

**Repetition suppression in the auditory cortex:
an interdisciplinary investigation**

Thesis for the degree of

doctor rerum naturalium (Dr. rer. nat.)

approved by the Faculty of Natural Sciences
of Otto von Guericke University Magdeburg

written by **Nina Härtwich, M.Sc.**

born on 16 July 1990 in Dresden, Germany

Examiners: PD Dr. Reinhard König
Prof. Dr. Jonas Obleser

submitted on 26 August 2024

defended on 04 March 2025



Abstract

"Repetition suppression in the auditory cortex: an interdisciplinary investigation"

Nina Härtwich, M.Sc.

Auditory sensory memory is essential for the interpretation of a sound in the context of the immediate past. While memory traces can be observed behaviourally, the underlying mechanisms in the auditory cortex (AC) are more elusive and can often only be observed indirectly. One such indirect observation is repetition suppression (RS), the attenuation of neural responses when a stimulus is presented repeatedly. This phenomenon indicates that the auditory system contains memory traces of recent stimulation that affect responses to incoming stimuli.

While the RS phenomenon is well established, the neuronal underpinning and factors that modulate its lifetime are not fully understood. This thesis investigates the emergence of RS in the AC through computational modelling and in-vivo measurements. Moreover, a robust analysis pipeline for the determination of in-vivo RS lifetimes and an automated approach for the optimisation of computational AC models are introduced.

Simulation results from a newly created computational model of the gerbil AC predict that the lifetime of RS does not simply reflect recovery from short-term synaptic depression (STSD). Instead, RS is a result of the interplay between STSD dynamics and network connectivity patterns. The same prediction also resulted from simulations with a simplified computational model of the human AC. When network connection strengths were altered, the lifetime of RS changed. Moreover, the gerbil AC model demonstrated that network interactions cause variations in RS lifetime as a function of stimulus audio-frequency. Electrophysiological recordings revealed audio-frequency-specific RS lifetimes in the AC of four out of six gerbils. Moreover, the RS lifetimes deduced from the intracortical recordings were shorter than lifetimes deduced from extracortical measurements of activity in the human AC.

In conclusion, the findings suggest that RS in the AC is a network effect. The resulting lifetime variations across neural populations and the audio-frequency of the stimulus might play a functional role in the context of temporal binding of sounds. The variation across species might reflect that AC response dynamics are respectively tailored to temporal binding across the shorter time scales of gerbil vocalisations and the longer time scales of human speech.

Zusammenfassung

"Wiederholungsunterdrückung im Hörkortex: eine interdisziplinäre Untersuchung"

Nina Härtwich, M.Sc.

Ein auditives sensorisches Gedächtnis ist essentiell, um Hörreize im Kontext der unmittelbaren Vergangenheit zu interpretieren. Während Gedächtnisspuren auf der Verhaltensebene gut beobachtet werden können, sind die zugrunde liegenden Mechanismen im Hörkortex schwerer zu fassen und meist nur indirekt zu beobachten. Eine solche indirekte Beobachtung ist Wiederholungsunterdrückung (WU), eine Verminderung neuronaler Aktivität bei wiederholtem Auftreten eines Stimulus. Dieses Phänomen deutet darauf hin, dass im auditiven System Gedächtnisspuren von vorhergehenden Stimuli existieren, die die neuronale Antwort auf folgende Stimuli beeinflussen.

WU ist ein gut beschriebenes Phänomen, jedoch sind die neuronalen Grundlagen sowie die Faktoren, die die Verfallszeit der Unterdrückung beeinflussen, nicht abschließend geklärt. Diese Dissertation nutzt Computersimulationen und In-vivo-Messungen, um das Auftreten von WU im Hörkortex zu untersuchen. Zusätzlich werden eine robuste Analyse-Pipeline für die Ermittlung von Verfallszeiten in vivo sowie ein Ansatz für die automatisierte Optimierung von Computermodellen des Hörkortex vorgestellt.

Die mit einem neuen Modell des Gerbil-Hörkortex erzielten Simulationsergebnisse sagen voraus, dass Verfallszeiten der WU nicht allein die Verfallszeiten synaptischer Kurzzeitdepression spiegeln. Stattdessen ergibt sich WU aus der Wechselwirkung zwischen synaptischer Kurzzeitdepression und dem Verbindungsmuster des neuronalen Netzwerkes. Dieselbe Prognose resultierte auch aus Simulationen mit einem vereinfachten Computermodell des menschlichen Hörkortex. Änderungen der Stärke der Netzwerkverbindungen führten zu einer Veränderung der WU-Verfallszeit. Zusätzlich war im Gerbil-Modell zu beobachten, dass aufgrund von bestimmten Netzwerk-Interaktionen die Verfallszeit abhängig von der Audiofrequenz des Stimulus ist. Elektrophysiologische Messungen im Hörkortex offenbarten audiofrequenz-spezifische Verfallszeiten bei vier von sechs Gerbilen. Außerdem waren die aus den intrakortikalen Messungen ermittelten Verfallszeiten kürzer als die von extrakortikalen Messungen abgeleiteten Verfallszeiten für den menschlichen Hörkortex.

Die Ergebnisse deuten somit darauf hin, dass WU im Hörkortex ein Netzwerkeffekt ist. Die sich daraus ergebenden Variationen der Verfallszeiten, über neuronale Populationen und Stimulus-Audiofrequenzen hinweg, könnten eine Rolle bei der Verarbeitung von längeren und komplexeren Tonsequenzen spielen. Die Variation zwischen den Spezies könnte bedeuten, dass die Reaktionsdynamik des jeweiligen Hörkortex an die Verarbeitung der kürzeren Gerbil-Laute beziehungsweise der längeren Sequenzen menschlicher Sprache angepasst ist.

Acknowledgements

PhD students are traditionally viewed as lone fighters, requiring a ‘my project, my data, my results’-attitude. A thesis carrying the words ‘interdisciplinary investigation’ in its title would have been impossible in such a climate. I had the privilege of being surrounded by colleagues who do not subscribe to the lone-wolf approach. To all of you: Thank you for your willingness to share, thank you for spirited discussions, and thank you for your enthusiasm for collaborative work. Special thanks go to my supervisors, who made this thesis possible and supported me through the highs and lows of my PhD journey. This journey also required encouragement and help outside of working hours. Thank you to all the wonderful people in my social support system. I am so grateful to have you in my life.

Danksagung

Menschen, die eine Promotion anstreben, gelten traditionell als Einzelkämpfer, die eine ‘mein Projekt, meine Daten, meine Ergebnisse’-Attitüde benötigen. Eine Dissertation mit den Worten ‘interdisziplinäre Untersuchung’ im Titel wäre in einem solchen Klima nicht möglich gewesen. Ich hatte das Privileg, von Kolleginnen und Kollegen umgeben zu sein, die keinen Einzelkämpfer-Ansatz verfolgen. An euch alle: Danke für eure Bereitschaft, zu teilen, danke für die spannenden Diskussionen und danke für euren Enthusiasmus bezüglich kollaborativer Arbeit. Besonderer Dank gilt den beiden Betreuern meiner Promotion. Sie haben diese Dissertation ermöglicht und mich über Höhen und Tiefen hinweg stets unterstützt. Auch außerhalb der Arbeitszeit waren Hilfe und Zuspruch essentiell. Danke an all die wunderbaren Menschen in meinem sozialen Netzwerk, es ist ein Geschenk, dass ihr Teil meines Lebens seid.

Table of contents

Abstract	iii
Acknowledgements	vii
Table of contents	ix
List of Abbreviations	xiii
List of Equations	xv
List of Figures	xvii
List of Tables	xix
I Introduction	1
Memory traces in the human auditory cortex	2
Quantification of repetition suppression lifetimes	4
Memory traces along the auditory pathway of other species	5
Mechanisms behind cortical repetition suppression	6
Insights from in-vivo recordings	6
Insights from computational modelling	7
Animal models of human auditory processing	9
Research questions addressed in this thesis	10
Collaborations	11
II Repetition suppression in the gerbil auditory cortex	13
1 Repetition suppression in silico	13
1.1 Materials and Methods	14
1.1.1 Model dynamics	14
1.1.2 Structure of the anatomy-based model	17
1.1.3 Structure of the single-column model	25
1.1.4 Stimulation paradigms	26
1.1.5 Analysis of simulated data	27
1.2 Simulation results and predictions	36
1.2.1 Response behaviour of the single-column model	36
1.2.2 Response behaviour of the gerbil AC model	46
2 Repetition suppression in vivo	59
2.1 Materials and pre-established methods	59
2.1.1 Surgical procedure and electrophysiological measurements	59

2.1.2	Stimulation paradigms	60
2.1.3	Measures of neuronal activity	61
2.2	Preliminary assessment	63
2.2.1	Frequency response	63
2.2.2	Repetition suppression stabilisation	65
2.3	A new analysis pipeline for subject-specific statistical inferences	69
2.3.1	The standard approach to quantifying repetition-suppression lifetime	69
2.3.2	A bootstrap-based expansion of the standard approach	71
2.3.3	SOI-specific peak latencies	75
2.4	Final results and comparison with model predictions	77
2.4.1	Repetition suppression lifetime	77
2.4.2	Peak latency shift	77
2.4.3	Median peak latency	80
2.4.4	Preliminary conclusions	80
III	Refining strategies for repetition suppression studies	83
3	An analysis pipeline to verify the robustness of lifetime estimates	83
3.1	Materials and Methods	84
3.1.1	Subjects	84
3.1.2	Experimental paradigm	84
3.1.3	Data acquisition and pre-processing	85
3.1.4	Computation of SOI-specific ERFs	85
3.1.5	Bootstrapping	86
3.2	Results	89
3.2.1	Repetition suppression lifetime and saturation amplitude across hemispheres and subjects	89
3.2.2	Relationship between adaptation lifetime and saturation amplitude	91
3.2.3	Summary and outlook	93
4	An automated approach to improve auditory cortex models	94
4.1	Materials and Methods	95
4.1.1	Human AC model and simulation of ERFs	95
4.1.2	Experimental data and paradigm	97
4.1.3	Optimisation algorithm	98
4.1.4	Shortening runtimes via a SOI-specific reduction of stimulus repetitions	100
4.2	Results	102
4.2.1	Auditory cortex model optimisation	102

4.2.2	Predictions regarding area-specific repetition suppression lifetimes . . .	103
4.2.3	Preliminary conclusions	106
IV	Discussion	107
Q1:	How can the challenges of repetition suppression lifetime quantification be addressed?	107
Q2:	What does the lifetime of repetition suppression reflect?	110
Q3:	What is the functional relevance of repetition suppression as a network effect?	112
Q4:	Is the lifetime of repetition suppression species-specific?	118
Q5:	How could the gerbil auditory cortex model be refined?	121
	Closing remarks	123
	References	125
	Declaration of honour	137

List of Abbreviations

- ΔF** distance of the stimulus audio-frequency from the BF at the measurement site. 25
- A1** primary auditory field. 5
- AAF** anterior auditory field. 17
- AC** auditory cortex. 1
- AV** anteroventral field. 17
- AVREC** average rectified current source density. 62
- BF** best frequency, stimulus audio-frequency that elicits the strongest response. 23
- CSD** current source density. 61
- D** dorsal field. 17
- DP** dorsoposterior field. 17
- EEG** electroencephalography. 2
- ERF** event-related field. 2
- ERP** event-related potential. 2
- FR** frequency response. 60
- i.i.d.** independent and identically distributed. 68
- IQR** interquartile range. 90
- l.h.s.** left-hand side. 103
- LFP** local field potential. 61
- MEG** magnetoencephalography. 2
- MGN** medial geniculate nucleus. 5
- MGv** ventral division of the medial geniculate nucleus. 17
- MMN** mismatch negativity. 3
- r.h.s.** right-hand side. 27

RS repetition suppression. [1](#)

SNR signal-to-noise ratio. [86](#)

SOI stimulus-onset interval. [2](#)

SPL sound pressure level. [23](#)

SSA stimulus-specific adaptation. [5](#)

STSD short-term synaptic depression. [6](#)

STSF short-term synaptic facilitation. [113](#)

V ventral field. [17](#)

VM ventromedial field. [17](#)

VP ventroposterior field. [17](#)

List of Equations

1	$P_{\text{fit}}(\text{SOI})$ – peak response amplitude as a function of stimulus-onset interval	4
2	$g(u)$ – temporal evolution of AC model firing rates	14
3	$\dot{\mathbf{u}}(t)$, $\dot{\mathbf{v}}(t)$, and $\dot{\mathbf{q}}(t)$ – temporal evolution of AC model state variables	15
4	$k(i)$ – intrafield column index as a function of column index i	21
5	Δi – distance in terms of intrafield column indices	21
6	$W(i, j)$ – connection weight value as a function of source and target index	21
7	$s(i)$ – scaling factor, afferent input as a function of target index	24
8	$\dot{\mathbf{q}}_{\text{aff}}(t)$ – temporal evolution of synaptic efficacy of afferent connections	24
9	$\dot{\mathbf{i}}_{\text{aff}}(t)$ – temporal evolution of afferent synaptic currents	24
10	$\xi_{\text{leak},i}(t)$ – leak current at excitatory population of column i	27
11	$\xi_{\text{ee},i}(t)$ – excitatory synaptic current at excitatory population of column i	27
12	$\xi_{\text{ei},i}(t)$ – inhibitory synaptic current at excitatory population of column i	27
13	$\xi_{\text{inp},i}(t)$ – afferent synaptic current at excitatory population of column i	28
14	$\xi_{\text{net},i}$ – net transmembrane current flow at excitatory population of column i	28
15	$\mathbf{u}_{\text{ee},i}^{\text{area}}$ – weights of connections originating in area and targeting column i	28
16	$Q^{\text{area}}(t)$ – synaptic efficacy of connections originating in area	28
17	$\mathbf{u}^{\text{area}}(t)$ – state variables of connection sources located in area	28
18	$\xi_{\text{ee},i}^{\text{area}}(t)$ – net excitatory current induced via connections originating in area	29
19	$\xi_{\text{ee},i}^{\text{recurrent}}(t)$ – excitatory current induced via intracolumnar ee-connection	29
20	A_{stable} – mean peak response amplitude for stabilised RS	30
21	$\ell_{\text{fit}}(A(\text{SOI}))$ – peak latency as a function of SOI-specific peak amplitude	35
22	ρ_{peak} – ratio between ξ_{aff} and ξ_{ee} at time t_{peak}	42
23	$a_{i,j}(t)$ – activity level at connection from excitatory population j to i	52
26	$\overline{t_{\text{fire}}}$ weighted arithmetic mean of firing onset latency	57
27	$\text{CSD}(z)$ – current source density as a function of cortical depth	62
28	$\text{AVREC}(t)$ – temporal evolution of average rectified CSD	62
29	$\Delta\tau_{\text{SOI}}$ – hemispheric difference in RS lifetime	89
30	$R(t)$ – in-silico MEG signal	96
33	$\Phi(x)$ – fitness function, overlap between in-vivo and in-silico MEG signal	99
35	\mathcal{S} – similarity between responses across stimulus repetitions	101
36	$R_{\text{area}}(t)$ – area-specific components of in-silico MEG signal	103

List of Figures

1	Illustration showing why memory is essential in auditory processing	1
2	Representative example of an ERF from the AC of a human subject	3
3	Connectivity of column in the AC model	15
4	Structure and connectivity pattern of the gerbil AC	18
5	Illustration of the connection weight matrix W_{ee}	20
6	Tonotopic map of MGv field in gerbil AC model	23
7	Diagram illustrating the single-column model	25
8	Responses to the first stimulus within each block	30
9	Response stabilisation in the gerbil AC model	31
10	Analysis steps for the gerbil AC model	33
11	Analysis steps for the single-column model	34
12	FR functions of the single-column model	36
13	RS lifetimes and peak latencies in the single-column model	38
14	Single-column model response components	39
15	Effect of time constant $\tau_{rec,aff}$ on afferent current ξ_{aff}	41
16	Ratio between afferent and recurrent excitatory current	42
17	Firing onset latency in the single-column model	44
18	Time courses of synaptic current ξ_{ee} in the single-column model	45
19	Overview of FR functions for the gerbil AC model	47
20	Response behaviour in the primary auditory field of the gerbil AC model . . .	48
21	Gerbil AC model response components	50
22	Activation of connections in the gerbil AC model	52
23	Interaction between excitatory and inhibitory synaptic current	54
24	Firing onset latency across the gerbil AC network	56
25	Conversion of measured LFP into AVREC of CSD	61
26	Frequency response functions observed for the nine gerbils	64
27	Stabilisation of response as a function of stimulus index	66
28	Analysis pipeline for animal-specific statistical inferences	71
29	Comparison of single-trial responses and average response	72
30	Resampling of LFP profiles	73
31	Peak latency as a function of SOI-specific peak amplitude	76
32	Comparison of in-vivo and in-silico results	79
33	Procedure for determining RS lifetime τ_{SOI} from MEG data	89

34	Hemispheric difference in RS lifetime	90
35	Relationship between RS lifetime and saturation amplitude	92
36	Structure and connectivity pattern of the macaque AC	96
37	In-silico and in-vivo MEG signals observed for the human AC	98
38	SOI-specific in-silico MEG signals and associated peak response amplitudes	106
39	Audio-frequency-specific RS lifetimes as a basis for CS	117

List of Tables

1	Time constants and firing rate parameter values used in gerbil AC model	16
2	Gerbil AC model: column indices per field	19
3	Weight values used in the weight matrices W of the gerbil AC model	22
4	Scaled pre-synaptic firing rates as a function of ΔF	24
5	Time constants of recovery from STSD used in the single-column model	26
6	Area-specific column indices for the gerbil AC model	28
7	Overview of experimental conditions investigated in our simulations	32
8	Shape of FR functions for different σ_{ie} -values	46
9	Number of active connections in the gerbil AC model	53
10	Ranges of firing onset latency and peak amplitude across SOIs	57
11	SOI-specific structure of in-vivo stimulus blocks	69
12	Hemispheric differences – summary of test statistics	91
13	Parameter values used in the simplified human AC model	97
14	SOI-specific reduced number of stimuli to decrease EA runtime	102
15	Overview of median τ_{SOI} -values across the ten EA runs	104

Part I: Introduction

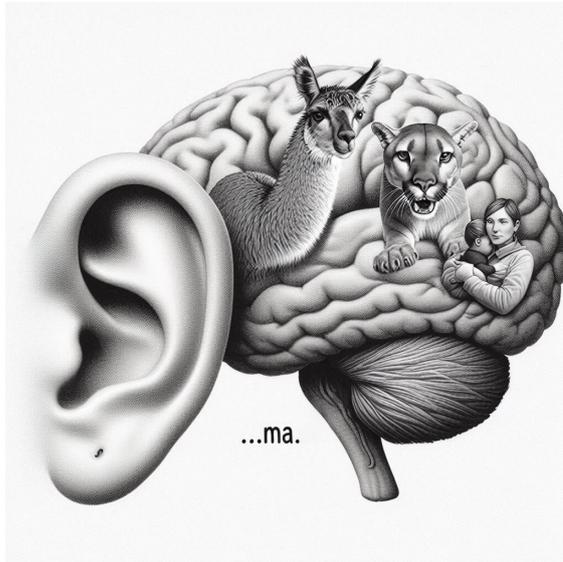


Figure 1: Someone just said ‘ma’ – is there a cuddly la-ma in sight, is a child looking for their ma-ma or did a dangerous pu-ma appear? Distinct temporal sequences of sounds modulate the meaning of a given sound. The soundscape at a specific moment in time can thus only be interpreted properly if the brain creates an auditory memory trace of the immediate past. Image generated using *Microsoft Designer AI*.

As exemplarily illustrated in Figure 1, auditory memory is essential for humans and other animals to make sense of speech or species-specific vocalisations and of other sounds with a temporal structure. It allows for incoming auditory stimuli to be perceived and interpreted in the seamless context of the immediate past. Auditory processing, from stimulus transduction to perception and interpretation, begins at the ear, where sound pressure waves are translated into neural signals that encode the audio-frequency spectrum of the incoming sound. These signals then travel through a number of subcortical stations before reaching the sensory area of the cortex tasked with further interpretation of the sound scape – the auditory cortex (AC). Higher-level cortical areas are also involved in this process. As summarised by [King et al. \(2018\)](#), several aspects of hearing, such as sound localisation and pitch discrimination, rely on processing steps that already take place subcortically, whereas the AC seems to perform processing steps that play a critical role in the perception and interpretation of sound *sequences*. As described above, this requires the formation of memory traces.

While memory traces in the brain can be observed at the behavioural level (for a critical review, see [Brady et al., 2023](#)), the underlying brain mechanisms, specifically basic mechanisms at the level of the AC, are more elusive and can often only be observed indirectly. One such indirect observation is repetition suppression (RS), also known as adaptation or habituation. RS refers to the phenomenon of neuronal responses diminishing in strength when a stimulus is repeated ([Megela and Teyler, 1979](#); [Malmierca et al., 2014](#)). It is also closely related to forward masking, which is the phenomenon that the response to a test tone is maximally

suppressed by a preceding probe tone if the probe and the test are identical (e.g. [Brosch and Schreiner, 1997, 2000](#)).

Given that the neural response to a specific stimulus depends on stimulus history rather than being fixed, RS indicates directly that the auditory system contains representations – memory traces – of the recent past that affect responses to incoming stimuli. But what exactly is the neural substrate of these memory traces, how are they formed, and which factors modulate their properties? While the RS phenomenon is a well-established feature of the brain (for a more general review, see [Webster, 2012](#), for a review regarding the auditory system, see [Pérez-González and Malmierca, 2014](#); [Malmierca et al., 2014](#)), the neuronal mechanisms *underlying* it are not fully understood. In this thesis, I attempt to explain the emergence of RS as a system property through computational modelling as well as extracranial and intracortical in-vivo measurements of neuronal responses. The following sections of the Introduction first review the existing literature on RS and then introduce the new studies carried out in the context of the thesis work.

Memory traces in the human auditory cortex

Barring some rare exceptions in the context of epilepsy treatment and other medical interventions (for a review, see [Mukamel and Fried, 2012](#)), ethical reasons limit in-vivo studies of the response behaviour of the human AC to non-invasive measurement methods. RS is observed via electroencephalography (EEG) and magnetoencephalography (MEG) by presenting blocks of repeating tones and varying the stimulus-onset interval (SOI) across the blocks. In recent years, the term ‘regular-SOI paradigm’ was coined for this type of stimulation pattern ([Hajizadeh et al., 2019](#); [Tomana et al., 2023](#)) and it will be used throughout this thesis. Responses to the regular-SOI paradigm are averaged across each stimulus block to reveal the event-related potentials (ERPs) and fields (ERFs). One example of such an ERF, from the left AC of a human subject, is shown in [Figure 2](#).

The long-latency components of ERPs and ERFs consist of a series of characteristic deflections called the P1(m), the N1(m), and the P2(m). The names refer to waves of opposite polarity – P for positive and N for negative in the context of ERPs obtained via EEG. A lower case ‘m’ (magnetic) is added for ERFs. The characteristic deflections of the auditory response occur at latencies of approximately 50 ms (P1m), 100 ms (N1m), and 150–180 ms (P2m) relative to stimulus onset.

Variations in SOI affect the amplitudes and latencies of these peaks (for a review, see [May and Tiitinen, 2010](#)). The effect of SOI on the magnitude of the N1(m), the most prominent peak of the response, is particularly well studied (e.g. [Lu et al., 1992a](#); [Mäkelä et al., 1993](#); [Sams et al., 1993](#); [McEvoy et al., 1997](#); [Rojas et al., 1999](#); [Cheng and Lin, 2012](#); [Zacharias et al., 2012](#)). Below a certain subject-specific threshold, stimulus history affects the response

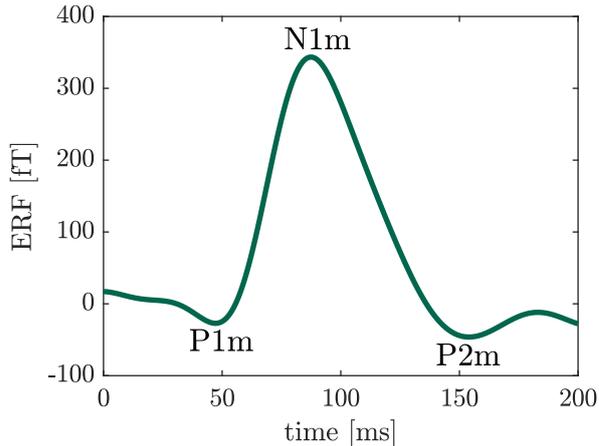


Figure 2: Representative example of an auditory ERF from the left hemisphere of a human subject (regular-SOI paradigm, 1-s SOI). The characteristic deflections – P1m, N1m, and P2m – are marked. Time relative to stimulus onset. For details regarding the measurements and analysis performed to obtain this ERF, please refer to Section 4.1.2.

and the N1(m) magnitude is smaller the shorter the SOI, i.e. RS is more pronounced the higher the repetition rate of the stimulus. Above this threshold, RS is not observed and the stimulus history does not seem to affect the response to the stimulus. This suggests that the magnitude of the suppressive effect caused by preceding stimuli decays over time and that RS has a distinct lifetime. How this lifetime can be quantified based on the regular-SOI paradigm is explained in more detail in the following section. In human MEG studies, RS lifetime mostly ranges from about 1 s to 3 s and is highly subject-specific (Lu et al., 1992a; Mäkelä et al., 1993; Sams et al., 1993; McEvoy et al., 1997; Rojas et al., 1999; Cheng and Lin, 2012; Zacharias et al., 2012).

Lu et al. (1992a) investigated the relationship between RS lifetime and the lifetime of sensory memory measured behaviourally. In a forced-choice task, subjects evaluated whether a probe tone was softer or louder than a preceding test tone and the interval between the two tones was varied. Additionally, RS lifetime was quantified via a regular-SOI paradigm. Lu and colleagues found that the longer the RS lifetime observed for the subject, the longer the maximum time span across which the subject could remember the loudness of the probe tone. This result suggests that RS is not only a fatigue effect but a reflection of a *psychological* memory process.

Moreover, the adaptation of the N1(m) to a repeated stimulus does not generalise to other stimuli. If the repeated sound is followed by a sound that deviates in some way from the repeated one, an enhanced ERP and ERF is observed at 100-ms to 200-ms latencies (Butler, 1968). This differential response is termed the mismatch negativity (MMN; Näätänen et al., 1978). It indicates the presence of a memory trace of the repeating stimulus as well as the brain’s ability to detect surprising events in the environment (Näätänen, 1990; May and Tiitinen, 2010). Indeed, MMNs appear to underlie the behavioural ability to detect unexpected sounds: When the subject’s task is to detect the deviating stimulus, the presence or absence of an MMN predicts whether the subject is able to make this detection (Näätänen et al., 1993) and the peak latency of the MMN predicts reaction time (Tiitinen et al., 1994).

An alternative interpretation of the MMN is expressed in the predictive coding framework (Friston, 2005; Rao and Ballard, 1999; Wacongne et al., 2012), where RS is not viewed as a reflection of a memory process in AC but, instead, attributed to a suppressive top-down signal. In consequence, the MMN is viewed as a prediction error. In a series of simulations, however, May and colleagues (May and Tiitinen, 2010; May, 2021) have shown that the RS-based interpretation of MMN phenomena in the AC is very robust. They demonstrated that it provides mechanistic explanations for experimentally observed responses to various MMN paradigms reported in the literature whereas explanations in the current predictive coding framework are mostly conceptual. Note, however, that the two hypotheses do not entirely exclude each other. As concluded by May, "it is possible that the brain uses local adaptation and predictive coding in tandem" (May, 2021).

Quantification of sensory memory lifetimes

As elaborated in the previous section, the neural responses recorded in the context of a regular-SOI paradigm are stimulus-history-specific – the amplitude of the stimulus-evoked response increases and ultimately saturates as a function of SOI. This indicates that the auditory system contains representations – memory traces – of the recent past and that these memory traces decay with increasing SOI. The dependence of the response peak amplitude on SOI is commonly described by an exponentially saturating function (Lu et al., 1992a; Mäkelä et al., 1993; Sams et al., 1993; McEvoy et al., 1997; Rojas et al., 1999; Cheng and Lin, 2012; Zacharias et al., 2012). This function first rises rapidly and then levels off as it approaches an asymptote:

$$P_{\text{fit}}(\text{SOI}) = A_{\text{sat}} \left[1 - \exp \left(- \frac{\text{SOI} - t_0}{\tau_{\text{SOI}}} \right) \right]. \quad (1)$$

Parameter A_{sat} defines the saturation level, i.e. the response amplitude that the function approaches asymptotically. It is given in units of the selected amplitude measure (for example magnetic field strength in the context of MEG measurements using magnetometers). Time point t_0 marks the SOI at which the function crosses the abscissa. It reflects the largest SOI-value at which the response is fully suppressed. Note that amplitude is measured in terms of absolute values and that Equation (1) is only a valid description of response amplitudes as a function of SOI for $\text{SOI} \geq t_0$. Thus, it does not suggest a reversal in the polarity of the response for $\text{SOI} < t_0$.

Parameter τ_{SOI} is a time constant describing the steepness of the initial, rapidly rising slope of the function and thus the time course of recovery from RS. It is given in units of time and technically reflects the SOI at which the response amplitude reaches 66% of the

asymptote. Time constant τ_{SOI} is used as a robust quantifier of the lifetime of RS and viewed as a reflection of the lifetime of a sensory memory trace in the AC that stores information on stimulus history.

Memory traces along the auditory pathway of other species

Human studies are predominantly limited to extracranial measurements. However, as reviewed by [Willmore and King \(2023\)](#), invasive electrophysiological recordings from single neurons or small populations of neurons in other species have shown that: (1) RS can be observed throughout the subunits of the entire auditory system, (2) RS does not generalise to all stimuli, just as observed via extracranial recordings. If the repeated stimulus (standard) is followed by a more rare stimulus deviating from the repeating one (deviant), a stronger response is observed. In the context of single-cell and multi-unit responses, this phenomenon is called stimulus-specific adaptation (SSA). Evidence suggests that SSA is the single-cell response behaviour involved in the generation of the MMN (e.g. [Ulanovsky et al., 2003](#); [Klein et al., 2014](#); [Yarden and Nelken, 2017](#)).

Intracranial recordings have shown that RS is more pronounced in non-lemniscal than in lemniscal parts of the auditory system and that the lifetime of RS increases when ascending along the auditory pathway – from auditory nerve to superior olivary complex to inferior colliculus to medial geniculate nucleus of the thalamus to AC ([Pérez-González and Malmierca, 2014](#); [Parras et al., 2017](#)). In the auditory nerve, RS can be observed for inter-stimulus intervals of up to 35 ms ([Yates et al., 1983](#), recordings from anaesthetised guinea pigs). For the superior olivary complex, an RS recovery time constant of 106 ms was reported ([Finlayson and Adam, 1997](#), recordings from anaesthetised rats, time constant similar but not equal to Equation (1)). In the inferior colliculus, RS for the standard tone in an SSA paradigm was usually only observed for SOIs below 250 ms ([Pérez-González et al., 2005](#); [Malmierca et al., 2009](#), recordings from anaesthetised rats). In the medial geniculate nucleus (MGN) of the thalamus, the lemniscal divisions exhibited RS for the standard tone at SOIs below 250 ms, whereas in the non-lemniscal MGB, it was observed with SOIs ranging up to 2 s ([Anderson et al., 2009](#); [Antunes et al., 2010](#), recordings from anaesthetised mice and rats, respectively). In the AC, RS for the standard tone is strong and occurs for SOIs of up to 2 s ([Taaseh et al., 2011](#), recordings from anaesthetised rats).

