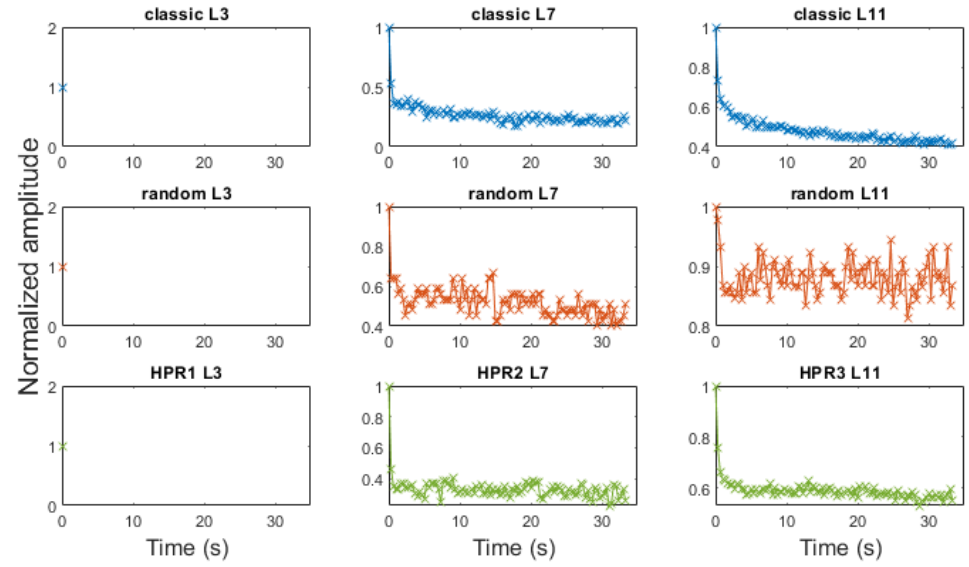


Figure C.11 – Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participants C70 and C46.

(a) C26

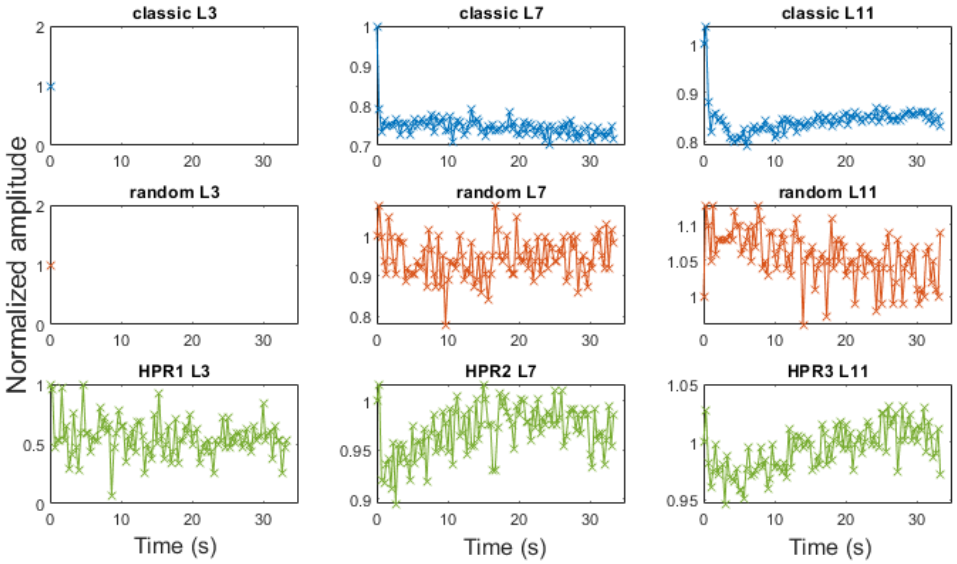


Figure C.12 – Evolution of the eCAP normalized amplitude for the different paradigms at L3, L7 and L11, for participant C26.