Functional determinants of spectrotemporal selectivity in mouse auditory cortex

GÆLLE A. CHAPUIS

SUPERVISOR: DR. PAUL T. CHADDERTON

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE
DEPARTMENT OF BIOENGINEERING
DATE: 11TH OF APRIL 2018
Submitted in fulfilment of the requirements for the award of Doctor of Philosophy in Bioengineering and the Diploma of the Imperial College.
Declaration

Originality

I hereby declare that the content of this thesis represents my own work, except for quotations or images from published and unpublished sources which are clearly indicated and acknowledged as such.

Copyright

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

What if I fall? Oh but darling, what if you fly?
Erin Handson
I would like to extend thanks to many people, offering their help and support countless, for some regardless of time or distance. Without those, the work presented in this thesis would not have been possible.

A special mention goes to my supervisor, Dr Paul Chadderton, who supported my ideas and scientific preferences by letting me explore subjects that intrigued me. Another mention goes to the Chadderton Lab members (Dr Joseph Sollini; Dr Matthew Brown; Dr Caroline Golden; Dr Susu Chen; Dr Alexander Morris; Jun Song; Tunvez Boulic; Altina Fazliu; Adel Haddad) who all offered tremendous help with experiments and conceptual dilemmas, and also importantly brought some fun and moral support in most needed times.

Luckily, these qualities extended to many people I have had the pleasure to meet throughout these past years. To only cite a few from Imperial College London, all members of the Clopath Lab (special mention to Dr Claudia Clopath; Dr Wilten Nicola; Dr Katha Wilmes), of the Schultz Lab (special mention to Dr Yu Liu; Dr Marie Tolkiehn; Dr Susanna Milolo; Dr Ann Go; Dr Stef Garasto; Dr Luca Anneckino), of the Krapp Lab (special mention to Dr Jiaqi Huang, a BIG THANKS for all the help with Python; Dr Ben Hardcastle; Dr Silvia Jimenez), of the Brickley Lab (special mention to Dr Stephen Brickley and Dr David Uygun who provided the DREADD construct; Diana Lucaci), and of the Overby Lab (special mention to Dr Jason Chang; Alice Spencehauer; Pei Jung Wu) have been particularly present and helpful. The staff of Imperial College London was also very helpful and professional throughout my PhD. A special thanks goes to Paschal Egan, Daniel Nardini, Nazib Ahmed and Allan Nyunt, for your tremendous help with the making and ordering of equipments. A special thanks goes to Britta Ross for the help with administrative tasks.

Beyond Imperial College London, I have encountered generous and helpful people in several universities. A special thanks goes to all members of the Hauser Lab at UCL (special mention to Prof Michael Hauser; Dr Arnd Roth; Dr Mehmet Fisek; Dr Adam Packer; Lloyd Russel), who broadened my perspective on neuroscience, and now constitute a strong net of friends with common scientific and moral interests. Thanks to Dr Edward Bracey from the Margrie Lab in Mill Hill, Raphael Steinfeld from the Renart Lab in Champalimaud and Dr Eli Meyer from the Luthi Lab in Basel for all the tips given on mice behaviour and pharmacology. Finally, I have learnt that friends from EPFL are always here to listen and advise (given the right call - Oh santé!...; special mention to Dr Andreas Thomik).

Whilst scientific research is in itself an expedition, I feel grateful to have been surrounded with people who were always up to a different kind of adventure. For all my flatmates with whom I spent years in 53 Averill street (Dr Martina Fonseca; Dr Marie Tolkiehn; Dr Jonathan Ball; Ana Maria Gomez-Ball), a deep thanks for all the fun we have had. Weekends would have been different without the climbing group (the italic font emphasises that climbing was just an excuse to share many precious moments). Many thanks to the members of the San Francisco team formed years ago, with whom I travelled to all places (special mention to Tanguy Cetelem, for your help with coding as well). A vivid thanks to all my friends left behind in Switzerland, for not forgetting about me whilst I was giving so little sign of life. Finally, I would like to mention you HH, my dear companion, and acknowledge the great home and life we are building. To you, I can only say: xxxx.

My family played a major role in keeping me up and afloat. Their moral and financial support was limitless, hence this last acknowledgement section is dedicated to them. Je ne citerai que des noms de famille, pour faire court, car chacun d’entre vous apporte un touche spéciale à notre grande smala. Un grand merci aux familles Chapuis, Berthoin et Thouanel, pour votre amour, votre affection et votre soutien, qui se manifestaient bien au-delà des frontières!

And for all the Georges I cherished, R.I.P.
Abstract

When we interact with our surroundings, our brain must generate a representation of the world that we can interpret and act upon. How such a representation emerges, and how can it best support behavioural needs have been open questions in the field of neuroscience for decades. Throughout the years, theories on sensory processing have often been challenged, owing to new discoveries, which better captured the brain’s complex organisation. The mammalian brain possesses a substantial diversity of molecules, cells, and connections, which may serve distinct roles in shaping sensory representation. Such heterogeneity presents a challenge when aiming to study the underlying basis of neural circuit computation and behaviour.

Recent technologies have facilitated the functional dissection of brain circuits by enabling us to record and manipulate the activity of distinct cell types based on molecular signatures or anatomical projections. Taking advantage of these technologies, I aimed to decipher the functional determinants of a neuron's selectivity to sound features in the mouse primary auditory cortex (AC), a region of the brain believed to be important for sound perception. Specifically, I studied the principles underlying neural encoding and perception of frequency modulation (FM), an important characteristic of natural sounds such as vocal calls. Using electrophysiological, behavioural and pharmacogenetic techniques in mice, I showed that PV-expressing (PV+) interneurons in AC do not mediate the perception nor the encoding of slow FM direction. Upon the reduction of PV+ interneurons activity in vivo, AC cells remained selective to the direction of FM sounds. Also, mice were able to accurately discriminate the direction of slow FM sounds even when cortical inhibition was reduced. It is proposed that instead, the organisation of ON and OFF receptive fields contributes towards slow FM direction selectivity in cortex.

Furthermore, I investigated how sensory processing may adapt to support behavioural needs in scenarios where only a subset of sensory inputs is relevant. By recording in the auditory cortex and thalamus of mice performing auditory tasks, I was able to show that neural responses to sounds adapted according to behavioural demand. Such enhancement in sensory processing appeared as early as in the thalamus, suggesting that this region might be the prime anatomical locus of ‘top-down’ executive control. Nonetheless, AC neurons displayed complex responses during behaviour, which were highly informative on the animal’s behavioural responses. This suggests that the AC encodes high-level aspects of the task alongside representing stimulus features.

As the mouse shares a remarkable genetic resemblance and aspects of behaviour with humans, these studies may shed light on auditory processing properties relevant to our species, thereby fostering further translational research between mice and men.
# Contents

<table>
<thead>
<tr>
<th>Abstract</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1: Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Challenges in hearing</td>
<td>1</td>
</tr>
<tr>
<td>1.2 On the encoding and perception of frequency modulation</td>
<td>2</td>
</tr>
<tr>
<td>1.2.1 Studying frequency modulation encoding: Breaking down the problem</td>
<td>2</td>
</tr>
<tr>
<td>1.2.2 Representation of frequency content along the ascending auditory pathway</td>
<td>3</td>
</tr>
<tr>
<td>1.2.3 Cellular diversity and organisation in primary auditory cortex</td>
<td>4</td>
</tr>
<tr>
<td>1.2.4 Sound processing within cortical circuits - A review on inhibitory and excitatory contributions</td>
<td>5</td>
</tr>
<tr>
<td>1.2.5 Techniques to selectively manipulate neural activity</td>
<td>6</td>
</tr>
<tr>
<td>1.2.6 Cortical processing beyond inhibitory-excitatory interaction</td>
<td>7</td>
</tr>
<tr>
<td>1.2.7 The role of auditory cortex in sound perception</td>
<td>9</td>
</tr>
<tr>
<td>1.2.8 Summary of research aims</td>
<td>10</td>
</tr>
<tr>
<td>1.3 Sound perception in noisy environment</td>
<td>11</td>
</tr>
<tr>
<td>1.3.1 Auditory scene analysis: The challenge of forming sound streams</td>
<td>11</td>
</tr>
<tr>
<td>1.3.2 Attention modulates perception and sensory representation</td>
<td>13</td>
</tr>
<tr>
<td>1.3.3 Anatomical circuits supporting sensory representation modulation and attention</td>
<td>17</td>
</tr>
<tr>
<td>1.3.4 Techniques to decipher the contribution of specific anatomical circuit components</td>
<td>20</td>
</tr>
<tr>
<td>1.3.5 Summary of research aims</td>
<td>21</td>
</tr>
<tr>
<td>Chapter 2: Materials and Methods</td>
<td>35</td>
</tr>
<tr>
<td>2.1 Animals</td>
<td>35</td>
</tr>
<tr>
<td>2.2 In vivo electrophysiology</td>
<td>36</td>
</tr>
<tr>
<td>2.2.1 Electrodes</td>
<td>36</td>
</tr>
<tr>
<td>2.2.2 Targeting the auditory cortex of anaesthetised mice</td>
<td>36</td>
</tr>
<tr>
<td>2.2.3 Targeting the auditory cortex of awake mice</td>
<td>37</td>
</tr>
<tr>
<td>2.2.4 Targeting the medial geniculate body of awake mice</td>
<td>39</td>
</tr>
<tr>
<td>2.2.5 Simultaneous recording of prefrontal and auditory cortex activity</td>
<td>39</td>
</tr>
<tr>
<td>2.3 Auditory stimulation</td>
<td>39</td>
</tr>
<tr>
<td>2.3.1 Stimulus generation</td>
<td>39</td>
</tr>
<tr>
<td>2.3.2 Stimulus presentation</td>
<td>39</td>
</tr>
<tr>
<td>2.3.3 Stimulus properties</td>
<td>40</td>
</tr>
<tr>
<td>2.4 Behavioural training and assessment</td>
<td>42</td>
</tr>
<tr>
<td>2.4.1 Setup</td>
<td>42</td>
</tr>
<tr>
<td>2.4.2 Paradigms</td>
<td>42</td>
</tr>
<tr>
<td>2.4.3 Measure of performance</td>
<td>43</td>
</tr>
<tr>
<td>2.4.4 Training protocols</td>
<td>43</td>
</tr>
<tr>
<td>2.5 Data acquisition</td>
<td>46</td>
</tr>
<tr>
<td>2.5.1 During passive listening experiments</td>
<td>46</td>
</tr>
</tbody>
</table>
2.5.2 During active behaving experiments ........................................... 47
2.6 Pharmacogenetics ................................................................. 47
  2.6.1 Viral vectors ................................................................. 47
  2.6.2 Viral injection ............................................................... 48
  2.6.3 Auditory cortex inactivation ............................................. 49
  2.6.4 Recording paradigms ....................................................... 49
2.7 Histology .......................................................................... 50
2.8 Data analysis ..................................................................... 50
  2.8.1 Spike sorting ............................................................... 50
  2.8.2 Selection of auditory responsive cells .................................. 50
  2.8.3 Receptive field analysis .................................................. 51
  2.8.4 Direction and rate selectivity analysis .................................. 51
  2.8.5 Discrimination performance ............................................. 51
  2.8.6 Circular data analysis ..................................................... 52
  2.8.7 Local field potential and multi-unit analysis ......................... 52
  2.8.8 Modelling of experimental data ........................................ 53
  2.8.9 Statistics ...................................................................... 54

Chapter 3 Functional determinants of response selectivity to frequency modulated sounds 63
3.1 Introduction ............................................................................ 63
  3.1.1 Mechanistic basis of ON and OFF receptive field arrangement .... 63
  3.1.2 ON/OFF receptive field arrangement conferred direction selectivity to frequency-modulated sounds \textit{in silico} ................................................................. 64
  3.1.3 Research aims: Tests for the functional significance of ON and OFF receptive fields \textit{in vivo} ................................................................. 65
3.2 Results .................................................................................. 66
  3.2.1 ON/OFF cells discriminated best the direction of slow frequency modulations ... 66
  3.2.2 Relationships between FM direction selectivity and frequency tuning .... 67
  3.2.3 ON/OFF receptive field arrangement was a good predictor of direction selectivity to slow frequency-modulated sounds ........................................... 68
  3.2.4 Fast cortical inhibition was not required for direction selectivity .......... 69
  3.2.5 Direction selectivity of Young A1 neurons \textit{in vivo} ....................... 70
  3.2.6 Direction selectivity of Adult A1 neurons recorded in awake passive listening condition 71
  3.2.7 Mice were able to discriminate the direction of slow FM sweeps under reduced cortical inhibition ................................................................. 71
3.3 Discussion ............................................................................. 74
  3.3.1 ON/OFF segregation as a specialised mechanism for enhancing direction selectivity to slow FM speeds .................................................. 74
  3.3.2 ON/OFF arrangement relates to functional topography in Adult and Young mice ...................................................................................... 75
  3.3.3 Delineating the dependence of FM encoding on cortical state ............ 75
  3.3.4 The role of auditory cortex in the perception of FM sounds ............... 76
  3.3.5 Concluding remark ................................................................ 76
3.4 Future work ................................................................. 77

Chapter 4 Anatomical circuits underlying task-related modulation of spectrotemporal
receptive fields ......................................................... 95
4.1 Introduction .............................................................. 95
4.2 Results ....................................................................... 96
  4.2.1 Mice were sensitive to spectrotemporal cues ................. 96
  4.2.2 Spectrotemporal expectation modulated neural responses in auditory cortex ... 98
  4.2.3 Spectrotemporal expectation modulated neural responses in auditory thalamus ... 107
  4.2.4 The tuning profile of auditory thalamic cells was not affected by TRN manipulation 109
4.3 Discussion ................................................................. 111
  4.3.1 Auditory attention and streaming could be inferred from the timing of false alarms 111
  4.3.2 Activity in auditory cortex depended on attentional state – An effect mediated by
      the thalamus? ............................................................... 111
  4.3.3 Comparison of local field potential signals in the auditory cortex of mice and primates 12
4.4 Future work ............................................................. 113
  4.4.1 Delineating the effect of PFC and TRN on sound processing ............... 113
  4.4.2 Study of recurrent thalamo-cortical circuits ..................... 113
  4.4.3 Behavioural paradigm refinement .................................. 113
  4.4.4 Assessing the effect of movement on neural processing .......... 114

Chapter 5 Conclusion ......................................................... 135

Bibliography ................................................................. 137

Copyrights ................................................................. 153

Appendix ........................................................................... 178
I.1 Supplementary Information ............................................. 179
  I.1.1 Assessing contribution of functional neuronal properties to slow direction selectivity
        in A1 ................................................................. 179
I.2 Supplementary Figures ................................................... 181
List of Figures

1.1 Example spectrograms of natural sounds. ........................................... 22
1.2 Mouse vocalisation contains slow frequency modulations. ...................... 22
1.3 Example selective neural response to FM sweep direction. ....................... 23
1.4 Approximation of FM by pure tones. .................................................. 23
1.5 Spectral and temporal history affect spiking response in auditory cortex. .......... 24
1.6 Ascending auditory pathway and tonotopy. ......................................... 25
1.7 Correlation between frequency tuning and FM direction selectivity in primary auditory cortex. ............................................................... 26
1.8 Columnar organisation in primary auditory cortex. .................................. 27
1.9 Classes of inhibitory interneurons. ...................................................... 28
1.10 Non-exhaustive schematic of a cortical microcircuit comprising inhibitory and excitatory cells. ................................................................. 28
1.11 Identified mechanism for FM direction selectivity in A1 based on inhibitory-excitatory interaction. .......................................................... 29
1.12 Two distinct mechanisms can underlie FM direction selectivity. ................... 30
1.13 Neurons in auditory cortex exhibit distinctly tuned excitatory responses to the onset and offset of sound. ......................................................... 31
1.14 ‘ABA’ streaming paradigm. ................................................................ 31
1.15 Attention influences cortical selectivity to sound features. ....................... 32
1.16 Dynamic increase in A1 neuronal response during an auditory attention task involving temporal expectation. ................................................. 32
1.17 Non-exhaustive schematic of anatomical circuits involved in attentional processes. ................................................................. 33
1.18 Techniques to selectively manipulate projecting neurons. ......................... 33

2.1 Multi-electrode array used for in vivo recording. ...................................... 55
2.2 Apparatus positioning during in vivo electrophysiology in the auditory cortex of anaesthetised mice. .......................................................... 55
2.3 Apparatus positioning during in vivo electrophysiology in the auditory cortex of awake mice. ................................................................. 56
2.4 Head-implant for in vivo electrophysiological recording in awake mice. ........ 56
2.5 Software for auditory stimulation in passive listening conditions. ................ 57
2.6 Software for auditory stimulation in active behaving conditions. ................. 57
2.7 Hardware used during behavioural training. .......................................... 58
2.8 User interface used to monitor mouse performance online during the FM direction discrimination task. ......................................................... 58
2.9 User interface used to monitor mouse performance online during the oddball task (T2). 59
2.10 FM discrimination task design. ............................................................ 59
2.11 Oddball task design (T1). ................................................................. 60
2.12 Hardware for auditory stimulation combined with electrophysiological recording in vivo in passive listening conditions. ......................... 60
2.13 Hardware for auditory stimulation combined with electrophysiological recording in vivo in active behaving conditions. .......................................................... 61
2.14 Example single cells sorted using Klusta. ................................................. 61

3.1 Hebbian plasticity could underlie developmental divergence of ON and OFF receptive fields in auditory cortex. ................................................................. 79
3.2 Proposed functional significance of segregated ON/OFF RFs. ................. 80
3.3 ON/OFF cells discriminated best the direction of slow FM sweeps. .......... 81
3.4 A1 cells discriminated best the direction of slow FM sweeps. ............... 82
3.5 The relationship between DSI and CF ON depended on the FM speed. ........ 83
3.6 ON/OFF arrangement was a good predictor of direction selectivity at slow FM speed. ..................................................... 84
3.7 Details and validations of pharmacogenetic experimental procedure. ........ 85
3.8 Neural activity increased Post-CNO ..................................................... 86
3.9 ON/OFF cells conserved their characteristics Post-CNO. ..................... 87
3.10 The relationship between DSI and CF ON was conserved Post-CNO. .... 88
3.11 Direction selectivity remained correlated with ON-OFF arrangement Post-CNO. ................................................................. 89
3.12 Direction selectivity in Young ON/OFF A1 neurons. ......................... 90
3.13 Direction selectivity of ON/OFF A1 cells in awake adult mice under pharmacogenetic manipulation. ......................................................... 91
3.14 Perception of FM direction was preserved following reduction of cortical inhibition. ................................................................. 92
3.15 Multi-unit direction selectivity was correlated with BF ON during behaviour. ........................................................................ 93

4.1 Complex discrimination task involving spectrotemporal bias. ................ 115
4.2 Mice were sensitive to temporal bias as revealed by their false alarm timing. 115
4.3 Mice timed their false alarm responses to tone onsets. ......................... 116
4.4 Mice were sensitive to spectral bias as revealed by their false alarm timing. 116
4.5 Spiking activity was modulated by task specificity. .............................. 117
4.6 Increase in spiking activity prior to target occurrence irrespective of frequency tuning. ............................................................... 118
4.7 Dynamic increase in spiking activity prior to target occurrence. ............... 119
4.8 Modulation in evoked and baseline spiking activity prior to target occurrence. ..................................................... 120
4.9 Predicting FA specificity using neural responses during Hit trials. ............. 121
4.10 Dynamic enhancement of signal to noise ratio prior to target occurrence. .... 122
4.11 Modulation of firing rate prior to target occurrence was correlated with behaviour. ............................................................. 123
4.12 Tone-Target discrimination performance of single cells in auditory cortex was correlated with behaviour. ......................................................... 124
4.13 LFP phase-reset was moderately visible in anaesthetised mice. ............. 125
4.14 LFP phase-reset was moderately visible in active behaving mice. .......... 126
4.15 Facilitative modulation of the LFP phase distribution following repetitive stimulus presentation. ................................................................. 127
4.16 Facilitative modulation of the LFP phase distribution prior to target occurrence. ............................................................. 128
4.17 Modified discrimination task involving spectrotemporal bias. ................. 129
4.18 Mice were sensitive to the spectrotemporal bias in the modified discrimination task as revealed by their false alarm timing. ......................................................... 130
4.19 Thalamic cells displayed excitatory and inhibitory responses.

4.20 Thalamic cells displayed an increase in firing rate prior to target occurrence irrespective of frequency tuning.

4.21 MGB cells did not change tuning profile following reduction of PV-expressing TRN cells activity.

4.22 Testing the CAV preparation.

I.26 The direction selectivity index was invariant to the underlying measure.

I.27 Autocorrelograms and waveforms of example single cells recorded to quantify ON-OFF CF difference and DSI in vivo.

I.28 Poor relationship between DSI and CF ON for Young ON/OFF A1 neurons.

I.29 Measured and predicted DSI values of Young ON/OFF A1 neurons.

I.30 Relationship between DSI and CF ON for Multi-unit data recorded in awake passive-listening mice.

I.31 Spike in LFP power as a marker of hM4i activation in vivo.

I.32 Behavioural performance of hm4i-PV+ mice.

I.33 Multi-unit activity was not modulated by task outcome.

I.34 Behavioural performance of hm3-PV+ mice.

I.35 Post mortem microscopy confirmed the expression of DREADD in auditory cortex of mice used in behaviour.

I.36 Future work: Assessing the selectivity of cortical inputs.

I.37 Future work: Mapping ON/OFF responses in auditory cortex using Calcium imaging.

I.38 Beneficial future changes in the FM discrimination behavioural paradigm.

I.39 Post mortem microscopy confirmed the electrode placement in prefrontal cortex during recording in behaviour (oddball task T1).

I.40 Post mortem microscopy confirmed the electrode placement in auditory cortex during recording in behaviour (oddball task T1).

I.41 Comprehensive schematic of analytical procedure undertook to compute the difference in LFP phase distribution between Early and Late trial blocks (relates to figure 4.16).
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>primary auditory cortex</td>
</tr>
<tr>
<td>AC</td>
<td>auditory cortex</td>
</tr>
<tr>
<td>BF</td>
<td>best frequency</td>
</tr>
<tr>
<td>CF</td>
<td>characteristic frequency</td>
</tr>
<tr>
<td>CNO</td>
<td>clozapine-N-oxide</td>
</tr>
<tr>
<td>DP</td>
<td>discrimination performance</td>
</tr>
<tr>
<td>DREADD</td>
<td>designer receptor exclusively activated by designer drug</td>
</tr>
<tr>
<td>DSI</td>
<td>direction selectivity index</td>
</tr>
<tr>
<td>EZ</td>
<td>early zone</td>
</tr>
<tr>
<td>FM</td>
<td>frequency modulation</td>
</tr>
<tr>
<td>FRA</td>
<td>frequency response area</td>
</tr>
<tr>
<td>FSI</td>
<td>frequency selectivity index</td>
</tr>
<tr>
<td>GABA</td>
<td>gamma-Amino butyric acid</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>IC</td>
<td>inferior colliculus</td>
</tr>
<tr>
<td>IP</td>
<td>intraperitoneally</td>
</tr>
<tr>
<td>KS</td>
<td>Kruskal Wallis test</td>
</tr>
<tr>
<td>LFP</td>
<td>local field potential</td>
</tr>
<tr>
<td>MGB</td>
<td>medial geniculate body</td>
</tr>
<tr>
<td>MU</td>
<td>multi-unit</td>
</tr>
<tr>
<td>n.s.</td>
<td>not significant</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffer saline</td>
</tr>
<tr>
<td>PFA</td>
<td>paraformaldehyde</td>
</tr>
<tr>
<td>PFC</td>
<td>prefrontal cortex</td>
</tr>
<tr>
<td>PSTH</td>
<td>peri-stimulus time histogram</td>
</tr>
<tr>
<td>PV</td>
<td>parvalbumin</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristic</td>
</tr>
<tr>
<td>RT</td>
<td>reaction time</td>
</tr>
<tr>
<td>SC</td>
<td>subcutaneously</td>
</tr>
<tr>
<td>SNR</td>
<td>signal to noise ratio</td>
</tr>
<tr>
<td>SOA</td>
<td>stimulus onset asynchrony</td>
</tr>
<tr>
<td>TRN</td>
<td>thalamic reticular nucleus</td>
</tr>
<tr>
<td>V1</td>
<td>primary visual cortex</td>
</tr>
<tr>
<td>VS</td>
<td>vector strength</td>
</tr>
<tr>
<td>WRS</td>
<td>Wilcoxon rank sum test</td>
</tr>
<tr>
<td>WSR</td>
<td>Wilcoxon signed rank test</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Challenges in hearing

The sense of hearing is fundamentally important for survival, maintenance, and reproduction, as it provides animals with rich information about the surrounding environment. From the pattern of acoustic waves reaching the ears, it is possible to detect, identify, and locate a multitude of objects or events. The sense of hearing also serves as one half of an essential communication loop, integral for mating and warning within species. Whilst it is generally accepted that the auditory cortex (AC) bears an essential role in these processes\[^1\], the mechanistic basis remains unclear. To uncover the role of AC in sound processing and perception, it is important to review some of the challenges faced by the auditory system when aiming to make sense of sounds.

Sounds are pressure waves that propagate through a medium, most commonly air, and as such display two fundamental characteristics: frequency and amplitude. These properties can vary dynamically, as captured in spectrograms (see figure 1.1a for examples). Whilst it is possible to determine the nature of a sound given its isolated spectrogram, the auditory system only receives this information sequentially. Integrating past and present events to form a complete object representation is a hallmark of auditory processing, in which the AC might be greatly involved.

In addition, natural scenes often comprise multiple sound sources. In this case, the acoustic wave reaching the ears is a mixture of all sounds (see figure 1.1b for example spectrograms of a sound presented in isolation and in a noisy environment), and the auditory system is faced with the challenge of segregating relevant information from background noise. The auditory system is particularly well-versed in completing this task, as demonstrated in the classic "cocktail party" situation, where two people are able to have a conversation in a crowded, noisy place\[^2\]. As it is difficult to discern the contributing components of a sound recorded in ambient noise by simply inspecting its spectrogram, the auditory system must employ complex mechanisms for animals to comprehensively perceive their surroundings. Moreover, since noisy environments contain by definition a vast amount of information of which only a subset is of relevance (such as the conversation of a single person amongst other speakers), the brain must be able to preferentially extract and amplify features of the auditory scene based on behavioural interest.

These two fundamental processing challenges, integration of information over time and preferential feature extraction, are at the core of the studies undertaken as part of this thesis. In particular, I aimed to understand what role the auditory cortex plays in solving each of these challenges. The next sections are thus arranged to reflect these two major questions. First, I review our current understanding of the potential mechanisms underlying the encoding of frequency modulation, a property necessitating temporal integration and of vital importance for vocalisation perception. Next, I review experimental findings demonstrating the adaptive nature of auditory processing, and discuss potential mechanisms. These reviews provide the relevant background for the research presented in chapter 3 Functional determinants of response selectivity to frequency modulated sounds and chapter 4 Anatomical circuits underlying task-related modulation of spectrotemporal receptive fields respectively.
1.2 On the encoding and perception of frequency modulation

Most natural sounds contain variations (or modulations) in frequency. In several species, these modulations define the principal syllables of vocalisation, as seen in mouse calls\(^3\) (see figure 1.2a), cat calls\(^4\), monkey calls\(^5\), bird songs\(^6\), and human speech\(^7\). Frequency modulations (FM) are defined by their speed (rapid or slow) and direction (upward or downward). Smooth transitions in frequency are called 'sweeps'. Sweeps can be measured either in Hz per second (Hz/s; linear sweep) or in octaves per second (oct/s; logarithmic sweep). For example, in the mouse call presented in figure 1.2b, the change in the main spectral energy distribution corresponds to a relatively slow, upward logarithmic FM sweep. Because FM sweeps are such essential components of vocalisation, it is crucial to understand how they are represented in the brain. Furthermore, as vocal calls contain FM sweeps of different speed and direction (for example, see the distribution of FM in mouse calls in figure 1.2c), the brain must develop mechanisms to encode both of these aspects.

Considered as the fundamental building block of neural coding\(^8\), neurons are electrically excitable cells which produce impulses, called spikes, to transmit information. They are organised in large networks, connecting with each other via synapses. Variations in the number of generated spikes is one of the many strategies used by neurons to encode information. Frequency modulation could thus be represented by specialised auditory neurons using variations of spiking activity.

Accordingly, previous studies investigating neural responses to FM sweeps in the auditory system of a number of vertebrates (including birds, rats, ferrets, cats, and monkeys\(^5,9,10\)) have revealed that the strength of neural spiking response depends on FM speed and direction. An example neuron exhibiting selective spiking response to FM direction is displayed in figure 1.3. In this example, a strong response was elicited only after hearing an upward, but not downward, FM sweep. This selective property is quite remarkable, as it indicates that the neuron must have integrated the incoming sound information over time to generate its spiking output. Indeed, if the neuron had simply responded upon the presentation of a specific frequency contained in the FM sweep, upward and downward FM sweeps would have elicited responses of similar magnitude as they covered the same frequency range.

How does this selectivity emerge? Does this selectivity relate to our perception of FM sounds? In the next sections, I discuss how the study of neural responses to a simpler type of stimulus (a pure tone) may enrich our understanding of FM processing. Then, I detail how neural responses to pure tones are conveyed along the auditory pathway and integrated in auditory cortex. Further, I present theories on how integration properties may give rise to FM encoding, as well as techniques to test these theories \textit{in vivo}. Finally, I review evidences supporting the necessity of auditory cortex processing in the perception of FM sounds.

1.2.1 Studying frequency modulation encoding: Breaking down the problem

One approach to examine how complex sound features may be encoded consists of decomposing the complex property in its fundamental attributes. FM sweeps can be at first approximated by successive pure tones of different frequencies, as presented in figure 1.4. Using solely two tones (‘Two tones’ paradigm), it is possible to demonstrate the sensitivity of a neuron to stimulus context\(^9,11,12\). Context here refers to the recent history of auditory stimuli, including for example the order in which sounds are presented, and the interval between presentations. In the example paradigm presented in
figure 1.5a, the frequency of the second tone is kept constant (the constant sound is called a probe; see orange square in figure 1.5a) whilst the frequency of the preceding tone is varied (this sound is called a masker; see black squares in figure 1.5a). Notably, if the probe is of middle frequency, presenting a high and low masker respectively can approximate downward and upward FM sweeps.

If neurons were insensitive to stimulus context, the response to the probe should be similar following any of the masker. In AC, however, this is not the case, as demonstrated by the example neuron of figure 1.5a. Spiking activity in response to a tone can be facilitated or suppressed depending on the spectral content of the preceding sound [9, 11, 12]. Similarly, spiking activity in AC depends on the timing between successive sound presentations [13] (see example neuron in figure 1.5b). These dependencies are reminiscent of those observed when measuring neural responses to FM sweeps [5, 9, 10]. Thus, it appears that our understanding of neural responses to pure tones may elucidate the mechanisms underlying selectivity to FM properties. In the next section, I review how the representation of pure tones emerges in the auditory system, and discuss how the processing of tone may relate to that of FM sweeps.

### 1.2.2 Representation of frequency content along the ascending auditory pathway

A common feature of sensory cortical areas is that they contain topographic representations of the receptor surface [14]. In the auditory system, the cochlea acts as receptor and transducer, transforming sound waves into electrical impulses. The cochlea is a spiralled, hollow chamber of bone filled with fluid, in which waves propagate from the base to the apex (a schematic of the cochlea is given in figure 1.6a). Upon reaching the ear, sound waves travel through the ear canal and hit the tympanic membrane, causing it to vibrate. These vibrations are transferred by the auditory ossicles (notably by the stapes in contact with the oval window) to the fluid in the cochlea, resulting in the displacement of the basilar membrane contained within that latter structure. This displacement pattern peaks at a distance from the oval window depending upon the sound wave frequency. For example, a low frequency sound will generate maximal displacement at the apex, as displayed in figure 1.6a, whilst a high frequency sound with generate maximal displacement at the base. This displacement causes the depolarisation of hair cells, which send signals to the first auditory neurons [15]. This spatial arrangement of frequency-tuned hair cells along the basilar membrane generates a place code, otherwise known as cochleotopy or tonotopy.

In both humans and mice, this arrangement is preserved throughout the different stages of the auditory pathway (see figure 1.6a, b). Tonotopic organisation is visible at the level of the cochlear nucleus (CN), superior olive (SO), inferior colliculus (IC), medial geniculate body (MGB) and auditory cortex (AC) [16–18]. Note that the ascending auditory system comprises two parallel pathways, one topographically organised (lemniscal pathway) and one where tonotopy is less apparent (non-lemniscal pathway) [14, 19, 20]. In mice, the AC can be further divided into sub-regions depending on the frequency-preference organisation [16–18] (see figure 1.6c). The primary auditory cortex (A1) presents a low-to-high frequency gradient on the posterior-to-anterior axis (although note that the definition of A1 based on tonotopy is currently disputed [18]).

Interestingly, the neural selectivity for FM direction has been found to be organised along this tonotopic axis in the primary auditory cortex of rats [9] (see figure 1.7) and squirrel monkeys [10], as
well as in sub-cortical regions such as the IC and MGB\textsuperscript{21}. Low-frequency tuned neurons often preferred sweeps with FM from low to high frequencies (UP), and vice versa for high-frequency tuned neurons and DOWN FM sweeps. This degree of organisation suggests that FM sound processing may depend on the mechanisms contributing towards a neuron’s tuning to pure tones\textsuperscript{14}. With the aim to delineate the anatomical basis of these contributions, a brief review of auditory cortical circuit connectivity is outlined below.

1.2.3 Cellular diversity and organisation in primary auditory cortex

The cerebral cortex consists of six layers: the molecular layer (L1), the external granular layer (L2), the pyramidal layer (L3), the inner granular layer (L4), the inner pyramidal layer (L5) and the multiform layer (L6). The figure \textbf{1.8a,b} presents an outline of the layer boundaries in mouse A1. Each of these layers comprises distinct cell types, examples of which are given in figure \textbf{1.8c}, which serve different roles in sensory information encoding and transmission as described below.

The layers of cerebral cortex can be divided into three functional components: the supragranular layers (L1-3), the inner granular layer (L4) and the infragranular layers (L5-6). A comprehensive map of anatomical connections is presented in figure \textbf{1.8d}. The supragranular layers are the primary origin and termination of intracortical connections, which are either associational (i.e. forming connections with other areas of the ipsilateral hemisphere) or commissural (i.e. forming connections with the contralateral hemisphere, primarily through the corpus callosum). The inner granular layer receives thalamo-cortical connections, especially from specialised thalamic regions, such as the MGB for A1. The infragranular layers primarily connect the cerebral cortex with subcortical regions. In A1, sensory stimulation evokes an information flow across this “canonical” microcircuit: the thalamic input recipient layer (L4 and upper L6) sends information to the supragranular layers (L2/3), which in turn drives the infragranular, output layers (L5/6). It is important to highlight that evoked and spontaneous activity propagate differently across cortical layers and columns\textsuperscript{22,23}, with spontaneous patterns showing earliest activation in deep, infragranular layers and spreading upwards across layers.

Cortical neurons can be divided into two major classes. The first class comprises principal neurons, defined by their use of the excitatory neurotransmitter glutamate. Because of their shape, these cells are also named pyramidal neurons, and constitute \textasciitilde80\% of cortical neurons in rodents. Principal cells contact both local and distant targets through extensive axonal projection\textsuperscript{22,24,25}, as presented in figure \textbf{1.8d}.

The remaining \textasciitilde20\% are neurons that release the inhibitory neurotransmitter gamma-Amino butyric acid (GABA)\textsuperscript{26–28}. In contrast to principal cells, inhibitory neurons mainly synapse with local targets\textsuperscript{25,27}. Because of this property, inhibitory cells are also called interneurons. The mammalian cortex comprises four major classes of interneurons: Somatostatin (SOM)-expressing cells, Parvalbumin (PV)-expressing cells, vasoactive intestinal peptide (VIP)-expressing cells, and neurogliaform (NG) cells, as displayed in figure \textbf{1.9a} (although VIP and NG cells can be seen as belonging to the same class as they both express a similar type of serotonin receptor\textsuperscript{29}). Each of these classes can be further divided into multiple subclasses, based on molecular expression, morphology or synaptic target. For example, PV+ interneurons can be categorised into two main subgroups: basket and chandelier cells, each with distinct projection targets and morphologies, as displayed in figure \textbf{1.9a,b}. 

4
In accordance with their morphological heterogeneity, neurons from each class exhibit various synaptic integration and spike firing abilities\[^{22,24,25,27,28}\]. For example, in the AC of rodent, principal cells in L2/3 generate spikes in a sparse manner\[^{22}\], whilst PV-expressing cells are capable of firing rapidly\[^{30}\].

Pyramidal cells and interneurons form an intricate network (a non-exhaustive diagram is presented in figure 1.10), with various feedforward and feedback loops. The innervation of thalamic cells onto cortical neurons is one example of a feedforward loop: an incoming signal from the thalamus excites both a L4 pyramidal cell and a PV interneuron, which in turn inhibits the L4 pyramidal cell (this effect is called feedforward inhibition). All neurons in the cortex receive both excitatory and inhibitory inputs\[^{27}\]. In the case of the auditory cortex, the inhibition has a purely cortical origin\[^{9}\].

Through recruitment of interneurons via feedforward and feedback loops, the inhibition generated in a cortical network is approximately proportional to the incoming excitation\[^{26,27,31}\]. This balance is maintained over a wide range of input properties, as seen for example in the AC: a change in the stimulus frequency leads to a coordinated change in the strength of both excitatory and inhibitory inputs onto a neuron\[^{9,31–33}\]. Whilst it is acknowledged that co-occurrence of excitation and inhibition is necessary for proper cortical functions\[^{26,27,34}\], the physiological advantages of such a system are not fully understood. In the next section, I review theories of the dependence of auditory processing on these interactions, especially in response to frequency modulated sounds.

1.2.4 Sound processing within cortical circuits - A review on inhibitory and excitatory contributions

How the interplay between excitation and inhibition shapes cortical functions is a major question in sensory neuroscience\[^{35}\]. Two main schemes have been proposed. First, the difference (or similarity) in tuning between excitatory and inhibitory inputs can contribute to stimulus selectivity, by enhancing responses in a linear or multiplicative manner\[^{32,35,36}\]. Second, the difference in timing between excitatory and inhibitory inputs can contribute to stimulus selectivity\[^{9,31}\]. Whilst they are two separate mechanisms, the variations in both the input’s strength and delay are likely to jointly affect sensory processing.

In AC, the inhibitory input always follows the excitatory input by a brief temporal delay (in the order of several milliseconds), resulting in a reduction of excitation through inhibition\[^{9,32}\]. In response to pure tones, inhibitory and excitatory synaptic inputs manifest similar tuning, and their balance is approximatively maintained across a large range of stimuli\[^{9,31–33}\]. The slight mismatch between inhibition and excitation in the flanking (i.e. side) regions of the tuning curve has been proposed to underlie ‘inhibitory sidebands’\[^{32}\], i.e. zones of the receptive field where the inhibitory response to a first tone can putatively prevent any spiking response to a following tone\[^{9,11,12}\]. Further contributing towards the formation of inhibitory sidebands, AC neurons receive spectrally skewed inputs\[^{9,11}\], as displayed in figure 1.11a. In this example, a low frequency tuned neuron receives the strongest synaptic inputs upon the presentation of a low frequency tone (defined here as its best frequency), but also smaller non-null inputs following the presentation of slightly higher (but not lower) frequency tones than its best frequency.

A correlation between inhibitory sideband and FM direction selectivity has been observed\[^{9}\], which led to the proposition of a model for FM direction selectivity shaping in AC. This model (presented in
Figure 1.11 relies on four parameters: the spectral asymmetry of the synaptic input tuning profile, the co-tuning of excitatory and inhibitory inputs, the rough proportionality of these inputs' strengths, and the out-run of excitation by inhibition (i.e. the effect of inhibition must last longer than excitation). In this model, the presentation of two successive tones of different frequencies (a proxy for FM) can facilitate or depress a neuron's response depending on its inputs' tuning skewness, as shown in figure 1.11b. Based on the measured asymmetry of AC neurons' tuning\cite{9,11}, the response to upward and downward FM should in this scheme be facilitated in low and high frequency tuned neurons respectively.

Whilst this model has received considerable supporting evidence, it cannot account for all effects observed when stimulating AC neurons with FM sweeps\cite{9}. In particular, cells in AC can be sensitive to the direction of slow (< 25 oct/s) FM sweeps\cite{37–39}, whilst in this model FM speeds of ∼40 oct/s and above are required to impact on direction selectivity. Moreover, given the heterogeneity of interneurons and connections, the anatomical source of inhibition is presently unclear. Further, inhibition does not always out-run excitation when measured in response to pure tones\cite{31–33}, raising the question how FM direction selectivity may otherwise be implemented.

To answer to these questions, it is important to first delineate the contributions of different inhibitory cell types towards FM encoding, and second to consider the interactions of other neural properties which occur in response to sounds. The two following sections provide an overview of respectively new techniques and theories, which may enable the further understanding of auditory cortical functions.

1.2.5 Techniques to selectively manipulate neural activity

In recent years, genetic techniques have enabled researchers to selectively manipulate the activity of specific, targeted cells\cite{40–42}. In this aspect, they differ from the previously available techniques for neural activity manipulation, such as ablation, thermal cooling or electrical stimulation, which affect all cells in the vicinity of the perturbation locus. The two main types of genetically targeted manipulation currently are optogenetic and pharmacogenetic (or chemogenetic). Both techniques rely on the activation of engineered ion channels, whose expression can be restricted and genetically targeted to a specific neural population. These channels are sensitive to a particular agent, i.e. light for optogenetics and drug for pharmacogenetics, and as such are given the name of receptors.