SSA measured in AC is likely to be cortically generated rather than inherited from sub-cortical processing, given that the lemniscal midbrain, which feeds into the primary auditory field (A1) of the AC, displays weaker SSA with a shorter lifetime ([Ulanovsky et al., 2003](#); [Pérez-González et al., 2005](#), recordings from anaesthetised cats and rats, respectively). This proposition is also supported by layer-specific recordings of SSA in A1 ([Szymanski et al., 2009](#), recordings from anaesthetised rats).

Mechanisms behind cortical repetition suppression

Insights from in-vivo recordings

The fundamental mechanism giving rise to cortical RS is suspected to be short-term synaptic depression (STSD), a transient weakening of synaptic efficacy caused by previous activity. This form of plasticity occurs due to a mixture of stimulus-evoked effects: depletion of readily releasable synaptic vesicles, inactivation of release sites, and inactivation of presynaptic calcium channels (Fioravante and Regehr, 2011). STSD has been linked to neuronal information processing, specifically gain control (Abbott et al., 1997), temporal filtering (Fortune and Rose, 2001; Rosenbaum et al., 2012), and the efficiency of information transfer neurons (Benda et al., 2005; Salmasi et al., 2019; Kohashi et al., 2021). It affects cortical as well as subcortical synapses (for a review, see Friauf et al., 2015). Ulanovsky et al. (2004), observing SSA in cat A1, noted that the decay time of STSD in cortex, being hundreds of milliseconds to several seconds, coincides with the lifetime of SSA measured in A1, which would be consistent with cortical STSD driving cortical SSA.

Another candidate mechanism for cortical RS is inhibition. However, Wehr and Zador (2003, 2005), conducting intracortical recordings in rat A1, noted that the suppressive effects of repeating a stimulus outlast the 100-ms effects of cortical inhibition. They concluded that RS is likely to be caused by cortical STSD rather than by inhibition.

Although invasive intracortical recordings provide greater insights into the mechanisms behind RS in the AC than the extracranial measurements commonly used in human studies, they also come with limitations. For example, measurement durations and locations need to be restricted to preserve the viability of the animal during the recordings. Moreover, it is difficult to isolate and alter individual factors that might affect RS. As summarised by Teufel and Fletcher (2016), in modelling, "*certain aspects of reality can be ignored because they are irrelevant to the bit of reality that the model attempts to explain*". Computational neuroscience can explore how interactions between biological neurons implement computational functions. As reviewed by Kriegeskorte and Douglas (2018), computational modelling endeavours of the last decades have led to a range of insights regarding sensory coding (Simoncelli and Olshausen, 2001; Olshausen and Field, 2004), working memory (Chaudhuri and Fiete, 2016), decision mechanisms (Newsome et al., 1989; Wang, 2008; Shadlen and Kiani, 2013), motor control (Diedrichsen et al., 2010), and even high-level sensory and cognitive brain representations (Yamins and DiCarlo, 2016; Kriegeskorte and Douglas, 2018). Therefore, computational modelling is also an important complementary tool for the exploration of AC function.

Insights from computational modelling

The role of STSD in the context of the SSA observed in the AC has already been further delineated through computational modelling. [David and Shamma \(2013\)](#) modelled a single A1 neuron receiving input via a set of synapses, each with different STSD dynamics. Responses were then compared to the spiking activity recorded from single neurons in A1 of awake ferrets. To optimise overlap, the temporal response function of the model was adjusted via a boosting algorithm. Specifically, the connection weights of the different synapses were altered. David and Shamma were able to demonstrate that synapse-specific STSD dynamics can explain the adaptive response dynamics of neurons in A1 in response to the stimulus envelope over a time scale of several hundreds of milliseconds.

[Loebel et al. \(2007\)](#) constructed a model of A1 as a chain of columns, each tuned to a specific stimulus audio-frequency and comprising two units, one representing the column's excitatory neurons and the other representing the inhibitory neurons. STSD was incorporated as a transient weakening of the excitatory synapses, driven by the pre-synaptic firing rate. Model parameters governing STSD dynamics had identical values across all connections. The authors found that this model reproduced forward masking and it demonstrated how frequency tuning curves depend on the balance between excitation and inhibition. In addition, the model replicated non-linear effects present in responses to subthreshold tones mixed in with noise stimuli.

[Mill et al. \(2011\)](#) represented A1 not by columns but by a small network of individual spiking neurons connected by dynamic synapses. The input layer of the network consisted of audio-frequency-tuned Poisson neurons which formed connections with sets of adaptive exponential integrate-and-fire neurons. Synapses belonged to one of three categories: fast excitatory, fast inhibitory, and fast excitatory with rapid depression and slow recovery. Model parameters governing STSD dynamics were constant across all synapses in the third category. Mill and colleagues demonstrated that a wide range of published SSA data could be replicated if more complex circuits with more than one layer are formed from the building blocks described above.

[Yarden and Nelken \(2017\)](#) created a model of A1 that closely followed the approach of [Loebel et al. \(2007\)](#) but where both feedforward thalamocortical connections and recurrent intracortical connections were affected by STSD. Recovery from STSD had identical dynamics across thalamocortical and intracortical connections, respectively, but was assumed to be faster for the former. Yarden and Nelken demonstrated that modelling STSD at the feedforward connections did not suffice to explain experimentally observed properties of SSA. STSD at the recurrent connections played an important role. Moreover, they found that feedforward inhibition was not necessary for the model to exhibit SSA. It did, however, expand the range of model parameters where SSA could be observed.

[Kudela et al. \(2018\)](#) created a model that represents a 3.6 mm² multilayer patch of A1, made up of multicompartmental cells organised in an audio-frequency-specific manner. The connection strength between these cells was modulated by STSD while the afferent thalamocortical connections were not affected. In addition to single-cell and population spiking activity, which were the focus of the previous modelling studies reviewed above, this model was able to simulate local field potentials. These could be compared to auditory evoked responses recorded from human cortex. The model's responses were validated relative to intracranial electrocorticographic recordings from a human subject undergoing intracranial monitoring for clinical purposes. The authors demonstrated that cortical sources suffice for SSA phenomena to arise but also noted that this does not exclude the contribution of other sources.

The modelling studies reviewed above suggest that SSA is a network effect arising from a tonotopically organised system of interconnected neural populations where synapses are affected by STSD. The replication of experimental data did not require a plethora of synapse-specific STSD dynamics, rather, the corresponding model parameters were constant across cortical connections and the connectivity pattern shaped the diverse adaptive response behaviour of the model units.

In this context, it is important to note that, by only describing multiunit behaviour in a single AC field, namely A1, these modelling studies omit a large part of the network pattern that characterises the AC. It is hierarchically organised and consists of a core, belt, and parabelt area (for a review, see [Hackett, 2011](#)). Each of these areas is further subdivided into fields and the number of fields as well as their connectivity pattern is species-specific.

An approach to modelling AC that reflects this network organisation was introduced by May and colleagues ([May and Tiitinen, 2013](#); [May et al., 2015](#)). In their model, the computational unit is the cortical column, described as two interacting mean-field units, one representing the excitatory neurons and the other the inhibitory neurons of the column. The excitatory connections are modulated by STSD as in [Loebel et al. \(2007\)](#) and as in the modelling studies reviewed above, simulations were run with identical STSD parameter values across all cortical connections. What sets the model apart is that it emulates the mesostructure of the macaque AC by including 13 cortical fields, each one tonotopically organised. These represent three core fields, eight belt fields, and two parabelt fields, as described by [Hackett et al. \(1998\)](#) and [Kaas and Hackett \(2000\)](#).

A unique feature that results from this modelling approach is the model's ability to replicate intracortically as well as non-invasively observed phenomena (see also [May, 2021](#)): SSA (e.g. as observed intracortically in cat AC by [Ulanovsky et al., 2003](#)), context sensitive responses to species-specific vocalisations/speech (e.g. as observed intracortically in macaque AC by [Recanzone, 2008](#)) and tone sequences (e.g. as observed intracortically in marmoset AC by [Sadagopan and Wang, 2009](#)), forward masking and forward facilitation in two-tone stimulation (e.g. as observed intracortically in cat AC by [Brosch and Schreiner, 2000](#)), as well

as the adaptation of the N1(m) response of the ERP and ERF (e.g. as observed non-invasively in human AC by Zacharias et al., 2012) and the generation of the MMN (e.g. as observed non-invasively in human AC by Näätänen et al., 1978).

Simulations, also those by Westö et al. (2016), were used to investigate the time window over which past stimulation affects the responses to incoming signals. The results indicate that this time window widens as one moves from the core to the belt and then to the parabelt and that it exceeds the time window of recovery from STSD. Hajizadeh et al. (2022) created a simplified and analytically tractable version of the model used by May and colleagues, and found that RS as observed in the adaptation of the N1m response results from a complete reorganisation of AC dynamics in terms of the normal mode oscillations that AC supports. The authors also observed that, in the simulated AC, τ_{SOI} increases as one moves up the core-belt-parabelt hierarchy. In summary, these modelling studies indicate that the structure of the neural network itself shapes the sensory memory traces formed in AC.

Animal models of human auditory processing

As showcased by the literature reviewed in the previous sections, various mammals, such as cats, mice, rats, macaques, marmosets, and ferrets, are treated as model organisms of human auditory processing. Another popular animal model is the Mongolian gerbil (*meriones unguiculatus*, hereafter simply referred to as ‘gerbil’), a small rodent with a body size that lies between those of mice and rats. Brain-related studies across multiple topics, such as for example epilepsy (as reviewed as early as 1995 by Bertorelli et al.), age-dependent hearing loss (for a review, see Gleich and Strutz, 2012), and social cognition (e.g. Tchabovsky et al., 2019), have turned to the gerbil over the last decades. Most importantly, the gerbil has advanced to one of the key species in auditory processing research (e.g. Scheich, 1991; Ohl and Scheich, 1997a; Happel et al., 2010; Wrobel et al., 2018).

One general advantage of the gerbil is its suitability as a laboratory animal. Gerbils are small in size and easy to breed, good care and living conditions can be provided at relatively low cost, and the animals take relatively well to caged living. One of the main reasons for its preferential use in auditory research is the gerbil’s high sensitivity to sounds with a relatively low audio-frequency. As described by Gleich and Strutz (2012), this likens it to humans, who are highly receptive to the frequency-range of the normal human voice (from about 100 Hz to 2 kHz). In contrast, other laboratory animals like mice and rats are most sensitive to much higher audio-frequencies (> 6 kHz). Gerbil audiograms are thus particularly well aligned with those of humans. Moreover, the middle ear cavities of the gerbil are large enough to easily reach the cochlea – this is beneficial in the context of experimental intervention. The gerbil cochlea covers an audio-frequency range from about 200 Hz to 50 kHz which is mapped to tonotopically organised fields in the AC (Thomas et al., 1993). The field structure and

connectivity pattern of the gerbil AC is very well described (e.g. [Budinger et al., 2000a,b](#); [Saldeitis et al., 2014](#), for a review see Section 1.1.2 of this thesis), an advantage for employing the computational modelling approach introduced by May and colleagues [2013](#); [2015](#). Thus, we selected the gerbil as a focal point of our in-vivo and in-silico investigation of RS.

Research questions addressed in this thesis

As reviewed above, STSD seems to be a central mechanism driving RS, yet the meso-structure of AC could also be an important factor in generating context sensitivity in the AC. It is, however, unclear, how the network pattern shapes RS dynamics. In this thesis, we study the emergence of RS in vivo and in silico in the gerbil as well as the human AC. We used the modelling approach developed by May and colleagues as a starting point for building our gerbil AC model – a model that captures the gerbil-specific field structure and connectivity pattern of the AC. As in the studies by May and colleagues, STSD dynamics were kept constant across all model units. We then used our model to explore RS at the level of individual neural populations in A1.

Specifically, we employed a regular-SOI paradigm to investigate how the lifetime of RS might vary as a function of stimulus audio-frequency and traced back the network effects giving rise to this variation. The in-silico investigation was complemented by an analysis of in-vivo electrophysiological data from gerbil A1, recorded during a regular-SOI paradigm where stimulus blocks were presented at two different audio-frequencies. The explanatory power of our results depended on two important factors: (1) the quality and robustness of the analysis pipeline from raw experimental data to the final subject- and condition-specific τ_{SOI} -values, i.e. RS lifetimes, and (2) the quality of our gerbil AC model. Therefore, beyond examining RS itself, we also worked on these two factors.

To address factor 1, we scrutinised the data set from each animal to verify its suitability for τ_{SOI} -computation (Section 2.2) and developed a bootstrap-based approach for the determination of subject- and condition-specific confidence intervals for these RS lifetimes (Section 2.3). The bootstrap-based approach is also transferable to data sets collected using other measurement methods and we employed it to verify the robustness of a hemispheric difference observed for RS lifetimes in the human AC (Chapter 3).

To address factor 2, we took great care in reflecting a plethora of experimentally established properties of the gerbil AC in our computational model (Sections 1.1.1 and 1.1.2). With the future improvement of this first version of our model in mind, we also embarked on a project where we pioneered an algorithm-based optimisation pipeline for a simplified version of the AC model by May and colleagues (Chapter 4). This optimised model, in turn, provided predictions regarding widespread RS generation and area-specific RS lifetimes in the human AC.

Thus, in total, this thesis consists of four intertwined projects:

- P1:** The in-silico exploration of audio-frequency-specific RS in the gerbil AC via computational modelling
- P2:** The in-vivo exploration of audio-frequency-specific RS in the gerbil AC via electrophysiological recordings, including the development of a new analysis pipeline
- P3:** The application of this new pipeline to a human MEG data set to verify the robustness of the observed hemispheric difference in RS lifetime
- P4:** The development of an automated approach for the improvement of computational AC models, also resulting in predictions regarding RS generation and lifetimes

The following chapters, each comprising their own ‘Materials and Methods’ and ‘Results’ sections, cover these projects in the order given above. Each chapter ends with a short set of preliminary conclusions or predictions. The main ‘Discussion’ ensues after the results from all four projects have been introduced and aims to answer a series of questions by linking these results. Specifically, I will discuss how the challenges of RS lifetime quantification can be addressed, which underlying processes this lifetime reflects, and which role network-modulated RS might play as a memory trace facilitating the process of temporal binding, such as, for example, the binding of the syllables ‘pu’ and ‘ma’ in our opening example (Figure 1).

As I aim to link our results across projects, the electronic form of this thesis contains many clickable cross references, some of them taking the reader to a location far away from the respective current page. To improve the reading experience, I would like to make the following recommendation: In most PDF readers, the viewing path can be retraced. After following a cross reference, rather than scrolling through the document in search of your previous location, simply use the shortcut `Alt` + `←` (or `cmd` + `[` for Mac users).

Collaborations

Most of the interdisciplinary research I present below was only possible via collaborations. Therefore, I use the ‘we’-perspective throughout the next four chapters.

To investigate in-vivo RS dynamics of the gerbil AC (Project 2), I collaborated with former PhD candidate Dr J. Ma, who carried out electrophysiological recordings in gerbil A1. For her thesis, J. Ma analysed the resulting data set with a focus on layer-specific RS in gerbil AC, whereas, for my thesis, I investigated audio-frequency-specific RS using the same recordings. For both approaches, I developed a bootstrap-based analysis pipeline. The results from J. Ma’s thesis about layer-specific RS in the gerbil AC are presented in a joint manuscript (preprint: Ma & Brunk et al., 2021).

The human MEG data presented in Project 3 was recorded during a previous research project at the Leibniz Institute for Neurobiology Magdeburg. PhD candidate A. Dar and I then collaborated to create a refined analysis pipeline for the computation of RS lifetimes in the human AC. To assess the robustness of a potential hemispheric difference in RS lifetime, we built on the bootstrap-based approach I established for the gerbil data. A. Dar developed the scripts required to adapt the approach to MEG data and I introduced additional analysis steps. The collaboration resulted in a shared first-author publication about hemisphere-specific RS in the human AC (Dar & Härtwich et al., 2025). The exploration of underlying mechanisms giving rise to this difference, as well as its potential functional relevance, will be the central focus of A. Dar's thesis. In my thesis, the focus lies on demonstrating how to produce robust subject-specific estimates of RS lifetimes.

To carry out Project 4, the development of an automated approach for the improvement of computational AC models, I collaborated with PhD candidate E. Tomana and her supervisor, Dr C. Sielużycki (Department of Biomedical Engineering, Wrocław University of Science and Technology, Poland). Supported by her supervisor, E. Tomana developed and implemented an optimisation algorithm suitable for a simplified model of the human AC. I prepared this computational model and provided neuroscientific advice and general assistance during the conception and refinement of the optimisation. E. Tomana's analysis of the resulting data focussed on network connectivity patterns, whereas I characterised the model in terms of RS response dynamics. The collaboration lead to a joint publication about connectivity patterns and RS in the human AC (Tomana et al., 2023). Connectivity patterns in the human AC will be a central focus of E. Tomana's thesis, whereas my thesis introduces predictions regarding RS lifetimes.

Part II: Repetition suppression in the gerbil auditory cortex

"All models are wrong, but some are useful"

– George Box

Project 1: Repetition suppression in silico

As reviewed in the Introduction, while STSD seems to be a central mechanism in driving RS, the network structure of the AC could also play an important role in generating context sensitivity in the AC. To unveil how the network pattern might shape RS dynamics, we turned to computational modelling. After building a gerbil AC model, we employed a regular-SOI paradigm to investigate how the lifetime of RS in individual neural populations of the A1 field might vary as a function of stimulus audio-frequency and traced back the network effects giving rise to this variation.

The Materials and Methods section of this chapter first describes the computational gerbil AC model we created. Additionally, a model consisting of a single node – the single-column model – is introduced and used to contrast its response behaviour with that of the nodes in the complex gerbil AC network. Next, the regular-SOI paradigm used to stimulate the two models is specified. Finally, the analysis pipeline applied to the in-silico responses recorded from both models is presented.

The Results section first describes the response behaviour of the single-column model, in terms of frequency response functions, RS lifetimes, and response peak latencies. Then the mechanisms behind this response behaviour are unveiled. Next, the single-column model's response behaviour is contrasted with that of the gerbil AC model and the network effects leading to these different responses are explored.

Finally, the predictions deduced from the simulations are summarised and then tested in vivo in the following chapter (Project 2).

1.1 Materials and Methods

We constructed a computational gerbil AC model based on the modelling approach for the human AC developed by May et al. (2013; 2015). The basic building block in this approach is a simplified version of the cortical column, encompassing a population of excitatory cells (pyramidal neurons) and a population of inhibitory cells (interneurons) and thus having antecedents in the model of Wilson and Cowan (1972). The following subsections first introduce the dynamic equations that govern the interactions between any assembly of these building blocks (Section 1.1.1) and then explain how they were assembled to reflect the gerbil AC (Section 1.1.2). Next the experimental paradigms that we simulated (Section 1.1.4) are described and, finally, the analysis steps applied to the data ‘recorded’ in silico are explained (Section 1.1.5). Simulations as well as data analysis were performed using MATLAB (The Math Works, Inc., Version 2021b).

1.1.1 Model dynamics

The state of each neural population is described by a state variable, u for excitatory and v for inhibitory populations, and the output of each population is reflected by a mean firing rate g , which is a non-linear function of the state variable:

$$g(u) = \begin{cases} 0 & u \leq \theta \\ \tanh[\kappa(u - \theta)] & u > \theta \end{cases} \quad (2)$$

Note that the same equation applies for state variable v . The population starts to fire when threshold θ is crossed and, as the state variable increases further, the firing rate asymptotically approaches a maximum value of 1. The factor κ scales the sensitivity of the firing rate to the state variable.

The dynamic equations describing the temporal evolution of the state variables of the neural populations are based on the leaky-integrator model, specifically as described by Hopfield and Tank (1986). They expressed neural interactions at the single neuron level based on approximations of the cross-membrane current flow (mean-field leaky integrator neuron). In this case, the state variable of the individual units in the network is an approximation of the membrane potential. As a simplifying approximation, the May-et-al. modelling approach assumes that individual neurons within a population are identically and symmetrically connected with each other, and that they all receive the same external input. Under these conditions, the individual units of a population all behave identically and can be viewed as subunits of the population which in turn can be treated as a unit in the mean-field leaky

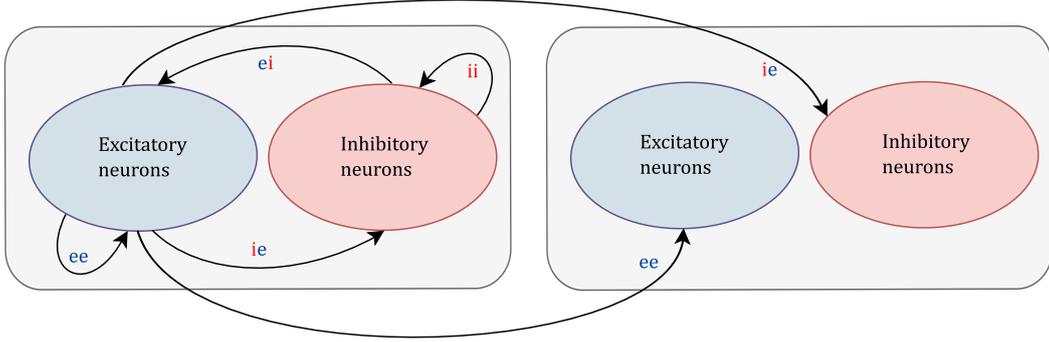


Figure 3: Diagram illustrating the excitatory (blue) and inhibitory (red) neural population of two columns (grey) in the May-et-al. modelling approach. The connections the left column forms internally and across its boundaries are illustrated by arrows.

integrator neuron model.

For a network with a total of N cortical columns, the two vectors $\mathbf{u}(t) = (u_1(t), \dots, u_N(t))^\top$ and $\mathbf{v}(t) = (v_1(t), \dots, v_N(t))^\top$ respectively denote the state variables of the N excitatory and inhibitory populations in the network. The dynamic equations describing neural interactions as a function of time t are:

$$\tau_m \dot{\mathbf{u}}(t) = -\mathbf{u}(t) + W_{ee} Q(t) g[\mathbf{u}(t)] - W_{ei} g[\mathbf{v}(t)] + \mathbf{i}_{\text{aff}}(t), \quad (3a)$$

$$\tau_m \dot{\mathbf{v}}(t) = -\mathbf{v}(t) + W_{ie} Q(t) g[\mathbf{u}(t)] - W_{ii} g[\mathbf{v}(t)], \quad (3b)$$

$$\dot{\mathbf{q}}(t) = -\frac{\mathbf{q}(t) \circ g[\mathbf{u}(t)]}{\tau_{\text{on}}} + \frac{\mathbf{1} - \mathbf{q}(t)}{\tau_{\text{rec}}}, \quad Q(t) = \text{diag}(\mathbf{q}(t)), \quad (3c)$$

where τ_m is the membrane time constant and $\mathbf{i}_{\text{aff}}(t)$ is a vector of length N reflecting the amplitudes of afferent input arriving at individual populations via the auditory pathway. Afferent input is assumed to target the excitatory populations only. The matrices W specify the weights of

- Excitatory connections within and between populations of pyramidal neurons (W_{ee}),
- Inhibitory connections from interneuron to pyramidal neuron populations (W_{ei}),
- Inhibitory connections within and between populations of interneurons (W_{ii}),
- Excitatory connections from pyramidal neuron to interneuron populations (W_{ie}).

Figure 3 illustrates these connection types. Note that ii-connections are effectively excitatory as they dampen further inhibition whereas ie-connections are effectively inhibitory as they drive the inhibitory action of inhibitory populations.

description	symbol	value
firing rate threshold	θ	0.050
firing rate sensitivity	κ	$\frac{2}{3}$
membrane time constant	τ_m	0.030 s
time constant of synaptic depression onset	τ_{on}	thalamus: 0.020 s cortex: 0.100 s
time constant of recovery from synaptic depression	τ_{rec}	thalamus: 0.100 s cortex: 1.000 s

Table 1: Time constants and firing rate parameter values used in the gerbil AC model.

A further expansion of the [Hopfield and Tank](#) model is the inclusion of STSD. Following [Loebel et al. \(2007\)](#), this is implemented by introducing a term that scales the weights of excitatory connections as a function of pre-synaptic firing rate. The modulation is expressed by the matrix multiplication of the elements of W_{ee} and W_{ie} with the time-dependent diagonal matrix $Q(t) = \text{diag}(\mathbf{q}(t))$. The elements of \mathbf{q} reduce in amplitude as a function of increasing pre-synaptic firing rate, with distinct time constants τ_{on} and τ_{rec} for onset of and recovery from this reduction in synaptic efficacy (see Equation (3c), note that \circ indicates element-wise multiplication, i.e. the Hadamard product). Onset of and recovery from a decrease in synaptic efficacy are competing effects and full recovery can only occur if the pre-synaptic neural population is not firing. For simplicity, only excitatory connections are affected by synaptic depression. However, note that, due to the effectively inhibitory ie-connections, both excitation and inhibition of the network is affected by synaptic depression.

For our simulations, we set the time constants of the model to the values listed in Table 1. The order-of-magnitude difference of the time constants of synaptic depression in the parts of the model representing MGv and AC reflects the different time scales of RS observed along the lemniscal part of the auditory pathway ([Ulanovsky et al., 2004](#); [Asari and Zador, 2009](#); [Pérez-González and Malmierca, 2014](#)). Across all our simulations, the firing rate threshold θ was equal to 0.050 and we set κ to $\frac{2}{3}$. When the structure of the model (i.e. the weight matrices W_{ee} , W_{ei} , W_{ie} , and W_{ii}) as well as the stimulus sequence (in terms of \mathbf{i}_{aff}) is defined, the nonlinear system described by Equations (3a)–3c can be solved numerically to provide a picture of the spatiotemporal activity patterns of the network in response to the chosen paradigm.

1.1.2 Structure of the anatomy-based model

In our computational model, the weight matrices W represent the structure and connectivity of the gerbil AC. Its functional organisation has been investigated intensively, using electrophysiological (e.g. [Thomas et al., 1993](#)) as well as metabolic (e.g. [Scheich et al., 1993](#)) and anatomical (e.g. [Budinger et al., 2000a,b](#)) mapping techniques. These studies have identified eight functional fields. Each field can be characterised in terms of its position in the structural hierarchy of the core, the belt, and the parabelt. The core consists of two fields – the primary auditory field (A1) and the anterior auditory field (AAF). Three fields make up the gerbil AC belt area – the dorsoposterior field (DP), the ventroposterior field (V), and the ventral field (V). The gerbil AC parabelt area encompasses three fields – the dorsal field (D), the ventromedial field (VM), and the anteroventral field (AV).

Properties of the gerbil AC fields

Auditory fields are characterised by their connectivity at the intracortical and subcortical level ([Budinger et al., 2000a,b, 2006, 2008](#); [Budinger and Scheich, 2009](#); [Budinger et al., 2013](#); [Saldeitis et al., 2014](#); [Henschke et al., 2015](#); [Kurt et al., 2008](#)) as well as by cyto-, fiber-, and chemoarchitectural criteria ([Budinger et al., 2000a](#); [Radtke-Schuller et al., 2016](#)).

Like in cats, monkeys, humans, and other mammals (for a review, see [Hackett, 2011](#)), fields in the gerbil AC have the following characteristics: The core fields are tonotopically organised, have short-latency on-responses to pure tones, a koniocortical architecture, and dense myelination. They form connections with the thalamus, specifically the tonotopically organised ventral division of the medial geniculate nucleus (MGv). Additionally, core fields form extensive local connections with each other and the surrounding belt fields.

Belt fields are also tonotopically organised, although with a lesser spatial resolution of frequencies, and they have a prokoniocortical architecture. Their connections with thalamus primarily target the MGv but they are also connected to non-tonotopic parts of the auditory thalamus. Furthermore, strong connections exist between the belt and the core as well as between the belt and the parabelt, and the belt fields also share connections with other sensory and non-sensory cortical areas.

The non-tonotopic parabelt fields are isocortical with a lesser cell density and fiber myelination than the belt fields. They are not tonotopically organised and respond to a large range of audio-frequencies. Connections project to non-tonotopic auditory and non-auditory thalamic nuclei and remote cortical areas.

Taken together, the studies cited above provide a detailed picture of the field parcellation, connectivity, and hierarchical organisation in the gerbil AC. We summarised this information in the diagram shown in [Figure 4a](#). It forms the basis of the more formalised connectivity matrix used for the gerbil AC model, shown in [Figure 4b](#). The connections listed in this matrix

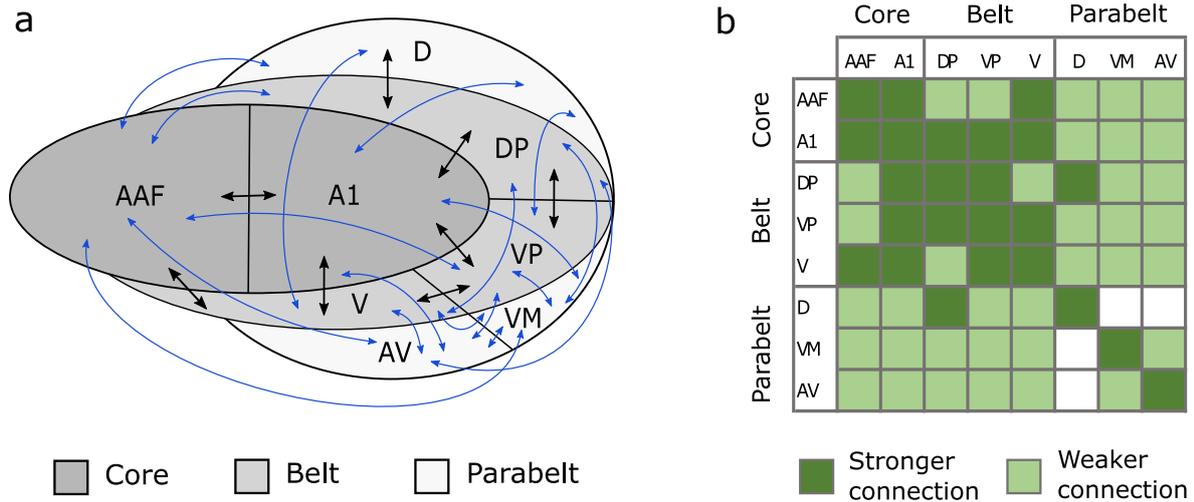


Figure 4: Gerbil AC connectivity diagram (a) and translation of the diagram into a connectivity matrix (b), mainly based on [Budinger et al. \(2000a,b\)](#) and [Saldeitis et al. \(2014\)](#), as well as on assumptions 1, 2, and 3. The thicker black arrows in the diagram (a) indicate that experimentally observed neuronal connections between the respective fields were more numerous than between fields linked by blue arrows. In b, entries along the main diagonal of the matrix reflect intra-field connections and off-diagonal entries describe inter-field connections. In analogy to the arrows in the diagram, the darker green shade marks net connection strengths classified as stronger and the lighter green shade those classified as weaker. Intra-field connections are also strong. The white matrix elements indicate that these particular inter-field connections have, thus far, not been reported in the literature.

are mainly based on results of anterograde tracer injections (biocytin, dextran amines) into the two core fields as well as the three belt fields and the MGv ([Budinger et al., 2000a,b](#); [Saldeitis et al., 2014](#)). Additionally, the following assumptions were made:

1. Fields that share a common border (i.e. neighbouring fields) are connected.
2. Interfield connections occur in a reciprocal manner (the available studies did not always test for reciprocity).
3. Reciprocal connections have equal weight in both directions.

Connections within auditory fields

To model the functional organisation of the gerbil AC introduced above, we reflected each of the eight cortical fields by a total of $N_f = 16$ model columns. Furthermore, we included 16 MGv columns to be used as an input area, relaying activation to the AC. In total, there were thus $N = (8 + 1) \cdot N_f = 144$ columns and over $20 \cdot 10^3$ potential connections per connection type ($N^2 = 144 \cdot 144 = 20,736$). To identify each column, we assigned column indices i from 1 to 144. Table 2 lists the columns' distribution across cortical fields in terms of this index.

The fine structure of intra-field connections originating from excitatory neural populations (ie, ee) was based on findings from the mouse AC. [Levy and Reyes \(2012\)](#) showed that neurons in A1 with somata in close proximity of each other have a higher connection probability than neurons located further apart. The connection probability decreases as a Gaussian function

Location	MGv	Core		Belt			Parabelt		
		AAF	A1	DP	VP	V	D	VM	AV
Column indices	1	17	33	49	65	81	97	113	129
	\vdots								
	16	32	48	64	80	96	112	128	144

Table 2: Distribution of columns in the gerbil AC model, in terms of column index i , across MGv and cortical fields in the core, belt, and parabelt area.

of radial distance from the soma of the respective cell. In past implementations of the May-et-al. modelling approach (reflecting the macaque/human AC), this finding was reflected by generating weight matrices W via a stochastic process. The occurrence (binary, 'on' vs 'off') of connections beyond the leading diagonal was governed by a Gaussian probability density function (May and Tiitinen, 2013; May et al., 2015). Thus, each regeneration of the matrices yielded the same general connectivity pattern at the field level but different patterns at the microstructure (single-column) level. Here, we opted for a deterministic process generating highly organised weight matrices – even at the single-column level – to improve the traceability of single-column response patterns and identify basic mechanisms behind RS dynamics.

Weight values for all four connection types were largest along the main diagonal of the weight matrices W (i.e. for intra-column connections). Connections originating from inhibitory neural populations (ei, ii) were assumed to be very localised and only occurred within columns. On either side of the main diagonal (but within the intra-field limit), values decreased with increasing difference between column indices i (target of connection) and j (source of connection) for ee- as well as ie-connections. Note that, for columns within the same field, the difference between column indices reflects, at least conceptually, the physical distance between neurons as investigated by Levy and Reyes (2012). The *effectively inhibitory* ie-connections (c.f. Section 1.1.1) between these columns reflect a form of lateral inhibition observed experimentally in gerbil A1 (Kurt et al., 2008; Moeller et al., 2010).

Connections across auditory fields

Connections across fields, i.e. inter-field connections, only originated from excitatory populations, reflecting the considerable reach of the long axons of pyramidal neurons (de la Mothe et al., 2006). For connections across auditory fields, these axons predominantly target cortical layer III (Thomas and López, 2003). In this layer, inhibitory neurons are not very abundant, making up only a quarter of the neuronal population, and pyramidal neurons receive the majority of ipsilateral as well as contralateral cortical connections (Budinger and Kanold, 2018). Based on these findings, our model assumed that connections driving inhibition (ie-connections) only extend across columns within the same field.

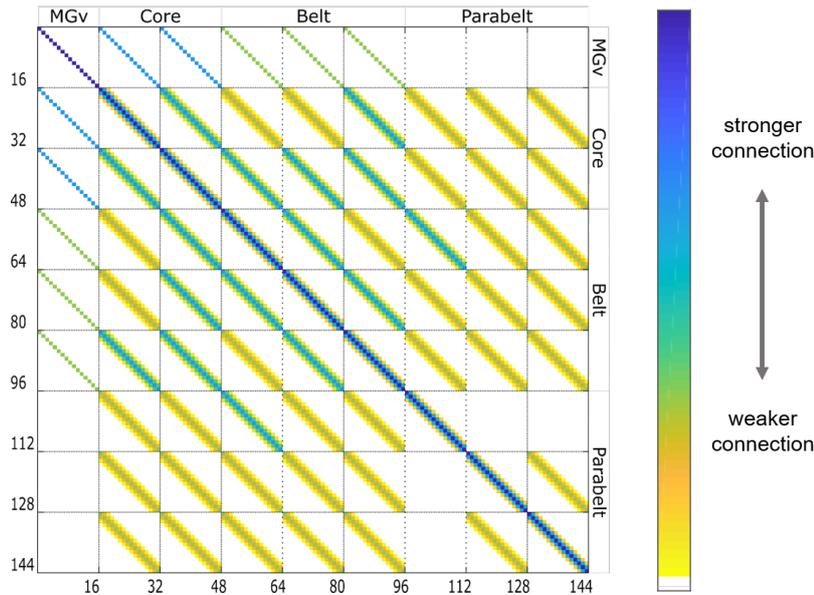


Figure 5: Illustration of the connectivity pattern defined by weight matrix W_{ee} . The network reflecting gerbil AC has a total of $N = 144$ ‘nodes’, namely columns comprising an excitatory and an inhibitory neural population. Thus, matrix W_{ee} is a square matrix of size 144×144 . It reflects connections between excitatory populations (ee). Indices i (1 to 144) of the connection *targets* are listed along the vertical axis and indices j (1 to 144) of the connection *sources* along the horizontal axis. Each entry $W_{ee}(i, j)$ reflects the strength of the connection between source j and target i , one of the over $20 \cdot 10^3$ possible ee-connections ($N^2 = 20,736$). Weight values are mapped out according to a colour gradient – the darker the colour the stronger the connection. Recurrent excitatory connections, i.e. intracolumn connections, are listed along the main diagonal of the matrix, where $i = j$. They are the strongest of all connections, as reflected by the dark blue colour. As each field comprises a total of 16 columns, grey lines divide the matrix into subdivisions of size 16×16 . Each subdivision reflects connections between two specific fields (or, for subdivisions along the main diagonal, one and the same field), as summarised in Figure 4b. Weight values are largest on the respective diagonal of each subdivision and decrease with distance from this diagonal (Equation (6)), thus creating a topographic connectivity pattern. The distinction between stronger and weaker connections at the field level, as mapped out in Figure 4b, is reflected by the stronger and weaker connection weights in the respective subdivisions of W_{ee} .

We arranged the connections driving excitation across fields (ee-connections) in a topographic manner. This is illustrated in Figure 5. Each 16×16 subdivision of weight matrix W_{ee} , mapped out by grey lines, reflects connections between two specific fields. Weight values were largest on the diagonals of these subdivisions and continuously decreased with distance from the respective diagonal. This distance Δi is given by the difference between *intrafield* indices $k(i)$, i.e. the difference between the columns' respective positions *within* a field:

$$k(i) = i - \lfloor i / (N_f + 1) \rfloor \cdot N_f, \quad (4)$$

where $\lfloor \cdot \rfloor$ denotes the floor function. For example, the column with index $i = 39$ is the *seventh* column in A1, the third of the model fields, where indices range from 33 to 48 (c.f. Table 2). Consequently, the intrafield index k of column 39 is equal to $(39 - \lfloor 39 / (16 + 1) \rfloor \cdot 16) = 7$. For any connection from column j to column i , distance Δi from the diagonal of the corresponding weight matrix subdivision is thus

$$\Delta i = |k(i) - k(j)|. \quad (5)$$

The decrease in connection weights as a function of Δi was governed by a truncated Gaussian function. For $i \neq j$, connection weights were defined by

$$W(i, j) = \begin{cases} \alpha \cdot \exp \left[-\frac{1}{2} \left(\frac{\Delta i}{\sigma} \right)^2 \right] & \text{if } \Delta i \leq \sigma \\ 0 & \text{otherwise} \end{cases}, \quad (6)$$

where the connection-type-specific standard deviation σ is expressed in units of Δi and α is a connection-type-specific scaling factor. The weight values for connections within the column ($i = j$) were assigned separately and were always larger than for all other connections.

By imposing the topographic connectivity reflected by Equation (6), we created a network that preserved the tonotopic organisation of the input (described in the following section), thus reflecting the tonotopic maps found in core and belt fields of the gerbil AC. For simplicity, parabelt field connections in our model were also tonotopically organised. Parameters defining the specific connection weight values we used across all four weight matrices are summarised in Table 3. Weight values describing connections within columns and connections between fields remained constant across our simulations. However, as part of our investigation, we varied the lateral reach σ of ie-connections within fields. [May and Tiitinen \(2013\)](#) demonstrated that this lateral inhibition modulates levels of spectral selectivity. Thus, we used parameter σ_{ie} to investigate the relationship between spectral selectivity and audio-frequency-specific RS.

connection type	parameter setting		
	within column	within field	across fields
ei – from inhibitory to excitatory population	$W_{ei}(j, j) = 3.5$	$W_{ei}(i, j) = 0$	$W_{ei}(i, j) = 0$
ii – from inhibitory to inhibitory population	$W_{ii}(j, j) = 1.0$	$W_{ii}(i, j) = 0$	$W_{ii}(i, j) = 0$
ie – from excitatory to inhibitory population	$W_{ie}(j, j) = 3.5$	thalamus: $W_{ie}(i, j) = 0$	$W_{ie}(i, j) = 0$
		cortex: $\alpha = 1$ $\sigma \in [4, 5, 7]$	
ee – from excitatory to excitatory population	$W_{ee}(j, j) = 6$	thalamus: $W_{ee}(i, j) = 0$	thalamus \leftrightarrow cortex: $\alpha = 0.5, \sigma = 0$
		cortex: $\alpha = 0.6$ $\sigma = 2$	cortex \leftrightarrow cortex: $\alpha = 0.6$ (strong) $\alpha = 0.06$ (weak)

Table 3: Weight values used in the weight matrices W of the gerbil AC model. Connections originating from inhibitory populations (**ei**, **ii**) are very localised and only occur within columns. All other entries of the respective weight matrices are equal to zero. Connections driving the inhibition (**ie**) are strongest within the column but also extend across columns within the same field. Similarly, connections driving excitation (**ee**) are strongest within the column and also extend across columns within the same field. Additionally, they are the only connections that extend across fields. The larger the distance Δi (Equation (5)) between intrafield indices k (Equation (4)) of the source and target population, the weaker the connection weight (Equation (6)). Exact weight values are defined by the respective parameter values α and σ . Where the table lists weight values equal to zero, α was set to zero. The network defined by the four weight matrices given above is characterised by a topographic connectivity pattern that preserves the tonotopic organisation of the input.

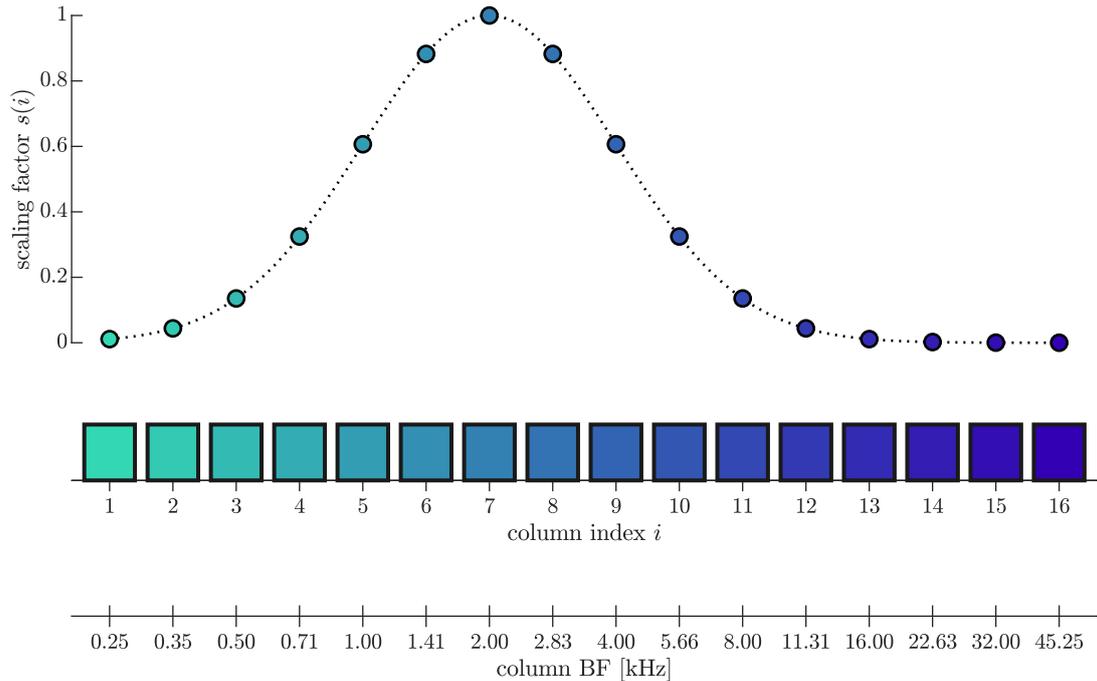


Figure 6: Diagram illustrating the tonotopic map across the 16 columns reflecting thalamus (MGv) in the gerbil AC model. The bottom axis lists the BFs assigned to these columns, with respective column indices i listed along the middle axis. As an example for one of the 16 audio-frequency options, the plotted curve depicts the scaling factor $s(i)$ (Equation (7)) applied to the presynaptic firing rate $r(t)$ when a 2-kHz pure tone ($i_{\text{BF}} = 7$) is simulated.

Connections involving thalamus

Thalamic connections in the model were purely intra-columnar and both thalamocortical and corticothalamic connections were limited to their direct counterparts ($\Delta i = 0$) in the respective fields. MGv and core fields were more strongly connected than MGv and belt fields, whereas parabelt fields were not connected to MGv at all. Moreover, our model did not include a proper description of peripheral processing. Instead, the front end of the model was an abstract tonotopic map whereby each of the 16 columns in the MGv could be maximally tuned to a single stimulus audio-frequency or frequency range and thus each had a different best frequency (BF, the stimulus audio-frequency that elicits the strongest response). The tonotopic map we specified to facilitate comparisons with experimental results is illustrated by the bottom axis of Figure 6. We assigned a specific BF to each MGv column, spanning the hearing range of the gerbil in half-octave steps from 0.25 kHz to 45.2 kHz (the hearing range of the gerbil extends from about 0.2 kHz to 50 kHz at 30 dB sound pressure level; e.g. [Ryan, 1976](#); [Gleich and Strutz, 2012](#)).

Sequences of identical pure-tone stimuli were expressed in terms of pre-synaptic firing rate time courses $r(t)$. A firing rate of 0 reflected no stimulation, a firing rate of 1 the maximum sound pressure level (SPL) of the tone. Any given pre-synaptic firing rate was then scaled as a function of the target field and the difference between the BF of the target column and the

ΔF [oct]	0	0.5	1	1.5	2	2.5	3	3.5	4
$s(i)$	1.000	0.969	0.882	0.755	0.607	0.458	0.325	0.216	0.135

Table 4: Scaled pre-synaptic firing rates $s(i)$ reflecting pure tone stimulation, as a function of the distance ΔF of the audio-frequency of the simulated stimulus from the targeted MGv column’s BF ($\sigma_{in} = 2$).

audio-frequency of the stimulus. Cortical columns were never directly activated by afferent input, hence the corresponding scaling factors were always equal to zero. For the MGv, the scaling factor was governed by a Gaussian function, with a maximum value of 1 centred on the index i_{BF} of the MGv column where $BF =$ stimulus audio-frequency ($1 \leq i_{BF} \leq N_f$), and with a standard deviation $\sigma_{in} = 2$ columns. Thus, in summary, scaling factor s for target column i was described by:

$$s(i) = \begin{cases} \exp \left[-\frac{1}{2} \left(\frac{i - i_{BF}}{\sigma_{in}} \right)^2 \right] & i \leq N_f \\ 0 & i > N_f, \end{cases} \quad (7)$$

where $1 \leq i \leq N$. The round markers in Figure 6 illustrate, as an example, $s(i)$ for the 16 MGv columns and a stimulus audio-frequency of 2 kHz.

We listed the products of column-specific scaling factor $s(i)$ and firing rate time course $r(t)$ in the vector $\mathbf{g}_{aff}(t) = (s(1) \cdot r(t), \dots, (s(N) \cdot r(t)))^\top$. Via vector \mathbf{i}_{aff} (c.f. Equation (3a)), these column-specific scaled pre-synaptic firing rates ultimately induced excitatory synaptic currents in the AC model. Akin to all other connections in the network, afferent connections were affected by STSD, with the temporal evolution of synaptic efficacies described by (c.f. Equation (3c)):

$$\dot{\mathbf{q}}_{aff}(t) = -\frac{\mathbf{q}_{aff}(t) \circ \mathbf{g}_{aff}(t)}{\tau_{on,aff}} + \frac{\mathbf{1} - \mathbf{q}_{aff}(t)}{\tau_{rec,aff}}. \quad (8)$$

Time constants $\tau_{on,aff}$ and $\tau_{rec,aff}$ were set to values equal to those used for thalamic connections (see Table 1). Entries in the vector $\mathbf{i}_{aff}(t)$, i.e. the afferent excitatory synaptic currents (see Equation (3a)) as a function of time, were ultimately computed as:

$$\mathbf{i}_{aff}(t) = w_{aff}(\mathbf{q}_{aff}(t) \circ \mathbf{g}_{aff}(t)), \quad (9)$$

where w_{aff} is the connection weight for the afferent connections. We used $w_{aff} = 1$ in all our simulations. Due to the Gaussian function in Equation (7), input representing pure tone stimuli at the BF of a specific MGv column also activated neighbouring columns, but to a lesser extent. This approach reflects the fact that neuronal tuning curves have a certain

width (see, for example, [Butts and Goldman, 2006](#)) and that there should thus not be a sudden cut-off of input across columns. To further facilitate comparisons with experimental results, we defined the term ΔF . It reflects the distance of the stimulus audio-frequency from a column's BF in units of octaves. For a pre-synaptic firing rate reflecting the maximum SPL of a simulated pure-tone stimulus ($r = 1$), [Table 4](#) summarises, as a function of ΔF , the scaled pre-synaptic firing rate $s(i)$ at afferent connections targeting MGv.

1.1.3 Structure of the single-column model

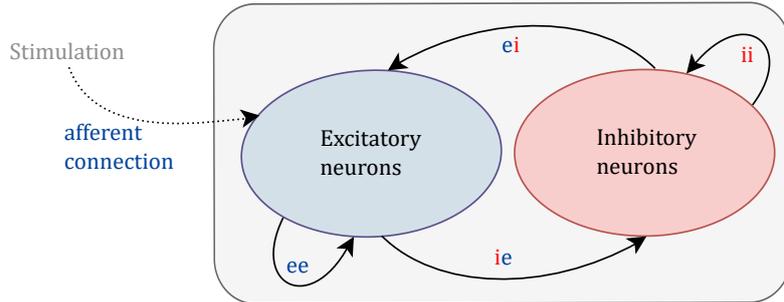


Figure 7: Diagram illustrating the single-column model, where $N = 1$. Afferent input (reflecting auditory stimulation) targets the excitatory population via the afferent connection. Just as for the gerbil AC model, the dynamic response of the network is governed by Equations (3a-c). Afferent and excitatory intracolumnar connections are subject to STSD (c.f. Equations (3c) and (8), respectively).

In addition to the gerbil AC model, we implemented a model consisting of a single excitatory and inhibitory population (see [Figure 7](#)). The aim was to compare this column's response behaviour to the RS dynamics observed for columns embedded into the larger gerbil AC network and thus identify potential network effects. Just like for the gerbil AC model, the dynamic response of the single-column model was governed by Equations (3a-c). Note that, for a model of size $N = 1$, these equations are reduced to scalar form and only a single target for afferent input remains. Such a single-column model could, of course, not be tonotopically organised. However, akin to the approach taken for the gerbil AC model, we reflected different stimulus audio-frequencies by scaling the pre-synaptic firing rate of the afferent connection. BF stimulation was reflected by a pre-synaptic firing rate of 1 and for each half-octave step away from the BF, the maximum value decreased according to the previously described Gaussian function with $\sigma_{\text{in}} = 2$ (see [Table 4](#)).

With the exception of $\tau_{\text{rec,aff}}$, parameter values in the single-column model were always equal to those in the gerbil AC model, i.e. all cortical values summarised in [Table 1](#) as well as the intracolumnar connection weights summarised in [Table 3](#) were applied. We used three different values for $\tau_{\text{rec,aff}}$ to investigate how RS lifetime varies when STSD recovery at the afferent connection is faster than, equal to, or slower than at the recurrent connections. Exact values are summarised in [Table 5](#).

afferent vs. intracolumnar STSD recovery			
parameter	faster	equal	slower
$\tau_{\text{rec,aff}}$	0.500 s	1.000 s	1.500 s
τ_{rec}	1.000 s	1.000 s	1.000 s

Table 5: Time constants of recovery from STSD for afferent ($\tau_{\text{rec,aff}}$) and intracolumnar (τ_{rec}) connections, used for simulations with the single-column model.

1.1.4 Stimulation paradigms

Repetition suppression paradigm

We investigated audio-frequency-specific RS in the gerbil AC and in the single-column model by presenting a classic regular-SOI paradigm. Each stimulus had a duration of 100 ms, with a respective linear rise and fall time of 5 ms. Within a stimulus block, the SOI was constant and the stimulus was repeated a total of $n_{\text{stim}} = 20$ times. Across blocks, the following 10 SOI-values were used: 0.219 s, 0.328 s, 0.438 s, 0.656 s, 0.875 s, 1.313 s, 1.750 s, 2.626 s, 3.500 s, and 7.000 s. These values were selected to match those employed for the in-vivo experiments (see Section 2.1.2). For the single-column model, the blocks were presented at nine different audio-frequencies, with ΔF ranging from 0 oct to 4 oct in half-octave steps.

For the gerbil AC model, we found that, at a SOI of 219 ms, model columns were still responding to the previous stimulus when the subsequent stimulus was presented (c.f. Figure 8). Stimulus-specific responses could thus not be identified. Therefore, we excluded data recorded for the 0.219-s SOI from further analysis and added a 5-s SOI to the paradigm instead. The 10 stimulus blocks were each presented at two different audio-frequencies – 2 kHz and 8 kHz.

The initial conditions before the presentation of each stimulus block were equal to the model’s resting state (i.e. all state variables in \mathbf{u} and \mathbf{v} set to zero, all elements of \mathbf{q} equal to 1, no input), postulating full recovery from synaptic depression after the presentation of the previous stimulus block. The order of presentation of the different blocks was thus irrelevant. During the presentation of each block, we continuously recorded the state variables of all populations in the network as well as the synaptic efficacies for all connections. The sampling rate was 1 kHz.

Frequency response paradigm

To characterise the frequency response of the gerbil AC model, we presented the 0.656-s-SOI stimulus block at additional audio-frequencies, ranging from 0.25 kHz to 42.5 kHz, in half-octave steps (thus covering the full tonotopic map defined for the MGv columns). For the single-column model, the frequency response paradigm was already included in the RS paradigm, where the 0.656-s-SOI stimulus block was presented at nine different audio-frequencies, with ΔF ranging from 0 oct to 4 oct in half-octave steps.

1.1.5 Analysis of simulated data

Measures of neuronal activity

In our modelling approach, the state of excitatory and inhibitory neuronal populations is reflected by state variables (u and v , respectively) and population output is reflected by firing rates ($g(u)$ and $g(v)$). Moreover, as the state variables can be thought of as the populations' mean membrane potentials, $-\dot{u}$ and $-\dot{v}$ represent net current flow across the respective membrane. The chosen polarity matches the sign convention used in electrophysiology (see, for example, [Hammond, 2015](#)). Given that our in-vivo data reflected changes in current flow across the membranes of pyramidal neurons (see Section [2.1.3](#)), we based our analysis of in-silico RS dynamics on the net current flow reflected by $\xi_{\text{net}} = -\dot{u}$. We used Equations [3a-3c](#) to compute column-specific time courses $\xi_{\text{net}}(t)$ from the data recorded during the simulations ($\mathbf{u}(t)$, $\mathbf{v}(t)$, and $\mathbf{q}(t)$).

To investigate the response behaviour of individual columns in further detail, we also decomposed $\xi_{\text{net}}(t)$ into the different types of transmembrane current that make up this net flow. These types are reflected by the individual terms on the right-hand side (r.h.s.) of Equation [\(3a\)](#). The *first* term on the r.h.s. of Equation [\(3a\)](#) reflects a passive leak current to account for the membrane's finite resistance. For any given column i , the leak current associated with the excitatory population is

$$\xi_{\text{leak},i}(t) = \frac{u_i(t)}{\tau_m} . \quad (10)$$

The excitatory synaptic current caused by ee-connections (c.f. *second* term on the r.h.s. of Equation [\(3a\)](#)) is

$$\xi_{\text{ee},i}(t) = -\frac{1}{\tau_m} \mathbf{w}_{\text{ee},i} Q(t) g[\mathbf{u}(t)] , \quad (11)$$

where the weights of the ee-connections targeting column i are given by $\mathbf{w}_{\text{ee},i} = (W_{\text{ee}}(i, 1), \dots, W_{\text{ee}}(i, N))^{\top}$. The inhibitory synaptic current caused by ei-connections (c.f. *third* term on the r.h.s. of Equation [\(3a\)](#)), is

$$\xi_{\text{ei},i}(t) = \frac{1}{\tau_m} \mathbf{w}_{\text{ei},i} g[\mathbf{v}(t)] , \quad (12)$$

where vector $\mathbf{w}_{\text{ei},i} = (W_{\text{ei}}(i, 1), \dots, W_{\text{ei}}(i, N))^{\top}$ contains the weights of the ei-connections targeting column i . Finally, synaptic current induced via the afferent connection to reflect stimulus presentation (c.f. *fourth* term on the r.h.s. of Equation [\(3a\)](#)), is

area	index vector	column indices
MGv	\mathbf{j}_{MGv}	$(1, \dots, N_f)$
core	\mathbf{j}_{core}	$(N_f + 1, \dots, 3N_f)$
belt	\mathbf{j}_{belt}	$(3N_f + 1, \dots, 6N_f)$
parabelt	$\mathbf{j}_{\text{p.belt}}$	$(6N_f + 1, \dots, 9N_f)$

Table 6: Area-specific column indices for the gerbil AC model. N_f is the number of columns per field and was set to 16. The core area has two fields, the belt areas has three fields, and the parabelt area also has three fields (c.f. Figure 4).

$$\xi_{\text{aff},i}(t) = -\frac{i_{\text{aff},i}(t)}{\tau_m}. \quad (13)$$

Note that, for the single-column model, where $N = 1$, the equations introduced above are reduced to scalar form. Net transmembrane current flow for any column i is given by

$$-\dot{u}_i = \xi_{\text{net},i} = \xi_{\text{leak},i} + \xi_{\text{ee},i} + \xi_{\text{ei},i} + \xi_{\text{inp},i}. \quad (14)$$

To explore the response behaviour of individual columns in the gerbil AC model even further, we also decomposed the excitatory synaptic current $\xi_{\text{ee},i}$ according to the following regions of connection origin: MGv, core, belt, parabelt. Moreover, we treated recurrent excitation (where the index of the target column is equal to the index of the source column, $i = j$), as a separate category. Columns located in a given area have indices $j \in \mathbf{j}_{\text{area}}$. Table 6 specifies these vectors \mathbf{j} for the different areas. Thus, for a given column i , the vector

$$\mathbf{w}_{\text{ee},i}^{\text{area}} = (W_{\text{ee}}(i,j))_{j \in \mathbf{j}_{\text{area}}, j \neq i} \quad (15)$$

lists the weights of the intercolumnar connections originating in the respective area and targeting column i . The matching levels of synaptic efficacy are summarised in matrix

$$Q^{\text{area}}(t) = \text{diag}[(q_j(t))_{j \in \mathbf{j}_{\text{area}}, j \neq i}], \quad (16)$$

and the state variables of the connection sources are reflected by entries of vector

$$\mathbf{u}^{\text{area}}(t) = (u_j(t))_{j \in \mathbf{j}_{\text{area}}, j \neq i}. \quad (17)$$

The net excitatory synaptic current induced via intercolumnar connections originating in a given area is thus

$$\xi_{ee,i}^{\text{area}}(t) = -\frac{1}{\tau_m} \mathbf{w}_{ee,i}^{\text{area}} Q^{\text{area}}(t) g[\mathbf{u}^{\text{area}}(t)], \quad (18)$$

and the excitatory synaptic current induced via the recurrent, intracolumnar ee-connection is

$$\xi_{ee,i}^{\text{recurrent}}(t) = -\frac{1}{\tau_m} W_{ee}(i, i) q_i(t) g[u_i(t)]. \quad (19)$$

Preliminary assessment

In the gerbil AC model, our investigation of response behaviour was focussed on the primary auditory cortex, i.e. the 16 columns making up A1. We found that the responses that our stimuli evoked in these columns evolved over a time scale of about 300 ms. This is illustrated in Figure 8, where examples of time courses $\xi_{\text{net}}(t)$ recorded in response to the first stimulus within each block are plotted in a time window from -50 ms to 350 ms relative to stimulus onset. Given that our model was noise-free and the initial conditions for the presentation of each block were equal to the model's resting state, the response to the first stimulus within each block was identical across SOIs. After stimulus presentation, current flow $\xi_{\text{net}}(t)$ returned to baseline (zero) shortly after the 300-ms mark.

As shown in Figure 8, peak response amplitudes in the time window from 0 ms to 150 ms declined with increasing difference between i_{BF} (index of the MGv column where BF = stimulus audio-frequency, c.f. Figure 6) and $k(i)$ (Equation (4), intrafield index of the A1 columns). We found that $|i_{\text{BF}} - k(i)| = 4$ was the largest difference where the A1 columns showed a distinct response peak across all three settings of model parameter σ_{ie} . Thus, we limited our analysis of RS response behaviour to columns where $|i_{\text{BF}} - k(i)| \leq 4$. Note that, while $|i_{\text{BF}} - k(i)| = 0$ only applies to one A1 column, the other distances each apply to two columns. Thus, a total of nine A1 columns satisfied the condition. The complete overlap of the response curves for column pairs with the same $|i_{\text{BF}} - k(i)|$ -value reflects the symmetry of the connectivity patterns defined in the weight matrices.

After identifying the subset of responsive columns, we assessed the evolution of the peak amplitude $A_{i_{\text{stim}}}$ of the response in the time window from 0 ms to 150 ms for stimulus indices $i_{\text{stim}} = 1$ to n_{stim} . For SOIs < 3.5 s, stimulus repetition led to, at least, a 10% decrease in $A_{i_{\text{stim}}}$ relative to the first stimulus, i.e. for $i_{\text{stim}} > 1$, $(A_1 - A_{i_{\text{stim}}})/A_1 \geq 0.100$. The smaller the SOI, the larger the decrease. The model was thus exhibiting SOI-dependent RS. In 60% of the 540 cases (2 stimulus audio-frequencies \times 10 SOIs \times 3 σ_{ie} -values \times 9 columns), $A_{i_{\text{stim}}}$ was fully stable from the fifth stimulus onward ($A_5 = A_6 = A_7 = \dots = A_{n_{\text{stim}}}$). For an example of this response behaviour, see Figure 9a.

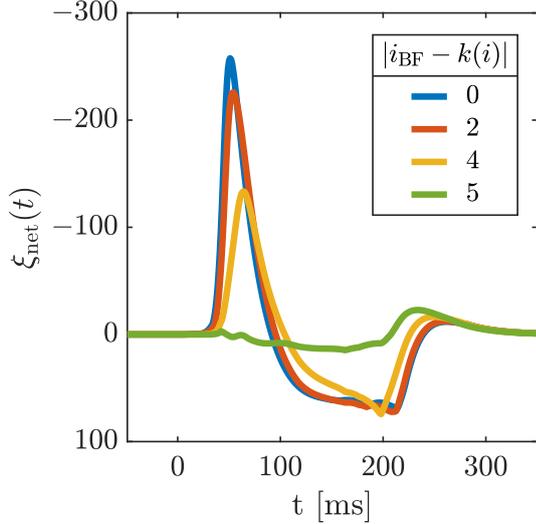


Figure 8: Column-specific responses, in terms of ξ_{net} , to the first stimulus in all blocks for $\sigma_{\text{ie}} = 7$ and $i_{\text{BF}} = 7$. Evoked responses unfold within a time window of about 300 ms and the amplitude of the response peak in the time window from 0 ms to 150 ms decreases with increasing difference $|i_{\text{BF}} - k(i)|$. At $|i_{\text{BF}} - k(i)| = 5$, columns are virtually unresponsive. Blue curve: A1 column 7; orange curve: A1 columns 5 and 9; yellow curve: A1 columns 3 and 11; green curve: A1 columns 2 and 12.