Several techniques exist to selectively express these receptors onto the membrane of specific cells. One of them consists in manipulating an organism's genome, by adding another DNA strand encoding for an engineered protein (such as the receptor) next to a gene which is natively expressed in the cells of interest. This type of genetic manipulation is successfully achieved in animals which replicate rapidly, such as mice or drosophilae. As genetically modified animals are bred, it is common to refer to them as part of a ‘line’ of transgene. Of importance, all cells in an organism contain the same DNA code, but express only a finite portion of the encoded genes. It is thus necessary to initially identify a gene selectively expressed in the cells of interest prior to using these techniques. For example, in order to manipulate the activity of all interneurons simultaneously, the vesicular GABA transporter (VGAT) is a suitable target gene as it is expressed in all interneurons. However, to manipulate only the interneurons with axonal branches terminating onto the soma of pyramidal cells, parvalbumin (PV) would be a more suitable gene to select.

Once a gene of interest has been identified, the choice of receptor remains. In optogenetics as
in pharmacogenetics, two main types of receptors exist, with opposite functional effects: excitatory or inhibitory receptors. Channelrhodopsin (ChR) and archaerhodopsin (Arch) are two examples of light-sensitive proteins, which excite or inhibit a cell if activated\textsuperscript{42}. One prominent type of pharmacoreceptor is the DREADD\textsuperscript{43} (designer receptor exclusively activated by designer drug), which is activated by application of CNO (clozapine-N-oxide), converted in the brain into clozapine\textsuperscript{44}. It is important to highlight that the activating drug used in pharmacogenetics must minimally interfere with other brain functions to avoid potential confounds. The choice for the use of opto- or pharmacoreceptor mainly depends on two parameters: firstly, the temporal dynamics required; and secondly, the level of surgical invasiveness permitted. The use of opsins enables the temporally precise control of neural activity, but often requires a surgically invasive preparation (such as the implantation of an optic fibre in the brain), as light needs to be delivered directly onto the neurons with sufficient power. In contrast, DREADD acts over an extended period, with a slow rising phase and a prolonged sustain phase (in the order of tens of minutes and hours respectively), and can be activated minimally invasively by injection of CNO. Another major difference between DREADDs and opsins appears is their potential to restrict the neural manipulation to a specific locus. While it is possible to focus light onto a particular spot, it is not possible to constrain the effect of CNO onto specifically located receptors after injection.

However, engineers have developed an ingenious system enabling the selective and local expression of specific proteins\textsuperscript{40,42}. This system is based on two mechanisms. Firstly, it relies on the Cre/loxP recombination mechanism, i.e. the ability of Cre-recombinase to induce the recombination of loxP recognition sequences\textsuperscript{40,42}. The loxP sites can be inserted into a DNA strand so as to flank the exons of a gene of interest (referred to as a ‘floxed’ gene)\textsuperscript{40}, preventing the gene expression under the absence of Cre-recombinase. Secondly, this system relies on viral strategies to deliver the ‘floxed’ gene strand into cells locally\textsuperscript{41,42}. As it is possible to breed animals to express Cre-recombinase in specific cell types (such as those expressing PV), this system offers a versatile means by which one can achieve the selective and local expression of an engineered protein. For example, in order to specifically manipulate PV-expressing interneurons in auditory cortex, one could do the following. First, use the transgenic mouse line PV-Cre, in which PV-expressing cells co-express Cre-recombinase. Second, inject a virus that contains the floxed-gene encoding for DREADD directly into the auditory cortex. As the virus infects all cells near the injection site, while only cells with Cre-recombinase can express the floxed-gene delivered by the virus, the receptor DREADD will only be expressed in the PV-expressing cells of auditory cortex.

Taking advantage of these techniques, it is possible to causally determine the functional role of specific cell types in neural computation. I will leverage these techniques to study the contribution of targeted interneurons towards the encoding of FM sounds, as presented in chapter 3 \textit{Functional determinants of response selectivity to frequency modulated sounds}.

1.2.6 \textbf{Cortical processing beyond inhibitory-excitatory interaction}

The interaction of inhibitory and excitatory synaptic inputs are at the core of several computational theories for cortical processing\textsuperscript{45}. Yet, a multitude of other factors also play a role in shaping a neuron’s response to sensory inputs, such as for example short term-plasticity of synaptic strength, active dendritic properties, and non-linearities in synaptic integration and spike generation\textsuperscript{45–48}.
Accordingly, neurons in the AC exhibit substantially diverse response profiles following sound presentation. For example, one neuron may display a sustained response throughout the whole duration of a sound, whilst another neuron may respond only transiently. Rather than the sole interaction between excitation and inhibition, a mixture of these response properties might be necessary to produce selectivity to high order sound features such as FM.

In the visual cortex, direction selectivity has been suggested to emerge not only from the interaction of excitatory and inhibitory inputs (Exc-Inh model), but also from pure excitatory contributions\(^47\) (Exc-Exc model). Recently, these models have been reviewed in the context of FM direction selectivity shaping in the auditory cortex\(^{46}\), as evidence supporting both models has been observed in vivo\(^{9,21,46,49}\). These models are presented in figure 1.12. However, whilst the source of inhibition in AC can be ascribed to interneurons, the source of excitation presented as part of the Exc-Exc model is currently unclear.

Since the Exc-Exc model for direction selectivity has first been advanced by studies in the visual system, it is important to review some properties of this latter system. In vision, the retina transforms the incoming light into electrical signals. Axons of retinal ganglion cells form the optic nerve, which transmits these signals to neurons in the lateral geniculate nucleus (LGN), which in turn projects to the primary visual cortex (V1). V1 neurons are tuned to the orientation and movement direction of visual gratings, the latter property being an emergent feature of V1. Cells in V1 are excited by changes in light brightness, either when it turns on (ON response) or off (OFF response). ON and OFF responses are organised in subfields, which are spatially distinct (see example in figure 1.13c). If these subfields had different excitatory latencies, they could contribute towards the shaping of direction selectivity as proposed in the Exc-Exc model\(^{47}\).

When comparing the visual to the auditory system, V1 appears more similar to the IC rather than the AC based on anatomical hierarchy and processing capacity. Nonetheless, theories derived from V1 computations were often proven useful in understanding AC processing\(^1\), which is why, in the next paragraph, I assess whether the Exc-Exc model could be supported in A1 via the same mechanisms as proposed in V1.

Cells in the AC respond to the onset of sounds, but also to their offset\(^{50–52}\), an example of which is displayed in figure 1.13a. This property has been reported in a variety of species (such as in mouse, cat and monkey)\(^{50–52}\), and was at first considered a rebound from inhibitory release following sound termination\(^{50–52}\). It has now been showed that offset responses are not just by-products of inhibition, but are emerging responses that are transmitted from early levels of the auditory system, visible as early as in the superior olive in mice\(^{50–53}\). Interestingly, distinct anatomical channels are sensitive to onset and offset, as demonstrated in the MGB by studies in mice\(^{54}\) and guinea pig\(^{55}\). In the AC, these channels converge onto single cells in a non-overlapping manner, so that onset and offset responses are elicited via different sets of synapses\(^{51}\). As seen in the visual modality, cortical neurons of the auditory system exhibit distinct tuning properties in response to stimulus onset and termination, as shown in figure 1.13a-c. Whilst such a difference has been observed in several species\(^{50–52}\), its functional significance remains to be determined. To assess whether responses to onset and offset may contribute to the encoding of high order sound feature such as FM is one of the main aims of the research presented in chapter 3 Functional determinants of response selectivity to frequency modulated sounds.
1.2.7 The role of auditory cortex in sound perception

There are two fundamental quests in sensory neuroscience. One is to link a specific sensory processing function to an anatomically defined circuit within the sensory pathway, and the other is to link sensory processing functions to perception and behaviour\[35\]. In the auditory system, the AC might hold a special role in the formation of sound perception\[1\]. In this section, I present the concept of perception, outline methods to uncover its underlying anatomical basis, and review the role of AC in the perception of sounds, in particular of FM sweeps.

“Ni xiang he dian cha ma?” Promptly answered by 1.4 billion people, this question might hold more than one's tongue on the European continent. This discrepancy illustrates one major determinant of our interactions with our surroundings: whilst we may receive comparable sensory inputs, we rely on the perception of these inputs to take action. Since perception may define our behaviour, it is crucial to understand how it emerges, and how it can be manipulated.

Sensory perception involves the detection of a stimulus, and the subsequent recognition and characterisation of it. For centuries, the anatomical basis of these abilities was a puzzle to a number of philosophers and scientists. Based on two main findings, the brain is nowadays viewed as the major orchestrator of perception. Firstly, stimulus representations are observable in this organ, for example in neurons as previously described. Secondly, changes in neural signals can correlate with changes in perception and behaviour - striking examples of which can be found in the reports of O. Sacks\[56\]. However, these correlations do not imply causations (a fallacy as old as the Roman Empire; cum hoc ergo propter hoc). A fundamental quest in sensory neuroscience is thus to determine whether neural processing functions causally relate to perception.

One rigorous approach to link brain signals to perception is pertubing the activity of the brain in a controlled manner, and assessing corresponding changes in behaviour and perception. Such techniques are available to humans (an example of which being transcranial magnetic stimulation), but are more extensively developed in animal models, permitting more invasive preparations (some of which were presented in the previous section). As animals can be trained to perform complex tasks, such as stimulus detection and discrimination, it is possible to relate neural processing to their perception via behavioural assessments.

Such an approach has been used to determine the role of the auditory cortex in FM sweep perception. Mongolian gerbils were trained to discriminate the direction of FM sweeps\[57–59\], using an avoidance-conditioning paradigm. Ablation of their auditory cortex resulted in a net decrease in performance, suggesting a critical role for this brain region in FM sound perception. Further support for the necessity of intact AC processing in sound perception was reported in rats using the silencing agent muscimol\[60,61\]. As more refined techniques are now available to manipulate cortical activity\[40\], it is an exciting time to aim at linking specific neural circuitry to sensory perception.

In chapter 3 Functional determinants of response selectivity to frequency modulated sounds, I take advantage of the genetic tools developed in mice, and the fact that these animals are great learners of complex tasks\[62\], to investigate whether specific interneurons in AC are involved in the perception of FM direction.

\[1\] “Would you like some tea?” in Chinese.
1.2.8 Summary of research aims

The main aims of the research presented in chapter 3 *Functional determinants of response selectivity to frequency modulated sounds* are the following:

1. To test for the functional significance of segregated ON and OFF receptive fields in A1, specifically towards the encoding of FM sound direction.

2. To test the role of specific A1 interneurons in FM sound direction encoding.

3. To test whether the inhibition mediated by specific A1 interneurons is necessary to the perception of FM sound direction.
1.3 Sound perception in noisy environment

A natural scene often comprises several acoustic sources, but at given times, attending to only one of them may be of interest. For example, in a cocktail party[^2], you might want to listen to the person on your right, and not be distracted by what the person on your left is saying. In such a noisy situation, attention allows us to direct a ‘searchlight’ towards an object of interest in the environment[^63]. This searchlight can be directed voluntarily (‘top-down’ attention), or be captured following a salient event (‘bottom-up’ attention).

Voluntary attention is known to enhance sensory processing and perception in a selective manner[^2,63–65]. However, the underlying anatomical mechanisms are unknown. This is potentially due to the fact that attention is an intricate phenomenon to study, as it involves at least three complex processes: competitive selection, top-down sensitivity control, and working memory[^64,66]. Here, I focus on studying the top-down sensitivity control aspect of attention, more specifically how sound representation can be selectively modified to facilitate the achievement of behavioural goals. In particular, I aim to unravel the mechanisms leading to such changes at the level of the auditory cortex (AC), a region of the brain important for sound perception, and whose processing can be strongly modulated by attention (see below).

The next sections provide the relevant background to answer this question. Firstly, I give an overview of the challenges faced when aiming at making sense of sounds in a natural environment, a problem known as the ‘auditory scene analysis’[^67]. Next, I present evidence for the modulation of neural activity and perception by attention, and review the role of the AC and and other brain regions in mediating attentional effects. Further, I discuss potential anatomical circuits for the selective enhancement of specific sounds by attention. Finally, I present techniques to study the functional role of these anatomical circuits.

1.3.1 Auditory scene analysis: The challenge of forming sound streams

A natural scene often comprises several acoustic sources (e.g. someone playing the piano or walking), which usually generate acoustic events (e.g. a single piano note or the sound of a footstep) in a sequential manner. These sequences may overlap in time, in which case the sound wave reaching the ears constitutes a mixture of all simultaneous acoustic events. To form a useful representation of the scene, the auditory system must assign the individual sound elements within this incoming mixture to the correct auditory sources - in other words, it must ‘analyse the auditory scene’ to create the perception of distinct ‘auditory objects’[^67]. The sequence of sound perceived as part of a single coherent entity (or object) is named a ‘stream’ of sound[^67].

The process of stream formation often requires the segregation of overlapping sound features. This is a highly complex problem, as illustrated in figure 1.1: based on a spectrogram representation, it is difficult to distinguish the word “shoe” from the ambient noise. This is noteworthy, as the spectrogram contains most of the information the auditory cortex receives. Moreover, the auditory system only receives this information in a successive manner, and thus must link the sound events over time to form a stream. How this may be achieved is the topic of the next section.
Neural models of auditory stream formation

Several hypotheses for the neural underpinning of stream formation have been proposed\textsuperscript{[67–71]}. Firstly, the ‘population separation’ hypothesis suggests that sound elements may be attributed to separate streams if they activate distinct neural populations. Since tonotopy is the most prevalent organisational feature in the auditory system, most paradigms testing for this hypothesis have focused on the frequency dimension (and have been later extended to other features, such as spatial location)\textsuperscript{[69]}. Several perceptual effects can be explained by the ‘population separation’ hypothesis, of which the famous ‘ABA’ streaming effect\textsuperscript{[70]} (presented in figure 1.14a,b): two alternating sequences of pure tone of different frequencies can either be perceived as one stream if the frequency difference is small, or two streams if the frequency difference is large. Presumably, this effect stems from the fact that the two streams are represented by non-overlapping neuronal populations in the case of a large frequency difference.

However, the ‘population segregation’ hypothesis cannot account for the dependence of stream formation on sound timing\textsuperscript{[69]}. For example, when the two sequences of pure tones previously used in the ‘ABA’ paradigm are presented concurrently rather than in alternating fashion (such as presented in figure 1.14c), they are perceived as one stream even if their frequency difference is large. This contradiction led to the formation of another model for stream formation: the ‘temporal coherence’ framework\textsuperscript{[69]}. This framework comprises two main stages. Firstly, the incoming mixture is decomposed according to several features (such as frequency content or spatial location), and represented by specialised neurons with corresponding tuning profiles. This first stage is labelled ‘feature-analysis’, and is reminiscent of a ‘population separation’ process. Secondly, the amount of coherence between the temporal outputs of the different feature-tuned neurons determines the grouping.

Several psychophysical and neurophysiological effects can be explained using this framework\textsuperscript{[69]}, such as the effect of comodulation-masking release\textsuperscript{[72]}, or the fact that temporal coherence across different sensory modalities can enhance the segregation of auditory inputs\textsuperscript{[73]}. Moreover, the ‘coherence-analysis’ stage would require computations to be carried over relatively long time windows\textsuperscript{[69]}. Auditory cortex (AC) neurons possess ideal characteristics for the implementation of this stage: their responses contain dynamics consistent with this time range\textsuperscript{[69]}, and they present sensitivity to stimulus history over multiple time scales\textsuperscript{[74]}. However, whilst responses in the cortex have been extensively studied\textsuperscript{[71,75]}, the specific neural implementation of this second stage remains unclear.

Regularities in sound sequences can also markedly affect auditory scene analysis\textsuperscript{[68,76,77]}. Notably, pattern regularity can stabilise stream perception\textsuperscript{[76]}, and enhance stream segregation\textsuperscript{[77]}. Moreover, regularity in the timing of sounds can also influence stream formation\textsuperscript{[76,78]}. The identification of environmental regularities is believed essential for goal-directed behaviour\textsuperscript{[79]}, as it enables the generation of reliable predictions, and the optimisation of sensory processing. This process might be particularly important in the detection of ‘novel’ events, which may require rapid behavioural adjustment\textsuperscript{[68,79–81]}. Notably, the auditory cortex is sensitive to several forms of regularities\textsuperscript{[68,82,83]}. For example, cortical cells adapt to sounds presented in a regular sequence, and respond more strongly to rarely presented sounds embedded within such sequence (a phenomenon known as stimulus-specific adaptation; SSA)\textsuperscript{[68,81–83]}. One prominent theory\textsuperscript{[68,83]} suggests that predictive
representation of temporal regularities constitutes the core of auditory object formation. Whether predictive representation is an emergent feature of AC is debated\cite{ref1, ref83}, as SSA has been observed in sub-cortical regions\cite{ref84–ref86}.

**Relationship between attention and streaming**

Neural correlates of stream segregation have been observed under anaesthetised preparations in AC\cite{ref71–ref74, ref87–ref89}, suggesting that a certain amount of sound processing occurs unconsciously. Whether attention is necessary for the formation of a stream is an ongoing debate\cite{ref68, ref69}. However, it is clear that attention can influence auditory stream formation\cite{ref69}. For example, in the ‘ABA’ paradigm, the frequency difference necessary to elicit the perception of two streams is reduced if the listener actively aims at hearing out the high-pitch tones\cite{ref69}. Whether attention acts by sharpening neural tuning (i.e. enhancing the ‘population analysis’ stage), by influencing the temporal coherence of sound representation, and/or by affecting predictive coding is the topic of the next section.

1.3.2 **Attention modulates perception and sensory representation**

Attention is generally inferred from the level of engagement in a task, which can be assessed via the measurement of behavioural performance\cite{ref69, ref90, ref91}. For example, subjects may be required to listen to a story without reacting to it (passive listening task) or to detect a specific word within sentences (active listening task) – in the latter case, it is assumed that subjects ‘pay attention’ to the stimuli more intensively than in the first scenario. In other tasks, subjects may be required to detect specific words from the story teller whilst the story is played amongst background noise (active listening task with competitive stimuli), or to switch from listening to one person to another (active attention-switching task). Whilst all types of tasks are important to understand the effect and anatomical basis of attention, it is important to highlight that the effect of attention on perception and sensory processing may largely depend on the specific design of the task a subject has to perform\cite{ref90–ref92}. Nonetheless, attention is generally viewed as having a beneficial effect, as it enhances behavioural sensitivity and shortens reaction time\cite{ref61, ref63, ref69, ref90, ref91, ref93–ref95}.

Several tasks have now been developed in humans and animals models to investigate the effect of task-engagement and task-switching on sensory processing and perception. A few of them are detailed in the following sections, as they provide critically different insights.

**Attention changes neural tuning profile**

In pioneering studies, Hubel et al.\cite{ref96} reported that AC neurons responded differently to sounds depending on the apparent attentiveness of the subject. This contradicted the intuitive idea that neural responses to sensory stimuli should remain stable, and mainly represent the incoming stimulus features\cite{ref97}. Inspired by this finding, several studies in humans and animals have investigated the effect of attention on sound representation and perception\cite{ref63, ref71, ref90, ref93, ref94}.

Notably, a set of animal studies conducted by Fritz et al.\cite{ ref65, ref98–ref100} demonstrated that sound representation can be rapidly shaped by task demand. Ferrets were trained to detect a tonal stimulus (named target) in ripple-noise stimuli. The latter stimuli enabled the computation of neurons’ spectro-temporal receptive fields (STRF). When animals became engaged in the task, the STRF of AC neurons changed, so as to specifically enhance the representation of the target frequency content,
as displayed in figure 1.15a. When the target frequency content was changed, the STRF adapted correspondingly. Similar effects were found using another task design, where animals attended either to a tonal mixture, or to a specific tone sequence within that mixture. As seen in figure 1.15b, neuron tuning profiles adjusted according to task demand.

By sharpening and modifying neuronal tuning, attention could hypothetically affect the ‘population segregation’ stage during auditory streaming.

**Attention affects cortical state and local field potential phase coherence**

Neurons in sensory cortices are organised in complex networks. Because of this anatomical configuration, a group of simultaneously recorded neurons often presents correlation in spiking activity. Correlated variability has a strong influence on how a population of neurons can encode sensory information\[^{101–104}\]. The state of the network influences these correlations, classic examples of which being the ‘UP’ and ‘DOWN’ states\[^{101,103,105}\]. Attention has been proposed to involve processes that are similar to state changes\[^{102,106}\]. Therefore, during attention tasks, it may be critical to measure signals informative on cortical states. Notably, the local field potential (LFP) signal has been proposed to reflect network state\[^{105,107}\].

In a series of experiments in primates, Lakatos et al.\[^{95,106,108,109}\] analysed the variations of LFP signals during attention tasks. Animals were presented with two simultaneous but out-of-phase rhythmical stimulus sequences, and had to detect a target presented in the cued sequence only. The task was either cross-modal (using a visual and an auditory sequence) or uni-modal (using two auditory sequences of different frequencies). The sequences had a stimulus presentation period of \(\sim 500-800\) ms, i.e. a rhythm falling into the delta range (1-4 Hz) of neural oscillations.

Rhythmical stimuli can rapidly entrain ongoing cortical oscillations and shape spiking responses\[^{95}\], as neural excitability in AC correlates with the phase of LFP signals\[^{105,110,111}\]. Lakatos et al.\[^{95,108,109}\] found that whilst LFP entrainment to external stimuli may not require attention, its phase can be strongly modulated according to task demands. Specifically, in the cross-modal attention task, the phase of delta-range LFP oscillations in V1 synchronised (i.e. was coherent) with the rhythm of the attended stream only. This phenomenon was independent of the cued modality.

This finding was recapitulated in the uni-modal (auditory only) task\[^{109}\]. The phase of delta-range LFP oscillations in A1 were synchronised to a stimulus sequence only if it was attended. Moreover, the phase of the oscillation depended on the frequency preference of the site: in regions best tuned to process the attended frequency, the high-excitability phase of entrained oscillations was aligned with the occurrence of stimuli. Regions tuned to process different frequencies (specifically two octaves above or below the attended frequency) were also entrained, but in counter-phase (i.e. a low-excitability phase was aligned to stimuli occurrence). This difference in phase resulted in the amplification and sharpening of responses at attended time point, and was thus proposed to act as a spectro-temporal filter\[^{109}\].

Furthermore, Lakatos et al.\[^{106}\] revealed that lapses of LFP entrainment reflected lapses of attention. Specifically, periods with attentional lapses were characterised by strong alpha oscillations affecting neural responses, and were anti-correlated (i.e. alternated) with the periods of entrainment. The neural mechanisms underlying these two distinct brain states are currently unknown, but have been proposed to rely on interactions between cortical and sub-cortical structures\[^{106}\].
As the phase of the LFP has been shown to impact on sensory processing and perception in several species (including humans\cite{95,112}, monkeys\cite{106,108,110}, ferrets\cite{73} and rats\cite{105,111}), the modulation of low-frequency LFP oscillations appears as a widely-conserved substrate for attentional orienting throughout animal evolution\cite{95}.

**Directing attention in time modulates neural activity dynamically**

Auditory scenes are intrinsically dynamic, as they comprise fast-changing and relatively brief acoustic events\cite{69}. Similarly, attention is not a static process, but rather displays fluctuations over time, with moments of heightened attention and moments of lapses\cite{91,106}. Therefore, the ability to effectively direct attentional resources in time is critical to capture relevant sensory events\cite{91,113}. In order to know when to direct these resources, the brain may generate predictions on the timing of sensory events\cite{79,90,91,93,113}. These predictions can be formed via the identification of regularities present either in the stimuli or in the context in which they are presented.

Valid predictions (or expectations) have been shown to enhance sensory perception and reduce the time to react to sensory events\cite{61,90,91,93,94,113}. These effects have been observed in several species (such as in humans, monkeys and rats\cite{61,90,91,93,94,113}), indicating that most animals have evolved mechanisms to anticipate future events which impact on motor control and sensory processing. Whilst the neural basis of accelerated reaction time has been extensively studied\cite{114}, little is known about the mechanisms underlying enhanced sensory perception\cite{61,91,113}.

In a study conducted by Jaramillo and Zador\cite{61}, valid temporal expectation was shown to modulate auditory sensory processing as early as the primary auditory cortex (A1). Rats were trained to detect and identify a modulated sound (named target) in a sequence of tones of random frequencies. The timing of target presentation (early or late) was conserved across several consecutive trials, enabling a prediction to be formed. Valid expectation decreased reaction time and increased perceptual reports, consistent with previous studies\cite{61,90,91,93,94,113}. Activity in A1 was tightly linked to behaviour: increased sound-evoked activity led to faster reaction time, and silencing of cortical activity impaired behaviour. Moreover, expectation modulated neural responses dynamically: over the course of a single trial, neurons displayed a ramping-like increase in evoked response, peaking specifically near the timing of expected target occurrence (see figure 1.16). The neural mechanisms underlying this increase are currently unknown.

**During attention tasks, activity in auditory cortex reflects stimulus features – but also task rule, stimulus meaning, and behavioural decision**

A few recent studies\cite{115–117} began to paint the AC as more than a mere stimulus feature encoder, but rather as a fundamental processing stage in cognitive processes – a proposal already advanced in the visual system\cite{118,119}. These findings question the role of sensory cortices in mediating attentional effects, as presented below.

Cognitive aspects of attention are mostly believed to be encoded in higher-order brain structures than primary sensory areas, such as for example in the prefrontal cortex (PFC)\cite{120–126}. Notably, PFC neurons encode task rule\cite{115,120,121,126}, and disruption of their activity impairs attentional control\cite{123,124,126}. The role of PFC in supporting attentional and other cognitive processes is not disputed here; however, new evidences point towards AC as encoding for similar features.
Rodgers et al.\textsuperscript{[115]} found that A1 neurons encoded selection rule during an attention-switching task. Rats were trained in two tasks: first, to discriminate between the location of a noise burst, and second to discriminate the pitch of a FM sound. When rats were presented with both noise burst and FM sound simultaneously, they had to respond based on the property of only one of the two stimuli — in other words, animals had to follow a task selection rule. The cue for the rule was given on a trial-block basis, i.e. the rule was conserved for a set amount of trials, and then switched. Rodgers et al.\textsuperscript{[115]} found that prior to receiving the noise and FM stimuli, neurons in both A1 and PFC encoded for task selection rule. It has to be highlighted that the proportion of rule-encoding neurons was nearly double in PFC (63 %) compared to A1 (36 %). Moreover, the time course for rule-encoding was different in PFC and A1 neurons, with PFC neurons exhibiting selective activity earlier than A1 neurons. This was consistent with the hypothesised role of PFC neurons as the source of top-down modulation\textsuperscript{[115]}. Nonetheless, the responses of A1 neurons were intriguing in several ways. Firstly, it was surprising to observe a modulation in A1 activity preceding stimuli onset, as no stimuli were presented during this period to actively drive cortical responses (i.e. the activity was spontaneous). Moreover, this change in activity reflected a ‘high-order’ cognitive rule. Furthermore, spontaneous activity in A1 increased in a dynamic manner, as seen in evoked responses during a temporal expectancy task\textsuperscript{[61]}. However, it was not tested whether A1 spontaneous activity causally guided behaviour. The anatomical basis and importance of this modulation remain unknown.

Similar types of findings were uncovered by Bagur et al.\textsuperscript{[116]}. A1 and dorsolateral frontal cortex (dlFC) neurons were recorded in ferrets engaged (active) or not (passive) in Go/No-Go tasks. Two stimuli were used in all Go/No-Go tasks, one of them being a target to respond to (either by licking or by stopping licking). The target valence (i.e. aversive or appetitive) varied across tasks and animals. When the animals engaged in tasks, A1 neural population code switched from a sensory-driven representation of the Go/No-Go stimuli to a behaviourally-relevant representation. Specifically, the representation of the target stimulus was enhanced during behaviour. Interestingly, this asymmetry in stimulus encoding partly relied on changes in spontaneous activity\textsuperscript{[116]}. Target-driven stimulus encoding was also found in single dlFC cells, mirroring the encoding by A1 population. Based on dynamics of A1 population and single dlFC cells, the authors\textsuperscript{[116]} suggest that A1 and dlFC might form a recurrent loop, in which A1 neurons feed onto single dlFC cells, which integrate the task-related information present at the A1 population level, and feed it back onto A1 neurons for further processing. The anatomical basis of such a circuit is yet however hypothetical.

Behavioural decision is another cognitive aspect generally thought to emerge from processing in high-order areas such as frontal cortices\textsuperscript{[127–129]}. However, studies by Rodgers et al.\textsuperscript{[115]}, Bagur et al.\textsuperscript{[116]} and Francis et al.\textsuperscript{[117]} revealed that A1 activity also encodes for behavioural decisions made during task engagement. Behavioural choice could be decoded from activity prior to stimuli\textsuperscript{[115, 116]}, and did not depend on the animals’ movement\textsuperscript{[116]}. This latter verification is important, as it is known that engaged animals can adopt stereotypical movements indicative of their subsequent behaviour\textsuperscript{[130]}, and that A1 activity can be rapidly modulated by motor-related areas\textsuperscript{[131–133]}. Francis et al.\textsuperscript{[117]} further showed that the functional connectivity of A1 neural network changed between passive and active engagement, and that behavioural decision was encoded by small networks of neurons. This result is consistent with the findings of Bagur et al.\textsuperscript{[116]}, who suggest that decision is encoded in A1 via the activity of neural population rather than by that of single neurons.
To develop a system enabling the simultaneous recording of activity in PFC and A1 during ongoing behaviour is one of the main aims of the research presented in chapter 2 Materials and Methods, as it would enable the comparison of neural dynamics between these two regions.

1.3.3 Anatomical circuits supporting sensory representation modulation and attention

From the above literature, it appears that attention can modulate cortical processing in various ways. One plausible conjecture is that different effects may be mediated by distinct sets of anatomical regions and connections. The following sections aim at providing an overview of different brain regions and projections known to be involved in the modulation of sensory processing. The involvement of these regions in attentional processes is also discussed. The diagram presented in figure 1.17 draws connections between the main structures reviewed below.

Sub-cortical structures: Mediators of attention?

Whilst there is considerable evidence for the modulation of auditory cortical processing by attention, a large body of literature suggests that sub-cortical nuclei may play a pivotal role in mediating attentional processes\[19,123,134–136\]. One prominent theory\[135,137\] argues that attention may modulate sensory signals before they even reach cortex, specifically by altering responses in the sensory thalamus. Consistently, cells in the inferior colliculus (IC) and medial geniculate body (MGB) can undergo rapid receptive field re-tuning\[19\] as seen in cortex\[99\]. Moreover, non-lemniscal MGB neurons exhibit climbing discharge prior to reward delivery during appetitive tasks, potentially reflecting anticipatory and attentional processes\[19\]. It is thus unclear whether properties observed in cortex are emergent or inherited from sub-cortical stations\[1\].

Crick\[137\] stated that “if the thalamus is the gateway to the cortex, the reticular complex might be described as the guardian of the gateway”. The thalamic reticular nucleus (TRN) is a thin sheet of GABAergic cells surrounding the thalamus\[137–140\]. TRN cells receive both cortical and thalamic inputs, but projects only onto thalamic cells\[138–140\]. Because of this unique connectivity, TRN cells can strongly modulate the sensory information passing through the thalamus en-route to cortex\[123,135–137,139,140\]. Consistent with its theoretical role of ‘gatekeeper’, the TRN exhibits differential activity depending on attentional state, in a manner consistent with specific enhancement and dampening of thalamic activity supporting behavioural needs\[123,135\]. Moreover, disruption of the TRN leads to markable attention impairments\[123,134,136\].

Whilst it is evident that the TRN can strongly affect sensory processing and behaviour, the underlying mechanisms are unclear. Recent research\[123,136,140–142\] endeavoured to uncover how the TRN is modulated by attention, and how it can affect sensory processing. A selection of critical findings is presented below.

Halassa et al.\[123,126\] developed a paradigm to study cross-modal attention in freely-moving mice. Visual and auditory cues were simultaneously presented to the animals, who had to respond based on either the visual or auditory stimulus feature. The cue for modality-selection (i.e. rule-selection) was a brief sound presented prior to the two competitive stimuli, changing on a trial-basis. The period in between the cue reception and stimuli presentation is called here ‘pre-stimuli period’. Using optogenetic techniques, Halassa et al.\[123,126\] demonstrated that whilst visual cortex was necessary
for the perception of visual stimuli, the visual thalamus was responsible for mediating the attentional effect of stimulus selection. Upon the reception of the competitive stimuli, visual thalamic cells were enhanced during the ‘select-visual’ trials, and suppressed during the ‘select-auditory’ trials. This effect was mediated by TRN cells projecting onto the visual thalamus (TRNvis): during the pre-stimulus period, TRNvis cells displayed increased activity when the auditory modality had to be selected, and vice-versa for visual selection, acting therefore as gain modulator. Optogenetic silencing of PFC during that period drastically diminished selection rule encoding by TRNvis cells, as well as behavioural performance. Such perturbation did not impair visual perception, suggesting that the PFC→TRN→thalamus circuit mediates the ‘selective’ aspect of the task (the precise anatomical connections between PFC and TRN are yet unknown in mice, hence the dashed arrow).

Whilst the studies of Halassa et al. provide remarkable insight on the anatomical basis of cross-modal attention, it remains to be seen whether similar mechanisms are used during competitive, auditory-only, stimulus selection tasks. Moreover, it is not known whether the PFC acts directly onto the TRN, or via another anatomical station that projects onto the TRN.

Ahrens et al. used both cross-modal (visual/auditory) and uni-modal (auditory-only or visual-only) attention paradigms in freely-moving mice to study the role of TRN in mediating attention. The TRN was perturbed by the selective deletion (knock-out, KO) of the ErbB4 receptor onto SOM+ cells. KO mice performed better than controls during uni-modal tasks, but worse during the cross-modal task. Deletion of ErbB4 specifically enhanced cortical drive onto TRN cells. By dampening cortical drive onto TRN cells, the authors were able to reverse the abnormal behaviour in KO mice. Ahrens et al. showed that the anatomical spread of ErbB4 deletion related to the animal’s performance, suggesting a distinct role for different areas of the TRN in mediating attention. Accordingly, cells in specific regions of the TRN present organised and well-defined anatomical projection patterns. Enhancing cortical drive onto different TRN populations may result in distinct behavioural effects, as observed between cross-modal and uni-modal tasks (see supplementary figure 10 of Ahrens et al. for a comprehensive diagram).

In the auditory system, the TRN is part of a large thalamo-cortical loop (presented in figure 1.17). The AC sends dense corticofugal projections to several sub-cortical nuclei; notably to the IC, MGB and TRN. Interestingly, corticofugal projections originating from Layer 6 innervate both lemniscal MGB and TRN neurons. This pattern of innervation has been recently proposed as a substrate for spectral tuning sharpening via “center-surround” suppression. However, only few studies endeavoured to uncover the impact of TRN and corticofugal projections on MGB and AC processing. The role of thalamo-cortical recurrent circuits in shaping the representation of sensory inputs during attention is yet unknown. To investigate the role of the TRN in shaping sound representation is one of the main aims of the research presented in chapter 4 Anatomical circuits underlying task-related modulation of spectrotemporal receptive fields.

Frontal cortices: Top-down origin of sensory representation modulation?

In mice, the prefrontal cortex (PFC) can be divided into several subregions: the medial prefrontal cortex (mPFC; comprising limbic and pre-limbinc zones), the orbito-frontal cortex (OFC) and the anterior cingulate cortex (ACg).

Halassa et al. showed that solely mPFC was critical for the maintenance of rule-selection in
a cross-modal attention task. Disrupting the activity of mPFC neurons modulated TRN cells targeting sensory thalamic nuclei\[123\]. However, the precise anatomical connection between PFC and TRN is yet unknown.

The OFC has recently been shown to modulate auditory cortex tuning\[145,146\]. OFC neurons project onto the supragranular and infragranular layers of A1, and provide excitatory inputs to both A1 interneurons and pyramidal cells\[146\]. The OFC has been implicated in numerous aspects of goal-directed behaviour, however its precise role is currently debated\[147\].

In the visual modality, top-down modulation of sensory inputs processing can be mediated by the projection of ACg onto primary visual cortex (V1) and superior-coliculus (SC) cells\[148\]. Activation of ACg neurons resulted in a sharpening of V1 cell tuning, by acting on local interneuron microcircuits\[148\]. Specifically, SOM+ interneurons contributed to center-surround suppression, whereas VIP+ interneurons were involved in center facilitation. Moreover, activation of ACg cells improved behavioural performance in a visual discrimination task\[148\]. Whilst interneurons can drastically affect the processing of sound in auditory cortex as well as behavioural outcome during auditory-based tasks\[36,149,150\], it is not known whether the ACg projects to auditory processing stages in a similar manner as in the visual modality.

**Neuromodulators**

Neuromodulators have been strongly implicated in supporting attentional processes and sensory representation\[99,151–156\]. Neuromodulators are chemicals that diffuse through neural tissue, with the potential to affect neurons at large spatial scale. They usually bind to receptors with slow kinetics, and can have prolonged effect on cellular metabolism\[157,158\]. In the central nervous system, major neuromodulators include acetylcholine (ACh), dopamine, serotonin, histamine and noradrenaline. Distinct brain regions produce these neuromodulators; for example, the nucleus basalis (NB) of Meynert is responsible for acetylcholine production, whilst the tuberomammillary nucleus is the sole source of histamine in the mammalian brain\[157,159\].

The mechanisms underlying sensory representation re-shaping or attention modulation by neuromodulators are not well understood, probably owing to the fact that neuromodulatory kinetics are extremely intricate. Notably, there are different types of receptor a single neuromodulator can bind to (e.g. muscarinic and nicotinic receptor for ACh, and five distinct receptors for dopamine\[157,158\]). Moreover, these receptors are not expressed uniformly across neurons\[157,156\]. Similarly, the structure of a single neuromodulatory nucleus is not homogenous; for example, the NB comprises several types of neurons – some may be cholinergic, but some others are GABAergic\[157,160\]. Furthermore, the activity of one neuromodulatory structure may cause the release of another type of neuromodulator\[161\]. Finally, neuromodulators have diffuse anatomical targets in the brain. For example, cells in the NB can project to the AC, but also to the PFC\[157\].

To address the role of neuromodulators in attention and sensory representation is beyond the scope of this thesis. Nonetheless, it is important to mention their existence as they are clearly involved in these processes. For example, pairing NB stimulation with a tone results in receptive field re-tuning in AC\[151\], and the activation of NB cholinergic terminals in sensory cortex results in decreased spike train correlation amongst neurons\[152\], as well as in an enhancement in sensory discrimination perception – all of these effects are reminiscent of those seen during attention tasks. Correct balance
of ACh in PFC is also critical in maintaining high performance during attention task\textsuperscript{[154]}.  

1.3.4 Techniques to decipher the contribution of specific anatomical circuit components  

Long-range projecting neurons are a critical component of large-scale networks (such as the one presented in figure 1.17), as they provide distant brain regions with the outcome of local circuit computation\textsuperscript{[162]}. To delineate their specific contribution requires selective manipulation techniques. Some projecting neurons have identified gene expression that differs from surrounding cells, allowing the generation of specific driver lines. For example, Ntsr1-positive neurons are restricted to Layer 6 in the mouse sensory cortex, enabling the generation of the transgenic mouse line Ntsr1–Cre\textsubscript{}\textsuperscript{[163,164]}. However, specific genetic access is not always available\textsuperscript{[162,165]}.  

Two different approaches enable to circumvent the lack of specific driver lines (presented in figure 1.18). The first one relies on activating projecting neurons by the stimulation of their axon terminals via optogenetic means (see figure 1.18b for a schematic). For example, to delineate the impact of OFC onto AC processing \textit{in vivo}, one can induce the expression of ChR in OFC, and stimulate axon terminals reaching AC specifically by shining light onto AC only\textsuperscript{[146]}. Such a technique however requires to deliver light onto the region of interest (e.g. AC in the previous example), either via guiding-cannula which hinders access to that latter region, or through the use of ‘optoelectrodes’ which are yet rather expensive (e.g. a 4x2 NeuroNexus tetrode coupled with single optical fibre costs over 1000 GBP; quote obtained in 2017).  

The second approach relies on retrograde viral techniques to induce the expression of engineered ion channels into projecting neurons selectively\textsuperscript{[41,162,165]}. When a retrograde virus is injected in a region of interest, any neuron (with appropriate tissue tropism) projecting onto that region is infected\textsuperscript{[41,162,165]}. To specifically manipulate projecting neurons emerging from a single region, two approaches are possible. The first one relies on retrograde virus encoding for light-gated ions channels, and on stimulating with light a specific origin region (see figure 1.18c for a schematic). The second approach consists on injecting a retrograde virus encoding for Cre-recombinase in the region of interest, and on injecting a second virus encoding either for floxed-DREADD (as presented in figure 1.18d) or floxed-opsin in the origin region. This latter approach is attractive, as it permits the use of DREADD, which requires minimally-invasive activation methods.  

In chapter 4 \textit{Anatomical circuits underlying task-related modulation of spectrotemporal receptive fields}, I take advantage of newly developed viral techniques\textsuperscript{[165]} to investigate how specific projecting neurons from the TRN to the MGB may modulate sound representation.
1.3.5 Summary of research aims

The main aims of the research presented in chapter 2 Materials and Methods and chapter 4 Anatomical circuits underlying task-related modulation of spectrotemporal receptive fields are the following:

1. To setup a system enabling the recording of A1, MGB and PFC activity in mice during ongoing behaviour.

2. To design and validate a behavioural paradigm for mice involving different behavioural responses (i.e. task-switching) to otherwise identical auditory cues.