Cases where response amplitudes were not fully stable by the fifth stimulus were clustered at shorter SOIs. Stabilisation required up to $n_{\text{stable}} = 10$ stimulus repetitions – one example is illustrated in Figure 9b. Moreover, stabilisation was not absolute and we observed small cyclic patterns. One clear example of such a pattern is shown in Figure 9c. Compared to the change in amplitude relative to the first response, however, variations in peak amplitudes across the second half of the stimulus block were small. We defined the mean peak amplitude

$$A_{\text{stable}} = \left(\frac{1}{n_{\text{stim}} - n_{\text{stable}}} \right) \sum_{(n_{\text{stable}}+1)}^{n_{\text{stim}}} A_{i_{\text{stim}}} \quad (20)$$

and found that, at most, deviations from this mean were equal to 16% of the difference between A_1 and A_{stable} , i.e. for $i_{\text{stim}} > 10$, $|A_{\text{stable}} - A_{i_{\text{stim}}}| / (A_1 - A_{\text{stable}}) \leq 0.16$. 53% of the deviations $|A_{\text{stable}} - A_{i_{\text{stim}}}|$, were equal to less than 10% of the difference $(A_1 - A_{\text{stable}})$. Given that our gerbil AC model was noise free, these small variations were part of the network’s response to the RS paradigm. To determine lifetimes of RS for this model, we characterised stabilised RS for each SOI, column, stimulus audio-frequency, and σ_{ie} by the average response, in terms of ξ_{net} , across stimuli $n_{\text{stable}} + 1$ to n_{stim} . Thus, with time t expressed relative to stimulus onset, single trials in the time window $-50 \text{ ms} \leq t \leq 350 \text{ ms}$, or $-50 \text{ ms} \leq t \leq \text{SOI}$ if $\text{SOI} < 350 \text{ ms}$, were averaged across stimuli 11 to 20.

Examples of such condition-specific average time courses $\overline{\xi_{\text{net}}(t)}$ are shown in Figure 10a and b. We determined the absolute amplitude $A(\text{SOI})$ of the associated response peak in the time window from 0 ms to 150 ms. For each model and stimulus parameter setting employed in our simulations, $A(\text{SOI})$ gradually increased with SOI and converged towards a saturation value, as is characteristic of the RS phenomenon. Table 7 provides an overview of all conditions investigated in our simulations.

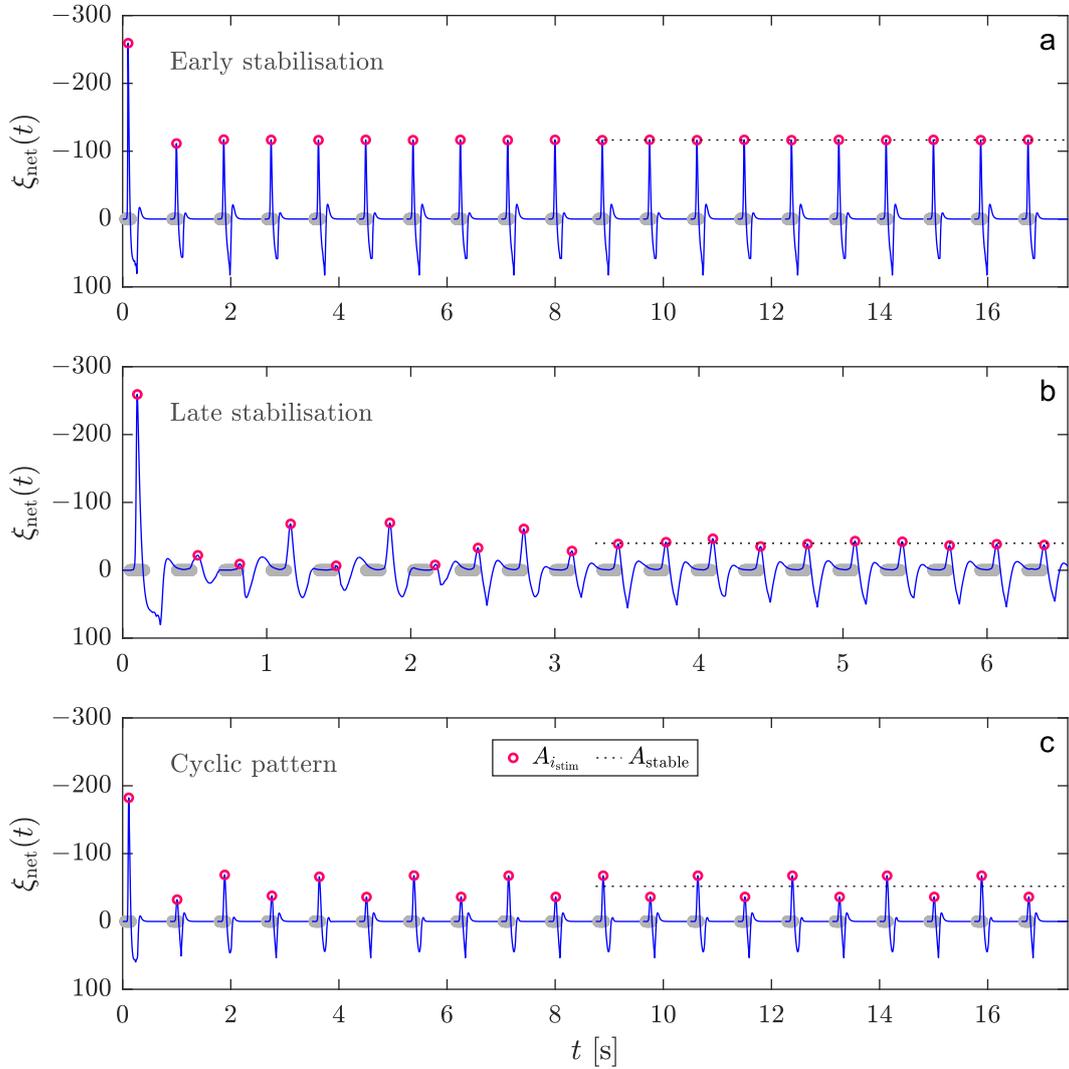


Figure 9: Examples of the response behaviour of A1 column 7, in terms of ξ_{net} , observed for different stimulus blocks. The grey horizontal markers reflect stimulus onset and duration, red circles mark response peaks, and the dashed line indicates mean peak amplitude A_{stable} . **a:** For a SOI of 0.875 s and a stimulus audio-frequency of 2 kHz, the column's response peaks stabilise within the first five stimulus repetitions. **b:** For a SOI of 0.328 s and a stimulus audio-frequency of 2 kHz, the column's response peaks only stabilise after ten stimulus repetitions. Stabilisation is not absolute but deviations from the mean amplitude A_{stable} are small. **c:** For a SOI of 0.875 s and a stimulus audio-frequency of 8 kHz, the column exhibits a clear cyclic response pattern but response peaks remain close to the mean amplitude A_{stable} when compared to the difference between A_1 and A_{stable} . Note that while peak amplitudes have negative polarity in this figure, peak amplitude $A(\text{SOI})$ determined from the mean across stimulus repetitions was an absolute value without polarity in subsequent analysis steps.

model type	stimulus audio-frequency	SOI [s]	model parameter setting	examined columns	total
gerbil AC	two options: 2 kHz, 8 kHz	ten options: 0.328, 0.438, 0.656, 0.875, 1.313, 1.750, 2.626, 3.500, 5.000, 7.000	three options: intrafield $\sigma_{ie} \in [4, 5, 7]$	nine options: A1 columns where $ i_{BF} - k(i) \leq 4$	540 conditions
single column	nine options: 0 oct \leq $\Delta F \leq 4$ oct (0.5-oct steps)	ten options: 0.219, 0.328, 0.438, 0.656, 0.875, 1.313, 1.750, 2.626, 3.500, 7.000	three options: $\tau_{rec,aff} \in$ [0.5, 1.0, 1.5] s	one option	270 conditions

Table 7: Overview of experimental conditions investigated in our simulations

In the single-column model, we found that the responses evoked by the stimuli evolved over a time scale of about 200 ms. Thus, unlike for the gerbil AC model, we included data recorded for the 219-ms SOI in further analysis. Moreover, even stimulation at the largest ΔF (4 oct) used for the RS paradigm showed a clear response peak across all SOI- and $\tau_{rec,aff}$ -values. Consequently, all recorded data was included in further analysis. For SOIs < 2.626 s, stimulus repetition led to, at least, a 10% decrease in $A_{i_{stim}}$ relative to the first stimulus (for $i_{stim} > 1$, $(A_1 - A_{i_{stim}})/A_1 \geq 0.100$) and the smaller the SOI, the larger the decrease. The model was thus exhibiting SOI-dependent RS. $A_{i_{stim}}$ was always fully stable from the fifth stimulus onward ($A_5 = A_6 = A_7 = \dots = A_{n_{stim}}$). To harmonise the analysis pipelines, we characterised stabilised RS in the single-column model via the same approach as used for the gerbil AC model. Thus, although technically unnecessary, we computed the average response $\overline{\xi_{net}}$ across stimuli $n_{stable} + 1$ to n_{stim} , i.e. stimuli 11 to 20, separately for each SOI, stimulus audio-frequency, and $\tau_{rec,aff}$ -value.

Examples of the resulting condition-specific average time courses $\overline{\xi_{net}(t)}$ are shown in panels a and b of Figure 11. Again, as is characteristic of the RS phenomenon, peak response amplitude $A(\text{SOI})$ observed in the time window from 0 ms to 80 ms gradually increased with SOI and converged towards a saturation value at the largest SOI. The SOI-dependence of the peak amplitude exhibited this characteristic increase for each of the model and stimulus parameter settings employed in our simulations. Table 7 provides an overview of all the investigated conditions.

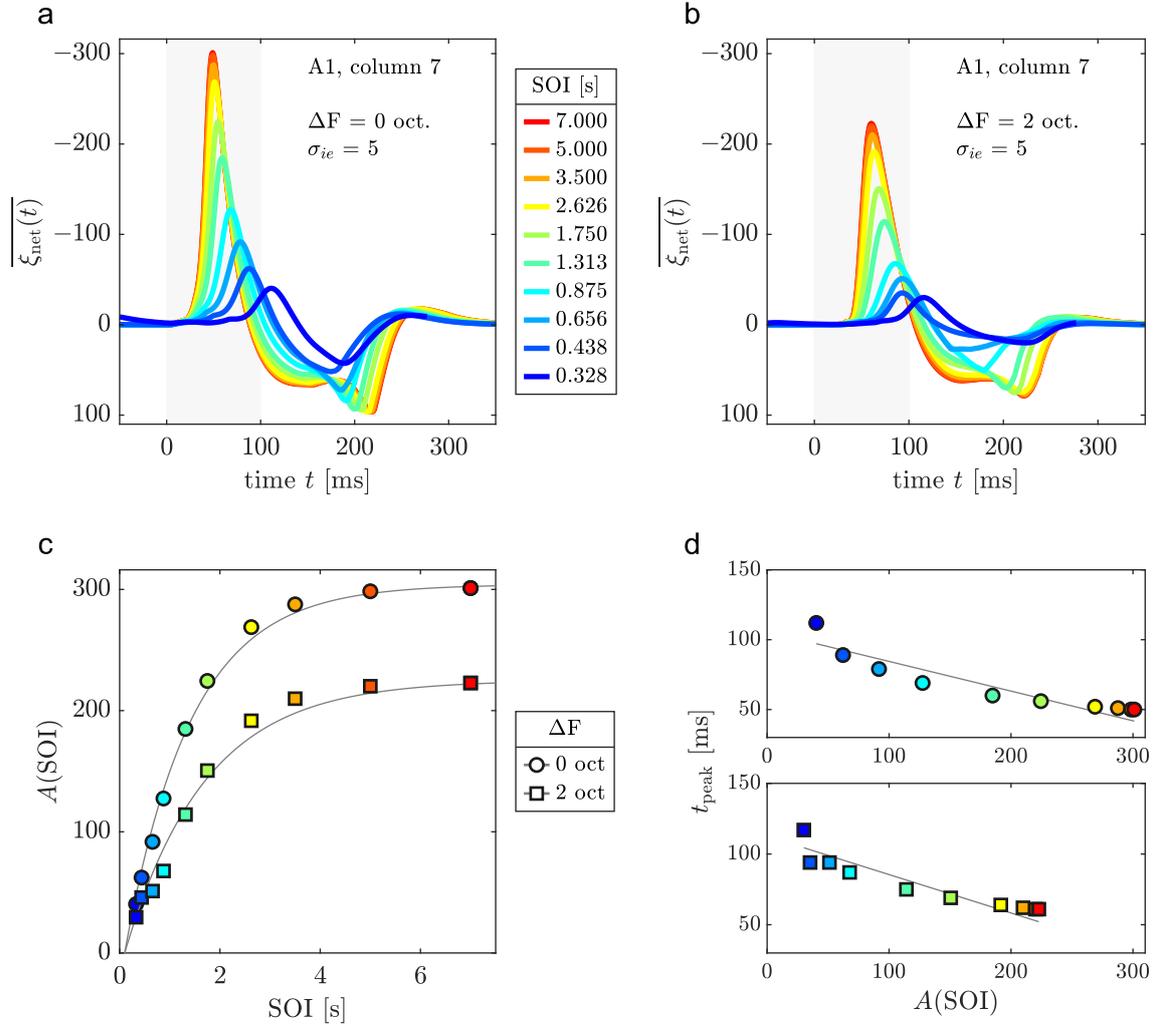


Figure 10: Overview of results for column 7 in A1 of the gerbil AC model with lateral range of effectively inhibitory connections, σ_{ie} , equal to 5. **a** and **b**: SOI-specific response, in terms of ξ_{net} , for stimulation at $\Delta F = 0$ oct (left) and at $\Delta F = 2$ oct (right). **c**: Circles indicate peak response amplitude plotted as a function of SOI for $\Delta F = 0$ oct, squares for $\Delta F = 2$ oct. The two grey curves are plots of Equation (1) fitted to the respective data set. **d**: Circles indicate peak latency plotted as a function of SOI-dependent peak amplitude for stimulation at $\Delta F = 0$ oct, squares for stimulation at $\Delta F = 2$ oct. The dark grey line in each panel is a linear regression (c.f. Equation (21)).

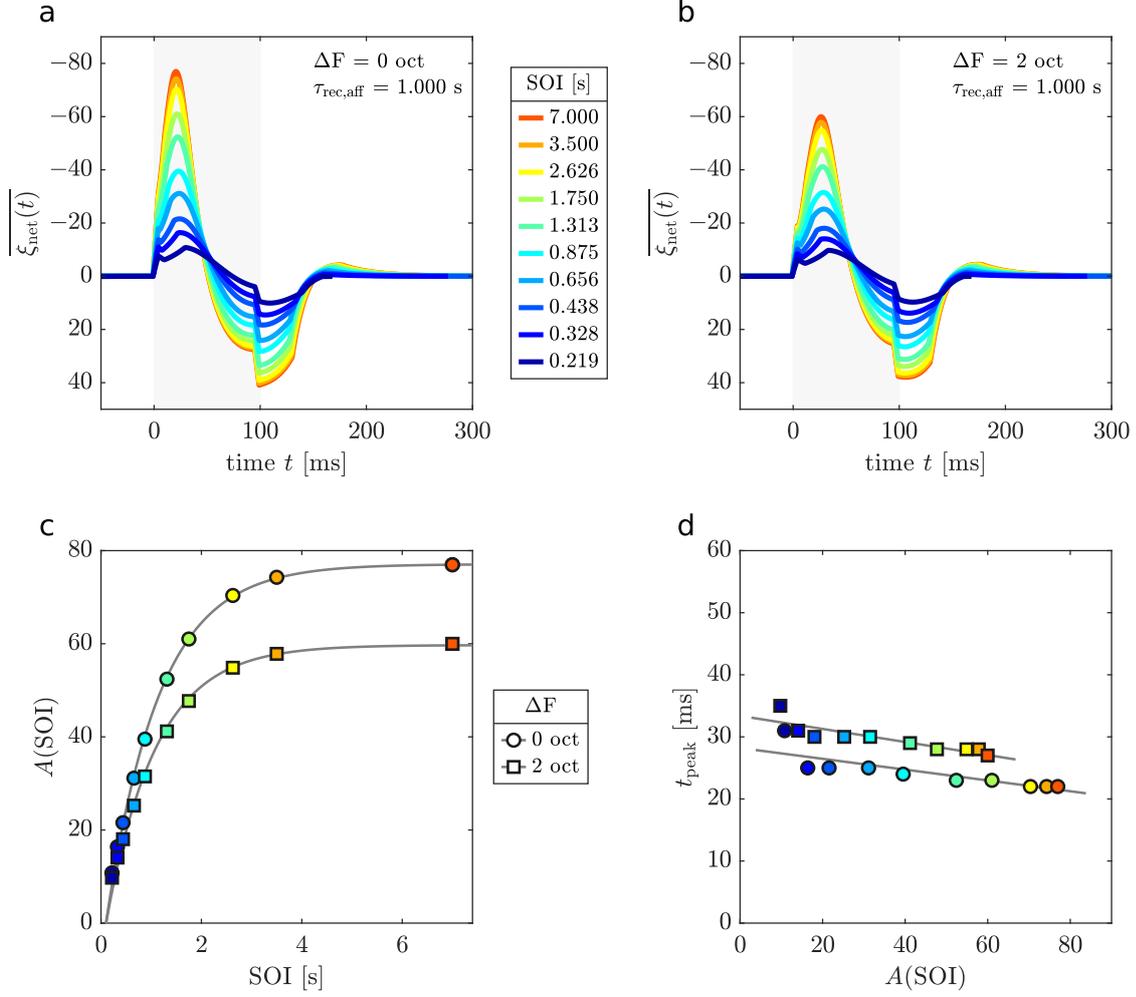


Figure 11: Overview of results for the single-column model when $\tau_{\text{rec,aff}} = 1.000$ s. **a** and **b**: Time course of $\overline{\xi_{\text{net}}}(t)$ per SOI for stimulation at $\Delta F = 0$ oct (left) and at $\Delta F = 2$ oct (right). **c**: Circles indicate response amplitude plotted as a function of SOI for $\Delta F = 0$ oct, squares for $\Delta F = 2$ oct. The two grey lines are plots of Equation (1) fitted to the respective data set. **d**: Circles indicate peak latency plotted as a function of SOI-dependent peak amplitude for stimulation at $\Delta F = 0$ oct, squares for stimulation at $\Delta F = 2$ oct. The dark grey curves are linear regressions (c.f. Equation (21)) for the respective data set.

Quantification of repetition suppression lifetime

To quantify the lifetime of RS, we fitted Equation (1) to the amplitude values $A(\text{SOI})$ observed for $\overline{\xi_{\text{net}}}$ (time window 0 ms to 150 ms w.r.t. stimulus onset for the gerbil AC model, 0 ms to 80 ms for the single-column model). The parameter t_0 was set to the stimulus duration (0.100 s), reflecting the logic that once SOI equals stimulus duration and individual tones merge into a continuous tone, no evoked response peak should be detectable beyond the first stimulus. To ensure robust fitting results, the fitting algorithm consisted of a two-step process. First, we applied a non-iterative regression method based on an appropriate integral equation (Jacquelin, 2009). Next, the values for τ_{SOI} and A_{sat} obtained via this approach were used as the starting points for a further improvement of the fit via an iterative least squares approach (lsqnonlin function, MATLAB R2021b). Ultimately, we thus obtained six τ_{SOI} -values per column for the gerbil AC model (two options for stimulus audio-frequency \times three options for model parameter σ_{ie}) and 27 values for the single-column model (nine options for stimulus audio-frequency \times three options for model parameter $\tau_{\text{rec,aff}}$). As an example, Figures 10c and 11c show the peak amplitudes from respective panels a and b as a function of SOI along with the corresponding fits of Equation (1) (grey curves).

Quantification of peak latency shift

In both the gerbil AC and the single-column model, we observed a shift in peak latency as a function of SOI. In most cases, peak latency t_{peak} increased with decreasing SOI and thus with decreasing peak response amplitude $A(\text{SOI})$ (see, for example, Figures 10a and b). For all simulation conditions, the relationship between SOI-specific peak amplitude $A(\text{SOI})$ and peak latency t_{peak} could be coarsely summarised by the linear function

$$\ell_{\text{fit}}(A(\text{SOI})) = \mu A(\text{SOI}) + c, \quad (21)$$

where fitting parameter μ reflects direction and magnitude of the peak latency shift and c is the x-axis intercept. As an example, Figures 10d and 11d show plots of t_{peak} as a function of $A(\text{SOI})$, deduced from the data in panels a and b, and the corresponding linear regressions (Equation (21), grey lines). Ultimately, we normalised peak amplitudes relative to the respective 7-s SOI peak amplitude before carrying out linear regressions to achieve better comparability of μ -values across conditions.

Computation of frequency response functions

For the FR paradigm, we characterised the audio-frequency-specific responses of each column via the same single-trial averages as the SOI-specific responses for the RS paradigm, i.e. we computed the average response, in terms of ξ_{net} , across stimuli $n_{\text{stable}} + 1$ to n_{stim} for each audio-frequency and model parameter setting. Next, we determined the amplitude of the response peak (time window 0 ms to 150 ms w.r.t. stimulus onset for the gerbil AC model, 0 ms to 80 ms for the single-column model). Finally, FR functions for individual columns were obtained by plotting the respective peak amplitudes as a function of stimulus audio-frequency (c.f. Figures 12 and 19).

1.2 Simulation results and predictions

1.2.1 Response behaviour of the single-column model

Frequency response

Figure 12 shows the FR functions obtained for the single-column model by varying ΔF , i.e. the amplitude of the presynaptic firing rate of the afferent connection (c.f. Table 4). Overall, Figure 12 confirms that the approach we devised to reflect stimulus audio-frequency in the single-column model leads to single-peaked FR functions comparable to the average FR functions the literature reports for neural populations in the gerbil AC (e.g. Happel et al. 2010, Figure 3A; Deane et al. 2020, Figure 3B; Ma & Brunk et al. 2021, Figure 1G). The model parameter $\tau_{\text{rec,aff}}$, the time constant of recovery from STSD at the afferent connection, has an effect on the model’s FR function. With increasing $\tau_{\text{rec,aff}}$, the peak amplitude at $\Delta F = 0$ oct decreases from 38.9 to 31.1 to 26.3. Additionally, the FR function becomes less sharp. FR function width, in terms of the ratio between the peak response amplitude at $\Delta F = 2$ oct and $\Delta F = 0$ oct increases from 0.760 to 0.811 to 0.846.

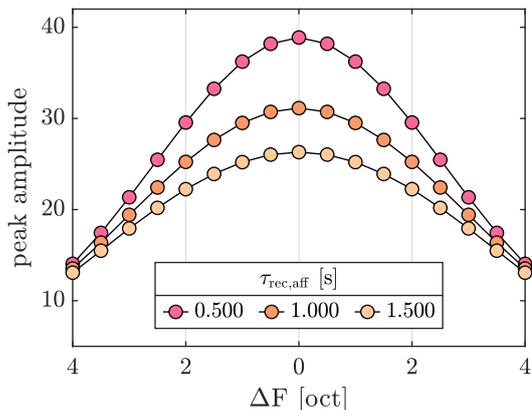


Figure 12: FR functions of the single-column model for the three different $\tau_{\text{rec,aff}}$ -values used during simulations. Results are mirrored at the axis $\Delta F = 0$ oct to reflect stimulus audio-frequencies above as well as below the BF.

Lifetime of repetition suppression and peak latency shift

Figure 13 summarises the RS response dynamics of the single-column model for a $\tau_{\text{rec,aff}}$ -value of 0.500 s (black markers), 1.000 s (grey markers), and 1.500 s (white markers). Compared to STSD at the recurrent connections, this corresponds to faster, equal, or slower recovery from STSD at the afferent connection (c.f. Table 5). As we describe our results in this context, we refer to the three parameter settings with these descriptive terms in the text below as well as in the legends of Figures 13, 16, and 17.

In Figure 13a, RS lifetimes τ_{SOI} are plotted as a function of ΔF . As illustrated by the white triangles, τ_{SOI} *decreases* with increasing distance from the BF (i.e. with increasing ΔF) when recovery from STSD is slower at the afferent connection. The τ_{SOI} -value starts out at 1.284 s for $\Delta F = 0$ oct and decreases to 0.989 s for $\Delta F = 4$ oct. In contrast, the black triangles show that τ_{SOI} *increases* with increasing distance from the BF when STSD recovery is faster at the afferent connection. For this scenario, the τ_{SOI} -value starts out at 0.762 s for $\Delta F = 0$ oct and reaches 0.883 s for $\Delta F = 4$ oct. When τ_{rec} and $\tau_{\text{rec,aff}}$ are equal, τ_{SOI} shows a milder decrease with increasing distance from the BF. The τ_{SOI} -value starts out at 1.049 s for $\Delta F = 0$ oct and decreases to 0.939 s for $\Delta F = 4$ oct.

Figure 13b presents results regarding response peak latencies. The top panel summarises peak latency shift μ (c.f. Equation (21)), as a function of ΔF . The absolute value of μ always exhibits a mild increase with distance from the BF, however the polarity and magnitude varies. For slower afferent STSD recovery, latency shift μ is positive (black circles). This means that peak latency is delayed with *increasing* SOI. Note, however, that this shift is very minor, with $\mu \leq 5$ ms for $\Delta F < 3$ oct and a maximum value of 11 ms. For faster afferent STSD recovery, latency shift μ is negative peak (white circles). This means that peak latency is delayed with *decreasing* SOI. The extent of this delay increases with ΔF . The μ -value starts out at -13 ms for $\Delta F = 0$ oct and reaches -32 ms for $\Delta F = 4$ oct. When STSD recovery dynamics are equal for all connections, there is a mild delay in peak latency with decreasing SOI. The extent of this delay is fairly stable across all but the largest ΔF : for $\Delta F < 4$ oct, μ -values range from -6 ms to -8 ms and for $\Delta F = 4$ oct, $\mu = -15$ ms.

The bottom panel of Figure 13b shows the median response peak latency, $m(t_{\text{peak}})$, across all ten SOI values as a function of ΔF . Peak latencies increased by about 40 ms across ΔF -values, and there was little difference between $\tau_{\text{rec,aff}}$ -settings (3 ms at most).

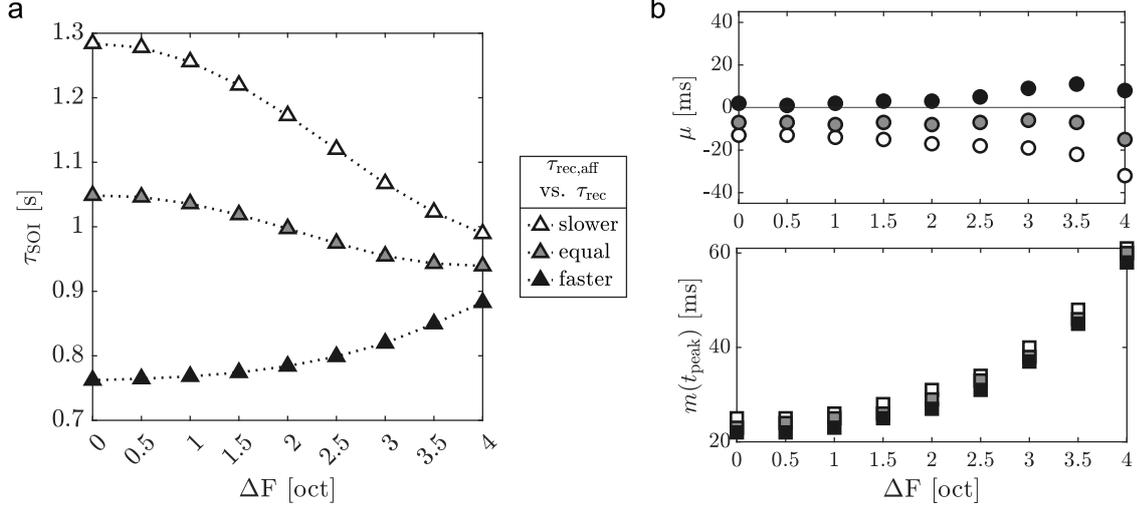


Figure 13: RS lifetimes and peak latency variations in the single-column model. **a:** RS lifetime τ_{SOI} as a function of ΔF when, relative to the recurrent connections, STSD recovery at the afferent connection is slower (black markers), equal (grey markers), or faster (white markers). For the faster condition, τ_{SOI} increases, whereas for the slower condition, τ_{SOI} decreases with distance from the BF. **b:** The top panel summarises RS-induced peak latency shift, reflected by fitting parameter μ , and the bottom panel shows the median response peak latency $m(t_{\text{peak}})$, across all ten SOI values, as a function of ΔF for the three different $\tau_{\text{rec,aff}}$ -settings. The same colour code as in panel a applies.

In summary, we thus observed:

1. a decrease or an increase in τ_{SOI} as a function of ΔF , depending on STSD recovery dynamics.
2. an increase in median peak latency as a function of ΔF
3. an increase in the magnitude of the peak latency shift μ as a function of ΔF
4. a positive or a negative the peak latency shift μ , depending on STSD recovery dynamics

Response measures μ , $m(t_{\text{peak}})$, and τ_{SOI} were based on the net transmembrane current associated with the excitatory population of the model. In the following sections, we will explore the model's response behaviour in terms of the different subtypes of transmembrane current that make up this net flow (c.f. Equations (10) to (13)). This was done in order to study which neuronal interactions are contributing to the variations in RS lifetime and peak latency.

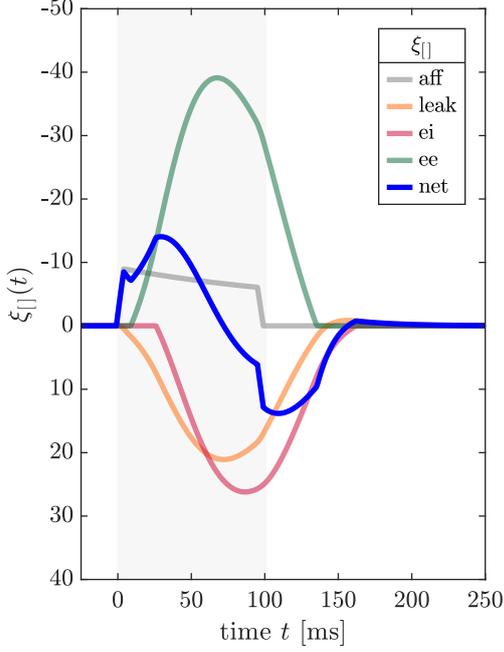


Figure 14: Break-down of the time course of the single-column model’s stabilised response into the components that make up the net current flow. In this example, $\tau_{\text{rec,aff}} = \tau_{\text{rec}}$, $\text{SOI} = 0.328$ s, and $\Delta F = 2$ oct. Stimulus onset and duration are mapped out via the grey shading in the background. Grey curve: afferent input, ξ_{aff} , reflecting stimulus presentation (Equation (13)); orange curve: leak current, ξ_{leak} (Equation (10)); red curve: inhibitory synaptic current, ξ_{ei} (Equation (12)); green curve: excitatory synaptic current, ξ_{ee} (Equation (11)); blue curve: net current flow, ξ_{net} (Equation (14)).