3. To test for the functional role of the PFC and TRN in modulating sound encoding in MGB and A1.
Figure 1.1: Example spectrograms of natural sounds. a. Right: Spectrogram of the word “Waldo” pronounced by a human, adapted from Pasley et al.\cite{166}. Left: Spectrogram of a cat call, adapted from Hubka et al.\cite{4}. Red–blue colors indicate high–low energy. b. Spectrograms of the spoken word “Shoe” recorded in isolation (left) and in ambient noise (right). Images adapted from Bregman\cite{67}.

Figure 1.2: Mouse vocalisation contains slow frequency modulations. a. Example spectrograms of characteristic ultrasonic syllable types (bottom). Blue box indicates syllable type presented in b. Top: Associated sound waveform. Data shared with courtesy by J. Grimsley\cite{167}. b. Example spectrogram of UP frequency-modulated (FM) sweep from mouse vocalisation (grey), with overlayed synthesised FM sweeps of different frequency modulation speeds (blue). Image adapted from Musolf et al.\cite{3}. c. Mouse vocalisation contains mainly low FM speeds (< 25 oct/s). Distributions of FM speed measured at the start, middle and end portions of ultrasonic syllables. Data shared with courtesy by J. Grimsley\cite{167}.
Figure 1.3: Example selective neural response to FM sweep direction. Neural responses (middle panel: raster of spiking activity, top panel: corresponding peri-stimulus time histogram) showing preference for upward (UP) versus downward (DOWN) frequency modulated (FM) sweeps in the rhesus monkey auditory cortex. The FM sweep is displayed in the bottom panel. Image adapted from Kanwal and Rauschecker.\[168\]

Figure 1.4: Approximation of FM by pure tones. FM sweep (right) can be approximated by successive pure tones (middle), and further by presenting two tones only (right).
Figure 1.5: Spectral and temporal history affect spiking response in auditory cortex. 

a. The spectral history of sound affects spiking response in auditory cortex, as shown by the use of ‘Two tones’ paradigm. Right: Schematic of the auditory stimuli. The masker varies in frequency, whilst the probe does not. Essentially, presenting a lower masker than a probe can be viewed as an approximation for an upward FM, and vice-versa for higher masker and downward FM. Left: Example of single neuron spiking activity recorded in mouse AC in response to the two tones. The red arrow indicates the probe frequency. Images adapted from Phillips et al.\cite{12}.

b. The temporal history of sound affects spiking response in auditory cortex, as shown by the use of ‘Two clicks’ paradigm (clicks are very brief sounds). Example of single neuron spiking activity in rat AC in response to click pairs of different intervals (top: raster plots, bottom: firing rates, stimuli indicated by ticks on abscissa, colors indicate different intervals). The response to the second click was completely suppressed for short intervals, and progressively recovered for longer intervals. Image adapted from Wehr and Zador\cite{13}.

\cite{12}
Figure 1.6: Ascending auditory pathway and tonotopy. a. Right: Schematic of the cochlea (copyright: Pearson Education, 2011). When a sound is received, the staples transmit the vibrations to the fluid inside the cochlea (symbolised as red arrow), which provokes a local displacement of the basilar membrane. If the sound frequency is low (red), the displacement is located at the apex, whilst if it is high (green) the displacement is located near the oval window, at the base of the basilar membrane. This gives rise to tonotopy, an organisation of response in space according to frequency. Left: Such organisation is conserved throughout the different stages of the ascending auditory pathway in humans. These stages are: the cochlear nucleus (CN), the superior olive (SO), the inferior colliculus (IC), the medial geniculate body (MGB) and the auditory cortex (AC). b. The architecture of the ascending auditory pathway of the mouse is similar to humans. Labels same as in a. Image adapted from Mueller[169]. c. The auditory cortex (AC) can be further divided in multiple subfields according to the spatial organisation of frequency tuning. These fields are: the primary auditory cortex (A1), the anterior auditory field (AAF), the dorsomedial field (DM), and the secondary auditory field (AII). D/A arrows indicate dorsal/anterior orientations. Image adapted from Tsukano et al.[18].
Figure 1.7: Correlation between frequency tuning and FM direction selectivity in primary auditory cortex. **Left panel:** Example frequency tuning map recorded in the rat auditory cortex. Frequency preference is spatially arranged in a tonotopic fashion, with posterior/anterior sites responding preferentially to low/high frequencies. The A/D arrows indicate anterior/dorsal axis. Scale bar: 0.5 mm. **Right panel:** Example FM direction selectivity in the same animal. Direction selectivity is spatially arranged, with posterior/anterior sites responding preferentially to UP/DOWN frequencies. Positive/negative direction selectivity index (DSI) indicates preference for UP/DOWN FM sweeps respectively. Images adapted from Zhang et al.\textsuperscript{9}. 

---

**Figure 1.7:** Correlation between frequency tuning and FM direction selectivity in primary auditory cortex. **Left panel:** Example frequency tuning map recorded in the rat auditory cortex. Frequency preference is spatially arranged in a tonotopic fashion, with posterior/anterior sites responding preferentially to low/high frequencies. The A/D arrows indicate anterior/dorsal axis. Scale bar: 0.5 mm. **Right panel:** Example FM direction selectivity in the same animal. Direction selectivity is spatially arranged, with posterior/anterior sites responding preferentially to UP/DOWN frequencies. Positive/negative direction selectivity index (DSI) indicates preference for UP/DOWN FM sweeps respectively. Images adapted from Zhang et al.\textsuperscript{9}. 

---
Figure 1.8: Columnar organisation in primary auditory cortex.  

a. The primary auditory cortex (A1) is located in the temporal lobe. This image displays a coronal slice from a mouse brain, adapted from the Allen brain atlas. The outline represents the layers shown in b.  
b. A1 is organised in six layers (labelled 1 to 6). Each layer of cortex contains different cells, with distinct morphologies. This image displays two principal cells (PC) originating from different layers (L2/3 and L5) in mouse A1. Image adapted from Sakata and Harris [170].  
c. Anatomical connections and information flow within a column of primary auditory cortex. Left face of cube shows thalamic (MGB) and cortico-cortical (Ctx; ipsi: ipsilateral, contra: contralateral) inputs; right face displays interlaminar connections as well as thalamic, collicular (IC), and cortico-cortical outputs. Shading indicates layers that receive thalamocortical input (left) or produce cortico-thalamic outputs (right). Layers are indicated with roman numbers. MGB1 and MGB2-3 represent lemniscal and non-lemniscal thalamic inputs respectively. Within the column, pyramidal cells in L4 and lower L3 receive the major lemniscal thalamic input, initiating a flow of information into the supragranular layers (L1-3) and then down to the infragranular layers (L5-6). Pyramidal cells in L2 and L3 also extend long-range lateral projections to form horizontal connections with other cortical columns (symbolized by neuronal projection on top of cube). Image and caption adapted from Linden and Schreiner [24].
Figure 1.9: Classes of inhibitory interneurons. a. Current understanding of the synaptic targets of four classes of cortical interneurons. The classes are: neurogliaform (NG), vasoactive intestinal peptide (VIP)-expressing, somatostatin (SOM)-expressing, parvalbumin (PV)-expressing cells. The green cell represents a principal cell. Image adapted from Harris and Mrsic-Flogel[25] and from Naka and Adesnik[171]. b. PV-basket and PV-chandelier cells have distinct morphologies. Image adapted from Raghanti et al.[172].

Figure 1.10: Non-exhaustive schematic of a cortical microcircuit comprising inhibitory and excitatory cells. Inhibitory and excitatory cells form a vast network, with feedback and feedforward connections or loops. Pyramidal (Pyr.) cells and excitatory connections are represented by a triangle, whilst interneurons and inhibitory connections are represented by circles. Only a subset of all known anatomical connections is represented for simplification. Image adapted from lavorska and Wehr[173] and from Naka and Adesnik[171].
Figure 1.11: Identified mechanism for FM direction selectivity in A1 based on inhibitory-excitatory interaction. 

a. Excitatory currents evoked in a low frequency tuned neuron to pure tones of various frequencies and intensities. The receptive field is asymmetric, as outlined by the green dashed line: currents are strongest following low frequency tones presentation (black star), and progressively diminish following higher frequency tones presentation (grey star). Scale bars: 200 pA (vertical), 80 ms (horizontal). The voltage was clamped to a value close to the expected cell resting potential (-70 mV).

b. Model for amplification of frequency response in a Two Tone paradigm. Top: A neuron tuned to low frequency has a higher response to low (black) than high (grey) frequency tone, with inhibitory currents (red) following excitatory currents (green). Middle and bottom: When a high frequency tone is successively presented after a low tone (UP, orange), the neuron can reach spike threshold (dashed line), as the strong initial excitatory response is not immediately dampened by the inhibitory input. Vice-versa, when a low frequency tone is successively presented after a high tone (DOWN, brown), the inhibition elicited by the high tone remains and prevents the strong excitatory response to the low tone to reach threshold. Images adapted from Zhang et al. [9].
Figure 1.12: Two distinct mechanisms can underlie FM direction selectivity.  

a. A neuron is part of a neural network, receiving inputs from other neurons which are integrated to generate an output. 

b. Excitatory-Excitatory model for FM direction selectivity. Two excitatory inputs (Exc) to a neuron can be tuned to different frequencies ($f_1$, $f_2$), with different excitatory peak delay. The difference in the delay between these inputs can account for preferential response to a specific frequency presentation order. The preferred order in this schematic is $f_1 \rightarrow f_2$. In this model, the integration of both excitatory inputs facilitates the output neuron response. 

c. Excitatory-Inhibitory model for FM direction selectivity. Excitatory (Exc) and inhibitory (Inh) inputs to a neuron can be tuned to different frequencies ($f_1$, $f_2$), with different excitatory/inhibitory peak delay. The difference in the delay between these inputs can account for preferential response to a specific frequency presentation order. Preferred order same as in a. In this model, the non-preferred order generates a depression in the output response of the neuron. Images adapted from Ye et al.\cite{Ye2010}. 

\cite{Ye2010}
Figure 1.13: Neurons in auditory cortex exhibit distinctly tuned excitatory responses to the onset and offset of sound. a. Example raster and PSTH of a neuron recorded in monkey primary auditory cortex (A1) in response to pure tones of different frequencies (black square; left: 4.0 kHz, right: 12.5 kHz). The neuron exhibits excitatory response to the onset (ON; green) and offset (OFF; red) of the tone. Notably, the magnitude of these responses depends on the pure tone frequency. b. Onset and offset responses are spectrally segregated. Example tuning of onset and offset response magnitudes of a monkey A1 neuron following the presentation of bursts of noise (bandwith: 1/6 octave) centered around several frequencies (x-axis). The tuning of the onset response (green) does not overlap with the tuning of the offset response (red). Images in a-b adapted from Tian et al. \[52\]. c. In the primary visual cortex (V1), ON and OFF responses are also distinctly tuned. Moreover, this difference is invariant within cortical column. Lines connect the ON subfield (green) and the OFF subfield (red) centres of individual V1 cell receptive fields measured in a single cortical column. Image adapted from Lee et al. \[174\].

Figure 1.14: ‘ABA’ streaming paradigm. a. Images adapted from Shamma et al. \[69\].
Figure 1.15: Attention influences cortical selectivity to sound features. a. Spectrotemporal receptive field (STRF) of an AC neuron recorded in ferret performing a sound detection task. When the animal was engaged in the task, the STRF changed to enhance the representation of the target spectral content. STRFdiff (right panel) indicates the difference between active (Detection; middle panel) and passive listening (Passive; left panel) STRFs. Red–blue areas indicate enhanced–suppressed regions respectively. Image adapted from Fritz et al.\[99\].

b. Left panel: Schematic of stimuli used during an auditory task involving the detection of a target sequence (red) in a mixture of pure tones (maskers; black). Right panel: Illustration of the effect of task engagement on neural tuning. When ferrets were engaged in the target detection task, the tuning curve of AC neurons sharpened (red) compared to during another task where the animals had to use the complete sound mixture to respond (grey; global task). Image adapted from Shamma et al.\[69\].

Figure 1.16: Dynamic increase in A1 neuronal response during an auditory attention task involving temporal expectation. a. Sharpening of frequency tuning (here for one example cell) as the time of the expected (late) target approaches. For each time slot preceding the late target, the estimated tuning curve is plotted in a different color. The color bar shows the time of each time slot with respect to the target onset. b. Gradual increase in neuronal response (measured at each cell’s preferred frequency; PF) as the time of the expected (late) target approaches. Responses are normalized with respect to the response to the seventh tone prior to the first possible late target occurrence (-1050 ms). Each point corresponds to the median across cells, with error bars proportional to the median absolute deviation. Images and captions adapted from Jaramillo and Zador\[61\].
Figure 1.17: Non-exhaustive schematic of anatomical circuits involved in attentional processes. Arrow-heads indicate the type of connection (inhibitory or excitatory). Dashed lines represent connections which are either not yet anatomically defined, or which effects are complex. FC: frontal cortex; AC: auditory cortex; TRN: thalamic reticular nucleus; MGB: medial geniculate body; IC: inferior colliculus.

Figure 1.18: Techniques to selectively manipulate projecting neurons. a. Schematic of possible connections between origin regions (A₁ and A₂) and target regions (B₁ and B₂). Neurons in A₁ target B₁ and/or B₂, and neurons in A₂ target B₂ only. When willing to selectively manipulate the activity of A₁ neurons projecting onto B₂, several approaches are possible (presented in b-d). b. First approach: To inject a virus encoding for an opsin and infecting neurons locally in A₁, and to optically stimulate B₂. Only infected cells with axons in B₂ can be activated. c. Second approach: To inject a virus encoding for an opsin and infecting neurons retrogradely in B₂, and to optically stimulate A₁. Only infected cells with cell bodies in A₁ can be activated. d. Third approach: To inject a first virus encoding for Cre-recombinase and infecting neurons retrogradely in B₂, and to inject a second virus encoding for floxed-DREADD and infecting neurons locally in A₁. Only cells infected with both viruses can be activated.
Seven types of experiment were conducted in this work. The first three formed part of chapter 3 *Functional determinants of response selectivity to frequency modulated sounds*. This incorporated both anaesthetised (Exp 1), awake passive-listening (Exp 2), and awake active-behaving (Exp 3) experiments. Four other types of experiment were performed for the work in chapter 4 *Anatomical circuits underlying task-related modulation of spectrotemporal receptive fields*, incorporating anaesthetised (Exp 4), awake active-behaving (Exp 5), awake passive-listening (Exp 6), and virus construct testing (Exp 7) experiments.

2.1 Animals

Mice were housed under a 12/12h light/dark cycle with food and water available *ad libitum*, except during behavioural training (see section 2.4.4 Training protocols). Electrophysiological recordings and behavioural training were performed during the dark phase of the cycle. All experiments were conducted under the UK Animals (Scientific Procedures) Act 1986.

In passive listening experiments, mice were aged 8-16 weeks and weighed 20-30g on the day of electrophysiological recording. In active behaving experiments, mice were aged 16-18 weeks and weighed 25-35g on the day of electrophysiological recording.

Exp 1 14 adult mice (7 C57BL/6, Charles River UK Ltd & Harlan UK Ltd; 7 homozygous Pvalb-IRES-Cre, JAX stock number 008069) of both genders were used in anaesthetised experiments. From the 7 C57BL/6 mice, 4 were used for baseline experiments and 3 were used for sham control during pharmacogenetic experiments (see section 2.6 Pharmacogenetics).

Exp 2 14 adult mice (4 C57BL/6, Charles River UK Ltd & Harlan UK Ltd; 10 homozygous Pvalb-IRES-Cre, JAX stock number 008069) of both genders were used in passive listening experiments. The C57BL/6 mice were used for sham control.

Exp 3 14 adult mice (Pvalb-IRES-Cre, JAX stock number 008069) were used in behavioural experiments.

Exp 4 20 adult female mice (C57BL/6, Charles River UK Ltd & Harlan UK Ltd) were used in anaesthetised experiments.

Exp 5 8 adult female mice (C57BL/6, Charles River UK Ltd & Harlan UK Ltd) were used in behavioural experiments, split equally into two batches of 4 mice. Each batch was trained on a specific task (T1, T2, see section 2.4.2 Oddball tasks).

Exp 6 8 adult mice (2 C57BL/6, Charles River UK Ltd & Harlan UK Ltd; 6 homozygous Pvalb-IRES-Cre, JAX stock number 008069) of both genders were used in passive listening experiments. The C57BL/6 mice were used for sham control.

Exp 7 5 adult female mice (C57BL/6, Charles River UK Ltd & Harlan UK Ltd) were used in experiments testing for viral labelling strategies.
2.2 *In vivo* electrophysiology

2.2.1 Electrodes

All recordings were made using silicon microelectrodes comprising multiple tetrodes (A32, 4x2Tet; NeuroNexus, USA; see figure 2.1), advanced in the brain using a micromanipulator (IVM, Scientifica, UK).

2.2.2 Targeting the auditory cortex of anaesthetised mice

**Surgical preparations**

**Anaesthesia and surgical preparation (non-recovery)** The mouse was anaesthetized with fentanyl/midazolam/medetomidine mixture (0.05, 5.0 and 0.5 mg/kg) injected intraperitoneally (IP). The body temperature of the mouse was maintained throughout the procedure, by placing the animal either in a heating chamber or onto a heat mat (DC Temperature Control System, FHC). The mouse eyes were covered with lubricant gel (Lacrilube, Allergen, UK) as soon as the animal was judged immobile. After 15 min following anaesthetic injection, the depth of anaesthesia was ascertained by testing the pedal reflex. The skin between the toes was pinched using forceps. The absence of a pedal-withdrawal reflex and a decreased breathing rate and amplitude compared to normal were used as indication that the animal had reached a suitable level of anaesthesia. The pedal-withdrawal reflex was periodically tested to ensure that this level of anaesthesia was maintained. An anaesthetic top-up (1/4 of the initial dosage) was administered IP every hour if not otherwise required. The animal was never recovered.

**Head-implantation** Once the animal was anaesthetised, the hair covering its head was shaved. The mouse was then placed on an aseptic drape, and iodine was applied on the exposed skin. The skin was incised along the midline, and connective tissues were cleared from the cranium. Hatched score lines were made on the cranium using a surgical blade to increase the surface area between bone and adhesive material. A custom-made stainless steel head-implant was fixed using tissue glue (Histoacryl, Braun Corporation, USA) and dental cement (Associated Dental Products Ltd). The animal was then solidly fixated by sliding the head-implant into the head-post (see figure 2.2 for illustration). A small burr hole was made in the cranium, putatively above the right hemisphere visual cortex, and a grounding pin was inserted until it touched the brain. The grounding pin was secured using dental cement.

**Craniotomy** Once the animal was head-implanted, the skin above the temporal plane was removed, and the temporal muscle resected. Using the squamosal fissure as an anatomical marker, a craniotomy was made directly above auditory cortex (typically 1x1mm, centred 2.7 mm posterior of Bregma, see section 2.7 Histology for histological confirmation). The dura was removed, and the brain was lubricated using phosphate buffer saline (PBS, 0.1M).

**Electrophysiology**

All recordings were made in the left hemisphere. Note that whilst previous studies\(^9\) conducted recordings in the right hemisphere, the left hemisphere was chosen here for practical reasons (easier access given the setup configuration). NeuroNexus 4x2 tetrodes were advanced into the auditory cortex using a 90° angle from the vertical line (see figure 2.2). The angle of the recording was
ascertained by mounting the micromanipulator on a stereotaxic frame (Leica Microsystems, D). The recording depth zero-reference was set when the bottom four tetrodes penetrated the brain, as visually assessed using a microscope (Leica Microsystems, D) and by observing drastic diminution in the local field potential noise level. The mean depth of recording was $568 \pm 47 \mu m$ (mean ± standard deviation; $N = 14$).

2.2.3 Targeting the auditory cortex of awake mice

Challenges and approaches
Due to the location of the auditory cortex laying close to the animal's eye, it was not possible to insert the electrode horizontally as done during anaesthetised preparation (see section 2.2.2 Targeting the auditory cortex of anaesthetised mice). Indeed, the animal's behaviour was affected by seeing the electrode approaching or the microscope light. Moreover, the metal implant used during anaesthetised recording was not rigid enough to maintain the animal's head fixed upon sudden movements. Therefore, a new targeting approach was developed, inspired by the work of Sakata\textsuperscript{22} in rats. Instead of a horizontal insertion, the multi-electrode array was inserted at $35^\circ$ angle from the vertical line (see figure 2.3), entering the brain over the primary visual cortex (coordinates: 2.7mm posterior and 2.8mm lateral from Bregma). Typically, an insertion depth of 2.6mm was required to reach the primary auditory cortex. Because of such angle and penetration depth, it was critical to position the head-implant perfectly flat compared to the mouse head. The head-implant was designed so as to allow craniotomies to be done laterally, and simultaneously prevent the animal from seeing any operations undertaken above its head. The design (see figure 2.4) was inspired from a head-implant model developed in the Royer Lab (KIST), routinely used for in vivo electrophysiological recording in the hippocampus of awake mice. The head-implant was 3D-printed using polymer material, and weighed less than 2g after the addition of two bolts used for fixation. During an experiment, the mouse was head-fixed by screwing the implant onto a rigid metal platform. A hole was cut in the metal platform to allow for the electrode to be inserted (visible in figure 2.3c).

Surgical preparations

Anaesthesia and surgical preparation (recovery) The mouse was placed in an induction chamber (Harvard apparatus) and anaesthetised via inhalation of isoflurane (5% v/v, mixed with oxygen, 1.2-1.6 L/min). Once the breathing of the animal had slowed from a normal level, the mouse was removed from the chamber and the top of the animal's head was shaved. The mouse was subsequently placed in a stereotaxic frame (Angle 2, Leica Microsystems, Germany), and covered with aseptic drapes. The isoflurane concentration was lowered to 1-2% v/v, and the depth of anaesthesia was ascertained by testing the pedal reflex. The body temperature of the mouse was maintained throughout the procedure at $37 \pm 0.5^\circ C$ using a homeothermic heat mat (DC Temperature Control System, FHC).

Analgesia Analgesia was provided by injecting carprofen (5mg/kg) sub-cutaneously (SC) 20mn prior to recovery. During the 2-5 days following the operation, the animal was administered orally with buprenorphine jelly (Vetergesic containing 0.3mg/ml buprenorphine in Hartley's Jelly; 0.8mg/kg every 8-12 hours as required) and antibiotics (Baytril; 7mL/250mL water).
**Head-implantation**  The animal was anaesthetised and surgically prepared as described in section 2.2.3 *Anaesthesia and surgical preparation (recovery)*. Iodine was applied on the exposed head skin. The skin was incised along the midline, resected (typical skull exposure: 1x1cm), and the connective tissues were cleared from the cranium. Hatched score lines were made on the cranium over the cerebellum and frontal zones using a surgical blade to increase the surface area between bone and adhesive material. A custom-made plastic head-implant (see figure 2.4) was fixed using tissue glue (Histoacryl, Braun Corporation, USA) and dental cement (Associate Dental Products Ltd). During implantation, the head-implant was screwed onto a custom-made head-post which was connected to the stereotaxic frame. This enabled the head-implant to be implanted perfectly horizontally onto the mouse’s head. Once the cement had dried, the head-post was detached and removed. A small burr hole was made in the cranium above the cerebellum, and a grounding pin was inserted until in contact with the brain. The grounding pin was secured using dental cement. The location of the future craniotomy was measured using a pipette referenced to Bregma, and marked by a cross marked on the skull using a surgical blade. A permanent marker was used to enhance the visibility of the cross. The exposed skull was then covered with Kwikcast (World Precision Instruments). Analgesia was administered as described in section 2.2.3 *Analgesia*.

**Craniotomy**  The animal was anaesthetised using isoflurane and surgically prepared as described in section 2.2.3 *Anaesthesia and surgical preparation (recovery)*, except for shaving the head as the animal was already head-implanted. The Kwikcast was removed, exposing the skull over both hemispheres. A right- or left-hemisphere craniotomy (1x1mm) was made over the cross-marked location (see section 2.2.3 *Head-implantation*). The dura was removed, and the brain was lubricated with PBS. Agar (1% in PBS) was applied over the PBS as a moisturising sealant. Once hardened, the agar was covered with a layer of Kwikcast. Nail varnish was applied at the boundary between the Kwikcast and the head-implant as a further sealant. The mouse was administered with analgesics subcutaneously (see section 2.2.3 *Analgesia* for details), and left to recover in a heating chamber until locomotor and grooming activity were fully recovered. Typically, the animal was exposed to isoflurane (starting in the induction chamber and ending when removed from the surgical table) for 15 minutes.

This surgical preparation was done typically 8 weeks following head-implantation, except for control animals in passive listening experiments where it was done typically after 3 weeks following head-implantation.

**Electrophysiology**

Once the animal was recovered from the craniotomy, it was head-fixed in the apparatus using zinc screws. The back of the animal was gently restrained using a half-tube composed of soft fabric, clamped down by a grounded metal plate (visible in figure 2.3c).

Once a craniotomy was made, up to five subsequent recordings were made in that hemisphere. Recordings were made in the other hemisphere successively. Mice underwent left or right craniotomies first in balanced proportion across the cohort.

The same apparatus was used as when conducting recording in anaesthetised animals, but the configuration was changed (see figure 2.3). The stereotaxic frame was used to tilt the micro-manipulator precisely by 35° away from the vertical line. As with anaesthetised experiments, the recording depth zero-reference was set when the bottom four tetrodes penetrated the brain, as
visually assessed using a microscope and by observing drastic diminution in the local field potential noise level.

2.2.4 Targeting the medial geniculate body of awake mice

Surgical preparations and electrophysiological recordings were performed as described in section 2.2.3 Targeting the auditory cortex of awake mice, with the exception that the probe was inserted vertically (0° angle compared to 35° angle for targeting the auditory cortex) and parallel to the coronal plane. The probe was inserted typically -3.05 mm deep in the brain (measured from the pia), using the following coordinates: AP -3.0mm AP, ML ±2.0 mm from Bregma. Click train stimuli (similar as presented in T2, see section 2.3.3 Stimulus properties for details) were presented to the mouse when inserting the electrode in order to visually assess evoked spiking activity online. The electrode was always placed in a region with evident locking to click stimuli.

2.2.5 Simultaneous recording of prefrontal and auditory cortex activity

Surgical preparations and electrophysiological recordings were performed as described in section 2.2.3 Targeting the auditory cortex of awake mice, with the exception that two probes were used. The first probe was inserted in auditory cortex as previously described (see section 2.2.3 Targeting the auditory cortex of awake mice). The second probe was inserted in the prefrontal cortex, vertically (12° tilt angle from vertical axis) and parallel to the sagittal plane. The probe was inserted typically -2.00 mm deep in the brain, using the following coordinates: 2.10mm AP, ±0.25 ML from Bregma.

2.3 Auditory stimulation

2.3.1 Stimulus generation

Auditory stimuli were pre-generated and calibrated (5-100kHz flat spectrum ±1.5dB SPL) in Matlab (Mathworks, USA) and stored on a PC in text format. The stimuli were presented free-field (ES1; Tucker Davis Technologies, USA) via an RZ6 Processor (using RPvdsEX software; 195312Hz sampling frequency; Tucker Davis Technologies, TDT). A microphone (Alternative IEC61094-4, 1/4 inch, ACO Pacific Inc.) and acoustic calibrator (CR:515, 94dB at 1kHz, Cirrus Research Plc.) were used to measure and calibrate the EZ1 speaker output function. The calibration was implemented by Dr Sollini using custom-written Matlab software.

2.3.2 Stimulus presentation

During passive listening experiments

The RZ6 was controlled using Matlab (ActiveX Control) by a standard alternating-buffer strategy (code provided online by TDT, see figure 2.5 for an illustration). Briefly, the pre-generated signal was chunked into segments and loaded in partitions of the RZ6 buffer. While the RZ6 was reading a buffer partition and sending the voltage to the speaker, the reading index was monitored by the PC. A new stimulus segment was loaded in the RZ6 buffer partition once the reading index had reached the end of that buffer partition. This strategy enabled the continuous read and play of signals whose length exceeded the RZ6 buffer size.

1V square pulses with rise and fall times marking the stimuli onset and offset respectively were pre-generated using Matlab and presented simultaneously to the data acquisition system (see section
During active behaving experiments

An alternative RZ6 controller scheme was used, so as to trigger stimulus presentation in a versatile manner. The RZ6 buffer was partitioned, typically in 8 partitions of 300ms long, and each partition was loaded with a different stimulus. A stimulus could be shorter or equal in length to the partition size. A stimulus presentation was triggered upon detecting a rising edge (0-5V) in the associated digital input signal (RZ6 digital I/O port, byte C), see figure 2.6 for a schematic of the software. Once triggered, a stimulus could only be triggered again after the partition (read) time had elapsed. If multiple stimuli were triggered consecutively, such as shown in figure 2.6, their waveforms were summed before being presented. This strategy enabled the presentation of short (≤ buffer partition size) and potentially overlapping stimuli upon external input triggers. One can think of it as a piano keyboard: any key press will generate a stimulus, but its versatility lies in the possibility to generate chords.

The computation time required by the system to generate an output upon receiving an input was measured using 100ms square volt pulses (0-1V, unramped) as output stimuli. The TDT RZ6 buffer was partitioned in eight 300ms partitions, and each partition was loaded with a pulse signal. Each partition was triggered by a different Arduino pin (used in the behavioural setup, see section 2.4.1 Setup). The Digital Lynx acquisition system (see section 2.5.2 During active behaving experiments) was used to record the input and output signals. The Arduino input signals were recorded via digital ports, whilst the TDT RZ6 output signal was recorded via an analog port. 50 consecutive activation trials were made. One trial consisted of the sequential, isochronous triggering of each partition in a random order, using a 300ms stimulus onset asynchrony, followed by an additional sham trigger generated on a non-connected pin. Threshold crossing was used on the recorded output trace to detect the onset of each pulse. These onset times were compared to the event timestamps recorded from the digital channels (see section 2.5.2 Alignment of events with neural traces). The computation time of the system was 0.7914±0.0136ms (mean±std). This variance was considered negligible when assessing neural responses to auditory stimuli.

2.3.3 Stimulus properties

The start and end of all stimuli were ramped with a 3ms cosine ramp.

During passive listening experiments

Pure-tone frequency response areas (FRAs) were measured using 25 different frequencies (400 ms duration, 7-56 kHz, with 0.125 octave spacing) and 8 different sound levels (10-80 dB SPL at 10dB steps), using a 1 second inter-stimulus interval. Stimulus order was randomly selected with each stimulus repeated 15 times in standard-, and 5 times in each pharmacogenetic-, recording session.

Direction selectivity was measured using frequency-modulated (FM; logarithmic) sweeps spanning 3 octaves, varying in velocity (14 different velocities: ± 2.2; 4.4; 8.8; 17; 35; 70; 140 octave/s). FM sweeps were created using the chirp Matlab function, and were presented at 60dB SPL. Positive velocity represents an UP sweep (from 7 to 56 kHz), and a negative velocity represents a DOWN sweep.
sweep (from 56 to 7 kHz). The order of the stimuli was randomly selected with each stimulus repeated 30 times.

Sequences of five pure tones were used to study LFP phase distribution change (50ms tone, 150-225-300ms stimulus onset asynchrony). The tones varied in frequency and intensity so as to construct a FRA (25 frequencies: 7-56 kHz, with 0.125 octave spacing; 8 sound levels: 10-80 dB SPL at 10dB steps).

**During active behaving experiments**

All paradigms (FM discrimination task, T1, and T2) are presented in section 2.4.2 Paradigms. Schematics of the stimuli and task designs are shown in figure 2.10 (FM discrimination task), figure 2.11 (T1) and figure 4.17 (T2).

For all tasks, a broad-band noise burst (300ms; 1-50kHz, 1/10 octave spacing; 60dB) was used as a punishment reinforcer upon false alarms.

**FM discrimination task** A rising (UP) and a falling (DOWN) FM sweeps were used (250 ms; ±4.4oct/s; spanning 10-21kHz; 60dB) as Go and No-Go cues respectively. At the end of the recording, FRAs were measured using 17 different frequencies (400 ms duration, 7-56 kHz, with 0.25 octave spacing) at 60dB, using a 1 second inter-stimulus interval. Stimulus order was randomly selected with each stimulus repeated 5 times.

**T1** All sounds presented were 100ms long. A High (14kHz) and Low (5kHz) carrier frequency oddball streams were generated, by presenting pure tones isochronously and by introducing a target instead of a pure tone. Per trial, only one target could be presented, except during reversal learning (see section 2.4.2 Oddball tasks). A target was a rising frequency modulated sweep (4.4 octave/s) starting at the stream carrier frequency (i.e. at either 5kHz or 14kHz). Unless stated otherwise, the stimulus onset asynchrony (SOA) of the low and high streams was fixed at 617ms and 400ms respectively. At the beginning of a trial, the two streams started concurrently.

**T2** All sounds presented were 100ms long. A single carrier frequency (10kHz) oddball stream was generated, by presenting pure tones isochronously (600ms SOA) and by introducing a target instead of a pure tone. The silent portion of that stream was filled with an isochronous sequence (150ms SOA) of three distracting pure tones randomly selected from a distribution of surrounding frequencies (±1.5 and ± 0.75 octave from 10kHz, i.e. [4, 6, 17, 28] kHz). 8 different distractor sequences were used (see table 2.1). Per trial, only one target could be presented. A target was a rising frequency modulated sweep (4.4 octave/s) starting at the stream carrier frequency. At the beginning of a trial, the first sound was always a carrier frequency pure tone. The beginning of a trial was marked by a 3-click train (click duration 1ms, SOA 100ms, 60dB).

<table>
<thead>
<tr>
<th>Order</th>
<th>Sound ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>first</td>
<td>1 4 1 2 4 3 3 2</td>
</tr>
<tr>
<td>second</td>
<td>4 1 2 1 3 4 2 3</td>
</tr>
<tr>
<td>third</td>
<td>2 2 3 3 1 1 4 4</td>
</tr>
</tbody>
</table>

**Table 2.1:** Sequences of distractor sound used in T2.
2.4 Behavioural training and assessment

2.4.1 Setup

During behavioural training, mice were headfixed to the same platform as used during electrophysiological recording in awake mice. An Arduino UNO (www.arduino.cc) was used as a controller, interfacing with a PC via the serial port (see figure 2.7a). Outputs from the Arduino were sent to the PC and saved in text format using custom-written scripts in Python. Data from the text file were further analysed and plotted using custom-written Python scripts so as to display the mouse performance online (with a single trial delay). Examples of such interface can be seen in figure 2.8 and figure 2.9.

A lick port was used to monitor the animal’s behaviour (see figure 2.7b). The lick port was composed of an infra-red LED and sensor electrical circuit, a water delivery system and a vacuum system used for water removal. Briefly, a drop of water was delivered upon the opening of a solenoid valve clamping the water delivery tube. The control signal was a digital square pulse sent by an Arduino. Similarly, the vacuum removal was triggered upon the activation of a solenoid valve, connected to a positive pressure. The lick port was mounted onto micromanipulators so as to position it finely compared to the mouse jaw. The speaker was positioned in front of the animal.

2.4.2 Paradigms

Mice were water deprived, and used licking to indicate a response. A single lick was considered as a Go-response.

Go/No-Go frequency modulation discrimination task

The prime interest of this study was to assess whether mice were able to perceptually discriminate the direction of slow FM under conditions of reduced cortical inhibition. Therefore, the sounds used in the task varied only in their FM direction. The FM rate was fixed at 4.4 oct/s. Two sounds were used in this study: a rising (UP) FM sound, and a falling (DOWN) FM sound (see section 2.3.3 Stimulus properties for details).

The discrimination task was designed as a Go/No-Go task, where a Go-response (lick) to the UP FM sound was reinforced with a reward (water drop), and a Go-response to the DOWN FM sound was associated with a mild punishment (noise and time out). Mice had to refrain licking for a variable period of time (600-1200ms) for a sound to be presented. UP and DOWN sweep were pseudo-randomly presented with equal probability, such that one sound type was not presented more than five times consecutively. The response window was set as 200-650ms post-stimulus onset.

Oddball tasks

The prime interest of this study was to develop selective-listening paradigms in mice. Inspired by a previous study in primates\cite{109}, mice were trained on two different tasks (T1,T2) based on oddball streams, i.e. sequences of pure tones containing a deviant sound.

In all tasks, a vacuum pump was turned on for 500ms at the end of every trials to remove the excess water present in the lick port. The noise and sensory feedback from the deactivation of the vacuum pump could be used by the mice as a cue for trial onset. A random delay (100-150ms) was applied before presenting any stimulus following vacuum offset. The reward window started 100ms after target onset, and ended when the next sound of the corresponding stream occurred, i.e. the

42
length of the reward window was the SOA minus 100ms.

**T1** Mice were trained on a dual-stream oddball paradigm (high and low frequency stream). An oddball stream was composed of a sequence of pure tones of a given frequency (High or Low), containing a target (frequency modulated) sound (see section 2.3.3 Stimulus properties for details). Mice were cued to expect a target in a particular stream using a trial block design. Per block, the target alone was first presented (S0 condition) until the mouse reached 5 Hit trials. Subsequently, the oddball stream associated with that target was presented (S1 condition) until the mouse performed 10 Hit trials. Finally, the two streams were presented concurrently (S2 condition) until the mouse performed 20-25 Hit trials.

In an effort to match previous studies in macaques\(^{[109]}\), the ability for reversal learning was tested. Following reversal learning testing, temporal expectation was modulated: a target could occur Early or Late depending on its frequency content. Specifically, High frequency target occurred Early (~2s from trial onset), whilst Low frequency target occurred Late (~6s from trial onset).

**T2** Mice were trained on a single-stream oddball paradigm, whose silent portion contained distracting pure tones (see section 2.3.3 Stimulus properties for details). Temporal expectation was modulated as in T1, using ~2s and ~4s for early and late temporal window respectively.

2.4.3 Measure of performance

In the FM discrimination task, one trial could give 2 outcomes: 1) in the case of a Go cue being presented, the mouse could lick (Hit) or not (Miss); 2) in the case of a No-Go cue being presented, the mouse could lick (false alarm; FA) or not (correct rejection; CR).

In oddball tasks, all sounds presented before the target window were considered as one No-Go cue, and all sounds presented after as another No-Go cue. One trial could thus give 4 outcomes depending on the mouse response timing: 1) if the mouse licked before the target, a FA was counted; 2) if the mouse licked in the target window, a CR + Hit was counted; 3) if the mouse licked after the target window, a CR + Miss + FA was counted; 4) if the mouse did not lick during the trial, a CR + Miss + CR was counted.

In all tasks, performance was measured using the sensitivity index \(d'\), defined as \(z(P_{Hit}) - z(P_{FA})\), where \(z\) is the \(z\)-transform, \(P_{Hit}\) and \(P_{FA}\) the probability of Hit and FA respectively. \(P_{Hit}\) was defined as \(N_{Hit} / N_{Go}\), i.e. the number of Hit trials divided by the number of Go cues presented (note that \(N_{Go} = N_{Miss} + N_{Hit}\)). Similarly, \(P_{FA}\) was defined as \(N_{FA} / N_{No-Go}\), i.e. the number of FA trials divided by the number of No-Go cues presented (note that \(N_{No-Go} = N_{CR} + N_{FA}\)). In the case of \(P_{Hit} = 0\) or \(P_{FA} = 0\), the value was replaced with \(1/N_{Go}\) or \(1/N_{No-Go}\) respectively. In the case of \(P_{Hit} = 1\) or \(P_{FA} = 1\), the value was replaced with \((N_{Go} - 1)/N_{Go}\) or \((N_{No-Go} - 1)/N_{No-Go}\) respectively.

2.4.4 Training protocols

Typically, mice required the following total number of training sessions to achieve high performance in the different tasks: 30 sessions for the FM discrimination task, 50 sessions for T1, and 40 sessions for T2. The total number of sessions is counted from the first day of water restriction, and until high behavioural performance is reached for three consecutive sessions.
**Water restriction**

Following a minimum of 5 days post-surgery recovery, the mouse was placed under water restriction. On the first day of water restriction, no water was given. On the second day, 1mL of water was given. Subsequently, the mouse began its training protocol on the behavioural setup, where it received most of its daily water dose. The minimal water dose given per day was fixed to 1mL. If the mouse did not perform enough trials in a single session so as to reach the daily dose, the complementary water volume was given at the end of the training session once the mouse had been removed from the head-fixing apparatus. Mice were typically trained 6 to 7 days per week.

Once under water restriction, the weight of the mouse was measured daily, prior to the delivery of water. In the case of the animal's weight decreasing below 80% of an aged-match, water unrestricted litter-mate, the animal was given water ad libitum for a day. In order to give a precise dose to an animal in isolation, the mouse was placed on a scale, and its weight recorded. The dose of water was then administered either directly to the mouse's mouth using a pipette, or by squirting the water out onto the scale's plastic flooring. The difference in body weight measurement pre- and post- water delivery ensured that the animal had drunk its daily dose.

**Habituation**

The first 3 training sessions consisted of habituating the mouse to become head-fixed in the apparatus and to associate the lick port with water delivery. During these habituation sessions, the reward delivery valve was replaced by a syringe full of water that could be manipulated by the experimenter so as to deliver water into the lick port on demand.

During a session, the mouse was fixed into the setup, and a pipette full of water was first advanced to the animal's mouth so as to prompt licking behaviour. Once the mouse had started to express exploratory licking behaviour, the lick port was filled with water and advanced so as to be reached by the animal's tongue. The lick port was filled with water until the animal displayed signs of satiation, i.e. stopped licking and attempted to push away the lick port with its forelimbs. The mouse was then removed from the apparatus, and supplementary water was given if necessary.

**Target-reward association**

Following the habituation sessions, the mouse underwent typically 7 sessions to associate the water delivery with target sound(s). During the first 4 sessions, an automatic reward (water drop of 5µL) was given 100ms after sound onset. During the 3 consecutive sessions, the mouse had to trigger a reward by licking in the reward time window. During training for the FM discrimination task, the proportion of Go cues was decreased within and across sessions. A fixed amount of Hit trials were required to subsequently present No-Go cues (see table 2.2). For the oddball tasks, the target sound was presented in isolation. In T1, the two targets were presented, using a 20-30 trials block design.

<table>
<thead>
<tr>
<th>Sessions</th>
<th>1</th>
<th>...</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Hit trial at start of session</td>
<td>100</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>Percentage of No-Go trial</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

*Table 2.2:* Proportion of Go and No-Go cues across sessions during task-reward association training in the FM discrimination task.
**Refrain licking prior to target onset**

During 5 sessions following the target-reward association sessions, the mouse was required not to lick before the occurrence of a sound (note that unlike in the FM discrimination task, all sounds presented at this stage are targets in the case of oddball tasks). The length of time the mouse was required to withhold licking for was randomly selected from a uniform distribution whose boundaries were incrementally increased over the sessions ([200-500], [400-700], [600-1000], [800-1200], [1000-1400] ms). Following these sessions, mice trained in the FM discrimination task went onto the stage presented in section **2.4.4 Reinforcing false alarms**, whilst mice trained in oddball tasks went onto the stage presented in section **2.4.4 Introduction of single stream**.

**Introduction of single stream**

For the following 8 sessions, pure tones were introduced so as to form a single-stream oddball paradigm. Within a training session, the intensity of the pure tone was augmented from 0dB up to the intensity level used for the target sound using 10dB increment steps. The amount of hit trials required to augment the intensity was diminished across sessions (50-30-20-15-10-7-5-3). The mouse was required not to lick before the occurrence of the target, using the longest interval used in the section **2.4.4 Refrain licking prior to target onset** training phase. In T1, the SOA varied depending on the frequency of the pure tones (see section **2.3.3 Stimulus properties**). Following these training sessions, mice went onto the stage presented in section **2.4.4 Introduction of distractor sounds**.

**Introduction of distractor sounds**

For the following 4 sessions, pure tones were introduced so as to form either another single-stream (T1) or a cloud of distractor tones (T2). Within a training session, the intensity of the extra tones were augmented after each 10 hit trials from 40dB up to the level used for the target sound using 10dB increment steps. The amount of hit trials required to present extra tones was diminished across sessions (20-15-10-7). Following these training sessions, mice went onto the stage presented in section **2.4.4 Reinforcing false alarms**.

**Reinforcing false alarms**

Mice trained in T1 underwent the training presented in this section for 2 sessions prior to the training described in section **2.4.4 Reversal learning**. Mice trained in T2 underwent the training presented in this section for 10 sessions prior to the training described in section **2.4.4 Introduction of a temporal bias**.

In all tasks, response to a No-Go sound (FA) was negatively reinforced by presenting a short noise burst upon the response, stopping any upcoming sound presentation, and subsequently applying a silent time out (random delay of 4-6s). The vacuum valve was turned on after the time out had passed. Such punishment was not applied in oddball tasks when the response was made after having missed target. Instead, the oddball sequence was terminated after presenting four regular tones past the target, and the trial was ended regularly (i.e. without time out).
Reversal learning

Only mice trained in T1 underwent the training sessions described in this section. Mice had to refrain licking prior to target onset as presented in section 2.4.4 Refrain licking prior to target onset. Mice underwent the training presented in this section for 7 sessions. Note that mice did not reach performance criterion in this task (data not discussed as part of this thesis), and were transferred to another training paradigm following these sessions (see section 2.4.4 Introduction of a temporal bias). The details outlined in this paragraph are merely for the reader to gain a complete overview of the training that the animals received.

A trial block comprised three different types of trials. Firstly, trials where only the target of a specific frequency would be presented (target alone condition; S0). After a total of 5 correct responses (Hit) in this type of trial, the pure tone of the specific frequency was also presented (single stream condition; S1). After a total of 10 correct responses in this type of trial, the second stream of the other frequency was also presented (dual-stream condition; S2). The sounds in S0 and S1 conditions are considered as cues.

Per trial block, mice were required to respond to the FM sound target (Go) only if its frequency content matched those of the cued stream. Both targets were presented in the dual-stream (S2) condition only. Responses to the non-cued (No-Go) target were negatively reinforced by presenting a noise burst upon response detection. Following a No-Go response, the trial was continued, i.e. pure tones were presented, and the mouse had to refrain licking prior to a following target to be presented. The identity (i.e. cued or non-cued) of the following target depended on the response of the animal on the prior target. If the mouse generated a FA, the following target was un-cued with a 70% probability. If the mouse generated a Miss, the following target was cued with a 70% probability. The trial was terminated once the animal generated a correct (Hit) response, or after 3 consecutive Misses. Across sessions, the probability of presenting a non-cued stream as first target in the trial augmented (10-20-30-40-50).

Introduction of a temporal bias

Only mice trained in oddball tasks underwent the training sessions described in this section. Reinforcement of FA (see section 2.4.4 Reinforcing false alarms) was always applied during this training stage.

For the following 5 sessions, the time window in which a target could occur was modified across session. For T1, only the window of the Low target varied so as to occur gradually later in the trial. Per session, the target window was fixed according to a trial block design (Early or Late block). On the final day of training, the Late window was \( \sim 6s \) and \( 4s \) in T1 and T2 respectively, and the Early window was \( \sim 2s \) in both task.

2.5 Data acquisition

2.5.1 During passive listening experiments

Data were acquired via Digital Lynx 16SX system (Neuralynx, USA; 64 analog channels; 32552 Hz sampling frequency) and stored on a PC. The 32 channels of the multi-electrode array were recorded, using either the intrinsic filter of the system (0.5-2000Hz), or using a high-pass filter (300-2000Hz) in order to better visualise spiking activity online. In the latter case, the local field potential (LFP) was
recorded by duplicating and applying a low-pass filter (0.5-300Hz) to one channel of each tetrode. 4 analog channels out of the 64 available were set as external analog channels, and 2 of them served to record the analog outputs of the RZ6 (stimulus and trigger traces; see figure 2.12a for a schematic diagram of the hardware connections.

**Alignment of stimuli with neural traces**

The recording system Digital Lynx received square voltage pulses from the stimulus generator RZ6 as analog input, enabling alignment of stimuli with neural traces during post processing. Threshold crossing was applied on the trigger trace during post-processing to detect stimulus occurrences and align them with neural traces. See figure 2.12b for example neural and stimuli traces recorded simultaneously.

### 2.5.2 During active behaving experiments

The data were acquired using the Digital Lynx system as presented in section 2.5 Data acquisition, except that inputs other than electrodes were recorded via digital rather than analog ports. This was mainly due to the large amount of inputs needed (11 in total, see section 2.4.1 Setup), as only 4 channels of the system could be converted to external analog input channels. The figure 2.13a presents the different inputs sent to the acquisition system. The output signal of the TDT RZ6 was recorded using the Digital Lynx only for calibration purposes.

**Alignment of events with neural traces**

The Digital Lynx recording system received digital inputs marking stimulus or behaviour events, enabling the alignment of such events with neural traces during post processing. The figure 2.13b presents a schematic of neural and stimulus traces recorded simultaneously. The top plot represents a fluctuating voltage trace (black), sampled at a regular interval (grey circles). Each data sample is associated with a timestamp (TS, stored in 10^6 samples/sec format and referenced to the internal computer clock), see figure 2.13b grey inset. The two bottom plots represent stimulus inputs (voltage change and associated sound output using the same nomenclature as in section 2.3.2 During active behaving experiments). To save storage space, digital input events (EV) were stored upon a state change. The identity of an event was stored in binary format, and the timestamp of an event was stored in the same format as the data sample timestamp. Timestamp matching was sufficient for the alignment of the local field potential with the input events. However, as spiking data was detected during post processing, the timestamps of spike was in data increment format (IN). Therefore, the IN had to be aligned with the TS in order to align the EV with the spiking data.

### 2.6 Pharmacogenetics

#### 2.6.1 Viral vectors

**Exp 1** At 8 weeks of age, the auditory cortex of Pvalb-IRES-Cre (N = 7) mice was unilaterally injected with AAV1/2-hSyn-DIO-hM4Gi-mCherry to selectively express the inhibitory designer receptor exclusively activated by designer drug (DREADD) in parvalbumin-positive (PV+) interneurons. For sham control experiments, C57/BL6 adult mice were used (N = 3) in lieu of transduced PV-Cre animals. Sham control animals did not undergo viral injection.
Exp 2 10 Pvalb-IRES-Cre mice were injected bilaterally in auditory cortex with AAV1/2-hSyn-DIO-hM4Gi-mCherry to selectively express the inhibitory DREADD in PV+ interneurons.

Exp 3 14 Pvalb-IRES-Cre mice were used in behavioural experiments, out of which 10 were injected bilaterally in auditory cortex with AAV1/2-hSyn-DIO-hM4Gi-mCherry, and 4 were injected bilaterally in auditory cortex with AAV1/2-hSyn-DIO-hM3-mCherry to selectively express the inhibitory (hM4) or excitatory (hM3) DREADD in PV+ interneurons.

Exp 6 6 Pvalb-IRES-Cre mice were injected bilaterally in the thalamic reticular nucleus with AAV1/2-hSyn-DIO-hM4Gi-mCherry to selectively express the inhibitory DREADD in PV+ interneurons.

Exp 7 4 WT mice were injected unilaterally (2 mice) and bilaterally (2 mice) with CAV-Cre-GFP in the medial geniculate body. 1 WT mouse was injected in the left hemisphere with AAV1/2-hSyn-DIO-hM4Gi-mCherry, and in the right hemisphere with CAV-Cre-GFP.

DREADD activation

The DREADD agonist Clozapine-N-Oxide (CNO; 5mg/kg body weight; 5mg CNO dissolved in 5ml saline) was administered IP in all pharmacogenetic experiments.

2.6.2 Viral injection

In auditory cortex

Once the animal was anaesthetised and placed in the stereotaxic frame such as presented in section 2.2.3 Anaesthesia and surgical preparation (recovery), a small craniotomy was performed -2.7 mm posterior and 4.3 mm lateral of Bregma. A glass pipette was lowered vertically in auditory cortex, and 500 nL of virus was injected at two injection sites (injection rate of 50 nL/min). The pipette was initially inserted to a depth of 1100 µm and 250 nL of virus was injected; 5 min later, the pipette was retracted to a depth of 800 µm where the remaining 250 nL of virus was injected. The pipette was then removed, and the tissue sutured. The animal was given analgesics and recovered. The virus was left to express for 8 weeks prior to electrophysiological recording.

If the mouse was scheduled to undergo awake passive-listening experiment (Exp 2) or behavioural training (Exp 3), it was injected bilaterally in the ACx, using a similar procedure as previously described, with the exception of the suture as the mouse was subsequently head-implanted as described in section 2.2.3 Head-implantation.

In thalamic reticular nucleus

Once the animal was anaesthetised and placed in the stereotaxic frame such as presented in section 2.2.3 Anaesthesia and surgical preparation (recovery), three small craniotomies were performed on each hemisphere. Several injection sites were used per craniotomies. The coordinates for the craniotomies and injection sites were: 1) AP -1.85 mm, ML ±2.15 mm, DV: 3.4 / 3.2 / 3.0 / 2.8 mm ; 2) AP -1.35 mm, ML ±2.15 mm, DV: 2.9 / 2.7 / 2.5 mm ; 3) AP -0.85 mm, ML ±1.80 mm, DV: 3.4 / 3.1 / 2.8 mm (AP: anterior-posterior, measured from Bregma ; ML: medio-lateral, measured from Bregma ; DV: dorso-ventral, measured from the pia). A glass pipette was lowered vertically in the thalamic reticular nucleus and 100 nL of virus was injected per site (injection rate of 50 µnL/min). The pipette was initially inserted to the deepest depth; following 3 min after injection, the pipette was retracted to
the next deepest depth. Once all injections were completed, the pipette was then removed, and the animal was head-implanted as described in section 2.2.3 Head-implantation. The animal was given analgesics and recovered. The virus was left to express for 8 weeks prior to electrophysiological recording.

**In medial geniculate body**

Once the animal was anaesthetised and placed in the stereotaxic frame such as presented in section 2.2.3 Anaesthesia and surgical preparation (recovery), a small craniotomy was performed -3.00 mm posterior and 1.85 mm lateral of Bregma. A glass pipette was lowered vertically in the medial geniculate body and 500 nL of virus was injected at one injection sites (injection rate of 50 nL/min). The pipette was then removed, and the tissue sutured. The animal was given analgesics and recovered. The virus was left to express for 6 weeks prior to sacrificing the animal.

2.6.3 Auditory cortex inactivation

To transiently inactivate A1, mice were injected intracortically and bilaterally in A1 with the gamma-aminobutyric (GABA) agonist muscimol (product number M23400, ThermoFisher, USA; 0.6 µl of 1µg/µl solution dissolved in 0.9% saline) prior to behavioural session. Mice were anaesthetised (as described in section 2.2.3 Anaesthesia and surgical preparation (recovery)), and a small hole was made in the skull bilaterally above A1. A glass pipette was inserted into A1 vertically, and the muscimol was injected following the same procedure as for viral injection (see section 2.6.2 Viral injection). The left hemisphere was injected first, and 10 min elapsed between the injections in left and right hemispheres. Mice were placed in the behavioural apparatus 40 min following the muscimol injection in the right hemisphere. As sham-control, saline was injected intracortically and bilaterally following the same procedure one day prior to muscimol injection (within animal control).

In pilot experiments, extracellular population recordings in anaesthetised mice were used to confirm that A1 multi-unit activity was abolished within 10 min of injection. In such experiments, WT mice were implanted unilaterally in A1 (left hemisphere) with a guide cannula, vertically inserted 700 µm from the pia and fixed with dental cement during head-implantation. During recording, the muscimol was delivered at a rate of 0.1 µl/min via a Hamilton syringe into the guide cannula. The experiments terminated 20 min following muscimol injection.

2.6.4 Recording paradigms

**Exp 1** The recording paradigm in pharmacogenetic experiments consisted of a 25min baseline recording where auditory stimuli were presented, followed by an IP injection of CNO. Following a silent 20min period post-injection, auditory stimuli were again presented for 25min. Thereafter, the probe was advanced deeper in the brain (typically 100µm), and another 25min recording was performed.

**Exp 2** The recoding paradigm was similar as in Exp 1, with the exception that the mouse was briefly anaesthetised with isoflurane following the 25mn baseline recording and prior to IP injection. The mouse was kept fixated in the apparatus for this procedure, and the electrode remained implanted. Following CNO administration, the isoflurane was removed and the mouse recovered. The mouse was always conscious following the 20mn silent period prior to the auditory stimulation post-CNO. Only one recording depth was used.
Exp 3  The mouse was briefly anaesthetised with isoflurane, following which it received an IP injection of CNO. After 20 min post-injection, the mouse was placed in the behavioural setup where it began its behavioural session (with or without concurrent electrophysiological recording).

2.7 Histology

Mice were deeply anaesthetised with Euthatal and transcardially perfused with PBS followed by 4% paraformaldehyde solution (PFA, wt/wt in PBS). Brains were extracted, placed in PFA solution and stored at 4°C overnight. Coronal brain sections (60-100µm) were cut using a vibrating microtome (VT1000 S, Leica Microsystems, Germany). Immunohistochemistry on mouse brain section was performed using rabbit anti-PV antibody (Swant, code no. PV 27) and Alexa Fluor 488 goat anti-rabbit IgG antibody (Life Technologies, code no. A11008). Fluorescence images were acquired on a Leica SP5 laser-scanning confocal microscope. Multi-tile scanned images were performed using a x10 air objective and x40 oil objective. No zoom was used. The laser power and gain of the PMTs were adjusted to avoid photobleaching and image saturation.

Tissue preparation and imaging were performed by both A. Haddad and G. Chapuis. All grey-shaded brain coronal sections presented in the supplementary materials are unstained tissue.

Whilst Di-I was consistently applied onto the electrodes to enable their localisation post-mortem, coronal sections were not prepared if the neural responses appeared as clearly auditory responsive, tuned to frequency (typically V-shaped), and with short first-spike latency (first spike at 20-30ms following stimulus onset).

2.8 Data analysis

All data was analysed using Matlab (Mathworks; USA), with the exception of spike sorting.

2.8.1 Spike sorting

Klusta (https://github.com/kwikteam/klusta) was used to detect and sort single cells based on their waveforms measured per tetrode. Waveform clusters were manually inspected using Klustaviewa and reclustered when necessary. Clusters that contained >1% of spikes within a 3 ms interspike interval, or that had low isolation quality (≤ 0.8) were rejected (http://neurosuite.sourceforge.net). Five example well-isolated single cells recorded from one tetrode are displayed in figure 2.14. Typically, a total of 30 cells were classified as well-isolated across all tetrodes for a given penetration (29.8 ± 8.3, mean ± standard deviation; N=9 recordings).

2.8.2 Selection of auditory responsive cells

Once a cell was categorised as well-isolated, it was further classified as auditory responsive if it produced a transient response to any of the stimuli presented. The classification was done manually and was based on visual inspection of raster plots displaying spiking activity pre- (400ms), during- and post- (500ms) stimuli presentation. On average, 69 ± 10% (mean ± standard deviation; N = 9 recordings) of cells were classified as auditory responsive.
2.8.3 Receptive field analysis

Exp 1-3 ON and OFF FRAs were generated by computing the mean firing rate in the 0-300ms window following tone onset and offset respectively. Raw FRAs were smoothed (3x3 pyramidal window) for analysis, and further resized (scale factor = 4) for display.

A response ($R$) was classified as evoked and excitatory if it was above a threshold value ($Threshold$; 30%) set between the global maximum response ($Max$) and the baseline evoked response ($Baseline$; computed as the mean firing rate response to all tones presented at the lowest intensity): $R > Baseline + Threshold \times (Max - Baseline)$.

The FRA edges were defined as the iso-response curves$^{[72]}$ delimiting the evoked region of the FRA. The characteristic frequency (CF) was taken as the frequency yielding an evoked response at the lowest intensity level (dB threshold). The FRA bandwidth was defined as the width of the evoked region of the FRA at 30dB above dB threshold. The best frequency (BF) was measured at the frequency-intensity level combination that yielded the global maximum response.

In order to select cells possessing ON and OFF responses (ON-OFF), raster plots, peri-stimulus time histograms (PSTHs) and FRAs were visually assessed. A cell was manually classified as ON-OFF if it produced a transient response at the beginning and at the end of pure tones. Cells that did not contain clear structure for both onset and offset FRAs or that were considered noisy were rejected.

Exp 5 In T1, the frequency selectivity index (FSI) was defined as $(r_1-r_2)/(r_1+r_2)$, where $r_1$ is the response to the Low frequency tone and $r_2$ the response to the High frequency tone. The response was measured in the 0-100ms window following sound onset. For LFP, the response was measured as the minimum voltage value in that window.

Significant spiking response to a tone was measured using the activity in the 100ms window pre-tone (spontaneous) and post (evoked) tone onset. A 10ms bin PSTH was generated for each trial, and the evoked response at each bin was compared (KS test) against the spontaneous response. A cell was classified as significantly evoked if either 3 bins passed the threshold of $p < 0.01$ or if one bin passed the threshold of $p < 0.001$ (Bonferroni correction factor 10).

2.8.4 Direction and rate selectivity analysis

Exp 1-3 The direction selectivity index (DSI) was defined as $(r_1-r_2)/(r_1+r_2)^{[8]}$, where $r_1$ is the response to the UP sweep at a given FM speed, and $r_2$ the response triggered by the DOWN sweep at the same speed. The response was defined as the mean number of spikes triggered by the stimulus unless stated otherwise. The time window for measuring a response started at sound onset and ended 100 ms after sound offset.

The maximum instantaneous firing rate was used when assigning best FM rate. The maximum firing rate evoked upon a given FM sweep was measured over the mean smoothed-PSTH (5ms bin, convolved with 20ms std Gaussian function for smoothing). The same time window was used as for the DSI.

2.8.5 Discrimination performance

Exp 1-3 The discrimination performance (DP) for UP and DOWN FM classification was computed using logistic regression and receiver operating characteristic (ROC) analysis. Responses on single
trials were measured for UP and DOWN sweeps of similar FM speed, using the metric used to compute the DSI (e.g. mean number of spike). A logistic regression model was fit on the single trial responses. The ROC curve was computed using the probability estimates from the logistic regression model. The DP was defined as the area under the ROC curve (AUROC). The criterion for informative AUROC was arbitrarily set at 0.65.

Exp 5  The discrimination performance (DP) for trial type (such as Early/Late, Tone/Target) classification was computed using logistic regression and receiver operating characteristic (ROC) analysis. Smoothed-PSTH on single trials were generated for the two conditions to compare (5ms bin, convolved with 20ms std Gaussian function for smoothing). A logistic regression model was fit at each bin of the single trial responses. The ROC curve was computed using the probability estimates from the logistic regression model. The DP was defined as the area under the ROC curve (AUROC). The criterion for informative AUROC was arbitrarily set at 0.6.

2.8.6  Circular data analysis

Functions from the Circular Matlab Toolbox\cite{176} were used when analysis circular data. The vector strength was defined as

\[
VS = \sqrt{\left(\sum_{i=1}^{N} \cos(\alpha_i)\right)^2 + \left(\sum_{i=1}^{N} \sin(\alpha_i)\right)^2} / N
\]

(2.1)

where \(\alpha_i\) is the phase angle at trial \(i\), and \(N\) the number of phases used to compute the VS.

2.8.7  Local field potential and multi-unit analysis

Signals were band-pass filtered (0.5-300Hz) to obtain local field potential (LFP) traces. Since the four electrodes comprising a tetrode are in close vicinity, one electrode was used per tetrode to measure the LFP, consistently chosen as the middle-left electrode of the tetrode for all experiments. The phase of the signal was computed using the Hilbert transform after appropriate alignment and filtering of the LFP.

Exp 1-3  The local field potential (LFP) and multi-unit (MU) comparative analysis in Pvalb-IRES-Cre and control animals were computed using only the recordings where baseline and post-CNO injection traces were acquired at the same penetration depth. Mean changes in evoked LFP and MU responses were calculated per tetrode, using only the responses to stimuli contained within the FRA edges of the baseline condition. The mean change for one tetrode was defined as mean(R2./R1), where R1 is a vector containing the mean responses obtained for each stimulus selected in the baseline period, and R2 the mean responses obtained in the post-CNO period. A value above 1 denotes an increase in activity post-CNO injection compared to baseline. The LFP response was quantified as the global minimum over the response window (0-300ms after sound onset for evoked ON responses). The minimum was chosen as LFP evoked responses were downwards deflections from the baseline voltage. A MU response was quantified as the number of spikes present in the response window (0-100ms after sound onset / offset for evoked ON / OFF responses).

Mean change in baseline MU activity for one tetrode was defined as S2./S1, where S1 and S2 are the mean number of spikes detected in the 100ms window preceding the pure tones presented in
the baseline and post-CNO periods respectively. Tetrodes with less than 1 spike on average in the baseline (S1<1) were excluded from the comparative analysis.

Changes in LFP and MU CF are computed as \( \log_2(CF2/CF1) \), i.e. the octave difference between the CF calculated in the baseline condition (CF1) and the CF calculated post-CNO injection (CF2).

**Exp 4**  In experiments using sequences of 5 sounds lasting 50ms, the LFP traces were aligned to the first stimulus onset and averaged. The average LFP for a given stimulus sequence was then band-pass filtered using \( \pm 2\)Hz around the stimulation frequency. The phase of the averaged and filtered trace was computed. A tetrode was included in the analysis if at least one sound-responsive single cell had been detected on that tetrode. The ON-frequency response area (FRA) of the LFP was measured using the 0-50ms window following sound onset. Excitatory and inhibitory responses were computed as the minimum and maximum deflection of the averaged LFP in that window respectively. Baseline LFP was computed similarly, using the 100-500ms window preceding tone onset. LFP \( z \)-score was computed as \( z = (x - \mu) / \sigma \), where \( x \) is a given LFP response, \( \mu \) the average baseline response and \( \sigma \) the standard deviation of the baseline response. The criterion for excitatory evoked response was set as \( z < -4 \) for the first sound presented, and \( z < -2 \) for subsequent presentations. The criterion for inhibitory evoked response was set as \( z > 1.5 \). For each tone in a sequence, the phase values reported are measured at 20ms post stimulus onset.

**Exp 5**  To measure phase-reset in behavioural experiments, a periodic LFP trace was generated for each stream by aligning the LFP trace to all pure tone stimuli of a given frequency, except for the first tone of a trial (due to the close proximity in time of the vacuum sound ending). A Fourier interpolation (maximum 2 terms) was performed on the averaged LFP period after removal of the evoked period (0-200ms post stimulus onset) using the Matlab Interpolation App. The phase of the interpolated traces at stimulus onset was used in the comparative analysis. Only interpolated traces with R-square values above 0.85 were included in the analysis. The phase values reported are measured at stimulus onset.

To measure phase distribution change across time in behavioural experiments, the LFP trace at each trial was band-passed filtered (1-4Hz). At a given time point, significant difference in the phase distribution between different trial types was assessed using the Kuiper test from the Circular Toolbox. The difference between phase distribution histograms (Early minus Late) at a given time point was measured using the normalised histograms computed at 10 equally spaced bins between -\( \pi \) to \( \pi \). Linear interpolation was performed per tetrode on the unwrapped 95% percentiles of the difference distribution matrix. The interpolated phase was measure at 20ms post first Early target onset (i.e. at 1.620s after trial start).

### 2.8.8 Modelling of experimental data

**Exp 1**  A linear model was generated based on the equation \( y = a_1x_1 + a_2x_2 + \ldots + a_nx_n \), where \( y \) represented the variable to predict (i.e. DSI), \( x_j \) represented the different predictor properties, and \( a_j \) represented the estimated coefficients corresponding to each of these properties. Initially 10 functional neuronal properties were used, namely (1) the firing rate increase evoked by onset of a pure tone of CFON frequency at 60dB, (2) the firing rate increase evoked by offset of a pure tone of CFOFF frequency at 60dB, (3,4) Fano factor of each of these two responses, (5) spontaneous firing rate and (6) associated Fano factor, (7,8) the bandwidths of ON and OFF RFs
measured at 30dB, (9) the percentage overlap between ON and OFF RFs, and (10) the octave difference between ON and OFF CFs. The number of observations used to estimate the model coefficients was equivalent to the number of cells with ON and OFF RFs, i.e. 30 for adult mice and 20 for young mice. To measure the individual contribution of a property $x_j$ towards the model performance, the component $x_j$ was removed from the original equation. The truncated equation $y_{\text{trunc}} = a_1 x_1 + a_{j-1} x_j - 1 + a_{j+1} x_{j+1} + a_n x_n$ was then used to predict DSI values, from which the residual sum of squares $SSE_{\text{truncated}}$ were computed. The proportional reduction of error was defined as $PRE = (SSE_{\text{truncated}} - SSE_{\text{full}})/SSE_{\text{full}}$, where $SSE_{\text{full}}$ is the residual sum of squares of the model including all predictors. To retrieve normalized coefficients $a_j$, the data of each property was z-scored.

2.8.9 Statistics

Blinding and randomization of neurophysiological data were not performed. Unless stated otherwise, results are presented as mean ± standard error of the mean. KS refers to the Kruskal Wallis test. WRS refers to the Wilcoxon rank sum test, and WSR to the Wilcoxon signed rank test.

Statistics on circular data were performed using the Kuiper and the Rayleigh tests\[^{177}\]. The vector strength (VS) was considered significant if $N \cdot V S^2 > k$ at the $\alpha$ level, where $N$ was the number of phases used to compute the VS, and $k = [2.9957 4.6052 5.2983 6.9078]$ for corresponding $\alpha$ values \(0.05\ 0.01\ 0.005\ 0.001\)\[^{177}\].

Throughout this thesis, the strength of p-value is indicated by stars: *p<0.05, **p<0.01, ***p<0.001 unless stated otherwise.

All R-square values presented in chapter 3 Functional determinants of response selectivity to frequency modulated sounds are adjusted R-square values.
Figure 2.1: Multi-electrode array used for in vivo recording. Adapted from the NeuroNexus 2016 Catalog.

Figure 2.2: Apparatus positioning during in vivo electrophysiology in the auditory cortex of anaesthetised mice. 

a. The mouse was head-implanted and fixed onto a head-post. The speaker was placed in front of the animal. The auditory cortex is outlined in blue in the brain schematic. 
b. The multi-electrode array was positioned horizontally and inserted in the brain using a micromanipulator and a microscope for visual guidance.
Figure 2.3: Apparatus positioning during in vivo electrophysiology in the auditory cortex of awake mice. 

a. Atlas reference (Allen Mouse Brain Atlas) and electrode tract. 

b. The multi-electrode array was positioned at 35° from vertical and inserted in the brain using a micromanipulator and a microscope for visual guidance. 

c. Top view of the apparatus during recording. The stereotaxic frame was used to precisely tilt the micromanipulator from the vertical axis. Manipulations done above the animal’s head are hidden from the subject due to the large metal base used to fixate the animal.

Figure 2.4: Head-implant for in vivo electrophysiological recording in awake mice. 

a. 3D model of the head-implant. 

b. 3D model of the adaptor used during surgical implantation to secure the head-implant to the stereotaxic system. 

c. The head-implant was not obstructing the eyes nor the ears of the animal.
Figure 2.5: Software for auditory stimulation in passive listening conditions. The PC (top box) was communicating with the RZ6 (bottom left box) via Matlab (ActiveX Control). The RZ6 buffer was partitioned (B1 and B2 in this example). The pre-generated sound file (TXT File) was chunked into segments (S1-S4 in this example) fitting the buffer partition size. While the RZ6 was reading a buffer partition and sending the voltage to the speaker (bottom right plot), the reading index (Indx) was monitored by the PC. A stimulus segment was loaded in the RZ6 buffer partition once the Indx had reached the end of that buffer partition.

Figure 2.6: Software for auditory stimulation in active behaving conditions. The RZ6 buffer (top left; green box) was partitioned (in this example, in 3 partitions labelled B1-3; black boxes). Each partition was loaded with a stimulus (blue and grey squares; x-y axes represent time-frequency). Upon receiving the corresponding digital input (In) trigger (0-5V rise), the stimulus contained in the buffer was send to the speaker. If multiple In were received, the signal was summed online. The right column plots present a schematic of overlapping In triggers (x-y axes represent time-voltage) and the resulting sound output (x-y axes represent time-frequency).
Figure 2.7: Hardware used during behavioural training. **a.** The system was composed of a sound delivery system (blue: RZ6 and speaker; PC), and a controller system (grey: Arduino; PC; water delivery; vacuum delivery; lick detector). Black thick arrows represent several parallel I/O connections. Thin arrows represent single I/O connections. Dashed arrows represent interactions with the animal present in the setup. Mouse image adapted with permission[178]. **b.** The lick port was composed of the water delivery, vacuum delivery and lick detector (infra-red LED and sensor). **c.** When used during behavioural training, the lick port was placed below the mouse jaw, at a fair but comfortable distance so that the tongue only may break the LED beam.

Figure 2.8: User interface used to monitor mouse performance online during the FM direction discrimination task. The lick raster (left column plot) displayed lick events (black) aligned to stimulus onsets (green/red bars, representing Go/No-Go cues respectively). The yellow bar indicated a time out period. The performance was assessed using rolling $d'$ measure (middle column plot; number of trial used in rolling window = 15) and overall metrics (left column plots). For visualisation simplification, $d'$ values were plotted between 0 and 1.5.
Figure 2.9: User interface used to monitor mouse performance online during the oddball task (T2). The lick raster (left column plot) displayed lick events (red and green dots represent the first response in the case of false alarm (FA) and hit respectively; black dots represent lick events not considered as a response). The responses were aligned to trial onset. The yellow bar indicated the target window within a trial. The purple star indicated the presentation of distractor sounds. The performance was assessed using rolling $d'$ measure (middle column plot, blue; number of trial used in rolling window = 15; red: rolling probability of FA, $p(FA)$; green: rolling probability of hit, $p(hit)$) and overall metrics (left column plots). For visualisation simplification, $d'$ values were plotted between -1 and 2.

Figure 2.10: FM discrimination task design. Top row: Go (upward FM sweep) and No-Go (downward FM sweep) stimuli were presented consecutively. A period of no lick was required to present the stimulus. Second row: Upon the presentation of a stimulus, the mouse could respond by licking or not in the response window. The first lick detected in a trial was counted as the response. Third row: In the case of a response being detected upon a Go sound (Hit), the water valve was actuated, and water delivered. Fourth row: The vacuum pump was activated after each trial, regardless of the trial outcome. Bottom row: In the case of a response being detected upon a No-Go sound (False Alarm), a noise was presented, and a time out delivered. No reinforcer was presented if the mouse did not lick in the response window (Miss and Correct Rejection).
Figure 2.11: Oddball task design (T1). a. The stimulus consisted of two isochronous trains of 100ms pure tone, one with High frequency tones and 400ms stimulus onset asynchrony (SOA), and one with Low frequency tones and 617ms SOA. A target (frequency modulated sound) could occur at three positions with equal probability in a trial (the positions are marked by the colored squares). The occurrence timing and the frequency content of a target was jointly modulated, i.e. High frequency targets were presented Early, and Low frequency targets were presented Late. The target spectrotemporal bias was manipulated using a trial block design (Early or Late block). b. A block was composed of three trial types: target only (S0; 7 trials), single stream (S1; 10 trials), and dual stream (S2). Once the S2 condition was reached, mice had to perform 20-25 correct trials for the stimulus to switch to the other block type. The coloured squares indicate the response windows. c. The first lick detected in a trial was counted as the response. The mouse was required no to lick before the occurrence of a target. If the mouse responded in the target window, a drop of water was administered (Hit). If the mouse responded before (false alarm; FA), a punishment (noise and time out) was administered. No punishment was administered if the mouse licked passed the response window. After each trial, the vacuum pump was activated.

Figure 2.12: Hardware for auditory stimulation combined with electrophysiological recording in vivo in passive listening conditions. a. The hardware was composed of a sound delivery system (blue: RZ6 and speaker; PC) and an acquisition system (fuchsia: electrodes; Digital Lynx; PC). Black thick arrows represent several parallel I/O connections. Thin arrows represent single I/O connections. Dashed arrows represent interactions with the animal present in the setup. b. The Digital Lynx received square voltage pulses (named triggers) from the RZ6 (blue; y axis scaled by 0.5) as analog input, enabling alignment of stimuli (grey; y axis scaled by 0.75) with neural traces (black) during post processing. Threshold crossing (represented here by the dashed line) was applied on the trigger trace to delimit onset and offset of stimuli.
**Figure 2.13:** Hardware for auditory stimulation combined with electrophysiological recording *in vivo* in active behaving conditions.  

**a.** The system was composed of a sound delivery system (blue: RZ6 and speaker; PC), an acquisition system (fuchsia: electrodes; Digital Lynx; PC) and a controller system (grey: Arduino; PC; water delivery; vacuum delivery; lick detector). Black thick arrows represent several parallel I/O connections. Thin arrows represent single I/O connections. Dashed arrows represent interactions with the animal present in the setup.  

**b.** For each stimulus (bottom row plot), the Digital Lynx received 5V square pulses from the corresponding Arduino channel (third row plot). An event is detected when a voltage change occurs on a least one channel (second row plot, grey boxes). The event type (EV; binary format) and the timestamp (TS; time reference: computer clock; $10^6$ s format) of an event were saved (second row plot; see grey inset for a zoomed-in example). Neural traces (top row plot) were sampled, and the measured voltage and corresponding TS were saved. Spikes were detected from the sampled traces, and their time stamps were saved in sample index (IN; integer). In order to align spike times to event times, the TS of the events were matched to the sampled trace TS, and subsequently converted in sample index.

**Figure 2.14:** Example single cells sorted using Klusta. Top: mean spike waveforms from 5 cells (right) extracted from the 4 electrodes composing the tetrode (left; purple). Bottom: Auto- (colours) and cross-correlograms (white) for the corresponding cells (bin width = 1ms). The thin grey line on the auto-correlogram indicates contamination threshold.
3  Functional determinants of response selectivity to frequency modulated sounds

Sections of this chapter are subject to a joint first author publication in Nature Communications\(^{[179]}\). All results are presented with the kind consent of all authors.

3.1  Introduction

Appearances and disappearances are both salient events for sensory processing\(^{[180–183]}\), and neurons in many sensory systems exhibit robust responses to stimulus initiation and termination\(^{[50–52,174,184–186]}\). In the mature primary auditory cortex (A1), neurons respond to both the onset and offset of sound via activation of non-overlapping populations of synapses\(^{[51]}\). These synaptic inputs have discrete frequency tuning, ensuring that individual A1 neurons exhibit distinct frequency selectivity to sound onset and offset\(^{[50,51,187]}\). A1 is topographically organised, and the frequency selectivity of individual neurons is largely determined by their positions within this tonotopic gradient\(^{[188,189]}\). Given fixed positions within this map, it is surprising that neurons exhibit different ON and OFF selectivity, and the mechanistic basis and functional significance of this organisation are not understood.

In the visual cortex, it was suggested that ON/OFF receptive field (RF) arrangement may contribute towards movement direction selectivity\(^{[47]}\). An equivalent type of stimulus in the auditory modality consists of frequency modulated (FM) sounds. Whilst most models for FM direction selectivity rely on the interaction of inhibitory and excitatory inputs\(^{[9,21]}\), recent evidence suggests that the cortex may possess distinct mechanisms for FM selectivity enhancement\(^{[46]}\).

As part of a collaboration between the Chadderton and Clopath laboratories (Imperial College London), I conducted electrophysiological recordings to test for the functional significance of ON/OFF RF arrangement. The model by which ON/OFF RF arrangement may confer direction selectivity to FM sounds emerged from a previous study jointly conducted by Dr Sollini, Dr Chadderton and Dr Clopath. The main results from this latter study are outlined below as an introduction to this model, with the kind consent of all authors.

3.1.1  Mechanistic basis of ON and OFF receptive field arrangement

Individual neuron ON and OFF RFs are typically adjacent but non-overlapping with respect to frequency, indicating a high degree of specificity in ON/OFF inputs organisation. This specificity may have an anatomical substrate: for instance, ON and OFF inputs could arise from discrete, but adjacent, regions of the auditory thalamus. In this scenario, differences in ON and OFF tuning should be present at the onset of hearing (postnatal days 13-17;\(^{[190]}\)). Alternatively, pruning of synaptic inputs or changes in synaptic strength later in development could drive RF reorganisation.

By combining electrophysiological recordings and computational modelling, Sollini et al. revealed cortical mechanisms underlying the organisation of ON and OFF RFs. Specifically, they demonstrated that discrete ON/OFF frequency tuning developed following cortical exposure to sound, and that complementary Hebbian plasticity of ON and OFF inputs was sufficient for such RF reorganisation.
The computational model was then used to generate a prediction on the functional significance of ON/OFF RFs arrangement.

**ON and OFF receptive fields in primary auditory cortex diverged during development**

In order to compare ON and OFF RFs in A1 neurons from developing and adult animals, *in vivo* extracellular recordings were made from anaesthetised Young and Adult mice. In individual neurons from Young mice, ON and OFF RFs were similar, and characteristic frequency (CF) for onset and offset were commonly in agreement (see figure 3.1a for a schematic). In contrast, neurons in Adult mice exhibited segregated but adjacent aligned ON and OFF RFs (see figure 3.1b for a schematic). ON/OFF segregation was measured by calculating the absolute difference in octaves (oct) between ON and OFF CFs (|CFdiff|, marked at the bottom of figure 3.1b). Distributions of |CFdiff| were different between Young and Adult neurons, with smaller values in the Young population (see figure 3.1c). These results demonstrated reorganisation of the relative tuning of A1 neurons to sound onset and offset during development: ON and OFF RFs showed a high degree of similarity at the onset of hearing but diverged following cortical exposure to sound.

**Hebbian plasticity could account for the developmental divergence of ON and OFF receptive fields**

Using computational modelling, the authors examined how ON and OFF RFs may become segregated during development. Modelled neurons were part of a rate-based feedforward network, and received a series of excitatory and inhibitory ON and OFF inputs across a range of frequency channels (see figure 3.1d for a single-channel schematic). Importantly, in this model, both excitatory and inhibitory synaptic efficacies were plastic according to Hebbian learning: when pre- and post-synaptic neurons fired together, their weights increased, otherwise they decreased. Developmental exposure to sound was modelled by presenting neurons with sound sequences composed of fluctuating onsets and offsets (black trace in figure 3.1d). Within this scheme, summation of onset- and offset-evoked activity could occur when offsets rapidly followed onsets for transient sounds, and when onsets rapidly followed offsets during brief gaps. As seen for neurons *in vivo*, modelled exposure to sound lead to receptive field reorganisation (see figure 3.1e). It is important to highlight that whilst the model was able to replicate the increased propensity of neurons with large CF diff as seen in Adult compared to Young *in vivo*, it could not explain the decrease propensity of neurons with small CF diff observed experimentally.