Components of the single-column model’s response

For one example condition, Figure 14 breaks down the net response of the single-column model into excitatory and inhibitory synaptic current, leak current, and afferent current. The illustrated response sequence could be observed across all conditions:

1. Afferent input drives the response via ξ_{aff} and moves the column away from its resting state ($u = 0, v = 0$), which first activates the passive leak current, ξ_{leak} .
2. After a certain delay relative to stimulus onset, the excitatory population’s firing threshold is crossed and the population begins to respond actively ($g(u) > 0$), leading to an excitatory synaptic current, ξ_{ee} , via the population’s recurrent ee-connection.
3. Activity of the excitatory population also causes the activation of the inhibitory population ($g(v) > 0$) via the ie-connection, which in turn causes inhibitory synaptic current, ξ_{ei} , targeting the excitatory population via the ei-connection.
4. After stimulus offset, current flow steadily decreases and the column returns to its resting state.

While this basic response sequence was preserved across conditions, the relative contribution of the individual current components to the net response varied. This in turn affected the response measures τ_{SOI} and μ . The underlying mechanisms will be explored in the following sections.

Effect of short-term synaptic depression on the afferent current

Recovery from STSD at the afferent connection, i.e. the connection that relayed pre-synaptic activity reflecting stimulus presentation, was governed by time constant $\tau_{\text{rec,aff}}$. The effect of $\tau_{\text{rec,aff}}$ on ξ_{aff} is illustrated in Figure 15, where time courses $\xi_{\text{aff}}(t)$ for the first 4 s of a stimulus block are shown for the three different $\tau_{\text{rec,aff}}$ -values. The four panels summarise responses for two different SOI values and two different stimulus audio-frequencies, respectively. Stimulus repetition led to depression at the afferent synapse which in turn suppressed single-trial responses. Across all panels, the effect of $\tau_{\text{rec,aff}}$ on the level of this suppression becomes apparent. The longer the time constant, the slower the recovery from STSD and thus the more suppressed the single-trial response when the stimulus is repeated. Compare for example Figures 15b and d, where SOI is equal to 1.750 s: for a $\tau_{\text{rec,aff}}$ of 0.500 s, single-trial time courses of ξ_{aff} are close to identical across stimulus repetition, whereas for a $\tau_{\text{rec,aff}}$ of 1.500 s, there is a clear decline in response. For the shorter SOI of 0.438 s (Figures 15a and c), recovery from STSD is incomplete across all three $\tau_{\text{rec,aff}}$ -values and responses decrease continuously across several stimulus repetitions but the decline is again most pronounced for the longest $\tau_{\text{rec,aff}}$. Across condition, single-trial response were always stabilised by stimulus five.

Across stimulus audio-frequencies, peak response amplitudes were larger the smaller ΔF . This reflects the audio-frequency-specific scaling we imposed on the presynaptic firing rate (c.f. Table 4). Across the duration of each stimulus, a continuous decline in ξ_{aff} -amplitude could be observed. This occurred across all conditions and reflected the decline in synaptic efficacy brought on by STSD and governed by time constant $\tau_{\text{on,aff}}$ (see Equation (8)). Thus, in summary:

- The afferent current ξ_{aff} exhibits SOI-specific RS. Single-trial responses decrease across stimulus repetitions and stabilise after a few repetitions. The amplitude of the stabilised response is smaller the shorter the SOI.
- For a given SOI, the amplitude of the stabilised ξ_{aff} response decreases with increasing $\tau_{\text{rec,aff}}$, i.e. RS is more pronounced the longer $\tau_{\text{rec,aff}}$.
- ΔF affects response amplitudes, reflecting the audio-frequency-specific scaling imposed on the presynaptic firing rate (c.f. Table 4)

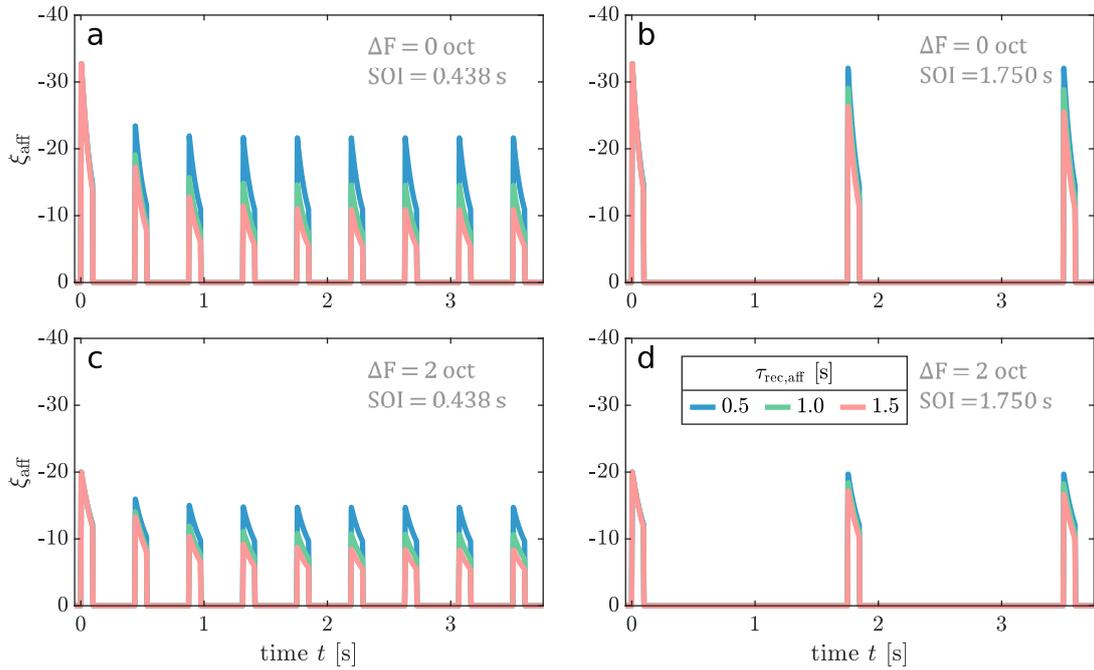


Figure 15: Afferent excitatory current ξ_{aff} induced at the excitatory population of the single-column model when simulating the presentation of pure tone stimuli with a SOI of 0.438 s (left) or a SOI of 1.750 s (right) and an audio-frequency characterised by $\Delta F = 0$ oct (top) or $\Delta F = 2$ oct (bottom). Blue curves reflect simulation results for $\tau_{\text{rec,aff}} = 0.500$ s, green curves results for $\tau_{\text{rec,aff}} = 1.000$ s, and pink curves for $\tau_{\text{rec,aff}} = 1.500$ s. The afferent current exhibits RS, single-trial responses decrease with stimulus repetition and the longer $\tau_{\text{rec,aff}}$, the more pronounced the suppression. Additionally, response amplitudes are affected by ΔF , reflecting the audio-frequency-specific scaling imposed on the presynaptic firing rate (c.f. Table 4).

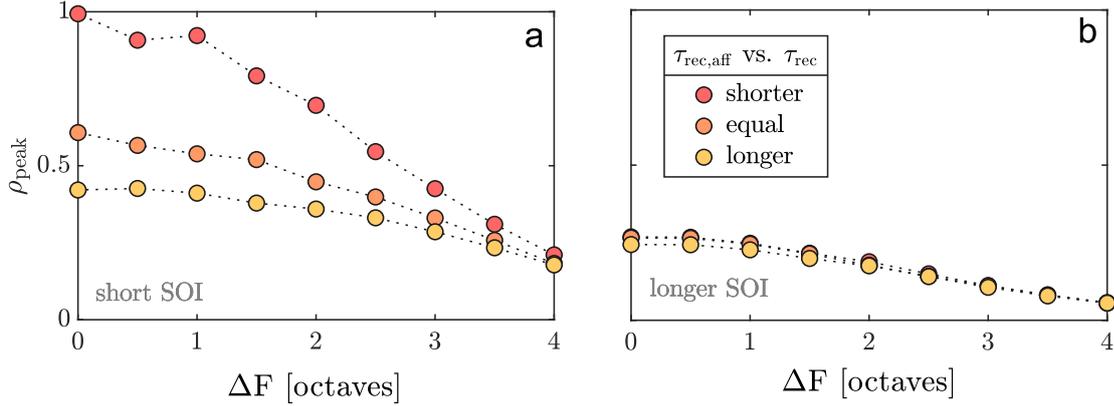


Figure 16: Single-column model response behaviour – ratio ρ_{peak} of currents ξ_{aff} and ξ_{ee} at t_{peak} (Equation (22)) as a function of ΔF for a short SOI (0.328 s, panel a) and a longer SOI (2.626 s, panel b) and all three $\tau_{\text{rec,aff}}$ -values (faster, red markers; equal, orange markers; slower, yellow markers). $\tau_{\text{rec,aff}}$ has a strong effect on ρ_{peak} when SOI is short and ΔF is small.

Mechanisms behind variations in the lifetime of repetition suppression

A response peak in ξ_{net} is observed at time t_{peak} because excitatory synaptic current, namely the sum of ξ_{aff} and ξ_{ee} , maximally exceeds the inhibitory current, $\xi_{\text{leak}} + \xi_{\text{ie}}$, at this moment (c.f. Figure 14). While this was the case across all $\tau_{\text{rec,aff}}$ -values, SOIs, and ΔF -values, the relative contribution of the two excitatory synaptic currents to the net response varied as a function of all these factors. To capture the variation, we defined the ratio

$$\rho_{\text{peak}} = \frac{\xi_{\text{aff}}(t_{\text{peak}})}{\xi_{\text{ee}}(t_{\text{peak}})}. \quad (22)$$

In Figure 16, ρ_{peak} is plotted as a function of ΔF for a short SOI (a) and a longer SOI (b) and for all three $\tau_{\text{rec,aff}}$ -values. For both SOIs, ρ_{peak} decreased with ΔF . This is easily explained by the decrease in presynaptic firing rates, i.e. maximum values of ξ_{aff} , that goes along with the increase in ΔF (c.f. Table 4).

For the shorter SOI, ρ_{peak} showed a large dependence on $\tau_{\text{rec,aff}}$, fanning out at smaller ΔF and converging at larger ΔF . In contrast, ρ_{peak} -values converged across $\tau_{\text{rec,aff}}$ -values in the longer SOI condition. This observation also applied to the other SOI-values. For SOIs ≥ 2.626 s, values of ρ_{peak} were very similar across $\tau_{\text{rec,aff}}$ -values, and differed by no more than 10%. For SOIs < 2.626 s and small ΔF , ρ_{peak} -values differed significantly across $\tau_{\text{rec,aff}}$ -values, remaining below 0.5 for the longer $\tau_{\text{rec,aff}}$ condition and approaching values of 1 for the shorter condition. A ρ_{peak} -value $\simeq 1$ means that afferent synaptic current ξ_{aff} made up about 50% of the net excitatory synaptic current at t_{peak} when STSD recovery was faster at the afferent than at the recurrent connection. This relative contribution to the net excitatory synaptic current decreased as a function of SOI and ΔF .

The observations described above can be related to the τ_{SOI} -distribution shown in Fig-

ure 13a. The τ_{SOI} -values reflect the amalgamation of the $\tau_{\text{rec,aff}}$ - and τ_{rec} -value influencing the time course of ξ_{aff} and ξ_{ee} . The amalgamation is partially based on the relative contribution of the respective current type to the net response peak. For $\Delta F = 4$ oct, the relative contribution of ξ_{aff} is small across all three $\tau_{\text{rec,aff}}$ -values ($\rho_{\text{peak}} < 0.300$ for all investigated SOIs) and thus τ_{SOI} is a close reflection of τ_{rec} irrespective of $\tau_{\text{rec,aff}}$. With τ_{rec} set to 1.000 s (c.f. Table 5), values of τ_{SOI} ranged from 0.882 s to 0.989 s.

At the other end of the ΔF -range, i.e. for $\Delta F = 0$ oct, the relative contribution of ξ_{aff} was larger and, additionally, varied as a function of $\tau_{\text{rec,aff}}$. For the shorter- $\tau_{\text{rec,aff}}$ condition and short SOIs, ξ_{aff} made up almost half of the net excitatory synaptic current at t_{peak} . The resulting τ_{SOI} -value, being equal to 0.762 s, lies about halfway between $\tau_{\text{rec,aff}}$ (0.500 s) and τ_{rec} (1.000 s). In contrast, for the longer- $\tau_{\text{rec,aff}}$ condition, the relative contribution of ξ_{aff} was equal to just under 30% for short SOI values. The resulting τ_{SOI} -value, being equal to 1.284 s, also lies about halfway between $\tau_{\text{rec,aff}}$ (1.500 s) and τ_{rec} (1.000 s). Here, we additionally need to consider that, unless ρ_{peak} is very small, recurrent excitation is limited in its ability to compensate for the STSD affecting ξ_{aff} . For $\tau_{\text{rec,aff}} = 1.500$ s, STSD is quite pronounced (c.f. pink curves in Figure 15).

In summary, the τ_{SOI} distribution observed for the single-column model (c.f. Figure 13a) can thus be explained as follows:

- Irrespective of $\tau_{\text{rec,aff}}$, τ_{SOI} -values converge towards a value close to τ_{rec} with increasing ΔF because the relative contribution of ξ_{ee} to the net response peak increases whereas the relative contribution of ξ_{aff} declines.
- For small ΔF , the relative contribution of ξ_{aff} to the net response peak is more significant and τ_{SOI} -values reflect an amalgamation of $\tau_{\text{rec,aff}}$ and τ_{rec} . Therefore, $\tau_{\text{SOI}} > \tau_{\text{rec}}$ for $\tau_{\text{rec,aff}} > \tau_{\text{rec}}$, $\tau_{\text{SOI}} \simeq \tau_{\text{rec}}$ for $\tau_{\text{rec,aff}} = \tau_{\text{rec}}$, and $\tau_{\text{SOI}} < \tau_{\text{rec}}$ for $\tau_{\text{rec,aff}} < \tau_{\text{rec}}$.
- In consequence, τ_{SOI} increases as a function of increasing ΔF when $\tau_{\text{rec,aff}} < \tau_{\text{rec}}$, stays fairly stable when $\tau_{\text{rec,aff}} = \tau_{\text{rec}}$, and decreases when $\tau_{\text{rec,aff}} > \tau_{\text{rec}}$.

Mechanisms behind variations in response peak latency

As described in Section 1.2.1, [Components of the single-column model's response](#), and illustrated in Figure 14, the column's excitatory population only starts to fire after a certain delay relative to stimulus onset, when the firing threshold θ is crossed due to the induced afferent current ξ_{aff} (c.f. Point 2 of the description of the column's response sequence). The population then starts to further excite itself via the recurrent ee-connection (ξ_{ee}). Figure 17 illustrates the firing onset latency, t_{fire} , as a function of ΔF for a shorter SOI (0.328 s, panel a) and a longer SOI (2.626 s, panel b), and all three $\tau_{\text{rec,aff}}$ options. For both SOIs, the firing onset latency increases by at least 13 ms from $\Delta F = 0$ oct to $\Delta F = 4$ oct. This increase in t_{fire}

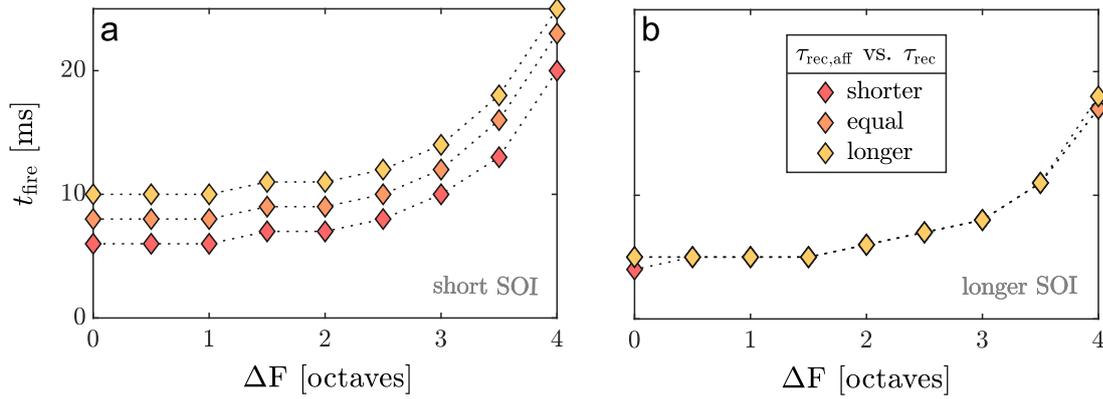


Figure 17: Single-column model response behaviour – firing onset latency t_{peak} of the excitatory population as a function of ΔF . In panel a, SOI is short (0.328 s) and in panel b, it is longer (2.626 s). An increase in ΔF delays firing onset.

as a function of ΔF was observed for all investigated conditions. The effect can be explained by the aforementioned decrease in ξ_{aff} -amplitude that accompanies the increase in ΔF . The smaller ξ_{aff} , and thus the resulting change in state variable u per increment of time, the longer it takes for the population to reach its firing threshold.

Additionally, decreasing ξ_{aff} -amplitudes also affected the rising slope of ξ_{ee} . The respective gradient was always steeper the smaller ΔF was. Figure 18 illustrates this for $\tau_{\text{rec,aff}} = \tau_{\text{rec}}$. An overlay of time courses of ξ_{ee} is shown for the different ΔF -values and the three panels reflect three different SOI-values. In the first 30 to 50 ms, the darker and lighter green curves do not overlap and the magnitude of the respective gradient is always larger for smaller ΔF . This difference further delays the latency of the ξ_{ee} peak, and therefore the net response peak, as a function of increasing ΔF . In summary, the delay in median peak latency as a function of increasing ΔF observed for the single-column model (c.f. lower panel of Figure 13b) can be explained as follows: The amplitude of ξ_{aff} decreases as a function of ΔF . This delays the firing onset of the excitatory population and decreases the gradient of the rising slope of ξ_{ee} , which in turn delays t_{peak} .

As discussed in Section 1.2.1, [Effect of short-term synaptic depression on the afferent current](#), ξ_{aff} -amplitudes for a given ΔF were affected by both SOI and $\tau_{\text{rec,aff}}$. For a shorter $\tau_{\text{rec,aff}}$, the effect of STSD on SOI-specific ξ_{inp} -amplitudes was comparatively mild (c.f. blue curves in Figure 15). In consequence, firing onset latencies did not differ by more than 5 ms across SOIs (e.g. compare red markers in Figure 17a vs. b). Peak latencies of the net response ξ_{net} were thus relatively stable across SOI-values and the magnitude of the latency shift μ was close to zero for $\Delta F < 3$ (c.f. black markers in the upper panel of Figure 13b). For large ΔF , we observed a mild delay in peak latency with *increasing* SOI. This can be explained by the slightly larger contribution of the early-peaking ξ_{aff} to the net peak for shorter than for longer SOIs (c.f. Figure 16, a vs. b).

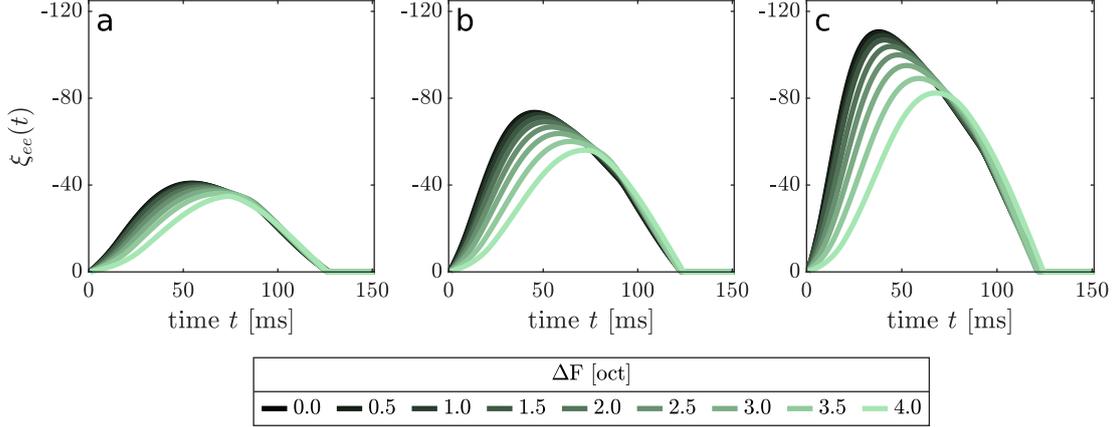


Figure 18: Time courses of synaptic current ξ_{ee} in the single-column model ($\tau_{rec}\text{-ratio} = 1$) for the nine different stimulus audio-frequencies ($\Delta F = 0$ oct to $\Delta F = 4$ oct, darkest to lightest green curve). To compare the gradients of the condition-specific time courses, they were aligned relative to the respective firing onset latency, i.e. $t = 0$ ms marks firing onset rather than stimulus onset. SOI-values increase from left to right, in panel a, SOI = 0.328 s, in panel b, SOI = 0.656 s, and in panel c, SOI = 1.313 s. For time points preceding the peak response, the magnitude of the gradient increases with decreasing ΔF . Moreover, it increases with increasing SOI.

For a longer $\tau_{rec,aff}$, the scenario was different. The effect of STSD on ξ_{aff} -amplitudes was comparatively strong (c.f. pink curves in Figure 15) and, in consequence, firing onset latencies and gradients of ξ_{ee} were shorter/shallower the shorter the SOI. The difference in t_{fire} between shortest and longest SOI was about 10 ms. This difference, along with the SOI-specific magnitude of the ensuing ξ_{ee} -gradient, caused a delay in peak latency with *decreasing* SOI and we thus observed negative μ -values, ranging from -13 ms to -32 ms (c.f. white markers in upper panel of Figure 13).

The condition $\tau_{rec,aff} = \tau_{rec}$ resulted in a milder delay in peak latency as a function of *decreasing* SOI (μ -values range from -7 ms to -15 ms, c.f. grey markers in the upper panel of Figure 13) due to a milder effect of STSD on ξ_{aff} -amplitudes (c.f. green curves in Figure 15). In summary, the peak latency dependence on SOI we observed for the single-column model can be explained as follows:

- When $\tau_{rec,aff}$ is longer, the effect of STSD on the afferent current is stronger and amplitudes of ξ_{aff} increase with SOI across the investigated range. In consequence, as SOI is increased, the firing onset latency decreases and the ensuing gradient of the rising slope of ξ_{ee} increases. This leads to a decrease in the peak latency of ξ_{net} .
- When $\tau_{rec,aff}$ is shorter, the effect of STSD on the afferent current is milder. In consequence, amplitudes of ξ_{aff} and therefore peak latencies of ξ_{net} vary less across SOI.

On the following pages, we will present simulation results from the gerbil AC model and compare the response behaviour of individual columns in core field A1 to the response behaviour of the single column model. Recall that for $\tau_{rec,aff} = \tau_{rec}$, i.e. when all connections were

parameter setting	$\sigma_{ie} = 4$	$\sigma_{ie} = 5$	$\sigma_{ie} = 7$
qualitative description of resulting FR functions	broader	intermediate	sharper

Table 8: Lateral reach of effectively inhibitory ie-connections σ_{ie} , and qualitative description of resulting FR functions for columns in A1

characterised by the same STSD recovery time, we observed $\tau_{\text{SOI}} \simeq \tau_{\text{rec}}$ in the single column model. In the gerbil AC model, all connections targeting A1 columns were characterised by a τ_{rec} of 1.000 s (c.f. Table 1). Does this mean that $\tau_{\text{SOI}} \simeq \tau_{\text{rec}}$ for these columns? How does embedding columns in a larger network affect their response behaviour? We aimed to answer these questions with our analysis.

1.2.2 Response behaviour of the gerbil AC model

Spectral selectivity

The FR functions of the individual A1 columns in the gerbil AC model were always single peaked, thus identifying a clear BF for each column. This BF increased with column index. Due to the tonotopic connectivity pattern we defined in the weight matrices, A1 tuning was a direct reflection of the MGv tuning we specified (c.f. Figure 6). Increasing σ_{ie} , the lateral reach of the effectively inhibitory ie-connections, from 4 to 5 to 7, lead to a sharpening of the FR functions.

As an example of the results described above, Figure 19 summarises the FR functions of A1 columns 7, 9 and 11. These have a BF of 2 kHz, 4 kHz, and 8 kHz, respectively. To quantify the ‘sharpness’ of the FR functions, we computed the ratio of the peak response amplitude two octaves below the BF and at the BF. This ratio decreased from about two thirds for $\sigma_{ie} = 4$ to one half for $\sigma_{ie} = 5$ to one third for $\sigma_{ie} = 7$. In contrast, peak response amplitudes at the BF differed by no more than 1%.

Our results are in agreement with the previously reported observation that strength and lateral reach of the effectively inhibitory ie-connections modulate spectral selectivity (May and Tiitinen, 2013). Using $\sigma_{ie} = 4$ results in broader and $\sigma_{ie} = 7$ in sharper audio-frequency tuning, whereas $\sigma_{ie} = 5$ results in an intermediate case. We will interpret the results presented below in this context of broader, intermediate, and sharper FR functions. Therefore, we will use these descriptive terms to refer to the respective σ_{ie} parameter setting. For easy referencing, Table 8 summarises the correspondence between parameter value and descriptive term.

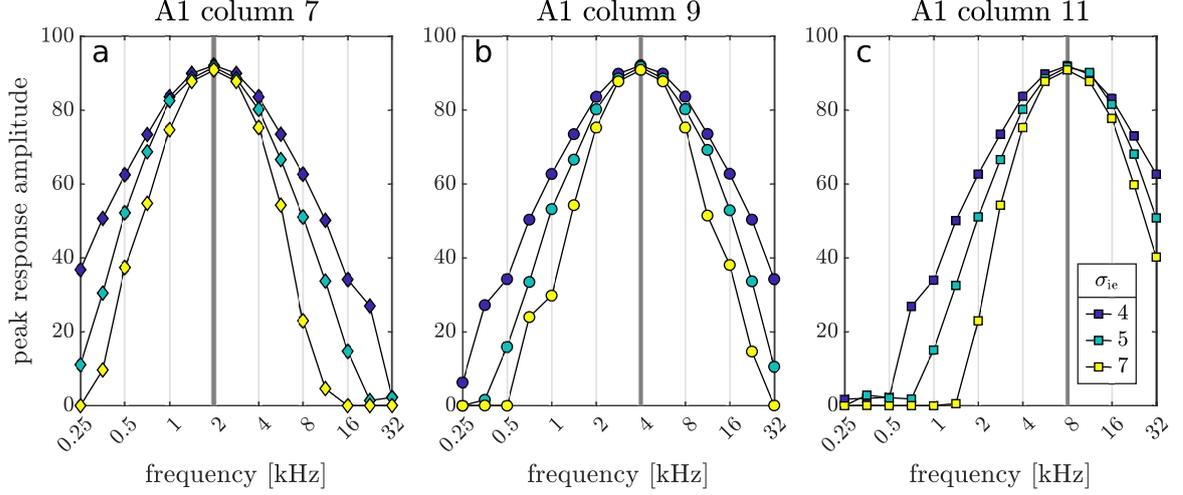


Figure 19: Overview of FR functions for A1 columns 7 (a), 9 (b), and 11 (c) of the gerbil AC model when σ_{ie} , the lateral reach of the effectively inhibitory ie-connections, is equal to 4 (dark blue markers), 5 (light blue markers), and 7 (yellow markers). For each column, the respective BF is highlighted by a thicker vertical grid-line. BF increases with column index and, due to the tonotopic connectivity pattern defined in the weight matrices, A1-tuning directly reflects the tuning defined for MGv. The increase in σ_{ie} , the lateral reach of the effectively inhibitory ie-connections, from 4 to 5 to 7 sharpens the FR function of each column.

Lifetime of repetition suppression and SOI-specific peak latency

In the gerbil AC model, τ_{rec} for all connections targeting the AC was equal to 1.000 s. Recall that, in the single-column model, we observed a mild *decrease* in τ_{SOI} as a function of increasing ΔF when this time constant of recovery from STSD was equal across connections but overall, τ_{SOI} was roughly equal to τ_{rec} . Thus, a similar result might be expected for the gerbil AC model.

Figures 20a to c summarise lifetimes of RS observed for A1 columns when FR functions were broader, intermediate, and sharper. Each bar chart reflects τ_{SOI} -values for individual columns in A1 for two different stimulus audio-frequencies: 2 kHz (purple bars) and 8 kHz (grey bars). These frequencies are the BFs of column 7 and 11, respectively. Each τ_{SOI} -distribution is u-shaped and centred on the BF column, i.e. the distribution shifts as a function of stimulus audio-frequency. The u-shaped distribution translates to a continuous *increase* in τ_{SOI} as a function of ΔF . This increase is steeper the sharper the FR function. For example, for a stimulus audio-frequency of 2kHz (the BF of column 7), τ_{SOI} -values for columns 7 ($\Delta F = 0$ oct) and 10 ($\Delta F = 1.5$ oct) differ by 89 ms when the FR function is broader but by 201 ms when the FR function is sharper.

Figures 20d to f summarise SOI-dependent peak latency shift in terms of fitting parameter μ . Each panel depicts results for one of the three different levels of FR function sharpness. Just as for τ_{SOI} , distributions of μ -values are u-shaped and centred on the BF column. Thus, latency shift also increases with ΔF . And just as for τ_{SOI} , the increase is steeper the sharper the FR functions.