The authors next explored which aspects of the model were necessary to reproduce the biological features of ON/OFF RF divergence. Notably, the authors assessed the contribution of evoked synaptic inhibition to RF divergence. Inhibitory ON and OFF inputs were removed from the model. Under this scenario, ON/OFF RF divergence still occurred (see figure 3.1f), indicating that evoked synaptic inhibition is not necessary for developmental RF reorganisation.

### 3.1.2 ON/OFF receptive field arrangement conferred direction selectivity to frequency-modulated sounds *in silico*

What is the functional significance of ON/OFF RF segregation in mature animals? It was hypothesised that the arrangement of ON and OFF RFs could influence how cortical neurons respond to
frequency-modulated (FM) sweeps (see figure 3.2a). FM sweeps will activate ON inputs as they enter a preferred frequency channel and activate OFF inputs when they exit. The rationale was that the relative arrangement of ON and OFF RFs could increase directional selectivity to FM sweeps: if a neuron is selective to higher frequency offsets than onsets, then ascending tone sweeps will exit the preferred-ON channel first and then enter the preferred-OFF channel (see figure 3.2b). Conversely, descending tone sweeps will exit the preferred-OFF channel and enter the preferred-ON channel coincidently (see figure 3.2c). In the latter case, the close temporal association of OFF and ON activation could lead to summation of ON/OFF inputs and therefore generate enhanced firing. The authors tested this prediction by presenting ascending (UP) and descending (DOWN) tone sweeps to the model neurons. The model predicted a clear relationship between the alignment of ON/OFF frequency selectivity and FM direction sensitivity (see figure 3.2d). As with developmental RF divergence, this property was preserved when evoked synaptic inhibition was removed from the model, suggesting that the tuning of excitatory ON and OFF synaptic inputs are sufficient to account for FM direction selectivity.

3.1.3 Research aims: Tests for the functional significance of ON and OFF receptive fields in vivo

Following the study of Sollini et al., my main aim was to test whether the predictions of the simulation were borne out in the in vivo characteristics of A1 neurons. By conducting recording in the auditory cortex of adult mice, I showed that ON/OFF RF arrangement was related to directional selectivity of slow, potentially ethologically relevant, frequency modulations. By taking advantage of a recent pharmacogenetic technique, I demonstrated that this property was conserved even when synaptic inhibition was reduced. Moreover, by comparing data collected from adult and young mice, I showed that direction selectivity and ON/OFF receptive field overlap jointly varied throughout development, and that direction selectivity was predictable based on neural properties in adult mice only. Finally, I trained mice in a FM direction discrimination task, and revealed that mice could discriminate slow frequency modulation direction accurately even when synaptic inhibition was reduced. Overall, these results validated the theoretical framework for segregated ON/OFF RFs significance proposed by Sollini et al.
3.2 Results

3.2.1 ON/OFF cells discriminated best the direction of slow frequency modulations

To test whether the predictions of the simulation were borne out in the in vivo characteristics of A1 neurons, I initially conducted electrophysiological recording in the A1 of anaesthetised adult mice. I measured evoked responses to ascending and descending FM sweeps at different velocities (range: ±2.2-140 octaves/second) as well as to pure tones. In a first set of experiment, 171 auditory responsive cells were recorded, out of which 30 possessed V-shaped ON and OFF RFs (see figure 3.3a,b for two example cells). 14 out of 30 cells had lower frequency ON versus OFF (ON lower than OFF), and the other 16 cells had higher frequency ON versus OFF (OFF lower than ON).

Before assessing the relationship between direction selectivity (DS) to FM sweep and ON/OFF RF arrangement, it was important to quantify the DS strength of ON/OFF cells. Indeed, these cells may be specialised in the encoding of FM sounds of a particular FM speed. In such case, the preferred FM speed should be selected when aiming to uncover any relationship between DS and RFs arrangement. The best FM speed was defined as the speed for which the maximum instantaneous firing rate was generated (see section 2.8.4 Direction and rate selectivity analysis for method details). For most ON/OFF cells, the best FM speed was 2.2 oct/s (see figure 3.3c), indicating that these cells were most strongly evoked by slow FM modulations. ON/OFF cells might thus be particularly selective for FM direction at slow FM speed. To test for this hypothesis, I computed the direction selectivity index (DSI) based on the mean number of spikes generated in response to UP and DOWN sweeps, at all FM speeds tested (see section 2.8.4 Direction and rate selectivity analysis for DSI formula). The rationale for choosing the mean number of spikes over the classically used maximum of instantaneous firing rate\[9,21\] is detailed in the next paragraph. Besides, the DSI value did not depend on the underlying metric used (see supplementary figure I.26). Based on the sign of the DSI (positive or negative) at 2.2oct/s, a cell was considered UP or DOWN selective (see figure 3.3a,b and supplementary figure I.27a,b for two example cells). In total, 10 cells preferred downward FM sweeps whilst 20 cells preferred upward FM sweeps. On average, it appeared that ON/OFF cells were indeed most selective for low (below 35 oct/s) FM speeds (see figure 3.3c).

To assess whether this selectivity trend was unique to ON/OFF cells, I counted the number of cells able to discriminate UP and DOWN sweep amongst the total cell population for each FM speed. For this analysis, all cells recorded in similar experimental conditions were used \(n = 171\) cells from original experiments in WT mice, \(n = 134\) cells from the Baseline period in pharmacogenetic experiments). The discrimination performance (DP) was computed using logistic regression, and defined as the area under the receiver operating characteristic curve (AUROC; see section 2.8.5 Discrimination performance for method details). Intuitively, the discrimination performance relied on the difference in response between the UP and DOWN sweeps. Whilst the magnitude of such difference was readily quantified by the DSI, the AUROC also depended on the reliability of the responses. Therefore, whilst the DSI was correlated with the AUROC (see figure 3.4a), a large DSI was not necessarily matched to a large AUROC value. Two metrics were used to compute the DSI and the AUROC. Firstly, the mean number of spikes, and secondly the maximum instantaneous firing rate. Compared to the latter metric, the first metric is more robust to jitter in spiking activity, however the response strength it provides may be dampened by a so-called ‘ice-berg effect’. Indeed, the window of time used to compute the DSI may encompass both evoked and non-evoked, baseline activity,
which are summed and averaged to calculate the mean firing rate. Nonetheless, the DSI based on the mean number of spike was correlated with the DSI based on the maximum firing rate at all FM speed tested (r>0.9, p < 0.001), and the DP was on average greater using the mean number of spike as measure, suggesting that the mean number of spike offered a more stable representation of the cell’s response. Only cells with high DP using both metrics (AUROC>0.65) were considered for further analysis (marked as red dots in figure 3.3a). Overall, cells with high DP were mainly found at slow FM speeds (35oct/s and lower; see top panel of figure 3.4b, red curve), suggesting that cells in A1 were best at discriminating UP and DOWN sweeps of slow FM speeds. Interestingly, the proportion of UP/DOWN selective cells depended on the FM rate (see top panel of figure 3.4b, orange and brown traces). When the same analysis was performed for the ON/OFF cells, the results recapitulated the findings presented in figure 3.3c, whereby cells discriminated best slow FM speeds, with a higher proportion of cell being UP-selective. It thus appeared that cells in the auditory cortex preferentially encoded for slow FM, and that ON/OFF cells might be particularly well-versed in encoding the slowest modulations. Moreover, cells in A1 were generally biased towards being UP-selective at most speeds tested. How could these biases for FM speed and direction emerge?

### 3.2.2 Relationships between FM direction selectivity and frequency tuning

Selectivity for the direction of FM sweeps had been previously linked to a cell’s position within the tonotopic axis, as demonstrated by the relationship between the DSI and the characteristic or best frequency (CF or BF) of the ON RF[9]. In this previous study conducted on anaesthetised rats, cells with low CF ON (<8kHz) were DOWN selective, whilst cells with high CF ON (>14kHz) were UP selective. However, in this study, the best FM rate was relatively high (~70 oct/s), contrasting with the earlier finding presented in figure 3.4. To test whether any relationship between DSI and CF ON was present for ON/OFF cells, I compared these two parameters at each FM speed (see figure 3.5a). Notably, the relationship between the DSI and CF ON varied with the FM speed, with the strongest correlations found at low FM speeds (below 35 oct/s). At low FM speeds, cells with low and high CF were more likely DOWN and UP selective respectively, whilst at higher FM speed this relationship was reversed. Since ON/OFF cells were not well-discriminating the direction of high FM speed sounds (see figure 3.3c), I also assessed whether such reversed relationship was visible in multi-unit data, a classical measure used when studying the correspondence between DS and ON RF characteristics[9] (see figure 3.5b). The multi-unit data was recorded in different animals than those where the ON/OFF cells were measured, but in similar experimental conditions (multi-unit recoded in the Baseline condition of pharmacogenetic experiments). Only sites with clear V-shape ON RFs were included in the analysis. BF was used instead of CF, as it better characterises multi-unit RF[9]. Overall, the relationship seen in the ON/OFF cells was conserved: at low FM speed, low and high BF sites were more likely DOWN and UP selective respectively. This relationship was reversed at higher speed (although note the poor interpolation R-square). It is important to highlight that these higher speeds are those for which A1 cells were worst discriminators (see figure 3.4).

Based on these findings, it appeared that the direction selectivity of A1 cells depended on the CF ON, with the strongest correlations found at low FM speeds. The reversal in the relationship between DSI and CF ON across FM speeds was not observed using the model, however it is important to highlight that the correlation between DSI and CF ON greatly depended on the FM speed used,
both in vivo and in silico, and that the correlation became poorer at fast speeds. Whilst A1 neurons may receive excitatory ON inputs that are already direction-tuned\cite{21}, strengthening mechanisms for direction selectivity have been proposed at all stages of the auditory pathway\cite{21}, as reported in other sensory modalities\cite{191}. Two major types of mechanisms have been suggested: firstly, differential delays of excitatory inputs across the spectral receptive field, and secondly spectral offset between excitatory and inhibitory inputs\cite{46}. It is important to highlight that the emergence of direction selectivity previously modelled by Sollini et al. relied mainly on the former mechanism, as the removal of inhibitory inputs did not affect direction selectivity (see figure 3.1). In that model, the delay between excitatory inputs arose from the spectral offset between ON and OFF excitatory receptive fields (see figure 3.2). Interestingly, this model contrasted with previous theories on direction selectivity shaping in A1\cite{9,192}, which argue for the latter excitatory-inhibitory mechanism. In the next sections, I explored the contribution of each mechanism to direction selectivity formation in vivo.

3.2.3 ON/OFF receptive field arrangement was a good predictor of direction selectivity to slow frequency-modulated sounds

By comparing the DSI of ON/OFF cells with the arrangement of ON and OFF RFs, I could directly assess whether the arrangement of ON/OFF RFs was a potential mechanism contributing towards direction selectivity shaping. ON/OFF had previously been reported as organised\cite{51}, so that cells with low CF ON had high CF OFF, and vice-versa. By plotting the CF ON against the difference in ON/OFF CFs (CFdiff), I revealed that such relationship prevailed for the 30 ON/OFF cells I recorded (see figure 3.6a). Noteworthy, this organisation mirrored the organisation seen between the DSI and CF ON (see figure 3.5a). The reversal point was similar when comparing the DSI and CF ON (interpolation curve intercept at 2.2oct/s = 11.85kHz; see figure 3.5a left panel), and when comparing the CF ON and CFdiff (interpolation curve intercept = 11.46kHz; see figure 3.6a). As a result, the DSI was correlated ($r = -0.61$, $p < 0.001$) with the CFdiff, as predicted in silico (see figure 3.2).

I further confirmed the contribution of ON/OFF RF alignment to direction selectivity in adult A1 using multi-variable linear regression models. Specifically, I tested for the relative contribution of different functional neuronal properties in predicting the DSI. The tested properties included the strength and variability of evoked responses to sound onset and offset, the bandwidth and overlap of ON and OFF RFs, and the octave difference between ON and OFF CF (see section 2.8.8 Modelling of experimental data for method details). The importance of each property for DSI prediction was assessed via the absolute value of their normalised coefficients, and the proportional reduction of error generated when adding the property as predictor. The performance of the model in predicting the DSI was assessed via the adjusted R-square, which considers the number of parameters used. For the population of adult A1 neurons, ON/OFF CF difference had the largest coefficient and accounted for a large proportion of the error reduction in DSI prediction (see supplementary table 5.1). By eliminating functional properties that degraded model performance, I generated a restricted model. The functional properties included in the restricted model comprised the magnitude of the ON-evoked response, the bandwidth of the OFF RF, %ON/OFF RF overlap, and ON/OFF CF difference. As seen in the full model, the normalised coefficient of ON/OFF CF difference was the largest (see supplementary table 5.2). Moreover, the other properties included in the model were poor in predicting the DSI when used in single parameter model (adjusted R-square<0.05). These results indicated that
the performance of the restricted model mainly relied on the linear relationship between the ON-OFF CF difference and the DSI, confirming that this parameter was the most important linear predictor of the DSI in adult mice.

3.2.4 Fast cortical inhibition was not required for direction selectivity

Rapid cortical inhibition has previously been implicated in the expression of FM direction selectivity\(^9\). However, the model developed by Sollini et al. suggested that the excitatory ON/OFF RF arrangement alone could contribute to FM direction selectivity. I tested this prediction by targeting one class of inhibitory interneurons, parvalbumin-positive interneurons (PV+), using a pharmacogenetic silencing approach. PV+ neurons are fast-spiking and preferentially target perisomatic regions of excitatory pyramidal cells, providing strong feedforward synaptic inhibition\(^{193}\). In auditory cortex, PV+ neurons are well-tuned for frequency and exhibit short response-latencies\(^{30}\), leading to the possibility that evoked PV+ inhibition can ‘out run’ and/or overwhelm synaptic excitation evoked by FM sweeps in non-preferred directions\(^{8,194}\) to enforce direction selectivity. I tested this proposal in A1 by selectively expressing the inhibitory DREADD receptor, hM4i, in PV+ cells. AAV-DIO-hM4i-mCherry virus was unilaterally injected into the auditory cortex of Pvalb-Cre mice to target the hM4i receptor to A1 PV+ cells (see figure 3.7a-c). After allowing sufficient time for transduction and expression (8 weeks, see section 2.6 Pharmacogenetics), I performed electrophysiological recordings to measure evoked responses in A1 to pure tones and FM sweeps (see figure 3.7d-e for electrode placement). Evoked activity was compared between an initial recording period (Baseline), and after intraperitoneal injection of the DREADD agonist, clozapine-N-oxide (CNO; 5mg/kg; see section 2.6.4 Recording paradigms for method details), when the excitability of PV+ neurons was reduced (Post-CNO; see figure 3.7f).

Following CNO injection, the magnitude of evoked local field potentials (LFPs) increased in hM4i-PV+ - but not sham control animals (see figure 3.8a-d - normalised ON response at BF post-CNO: 2.24 ± 0.14 in hM4i-PV+ animals versus 1.10 ± 0.02 in sham control animals; WSR test: p < 0.001; n = 56 and n = 24 respectively), consistent with reduced cortical inhibition in hm4i-PV+ animals only. However, the overall tuning profile of LFP FRAs were unchanged (see figure 3.8c; non-significant change in ON CF between Baseline and post-CNO: 0.04 ± 0.03 oct; WSR test: p = 0.26). Multi-unit (MU) recordings revealed that CNO injection produced a robust and significant increase in spontaneous and evoked firing rates in hM4i-PV+ but not control animals (see figure 3.8e-f; normalised spontaneous firing rate post-CNO: 3.14 ± 0.27 in hM4i-PV animals versus 1.13 ± 0.08 in sham control animals, p < 0.001; normalised evoked ON firing rate post-CNO: 2.48 ± 0.26 in hM4i-PV animals versus 1.17 ± 0.09 in sham control animals, p < 0.001 ; normalised evoked OFF firing rate post-CNO: 2.81 ± 0.29 in hM4i-PV animals versus 1.13 ± 0.11 in sham control animals, p < 0.001; WSR tests; N = 7 and N = 3 respectively). The overall structure of MU FRAs were also unchanged (figure 3.8f, non-significant change in evoked ON CF: 0.09 ± 0.05 oct, p = 0.41), indicating that PV+ inhibition plays a prominent role in regulating excitability but not sensory tuning of local A1 neurons\(^{30,194}\). I further found no significant differences in ON/OFF RF bandwidth and best frequency threshold level for single cells recorded pre- and post-CNO injection (p > 0.1 for all variables tested, WRS test; n = 30 and 29 cells pre- and post-injection).

Having confirmed the fidelity of the pharmacogenetic perturbation in hM4i-PV+ animals, I assessed
directional sensitivity in single cells following CNO injection and compared to ON/OFF RF arrangement in neurons that possessed V-shaped ON and OFF RFs (n = 29 cells in N = 7 mice; see figure 3.9a and supplementary figure I.27c for an example cell). Within this population, neurons exhibited direction selectivity (see figure 3.9b,c; n = 17 UP- and n = 12 DOWN- selective). Whilst most neurons in A1 were biased towards encoding UP sweeps Post-CNO (figure 3.9d), the ON/OFF cells conserved their characteristics. Notably, for the ON/OFF cells, the best FM speed remained low (see figure 3.9b,c and figure 3.9d), and the relationship between DS and CF ON at low FM speed persisted (see figure 3.10b). Importantly, the DSI values were independent of the metric tested (correlation between DSI based on the maximum firing rate and DSI based on the mean number of spike Post-CNO: r = 0.83, p < 0.001). The relationship between DS and CF ON was also conserved at the multi-unit level (see figure 3.10b). For ON/OFF cells, the relationship between CFon and CFdiff persisted Post-CNO (see figure 3.11a; r = -0.81, p < 0.001). As a result, the alignment of ON/OFF RFs in individual neurons remained correlated with direction selectivity at low FM speed (see figure 3.11b; r = -0.40, p < 0.05, DS measured at 2.2 oct/s). I therefore concluded that the segregated arrangement of ON/OFF RFs in A1 neurons can confer directional sensitivity to FM sounds, as predicted by the theoretical model. However, it is important to highlight that the persistence of the relationship between DSI and the arrangement of CF ON/OFF post-CNO may also suggest a subcortical origin for the generation of selectivity to FM sound direction, as previously proposed[21]. In this latter scheme, the arrangement of ON/OFF CF may contribute towards the sharpening of FM direction selectivity at the cortical level.

### 3.2.5 Direction selectivity of Young A1 neurons in vivo

If the difference between ON/OFF RF solely accounted for DS in A1, Young neurons should in theory display no DS for FM sweeps, since their ON/OFF RFs were shown to greatly overlap (see figure 3.1). Alternatively, inputs from subcortical structures may already confer DS to cortical neurons[21], which can be refined later during development. To test for this hypothesis, electrophysiological recording were conducted in the A1 of juvenile mice (P15, N = 3; recordings performed by Dr Sollini, data analysed by G. Chapuis). FM sweeps and pure tones were presented to the juvenile mice following the same procedure as in adult. 161 auditory responsive cells were recorded, out of which 20 possessed V-shaped ON and OFF RFs. The best FM rate of Young ON/OFF cells was biased towards slow (2.2 oct/s) but also higher (~17.5 oct/s) FM rates (see figure 3.12a) compared to Adult ON/OFF cells (see figure 3.3). For Young cells, DS was non-null (see figure 3.12b) and biased towards UP sweeps (see figure 3.12b,c). This bias was not predicted by the Adult model (see supplementary figure I.29). As previously reported by Sollini et al., the differences between ON/OFF CFs were small for Young cells, and the correspondence between CF ON and CFdiff was poor (see figure 3.12d). Moreover, the correspondence between DSI and CF ON was also poor (see figure I.28). The organisation of DSI compared to CFdiff visible in adult mice was thus non-existent in Young mice (see figure 3.12e). To test if other variables could account for the DS, a multi-variable linear model was created, using a similar procedure as for the Adult data (Full model, see supplementary table 5.3). In this model, %ON/OFF RF overlap was the most important predictor, however the linear relationship between this property and the DSI was poor (adjusted R-square<0.06; see supplementary table 5.3). Since the DS of A1 neurons in Young mice could not be explained by the ON/OFF arrangement, nor by the position in the tonotopic axis, nor
by any other parameters measurable in spiking activity, the mechanisms responsible for DSI bias at young age remained unclear.

3.2.6 Direction selectivity of Adult A1 neurons recorded in awake passive listening condition

As there is evidence for differences in the response pattern to stimuli in anaesthetised and awake preparations\(^\text{195,196}\), a main question is whether the coding of FM sweeps is fundamentally different in the awake animal. I thus extended the experimental techniques and conducted electrophysiological recordings and pharmacogenetic manipulations in awake passive-listening adult mice (N = 10 Pvalb-Cre, N = 3 WT used as Control; see chapter 2 Materials and Methods). As seen in anaesthetised mice, injection of CNO produced a significant increase in neural activity in hM4i-PV+ mice only, visible both at the LFP (see figure 3.13a; \(p < 0.001\) WRS test) and multi-unit (see figure 3.13b; \(p < 0.001\) WRS test) levels. ON/OFF cells were recorded in Baseline (n = 14 cells in 6 penetrations in WT mice) and Post-CNO (n = 10 cells in 8 (4th-11th) penetrations in hM4i-PV+ mice) conditions. The best FM speed was low for the majority of these cells (see figure 3.13c), consistent with the previous finding in anaesthetised mice. Most of ON/OFF cells were UP-selective (see figure 3.13d) in both Baseline and Post-CNO conditions. The arrangement between CF ON and CF OFF was also conserved (see figure 3.13e; correlation coefficient \(r = -0.83\), \(p < 0.001\) in Baseline). However, the relationship between DSI and CFdiff was different (see figure 3.13f; correlation coefficient \(r = 0.22\), \(p > 0.05\) in Baseline), potentially due to a difference in the organisation between CF ON and DSI, as shown at the multi-unit level (see supplementary figure I.30).

3.2.7 Mice were able to discriminate the direction of slow FM sweeps under reduced cortical inhibition

To further show that PV-mediated cortical inhibition was not required for the discrimination of slow FM direction, pharmacogenetic manipulations were carried out in mice trained to discriminate FM sweep direction. Specifically, mice were trained to report the occurrence of UP FM sweep only amongst UP and DOWN sweeps (Go/No-Go task design, see figure 3.14a). Both UP and DOWN sweeps had FM of 4.4 oct/s (see section 2.3.3 Stimulus properties for stimulus details), and were presented with equal probability (see section 2.4.2 Paradigms for training protocols). Mice reported their responses via licking actions, which were optically detected (see section 2.4.1 Setup for behavioural setup). Four outcomes were possible in this task, depending on the combination of sound presented and behavioural response: Hit, Miss, Correction Rejection and False Alarm (see figure 3.14b). The behavioural performance was measured using the sensitivity index \(d'\) (see section 2.4.3 Measure of performance for equation). A \(d'\) value over 1 indicated acceptable performance, and \(d'\) value over 2 indicated excellent performance. On average, mice reached \(d' = 1\) after 15 training sessions, and \(d' = 2\) after 26 sessions (N = 10 mice, see figure 3.14c). Following the 26th session, the performance was stable. Over sessions 26-40, mice did on average 153 ± 22 Hit trials per session (N = 10 mice; mean ± std).

Following 2 days of high performance (Baseline), perturbation experiments were conducted to assess the effect of cortical inhibition reduction on FM discrimination performance (figure 3.14d). Mice were injected IP with CNO to activate the DREADD (see section 2.6.4 Recording paradigms for
method details; all mice were injected with DREADD bilaterally in A1 at least 8 weeks prior to the first perturbation), following which their behavioural performance was assessed (see figure 3.14e for example behavioural sessions). Overall, mice maintained a high performance in Post-CNO condition (see figure 3.14f, mean $d' > 2$; sessions chosen as the first Post-CNO only session experienced by the animal, within the 25-40 session number window). This result was maintained when using the data from all Post-CNO sessions (presented in supplementary figure I.32; $d'$ Post-CNO = 1.81 ± 0.12, mean ± std). Electrophysiological recordings were performed in A1 to verify for the DREADD activation. Since the Baseline and Post-CNO recordings were unpaired during behaviour, it was important to determine a marker for DREADD activation. In awake passive-listening mice, spikes in LFP power were clearly visible Post-CNO (see supplementary figure I.33) in hM4i-PV+ mice only. These spikes were most detectable after filtering the LFP signal in the gamma band (30-100Hz; spikes detection threshold fixed at -75$\mu$V). In behaving mice, spikes were detected in all recorded LFP traces Post-CNO (n = 5), and none were detected in Baseline condition (n = 14) (see figure 3.14g for example LFP traces recorded in behaviour). During Post-CNO sessions with recording, the $d'$ remained high (1.66 ± 0.25, mean ± std; see supplementary figure I.32), as well as the number of Hit trials (101 ± 10, mean ± std). These results demonstrated that mice were able to perform the discrimination task under reduced cortical inhibition.

What code may underlie the ability of mice to discriminate FM sweeps? As previously discussed, the direction selectivity of A1 cells depended on the CF ON. A simple place-code may thus be sufficient for mice to distinguish between UP and DOWN FM sweeps. In order to test this hypothesis, the tuning of multi-unit activity was determined (measured as the ON best frequency, BF; see figure 3.15a for an example), and compared against direction selectivity (see figure 3.15b for an example). In behaving mice, the BF was correlated with DS ($r = -0.43$, p < 0.001; see figure 3.15c). Such organisation was also visible Post-CNO. The relationship between BF ON and DSI was enhanced when using the beginning portion of the sound to compute the DSI (see figure 3.15d, $r = -0.60$, p < 0.001), suggesting that the start frequency difference between UP and DOWN sweeps may be sufficient for discrimination. However, multi-unit activity did not significantly differ between correct and incorrect behavioural choices (p > 0.05, WSR test, comparing the mean number of spikes in correct and incorrect trials for all sites; see supplementary figure I.33), suggesting that the activity in A1 during the FM discrimination task mainly represented the stimulus features.

Intact activity at the level of A1 may however not be required to perform in a discrimination task relying on large frequency difference (i.e. in the range of an octave)\[58\]. To test such hypothesis, one must abolish cortical signalling. In an effort to suppress neural activity in A1, two different perturbations were undertaken. First, Pvalb-Cre mice were injected in A1 bilaterally with the excitatory DREADD (AAV-DIO-hM3-mCherry) so as to increase the activity of PV+ cells. Such a method had been previously shown to shut down cortical signalling carried by principal cells\[197\], although optogenetic rather than pharmacogenetic activation technique had been used. The hM3 DREADD approach was first tested as it was experimentally similar to the procedure using hM4i DREADD, thus constituting a suitable control. In hM3 mice, no drop in performance was seen Post-CNO (average $d' > 2$, N = 4 mice; see supplementary figure I.34). It was however difficult to confirm whether PV+ cell activation diminished the activity of principal cells, as electrophysiological recordings mainly captured the activation of fast-spiking interneurons. Therefore, a second silencing approach was undertaken.
Specifically, muscimol was injected intracortically and bilaterally in A1 40mn prior to behavioural assessment (N = 2 mice; see section 2.6.3 Auditory cortex inactivation for method details). In these mice, performance was maintained despite muscimol injection (see supplementary figure 1.34).
3.3 Discussion

3.3.1 ON/OFF segregation as a specialised mechanism for enhancing direction selectivity to slow FM speeds

Processing of sound onset and offset by the auditory system is particularly important for perception\cite{181–183}. The segregated arrangement of ON and OFF RFs within individual neurons allows these two different signals to be encoded independently by the same A1 population. In this study, I aimed at understanding the functional significance of ON/OFF RFs arrangement. The results from modelling and in vivo electrophysiology experiments indicated that interactions between ON and OFF inputs contribute to the selectivity of single neurons to higher-order features, particularly to the direction of frequency modulation (FM). FM is a common feature of natural sounds, and is particularly relevant in vocalisation\cite{198}, including speech\cite{199,200}.

Several mechanisms have been put forward to account for cortical FM directional selectivity, the most prominent of which require the involvement of cortical inhibition\cite{9,21,192}. In A1, inhibition rapidly follows excitation upon sound presentation\cite{9,13,31}, and the strength of these synaptic inputs is approximately balanced over a wide range of sound parameters\cite{9,13,31}. The tuning curve of these inputs are however skewed (as seen by the ‘inhibitory sidebands’ surrounding excitatory regions measured in ‘Two Tones’ paradigms\cite{9,11,12}), providing a mechanism by which inhibitory-excitatory interaction can influence DS in A1\cite{9}. However, DS has also been observed when GABAergic inhibition is pharmacologically blocked\cite{192}.

Here, I used pharmacogenetics to suppress inhibition provided by fast spiking cortical interneurons. These experiments revealed that CNO injection in control animals does not alter neural activity in A1 (see figure 3.8g,h and figure 3.13a,b), ruling out non-specific effects of metabolically converted CNO (i.e. clozapine acting upon other receptors in the brain) influencing these results\cite{44}. In hM4i-PV+ mice only, CNO produced increased neural excitability in A1, consistent with suppression of cortical inhibition. In these animals, the results show that specifically reducing the excitability of fast spiking PV+ interneurons does not abolish direction selectivity in the auditory cortex (see figure 3.9 and figure 3.13). This result was consistent with the role of PV+ cells in controlling cortical responses as seen in other sensory modality\cite{201}.

The former results do not rule out that alternative sources of cortical inhibition, such as somatostatin-expressing interneurons\cite{150,202}, may play a direct role in the expression of FM direction selectivity. However, it is suggested that while inhibitory side bands may play an essential role in the formation of directional selectivity at subcortical processing stations\cite{21}, and at rapid speeds\cite{9}, this mechanism may not be applicable at slower speeds in the auditory cortex\cite{48,49}. The majority of neurons in mouse A1 show preferences for slower speeds of FM\cite{37–39} (see figure 3.4), which is a common characteristic of mouse vocalisations\cite{198}. In mice, ON/OFF RFs are spectrally segregated, and are conveyed by distinct synapses onto AC neurons\cite{51}, thus comprising all attributes necessary to engender DS to FM sounds. In adult mice, I indeed found that the arrangement of ON/OFF RFs was correlated with DS to slow FM sweeps, in a manner that was consistent with the model’s prediction. It is thus suggested that the interplay of ON and OFF excitatory input channels can account for DS at slow FM speed in mouse AC.
3.3.2 ON/OFF arrangement relates to functional topography in Adult and Young mice

In this study, I investigated how the representation of FM sound by A1 neurons varies through development, by comparing the DS of cells in Adult and Young mice, and further relating it to ON/OFF RF properties. Most Adult neurons with measurable ON and OFF RFs displayed segregated ON/OFF RFs, and directional selectivity at the slowest FM speeds (see figure 3.3 and figure 3.13). In contrast, Young ON/OFF neurons had similar ON/OFF CFs, and were selective at slow but also faster FM speeds. Moreover, their selectivity was biased towards UP FM (see figure 3.12), as previously reported[38]. Interestingly, this bias towards UP and faster FM was recapitulated in Adult neurons following the reduction of PV-mediated cortical inhibition (see figure 3.9), suggesting that excitatory inputs to A1 neurons already convey direction selectivity at high FM speed[9,21]. Accordingly, the DS of single cell and multi-unit was dependent on the CF ON, the latter property being conveyed to A1 neurons by subcortical structures[17].

Whilst previous studies investigating ON/OFF responses in A1 and subcortical structures mainly observed neurons with higher CF OFF than ON (i.e. positive CFdiff)[51,54], a reversal in CFdiff was clearly visible along the tonotopic axis in the present study. Notably, the reversal frequency for ON/OFF CFs I measured (~11.5 kHz) matched well with the reversal frequency for DSI, and was consistent with those found in other studies when measuring reversal for tuning curve skewness (~12 kHz) or DSI along the tonotopic axis[9,203]. This relationship suggests a high degree of arrangement in A1 functional topography. Interestingly, recent studies in the primary visual cortex (V1) showed that the organisation of a multitude of properties (such as direction or orientation selectivity) relates to the organisation of ON/OFF RFs, which in V1 originates from the clustering of ON and OFF thalamic afferents[174,184]. It is thus possible that besides DS, the organisation of ON/OFF RFs apposition in A1 further relates to other sensory map topography.

3.3.3 Delineating the dependence of FM encoding on cortical state

Whilst the effect of suppressing PV+ cells was robust in both awake and anaesthetised animals, further experiments are required to comprehensively capture the differences in DS observed between awake and anaesthetised states. In awake mice, DS was correlated with CF ON, but following a different organisation than seen in anaesthetised mice. Anaesthetics are known to affect cortical states, influencing neural properties[101,103] such as ON and OFF evoked responses[53,204] or frequency tuning[205,206], and impacting on spike-sorting[207]. On the other hand, previous studies[196,208] showed that sensory cortical neurons recorded in anaesthesia present greater similarities of firing patterns with neurons under behavioural condition than those in awake passive condition. It is thus possible that the differences observed between anaesthetised and awake passive-listening preparations are due to drastic changes in physiological state, such as arousal, which were uncontrolled for in this study. The use of further experimental procedures, e.g. recording pupil dilatation[132,209], might shed light on these differences.

Besides, the use of two different electrode penetration techniques (i.e. perpendicular and parallel to cortical layers in anaesthetised and awake mice respectively) could be a source of potential confound. Indeed, any relationship with DS and CF ON and OFF may vary according to the electrode location, since the precision of A1 topography may vary according to the source and layer of the...
mapping signal\textsuperscript{[17]}. In future experiments, histological confirmation of all electrode sites should be performed, associated with staining of the cortical tissue to determine cortical layers boundaries.

### 3.3.4 The role of auditory cortex in the perception of FM sounds

Further results in awake behaving mice demonstrated that the suppression of PV-mediated inhibition does not abolish the encoding of FM sweeps in A1. In awake behaving mice, cortical inhibition was not required to perform FM direction discrimination (see figure 3.14). In those mice, FM sweeps selectivity was observed at the multi-unit level despite inhibition of PV+ cells. Activation of the DREADD by CNO was validated by recording spikes in LFP activity during behaviour Post-CNO. Whilst this pharmacological perturbation was successful, it is acknowledged that further controls are needed to define the role of A1 processing in this task. Causal relationship between neural activity and behaviour has previously been established using silencing or lesioning approaches\textsuperscript{[59,81,210]}. In this study, I used two different techniques to silence A1 activity. Firstly, I activated PV+ cells specifically in A1 using the excitatory DREADD (hM3), with the aim to silence principal cells as seen during optogenetic activation of PV+ cells in cortex\textsuperscript{[72,197]}. This perturbation did not result in a loss of performance, however it was difficult to measure its impact on principal cell activity. As enhanced drive on PV+ cells has been reported as beneficial to cortical coding\textsuperscript{[194,211,212]}, the effect of hM3 activation on behaviour should be carefully considered. Second, I performed intracortical injection of muscimol directly in A1. This perturbation did not result in a loss of performance. However, it was not possible to confirm the drug’s effect, as extracellular recordings were not conducted in A1 following muscimol application. Whilst muscimol has been reported to act for several hours\textsuperscript{[60]}, members of the Luthi Lab (CH), Hauser Lab (UK) and Renart Lab (P) reported in discussions that behavioural performance in sensory discrimination task (based on touch or sound) was affected by intracortical injection of muscimol only during the first hour following application. In published literature, A1 inactivation via muscimol has also been reported as having mixed effect on performance in a discrimination task\textsuperscript{[61]}. It is thus possible that muscimol did not act on A1 as expected.

Whilst lesion studies have shown that the auditory cortex is important for the performance of FM direction discrimination\textsuperscript{[59,210]}, animals with lesions were also able -although to a lesser extent- to discriminate between FM directions. However, it was suggested that animals with lesions discriminated FM mainly based on the sound starting frequency. In the FM discrimination task I designed, starting frequencies of UP and DOWN sweeps were an octave apart. It is thus possible that mice relied mainly on the starting frequencies to guide their decision, in which case a simple place-code may be sufficient to discriminate UP versus DOWN FM sweeps. Controls for this task aspect can be easily implemented in future research, as presented in the Future work section below.

### 3.3.5 Concluding remark

Despite reduction of PV-mediated inhibition, ON/OFF cells remained particularly selective at low speed (see figure 3.9). Therefore, the apposition of ON and OFF RFs during the functional maturation of A1 may enforce direction selectivity to slow FM sounds in a manner that is consistent with the model predictions. Such a contribution for ON/OFF RFs in auditory cortical direction selectivity may be analogous to their role underlying selectivity in primary visual cortex\textsuperscript{[174,184]}.
3.4 Future work

This study raised four fundamental questions. Firstly, does ON/OFF segregation accounts for slow FM selectivity at lower processing stations than A1, or does this property emerge in A1? In an attempt to answer to this question, I recorded extracellular activity in the Medial Geniculate Body of awake passive-listening mice (MGB; see supplementary figure I.36). As initial classification, MGB cells were labelled as displaying excitatory or inhibitory ON responses (see supplementary figure I.36a,b for example cells; note that some cells exhibited both types of responses, but one type was always clearly dominant as measured by RF size). In this pilot data, cells were mainly UP-selective (see supplementary figure I.36c,d), and cells with excitatory ON RF were best at discriminating slow but mainly extremely fast FM speed, contrasting with A1 neurons. Intriguingly, most OFF responses observed were inhibitory, regardless of the ON response type (2 cells with Exc-ON/Exc-OFF; 3 cells with Inh-ON/Exc-OFF; 3 cells with Exc-ON/Inh-OFF; 5 cells with Inh-ON/Inh-OFF). As the pathway responsible for OFF encoding might be spatially segregated at the level of the MGB\textsuperscript{[20,54,55]}, further experiments combining recording and histological confirmation of electrode placement are necessary to delineate the importance of ON/OFF RFs at subcortical levels.

Secondly, is OFF RF tuning spatially organised in mouse auditory cortex, as seen in other modality and species\textsuperscript{[174,184]}? Can such organisation be further related to other sound processing aspects (e.g. localisation), and/or to specific cell types? Multi-scale Calcium (Ca\textsuperscript{2+}) imaging has brought substantial insight on the spatial organisation of feature selectivity in cortex\textsuperscript{[174,184,213]}. I thus aimed to record Ca\textsuperscript{2+} transients from the auditory cortex of awake mice using wide field and two-photon Ca\textsuperscript{2+} imaging, so as to bridge macro- and micro-scale signals. Mice were injected with AAV encoding for GCamp6s in A1, and stereotaxically implanted with a glass coverslip over A1 (see supplementary figure I.37a). Pilot data recorded in anaesthetised mice (see supplementary figure I.37b-c) suggest that the surgical procedures and recording approach were stable, as functional mapping of ON receptive field was achieved in several animals at macro-scale. Moreover, two-photon Ca\textsuperscript{2+} imaging successfully recovered the activity of identified PV+, SOM+ and putative pyramidal cells in layer 2/3 of selected zones of auditory cortex (data not shown). Whilst this technique is promising, further experiments are required to fully map ON/OFF RFs in awake auditory cortex. This pilot data was collected in the laboratory of Prof Hausser, UCL.

Thirdly, can the contribution of other inhibitory cell types account for direction selectivity at high FM speeds? Similar experimental procedures as used in this study can be repeated using mice from a different transgenic line (e.g. SOM-Cre, VIP-Cre).

Finally, is activity in A1 necessary to the perception of FM direction? It is possible that the FM discrimination task design I used was such that processing at the level of A1 was not required for performance. Several enriching changes can be made to the current behavioural protocol, such as modifying the FM speed of start frequency (see supplementary figure I.38), which will enable to better determine what role A1 plays in supporting FM perception. As A1 inactivation via muscimol has been reported as having mixed effect on behaviour\textsuperscript{[61]}, other silencing protocols may be used, using e.g. optogenetic approaches.

Besides, it would be interesting to further demonstrate whether the model developed in this study accurately mimics in vivo computations, e.g. by using sustained stimuli which would engender greater adaptation and synaptic depression than the stimuli used in this study. Moreover, the relationship
between DSI and CF ON needs to be formally assessed at different speeds *in silico*. In such scheme, it would be particularly interesting to use a stimulation involving invariant and oddball sounds (e.g. by presenting repeatedly one sweep stimulus with a slow FM speed, and suddenly presenting another sweep stimulus with a faster FM speed, keeping the FM direction constant), as it would shed light on synaptic computations underlying invariance and irregularities detection.
Figure 3.1: Hebbian plasticity could underlie developmental divergence of ON and OFF receptive fields in auditory cortex. a-b. Schematic of ON and OFF receptive fields for Young (blue) and Adult (grey) A1 neurons. In Young neurons, ON/OFF RFs were in close proximity, whilst they were segregated in Adult. CF indicates characteristic frequency. c. Distribution of the absolute difference between ON and OFF CF (CFdiff) for A1 neurons recorded in vivo in Young (blue) and Adult (grey) mice. d. Model schematic of synaptic input to investigate influence of sound exposure on ON/OFF RFs. Top. Neurons received sound-evoked synaptic activity, driven by a sound input sequence (black) that randomly switched ON (green dashed lines) and OFF (red dashed lines) in separate frequency channels. Middle: Excitatory (Ex; upward) and inhibitory (In; downward) ON inputs were evoked by sound onsets. Bottom, Excitatory (Ex; upward) and inhibitory (In; downward) OFF inputs evoked by sound offsets. e. Distribution of CFdiff in modelled Young and Adult neurons. f. CFdiff did not vary for Adult neurons in excitation-inhibition (grey) and excitation-only (Ex only, orange) models. This figure was adapted from a study by Sollini et al. and is displayed with the kind consent of all authors.
Figure 3.2: Proposed functional significance of segregated ON/OFF RFs. 

**a.** Schematic of frequency-modulated (FM) sweep inputs. Modelled adult neurons are presented with sweeps either going UP (orange) or DOWN (brown).

**b-c.** Schematics of mechanism underlying sweep selectivity. 

- **b.** ON RF (green ellipse) is selective for lower frequencies than OFF RF (red ellipse): when a FM sweep is going UP (black arrow), ON (green) and OFF (red) inputs are activated at different times. The sum of the inputs (black) is only just above the neuronal threshold (dashed black line).

- **c.** when a FM sweep is going DOWN, OFF and ON inputs are activated in rapid succession, so they summate far above neuronal threshold.

**d.** Sweep selectivity as a function of the difference between OFF and ON for neurons modelled with excitatory and inhibitory inputs (black dots), and excitatory-only inputs (orange dots).

This figure was adapted from a study by Sollini et al. [179] and is displayed with the kind consent of all authors.
Figure 3.3: ON/OFF cells discriminated best the direction of slow FM sweeps. a. Response of a DOWN-selective cell to 3-octave FM sweeps. **Left panel:** Raster plot of spiking responses to different velocities. Orange and brown ticks indicate beginning and end of UP and DOWN sweeps respectively. **Top right panel:** Outline ON and OFF FRAs for the same cell. **Bottom right panel:** Mean spiking responses to sweeps at different speeds for UP and DOWN directions. b. Same as a, but for an UP-selective cell. c. Distribution of best FM rate for all ON/OFF cells. d. Direction selectivity index (DSI) at different speeds for the UP- (orange) and DOWN- (brown) selective cells recorded in adult A1. Positive DSI indicates selectivity for UP FM sweeps. Only cells with measurable ON and OFF receptive fields were included. In these cells, direction selectivity was mostly observed at slower speeds (below 35 oct/s).
Figure 3.4: A1 cells discriminated best the direction of slow FM sweeps. a. DSI versus discrimination performance (AUROC), calculated using the number of spikes as metric, at different FM speeds (marked on top of each panel). The cells in red are those with high DP (AUROC>0.65), and passing this criterion with all metric tested (i.e. number of spike and maximum firing rate). b. Percentage of cells passing criterion for FM direction discrimination at each FM speed (red; top panel: all cells, n = 305; bottom panel: ON/OFF cells, n=30). The orange and brown curves represent percentage of cells passing criterion that are UP- and DOWN-selective respectively.
Figure 3.5: The relationship between DSI and CF ON depended on the FM speed. 

**a**. DSI versus CF ON for ON/OFF cells at each FM speed (marked on top of the panel). Each dot represents a cell (n = 30). The red line represents a linear interpolation, from which the adjusted R-square is displayed on top of the panel ($R^2$). The Pearson correlation coefficient ($r$) between DSI and CF ON is marked on top of the panel, with stars indicating the significance level.

**b**. Same as in **a**, but for the DSI based on multi-unit activity versus multi-unit BF ON. Each dot represents a recording site (n = 35).

---

**Figure 3.5** reflects the dependency of DSI on CF ON at varying FM speeds for both ON/OFF cells and multi-unit activity. The relationship is quantified through linear interpolations and Pearson correlation coefficients, with adjusted R-squares and significance levels indicated.
Figure 3.6: ON/OFF arrangement was a good predictor of direction selectivity at slow FM speed. 

a. Relationship between CF ON and ON/OFF receptive field arrangement in adult A1. Octave spacing between the characteristic frequency of ON and OFF receptive field (CFdiff) is indicated on the bottom axis. Positive values indicate that OFF CF is higher than ON. Each dot represents a cell (n = 30). The red line represents a linear interpolation, from which the adjusted R-square is displayed on top of the Panel ($R^2$). The Pearson correlation coefficient ($r$) between CF ON and CFdiff is marked on top of the panel, with stars indicating the significance level (p<0.001). The grey background marks the zone of possible values when plotting CF ON versus CFdiff. 

b. Relationship between direction selectivity and ON/OFF receptive field arrangement in adult A1 for the cells presented in a. DSI values represented are calculated at the lowest speed (2.2 oct/s). In adult A1, the relationship between directional selectivity and ON/OFF receptive field arrangement corresponded as predicted in silico.
Figure 3.7: Details and validations of pharmacogenetic experimental procedure. a. AAV1/2-hSyn-DIO-hM4Di-mCherry was unilaterally injected into the auditory cortex of Pvalb-IRES-Cre mice. b. Post mortem microscopy confirmed the presence of mCherry expressing neurons (red) in auditory cortex. This image is the zoomed-in view of the inset in a. c. Specific infection of cortical parvalbumin-positive (PV+) interneurons was confirmed via co-expression (merged image, bottom) of mCherry (red, top) and PV-antibody (green, middle). d. Post mortem microscopy confirmed the placement of the electrode in A1. The electrode was coated with Di-I, and inserted horizontally into the auditory cortex. e. Zoomed-in view of the inset in d, displaying the electrode tract. All black scale bars in a-e represent 1000 µm. f. The experimental protocol consisted of recording neural responses to sounds during a Baseline condition (typically 30 min), followed by IP injection of the agonist CNO. The CNO was left to act for 20 min during which spontaneous neural activity was recorded, following which the neural activity was recorded in Post-CNO condition using the same sounds as during the Baseline condition.
Figure 3.8: Neural activity increased Post-CNO. a. Tone-evoked local field potential (LFP) recording from A1 at Baseline (black, top) and Post-CNO (orange, bottom). Horizontal grey bar indicates period of tone presentation. Black vertical lines indicate onset and offset. For presentation purposes, LFP responses are grouped by tone intensity (ascending from 10 to 80 dB SPL). Within each intensity group, responses are ordered by tone frequency (ascending from 7-56 kHz). Arrows indicate intensity/frequency combination shown in b. b. Tone-evoked LFP at Baseline (black) and Post-CNO (orange) for intensity/frequency combination indicated in a (average of 5 repeats). Arrows indicate peak ON-evoked LFP amplitude in each condition. c. On FRA at Baseline (top) and Post-CNO (bottom) for LFP recording shown in a,b. Overall RF structure is maintained but note overall increase in response magnitude post-CNO (calibration bar, right). d. Normalised change in ON-evoked LFP amplitude at BF following CNO injection for recordings from all sham- (Control, blue; N = 3) and AAV-DIO-hM4i-injected- (hM4i, orange; N = 7) mice. CNO injection is associated with a prominent increase in tone-evoked signals in A1 in hM4i-PV+ mice. e. Tone-evoked multi-unit (MU) activity from a single tetrode at Baseline (top) and Post-CNO (bottom). Display same as in a. f. ON and OFF FRAs at Baseline (top) and Post-CNO (bottom) for MU recording shown in a. Overall RF structure is maintained but note overall increase in MU firing rate following CNO injection. g-h. Normalised change in spontaneous (c) and evoked ON (d) MU firing rate Post-CNO for recordings from all sham- (Control, blue; N = 3) and hM4i-PV+ (hM4i, orange; N = 7) mice. Coloured open circles correspond to individual tetrodes (8 per recording). Filled black circles correspond to individual animal means. CNO injection is associated with elevated firing rates in hM4i-PV+ mice.
Figure 3.9: ON/OFF cells conserved their characteristics Post-CNO. a. Response of a DOWN-selective cell to 3-octave FM sweeps. This cell was recorded Post-CNO. **Left panel:** Raster plot of spiking responses to different velocities. Orange and brown ticks indicate beginning and end of UP and DOWN sweeps respectively. **Top right panel:** Outline ON and OFF FRAs for the same cell. **Bottom right panel:** Mean spiking responses to the sweeps at different speeds for UP and DOWN directions. b. Distribution of best FM rate for all ON/OFF cells. c. Direction selectivity index (DSI) at different speeds for the UP- (orange) and DOWN- (brown) selective cells recorded in adult A1 Post-CNO. Positive DSI indicates selectivity for UP FM sweeps. Only cells with measureable ON and OFF receptive fields were included. In these cells, direction selectivity was mostly observed at slower speeds (below 35 oct/s). d. Percentage of cells Post-CNO passing criterion for FM direction discrimination at each FM speed (red; top panel: all cells, n = 316; bottom panel: ON/OFF cells, n = 29). The orange and brown curves represent percentage of cells passing criterion that are UP- and DOWN-selective respectively. Note that A1 cells were mostly UP-selective Post-CNO.
Figure 3.10: The relationship between DSI and CF ON was conserved Post-CNO. Each dot represents a recording site (n = 35).

DSI and CF ON is marked on top of the panel, with stars indicating the significance level (p). Same as in a, but for the DSI based on multi-unit activity versus multi-unit BF ON. Each dot represents a cell. The red line represents a linear interpolation, from which the adjusted R-square is displayed on top of the panel (R²). The Pearson correlation coefficient (r) between DSI and CF ON is marked on top of the panel. Each dot represents a recording site (n = 35).

b. Same as in a, but for the DSI based on multi-unit activity versus multi-unit BF ON. Each dot represents a recording site (n = 35).

c. Same as in a, but for the DSI based on single-unit activity versus single-unit BF ON. Each dot represents a recording site (n = 35).
Figure 3.11: Direction selectivity remained correlated with ON-OFF arrangement Post-CNO. a. Relationship between CF ON and ON/OFF receptive field arrangement in adult A1 Post-CNO. Octave spacing between the characteristic frequency of ON and OFF receptive field (CFdiff) is indicated on the bottom axis. Positive values indicate that OFF CF is higher than ON. Each dot represents a cell (n = 29). The red line represents a linear interpolation, from which the adjusted R-square is displayed on top of the panel ($R^2 = 0.649$, $r = -0.81$ ***). The Pearson correlation coefficient ($r$) between CF ON and CFdiff is marked on top of the panel, with stars indicating the significance level ($p<0.001$). The grey background marks the zone of possible values when plotting CF ON versus CFdiff.

b. Relationship between direction selectivity and ON/OFF receptive field arrangement in adult A1 Post-CNO for the cells presented in a. Directional selectivity is correlated with the relative arrangement of ON and OFF receptive fields following pharmacogenetic inhibition of PV+ interneurons. Each point represents an individual cell (n = 29). DSI values represented are calculated at the lowest speed (2.2 oct/s).
Figure 3.12: Direction selectivity in Young ON/OFF A1 neurons. a. Distribution of best FM speed for all ON/OFF cells recorded in Young mice (n = 20). b. Direction selectivity index (DSI) at different FM speed for the UP-selective (orange) and DOWN-selective (brown) ON/OFF cells recorded in Young A1. Positive DSI indicates selectivity for UP FM sweeps. In these cells, direction selectivity was biased towards UP. c. Percentage of Young cells passing criterion for FM direction discrimination at each FM speed (red; left panel: all cells, n = 161; right panel: ON/OFF cells only, n = 20). The orange and brown curves represent percentage of cells passing criterion that are UP- and DOWN-selective respectively. d. Relationship between CF ON and ON/OFF receptive field arrangement in Young A1. Octave spacing between the characteristic frequency of ON and OFF receptive field (CFdiff) is indicated on the bottom axis. Positive values indicate that OFF CF is higher than ON. Each dot represents a cell (n = 20). The red line represents a linear interpolation, from which the adjusted R-square is displayed inside the panel ($R^2$). The Pearson correlation coefficient (r) between CF ON and CFdiff is marked inside the panel, with stars indicating the significance level (p<0.01). The grey background marks the zone of possible values when plotting CF ON versus CFdiff. b. Relationship between direction selectivity and ON/OFF receptive field arrangement in Young A1. DSI values represented are calculated at the lowest speed (2.2 oct/s). 

90
Figure 3.13: Direction selectivity of ON/OFF A1 cells in awake adult mice under pharmacogenetic manipulation.

a. Normalised change in evoked LFP amplitude Post-CNO for recordings from all sham- (Control, blue; N = 4) and hM4i-PV+ (hM4i, orange; N = 10) mice. Each column represents a recording (up to two recordings were performed in the same animal). Coloured open circles correspond to individual tetrodes (8 per recording). Filled black circles correspond to individual animal means. CNO injection is associated with elevated firing rates in hM4i-PV+ mice.

b. Same as in a, but for evoked ON MU firing rate.

c. Best FM rate for the ON/OFF cell population in Baseline (n = 14; recorded in Control mice) and Post-CNO (n = 10; recorded in hM4i-PV+ mice).

d. Direction selectivity index (DSI) at different FM speed for the UP-selective (orange) and DOWN-selective (brown) ON/OFF cells recorded in Baseline (n = 14) and Post-CNO (n = 10). Positive DSI indicates selectivity for UP sweeps.

e. Relationship between CF ON and CFdiff for ON/OFF cells in Baseline (blue) and Post-CNO (red). DSI values are calculated at the lowest speed (2.2 oct/s).
Figure 3.14: Perception of FM direction was preserved following reduction of cortical inhibition. a. Schematic of FM discrimination task paradigm. Per trial, one FM sweep was presented (10-21kHz), either UP (orange) or DOWN (brown). The mouse had to emit a Go (lick) response upon the presentation of the UP FM sweep, and to not respond (No-Go) upon the presentation of the DOWN FM sweep. b. Description of the four possible trial outcomes: Hit, Miss, False Alarm (FA) and Correct Rejection (CR). If a mouse responded upon the presentation of the UP sweep (Hit), a water drop was delivered. If a mouse responded upon the presentation of an DOWN sweep (False Alarm; FA), a noise and a time out were delivered. If the mouse did not respond (Miss or Correction Rejection, CR), nothing was delivered and the trial ended. c. Mice reached high performance in the FM discrimination task. Performance was measured as the $d'$ value over single session. The black curves indicate the mean ± std $d'$ across all mice (N = 10 PV-Cre mice bilaterally injected with hM4i). After 25 training sessions, mice reached high and stable performance on average (yellow box). Once the performance was stable, perturbation of cortical functions could occur. d. Following baseline sessions (i.e. without perturbation), three types of manipulations could occur: IP injection of CNO (Post-CNO), electrophysiological recording (Rec.), or both combined (Post-CNO + Rec.). The bottom panel displays the training curve of a mouse (ID: M122), with each manipulation color-coded. e. Behavioural performance during example sessions with and without CNO injection. Performance was measured as rolling $d'$ (n = 30 trials in rolling window). f. On average, mice maintained high performance Post-CNO. The yellow dots indicate the sessions presented in e. g. Validation of DREADD activation by detection of spikes in filtered (30-100 Hz) LFP trace. Example LFP traces (grey) in Baseline and Post-CNO condition recorded in behaviour. In Post-CNO, spikes were detected (red arrows, threshold -75µV symbolised by the horizontal line).
Figure 3.15: Multi-unit direction selectivity was correlated with BF ON during behaviour. 

a. ON/OFF tuning curves from an example multi-unit (MU) recording site. The BF ON is measured as the maximum of the ON tuning curve. The grey area indicates the frequency range covered by the FM sweeps. 

b. Responses of the MU site presented in a to UP and DOWN FM sweeps. The grey area indicates sound presentation. Orange and brown curves represent PSTHs in response to UP and DOWN sweep respectively. This site was DOWN selective. Only responses in Miss and correct rejection (CR) trials are presented. 

c. Relationship between direction selectivity index (DSI) and BF. The DSI was measured using either the maximum firing rate (Max FR; top panel) or the total number of spikes (Nspk; bottom panel) in the time window marked on top of the panel (0-250ms, encompassing the whole sound presentation). Positive DSI indicates selectivity for UP sweeps. The Pearson correlation coefficient between DSI and BF is marked inside each panel, with stars indicating the confidence level. Filled red dots indicate measurements during Post-CNO. 

d. Same as in c, but the window was truncated to 0-100ms.
4 Anatomical circuits underlying task-related modulation of spectrotemporal receptive fields

Sections of this chapter are subject to a first author publication in Frontiers Neuroscience[214] (article under review). All results are presented with the kind consent of all authors.

4.1 Introduction

In an acoustic scene comprising multiple auditory sources, humans and animals can readily identify and track a relevant sound source amongst the background noise, and switch from tracking that source to follow another[83,69]. Such ease however belies the complexity of the underlying processes: the auditory system must parse sounds from the various sources into discrete perceptual objects (a process known as ‘auditory stream segregation’[67]), and must further filter out irrelevant inputs to best support behavioural needs[25,63,69,137]. Switching attention from one source to another has been shown to modulate sensory signal representation at various levels of the auditory pathway, notably in the auditory cortex (AC)[83,96,109,115,215]. However, little is known about the underlying mechanisms. This is probably due to the large number of anatomical pathways potentially involved in modulating the activity in AC during attentional processes. For example, the activity in AC could be modulated via the prefrontal cortex (PFC)[115], which may either form a direct connection with the AC[146], or an indirect one via the targeting of sub-cortical structure such as the thalamic reticular nucleus (TRN)[123,137]. To understand the contribution of each pathway, it is necessary to develop experimental paradigms enabling their specific manipulation.

The mouse is an increasingly popular model for studying the neural mechanisms of perception and action because of the range of molecular tools enabling precise manipulation of neural circuitry. However, no behavioural paradigm currently exists in mice to study the effect of selective attention on auditory streaming. Current paradigms of selective attention in mice[115,123,136] mostly rely on briefly presented stimuli, which do not allow for auditory stream formation. In contrast, behavioural tasks are far more extensively developed in primates[95,109], owing to their greater capacity to learn complex rules. Translating behavioural task from primates to mice would enable an exquisite insight on the neural basis underlying sensory processing and perception.

Here, I present novel behavioural tasks inspired from primate models[109], that can be used to assess the effect of attention on auditory streaming in head-fixed mice. This latter preparation enables rigorous control over behavioural state, and to acutely measure electrophysiological signals in the brain[216]. By recording in the auditory cortex and thalamus of mice during ongoing behaviour, I reveal how sound processing is modulated by attentional states during these tasks. Taking advantage of recent genetic technologies, I further aim to unravel the anatomical basis of these modulations.
4.2 Results

Three types of experiments were carried out to evaluate the effect of attention on auditory stream- 
ing. Firstly, I conducted behavioural testing to assess whether mice were sensitive to spectrotemporal 
biases present in acoustic stimuli (i.e. whether their expectation could be manipulated). Guided atten-
tion was inferred from the level of behavioural expectation. Secondly, I performed electrophysiological 
recordings to measure the activity of single units in auditory cortex and thalamus of mice performing 
spectrotemporal expectation tasks. Single units in prefrontal cortex were also recovered, but their 
activity is not discussed as part of this thesis. These measurements allowed the characterisation 
of expectation-driven modulation of sensory processing. Finally, I used newly developed genetic 
technologies to assess whether these modulation originated from the interaction between thalamic 
reticular nucleus and auditory thalamus.

4.2.1 Mice were sensitive to spectrotemporal cues

Design of a new auditory task involving spectrotemporal expectation

A Go/No-Go auditory oddball paradigm was designed after a related study in primates\cite{109}, in which the influence of stimulus spectrotemporal bias on mouse behaviour could be assessed (see figure 4.1a). The oddball paradigm consisted of two simultaneously-presented isochronous sequences of pure tones (termed streams), one with high frequency tones and a rapid rhythm (labelled High), and one with low frequency tones and a slower rhythm (labelled Low). The rhythm period is called stimulus onset asynchrony (SOA) throughout this text. The SOA was 400ms for the High stream, and 617ms for the Low stream. Two rising frequency-modulated (FM) sounds were used as deviants, each starting at a stream frequency. These FM sounds represented the targets to detect. Only one 
target was presented per trial. Expectation of the spectral content of a target (High or Low frequency) 
was manipulated using a trial block design: in the first five trials of a block, only the target of a given 
stream was presented (target only condition; S0), then the single stream associated with that target 
was also presented (single stream condition; S1), and finally both streams were simultaneously 
presented (dual stream condition; S2) (see figure 4.1b). It is important to highlight that in the dual 
stream (S2) condition, the stimuli presented before a target were identical in both block types (i.e. 
both High and Low streams were presented), as displayed in figure 4.1a. Temporal expectation was 
manipulated conjointly, by presenting High frequency targets Early ($\sim$2s) and Low frequency targets 
Late ($\sim$6s). The target alone (S0) and single stream (S1) conditions were intended as cues indicating 
upcoming target spectral and temporal biases in the dual stream (S2) condition.

The water-restricted mice were head-fixed onto an apparatus where a lick port was mounted. The 
lick port comprised a water delivery and removal setup, and an infrared optical beam-break system to 
detect licks (see section 2.4.1 Setup for setup details). On a given trial, the animal’s response was 
declared as the first lick event detected in the trial (see figure 4.1c). If the response occurred outside 
of a target window, the response was considered incorrect (false alarm, FA). If it occurred within the 
target window, it was considered correct (Hit). The performance of the animal was measured using 
the sensitivity index $d'$ (see section 2.4.3 Measure of performance for equation). All mice reached the 
criterion of $d'>1$ for both block types before the electrophysiological recording, as displayed in figure 
4.1d. The lower performance in the Late block was due to the mice generating more false alarms 
in that block, rather than failing to respond to Low targets (the mean probability of FA was 32% in
Early block and 40% in Late block). Water-restricted mice were generally expected to show a high lick propensity as classical conditioned response\(^1\), and thus, indicate higher FA rates when required withhold a lick for a longer period. Therefore, the mice were considered equally trained on both block types despite the difference in \(d'\) values.

**Using false alarms to reveal spectrotemporal expectation in mice**

Having developed an auditory task where the spectrotemporal properties of the target sound were biased across trial blocks, I sought to examine whether mice were sensitive to these biases and if they adjusted their behaviour accordingly. Such change in behavioural strategy could not be assessed via correct responses (i.e. during Hit trials) alone, since a perfectly biased animal could conceptually respond equally efficiently during the Early or Late block. I thus assessed behavioural strategy change across trial blocks via the other response type, i.e. the FA. An example training session comprising Hit and FA responses is presented in figure 4.2.

Since target sound spectral and temporal biases were conjointly manipulated, it was important to disentangle the effects of either bias on behaviour. The animal's sensitivity to target temporal bias (i.e. Early or Late) was determined by comparing the median FA time (calculated from trial start) and the proportion of FA trials in the Early Zone (EZ; 0-2.4s from trial start) in each block type. These metrics were computed per session, using FA in the single stream (S1) and dual stream (S2) conditions only. As displayed in figure 4.2b, FA in the Late block were significantly delayed compared to those in the Early block (\(p < 0.001\), WSR test, \(N = 25\) sessions in 4 mice, only FA in S1 and S2 were used), indicating that animals were sensitive to target temporal bias. In other words, mice expected targets in the Early block to occur earlier than those in the Late block.

Furthermore, in the Early block, FA were markedly timed to the tone onset preceding the first target: the tone occurred at 1200ms from trial start, and the median FA time in the Early block was 1358±57ms (median±median absolute deviation, see figure 4.2b). Conceptually, an animal may try to respond upon the presentation of a sound whose properties are closest to the expected target, rather than at any time point during a trial. To investigate whether FA not only reflected the target temporal bias but also related to the target spectral bias, I compared FA timing patterns in the target alone (S0) and single stream (S1) conditions, expecting FA to be timed to tone presentation in the single stream (S1) condition, and randomly generated in the target alone (S0) condition. When displaying the responses in FA trials only, FA in the single stream (S1) condition appeared aligned to the stream tone onsets (see figure 4.3a for a mouse example session), contrasting with the lack of arrangement for the FA in the target alone (S0) condition. To quantify this alignment, I estimated reaction time (RT) distributions of FA measured from tone onset (example RT distributions are displayed in figure 4.3b for the example session of figure 4.3a). The distribution in the single stream (S1) condition presented a peak, whereas the distribution in the target alone condition (S0) did not. In order to assess peak strength and significance, I used the Vector Strength (VS) metric and associated Rayleigh statistic (see section 2.8.9 Statistics for equations). Briefly, the VS measures the coherence of the data points phase angle, with a value of 1 indicating perfect coherence (all data points have the same phase), and a value of 0 indicating no coherence. Here, the phase angle was measured as the RT reported to the stream SOA (i.e. the period of time between two tone onsets), converted into radian. For all mice, the VS was lower in the target alone (S0) condition compared with the single stream (S1) condition.
Moreover, significant VS was only found in the single stream (S1) condition (see figure 4.3c, starred distributions), indicating that mice did not use the trial onset to time their response with such specificity. This result confirmed that FA were not randomly generated, but instead timed to stream tone onsets. Furthermore, it provided a metric by which FA can be used to determine preferred response timing in the dual stream (S2) condition.

In the dual stream (S2) condition, the target was spectrally biased across Early and Late block, with High frequency targets presented Early and Low frequency targets presented Late. If a mouse were sensitive to the target spectral bias, one could imagine that it would selectively attempt to respond to the tones of the corresponding stream rather than to any tone presented, the former strategy increasing the chance of responding in the correct time window. If a mouse timed its response to the cued stream, the RT distribution for FA would be peaked when measuring the RT from the High tone onsets in the Early block, and vice versa for the Low tone onsets and the Late block. Correspondingly, the RT distribution measured from the Low tone onsets in the Early block should appear random (with a certain degree of non-randomness due to the residual beat between High and Low tone trains), and vice-versa for High tone onsets and Late block. Whilst this was not strikingly observed (see figure 4.4), it is important to highlight that the High and Low tones preceding the first Early target had near-overlapping onsets (High tone occurring at 1200ms from trial onset, and Low tone occurring at 1234ms). Given that mice timed their FA in the Early block mainly to the sound occurring at 1200ms (see figure 4.2), peaks in distribution were observed when reporting the FA in the Early block to both Low and High tone onsets. As displayed in figure 4.3d, the statistical difference between RT distributions indicated that one animal appeared sensitive to the target spectral bias (ID 2; p < 0.001 KS test). This means this animal timed its FA to the tones in the stream likely to contain the target. The other three animals timed their FA to tones of a specific frequency regardless of the block type.

The discovered FA timing patterns were consistent with mice aiming at maximising the chance of obtaining a reward when producing a behavioural response. Indeed, FA response timing changed across block types, so as to occur in close vicinity of a likely target. Although it was not possible to define the behavioural strategy of the animals during Hit trials, it seemed reasonable to assume that they would use a similar strategy as the one observed during FA trials. Mice were thus considered biased during Hit trials as revealed by their FA timing. Moreover, the fact that the false alarms were timed to the onset of the sounds present in the cued stream suggests that the animals could selectively follow the oddball stream despite the presence of a distractor stream.

4.2.2 Spectrotemporal expectation modulated neural responses in auditory cortex

Previously, I developed an auditory task in which mice displayed a change in behavioural response reflecting the target spectrotemporal biases. Specifically, mice timed their FA based on the likelihood of target occurrence (i.e. Early or Late; see figure 4.2), and in some cases also based on the likelihood of target spectral content (i.e. High or Low; see figure 4.4). These results indicated that mice expected the target to be presented at a specific time point, and within a specific frequency stream. Such an optimisation of behavioural strategy might be accompanied by an optimisation of sensory processing, promoting target detection. In support of this theory, previous studies demonstrated that both expectation of the timing and expectation of the spectral content of a target modulated sensory
processing to facilitate target representation\cite{61,65,109}. However, in these studies, the effects on neural processing due to temporal or spectral expectations were rather different. Temporal expectation did not modify the neural response tuning profile, but changed the neural response strength in a dynamic manner so that the response would be maximal near the expected target occurrence\cite{61}. In contrast, spectral expectation changed the neural response tuning so as to better represent the expected target frequency content\cite{65}. In natural scenarios, as well as in the dual-oddball task I designed, temporal and spectral expectations may be exploited simultaneously. In this case, it is unclear whether the previously observed effects of spectral and temporal expectations would combine, and enhance specific frequency responses at specific time points. By conducting electrophysiological recordings in mice during behaviour in the dual-oddball task, my first aim was to delineate how was sound processing modulated in a scenario combining temporal and spectral expectations.

Neural activity at the level of auditory cortex was shown to be correlated with performance in tasks requiring target detection in sequences of pure tones\cite{60,61,65,106,109}. Such relationship was found at both single cell\cite{61,65} and local field potential (LFP)\cite{60,61,106} level. To investigate this further, electrophysiological recordings were made in mouse auditory cortex during the oddball task using tetrodes, enabling the recovery of single cell and LFP signals. Recordings were performed in all animals, of which two (ID 2,4) completed sufficient amount of trials (>40) in the dual stream (S2) condition for both block types. Mouse ID 3 completed sufficient amount of trials in the S2 Late block only. Mouse ID 1 did not reach performance during recording sessions. Histological confirmation of electrode placement in mice ID 2,3,4 is presented in supplementary figure I.39 and supplementary figure I.40.

**Spiking activity was modulated by task specificity**

To investigate whether spectrotemporal expectation modulated spiking activity, I focused the analysis on the activity recorded during the Early time zone (EZ; 0-1.6s from trial onset) occurring before the first Early target onset in the dual stream (S2) condition. Since the stimuli in the dual stream (S2) Early and Late blocks were identical during the EZ, any difference in stimulus processing between Early and Late blocks could be attributed to top-down influences reflecting expectations. To quantify this difference, I used the PSTHs of spiking activity during the EZ in Early and Late blocks to generate a block type discrimination performance across time (see figure 4.5a,b top and second panels for example spiking activity and corresponding PSTHs). The discrimination performance was computed using logistic regression, and defined as the area under the receiver operating characteristic curve (AUROC; see section 2.8.5 *Discrimination performance* for method details). The discrimination performance relied on the difference in firing rates between the Early and Late blocks at each time point during the EZ (see figure 4.5a,b, third and fourth panels), with large and reliable differences generating large AUROC values. Multi-unit spiking activity was used in this analysis, as it encompassed all spiking information. In the EZ, spiking activity was indeed modulated by task specificity, as assessed by AUROC values (see figure 4.5c). Differences between Early and Late blocks occurred in both stimulus-locked evoked activity, and ongoing non-stimulus locked activity. The examples in figure 4.5a,b were illustrate these two types of modulation: whilst both sites were tuned to the High frequency sound, one site (a) displayed high AUROC values when spiking activity was driven by sound presentation, whereas the other site (b) displayed high AUROC values outside
of the sound presentation window (in this case during the silent portions preceding sound onset). Importantly, the sites in figure 4.5a and b were recorded from different animals, strengthening the belief that modulations in activity indeed emerged during task performance. Nonetheless, these results were at first surprising, as modulation of spiking response in auditory cortex had been reported as being specific to sound-evoked activity during a temporal expectation task[61]. However, such modulation was also shown to occur in silent periods prior to target onset in an attention task requiring behavioural responses biased to stimulus feature[119]. The results presented in figure 4.5 are thus consistent with the idea that auditory cortex activity reflects task-related features beyond solely the acoustic properties of the input[82,115].

Having shown that spiking activity in the EZ was different between Early and Late blocks, I characterised the underlying changes in firing rate responsible for this difference. In a previous study[61], temporal expectation provoked an increase in spiking activity prior to the likely occurrence of a target. This increase was restricted to evoked responses, and was observed in all auditory-responsive cells irrespectively of their frequency tuning. However, in this current task, the frequency content of the target could not be predicted. Contrastingly, in another study[65] where the target frequency content was predictable, changes in firing rate occurred in a frequency specific manner so as to enhance solely the representation of the target, rather than any irrelevant sounds. Based on these earlier findings, I speculated that in my study, the firing rate would increase in the EZ during the Early trials, as the presentation of a target was imminent at the end of the EZ (temporal expectation). Moreover, I hypothesised that this increase would be most prominent in cells tuned to the High frequency sound, as the target occurring Early is of High frequency (spectral expectation).

To test these hypotheses, each cell was assigned to a frequency-preference group (High or Low) based on its response to tones in the single stream (S1) condition. All cells recorded significantly responded (see section 2.8.3 Receptive field analysis for determination of significance threshold) to a single frequency, either to the High frequency (n = 10 cells) or to the Low frequency (n = 7 cells) tones only. Once the cell's tuning had been characterised, its responses in Early and Late blocks were compared (see figure 4.6a,b for the responses of two example cells). Modulations in firing rate fluctuated over time (see figure 4.6a,b bottom panels), thus I used the sum of the mean PSTH to assess firing rate change in the EZ. The figure 4.6d presents the modulation of the summed activity in the Early compared to Late block for the different cell types, with values above 0 indicating an increase in firing rate. Both High and Low cell populations displayed an increase in firing rate, as assessed by the average modulation index above 0. This increase was significant at the single cell (n = 5 cells, p < 0.05, WSR test) and population (p < 0.01, WSR test) levels.

To further confirm whether this increase in activity was not dependent on the cell's tuning, I measured the firing rate modulation of cells which were not significantly tuned to either Low or High tones, but which exhibited a high discrimination performance between Early and Late trials (AUROC>0.6, using the sum of the PSTH to assess firing rate change in the EZ). These cells were labelled as Other, see figure 4.6c for an example cell. Whilst these Other cells were not significantly responsive to High and Low tones, they might still be responsive to other sounds. As displayed in figure 4.6d, the average modulation index for the Other cell population was also above 0, indicative of an increase in firing rate in Early compared to Late trials. Most of these Other cells had a significant modulation (p<0.05, WRS test), which can be accounted for by the selection criterion based on high
To confirm whether the increase in firing rate was localised to the time prior to target occurrence, I measured the firing rate modulation of cells during Late trials at two different time intervals, one near trial onset and one near Late targets. These intervals were defined as early (0–1.6s from trial onset) and late (4–5.6s from trial onset), as is illustrated in figure 4.7a. Importantly, both time zones encompassed a similar amount of Low and High tones (although organised in a different pattern relative to each other), and had similar duration. This allowed for a direct comparison of the mean firing rate during these two periods. Since High and Low cells were mostly responsive to a single tone, the difference in tone presentation pattern was considered negligible when comparing the firing rates between the intervals. The figure 4.7a presents an example cell (classed as Other) displaying an increase in firing rate over the course of Late trials. This effect was not an artefact, as analysis of the autocorrelogram and spike waveforms presented in figure 4.7b confirmed. The minimal contamination in the refractory period argues against electrical noise generating the increase in activity visible in figure 4.7a. This observation was highly consistent across cells, as most cells displayed an increase in firing rate during the period close to the Late target (see figure 4.7c). This effect was significant both at the single cell (p < 0.05, WSR test) and population (p < 0.01, WSR) levels. However, the Other cells were selected based on their ability to discriminate late and early zones (AUROC>0.6), and might thus constitute a different population that the one presented in figure 4.6d.

Modulations in firing rate across trial types did not enhance sensory representation

These results were in agreement with temporal expectation generating an increase in the firing rate localised to the time prior to a likely target. However, it was surprising to observe an increase in Low and Other cells activity during Early trials, as these cells were not necessarily relevant in encoding the upcoming Early target sound. It thus seemed that the increase in firing rate observed between Early and Late trials might not be related to sound encoding per se, but might instead reflect behavioural response preparation or reward anticipation[116,118,218]. If this were the case, the increase in firing rate might manifest in both the spontaneous and evoked activity, which in turn might affect the contrast of sensory signals representation.

To address these points, I quantified the changes in spontaneous and evoked activity for High and Low cells between Early and Late trials (see figure 4.8). Spontaneous and evoked activity were measured in the 100ms windows preceding and following the preferred sound onset respectively (these windows were labelled as Evoked and Baseline windows, see figure 4.8a for a visual example). As shown in figure 4.8b, changes in firing rate (although not significant) were visible in both the Baseline and Evoked windows. This was consistent with the earlier finding showing that activity patterns in Early and Late trials could differ during both silence and sound presentation (see figure 4.5). To quantify how the modulation strength in Baseline and Evoked windows modulated the contrast of sound representation, I measured the signal to noise ratio (SNR) between Baseline and Evoked responses in Early and Late trials. I defined the SNR as the normalised absolute difference in firing rate between the Evoked and Baseline window (see figure 4.8c top panel). On average, the SNR was lower in Early compared to Late trials (see figure 4.8c bottom panel), indicating that the modulations in firing rate between Early and Late trials decreased signal contrast as defined by the SNR measure.

This previous finding indicated that the increase in firing rate observed preceding a likely target.
was not involved in the enhancement of sensory processing. Instead, this modulation might reflect response preparation or reward anticipation, and might thus occur prior to any response, be it a Hit or FA. To test this hypothesis, I assessed whether cells with a high discrimination performance between Early and Late Hit trials could accurately classify the FA response types (see figure 4.9 for a visual example). Throughout this section, the activity during the Early and Late Hit trials is labelled as training data, whilst the activity during the Early and Late FA trials is labelled as test data. Early FA were defined as the FA responses made in the 1.2-2.8s time window post trial onset in the Early S2 block, and Late FA were defined as the FA responses made in the 4s-7s time window post trial onset in the Late S2 block. For all trials, the activity considered was the average firing rate in the 0-1.2s time window post trial onset. Firstly, a cell was classified as informative if its discrimination performance (i.e. the AUROC) was above 0.6 when comparing the activity in Early and Late S2 Hit trials (training data; see figure 4.9a). Then, a logistic regression classifier was trained on the training data (removal of 5% of trials for cross-validation). The accuracy of this classification typically matched the AUROC value. Then, the activity during each FA trial (test data) was presented to the classifier (see figure 4.9b) and classification accuracy calculated. Since the number of Early and Late FA was generally unequal, it was important to build a baseline distribution for the classification accuracy. For example, a cell classifying any activity as originating from the class Early could obtain 90% classification accuracy if 9 out of 10 trials were truly Early. To generate a baseline classification accuracy distribution, the order of the true labels was shuffled, and these newly generated labels were compared against the predicted labels (process repeated 500 times). Out of the cells with high classification accuracy for Hit trials (n = 30), 9 cells were found to classify FA trial types above chance (p < 0.05; n = 1/3 for High cells; n = 1/3 for Low cells; n = 7/24 for Other cells). On average, classification accuracy was 66.5±2.2% for those cells, whilst the mean of the baseline distribution was 51.7±1.4%. These results confirmed that the spiking activity of a subpopulation of cortical neurons mirrored the behavioural response specificity.

Given these findings, it appeared that the prime difference in activity between Early and Late trials was not related to sensory processing per se, but rather reflected other aspects of the ongoing behavioural state. This was surprising, as the activity in auditory cortex is traditionally thought of as dedicated to the processing of auditory signals. Since the observed modulation was inconsistent with an enhancement in sensory processing, the role of sensory signalling at the level of the auditory cortex in guiding sensory-driven behaviour remained unclear.

**Spiking activity was modulated in a dynamic manner within trial to enhance stimulus representation**

The modulations in activity presented in the previous section appeared to diminish sensory signal contrast. Could sensory signals be nonetheless enhanced when critically important, i.e. in the vicinity of a likely target onset? As pointed out earlier, the modulations in firing rate between Early and Late blocks fluctuated over time. This indicated that within the time course of a trial, the firing rate of a cell varied in spite of the same sounds being repeatedly presented. If the dynamics of these variations differed between spontaneous and evoked activity, a change in signal contrast would be visible across time. This would represent a different type of modulation than the one previously uncovered between Early and Late trials.
To test this hypothesis, I computed the firing rate variability in spontaneous and evoked activity of cells classified as responsive to High or Low tones. This analysis was restricted to responsive cells since the key interest of this study was to assess changes in sensory processing. Specifically, I computed the mean firing rate for single cells at each sound presentation, in both the Evoked and preceding Baseline window (100ms windows as defined in the previous section, see figure 4.10a for an example). The cell’s firing rate at each sound presentation was normalised to the firing rate measured at the first sound presentation in a trial. Only cells with non-zero spontaneous activity prior to each sound, and non-null evoked activity at the first sound were included in the analysis. Cells from all experiments (Mouse ID 2,3,4) were included. On average, evoked firing rate increased over time, whilst baseline firing rate decreased (see figure 4.10b for the Low cell population). Comparing the modulations in baseline and evoked firing rate, an increase in signal to noise emerged in the Early compared to Late block at the last sound presented in the EZ (see figure 4.10c,d). Whilst the change in activity at the last sound was not significant between Early and Late block (p > 0.05, WRS test; p > 0.05 WSR using paired cells only), it was surprising to observe the same effect in both High and Low cell populations. The trend of this result was consistent with the effects observed in a previous study [61], where modulations prior to target onset resulted in a sharpening of tuning, which was maximal near target likely occurrence. Accordingly, both High and Low cell populations displayed a higher SNR at the Last sound compared to the First sound (p < 0.05, WSR test), see figure 4.10e,f. On average, the SNR was higher for Late trials compared to Early trials, consistent with the previous results (see figure 4.8).

Having shown that expectation of target timing modulated neural firing in a dynamic manner, I aimed to determine whether such modulation served behaviour. Enhanced sensory processing prior to target occurrence may result in faster reaction time. To test this hypothesis, I compared the amplitude of the firing rate during trials where the animal’s reaction time was either Fast (lower 20% of the reaction time distribution) or Slow (upper 20% of the reaction time distribution). Only Hit trials during the Early dual-stream (S2) condition were considered. The analysis was restricted to the activity of High and Low cells during the Last sound window (defined as 400ms prior to the first Early target; see figure 4.10a,b for visual examples), since the modulation of firing was shown to be maximal during that period (see figure 4.10). On average, an increase in firing rate was observed in faster trials (figure 4.10c). Whilst I did not observe significantly modulated cells, this could be due to the small amount of trials (~10 trials) used to generate the Slow and Fast response distributions. Nonetheless, this modulation was significant at the population level (p < 0.05 WSR, comparing average firing rate in Slow and Fast trials), indicating that most cells increased their firing rate prior to a Fast response.