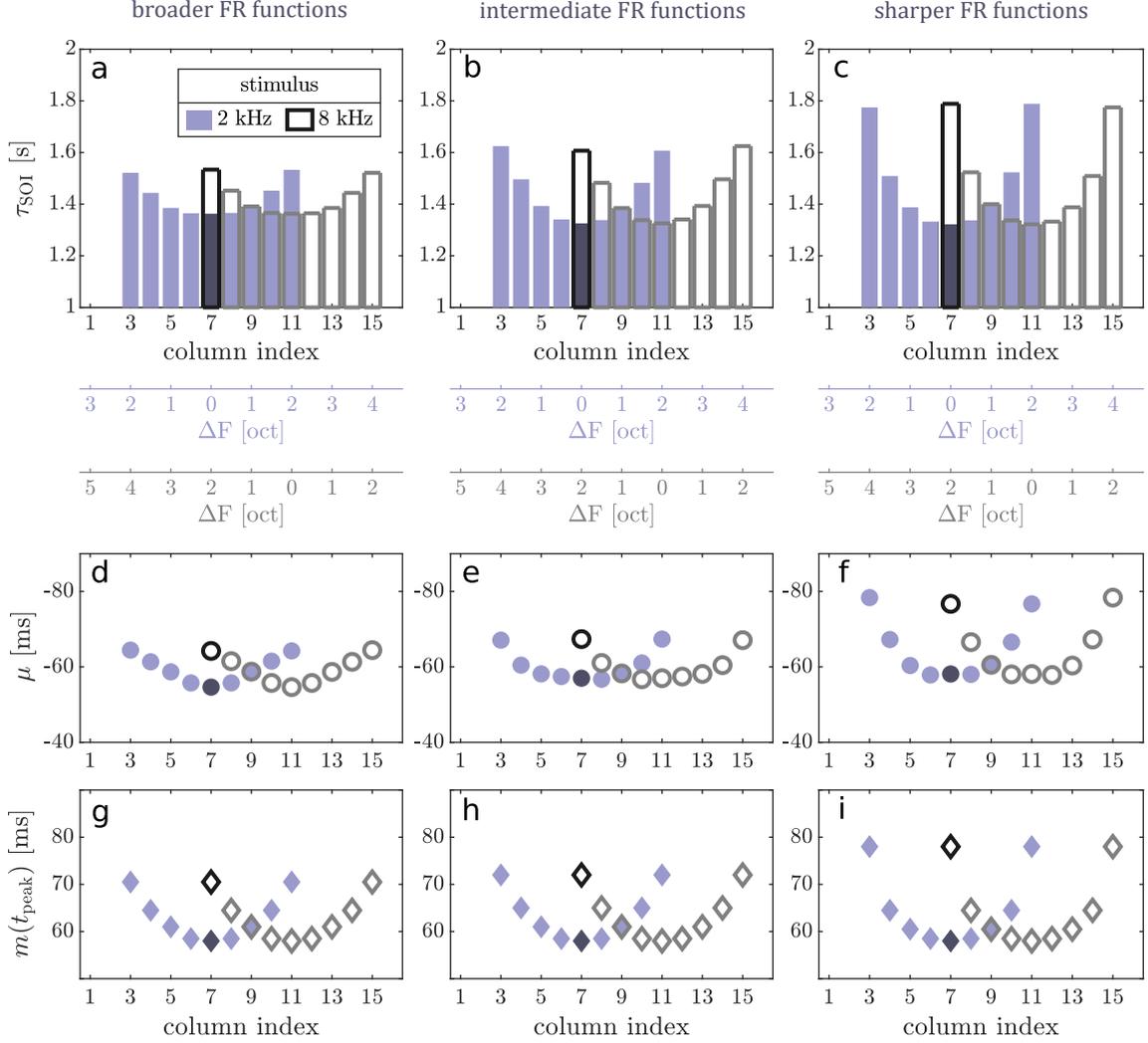


Figure 20: Response behaviour in the primary auditory field (A1) of the gerbil AC model for two different stimulus audio-frequencies and three levels of FR function sharpness (left: broader; middle: intermediate; right: sharper FR functions). Response measures are shown for columns that, given the stimulus audio-frequency, were responsive at all three sharpness levels. Results are plotted as a function of intrafield column index $k(i)$ (c.f. Equation (4)). In the first row (panels **a** to **c**), pale purple bars reflect the columns' τ_{SOI} -values for a stimulus audio-frequency of 2 kHz and grey bars reflect the columns' τ_{SOI} -values for a stimulus audio-frequency of 8 kHz. The secondary and tertiary x-axes in matching colours indicate the ΔF -values for the A1 columns given the stimulus audio-frequency. They also apply to the respective panels below. The larger the distance from a column's BF, the longer the respective τ_{SOI} -value and the sharper the FR function, the steeper the increase. In the second and third row (**d** to **i**), the same colour code reflecting stimulus audio-frequency as in the first row is used. Panels **d** to **f** reflect μ -values (circular markers) and panels **g** to **i** median peak latency (diamond markers) as a function of intrafield column index. The larger the distance from a column's BF, the larger the peak latency shift across SOIs (μ -values are increasingly negative) and the longer the median peak latency. The strength of this effect increases with FR function sharpness. Results for column 7 are emphasised by using more saturated colours – some of the underlying data is reflected in Figure 10 and we will compare the response behaviour of this column with our in-vivo observations.

Figures 20g to i summarises median peak latencies, $m(t_{\text{peak}})$, across SOIs for the three different levels of FR function sharpness. Again, the distribution is u-shaped and centred on the BF column, meaning that $m(t_{\text{peak}})$ also increases with ΔF and again, the increase is steeper the sharper the FR functions. In summary, the following observations can be made in comparison to the single-column model:

1. Unlike in the single-column model, an equal STSD recovery time constant for all connections resulted in a pronounced *increase* in τ_{SOI} as a function of ΔF for the A1 columns of the gerbil AC model.
2. For a given ΔF , response peaks in the gerbil AC model occurred at later times than in the single-column model. For $\Delta F = 0$ oct, for example, $m(t_{\text{peak}})$ in the single-column model was equal to about 25 ms whereas for A1 columns in the gerbil AC model, it was equal to 58 ms.
3. Latency shift μ was more pronounced in the gerbil AC model than in the single-column model. The smallest shift for an A1 column in the gerbil AC model exceeded the largest shift in the single-column model by 23 ms.

Response measures μ , $m(t_{\text{peak}})$, and τ_{SOI} were based on the net transmembrane current associated with the excitatory population of A1 columns in the model. In the following sections, we will explore these columns' response behaviour in terms of the different subtypes of transmembrane current that make up the net flow (c.f. Equations (10) to (13) as well as Equations (18) and (19)). This was done in order to study which neuronal interactions across the network are contributing to the variations in RS lifetime and peak latency.

Components of the gerbil AC model's responses

Figure 21 breaks down a response of A1 column 7 into the excitatory and inhibitory components that make up the net transmembrane current flow (panel a) and further divides the excitatory synaptic current into subcomponents according to connection origin (panel b). Taken together, Figures 21a and b illustrate the response sequence observed in A1 columns for BF-stimulation across all investigated SOI-values:

1. Stimulus presentation leads to excitatory synaptic current via the thalamocortical connection ($\xi_{\text{ee}}^{\text{MGv}}$, medium green dotted curve) and moves the excitatory population away from its resting state. This first activates the passive leak current, ξ_{leak} (orange curve).
2. After a certain delay relative to the onset of current flow $\xi_{\text{ee}}^{\text{MGv}}$, the excitatory population's firing threshold is crossed and it begins to respond actively. This in turn leads to an excitatory synaptic current, $\xi_{\text{ee}}^{\text{recurrent}}$ (light green dotted curve), via the population's recurrent ee-connection.

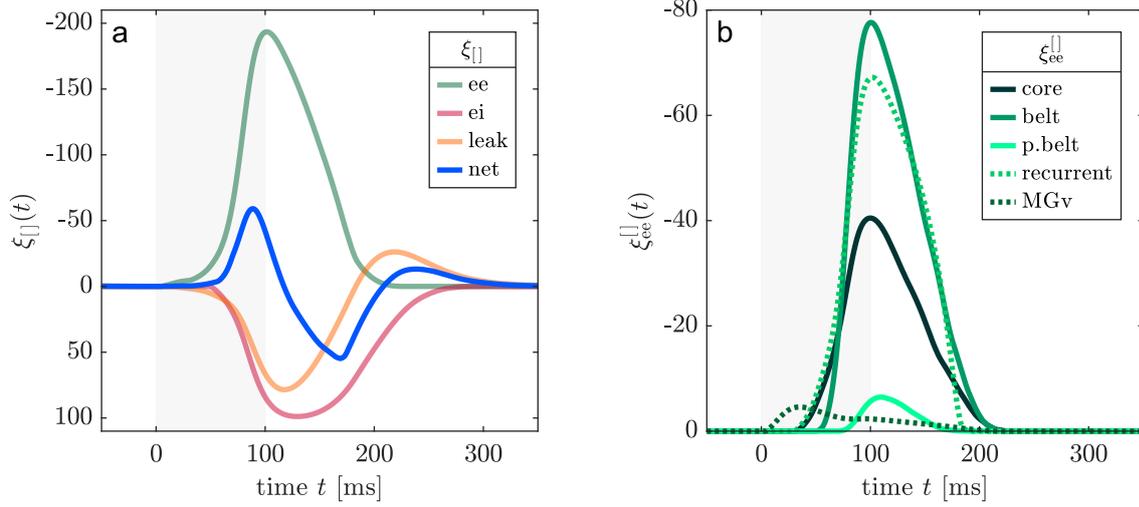


Figure 21: Components of transmembrane current flow in A1 column 7 for $\text{SOI} = 0.438$ s, $\Delta F = 0$ oct, and intermediate FR functions. **a:** The orange curve reflects the leak current ξ_{leak} (Equation (10)), the green curve reflects the excitatory synaptic current ξ_{ee} delivered via ee-connections (Equation (11)), and the red curve reflects the inhibitory synaptic current ξ_{ei} delivered via ei-connections (Equation (12)). The afferent synaptic current ξ_{aff} induced via an afferent connection during stimulus presentation (Equation (13)) is always zero beyond MGv (Equation (7)) and therefore not explicitly depicted. A response peak in ξ_{net} is observed at time $t_{\text{peak}} = 89$ ms because excitatory synaptic current ξ_{ee} maximally exceeds the inhibitory current ($\xi_{\text{leak}} + \xi_{ei}$) at this moment. **b:** Decomposition of excitatory synaptic current ξ_{ee} into regions of connection origin. Current components were computed using Equations 15 to 18 with area-specific column index vectors \mathbf{j}_{area} . The medium green dotted curve reflects ξ_{ee}^{MGv} , the excitatory synaptic current induced via intercolumnar connections originating in thalamus (MGv). Excitatory synaptic current induced via intercolumnar connections originating in the core area, ξ_{ee}^{core} , is shown in dark green. The medium green curve illustrates ξ_{ee}^{belt} , the excitatory synaptic current resulting from intercolumnar connections originating in the belt area. Finally, $\xi_{ee}^{\text{p.belt}}$, excitatory synaptic current caused by intercolumnar connections originating in the parabelt area, is mapped out in light green and $\xi_{ee}^{\text{recurrent}}$, the intracolumnar excitatory synaptic current, in dotted light green. Currents induced via the recurrent connection and connections originating in the belt are the dominant components of the excitatory synaptic current.

3. In parallel, other core columns become active and contribute input ξ_{ee}^{core} (dark green curve) via intercolumnar ee-connections.
4. Activity of the excitatory population and, to a lesser extent, neighbouring excitatory populations, also causes the activation of the inhibitory population via the ie-connections. In turn, this causes inhibitory synaptic current, ξ_{ei} , targeting the excitatory population via the ei-connection (red curve).
5. With a certain delay relative to onset of activity in the core, first the belt and then also the parabelt become active, leading to further excitatory synaptic current (medium green and light green curve) via intercolumnar ee-connections. Alongside recurrent excitation, $\xi_{ee}^{\text{recurrent}}$, synaptic current ξ_{ee}^{belt} reaching the target column due to activity in the belt fields makes up a particularly large component of the net excitatory synaptic current at the column. In contrast, while thalamocortical input drives the response, its direct contribution to the net response peak is negligible.
6. After stimulus offset, current flow steadily decreases and the column returns to its resting state.

With increasing ΔF , the sequence described above was slightly altered for longer SOIs. The onset of synaptic current via core connections started to precede the onset of input via the recurrent connection, because activation spreads laterally from the BF column (the first to activate) to neighbouring columns. The observations that, (1) while thalamocortical input drives the response, its contribution to the net response peak is very minor and that (2) $\xi_{ee,i}^{\text{recurrent}}$ and $\xi_{ee,i}^{\text{belt}}$ are the dominant components of $\xi_{ee,i}$, remained valid across SOI- and ΔF -values.

Effect of network activation pattern on the lifetime of repetition suppression

A response peak in ξ_{net} is observed at time t_{peak} because excitatory synaptic current ξ_{ee} maximally exceeds the inhibitory current, $\xi_{\text{leak}} + \xi_{ei}$, at this moment, just like in the single column model. Unlike in the single column, however, the amplitude of the response peak of a given excitatory population i in the gerbil AC model does not only depend on presynaptic firing rate and STSD-dependent synaptic efficacy at one afferent and one recurrent ee-connection. Instead, the SOI-dependent response behaviour of multiple individual network nodes has a decisive impact on the lifetime of RS in a given A1 column, as will be demonstrated in this section.

While the number of excitatory connections targeting a given excitatory population is constant (as defined in weight matrix W_{ee}), the effect a connection has on the target population depends on the activity level at that connection. This in turn depends on the firing rate $g(t)$

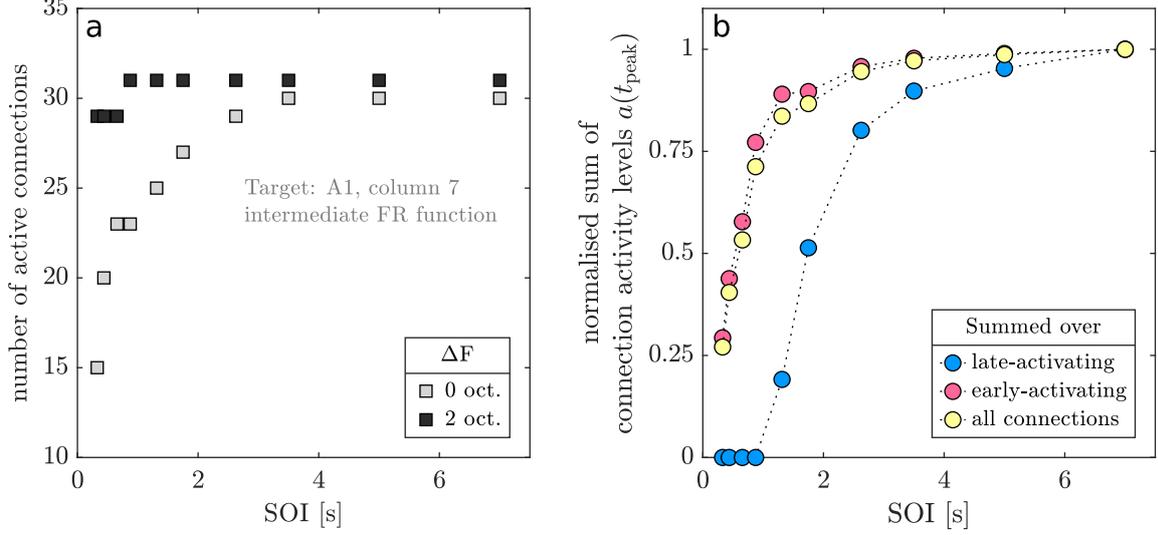


Figure 22: Activation of connections that target A1 column 7 in the gerbil AC model. **a:** Number of active connections when $\Delta F = 0$ oct (black markers) and when $\Delta F = 2$ oct. When stimulus audio-frequency deviates from the BF of the column, i.e. when $\Delta F = 2$ oct, the number of active connections increases with SOI. **b:** Summed connection activity levels as a function of SOI for $\Delta F = 2$ oct. Pink markers reflect summed activity for connections active across all SOIs (‘early-activating’), blue markers for connections only active if $\text{SOI} > 0.875$ s (‘late-activating’) and yellow markers for all connections. The late activation of additional connections decreases the rising slope of overall connection activity.

of the presynaptic population and the synaptic efficacy $q(t)$ for the given connection. We defined the activity level at a given connection as

$$a_{i,j}(t) = W_{ee}(i,j) q_j(t) g(u_j(t)) , \quad (23)$$

i.e. the product of connection weight, synaptic efficacy, and presynaptic firing rate. Note that this reflects the amplitude of the excitatory synaptic current transmitted via this connection, since (c.f. Equation (11))

$$-\frac{1}{\tau_m} \sum_{j=1}^N a_{i,j}(t) = \xi_{ee,i}(t) . \quad (24)$$

Given that minimal levels of activity persisted in the entire network, we defined an active connection as one where $a_{i,j}(t_{\text{peak}}) \geq 0.008$. The absolute number of active connections was, of course, sensitive to the selected threshold value. However, a coarse investigation, using threshold values as low as 0.001 and as high as 0.050, showed that the general response behaviour described below remained stable.

Figure 22a shows, as a function of SOI, the number of *active* ee-connections targeting column 7 in A1 when FR functions have intermediate sharpness and stimulus audio-frequency is equal to 2 kHz ($\Delta F = 0$ oct, black markers) or 8 kHz ($\Delta F = 2$ oct, grey markers). The number of active connections as a function of SOI is fairly stable for $\Delta F = 0$ oct. It is equal

ΔF	broader FR function			intermediate FR function			sharper FR function		
	active connections			active connections			active connections		
	min	max	for SOI \geq	min	max	for SOI \geq	min	max	for SOI \geq
0	29	31	1.750 s	29	31	0.875 s	29	31	0.656 s
2	23	35	5.000 s	15	30	3.5 s	17	23	2.626 s

Table 9: Overview of the minimum (min) and maximum (max) number of active ee-connections targeting A1 column 7 and SOI-values at which the stable maximum is reached.

to 29 for SOIs below 0.875 s and stabilises at 31 for SOIs ≥ 0.875 s. In contrast, the number of active connections as a function of SOI for $\Delta F = 2$ oct covers a much wider range. It starts out at 15 for SOI = 0.328 s and stabilises at 30 for SOI ≥ 3.500 s. We observed similar trends for both sharper and broader FR functions. The results are summarised in Table 9. In each case, the number of active connections as a function of SOI spans a wider range for $\Delta F = 2$ oct and saturates earlier for $\Delta F = 0$ oct.

Figure 22b illustrates how the SOI-dependent increase in the number of active connections observed for $\Delta F = 2$ oct affected activity levels at the connections targeting A1 column 7. Summed connection activity levels are plotted as a function of SOI for connections that activate at the shortest SOI, i.e. ‘early’ (pink markers), for connections that are only active for SOI > 0.875 s, i.e. ‘late’ (blue markers), and for all connections (yellow markers). The late activation of additional connections results in the rising slope of overall connection activity being less steep than for the activity of the early activating connections. In contrast, for $\Delta F = 0$ oct, the vast majority of connections activates early and this modulation is negligible.

The effect described above generalised to all three levels of FR function sharpness. As SOI was increased, levels of connection activity throughout the network increased because there was more time for recovery from STSD between subsequent stimuli. This, in turn, led to a more far-reaching spread of activation, i.e. an increasing total number of active connections. Thus, while all active connections recovered from STSD with the same τ_{rec} , some connections were only activated at longer SOIs. These connections were relevant for columns with a BF further away from the stimulus audio-frequency. The additional synaptic current (recall that a reflects amplitudes of synaptic current, c.f. Equation (24)) transmitted via the late-activating connections prolonged recovery from RS.

Beside the sheer number of active connections, it was of course also the amplitude of the excitatory synaptic current induced via these connections that shaped the RS recovery dynamics of a given column i . A connection that is only active for longer SOIs but that induces an additional current that is negligible in comparison to the total current ξ_{ee} will not prolong RS lifetime significantly. To verify how the observed differences in the number of active connections affected the SOI-dependence of the evoked excitatory synaptic current, we determined lifetimes of RS from the SOI-dependence of $\xi_{\text{ee}}(t_{\text{peak}})$. Note that here, t_{peak} still refers to the latency of the respective response peak observed for ξ_{net} . Equation (1) was

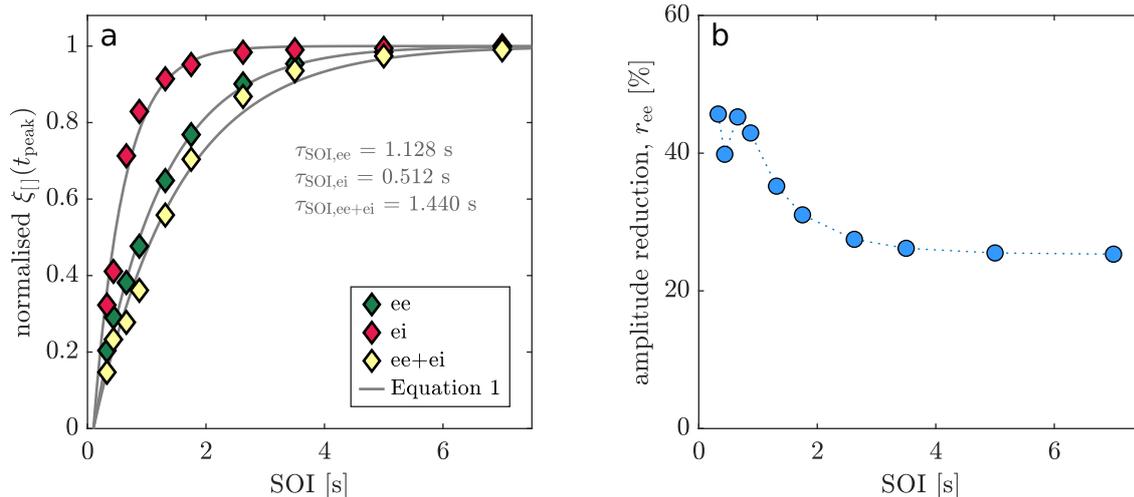


Figure 23: SOI-specific interaction between excitatory and inhibitory synaptic current for A1 column 7, given $\Delta F = 2$ oct and intermediate FR functions. a: SOI-dependence of the amplitude of excitatory (green markers) and inhibitory (red markers) synaptic current as well as the sum of these two currents (yellow markers) at time t_{peak} , along with respective fits of Equation (1). Amplitude values are normalised with respect to the saturation amplitudes of the fitted curves. b: Percentage reduction from amplitude $\xi_{\text{ee}}(t_{\text{peak}})$ to the amplitude of the summed currents $\xi_{\text{ee}}(t_{\text{peak}}) + \xi_{\text{ei}}(t_{\text{peak}})$ as a function of SOI. The reduction is equal to or exceeds 40% for SOIs ≤ 0.875 s and stabilises at about 25% for longer SOIs. This SOI-specific reduction has a flattening effect on the SOI-dependence of $\xi_{\text{ee}}(t_{\text{peak}}) + \xi_{\text{ei}}(t_{\text{peak}})$ compared to the SOI-dependence of $\xi_{\text{ee}}(t_{\text{peak}})$.

fitted using the same procedure as for ξ_{net} (c.f. Section 1.1.5, [Quantification of repetition suppression lifetime](#)) and the resulting lifetimes are referred to as $\tau_{\text{SOI,ee}}$. An example of such a fit is shown in Figure 23a (grey curve), where normalised values of $\xi_{\text{ee}}(t_{\text{peak}})$ are plotted as a function of SOI (green markers). Overall, we found that, in line with the relatively stable number of active connections, $\tau_{\text{SOI,ee}}$ -values for $\Delta F = 0$ oct differed by no more than 5% across levels of FR function sharpness. Comparing results for $\Delta F = 0$ oct and $\Delta F = 2$ oct, $\tau_{\text{SOI,ee}}$ -values were longer for the latter condition and differed more for the intermediate and sharper FR functions (23% and 24% difference, respectively) than for the broader FR function (15% difference). We thus conclude that, for $\Delta F = 2$ oct, the additional connections activated at longer SOIs made a more significant contribution to ξ_{ee} for the higher levels of FR function sharpness.

Effect of inhibition on repetition suppression lifetime

The inhibitory synaptic current ξ_{ei} was driven by excitatory populations via the lateral ie-connections. In consequence, the differences in activation level of the excitatory populations also affected the SOI-dependence of $\xi_{\text{ei}}(t_{\text{peak}})$. We again determined corresponding RS lifetimes, referred to as $\tau_{\text{SOI,ei}}$. One example of the approach is shown in Figure 23a, where normalised values of $\xi_{\text{ei}}(t_{\text{peak}})$ (red markers) are plotted as a function of SOI and the grey curve reflects Equation (1) fitted to the data. RS lifetimes for $\xi_{\text{ei}}(t_{\text{peak}})$ were generally short, with $\tau_{\text{SOI,ei}} < 0.530$ s, but increased as a function of ΔF . The sharper the FR functions,

the steeper the increase: comparing results for $\Delta F = 0$ oct and $\Delta F = 2$ oct given broader, intermediate, or sharper FR functions, we found that $\tau_{\text{SOI,ei}}$ -values differed by 12%, 31%, or 36%, respectively.

Inhibitory synaptic current opposes the excitatory current and thus reduces the amplitude of the response peak observed for the net current ξ_{net} (recall the time courses of the current components depicted in Figure 21a). As elaborated above, SOI-dependences of $\xi_{\text{ee}}(t_{\text{peak}})$ and $\xi_{\text{ei}}(t_{\text{peak}})$ were not characterised by the same lifetime. In consequence, relative to $\xi_{\text{ee}}(t_{\text{peak}})$, the reduction in amplitude for the sum of currents $\xi_{\text{ee}}(t_{\text{peak}})$ and $\xi_{\text{ei}}(t_{\text{peak}})$ was not characterised by a common scaling factor across SOIs. This is illustrated in Figure 23b, where

$$r_{\text{ee}} = \left[1 - \frac{\xi_{\text{ee}}(t_{\text{peak}}) + \xi_{\text{ei}}(t_{\text{peak}})}{\xi_{\text{ee}}(t_{\text{peak}})} \right] \times 100, \quad (25)$$

the percentage reduction from $\xi_{\text{ee}}(t_{\text{peak}})$ to the summed currents $\xi_{\text{ee}}(t_{\text{peak}}) + \xi_{\text{ei}}(t_{\text{peak}})$, is plotted as a function of SOI for A1 column 7, intermediate FR functions, and $\Delta F = 2$ oct. For SOIs ≤ 0.875 s, r_{ee} is equal to or exceeds 40% whereas for longer SOIs, it first decreases and then stabilises at about 25%. Such an effect was observed for all conditions, but to varying degrees. For BF stimulation, r_{ee} -values for shorter and longer SOIs differed by about 20 percentage points, irrespective of FR function sharpness. In contrast, for $\Delta F = 2$ oct, r_{ee} -values for shorter versus longer SOIs differed by 23, 24, and 35 percentage points for broader, intermediate, and sharper FR functions, respectively.

The difference in r_{ee} -values across SOIs resulted in a flattening effect on the the rising slope of the SOI-dependence of the sum of currents $\xi_{\text{ee}}(t_{\text{peak}})$ and $\xi_{\text{ei}}(t_{\text{peak}})$ compared to the SOI-dependence of current $\xi_{\text{ee}}(t_{\text{peak}})$ alone. This, in turn, caused RS lifetimes, $\tau_{\text{SOI,ee+ei}}$, that were longer than the lifetimes for the individual current components. The effect is visualised in Figure 23a, where the normalised sum of the opposing synaptic currents ξ_{ee} and ξ_{ei} at time t_{peak} is plotted as a function of SOI (yellow markers) for A1 column 7, given intermediate FR functions and $\Delta F = 2$ oct. The grey curve represents Equation (1) fitted to the data and the associated RS lifetime is referred to as $\tau_{\text{SOI,ee+ei}}$. Overall, we found that for BF stimulation, $\tau_{\text{SOI,ee+ei}}$ -values were about 19% longer than $\tau_{\text{SOI,ee}}$ values, irrespective of FR function sharpness. For $\Delta F = 2$ oct and broader, intermediate, and sharper FR functions, respectively, values of $\tau_{\text{SOI,ee+ei}}$ were about 20%, 28%, and 41% longer than the corresponding $\tau_{\text{SOI,ee}}$. Thus, lateral inhibition further steepened the increase in RS lifetime as a function of ΔF , in particular for sharper FR functions.

In combination, the network effects described above explain the τ_{SOI} distributions illustrated in Figures 20a to c. The distributions were characterised by an increase in RS lifetime τ_{SOI} as a function of ΔF and a change in the steepness of this increase across levels of FR function sharpness.

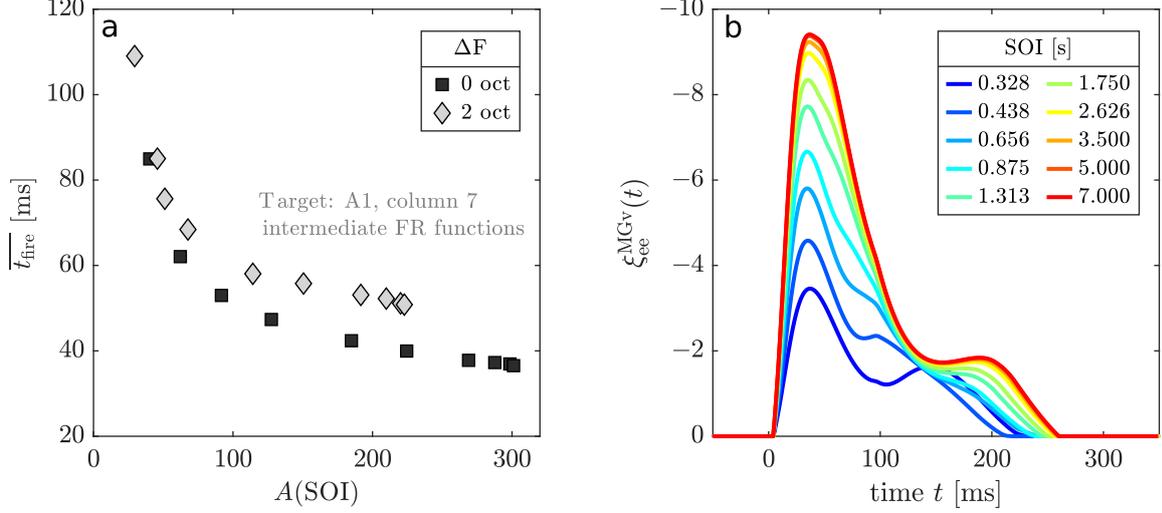


Figure 24: a: Mean firing onset latency $\overline{t_{\text{fire}}}$ (Equation (26)) of excitatory neural populations connected to A1 column 7, plotted as a function of SOI-specific response peak amplitude for stimulation at $\Delta F = 0$ oct (black markers) and at $\Delta F = 2$ oct (grey markers). Firing onset increases with distance from the BF and the longer the SOI, the earlier the firing onset. b: Time courses of thalamocortical synaptic current $\xi_{\text{ee}}^{\text{MGv}}$ (Equation (18)) reaching A1 column 7 for different SOI-values ($\Delta F = 0$ oct). The shorter the SOI, the smaller the current peak.

We found that

- The SOI-dependence of response measure ξ_{net} , the net transmembrane current, is affected by SOI-specific excitatory and inhibitory synaptic currents.
- The larger ΔF , the longer the SOI at which the increasing number of active excitatory connections ultimately stabilises. This prolongs RS lifetime.
- The sharper the FR function, the stronger the effect of the ‘late’ activation of additional connections on synaptic current flow. As a result, the longest RS lifetimes were observed for $\Delta F = 2$ oct and sharper FR functions.

Mechanisms behind variations in response peak latency

The previously described Figure 21b illustrates the response cascade that affected peak latencies in the gerbil AC model. Stimulus presentation first activated the MGv, which then relayed this activation to the core and, to a lesser extent, to the belt. Thus, columns in the core reached the firing threshold first, followed by belt, and then by parabelt columns. A1 columns were targeted by connections originating in each of these areas and, in consequence, the net response ξ_{net} peaked at later times than in the single-column model, where there was no comparable activation cascade.

We summarised the firing onset latencies of excitatory populations at the source of connections targeting a given A1 column by a weighted arithmetic mean $\overline{t_{\text{fire}}}$. To account for the differing contributions of individual connections to the net excitatory current, ξ_{ee} , weights

A1 column 7 FR functions	broader		intermediate		sharper	
ΔF	0 oct	2 oct	0 oct	2 oct	0 oct	2 oct
range of $\overline{t_{\text{fire}}}$ across SOIs	64 ms	90 ms	40 ms	58 ms	51 ms	53 ms
range of $A(\text{SOI})$ across SOIs	285	245	261	193	259	146
ratio of ranges	0.225	0.367	0.153	0.301	0.197	0.363
difference between ratios	0.142		0.148		0.166	

Table 10: For A1 column 7, this table list ranges of mean firing onset latency $\overline{t_{\text{fire}}}$ (Equation (26)) and response peak amplitudes ($A = |\xi_{\text{net}}(t_{\text{peak}})|$) across SOIs for the three levels of FR function sharpness and stimulation at $\Delta F = 0$ oct and $\Delta F = 2$ oct. The ratio between the two ranges increases with ΔF and the steepest increase was observed for the sharper FR functions.

were equal to the activity level $a_{i,j}(t_{\text{peak}})$ (c.f. Equation (23)) at the respective connection, i.e., for a given SOI and target population i :

$$\overline{t_{\text{fire}}} = \frac{\sum_{j=1}^N a_{i,j}(t_{\text{peak}}(i)) t_{\text{fire}}(j)}{\sum_{j=1}^N a_{i,j}(t_{\text{peak}}(i))}. \quad (26)$$

Figure 24a shows values of $\overline{t_{\text{fire}}}$ for A1 column 7 as a function of SOI-specific peak response amplitude $A(\text{SOI})$. The mean firing onset latency $\overline{t_{\text{fire}}}$ decreases as a function of $A(\text{SOI})$ and increases with ΔF , an observation that generalised to all simulation conditions. This in turn caused changes in the latency of the net response peak (recall that longer firing onset latencies result in longer peak latencies).