Whilst it has been shown that neural activity modulation prior to target onset is indicative of subsequent behaviour [61], how such modulation may benefit behaviour remained unclear. Changes in firing pattern before the expected target occurrence may improve the speed and accuracy of decoding a target relatively to the preceding pure tones, subsequently enabling faster reaction time. To test this hypothesis, I compared neural responses to target and last preceding tone during Hit trials where the animal’s reaction time was either Fast or Slow. Only cells tuned to the High frequency tone or target were considered in this analysis, as only High frequency targets occurred at the end of the Early zone. The figure 4.12a displays an example cell's response to tone and target in Fast (top panels) and Slow (bottom panels) trials respectively. Slow and Fast trials were only taken from the
dual stream (S2) condition. The discrimination performance (DP) (as quantified by the AUROC, see section 2.8.5 Discrimination performance for method details) was computed for tone and target in Slow and Fast trials for each cell responsive to either tone or target in the single stream (S1) condition. The figure 4.12b presents the DP between tone and target across time in Fast (red) and Slow (blue) trials for the example cell of figure 4.12a. In this example, the maximum DP value is reached earlier in Fast compared to Slow trials, and is larger in Fast than in Slow trials. However, the delay in maximum DP between Fast and Slow trials is an order of magnitude smaller than the actual reaction time (see figure 4.12c). In order to quantify whether changes in maximum DP value or timing were significant, I computed a baseline AUROC distribution for each High tuned cell using the remaining mid-RT trials (bootstrap n = 200; using the same number of trials as for the Fast and Slow trials). If a cell's maximum DP value or timing in Slow or Fast trials was above the 95% percentile of the baseline distribution, it was considered significantly higher than baseline (p < 0.05, marked as a blue dot in figure 4.12d,e). Similarly, if a cell's maximum DP value or timing in Slow or Fast trials was below the 5% percentile of the baseline distribution, it was considered significantly lower than baseline (p < 0.05, marked as a red dot in figure 4.12d,e). For most informative cells (average max AUROC>0.65; n = 5/8), the timing of maximum DP was significantly different from baseline in either Fast or Slow trials, so that target and tone were discriminated earlier in Fast trials compared to Slow trials (see figure 4.12d). The AUROC value appeared stable for most units tested, as displayed in figure 4.12e.

These findings indicated that the quality of sensory signal representation at the level of auditory cortex was related to behavioural outcome. Notably, target decoding dynamics mirrored the speed of reaction time during correct responses. Signalling at auditory cortex level thus appeared to correlate with sensory-driven behaviour in two different ways. Firstly, the broad strength of neural activity mirrored behavioural response specificity (as presented in the previous section). Secondly, fine enhancements in auditory input representation improved behavioural response. How do these responses emerge remained to be elucidated.

**Behavioural context modulated LFP phase to facilitate upcoming target representation**

Having shown that spiking activity was modulated by behavioural context in mice, an underlying mechanism for such modulation remained to be determined. In auditory cortex, the strength of neural firing is related to the phase of low-frequency oscillations present in the LFP\(^{[108–111]}\). Specifically, responses to incoming stimuli can be amplified or suppressed if aligned with the LFP phase reflecting peak or trough of neural excitability respectively\(^{[110,111]}\). Such a regulatory mechanism for sensory coding gain is believed to have a profound influence on sensory perception\(^{[95]}\). Accordingly, behavioural performance in target detection tasks has been shown to correlate not only with spiking activity\(^{[61,65,132]}\), but also with the phase of slow-frequency oscillations present in cortical signals\(^{[95,106,109,112]}\). However, most of the latter studies were performed in primates or humans, using signals reflecting broad network dynamics (such as LFP, current source density [CSD] or electroencephalogram [EEG]). It is yet unclear to what extent the LFP signal in mice resembles that of primates or humans, and whether it correlates in a similar fashion with behavioural performance. The following section of this study was thus dedicated to the characterisation of LFP phase modulations in various sensory scenarios, with the aim to differentiate stimulus-driven (bottom-up) from behavioural context-driven (top-down) modulations.
In primates, slow-frequency LFP oscillations were shown to be modulated by rhythmical stimuli in two distinct manners\[^{108,109}\]. Firstly, the LFP phase depended on the spectral content of the tones used to generate the rhythmical sequence. Specifically, sequences of tones at the recording site’s best frequency (BF) entrained the LFP so that the phase at tone onset became facilitatory, and vice-versa for non-BF tone sequences. In this study, a non-BF stimulus was defined as a stimulus eliciting a significantly inhibited response\[^{108}\]. Secondly, after sequence termination, the LFP was found to remain significantly entrained to the stimulation rhythm without stimulus presentation. This phenomenon was observed at most stimulation frequencies tested and at long time scales post-sequence termination\[^{109}\], arguing against a simple offset or rebound response.

To test whether these phenomena were present in mouse auditory cortex, I conducted in vivo electrophysiological experiments in anaesthetised and awake behaving mice. These two different states were deliberately chosen in order to differentiate bottom-up from top-down modulations. LFP signals (0.5-300Hz) were recorded from auditory cortex, and low-pass filtered offline to mainly encompass the stimulation frequency (see section **2.8.7 Local field potential and multi-unit analysis** for method details). The phase of the filtered LFP was computed using the Hilbert transform.

To characterise whether the LFP phase depended on the spectral content of stimuli, I presented isochronous sequences of five pure tones (duration 50 ms; SOA 150-225-300 ms) of various frequency and intensity to anaesthetised mice (see figure 4.13a). LFP and phase receptive fields were generated at each tone position, using the LFP and phase measured 20 ms after tone onset respectively (see figure 4.13b). An LFP response was classified as excitatory or inhibitory if its magnitude was higher or lower than during spontaneous activity (as assessed by a negative or positive z-scores). No significant inhibition (z-scores>1.5; see section **2.8.7 Local field potential and multi-unit analysis**) was observed upon tone presentation in LFP responses.

Non-BF stimuli were defined as those generating an LFP response decrease at the fifth tone onset (named adaptation; defined as LFP response \(z\)-score>-2). This response adaptation was not below spontaneous activity, as required when defining inhibitory responses, but was instead much weaker than excitatory evoked responses. Response adaptation was thus not considered to be a proxy for inhibitory response, but rather a way to compare changes in LFP response due to differences in the strength of excitatory drive. The figure 4.13b presents the LFP and phase receptive fields of the site presented in figure 4.13a at each tone onset, with the black square inset marking the adaptation zone of the receptive field at the fifth tone onset. The example site in figure 4.13a,b displays a phase-reset in the adaptation zone of the receptive field, i.e. the phase changed from being positive upon the first stimulus presentation, to negative upon the last presentation. This phenomenon was only observed in 2/20 animals, and was thus not visible on average (see figure 4.13c; this result was found for all SOA tested). Nonetheless, phase coherence remained significant at the last stimulus presentation for all SOA tested (as assessed by the Vector Strength, \(p < 0.001\); for the example presented in figure 4.13c, VS at the fifth tone was 0.83 and 0.53 for non-adapted and adapted responses respectively). Furthermore, the phase distribution of adapted responses at the last stimulus presentation significantly differed from the phase distribution of non-evoked responses (\(p < 0.001\), Kuiper test), which had a non-significant VS (\(p > 0.001\); since the VS p-value depends on the number of samples in the distribution, only the 0.1% percentile is considered as significance
Similarly, in awake behaving animals, LFP phase-reset was only moderately dependent on spectral content (see figure 4.14a,b for two example sites with different frequency tuning recorded simultaneously from one behaving animal). These sites presented a mild inhibitory response (i.e. an upward LFP deflection) upon non-preferred tone presentation. To study fluctuations present in the ongoing (baseline) LFP signal only, the LFP traces in the single stream (S1) condition were firstly aligned to tone onset and averaged (see figure 4.14a,b top panels). Then, solely the LFP trace outside of the evoked window (0-200 ms post tone onset) was considered. The baseline portion of the trace was used to generate an interpolated curve, whose phase was measured (see figure 4.14a,b bottom panels). The average onset responses for High and Low tones were used to determine a site's frequency selectivity index (FSI). Whilst recording sites presented widely different FSI, the shift in baseline oscillation phase was rather moderate (see figure 4.14c), contrasting with the primate model (see figure 4.14d). Overall, no relationship between frequency preference and the phase of LFP slow-oscillations was found (see figure 4.14e; only sites with interpolation R^2 > 0.85 were included).

Having established that the LFP phase only moderately depended on spectral content in mice, I investigated whether the second characteristic found in primates (LFP entrainment) was observable in mice. In anaesthetised mice, LFP signals displayed a strong entrainment to the rhythmical stimulation, which outlasted stimulus presentation (see figure 4.15a,b for one example site stimulated with pure tone trains of different rhythms). To quantify whether such entrainment had a facilitatory effect, the phase of filtered LFP traces was measured following stimulation end, specifically after one rythmical period (SOA) had passed. If the stimulation sequence had not stopped, a hypothetical tone would have occurred at that instant (this hypothetical tone is marked as 6 in figure 4.15a). The phase distribution at the hypothetical tone was significantly different between non-adapted responses and non-evoked responses (see figure 4.15c; p < 0.001 for all SOA tested, Kuiper test). The phase coherence for the non-adapted distribution was significant (VS = 0.42, p < 0.001) and of similar magnitude as previously found for adapted responses at the last tone presentation, contrasting with the lack of coherence for the non-evoked distribution (VS = 0.05, p > 0.001). The mean phase at the hypothetical tone was 2.27, a facilitative value as seen for evoked responses (see figure 4.15c; mean phase for evoked response at the fifth tone was 2.78 for non-adapted responses, and 2.10 for adapted responses). This suggested that rhythmical entrainment facilitated the upcoming representation of potential stimuli in a timely manner consistent with the stimulation SOA.

In scenarios where multiple rhythmical stimuli are competitively presented, rhythmical entrainment may adapt to promote the upcoming representation of specific behaviourally relevant stimuli. In behaving mice performing in the dual stream (S2) condition, one could expect entrainment to facilitate the upcoming High target representation in the Early block, and vice versa for Low target in the Late block. To test this hypothesis, I computed the change between Early and Late phase distributions in the Early Zone (EZ; 0-1.6 from trial onset; change computed as Early minus Late) across time, and quantified the change in significance via the p-value from the Kuiper test. The figure 4.16a,b presents the responses in Early and Late blocks of two example sites tuned to Low and High tones respectively (top panels), as well as the phase distribution change across time (middle panel) and corresponding p-value (bottom panel). For most sites, the difference in phase distribution was strongest in the 0.8s threshold).
time window prior to the first Early target (see figure 4.16c), consistent with previous studies showing that neural activity varies the most close to target occurrence\textsuperscript{[61]}. Having shown that the LFP phase was significantly modulated in the EZ, I aimed to decipher whether such modulation was beneficial to upcoming target encoding. To predict which LFP phase would be preferentially augmented at the time of the first Early target onset in the Early compared to Late condition, I interpolated the highest differences (above the 95% percentile; middle inserts in figure 4.16a,b; see figure I.41 for a comprehensive schematic of analysis procedure) between Early and Late phase distributions in the EZ. The interpolation curve was then used to compute a predicted phase augmentation at the time of target occurrence (specifically, at 20ms post target onset). Sites tuned to the Low tones displayed an enhancement in negative (non-facilitatory) phases, whilst sites tuned to the High tones displayed an enhancement in positive (facilitatory) phases (see figure 4.16d; correlation coefficient $r = -0.49, p < 0.001$). This difference can be attributed to top-down modulatory mechanisms, as the stimuli presented in the EZ were identical in the Early and Late blocks.

4.2.3 Spectrotemporal expectation modulated neural responses in auditory thalamus

Several theories have been advanced to account for receptive field modulation during behaviour\textsuperscript{[92,123,135,137]}, the most prominent of which suggesting that goal-driven modulation of sensory processing emerges from the influence of frontal cortex (FC) onto sensory cortex\textsuperscript{[125]}. Whilst there are evidences for the joint modulation of FC and A1 during complex behavioural tasks\textsuperscript{[115,219]}, the anatomical substrate for such influence remains unclear. In the visual system, the visual thalamus has been shown to be the locus of top-down modulation in demanding attention tasks\textsuperscript{[123,135,137]}. Frontal cortices may act onto thalamic reticular neurons (TRN), which in turn project onto sensory thalamic cells\textsuperscript{[123,135,138,140]}. To assess whether this principle may hold in the auditory modality, further recordings were made in the medial geniculate body (MGB) of mice during behaviour. The ventral subdivision of the MGB was primarily targeted, and recorded cells presented strong locking to click stimuli and low-latency responses typical of this subdivision\textsuperscript{[20,54]} (see section 2.2.4 Targeting the medial geniculate body of awake mice for methods).

Modification of the oddball task

Recording in the thalamus yielded less cells than in A1 per penetration, mainly due to the smaller region size. Typically, auditory responsive cells were found on two to four tetrodes out of the eight inserted in the thalamus, whilst in A1 all sites were auditory responsive. To maximise the chances of presenting a stimulus to which MGB cells responded, the oddball task was modified, thus encompassing a wider range of probe stimuli (see figure 4.17a). A single stream (S1) composed of isochronous tones (600 ms SOA; 10 kHz) was surrounded by a cloud of pure tones (named distractor tones; see SD condition in figure 4.17b). A target (4.4 oct/s FM UP sweep starting at 10kHz) was presented once per trial instead of a stream tone, either Early (∼ 2 s) or Late (∼ 4.5 s). The task thus contained spectrotemporal biases, as the target had fixed frequency content and occurred at determined moments. Mice reached high detection performance in the single-stream with distractor tones (SD) condition, as displayed in figure 4.17c. The decreased performance in the Late compared to Early block was due to an increase in FA rate (36.4% FA in Late block versus 29.2% in Early block),
To assess whether mice were sensitive to the spectrottemporal bias of the task, I analysed the FA timing during the task in the single-stream with distractor (SD) condition. As seen in figure 4.18a-b, mice timed their FA according to the likely occurrence of a target. FA were significantly delayed in the Late compared to Early block (p < 0.001, WRS test; see figure 4.18b). This indicated that mice adjusted their behavioural response according to the temporal bias of the target. It is important to highlight that the sole timing of the response according to the target spectral bias could not explain the high detection performance achieved by the animals (d’ > 1). Indeed, the d’ value reached during behaviour was always higher than the 95% percentile of the distribution of d’ values generated when randomly shuffling the animal’s response times across trials (n = 500 times, keeping the timing of the targets as presented during the behavioural session).

To assess whether mice were sensitive to the target spectral bias, I measured the RT distribution of FA in single-stream (S1) and single-stream with distractor (SD) condition. As displayed in figure 4.18c, all mice timed their FA according to tone onset in the S1 condition (top row). In the SD condition, two mice (ID 1,2) maintained such bias (VS > 0.30). This suggests that the animals could selectively follow the oddball stream despite the presence of distractor tones.

**Characterisation of MGB cell responses**

To determine whether these biases affected the responses of MGB cells, it was important to characterise the cells’ tuning profile, as expectation might differently affect responses to relevant (tone in stream and target) and irrelevant (distractor tones) sounds. In total, 29 auditory responsive cells were recorded in MGB in 4 mice. As seen in figure 4.19a-b, MGB cells displayed two distinct types of response, namely excitatory (a) and inhibitory (b). As a single cell could exhibit each response type depending on the tone frequency (see figure 4.19c-d), each response was treated separately in further analysis. The figure 4.19e displays the number of excitatory and inhibitory responses measured over the whole MGB cell population. Note that one cell can contribute towards responses in multiple frequencies, as the example cell shown in figure 4.19c-d. On average, responses were more prevalent in the frequency range 10-17 kHz (see figure 4.19e), encompassing the tone and target frequencies.

Having captured the frequency tuning of MGB cells, I assessed whether evoked responses were modulated in an informative manner between Early and Late blocks. The figure 4.20a displays the activity of an example cell in the Early and Late blocks in the Early Zone (EZ, 0-1.8 s from trial onset). As different sounds were presented as distractors across trials, it was not possible to assess changes in activity between Late and Early trials using as metric the summed activity in the EZ as was previously done in the A1 study. Instead, for each MGB response previously defined as significant, I measured the change in evoked ON firing rate between Early and Late blocks at the tone of corresponding frequency (see example grey boxes in figure 4.20a for responses to 10 kHz tones). The figure 4.20b displays the modulation index of responses with high discrimination performance (DP; see section 2.8.5 *Discrimination performance* for definition). On average, most responses exhibited an increase in Early compared to Late trials, as seen per positive modulation index values. This increase was significant at the single cell (p < 0.05 WRS test, comparing firing rate in Early and Late trials ; see filled circles in figure 4.20b) and population (p < 0.01 WRS test, comparing average firing
rate in Early and Late blocks) levels. This result recapitulated the earlier finding in A1 presented in figure 4.6, whereby A1 cells increased in firing rate prior to target occurrence irrespective of frequency tuning.

4.2.4 The tuning profile of auditory thalamic cells was not affected by TRN manipulation

The MGB receives inhibitory inputs from the thalamic reticular nucleus (TRN)\textsuperscript{[136,139]}, an important structure for sensory gain-control during attentional tasks\textsuperscript{[123,134,136,137,142]}. Despite its crucial role, how TRN affects MGB processing is poorly understood\textsuperscript{[141,144]}. It thus appeared necessary to study the impact of TRN onto MGB processing prior to assessing whether modulations observed in MGB during behaviour result from TRN interactions.

To manipulate the activity of TRN cells, I injected the TRN of PV-Cre mice with an AAV encoding for the inhibitory DREADD in a Cre-dependant manner (AAV-DIO-hM4-mCherry), as displayed in figure 4.21a. PV+ cells were targeted as they are known to preferentially project to sensory areas of the thalamus\textsuperscript{[138,140]}. Post-mortem histology confirmed the expression of DREADD in TRN PV+ cells (see figure 4.21b).

By recording MGB neurons’ response to pure tones before (Baseline) and post- CNO IP injection in awake mice (see electrode tract in figure 4.21c), I was able to assess whether PV+ TRN cells modulated MGB neurons’ tuning profile. The two example cells presented in figure 4.21d,e illustrate the main result of this experiment: the tuning profile of MGB neurons did not vary following the reduction of PV+ TRN cells activity. MGB cells whose spike waveform was stable across the whole recording session (n = 12; no significant change in largest spike amplitude between the first and last ten minutes of recording) did not display a significant change in ON CF (octave difference between Baseline and post-CNO: 0.04 ± 0.03, WSR p > 0.05). This effect was recapitulated at the multi-unit level (octave difference: 0.05 ± 0.02, WSR p > 0.05).

Two possible causes could underlie this lack of effect. Firstly, the DREADD may not have functioned as expected, and secondly the infected neurons may not have been part of a network comprising MGB cells. The second reason is the most likely, as the effect of DREADD had been validated \textit{in vivo} in A1 (see chapter 3 Functional determinants of response selectivity to frequency modulated sounds) and \textit{in vitro} in TRN (data recorded by Diana Lucaci, not shown). To specifically manipulate TRN cells projecting onto MGB neurons, I injected the retrograde canine adenovirus (type 2) encoding for GFP and Cre-recombinase (CAV-Cre-GFP) in the MGB of wild type (WT) mice (with the aim to further inject AAV-DIO-hM4-mCherry in the TRN). Whilst this viral construct has been showed to readily infect a large array of cortical and sub-cortical structures\textsuperscript{[165]}, its efficacy had never been tested in the auditory thalamus.

I validated the CAV titration by injecting a WT mouse with CAV-Cre-GFP and AAV-DIO-hM4-mCherry in AC in distinct hemispheres (see figure 4.22a for schematic and histological confirmation). mCherry-positive cells were observed in the AAV-injected hemisphere only. Subsequently, I injected CAV-Cre-GFP in the MGB of WT mice (n = 4), and assessed viral expression in areas projecting onto MGB (notably in the AC, IC and TRN). In two mice, GFP signals were observed in the MGB, but mainly surrounding the putative ventral zone (see figure 4.22b). Few GFP-positive cells were also observed in the IC, mainly concentrated in the dorsal (see figure 4.22c) and lateral cortex zone. This
pattern was consistent with non-central IC cells innervating preferentially non-lemniscal MGB\textsuperscript{[220]}. However, no GFP-positive cells were observed in deeps layers of AC, nor in TRN.
4.3 Discussion

In a world where the senses are continuously stimulated, optimisation of information processing is believed crucial for perception and resulting behavioural actions\cite{25,63,137}. To understand how such optimisation occurs is a major quest in sensory neuroscience. The mouse is an ideal model in which to study the neural processes underlying sensory refinement, as it is possible to precisely record and manipulate defined neural circuit components.

4.3.1 Auditory attention and streaming could be inferred from the timing of false alarms

Here, I developed novel behavioural tasks in mice to assess perceptual aspects of auditory streaming. Animals were trained to detect an oddball sound in one of two simultaneously presented, isochronous pure tone sequences. Temporal expectation was manipulated by presenting the target sound in a particular stream either early or late with respect to trial onset in blocks of trials. Animals reached high performance on this task, and notably their false alarms were very instructive of their behavioural state. False alarm timing was markedly delayed for late blocks compared to early ones, indicating that the animals associated a different context to an otherwise identical stimulus. More finely, I observed that the false alarms were timed to the onset of the sounds present in the oddball stream. This suggests that the animals could selectively follow the oddball stream despite the presence of a distractor stream. This effect was replicated in a second task design, where the distractor stream was replaced by a cloud of tones. Together, these results indicate that perceptual streaming can be inferred via the timing of false alarms in mice, and provide a new paradigm with which to investigate neuronal mechanisms of selective attention.

4.3.2 Activity in auditory cortex depended on attentional state – An effect mediated by the thalamus?

By recording in auditory cortex and thalamus during ongoing behaviour, I related modulation of sensory encoding to changes in behavioural states. Specifically, activity in the AC was modulated by expectation, in a manner that reflected the timing of the upcoming target, but not the spectral content of that target. Most cells displayed an increase in spiking activity prior to the first target during early trials, independently of their tuning profile. During single trial, pure tone representation was enhanced across time, so as to obtain the best SNR near likely target occurrence. Accordingly, mice responded faster when neural decoding of tone and target was increased. These results were reminiscent of earlier findings by Jaramillo and Zador\cite{61}, who showed that attention in time dynamically enhances AC processing.

Neural activity at the level of AC has recently been proposed to encode more than simply stimulus features\cite{115-117}. Here, both evoked and spontaneous firing rate modulations were informative on the type of trial (early or late). Moreover, a classifier trained based on the activity of a subset of AC cells during Hit trials was sufficient to decode response timing during FA trials. This implies that mice use similar behavioural strategies during Hit and FA, and that these strategies impact on neural processing at the level of the AC. Whilst AC firing rate modulations were informative on trial and response specificity, they were inconsistent with sound processing enhancement. These modulations might instead reflect the encoding of other task aspects, such as reward timing anticipation or task
Modulations in spiking activity were observed in both auditory cortex and thalamus, suggesting that the thalamus may act as the anatomical locus of sensory modulation during attentional processes\(^{123,137}\). Cells emerging from the thalamic reticular nucleus (TRN) can strongly inhibit sensory thalamic neurons\(^{139,141}\), providing an ideal anatomical substrate for sensory filtering in conflicting environment\(^{123,136,137}\). However, little is known about how the TRN affects MGB processing\(^{141}\). Here, I manipulated PV+ cells in the TRN during recording of MGB neurons in awake, passive-listening mice. The selective inhibition of this TRN cellular sub-population did not alter the response of MGB neurons to pure tones. This was surprising, as TRN PV+ cells are known to preferentially project to sensory thalamic nuclei\(^{140}\). Whether these cells are active during awake, passive-listening state is unknown\(^{139,140,142}\).

To ensure this lack of effect was not due to poor viral infection spread, I aimed to use a retrograde virus to specifically control the TRN cells projecting onto MGB. A new version of canine adenovirus (CAV type 2)\(^{165}\) was injected in the MGB of WT mice, and the projecting regions were inspected. Fluorescent cells were recovered in projecting areas (specifically in the IC), advocating CAV as a promising tool to study the role of TRN in modulating sub-cortical sensory processing. Further work is however required to ensure TRN cells possess appropriate tissue tropism for CAV.

### 4.3.3 Comparison of local field potential signals in the auditory cortex of mice and primates

The neural underpinnings of selective listening have mainly been studied in human and non-human primates\(^{109,215}\) via the analysis of local field potential (LFP) signals. LFPs are robust and long-lasting signals which can be retrieved from implanted electrodes, often used in human patients to monitor brain activity and control prosthetics\(^{221}\). LFP signals provide an integrated measure of several sources of electrical potential over time and space\(^{222,223}\), and as such reflect neural network state\(^{101,105,107}\). Whilst the mammalian brain possess cytoarchitectonic characteristics conserved throughout evolution\(^{224}\), it is not clear whether the LFP reflects similar processes across species.

Here, I recorded LFP signals in the mouse AC during the performance of an auditory task involving attention and streaming. Low-frequency LFP oscillations were affected by behavioural state, in a manner that was consistent with sensory encoding enhancement. Oscillations could entrain to the rhythm of a stimulus sequence, and remained entrained following stimulus termination – a property also observed in primate AC\(^{108,109}\). However, the phase of the LFP signal in mice was only modestly dependent on the spectral content of the stimulus, therefore casting doubt on the generality of the LFP acting as a spectrotemporal filter\(^{109,144}\).

It is important to highlight that the observation of ‘anti-phase’ entrainment in primate requires the use of tone sequences separated by at least two octaves\(^{108,109}\). In contrast, the octave difference between high and low tone sequences used in this study was less than two octaves. As the minimal octave difference required for the perception of two streams in the ‘ABA’ paradigm can be smaller than two octaves\(^{70}\), it was hypothesised that mice could perceive two segregated streams in this task. This was confirmed by the analysis of false alarms, as previously discussed. It is thus possible that the phase of the LFP signal does not reflect streaming per se, but rather attentive processes aiming at enhancing sensory signal representation.
4.4 Future work

4.4.1 Delineating the effect of PFC and TRN on sound processing

In this work, I developed a system to simultaneously record from PFC and AC in mice during behaviour. This method is believed crucial when willing to uncover the anatomical basis of sensory modulation during attention tasks. Notably, the analysis of LFP signals may inform on the coupling between PFC and AC\(^{[225]}\). Here, the analysis of PFC signals were not extended, as it was not clear which region of PFC would be implicated in the oddball task. Optogenetic silencing of precise regions of PFC such as done by Halassa et al.\(^{[123]}\) during behaviour is required to further define the region of interest for electrophysiological recording.

PFC has been shown to impact on the TRN in a cross-modal attention task\(^{[123]}\). However, the anatomical connections between PFC and TRN are unknown. Future anatomical tracer studies would be extremely beneficial in delineating the pathway between these two brain regions, using e.g. retrograde and anterograde viral labelling techniques\(^{[41]}\). Further work is also required to establish the role of the TRN in shaping sensory representation at the level of the auditory thalamus, as previously discussed.

4.4.2 Study of recurrent thalamo-cortical circuits

Modulations in neural activity have been observed in AC and MGB during behaviour. To uncover the anatomical origin of such modulation appears as an intricate endeavour, as the auditory cortex and thalamus are two brain regions heavily interconnected\(^{[17,19,24]}\). These recurrent circuits might be involved in shaping sound processing during behaviour\(^{[163]}\). Simultaneous recording in both brain regions would enable the assessment of these circuits’ dynamics. Neuropixels probes (UCL, UK) have been used for such a purpose in the motor and visual systems\(^{[226]}\), providing exquisite insight on the function of thalamo-cortical connections. As the auditory cortex and thalamus are located nearby in the mouse brain, the use of such probe appears experimentally feasible.

4.4.3 Behavioural paradigm refinement

It is acknowledged that the presently designed task could benefit from further refinement. To assess whether animals selectively attend to a specific stream, two target stimuli may be presented within the same trial (as done in primates in the original design of this task\(^{[109]}\)). Training in such a paradigm was attempted (see section 2.4.2 Paradigms), but mice did not manage to refrain from responding to a stimulus previously associated with a reward (data not shown as part of this thesis). To circumvent this issue, competitive selection task protocols currently developed in rodents\(^{[115,123]}\) present instead two simultaneous target stimuli, i.e. both stimuli are associated with a reward. The selection rule defines the action to obtain the reward. This requires however two different behavioural reporters (such as ‘go left’ and ‘go right’), which was not available in this study using the current setup (i.e. a single lick port). Dual lick ports have been developed to train head-fixed mice in two-alternative force choice (2AFC) tasks\(^{[216,217]}\). However, a dual lick port functions by the use of electrical current and the mouse acting as a switch, generating large electrical charges upon licking. An electrical dual lick port was tested in this study (design obtained from Lloyd Russel, UCL), but eventually judged not suitable as the quality of electrophysiological recording was dramatically affected. Optical dual lick ports are not yet developed\(^{[216,217]}\), potentially due to the fact that electrical components remain
large relative to mouse size. A rotary wheel may be an appropriate behavioural reporter apparatus in future studies\textsuperscript{[216,217]}.

4.4.4 Assessing the effect of movement on neural processing

As it is known that engaged animals can adopt stereotypical movements indicative of their subsequent behaviour\textsuperscript{[130]}, and that A1 activity can be rapidly modulated by motor-related areas\textsuperscript{[131–133]}, videos were recorded in a subset of behavioural sessions to assess the movement of animals during behaviour (60 fps, videos recorded by Gaelle Chapuis and analysed by Xi Chen as part of an MSc project; data not presented as part of this thesis). Mice exhibited movement prior to lick detection, however it was difficult to assess their stereotypy. Linking electrophysiological recording with high-speed imaging of mouse behaviour will be required in future studies to rule out movement as an origin for modulation of cortical signals. Such system would further permit to assess pupil diameter, an important read-out of behavioural state\textsuperscript{[132,216]}. 
Figure 4.1: Complex discrimination task involving spectrotemporal bias. a. The stimulus consisted of two isochronous trains of 100ms pure tone (termed stream), one with High frequency tones and 400ms stimulus onset asynchrony (SOA), and one with Low frequency tones and 617ms SOA. A target (frequency modulated sound) could occur at three positions with equal probability in a trial (the positions are marked by the colored squares). To generate a spectrotemporal bias, the occurrence timing and the frequency content of a target was jointly modulated, i.e. High frequency targets were presented Early, and Low frequency targets were presented Late. Spectrotemporal bias was manipulated using a trial block design. b. A block was composed of three trial types: target only (S0; 7 trials), single stream (S1; 10 trials), and dual stream (S2). Once the S2 condition was reached, mice had to perform 20-25 correct trials for the stimulus to switch to the other block type. c. Mice were water-deprived and head-fixed onto an apparatus enabling the detection of licking events. The first lick detected in a trial was counted as the response (represented as the darker filled dot on top of the stimulus trace). The mouse was required not to lick before the occurrence of a target. If the mouse responded in the target window, a drop of water was administered (Hit). If the mouse responded before (false alarm; FA), a punishment was administered. d. Averaged $d'$ over the 3 sessions prior to recording for each mouse in the S2 Early and Late condition. $d'$ values above 1 indicate good performance.

Figure 4.2: Mice were sensitive to temporal bias as revealed by their false alarm timing. a. Example session of a mouse (ID 3). Each dot represents a response (Hit or FA). FA are color-coded based on the block type (Early or Late). The Early zone (EZ) starts at trial onset and ends and the last Early target onset. b. Left panel: Significant increase in the median of FA time (measured from trial onset) in the Late versus Early condition. Error bars represent median absolute deviation. Right panel: Significant decrease in the percentage of FA in the EZ in the Late versus Early condition.
Figure 4.3: Mice timed their false alarm responses to tone onsets. a. Example FA during the Late block in a single session. Each dot represents a FA response, in the S0 (gray) or S1 (blue) condition. Vertical gray lines indicate sound onsets in S1 condition. b. FA reaction time (RT) distributions for S0 and S1 conditions (same colors as in a). The Vector Strengh (VS) indicates the peak strength of the distribution, and is reported in the inset (*p<0.05). c. Normalised RT distribution for FA in S1 conditions for each mice (one mouse per column, the mouse ID is presented in the square box) computed using all sessions. RT are measured either from High (top) or Low (bottom) tone onsets. The distributions are color-coded based on the block type (Early or Late). The RT distribution of a mouse generating FA timed to sound onsets presents a peak, as assessed by the VS (* p < 0.05). RT distributions are smoothed (1 bin std Gaussian filter) for display. d. The VS was significantly higher in the S1 than S0 condition. Each line indicates the VS change between S1 and S0 condition for a mouse generating FA in a specific block type (Early or Late).

Figure 4.4: Mice were sensitive to spectral bias as revealed by their false alarm timing. d. Normalised RT distribution for FA in S2 conditions for each mice (one mouse per column, the mouse ID is presented in the square box) computed using all sessions. RT are measured either from High (top) or Low (bottom) tone onsets. The distributions are color-coded based on the block type (Early or Late). RT distributions are smoothed (1 bin std Gaussian filter) for display. The p-value from a KS test comparing the RT distributions in Early and Late conditions is displayed as an indication of response bias (**p<0.01, ***p<0.001; Bonferroni correction factor 8). The first mouse times its FA mainly to the High tones. The second mouse times its FA to the cued tones. The third and fourth mice time their FA mainly to the Low tones.
Figure 4.5: Spiking activity was modulated by task specificity. a. Single recording site presenting multi-unit spiking activity tuned to the High frequency. Top and second panels: Raster plots and corresponding mean PSTHs of spiking activity during Hit trials in S2 condition, color-coded based on block type (Early or Late). The time axis is truncated so as to encompass only the Early zone (EZ; 0-1.6s from trial onset). The orange arrow marks the onset time of the first target in the Early block. The vertical lines indicate tone onset (black: High frequency, gray: Low frequency). Third panel: Firing rate difference modulation (measured as the difference) between the Early and Late PSTHs. Fourth panel: Discrimination performance (quantified by the AUROC from logistic regression). The horizontal line indicates an AUROC value of 0.6. The black arrows indicate local maximal AUROC value within a stimulus period. At this recording site, evoked activity is modulated by task context. b. Same as in a, except that spontaneous activity preceding High tone onset is modulated by task context. c. Heterogeneity in the timing of information across sites. AUROC curves for all sites recorded: each block represents one experiment, and one line represents a recording site (n=2 mice; Mouse 4: top 2 blocks, Mouse 2: bottom block). Only sites with significant evoked spiking responses were included. The black arrows indicate the sites presented in a and b.
Figure 4.6: Increase in spiking activity prior to target occurrence irrespective of frequency tuning. 

**a.** Increase in spiking activity in Early versus Late trials for an example cell tuned to the High frequency sound. **Top and second panels:** Raster plots and corresponding mean PSTHs of spiking activity during Hit trials in S2 condition, color-coded based on block type (Early or Late). The time axis is truncated so as to encompass only the Early zone (EZ; 0-1.6s from trial onset). The orange arrow marks the onset time of the first target in the Early block. The vertical lines indicate tone onset (black: High frequency, gray: Low frequency). **Third panel:** Firing rate modulation (measured as the difference) between the Early and Late PSTHs. Values above zero indicate an increase in firing rate in the Early versus Late condition.

**b.** Same as in a, but for a cell tuned to the Low frequency.

**c.** Same as in a, but for a cell not tuned to High nor Low tones.

**d.** Modulation of the firing rate between Early and Late conditions, displayed for each cell type. A value above 0 indicates an increase in the Early versus Late trials. Each dot represents a cell, and the dot is filled with grey color if the firing rate of the cell significantly differed in Early versus Late condition.
Figure 4.7: Dynamic increase in spiking activity prior to target occurrence. 
a. Example raster and mean PSTH of spiking activity for a cell (classified as Other) during Late trials (S2 Hit only). The raster displays trials from a single block judged most representative, whilst the PSTH represents the mean activity over all blocks. The early and late zones define the portions of time used to compute the modulation of firing rate presented in c. Both time zones are 1.6 second long. The early zone starts at trial onset, whilst the late zone ends prior to the first Late target. 
b. Left panel: Auto-correlogram of the cell presented in a. The 3ms refractory period is marked in yellow. This cell presents minimal contamination in the refractory period, and it thus considered well isolated. Right panel: Waveforms of the cell presented in a. The grey traces are individual waveform traces, and the black trace is the mean. Note that four sets of waveforms are presented since the cell was recorded using a tetrode comprising four electrodes. The horizontal bar indicates 0.4ms, and the vertical bar indicates 10µV. c. Modulation in the firing rate between the early and late time zones during Late trials, displayed for each cell type. A value above 0 indicates an increase in the late versus early zone. Each dot represents a cell, and the dot is filled with grey color if the firing rate of the cell significantly differed in the early versus late zone.
Figure 4.8: Modulation in evoked and baseline spiking activity prior to target occurrence. 

a. PSTHs in Early and Late trials for an example cell tuned to the High frequency sound. The Evoked (orange) and Baseline (grey) windows were defined around the tone onsets to which the cell was tuned.

b. Modulation of the firing rate between Early and Late trials in Evoked and Baseline windows for High and Low tuned cells. Each dot represents a cell (no significant modulation was found).

c. Change in signal to noise ratio (SNR) between Early and Late trials for High (purple) and Low (gold) tuned cells. The SNR is measured as the absolute value of the modulation in firing rate in the Evoked (E) and Baseline (B) windows. The cross indicates the mean SNR over the whole cell population.

d. Comparison of the modulation index for evoked (x-axis) and baseline (y-axis) activity for High (purple) and Low (gold) tuned cells (same metric and data as presented in b).
Figure 4.9: Predicting FA specificity using neural responses during Hit trials. a. The activity in the 0-1.2ms in Early and Late Hit trials (S2 condition only) was used to train the classifier in distinguishing between Early and Late responses. The blue and orange traces are the mean PSTHs in Early and Late Hit trials respectively, for an example cell tuned to the Low frequency. The green box at the end of the PSTH symbolises the moments when a mouse would generate Hit responses. b. The activity in the 0-1.2ms in Early and Late FA trials (S2 condition only) was used to test the classifier. The blue and orange traces are the mean PSTHs in Early and Late FA trials respectively, for the example cell in a. The red box at the end of the PSTH symbolises the moments when a mouse would generate FA responses.
Figure 4.10: Dynamic enhancement of signal to noise ratio prior to target occurrence. 

a. Example cell displaying a modulation of the firing rate across time. **Top panel:** Raster plot of spiking activity during S2 Early Hit trials. The First and Last sounds are those presented first in a trial and last before the Early Target (marked with an orange arrow). This example cell is tuned to the Low frequency. **Bottom panel:** Corresponding PSTH. The orange line links the maximum evoked firing rate at each sound presentation. 

b. Average normalised firing rate in Early and Late condition for cells tuned to the Low frequency. Normalisation was performed respective to the first sound presented. Solid lines represent the normalised evoked activity, whilst dashed line represent the normalised baseline activity. Only cells with significant response to the first tone were included in the evoked average. Only cells with non-null baseline firing rate were included in the baseline average.

c. Ratio of evoked/baseline normalised firing rate for Low cells. On average, the ratio increases prior to target occurrence in the Early block compared to Late block. Note that this difference in ratio was non significant (p>0.05, WRS test).

d. Same as in c, but for High cells.

e. Change in signal to noise ratio (SNR) between Last and First sound for Low cells. 

f. Same as in e, but for High cells.
Figure 4.11: Modulation of firing rate prior to target occurrence was correlated with behaviour. 

a. Top panel: Average responses from single cell example for trials with expected early targets grouped according to reaction time. Averages are taken over those trials with the 20% fastest (red) or the 20% slowest (blue) reaction times. To compare the activity between Fast and Slow trials, the activity is measured in the 400ms encompassing the Last sound presented before the first Early Target. Bottom panel: Difference in firing rate between Fast and Slow trials. On average, this cell presented a moderate increase in the firing rate between Fast and Slow trials in the Last sound window.

b. Same as in a, but for a cell displaying a more sustained increase in the firing rate between Fast and Slow trials in the Last sound window.

c. Modulation of the firing rate between Fast and Slow trials in the Last sound window for all responsive cells. A value above 0 indicates an increase in the Fast versus Slow trials. Each dot represents a cell (no significant modulation was found).
Figure 4.12: Tone-Target discrimination performance of single cells in auditory cortex was correlated with behaviour. 

**a.** Example PSTHs of a cell during Fast (top) or Slow (bottom) trials, in response either to the Target (right) or last preceding Tone (left). Trials are organised so that the fastest trial is at the top, and the slowest at the bottom. 

**b.** Discrimination performance (DP) was computed using logistic regression between Tone and Target trials at each PSTH bin. The DP accuracy and timing were defined as the maximum value of the AUROC curve and corresponding timing respectively. For this example cell, DP timing is delayed in Slow compared to Fast trials. 

**c.** Mean reaction time in Slow and Fast trials used to compute the AUROC presented in b. It is important to note that the time scale of the RT difference between Slow and Fast trials is one order magnitude greater than for single cell DP timing difference. (By design, RT in Slow and Fast trials are significantly different; p<0.001 WRS) 

**d.** Timing of maximum AUROC value for all cells passing selection criterion (n=8). Gray dots indicate values not significantly different from baseline distribution. Blue and red dots indicate timing significantly earlier or later than baseline distribution respectively. 5 cells are significantly modulated (3 faster, 2 slower) in accordance with behavioural outcome. Colored rectangles indicate mean ± sem timing for all cells. (Significance threshold: p<0.05; n.s. not significant, WRS test). 