The increase in $\overline{t_{\text{fire}}}$ with decreasing SOI can be explained by the decreasing levels of thalamocortical synaptic current $\xi_{\text{ee}}^{\text{MGv}}$ reaching the cortical fields. This decrease occurs due to STSD at the thalamocortical connections. As an example, Figure 24b illustrates the decrease in $\xi_{\text{ee}}^{\text{MGv}}$ reaching A1 column 7. Recall that thalamocortical synaptic current drives the response of the A1 columns. The larger current $\xi_{\text{ee}}^{\text{MGv}}$, the sooner the targeted cortical columns reach the firing threshold and the sooner the response cascade illustrated in Figure 21b is initiated. This is comparable to the effect of the SOI-specific amplitudes ξ_{aff} on firing onset latency in the single-column model. But while the range of t_{fire} across SOIs covered less than 20 ms in the single-column model, network interactions in the gerbil AC model amplified this range, with values of $\overline{t_{\text{fire}}}$ spanning a range of 40 ms or more across SOIs.

The peak latency shift μ can be related to the ratio between the range of $\overline{t_{\text{fire}}}$ and the range of A(SOI) across SOIs. For A1 column 7, Table 10 lists the respective ranges and resulting ratios observed for the three levels of FR function sharpness and the two different stimulus-audiofrequencies. We found that the ratio between ranges of $\overline{t_{\text{fire}}}$ and A(SOI) increased with ΔF . Moreover, the steepness of the increase scaled with FR function sharpness. In consequence, the difference in the μ -value from $\Delta F = 0$ oct to $\Delta F = 2$ oct was most pronounced for sharper FR functions. This observation generalised to other A1 columns.

In summary, while the basic mechanisms giving rise to variations in peak latency generalised from single-column to gerbil AC model, network interactions delayed peak latencies and modulated the magnitude of the peak latency shift.

Overview of gerbil AC model predictions

Our simulations resulted in the following predictions regarding RS response behaviour in gerbil A1:

1. The lifetime of RS does not simply reflect the lifetime of STSD. Instead, RS is a network effect resulting from the interplay of STSD dynamics and connectivity patterns.
2. RS lifetimes at a given measurement location are not constant across stimulus audio-frequencies. The further the distance from the BF of the measurement location, the longer the lifetime of RS and the sharper the FR function, the steeper the increase.
3. Peak response latencies are not stable across stimulus audio-frequencies. The further the distance from the BF of the measurement location, the longer the peak latency and the sharper the FR function, the steeper the increase.
4. Peak latency is not stable across SOIs. An increase in SOI moves the response peak to an earlier time point. The sharper the FR function, the steeper the decrease in peak latency as a function of SOI.

Next, we verified whether these predictions could be confirmed *in vivo*.

Project 2: Repetition suppression in vivo

To test the predictions of the gerbil AC model in vivo, a regular-SOI paradigm comprising stimuli with two different audio frequencies was presented to nine anaesthetised gerbils while carrying out intracranial recordings in A1. The resulting data was then analysed to quantify RS response behaviour, in particular in terms of lifetimes. The sections below first introduce how the electrophysiological measurements were carried out and describe the details of the experimental paradigm. Next, the analysis pipeline developed to obtain robust results and subject-specific statistical inferences on response measures is explained and finally, in-vivo results are compared with model predictions.

2.1 Materials and pre-established methods

The experiments were carried out with healthy adult Mongolian gerbils (*Meriones unguiculatus*, $n = 9$, all male, age 3–6 months, body weight 70–90 g). All experiments satisfied the ethical animal research standards defined by the German Law and were officially approved by an ethics committee of the German state of Saxony-Anhalt (license 42502-2-1394LIN).

2.1.1 Surgical procedure and electrophysiological measurements

Surgical procedures and electrophysiological measurements were conducted akin to the approach described by [Brunk et al. \(2019\)](#), [Deliano et al. \(2020\)](#), and [Deane et al. \(2020\)](#).

A linear multichannel electrode (Neuronexus A1x32-50-413, $n_{\text{chan}} = 32$ equidistant channels, channel distance $\Delta z = 50 \mu\text{m}$) was implanted into the right A1 of the anaesthetised animal (anaesthesia induced by intraperitoneal injection of 45% v/v ketamine (50 mg/ml, Ratiopharm GmbH), 5% v/v xylazine (Rompun 2%, Bayer Vital GmbH), and 50% v/v of isotonic sodium-chloride solution (154 mmol/l, B. Braun AG) with a dose of 0.004 ml/g body-weight and maintained by the same ketamine/xylazine infusion at a rate of 15 mg/kg/h). The level of anaesthesia was controlled by monitoring the hindlimb withdrawal reflex and the breathing rate (every 10 to 25 min). The multichannel electrode was positioned perpendicular to the cortical surface ([Happel et al., 2010](#)) and the implantation site was determined based on vascular landmarks ([Thomas et al., 1993](#); [Ohl and Scheich, 1997b](#)).

Measurements of depth profiles of local field potentials (LFPs) were performed in an acoustically and electrically shielded recording chamber. Stimuli were delivered by a speaker positioned 100 cm behind the animal (Tannoy arena satellite KI-8710-32). These sound stimuli were generated and controlled via Matlab (Mathworks, R2006b) and then converted into an analogue signal (sampling frequency 1 kHz, NI PCI-BNC2110, National Instruments). This signal was routed through an attenuator (gPAH, Guger Technologies), and then amplified

(Thomas Tech Amp75) before reaching the microphone. To calibrate the stimuli, a microphone and conditioning amplifier (G.R.A.S. 26AM and B&K Nexus 2690-A, Bruel & Kjaer, Germany) were used. The recorded LFPs were amplified via an Omnetics connector (HST/32V-G2O LN 5V, 20x gain, Plexon Inc.) into a PBX25 preamplifier (Plexon Inc.), pre-amplified 500-fold and band-pass filtered (0.7-300 Hz). Next, the signals were digitized at 2 kHz with a multichannel-recording system (Multichannel Acquisition Processor, Plexon Inc.).

2.1.2 Stimulation paradigms

For each animal, the experiment was divided into two phases: first an investigation of the audio-frequency response at the measurement site, and then an investigation of audio-frequency-specific RS recovery dynamics at that same site. All presented stimuli were pure tones at an SPL of 65 dB.

Frequency response paradigm

During the frequency response (FR) paradigm, a series of 350 pseudo-randomised pure tones was presented. Audio-frequencies of these stimuli ranged from 0.25 to 16 kHz in iso-octavial steps and the duration of each stimulus was 200 ms, including a linear rise and fall time of 5 ms each. Inter-stimulus intervals were randomised but limited to the 600 ms to 800 ms range and stimuli were repeated a total of 50 times for each audio-frequency. After the implantation of the electrode, the described stimulus sequence was presented repeatedly to monitor the stabilisation of the responses to the different tones (c.f. [Deane et al., 2020](#)). Across all animals, stabilisation occurred within the first hour after implantation. Once the response was stable, a final data set was recorded.

Next, the best frequency (BF) of the measurement site was identified. It was defined as the stimulus audio-frequency evoking the largest root mean square (RMS) response amplitude in the granular layer III/IV (time window from 15 ms to 64 ms after stimulus onset, response in terms of current source densities computed during the recording, for details see [Section 2.1.3](#) and [Equation \(27\)](#) therein). Response amplitudes for stimulation two octaves above and below the BF were then compared and the audio-frequency eliciting the smaller response was selected as the ‘non-BF’. The BF and non-BF determined online were used to investigate audio-frequency-specific RS recovery dynamics in the second phase of the experiment. During subsequent offline data analysis, we verified the suitability of the BF and non-BF selection.

Repetition suppression paradigm

To investigate audio-frequency-specific RS, sequences of identical pure-tone stimuli, hereafter referred to as stimulus blocks, were presented at the BF and non-BF. Each tone had a duration of 100 ms, including a linear rise- and fall-time of 5 ms. Within a stimulus block, the SOI was

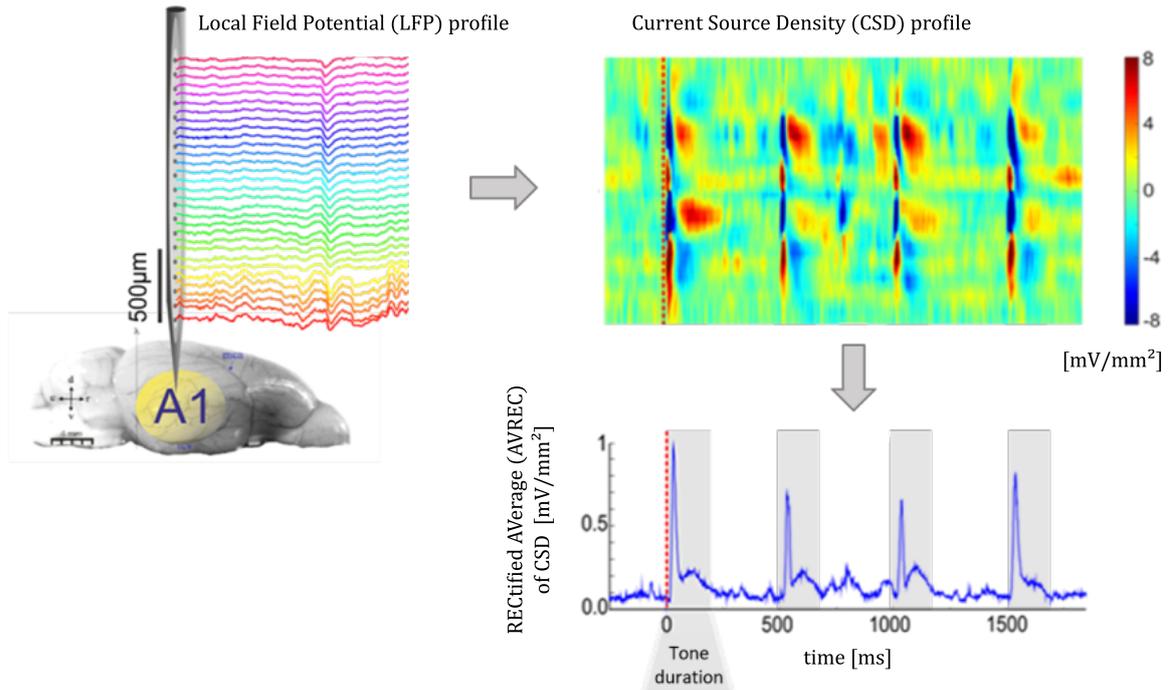


Figure 25: Illustration of analysis pipeline from measured local field potential (LFP) profile to current source density (CSD, see Equation (27)) profile to rectified average of the current source density profile (AVREC, see Equation (28))

constant. In total, we presented stimulus blocks with 10 different SOIs (values in seconds: 0.219, 0.328, 0.438, 0.656, 0.875, 1.313, 1.750, 2.626, 3.500, and 7.000) at both the BF and non-BF. Their presentation order was randomised and, for a given stimulus audio-frequency, each block was presented a total of 24 times. A silent interval of 14 s separated subsequent blocks. In order to harmonise the number of stimulus repetitions per block across SOI values, individual blocks had a duration of 7.2 s for SOIs < 2.626 s and of 14.2 s for SOIs ≥ 2.626 s. The last second of the silent interval preceding a stimulus block was treated as a baseline recording.

2.1.3 Measures of neuronal activity

The activity of neuronal populations leads to localised current flow of mostly positive ions through the cell membranes. On a more macroscopic scale, and viewed from the extracellular space, a lack of positive ions in the extracellular space leads to a negative charge (sinks) whereas a release of positive ions into the extracellular space leads to a positive extracellular charge (sources). The local field potentials (LFPs) measured by an electrode implanted perpendicular to the cortical layers are caused by these sinks and sources. They can be predominantly attributed to the activity of pyramidal neurons and their location and strength can be deduced from the LFP depth profiles by computing the corresponding current source density (CSD) profiles. The second spatial derivative of the LFP is an approximation of the spatiotemporal CSD profile at the measurement location. We used this well-established rela-

tionship (Mitzdorf, 1985) to compute CSD profiles from all our LFP data. At any time point t of the measurement, the CSD profile can be approximated as follows:

$$-\text{CSD}(z) \approx \frac{\partial^2 \Phi(z)}{\partial z^2} \approx \frac{\Phi(z + n\Delta z) - 2\Phi(z) + \Phi(z - n\Delta z)}{(n\Delta z)^2} \quad (27)$$

The spatial coordinate z reflects cortical depth (perpendicular to the cortical layers), $\Phi(z)$ denotes the field potential at depth z , Δz is the spatial sampling interval of the measurement (i.e. the distance between neighbouring electrode channels) and $n\Delta z$ specifies the differentiation grid. In our experiments, Δz was equal to 50 μm . We set $n = 1$ for CSD computation, resulting in a total of 32 CSD traces across the cortical depth covered by the electrode. Moreover, any LFP profile Φ used for CSD computation was spatially smoothed using a weighted Hamming window with a kernel size of 300 μm .

Given that our computational gerbil AC model did not include any resolution across cortical depth, we summarised neuronal activity across all six cortical layers in a single measure. The $N = 32$ traces of the CSD profiles were condensed into the average rectified current source density (AVREC):

$$\text{AVREC}(t) = \frac{\sum_{i=1}^N |\text{CSD}_i(t)|}{N} \quad (28)$$

where t is time, i denotes the trace index and $|\text{CSD}_i(t)|$ is the absolute value of the CSD for trace i at time t . While this rectification makes sinks and sources indistinguishable and thus causes a loss of information regarding transmembrane current flow direction, the approach allows for the computation of an average CSD trace across the full depth of the AC without sinks (current flow due to excitatory synaptic events) and sources (predominantly return currents) cancelling each other out (note that a small residual current remains, see for example Harding, 1992; Happel et al., 2010; Brunk et al., 2019). The AVREC is thus a good measure of the temporal evolution of the overall activation of a cortical column (Givre et al., 1994; Schroeder et al., 1998). All results presented here are based on AC response behaviour at the AVREC level. Figure 25 illustrates our LFP→CSD→AVREC analysis pipeline.

2.2 Preliminary assessment

2.2.1 Frequency response

For each animal, we segmented the LFP data recorded during the FR paradigm into responses to individual stimuli (time window -200 ms to 600 ms relative to stimulus onset) and sorted the segments according to the seven different stimulus audio-frequencies, resulting in 50 segments per frequency. For each audio-frequency separately, we then averaged across these 50 segments and spatially smoothed the resulting LFP profile. Next, we computed CSD profiles (Equation (27)) and baseline corrected each individual CSD trace with respect to the mean across the 200-ms time window prior to stimulus onset. Subsequently, we computed the AVREC of these baseline corrected CSD profiles (Equation (28)). For each stimulus audio-frequency, we then determined the peak response amplitude within the first 100 ms after stimulus onset from this AVREC trace and thus obtained animal-specific FR functions (peak response amplitude as a function of stimulus audio-frequency).

Figure 26 shows these FR functions for the nine animals in our study. The blue marker highlights the BF and the red marker the non-BF that were selected during the recording and used for the RS paradigm (c.f. Section 2.1.2). For two of the animals (4 and 9), the FR function is multi-peaked and the peak response amplitude for the 16-kHz stimulus is larger than that for the selected BF. We decided to omit these animals from further analysis and comparison with the gerbil AC model because we could not confirm the selected BF and because the model only produced single-peaked FR functions. Similarly, animal 2 was excluded because the selected BF did not match the peak of the FR function (BF selected online: 4 kHz; peak of AVREC-based offline FR function: 2 kHz).

The FR functions for the remaining six animals were single-peaked and the BF selected during the recordings could be confirmed offline – it corresponded to the stimulus audio-frequency at which the peak of the AVREC-based FR function was located. The BFs identified within the investigated range (250 Hz to 16 kHz) were either at 1 kHz (animals 1, 3, and 5), 2 kHz (animals 7 and 8) or 4 kHz (animal 6). The BF peak response amplitude ranged from 0.88 mV/mm² (animal 1) to 3.01 mV/mm² (animal 2). The FR functions were not equally broad – FR function width in terms of the ratio between non-BF and BF peak response amplitude ranged from 0.375 (animal 8, sharpest FR function) to 0.715 (animal 5, broadest FR function).

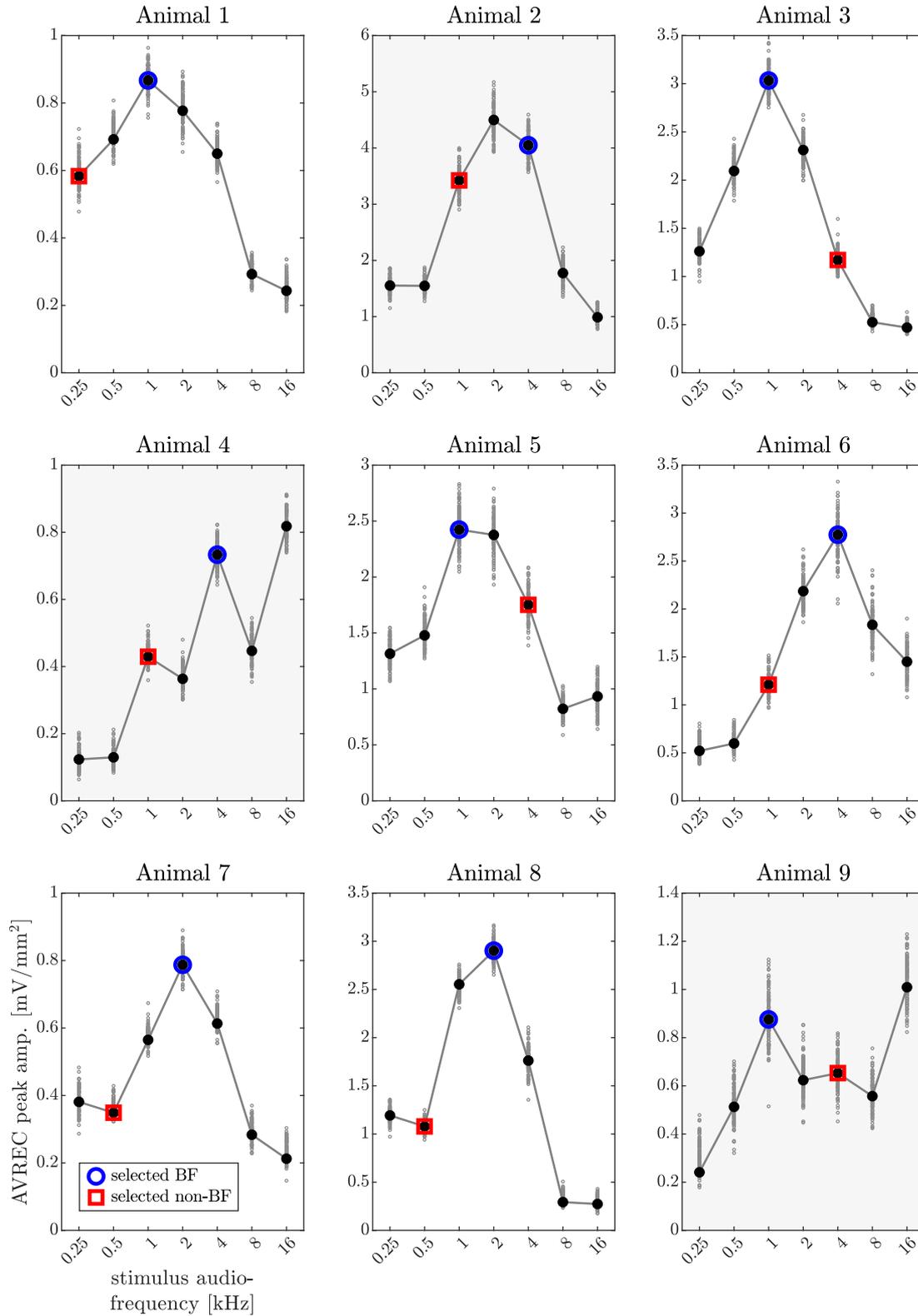


Figure 26: Overview of animal-specific FR functions. Black markers: AVREC peak response amplitude as a function of stimulus audio-frequency computed from the data recorded during the FR paradigm. The small grey markers map out associated 95% confidence intervals. How these were computed is described in the final paragraph of Section 2.3.2. Blue circles mark the BF and red squares the non-BF selected during ongoing recordings. Panels highlighted in light grey identify the three animals not considered for further analysis because FR functions were multi-peaked and/or the BF determined online did not correspond to the BF indicated by the offline FR function. Note that, across animals, the upper limit of the y-axis varies.

2.2.2 Repetition suppression stabilisation

To investigate how RS evolves across stimulus repetitions within stimulus blocks, we analysed the recorded response behaviour as a function of stimulus index i , i.e. of the stimulus position within a given stimulus block. For each animal separately, we first segmented the LFP data recorded during the RS paradigm into responses to individual stimulus blocks (time window -1 s to 18 s relative to the onset of the first stimulus in the block) and sorted the segments according to SOI value and stimulus audio-frequency (BF vs. non-BF, 24 segments per SOI, respectively – the sorting process is illustrated in Figures 30a and b). Each LFP trace within a profile (see Figure 25 for a visual representation of LFP traces making up a profile) was then separately baseline corrected with respect to the mean across the 1-s time window prior to onset of the first stimulus. Next, separately for each combination of SOI and audio-frequency, we averaged the LFP data across stimulus block repetitions. Using the approach described in Section 2.1.3, corresponding CSD profiles were then deduced from the resulting LFP profiles. Finally, we summarised the cortical activation reflected in each of the CSD profiles by computing an AVREC trace (Equation (28)). Figure 27a shows an example of such an AVREC trace (animal 8, SOI = 0.328 s, BF stimulation).

Next, we determined peak response amplitudes in time windows from 20 ms to 90 ms relative to the onset of the individual stimuli for each of the AVREC traces. Figure 27b maps out the stimulus-specific peak response amplitudes in panel a as a function of stimulus index i . Across the first few stimuli, response amplitude diminishes as a function of stimulus index. For subsequent stimuli, the amplitude reaches a plateau and only exhibits minor fluctuations for the remainder of the stimulus block. Visual inspection of the corresponding AVREC traces for all animals and SOIs, and for both BF and non-BF stimulation, suggested that the diminished peak response amplitudes caused by RS stabilise at most 2 s after the onset of the first stimulus within a block. In Figure 27b, the corresponding time window where responses were deemed ‘stable’, is highlighted in pale green.

To objectively verify the time point, t_{stable} , at which RS has stabilised across animals, SOIs, and stimulus audio-frequency, and to quantify the level of stability, we performed correlation tests (non-parametric Kendall rank correlation) for peak amplitude and stimulus index and assessed the level of fluctuation across stimulus-index-specific peak response amplitudes. Moreover, to assess the variance in the response across individual stimulus block presentations, we also computed AVREC traces from individual (non-averaged) LFP profiles (via Equations 27 and 28).

Correlation between peak response amplitude and stimulus index

We observed the most pronounced negative correlations between peak response amplitude and stimulus index for the stimulus block with the shortest SOI (0.219 s). For this block,

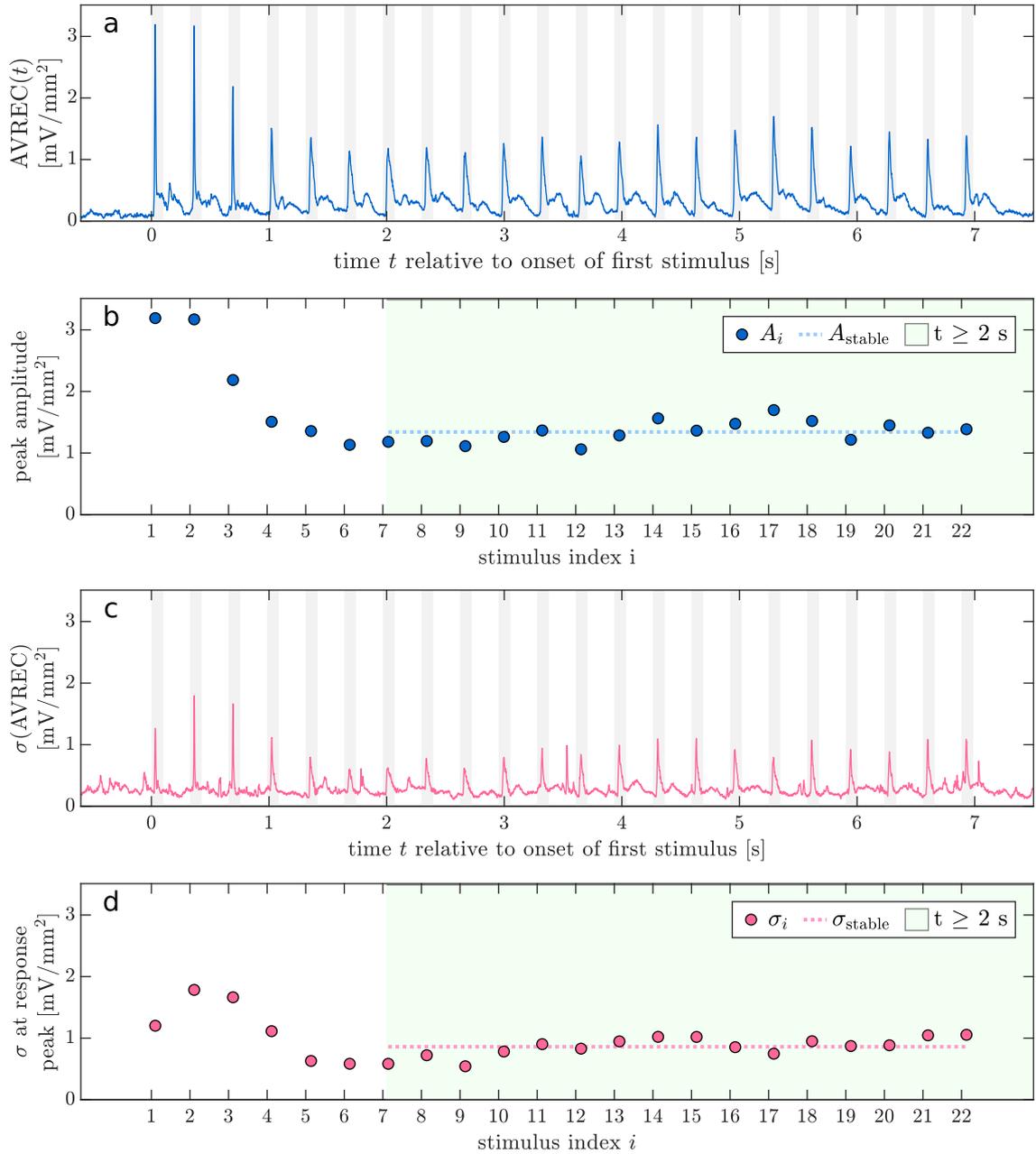


Figure 27: Stabilisation of response as a function of stimulus index for animal 8, SOI = 0.328 s and BF stimulation a: AVREC time course for 0.328-s SOI stimulus block, based on LFP profiles averaged as illustrated in Figure 30b. Grey shading maps out onset and duration of the stimuli. b: Stimulus-specific peak response amplitudes as a function of stimulus index i . Response amplitude initially diminishes as a function of stimulus index and then reaches a plateau for $t \geq 2$ s. The dotted line reflects A_{stable} , the mean peak amplitude for responses at $t > 2$ s. c: Standard deviation across AVREC traces computed from individual, unaveraged LFP profiles. d: Standard deviation at response peaks as a function of stimulus index i . As a function of time, σ exhibits pronounced fluctuations but at response peaks, σ -values stabilise after a few stimulus repetitions.

Kendall's tau for stimuli 1 to 5 was negative, with absolute values ≥ 0.6 , across all animals and both stimulus audio-frequencies. Moreover, in 9 out of the 12 cases (six animals \times two stimulus audio-frequencies), the correlation was negative and significant (Kendall's tau negative, absolute value ≥ 0.44 , p-value ≤ 0.03) up to stimulus index 9, i.e. across *all* responses occurring within the first 2 s of the stimulus block. However, using a sliding window encompassing nine responses, we found that these correlations between stimulus index and peak response amplitude quickly turned insignificant beyond the 2-s mark – on average, when the upper bound of the window reached the 12th response (i.e. when the window extended from about 0.7 s to 2.5 s) and, at the very latest, when the upper bound of the window reached the 16th response (i.e. when the window extended from about 1.5 s to 3.3 s). Across SOIs, 97.5% of the 120 separate cases (ten SOIs \times six animals \times two stimulus audio-frequencies), showed no significant negative correlation between peak response amplitude and stimulus index for the full set of responses occurring at $t > 2$ s. The very few cases with a significant correlation were not clustered at a particular SOI, animal or stimulus audio-frequency.

Stability of peak response amplitude

To assess the level of fluctuation across peak response amplitudes, we computed A_{stable} , the mean peak amplitude for responses at $t > 2$ s (separately for each animal, SOI, and stimulus audio-frequency). The dotted pale blue line in Figure 27b indicates A_{stable} in animal 8 for a SOI of 0.328 s and BF stimulation. Less than 3% of all peak amplitudes for responses occurring at $t > 2$ s (across animals, SOIs, and audio-frequencies) deviated from the respective A_{stable} -value by more than a quarter ($|A_i - A_{\text{stable}}|/A_{\text{stable}} \leq 1/4$). For the majority of these cases, the deviation was still small compared to the RS-related change in amplitude, being equal to less than a third of the difference between A_{stable} and the non-suppressed A_1 , i.e. the peak amplitude observed in response to the first stimulus within the respective block ($|A_i - A_{\text{stable}}|/(A_1 - A_{\text{stable}}) < 1/3$). Less than 1% of all cases exceeded both limits defined above and these cases were not clustered at a particular SOI or animal.

To further investigate the distribution of the evoked responses, we also evaluated the variance across individual stimulus blocks. For the same animal, SOI, and audio-frequency as in the panels above, Figure 27c shows an example of the standard deviation σ observed across AVREC traces computed from *individual* LFP profiles (i.e. computed without averaging profiles across stimulus block repetitions, as done for the trace in panel a). As a function of time, σ exhibits pronounced fluctuations. Relative to A_1 , values range from 4% up to 56%. Across all animals and SOIs, and both audio-frequencies, the maximum range even extended from 11% to 140% (non-BF stimulation and a 3.5-s SOI in Animal 6). Thus, the noise distribution for the AVREC traces was clearly heteroscedastic. However, when focussing on σ -values at times where peak responses occur in the respective AVREC traces computed from

the average LFP profile (c.f. Figure 27a), a more stable picture emerged. Figure 27d shows an example of σ -values plotted as a function of stimulus index i_{stim} . The dotted light blue line indicates σ_{stable} , the mean standard deviation for peak responses occurring at $t > 2$ s. Relative to σ_{stable} , the underlying σ_i -values deviate by no more than 35%. Less than 5% of all σ_i -values for responses occurring at $t > 2$ s (across animals, SOIs, and audio-frequencies) deviated from the respective σ_{stable} -value by more than a quarter ($|\sigma_i - \sigma_{\text{stable}}|/\sigma_{\text{stable}} \leq 1/4$).