**e.** Same as in d but for DP accuracy. Blue and red dots indicate values significantly lower or larger than baseline distribution respectively. (Significance threshold: p<0.05). The DP accuracy of most cells is not significantly modulated.
Figure 4.13: LFP phase-reset was moderately visible in anaesthetised mice. 

a. Example average of raw (dark gray) and band-pass filtered (4-9Hz; colored) LFP traces recorded upon the presentation of a sequence of five pure tones. The color of the filtered trace represents the phase of the signal (as in b). Note that the LFP signal adapts with each sound presentation, and that the phase is different if taken at the onset of the first or fifth tone. 

b. Example LFP receptive fields (z-score) at each tone (top), and corresponding phase receptive fields (bottom). The non-evoked zone is marked in gray. The adaptation zone is defined as the portion of the non-evoked zone at the fifth tone which was evoked at the first tone. In this example, an anti-phase change is observed in the adaptation zone (from positive at the first tone to negative at the fifth tone). 

c. Normalised phase distributions for adapted and non-adapted responses at the first (blue) and fifth (red) tone. Mean phases are represented by the vertical lines and reported in the inset. For adapted responses only, the LFP phase distribution at the fifth tone significantly differed (p<0.001; Kuiper test) from the distribution at the first tone. (Example distributions given for SOA 150ms).
Figure 4.14: LFP phase-reset was moderately visible in active behaving mice. **a.** Top panels: Average periodic LFP traces. Example average LFP trace of a site with positive FSI when aligned to the pure tone onsets presented in S1 condition. High or Low indicates the frequency of the pure tone. The response in the time window 0-0.2s post stimulus onset (gray) is removed from the interpolation analysis so as to tract fluctuation in the baseline signal only. The blue line displays the response amplitude used in computing the FSI. **Bottom panels:** Interpolated periodic baseline traces. The baseline part (black) of the average LFP trace is used to compute an interpolated curve (red), from which the phase is derived (colored line). Note that the period of the interpolated curve is consistent with the SOAs used for High and Low streams. **b.** Similar as in **a**, but for a site with negative FSI. The sites presented in **a** and **b** are from the same animal (Mouse 4). **c.** A moderate shift in the phase with respect to the FSI is visible across sites (example: Mouse 4). Each dot represents a site. The y-value corresponds to the site frequency preference (FSI). The color represents the stimulation frequency. The dashed lines are manually-drawn curves expressing the shift direction. Green and red circles display two example phases for preferred frequency (PF; site tuned to High, stimulation High) and non-PF (site tuned to Low, stimulation High). **d.** Expected model drawn from previous studies in primates\textsuperscript{108,109}, where PF and non-PF tone sequences generate baseline LFP phase near ±π and 0 respectively. Axes and colors are similar as in **c**. **e.** On average, the phase measured from the interpolated curve at stimulus onset was not consistently distributed based on frequency tuning (n=3 mice). Green and red histograms represent phase distributions at PF and non-PF respectively. The mean phase and number of site of each distribution are reported on top of the graph.
Figure 4.15: Facilitative modulation of the LFP phase distribution following repetitive stimulus presentation. 

a. Example average evoked LFP trace recorded upon the presentation of sequences of five pure tones (SOA 150ms). The sixth gray box represents a hypothetical sixth tone.

b. Same as in a, but using sequences of SOA 300ms.

c. Normalised phase distributions at the hypothetical sixth tone for non-adapted (measured at the fifth tone; top panel) and non-evoked (bottom panel) responses. The LFP phase distribution at the sixth tone for the non-adapted responses was significantly different (p<0.001; Kuiper test) from the non-evoked distribution. (Example distributions given for SOA 150ms).
Figure 4.16: Facilitative modulation of the LFP phase distribution prior to target occurrence. **a.** Top and second panels: Example average LFP trace in the Early zone (EZ; 0-1.6s from trial onset) in S2 condition for Early (top) and Late (bottom) blocks. This site is tuned to the Low frequency. The gray trace represents the average of the raw LFP, and the colored trace represents the average filtered trace (multiplied by a factor 2 for display). The color on the filtered trace represents the average phase gathered from the filtered trace at each trial (note that it is not equal to the phase of the average filtered signal). **Third panel:** Percentage change between the normalised phase distributions in the Early and Late blocks at each time point in the EZ. Only the 5% and 95% percentile values are plotted. **Bottom panel:** Log p-value originating from the Kuiper test between Early and Late phase distributions reported at each time point in the EZ. **b.** Similar as in a, but for a site tuned to the High frequency. **c.** For most sites, the strongest difference between the phase distributions (as assessed by the p-value) was found in the 0.8s window preceding the first Early target occurrence (n=2 mice; top: Mouse 4, bottom: Mouse 2). **d.** Interpolated LFP phase change at the first Early target is correlated with a site’s tuning (r=-0.49). Sites tuned to the Low frequency (positive FSI) show an increase in negative phases in the Early compared to Low block, whilst sites tuned to the High frequency (negative FSI) show an increase in positive phases. Each dot represents a site (n=16 in n=2 mice), and the color indicates the goodness of the interpolation (R-square). The gray line represents the linear fit on the data points, computed using the R-square values as weights. The top curve is a schematic of the LFP so as to visualise LFP trough and peak according to the definition of the phase.
Figure 4.17: Modified discrimination task involving spectrottemporal bias. 

a. The stimulus consisted of an isochronous train of 100ms pure tone (termed stream; 10kHz), surrounded by distractor tones. A target (frequency modulated sound) could occur at two positions instead of the stream tone with equal probability in a trial (the positions are marked by the colored squares).

b. A training session was composed of three types of trial: target only (S0), single stream (S1), and single stream with distractor tones (SD). Once distractor tones were presented, targets occurred either Late or Early using a trial block design.

c. Average performance, measured as $d'$, over the 3 sessions prior to recording for each mouse in the SD condition. $d'$ values above 1 indicate good performance.
Figure 4.18: Mice were sensitive to the spectrotemporal bias in the modified discrimination task as revealed by their false alarm timing. 

a. Example session of a mouse (ID 1). Each dot represents a response (Hit or FA; only responses in SD are displayed). FA are color-coded based on the block type (Early or Late).

b. Significant increase in the median of FA time (measured from trial onset) in the Late versus Early block. Error bars represent median absolute deviation.

c. Normalised reaction time (RT) distribution for FA in S1 (top) SD (bottom) conditions for each mouse (one mouse per column; the mouse ID is presented in the square box). FA in Early and Late blocks were used to generate the RT distribution. The RT distribution of a mouse generating FA timed to sound onsets presents a peak, as assessed by the vector strength (VS). The star indicates significance (* p < 0.001).
Figure 4.19: Thalamic cells displayed excitatory and inhibitory responses. a. Example raster (top) and corresponding PSTH (bottom) of an MGB cell in response to 10kHz pure tone. This cell displays an excitatory ON response. b. Same as in a, but for a cell displaying an inhibitory ON response (as well as an inhibitory OFF response, but such response is not considered for further analysis). c. Example raster of an MGB cell displaying excitatory and inhibitory ON responses depending on the frequency content of the pure tone presented. The frequency of the tone is indicated on the y-axis. Each raster displays 1000 trials. Baseline and ON windows used in d are indicated on top of the panel. d. Ratio of firing rate in ON over Baseline window for the cell presented in c at each tone frequency. Values above 1 indicate excitatory ON responses compared to Baseline. Stars indicate significant difference in firing rate between Baseline and ON windows. As seen in c, this cell displayed inhibitory response to the lowest frequency, and excitatory responses to the mid-range frequencies. e. Number of responses labelled as significantly excitatory or inhibitory. On average, most responses were tuned around the stream and target frequencies (10-17kHz).
Figure 4.20: Thalamic cells displayed an increase in firing rate prior to target occurrence irrespective of frequency tuning. a. Raster plots and corresponding mean PSTHs of spiking activity during Hit trials in SD condition, color-coded based on block type (Early or Late). The time axis is truncated so as to encompass only the Early Zone (EZ; 0-1.8 s). The orange arrow indicates the onset time of the first target in Early trials. The vertical lines indicate tone onsets (black: 10 kHz stream, grey: distractor tones). PSTHs in the distractor portion are shown as dashed lines as they represent average firing rate in response to tones of various frequencies. b. Modulation index of the firing rate between Early and Late trials conditions, displayed for each response type. A value above 0 indicates an increase in firing rate in the Early versus Late trials. Filled dots indicate significantly modulated responses (p < 0.05, WRS test).
Figure 4.21: MGB cells did not change tuning profile following reduction of PV-expressing TRN cells activity.

a. AAV-DIO-hM4-mCherry was injected bilaterally in the TRN of PV-Cre mice. Black squares indicate imaging region (green: Anti-PV; red: mCherry). Scale bar: 1000 µm. b. PV+ cells in the TRN expressed the DREADD as revealed by immunochemistry. c. Electrophysiological recordings were performed in the MGB of awake mice. Purple traces (Di-I) indicate electrode tracks. d-e. Example MGB cells that conserved their tuning profile (inhibitory in d and excitatory in e) post CNO injection. Throughout: Top row: Baseline; Bottom row: Post-CNO; Left panel: Raster; Middle panel: FRA ON; Right panel: FRA OFF. Black bar in Raster indicates tone presentation.
Figure 4.22: Testing the CAV preparation. a. AAV-DIO-hM4-mCherry was first injected unilaterally in the auditory cortex (AC) of a wild-type mouse. 10 days following the first injection, CAV-Cre-GFP was injected in AC in the contra-hemisphere. 5 weeks following the second injection, the animal was sacrificed and the brain was sliced. Neurons expressing GFP and mCherry were recovered near the first injection site (black square). b. CAV-Cre-GFP was injected in MGB unilaterally. 5 weeks following injection, the animal was sacrificed and the brain was sliced. GFP signal was low in MGB (black square). c. GFP-expressing cells were observed in the inferior colliculus (IC) of the same animal as in b. Arrowheads indicate example cells. Presumably, these cells projected onto the injected thalamic region. White scale bars: 500 µm. AAV-DIO-HM4-mCherry was first injected unilaterally in the auditory cortex (AC) of a wild-type mouse. 10 days following the first injection, CAV-Cre-GFP was injected in AC in the contra-hemisphere. 5 weeks following the second injection, the animal was sacrificed and the brain was sliced. Neurons expressing GFP and mCherry were recovered near the first injection site (black square). Figure 4.22: Testing the CAV preparation. a. AAV-DIO-hM4-mCherry was first injected unilaterally in the auditory cortex (AC) of a wild-type mouse. 10 days following the first injection, CAV-Cre-GFP was injected in AC in the contra-hemisphere. 5 weeks following the second injection, the animal was sacrificed and the brain was sliced. Neurons expressing GFP and mCherry were recovered near the first injection site (black square).
Hearing is one of the five traditional senses humans possess and share with several other species. Animals use sounds to communicate with each other, but also to identify threats or mating partners. Hearing thus appears as an essential sense for survival and reproduction conserved throughout evolution. It is because of its fundamental nature that I dedicated this thesis to the study of hearing function.

Sensation entails the ability to transduce, encode, and ultimately perceive the stimuli arising from the environment. Much of the brain is devoted to these tasks, however still little is known about how neural codes give rise to perception. This is a major quest in neuroscience, which I aimed to follow. Taking advantage of new genetic technologies, I uncovered several principles underlying the encoding and perception of sound, summarised below.

A natural acoustic environment is dynamic, comprising sounds which vary in frequency and amplitude over time. Frequency modulations are important cues, notably in vocal calls, however the mechanisms underlying the representation and perception of these modulations are unclear.

In chapter 3 Functional determinants of response selectivity to frequency modulated sounds, I studied the functional determinants underlying response selectivity to frequency modulation direction in the mouse auditory cortex. Specifically, I investigated the role of ON and OFF inputs, as well as the role of a particular type of interneurons (PV-expressing) suspected to be involved in shaping cortical direction selectivity. This work was part of a larger research plan\textsuperscript{[179]} aiming at understanding the functional role and underlying mechanisms of ON and OFF receptive fields arrangement.

Overall, I revealed that ON and OFF inputs may interact with one another \textit{in vivo} as predicted by theoretical models. ON and OFF RFs were arranged along the tonotopic axis, so that cells with low-frequency ON RF exhibited high-frequency OFF RF, and vice versa. The reversal point of such arrangement coincided with the reversal point for direction selectivity, suggesting that ON and OFF inputs may contribute to direction selectivity shaping in A1. In contrast, cortical inhibition mediated by PV-expressing cells was not required for A1 neurons to exhibit selectivity to the direction of slow frequency modulations, nor for mice to perceptually discriminate such direction. As these slow rates are most relevant for vocalisation in mice, the interaction of ON and OFF inputs may constitute a fundamental mechanism for neurons to encode ethologically relevant sounds. These results highlight that to understand the selectivity to higher-order stimulus features in A1, it is necessary to consider both ON and OFF RFs. Furthermore, as all neurons in the neocortex show selectivity to more than one stimulus feature, these results suggest that it is necessary to consider the rich tuning properties of individual cells to fully understand neuronal selectivity.

Natural environments are further challenging because of the multitude of events they comprise. An acoustic scene typically comprises several sound sources, which may generate sounds simultaneously. How animals are able to selectively follow a single sound source in such a noisy environment is a major puzzle for auditory scientists.

In chapter 4 Anatomical circuits underlying task-related modulation of spectrotemporal receptive fields, I investigated how the auditory system may adapt the representation of auditory inputs to facilitate
sound perception in a noisy setting. Notably, I successfully developed methods to train mice in complex auditory tasks, and to record electrophysiological signals in several regions of the awake mouse brain. These methods enabled me to relate neural code to perception and action.

Mice were presented with relevant and irrelevant sounds concurrently, and required to detect a specific sound within that mixture. Attention enhanced sound representation in a dynamic manner, as previously reported\cite{61,98,116}. Changes in neural activity by attentive state were visible at the level of both the auditory cortex and the auditory thalamus, suggesting that the thalamus may be the primary locus of executive ('top-down') control. However, several anatomical pathways may modulate the activity of auditory cortex and thalamus during behaviour. Retrograde viral techniques, some of which tested in this thesis, offer a powerful means by which to test for the functional role of these anatomical pathways.
Bibliography


T. Mueller. What is the thalamus in zebrafish? Frontiers in Neuroscience, 6(64), Feb 2012.


This Agreement between Imperial College London -- Gaëlle Chapuis ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

<table>
<thead>
<tr>
<th>License Number</th>
<th>432440882220</th>
</tr>
</thead>
<tbody>
<tr>
<td>License date</td>
<td>Apr 08, 2018</td>
</tr>
<tr>
<td>Licensed Content Publisher</td>
<td>Elsevier</td>
</tr>
<tr>
<td>Licensed Content Publication</td>
<td>Hearing Research</td>
</tr>
<tr>
<td>Licensed Content Title</td>
<td>Active listening: Task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex</td>
</tr>
<tr>
<td>Licensed Content Author</td>
<td>Jonathan Fritz, Mounya Elhilali, Shihab Shamma</td>
</tr>
<tr>
<td>Licensed Content Date</td>
<td>Aug 1, 2005</td>
</tr>
<tr>
<td>Licensed Content Volume</td>
<td>206</td>
</tr>
<tr>
<td>Licensed Content Issue</td>
<td>1-2</td>
</tr>
<tr>
<td>Licensed Content Pages</td>
<td>18</td>
</tr>
<tr>
<td>Start Page</td>
<td>159</td>
</tr>
<tr>
<td>End Page</td>
<td>176</td>
</tr>
<tr>
<td>Type of Use</td>
<td>reuse in a thesis/dissertation</td>
</tr>
<tr>
<td>Intended publisher of new work</td>
<td>other</td>
</tr>
<tr>
<td>Portion</td>
<td>figures/tables/illustrations</td>
</tr>
<tr>
<td>Number of figures/tables/illustrations</td>
<td>1</td>
</tr>
<tr>
<td>Format</td>
<td>both print and electronic</td>
</tr>
<tr>
<td>Are you the author of this Elsevier article?</td>
<td>No</td>
</tr>
<tr>
<td>Will you be translating?</td>
<td>No</td>
</tr>
<tr>
<td>Original figure numbers</td>
<td>Fig. 2</td>
</tr>
<tr>
<td>Title of your thesis/dissertation</td>
<td>Functional determinants of spectrotemporal selectivity in mouse auditory cortex</td>
</tr>
<tr>
<td>Publisher of new work</td>
<td>Imperial College London</td>
</tr>
<tr>
<td>Author of new work</td>
<td>Dr Paul Chadderton</td>
</tr>
<tr>
<td>Expected completion date</td>
<td>May 2018</td>
</tr>
<tr>
<td>Estimated size (number of pages)</td>
<td>1</td>
</tr>
<tr>
<td>Requestor Location</td>
<td>Imperial College London</td>
</tr>
<tr>
<td></td>
<td>South Kensington Campus</td>
</tr>
<tr>
<td></td>
<td>Royal School of Mines</td>
</tr>
<tr>
<td></td>
<td>Exhibition Road</td>
</tr>
<tr>
<td></td>
<td>London, SW7 2AZ</td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
</tr>
<tr>
<td></td>
<td>Attn: Imperial College London</td>
</tr>
<tr>
<td>Publisher Tax ID</td>
<td>GB 494 6272 12</td>
</tr>
<tr>
<td>Total</td>
<td>0.00 GBP</td>
</tr>
</tbody>
</table>

https://r100.copyrights.com/AppDispatch/Elsevier

Figure II.1: Permission to use images in Fritz et al. [99] (page 1/5).
INTRODUCTION
1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection
with completing this licensing transaction, you agree that the following terms and conditions
apply to this transaction (along with the Billing and Payment terms and conditions
established by Copyright Clearance Center, Inc. (*CCC*), at the time that you opened your
RightLink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS
2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to
the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has
appeared in our publication with credit or acknowledgement to another source, permission
must also be sought from that source. If such permission is not obtained then that material
may not be included in your publication/copies. Suitable acknowledgement to the source
must be made, either as a footnote or in a reference list at the end of your publication, as
follows:
"Reprinted from Publication title, Vol edition number, Author(s), Title of article / title of
chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE
SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The
Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with
permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which
permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be
altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions
and/or any other alterations shall be made only with prior written authorization of Elsevier
Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made
to any Lancet figures/tables and they must be reproduced in full.
6. If the permission fee for the requested use of our material is waived in this instance,
please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the
combination of (i) the license details provided by you and accepted in the course of this
licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment
terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed
immediately upon issuance of the license at the end of the licensing process for the
transaction, provided that you have disclosed complete and accurate details of your proposed
use, no license is finally effective unless and until full payment is received from you (either
by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If
full payment is not received on a timely basis, then any license preliminarily granted shall be
deemed automatically revoked and shall be void as if never granted. Further, in the event
that you breach any of these terms and conditions or any of CCC's Billing and Payment
terms and conditions, the license is automatically revoked and shall be void as if never
granted. Use of materials as described in a revoked license, as well as any use of the
materials beyond the scope of an unrevoked license, may constitute copyright infringement
and publisher reserves the right to take any and all action to protect its copyright in the
materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed
material.
10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and
their respective officers, directors, employees and agents, from and against any and all
claims arising out of your use of the licensed material other than as specifically authorized
pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed,
assigned, or transferred by you to any other person without publisher's written permission.
12. No Amendment Except in Writing: This license may not be amended except in a writing
signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

Figure II.2: Permission to use images in Fritz et al. [99] (page 2/5).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC’s Billing and Payment terms and conditions. These terms and conditions, together with CCC’s Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC’s Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

**LIMITED LICENSE**

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at: [http://www.sciencedirect.com/science/journal/xxxxx](http://www.sciencedirect.com/science/journal/xxxxx) or the Elsevier homepage for books at [http://www.elsevier.com](http://www.elsevier.com). Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at [http://www.elsevier.com](http://www.elsevier.com). All content posted to the web site must maintain the copyright information line on the bottom of each image.

**Posting licensed content on Electronic reserve:** In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

**Preprints:** A preprint is an author’s own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancements etc.). Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts:** An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

https://www.elsevier.com/ApplSupport/licensing
Authors can share their accepted author manuscripts:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- after the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (PJA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review, co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. For book authors the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. Posting to a repository: Authors are permitted to post a summary of their chapter only in their institution's repository.

19. Thesis/Dissertation: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

https://www.elsevier.com/AuthorCenterServices

Figure II.4: Permission to use images in Fritz et al. [99] (page 4/5).
Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any re-use of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated. The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC-SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC-ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. Other Conditions:

v1.9

Questions? customerservice@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.
Figure II.6: Permission to use images in Harris and Mrsic-Flogel[25] (page 1/3).
Terms and Conditions

Springer Nature Terms and Conditions for RightsLink Permissions

Springer Customer Service Centre GmbH (the Licensor) hereby grants you a non-exclusive, world-wide licence to reproduce the material and for the purpose and requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license re-use of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Where print only permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.

3. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

4. A licence for ‘post on a website’ is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.

5. Where ‘reuse in a dissertation/thesis’ has been selected the following terms apply:

Print rights for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatures to the STM Permissions Guidelines http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.

7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

8. The Licensor’s permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/table/illustrations or abstract, and must be hyperlinked to the journal/book’s homepage. Our required acknowledgement format is in the Appendix below.

9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.

10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor’s approval. However, the adaptation should be credited as shown in Appendix below.

Appendix — Acknowledgements:

For Journal Content:
Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

Figure II.7: Permission to use images in Harris and Mrsic-Flogel (25) (page 2/3).
Figure II.8: Permission to use images in Harris and Mrsic-Flogel[25] (page 3/3).
Figure II.9: Permission to use images in Jaramillo and Zador\textsuperscript{[61]} (page 1/3).
Figure II.10: Permission to use images in Jaramillo and Zador [61] (page 2/3).
Figure II.11: Permission to use images in Jaramillo and Zador. (page 3/3).
Figure II.12: Permission to use images in Lee et al.\textsuperscript{[174]} (page 1/3).
Figure II.13: Permission to use images in Lee et al.\cite{lee2023} (page 2/3).
Figure II.14: Permission to use images in Lee et al.\textsuperscript{[124]} (page 3/3).
Figure II.15: Permission to use images in Shamma et al. [69] (page 1/5).
INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your RightLink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you.
which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. Posting licensed content on any Website: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text link must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu. Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com. All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. For Journal authors: the following clauses are applicable in addition to the above:

Preprints: A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.). Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

https://www.copyright.com/AppDisclaimers

Figure II.17: Permission to use images in Shamma et al. [69] (page 3/5).
• immediately
  o via their non-commercial person homepage or blog
  o by updating a preprint in arXiv or RePliC with the accepted manuscript
  o via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  o directly by providing copies to their students or to research collaborators for their personal use
  o for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
• After the embargo period
  o via non-commercial hosting platforms such as their institutional repository
  o via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

• link to the formal publication via its DOI
• bear a CC-BY-NC-ND license - this is easy to do
• if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

**Published Journal article (PIA):** A published journal article (PIA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite and use the best available version.

Theses and dissertations which contain embedded PIAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

**Gold Open Access Articles:** May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. For book authors the following clauses are applicable in addition to the above:

Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PIAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

**Elsevier Open Access Terms and Conditions**

https://www.elsevier.com/termsandconditions

---

Figure II.18: Permission to use images in Shamma et al. [69] (page 4/5).
Figure II.19: Permission to use images in Shamma et al. [69] (page 5/5).
This Agreement between Imperial College London -- Gaelle Chapuis ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number: 4278820505166
License date: Jan 30, 2018
Licensed Content Publisher: Springer Nature
Licensed Content Publication: Biomechanics and Modeling in Mechanobiology
Licensed Content Title: Residual stress in the adult mouse brain
Licensed Content Author: Gang Xu, Philip V. Bayly, Larry A. Taber
Licensed Content Date: Jan 1, 2008
Licensed Content Volume: 8
Licensed Content Issue: 4
Type of Use: Thesis/Dissertation
Requestor type: academic/university or research institute
Format: print and electronic
Portion: figures/tables/illustrations
Number of figures/tables/illustrations: 1
Will you be translating?: no
Circulation/distribution: <501
Author of this Springer Nature content: no
Title: Functional determinants of spectrotemporal selectivity in mouse auditory cortex
Instructor name: Dr Paul Chadderton
Institution name: Imperial College London
Expected presentation date: May 2018
Order reference number: 1
Parts: Mouse (Fig 1a)
Requestor Location: Imperial College London
South Kensington Campus
Royal School of Mines
Exhibition Road
London, SW7 2AZ
United Kingdom
Attn: Imperial College London
Billing Type: Invoice
Billing Address: Imperial College London
South Kensington Campus
Royal School of Mines
Exhibition Road
London, United Kingdom SW7 2AZ
Attn: Imperial College London
Total: 0.00 GBP

Figure II.20: Permission to use images in Xu et al. [178] (page 1/3).
Figure II.21: Permission to use images in Xu et al. [178] (page 2/3).
Figure II.23: Permission to use images in Zhang et al. [9] (page 1/3).
Appendix — Acknowledgements:

For Journal Content:
Reprinted by permission from [the Licenser]: [Journal Publisher] [JOURNAL NAME] [REFERENCE CITATION] (Article name, Author(s) Name), [COPYRIGHT] (year of publication)
Figure II.25: Permission to use images in Zhang et al. [9] (page 3/3).
Appendix

I.1 Supplementary Information

I.1.1 Assessing contribution of functional neuronal properties to slow direction selectivity in A1

Adult data

To test the contribution of different neural properties in predicting the DSI, I generated linear multi-variable models based on the cell properties recorded in adult mice. The importance of each property in predicting the DSI was assessed via two parameters, namely the absolute value of the normalised coefficient, and the proportional reduction of error (PRE) generated when adding the property as predictor into the model (see also section 2.8.8 Modelling of experimental data). The performance of the model in predicting the DSI was assessed via the adjusted R-square, which takes into account the number of parameters used. Firstly, I built a model encompassing all neural properties judged relevant (n=10 properties, namely the firing rate of the ON response to tones of CFon frequency at 60dB, the firing rate of the OFF response to tones of CFoff frequency at 60dB, the Fano factor of each of these two responses, the spontaneous firing rate and associated Fano factor, the bandwidths of ON and OFF RFs (measured at 30dB), the percentage overlap between ON and OFF RFs, and the octave difference between ON and OFF CFs). This model was called the Full model. The adjusted R-square of the Full model was 0.279, which was worse than the adjusted R-square of the single-variable model built using the ON-OFF CF difference as sole predictor (adj. R-square=0.348). It is important to highlight that in the Full model, the difference between ON and OFF CF had the largest coefficient value (see table 5.1), and that it accounted for a large portion of the error reduction.

<table>
<thead>
<tr>
<th>Prop.</th>
<th>FRON</th>
<th>FOROFF</th>
<th>FFON</th>
<th>FFOFF</th>
<th>FRSP</th>
<th>FFSP</th>
<th>BCON</th>
<th>BCOFF</th>
<th>Overlap</th>
<th>CFdiff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff.</td>
<td>-0.021</td>
<td>0.007</td>
<td>-0.012</td>
<td>0.001</td>
<td>0.025</td>
<td>0.002</td>
<td>-0.090</td>
<td>0.151</td>
<td>-0.194</td>
<td>-0.355</td>
</tr>
<tr>
<td>PRE</td>
<td>1.257</td>
<td>0.100</td>
<td>0.365</td>
<td>0.002</td>
<td>0.100</td>
<td>0.009</td>
<td>0.733</td>
<td>2.536</td>
<td>0.774</td>
<td>0.907</td>
</tr>
</tbody>
</table>

Table 5.1: Normalised coefficients and proportional reduction of error for the Full model (Adult). PRE for variables considered highly predictive are marked in yellow. Adj. R-square 0.279.

The lower R-square in the Full model could originate from spurious variables which do not add any predictive information. By iteratively adding and removing variables from the Full model, the Optimal model was found to be composed of the following predictors: the firing rate of the ON response to tones of CFon frequency at 60dB, the bandwidth of the OFF RF, the percentage overlap between ON and OFF RFs, and the octave difference between ON and OFF CFs. The coefficients and PRE for the predictors of this model are presented in table 5.2. This model had an adjusted R-square of 0.419. Any addition or removal of predictor to this model decreased the adjusted R-square.

Young data

To further show that the properties found in Young neurons could not account for the DSI, I developed a Full linear multi-variable model (using the same parameters as for Adult neurons). In the case of Young neurons, the difference between ON-OFF CFs was not a good predictor of the
DSI (see table 5.3). Instead, the percentage overlap was found to be the variable with the best PRE. However, the percentage overlap was not a good predictor of DSI when used as single-variable model (adjusted R-square 0.059). I thus concluded that there was no clear linear relationship between DSI and any of the tested neuronal properties in Young mice.

<table>
<thead>
<tr>
<th>Prop.</th>
<th>FR\textsubscript{ON}</th>
<th>FR\textsubscript{OFF}</th>
<th>FF\textsubscript{ON}</th>
<th>FF\textsubscript{OFF}</th>
<th>FR\textsubscript{SP}</th>
<th>FF\textsubscript{SP}</th>
<th>BW\textsubscript{ON}</th>
<th>BW\textsubscript{OFF}</th>
<th>Overlap</th>
<th>CF\textsubscript{diff}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff.</td>
<td>-0.004</td>
<td>0.008</td>
<td>-0.010</td>
<td>0.028</td>
<td>0.207</td>
<td>0.025</td>
<td>-0.120</td>
<td>0.066</td>
<td>-0.638</td>
<td>-0.151</td>
</tr>
<tr>
<td>PRE</td>
<td>0.008</td>
<td>0.015</td>
<td>0.055</td>
<td>0.956</td>
<td>0.493</td>
<td>1.080</td>
<td>1.111</td>
<td>0.173</td>
<td>4.370</td>
<td>0.077</td>
</tr>
</tbody>
</table>

Table 5.2: Normalised coefficients and proportional reduction of error for the Optimal model (Adult). PRE for variables considered highly predictive are marked in yellow. Adj. R-square 0.419.

<table>
<thead>
<tr>
<th>Prop.</th>
<th>FR\textsubscript{ON}</th>
<th>FR\textsubscript{OFF}</th>
<th>FF\textsubscript{ON}</th>
<th>FF\textsubscript{OFF}</th>
<th>FR\textsubscript{SP}</th>
<th>FF\textsubscript{SP}</th>
<th>BW\textsubscript{ON}</th>
<th>BW\textsubscript{OFF}</th>
<th>Overlap</th>
<th>CF\textsubscript{diff}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff.</td>
<td>-0.012</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>0.123</td>
<td>-0.169</td>
</tr>
<tr>
<td>PRE</td>
<td>0.380</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>1.526</td>
<td>0.517</td>
</tr>
</tbody>
</table>

Table 5.3: Normalised coefficients and proportional reduction of error for the Full model (Young). PRE for variables considered highly predictive are marked in yellow. Adj. R-square -0.339.
I.2 Supplementary Figures

Figure I.26: The direction selectivity index was invariant to the underlying measure. a. Toy example of a cell’s PSTH in response to FM sweep. The grey box indicates the sound presentation, and the vertical dashed bars indicate when the response is considered. To measure the response, either the maximum instantaneous firing rate was considered (max FR) or the number of spikes (N spk). Note that computing the DSI using the N spk is equivalent to using the area under the PSTH. The red area marks the time zone following sound offset. In b-c, DSIs were measured using the full time window (i.e. time between dashed bars), whilst in d-e DSIs were measured using the activity during the sound presentation only (i.e. the activity in the red zone was not considered). b. The DSI measured via the mean number of spk (y-axis) was strongly correlated with the DSI measured via the maximum firing rate (x-axis). The Pearson correlation coefficient (r) and associated p-value (p) are displayed on top of the panel. DSIs were measured at the FM speed 2.2 oct/s. c. The relationship between the DSI and CF ON is conserved when using the DSI based on the maximum firing rate. The adjusted R-square (R^2) of the linear interpolation (red curve) and Pearson correlation coefficient (r, p<0.05) are indicated within the panel. d-e. Same as in b-c, but using a restricted time window to compute the DSIs (i.e. not comprising the red time zone in a).
Figure I.27: Autocorrelograms and waveforms of example single cells recorded to quantify ON-OFF CF difference and DSI in vivo. Autocorrelograms (left) and spike waveforms (right; 50 overlaid traces, mean in black) for single units shown in figure 3.3a (a), figure 3.3b (b), and figure 3.9a (c).
Young ON/OFF cells

Figure L28: Poor relationship between DSI and CF ON for Young ON/OFF A1 neurons. DSI versus CF ON for Young ON/OFF cells at each FM speed (marked on top of the panel). Each dot represents a cell (n = 20). The red line represents a linear interpolation, from which the adjusted R-square is displayed on top of the panel ($R^2$). The Pearson correlation coefficient ($r$) between DSI and CF ON is marked on top of the panel, with stars indicating the significance level.
Figure I.29: Measured and predicted DSI values of Young ON/OFF A1 neurons. Distributions of measured (blue) and predicted (purple) direction selectivity index (DSI) for Young A1 neurons exhibiting ON/OFF receptive fields (n = 20). Measured DSI were computed at 2.2 oct/s. Predicted DSI were computed using the Optimal model in Adult (see table 5.2), but inserting the predictor values of Young neurons. The dashed line represents the median of each distribution.
Figure I.30: Relationship between DSI and CF ON for Multi-unit data recorded in awake passive-listening mice.

a. Multi-unit Baseline

- 2.2 oct/s: $R^2 = 0.189$, $r = -0.45^{***}$
- 4.4 oct/s: $R^2 = 0.081$, $r = -0.31^*$
- 8.8 oct/s: $R^2 = 0.029$, $r = -0.21$
- 17.5 oct/s: $R^2 = 0.012$, $r = 0.06$
- 35 oct/s: $R^2 = 0.006$, $r = -0.10$
- 70 oct/s: $R^2 = 0.051$, $r = -0.26^*$
- 140 oct/s: $R^2 = 0.102$, $r = -0.34^{**}$

b. Multi-unit Post-CNO

- 2.2 oct/s: $R^2 = 0.026$, $r = 0.20$
- 4.4 oct/s: $R^2 = -0.011$, $r = 0.07$
- 8.8 oct/s: $R^2 = -0.010$, $r = -0.07$
- 17.5 oct/s: $R^2 = -0.013$, $r = -0.05$
- 35 oct/s: $R^2 = 0.006$, $r = -0.15$
- 70 oct/s: $R^2 = 0.005$, $r = -0.14$
- 140 oct/s: $R^2 = -0.007$, $r = -0.09$

Figure I.30: Relationship between DSI and CF ON for Multi-unit data recorded in awake passive-listening mice. a. DSI versus CF ON for Multi-unit (Baseline) at each FM speed (marked on top of the panel). Each dot represents a recording site ($n = 65$). The red line represents a linear interpolation, from which the adjusted $R$-square is displayed on top of the panel ($R^2$). The Pearson correlation coefficient ($r$) between DSI and CF ON is marked on top of the panel, with stars indicating the significance level. b. Same as in a, but for Multi-unit data recorded Post-CNO (same sites as in a).
Figure I.31: Spike in LFP power as a marker of hM4i activation in vivo. 

a. Example LFP spectrograms in Baseline and Post-CNO condition (same animal in both conditions). Note that in Post-CNO, spikes of high power emerged (visible as vertical red lines in the spectrogram).

b. Average change in power spectral density (PSD) in Post-CNO compared to Baseline. A strong increase in most frequency band is visible for the hM4i- (N = 10) but not for the sham- (Control; N = 3) mice. n = 6 and n = 11 recordings were used to generate the results in Control and hM4i respectively. Note that the Control average remains around 1 for all frequency bands.

c. The power distribution was not markedly different between Control and hM4i in Post-CNO. The traces indicate example power distributions measured in different scenarios (Post-CNO during either awake passive listening (N = 2 hM4i; N = 2 Control) or behaviour (N = 2 hM4i).
Figure I.32: Behavioural performance of hm4i-PV+ mice. a. Training curve for each individual hm4i-PV+ mouse (N = 10; ID marked in bold, e.g. M116). Performance is calculated as $d'$ over single session. The colored dots mark sessions with manipulations (see b for color labels). b. Average performance for each manipulation type: Baseline, Post-CNO, Recording (Rec.), and Post-CNO + Rec. The average is taken over all sessions of a type. Note that whilst the sessions were measured in N = 10 mice, some manipulations occurred multiple times in single animals, and some others did not occur in all mice (see a).
Figure I.33: Multi-unit activity was not modulated by task outcome. a. Responses of an example MU site to UP sweeps during correction rejection (CR) and false alarm (FA) trials (same site as in figure 3.15). The grey area indicates sound presentation. Blue and red curves represent PSTHs in response to UP sweep in CR and FA trials respectively. b. Responses of the MU site presented in a to DOWN sweeps during Miss and Hit trials. The grey area indicates sound presentation. Blue and green curves represent PSTHs in response to DOWN sweep in Miss and Hit trials respectively. c. Relationship between modulation index (MI) and BF. MI was measured using CR and FA trials. A positive MI indicates higher activity during FA trials. The MI was measured using either the maximum firing rate (Max FR; top panel) or the total number of spike (Nspk; bottom panel) in the time window marked at the bottom of the panels (0-100ms). The Pearson correlation coefficient between MI and BF is marked inside each panel, with stars indicating the confidence level. Filled red dots indicate measurements during Post-CNO. d. Same as in c, but the MI was measured using Miss and Hit trials. A positive MI indicates higher activity during Hit trials.
Figure I.34: Behavioural performance of hm3-PV+ mice. 

a. Training curve for each individual hm3i-PV+ mouse (N = 4; ID marked in bold, e.g. M137). Performance is calculated as $d'$ over single session. The colored dots mark sessions with manipulations (see b for color labels). 

b. Average performance for each manipulation type: Baseline, Post-CNO, Recording (Rec.), Post-CNO + Rec, Saline intra-cortical injection (Saline) and Muscimol intra-cortical injection (Muscimol). The average is taken over all sessions of a type. Note that whilst the sessions were measured in N = 4 mice, some manipulations occurred multiple times in single animals, and some others did not occur in all mice (see a).
Figure I.35: Post mortem microscopy confirmed the expression of DREADD in auditory cortex of mice used in behaviour.

Example coronal brain slices retrieved from four mice used in behaviour (ID and virus type used written on each slice). mCherry expressing cells (red) were imaged using a 40x objective (black square: imaging zone), and patched onto the anatomic image (x10 objective, grey) for display.
Figure I.36: Future work: Assessing the selectivity of cortical inputs. 

- **a.** Example frequency tuning of a cell recorded in the Medial Geniculate Body (MGB) of an awake mouse in passive listening condition. This cell displayed excitatory ON responses. In the top panel, the evoked firing rate is displayed. In the bottom panel, the response is z-scored compared to baseline activity. A positive z-score indicates an increase in firing compared to baseline.

- **b.** Same as in a, but for a cell displaying inhibitory ON responses. A negative z-score indicates a decrease in firing compared to baseline.

- **c.** Percentage of MGB cells with excitatory response passing criterion for FM direction discrimination at each FM speed (red, n = 58). The orange and brown curves represent the percentage of cells passing criterion that are UP- and DOWN-selective respectively.

- **d.** Same as in c, but for cells with inhibitory responses.
Figure I.37: Future work: Mapping ON/OFF responses in auditory cortex using Calcium imaging. a. Schematic and example specimen of cranial window developed for imaging Calcium transients in auditory cortex. b. Wide field signals exhibited the strongest ON responses for low frequency tones (here, 3.5 kHz at 70 dB SPL). Top inset: Example time course of average response intensity (30 repeats) of a ~200 µm² pixel to low frequency tone (3.5 kHz, 70 dB SPL). The grey box indicates the time window used to measure the mean response. Bottom inset: Mean response intensity map over the whole cranial window. The white color denotes high response intensity. c. Responses to low frequency tones were spatially organized (N = 2 mice displayed). d. The organisation seen in c was consistent with previous literature. Image modified from Tsukano et al. [18].

Figure I.38: Beneficial future changes in the FM discrimination behavioural paradigm. a. The current association rule is conserved (UP = Go, DOWN = No-Go). b. Schematic of stimuli during a protocol using different modulation rates. c. Schematic of stimuli during a protocol using different start frequency for UP and DOWN sweeps.
Figure I.39: *Post mortem* microscopy confirmed the electrode placement in prefrontal cortex during recording in behaviour (oddball task T1). Coronal brain section for each mouse (ID: 2,3,4 reported on top of the panel). The electrode was coated with Di-I (magenta).

Figure I.40: *Post mortem* microscopy confirmed the electrode placement in auditory cortex during recording in behaviour (oddball task T1). Coronal brain section for each mouse (ID: 2,3,4 reported on top of the panel). The electrode was coated with Di-I (magenta).
Figure I.41: Comprehensive schematic of analytical procedure undertaken to compute the difference in LFP phase distribution between Early and Late trial blocks (relates to figure 4.16). 

a. Schematic of an average LFP trace (black line) and the LFP phase values (colored lines) for each of the corresponding trials for Early (top) and Late (bottom) trial blocks. The phase distributions in the Early and Late blocks are compared at different times points (here two are marked, see circles). 

b. Top panel: Normalised phase distributions in the Early and Late blocks at each time point (circles). Bottom panel: The change between the phase distributions in Early and Late blocks is assessed by taking the difference between the two normalised distributions ($\text{Diff} = P_{\text{Early}} - P_{\text{Late}}$). Positive Diff values thus indicate an increase in Early compared to Late block. The overall (i.e., over all times points) 5% and 95% percentile values are marked by a triangle (blue and red respectively). As the LFP trace oscillates, it is expected to see a procession in percentiles phase values across time (see shift in triangle positions). 

c. Difference values measured in b are reported at each time points. Blue-to-red gradient indicates low-to-high percentile. 

d. The aim of the analysis is to assess whether a specific LFP phase is enhanced in Early compared to Late trial at the time of the first possible Early target. This corresponds to the phase value at which there is an augmentation in phase distribution difference at the time of the target. Thus, only the high percentiles are taken into consideration for this analysis, as the difference is measured as $P_{\text{Early}} - P_{\text{Late}}$ (see b). The phase values are unwrapped, and the difference values (measured as in b) are used to generate a linear interpolated curve over time (see section 2.8.7 Local field potential and multi-unit analysis for method details). The interpolation curve is used to estimate the enhanced phase.