Conclusions

In summary, our investigation showed that, in the context of the RS paradigm, peak response amplitudes initially decline as a function of stimulus index but then stabilise at $t > 2$ s, exhibiting only minor fluctuations (relative to the mean amplitude A_{stable} and to the RS-related change in amplitude $A_1 - A_{\text{stable}}$). Moreover, while the standard deviation associated with the AVREC traces varies greatly across the duration of the stimulus block and the data is thus not homoscedastic, the standard deviation associated with the peak response amplitudes at $t > 2$ s exhibits a similar level of stability as the peak amplitudes themselves.

From a neuroscientific perspective, the responses recorded during the RS paradigm are, strictly speaking, never independent as they reflect stimulus history. However, under the premise that a state of stable RS has been established and is continuously maintained, the observed peak response amplitudes can be viewed as independent and identically distributed (i.i.d.). Thus, we applied $t_{\text{stable}} = 2$ s and decoupled responses occurring at $t > t_{\text{stable}}$ from their respective stimulus indices in our subsequent analysis of peak amplitude as a function of SOI. Table 11 lists the specific number of responses for $t > t_{\text{stable}}$ at each SOI, along with the SOI-specific total number of presented stimuli within a block. Note that these numbers diverge because, when designing the paradigm, a compromise between competing demands had to be found:

- Reducing measurement duration by *reducing* the number of stimuli in a given block
- Reducing noise levels by *increasing* the number of stimuli in a given block
- *Equalising* the number of stimuli across blocks

SOI [s]	stimulus block duration [s]	number of stimuli	number of responses for $t > t_{\text{stable}}$	number of block repetitions
0.219	7.2	33	24	24
0.328	7.2	22	16	24
0.438	7.2	17	13	24
0.656	7.2	11	8	24
0.875	7.2	9	7	24
1.313	7.2	6	5	24
1.750	7.2	5	4	24
2.626	14.2	6	5	24
3.500	14.2	5	4	24
7.000	14.2	3	2	24

Table 11: Structure of stimulus blocks for the ten different SOI-values used for the RS paradigm, along with the respective number of evoked responses occurring after $t_{\text{stable}} = 2$ s. Time t is taken relative to the onset of the first stimulus in each block.

2.3 A new analysis pipeline for subject-specific statistical inferences

2.3.1 The standard approach to quantifying repetition-suppression lifetime

The first part of our analysis pipeline reflected the standard approach reported in RS studies based on extracranially recorded data from the human AC (see, for example Zacharias et al., 2012; McEvoy et al., 1997; Lu et al., 1992a). In these studies, one evoked response per SOI (and, if applicable, per condition), was determined for each subject by averaging across multiple trials. To determine evoked responses for the stabilised RS state in the gerbil AC, we first segmented the recorded LFP profiles into single trials (time window, -40 ms to 159 ms relative to stimulus onset), and baseline corrected each of the 32 individual LFP traces across cortical depth z with respect to the individual mean across the 20-ms time window preceding stimulus onset. Next, to reduce noise and thus reveal the evoked response, we computed the arithmetic average across responses recorded at $t > t_{\text{stable}}$ (for each animal, SOI, and BF vs. non-BF separately; t measured relative to the onset of the first stimulus within each block). Using the steps outlined in Section 2.1.3, we then computed CSD profiles and subsequently AVREC traces from these average LFP profiles. For one example animal, the resulting AVREC traces for all SOIs are shown for BF and non-BF stimulation in Figures 28a and b, respectively. The peak response amplitude increases with SOI. We determined amplitude and latency of the AVREC response peaks within a time window from 20 ms to 90 ms. Figures 28c and d map out the peak amplitudes from Figures 28a and b as a function of SOI for BF stimulation (circles) and non-BF stimulation (diamonds).

To quantify the lifetime of RS, we fitted Equation (1) to the SOI-specific peak response amplitudes. We decided to set parameter t_0 to the stimulus duration (0.100 s), reflecting the logic that once SOI equals stimulus duration and individual tones merge into a continuous

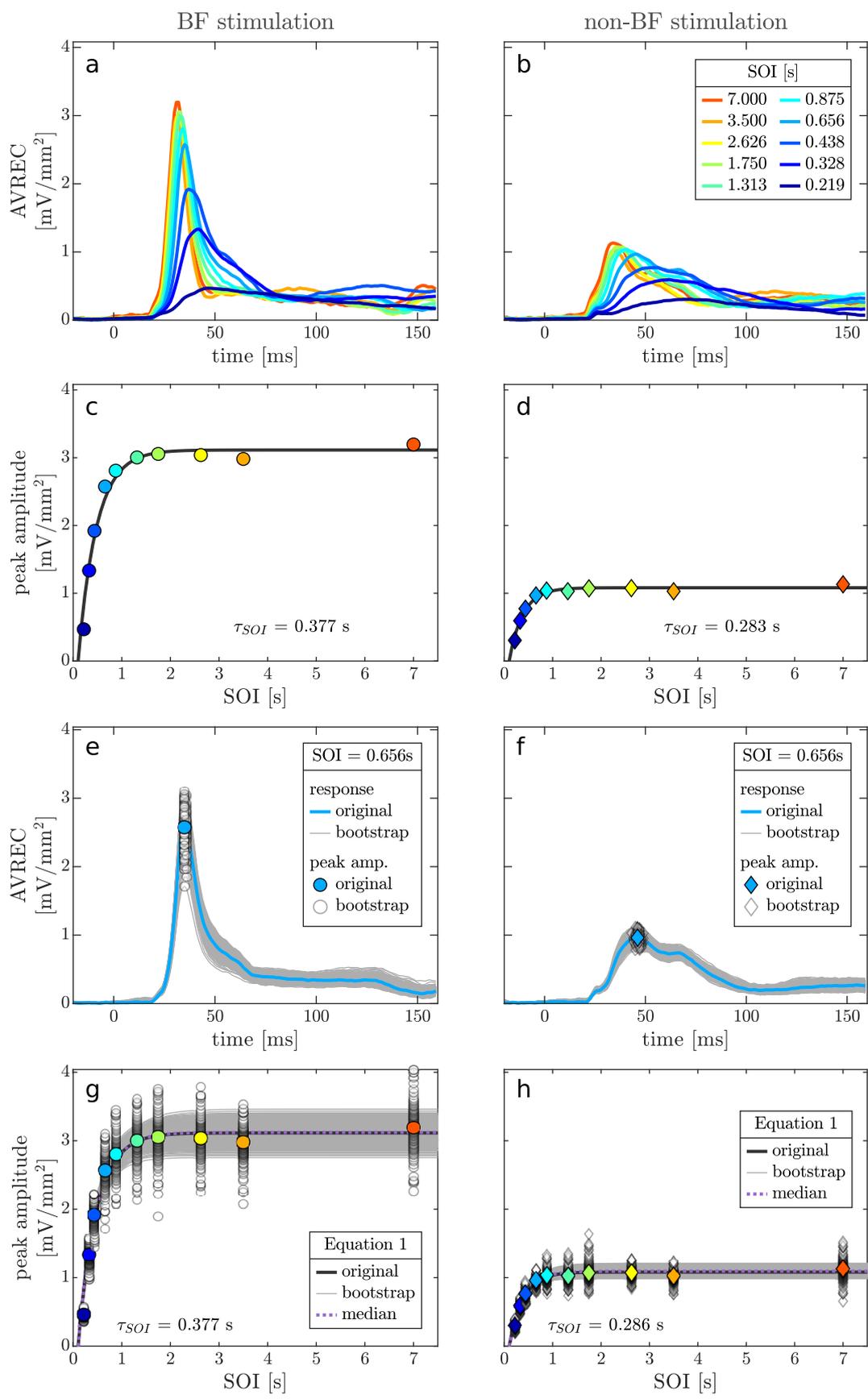


Figure 28: (Previous page.) Procedure for determining RS lifetime τ_{SOI} using the example of a single animal. Panels on the left reflect BF stimulation; those on the right non-BF stimulation. **a, b:** SOI-dependence of evoked responses for animal 8. The individual waveforms represent evoked responses in terms of the AVREC for SOIs between 0.219 s and 7 s. With increasing SOI, the peak amplitude gradually increases and converges towards a saturation value at the largest SOIs. Stimulus onset is at time $t = 0$ ms. **c, d:** Derivation of τ_{SOI} from the SOI-dependence of the peak amplitudes of the data shown in **a** and **b** for BF (circular markers) and non-BF (diamond markers) stimulation. The curves are fits of Equation (1). **e, f:** Generation of surrogate data using the bootstrap technique. AVREC time courses for the 0.656-s SOI are displayed as an example. The blue curves represent the original evoked response shown in **a** and **b**. The light grey waveforms represent the 999 bootstrap-based evoked response for each stimulus audio-frequency, with the corresponding peak amplitudes displayed as circular (**e**) and diamond (**f**) markers. Original peak amplitudes are highlighted in blue. **g, h:** Fits of Equation (1) to the SOI-dependence of the peak amplitudes from the original (coloured circles and diamonds, see **c** and **d**, respectively) and the resampled data (white circles and diamonds, see **e** and **f**, respectively). The fits to the resampled data are shown as light grey curves. The dotted curves depict Equation (1) computed with the median values of τ_{SOI} and A_{sat} across all fits.

tone, no evoked response peak should be detectable beyond the first stimulus. To ensure robust fitting results, our fitting algorithm consisted of a two-step process. First, we applied a non-iterative regression method based on an appropriate integral equation (Jacquelin, 2009). Next, the estimated parameter values obtained via this approach were used as the starting points for a further optimisation of the fit via an iterative least squares approach (lsqnonlin function, MATLAB R2021b). Ultimately, we thus obtained two τ_{SOI} -values per animal – one for BF and one for non-BF stimulation. In Figures 28c and d, the corresponding fitting results for the data displayed in **a** and **b** are shown as a black curve. For BF-stimulation, τ_{SOI} was equal to 0.377 s, whereas for non-BF stimulation, it was equal to 0.283 s.

Given the limited sample size, both in terms of SOI-values and in terms of stimulus repetitions per SOI, and the jitter of the data points around the fitted curve, questions regarding the robustness of the fitting results arise. How much would the subject-specific τ_{SOI} -values change if we were to repeat the experiment? Are RS recovery dynamics indeed affected by stimulus audio-frequency or are we simply observing two possible outcomes based on the same underlying distribution? To better answer these questions, we extended our analysis pipeline.

2.3.2 A bootstrap-based expansion of the standard approach

Figure 29 shows a representative example of the large variability across the single-trial data that forms the basis of the evoked response revealed via averaging. Panel **a** shows the 192 individual AVREC traces computed from the 192 single-trial LFP profiles recorded in animal 8 for the 0.656-s SOI and BF stimulation (24 block repetitions \times 8 responses at $t > t_{\text{stable}}$, c.f. Table 11). Panel **b** shows the evoked response computed from the arithmetic average across the 192 single-trial LFP profiles. This blue curve corresponds to the same-colour curve in Figure 28a. The range of peak amplitudes reflected by the single-trial data is very large. It exceeds the range of the peak amplitudes of the trial-averaged evoked responses observed

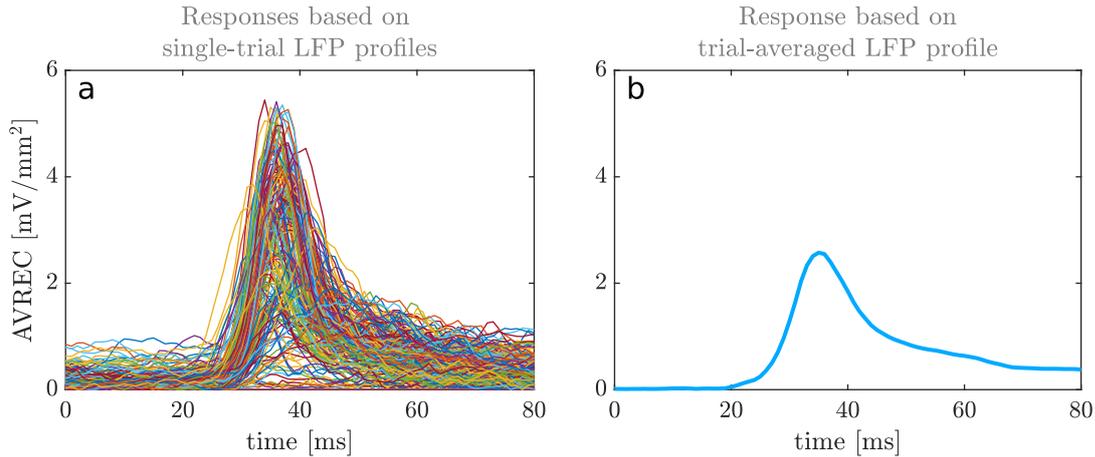


Figure 29: Illustration of variability across single-trial data. **a:** 192 individual AVREC traces computed from the 192 single-trial LFP profiles recorded in animal 8 for the 0.656-s SOI and BF stimulation (24 block repetitions \times 8 responses at $t > t_{\text{stable}}$, c.f. Table 11). **b:** Evoked response computed from the arithmetic average across the 192 single-trial LFP profiles. This blue curve corresponds to the same-colour curve in Figure 28a.

across SOI-values (c.f. Figure 28).

In an ideal world, we would extend our sample size (in terms of SOIs and stimulus repetitions) to increase the robustness of our estimates of RS lifetime and/or repeat our recordings to establish confidence intervals for these estimates. In the real world, the non-parametric bootstrap technique (Efron, 1979, for a review, see, for example Hesterberg, 2011) offers the next best thing: available data is resampled to create surrogate data sets. The power of this technique lies in its ability to yield statistical inferences – such as confidence intervals (CIs) or medians – without postulating a normal distribution or homoscedasticity for the original data. Therefore, bootstrap methods can be remarkably more accurate than other classical approaches relying on these assumptions.

The non-parametric bootstrap technique can be applied in a meaningful way when the set of observations, in our case the response peaks, can be assumed to originate from an independent and identically distributed (i.i.d.) population. In Section 2.2.2, we have shown that this is indeed a reasonable assumption for our data. From a neuroscientific perspective, the response peak amplitudes observed for the regular-SOI paradigm are, strictly speaking, never independent as they reflect stimulus history. However, we demonstrated that, after a few stimulus repetitions, a stable state of RS is established and maintained, and the observed peak response amplitudes can be viewed as i.i.d. Therefore, we applied the non-parametric bootstrap technique: for each animal, SOI, and stimulus audio-frequency, we resampled the corresponding 24 LFP profiles recorded during stimulus block presentation by randomly drawing one of the profiles a total of 24 times. Selected profiles were available again in subsequent draws, i.e. sampling occurred with replacement. For one SOI, this resampling is illustrated in Figure 30c. We then proceeded with the analysis described in the previous section (2.3.1) for

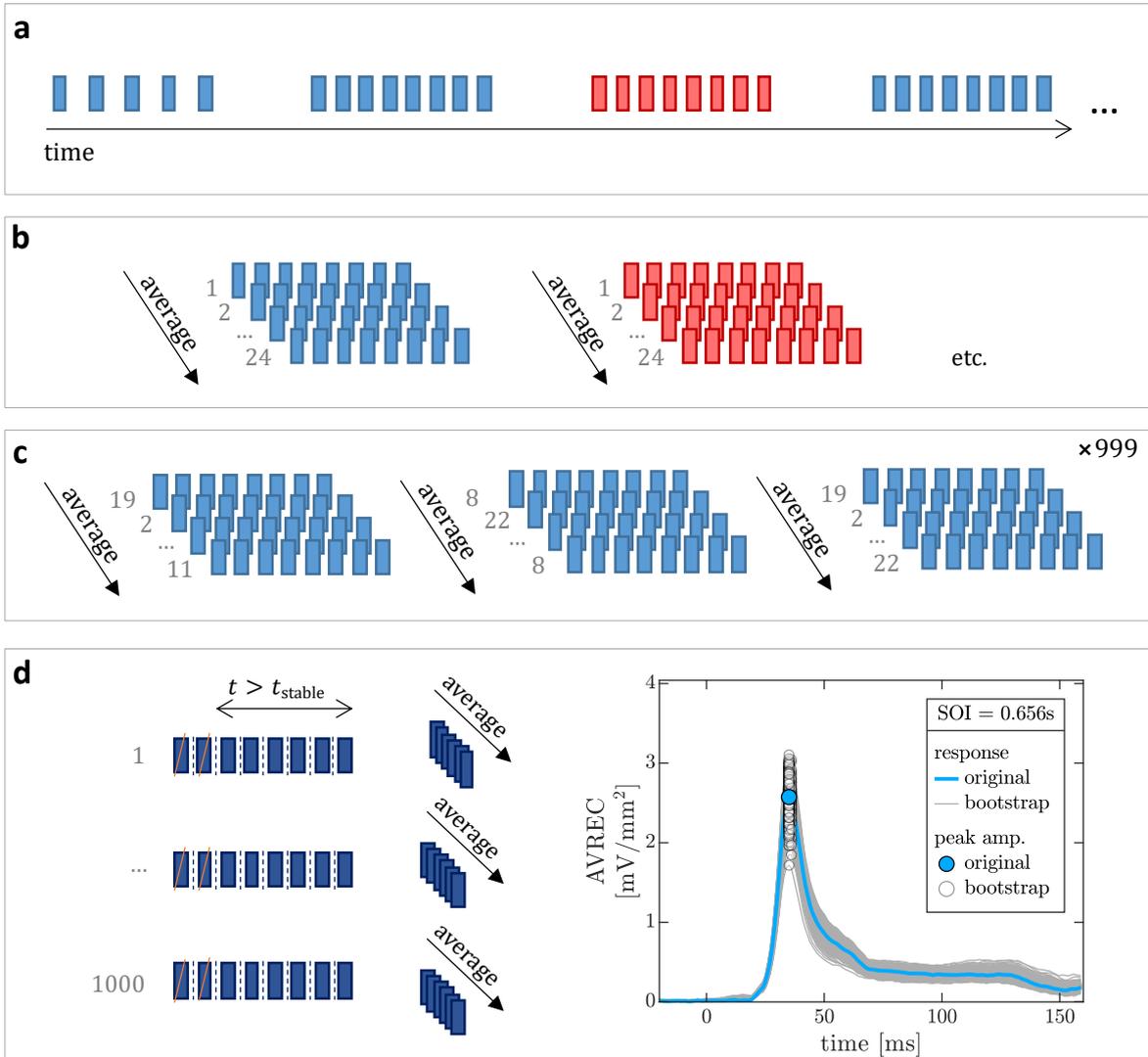


Figure 30: Illustration of resampling procedure applied to the LFP profiles recorded from gerbil A1. **a:** To investigate RS, a sequence of stimulus blocks, characterised by 10 different SOIs and two different stimulus audio-frequencies (BF and non-BF), was presented. In total, the sequence comprised 24 blocks per SOI and audio-frequency. During stimulation, an intracortical electrode recorded a continuous LFP profile. **b:** The continuous recording was segmented into responses to individual stimulus blocks. These were then sorted according to SOI and stimulus audio-frequency, resulting in 24 LFP profiles per condition. To assess response stabilisation across stimulus repetitions, AVREC traces were computed (Equations 27 and 28) based on the average across the respective 24 LFP profiles (for an example of a corresponding result, see Figure 27a). **c:** To assess the robustness of and confidence intervals for the evoked responses revealed via averaging, we first resampled the 24 LFP profiles per condition by drawing a random profile a total of 24 times. Selected profiles were available again in subsequent draws, i.e. sampling occurred with replacement. We repeated this procedure 999 times and computed an average LFP profile from each of the resampled sets. **d:** The resulting 1000 profiles (original plus resampled) were then each segmented into responses to individual stimuli and stabilised responses ($t \geq t_{\text{stable}}$, c.f. Section 2.2.2) were averaged to reveal LFP profiles reflecting the response evoked by the stimulus. Finally, 1000 AVREC traces were computed from these 1000 profiles. The plot on the right illustrates the obtained results for BF stimulation and a 0.656-s SOI in animal 8, as shown in Figure 28e.

this newly created data set: segmentation into single trials \rightarrow baseline correction \rightarrow arithmetic averaging across LFP profiles recorded at $t > t_{\text{stable}}$ \rightarrow computation of AVREC \rightarrow determination of response peak amplitude and latency. For each animal, SOI, and stimulus audio-frequency, we created 999 additional data sets via resampling, thus obtaining a total of 1000 AVREC traces reflecting evoked responses (1 original + 999 resampled) as well as the 1000 associated peak response amplitude and peak latency values (see Figure 30d for a visualisation).

Figures 28e and f show examples of the results obtained via this approach for the 0.656-s SOI with BF and non-BF stimulation, respectively. The blue curves represent the original AVREC traces, as shown in a and b. The grey curves represent additional evoked responses deduced from the resampled data sets. The respective response peaks are labelled with circular (BF stimulation) and diamond-shaped (non-BF stimulation) markers. The response peaks associated with the original data, as shown in b and c, are highlighted in blue. The grey band mapped out by the bootstrapped AVREC traces illustrates the variability across the single-trial responses that form the basis of the original evoked response, i.e. the spread of the sampling distribution is reflected by the spread of the bootstrap distribution. Similarly, the vertical scatter of the corresponding response peak amplitudes serves as an estimate of the confidence intervals (CIs, for example the 95-% CI) associated with the original peak amplitude.

In the final step of our analysis, we used these peak amplitudes to determine CIs for τ_{SOI} and A_{sat} . First, we created sets of data containing one peak response amplitude per SOI, i.e. 10 data points per set. The selection of the peak amplitudes from the pool of surrogate data occurred randomly and without replacement and we used up all data points, i.e. given the 1000 peak amplitudes \times 10 SOIs = 10,000 data points, we produced 1000 subsets with 10 data points each. Next, we fitted Equation (1) to each of these subsets. For the same example animal as in the panels above, Figures 28g (BF stimulation) and h (non-BF stimulation) show the final results obtained via this approach. The original peak amplitudes, along with the fit of Equation (1) to these amplitudes, are reflected by coloured markers and a black curve, just as in panels c and d. The vertical distribution of circular/diamond-shaped markers at each SOI maps out the peak response amplitudes deduced from the bootstrapped evoked responses (see Figures 28e and f) and the grey curves reflect the fits of Equation (1) to the described subsets of this data. The grey band mapped out by these curves demarcates ranges for the associated fitting parameter values. These ranges are a result of the variability across the original single-trial data and thus provide CIs for the τ_{SOI} - and A_{sat} - values associated with the original AVREC traces depicted in panels a and b.

We determined median values $m(\tau_{\text{SOI}})$ and $m(A_{\text{sat}})$ from the respective distributions for each animal and stimulus audio-frequency. These medians, along with the fixed intercept $t_0 = 100$ ms, were used in Equation (1) to compute the final RS recovery function. The

dotted curve in Figures 28g and h reflects the corresponding final result for animal 8. For BF stimulation, $m(\tau_{\text{SOI}})$ was equal to 0.377 s and for non-BF stimulation, it was equal to 0.286 s. In agreement with the FR functions (Figure 26), the saturation amplitude was larger for BF than for non-BF stimulation, with $m(A_{\text{sat}}) = 3.122$ mV/mm² and 1.095 mV/mm², respectively. This trend generalised to all animals.

Across all animals and conditions, original values and bootstrap-based medians were in good agreement, differing by less than 5% for τ_{SOI} and by less than 2% for A_{sat} . We were thus able to confirm the robustness of our original estimates despite the aforementioned limitations regarding sample sizes (see final paragraph of Section 2.3.1). A repetition of the resampling process also confirmed that the number of surrogate data sets was sufficient for the stabilisation of median values and confidence intervals, with respective median values differing by less than 1% and ranges of respective 95-% CIs by, at most, 8%.

We also used the bootstrapping technique to compute the CIs for the FR functions shown in Figure 26. For each stimulus audio-frequency, we resampled the available 50 single-trial responses to create surrogate data sets. Each surrogate data set contained 50 random samples, drawn with replacement, and we produced a total of 100 such data sets per stimulus audio-frequency. Each of the grey markers in Figure 26 reflects the peak amplitude of the AVREC trace computed from the average LFP profile across one of these surrogate data sets. The approach confirmed that, in the selected animals, there was a robust difference between the peak response amplitudes for BF and non-BF stimulation.

2.3.3 SOI-specific peak latencies

For both BF and non-BF stimulation, we observed a change in peak latency across SOIs. More specifically, peak latency decreased with SOI, asymptotically approaching a minimum value for the largest SOIs. Two examples of the shift in peak latency across SOIs can be seen in Figures 28a and b. Just as in our simulations, the relationship between SOI-specific peak latency and peak amplitude could be described by a linear function (Equation (21)), with latency shift μ quantifying the direction and magnitude of the change in peak latency across SOIs. Akin to the approach for RS lifetime, we used bootstrap-based peak amplitudes and latencies to determine CIs for μ . Note that, while the vertical scatter of the 1000 grey circular markers in Figures 28e and f represents CIs associated with the original peak amplitude observed for animal 8, the mild horizontal scatter reflects CIs for the original peak latency of the 0.656-s SOI.

For each animal and stimulus audio-frequency respectively, we created subsets of data containing one peak response amplitude and one peak latency per SOI, i.e. 10 value pairs per subset. The selection from the pool of surrogate data occurred randomly and without replacement, meaning that all available values were used up after creating 1000 subsets.

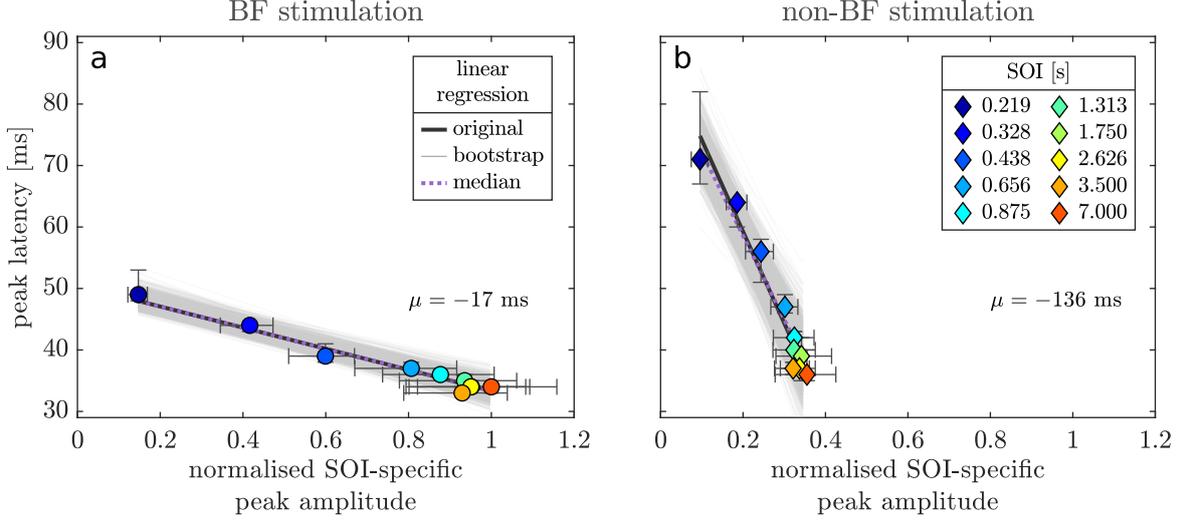


Figure 31: Peak latency as a function of SOI-specific peak amplitude for BF (a) and non-BF (b) stimulation in animal 8. Coloured markers reflect the original values and the black line is the associated linear regression. Error bars map out respective 95-% CIs across the resampled data. The fits to the resampled data are shown as grey lines and map out a grey band reflecting the ranges of the fitting parameters.

For better comparability of μ -values across animals, we normalised peak amplitudes with respect to the peak amplitude of the 7-s SOI at BF stimulation. Next, we performed a linear regression (Equation (21)) for each of these subsets. For the same example animal as in Figure 28, Figure 31 shows a summary of the results obtained via this approach. The original peak latencies are plotted as a function of the original peak amplitudes and the colour code of the markers reflects the associated SOI-value. The linear regression for these data points is reflected by a black curve. The error bars reflect respective 95-% CIs, deduced from the bootstrapped evoked responses (c.f. Figures 28e and f). Linear regressions to bootstrap-based subsets of data are reflected by grey lines. Just as for RS lifetime, the grey band mapped out by these lines demarcates ranges for the associated fitting parameter values (Equation (21)). These ranges are a result of the variability across the original single-trial data and provide CIs for μ -values associated with the original AVREC traces.

We determined median values for μ and the x-intercept c from the respective distributions for each animal and stimulus audio-frequency. The dotted line in Figures 31a and b reflects the final result for animal 8. For BF stimulation, $m(\mu)$ was -17 ms and for non-BF stimulation, it was -136 s. Across all animals and conditions, original values and bootstrap-based medians were in good agreement, differing by less than 1%. Just as for RS lifetime, we were thus able to confirm the robustness of our original estimates despite the aforementioned limitations regarding sample sizes (see final paragraph of Section 2.3.1). A repetition of the resampling process also confirmed that the number of surrogate data sets was sufficient for the stabilisation of median values and confidence intervals, with respective median values differing by less than 1% and ranges of respective 95-% CIs differing by, at most, 5%.

2.4 Final results and comparison with model predictions

In vivo, we presented stimuli at the BF and non-BF of the measurement location. This corresponds to conditions $\Delta F = 0$ oct and $\Delta F = 2$ oct in the gerbil AC model. Given the simulated stimulus audio-frequencies, 2 kHz and 8 kHz, these conditions were met for A1 columns 7 and 11 (c.f. Figure 19). Column 7 had a BF of 2 kHz and column 11 a BF of 8 kHz. The BFs observed in vivo ranged from 1 kHz to 4 kHz, with a mean of 2 kHz, therefore, we compared our in-vivo results to the response behaviour of A1 column 7. Figure 32 summarises in-vivo (left) and in-silico (right) RS response behaviour for the six animals and across the three different AC model parameter settings, respectively. The in-silico results presented in this figure correspond to the highlighted values in Figure 20.

2.4.1 Repetition suppression lifetime

Figure 32a summarises median τ_{SOI} -values and associated 95-% CIs across animals and across BF and non-BF stimulation. We found that RS lifetime in vivo was always shorter than 1 s. Median τ_{SOI} -values ranged from 0.197 s (animal 6, non-BF stimulation) to 0.465 s (animal 3, BF stimulation). For two animals (1 and 5), RS lifetime was longer for non-BF than for BF stimuli, with median values and 95-% CIs of τ_{SOI} above the line of equality, and for two animals (6 and 8), RS lifetime was shorter for non-BF stimuli (median values and 95-% CIs of τ_{SOI} below the line of equality). For the remaining two animals (3 and 7), RS lifetime was not significantly affected by the audio-frequency of the stimuli, with median values close to the line of equality and 95-% CIs extending on either side. Thus, in four out of six animals, the lifetime of RS at a given measurement location changed with stimulus audio-frequency. Time constants differed by up to 132 ms between the BF and non-BF condition (animal 1).

Compared to panel a, Figure 32b illustrates that, overall, RS lifetime was longer in silico than in vivo. In terms of relative differences between the BF and non-BF condition, simulations predicted an increase in RS lifetime from BF to non-BF stimulation, i.e. the scenario observed for animals 1 and 5. Moreover, our simulations with the gerbil AC model predicted that differences between BF and non-BF τ_{SOI} are more pronounced the sharper the FR function. We were not able to identify a similar correlation for the in-vivo data. Animals 1 and 5 had broader FR functions than animals 3 and 7 (c.f. Figure 26), yet exhibited larger differences between BF and non-BF τ_{SOI} .

2.4.2 Peak latency shift

Figure 32c summarises μ -values across animals as well as across BF and non-BF stimulation. For both stimulus audio-frequencies, median μ is negative across all six animals. Thus, peak latency was always delayed with decreasing SOI. The result is relatively robust, with only one

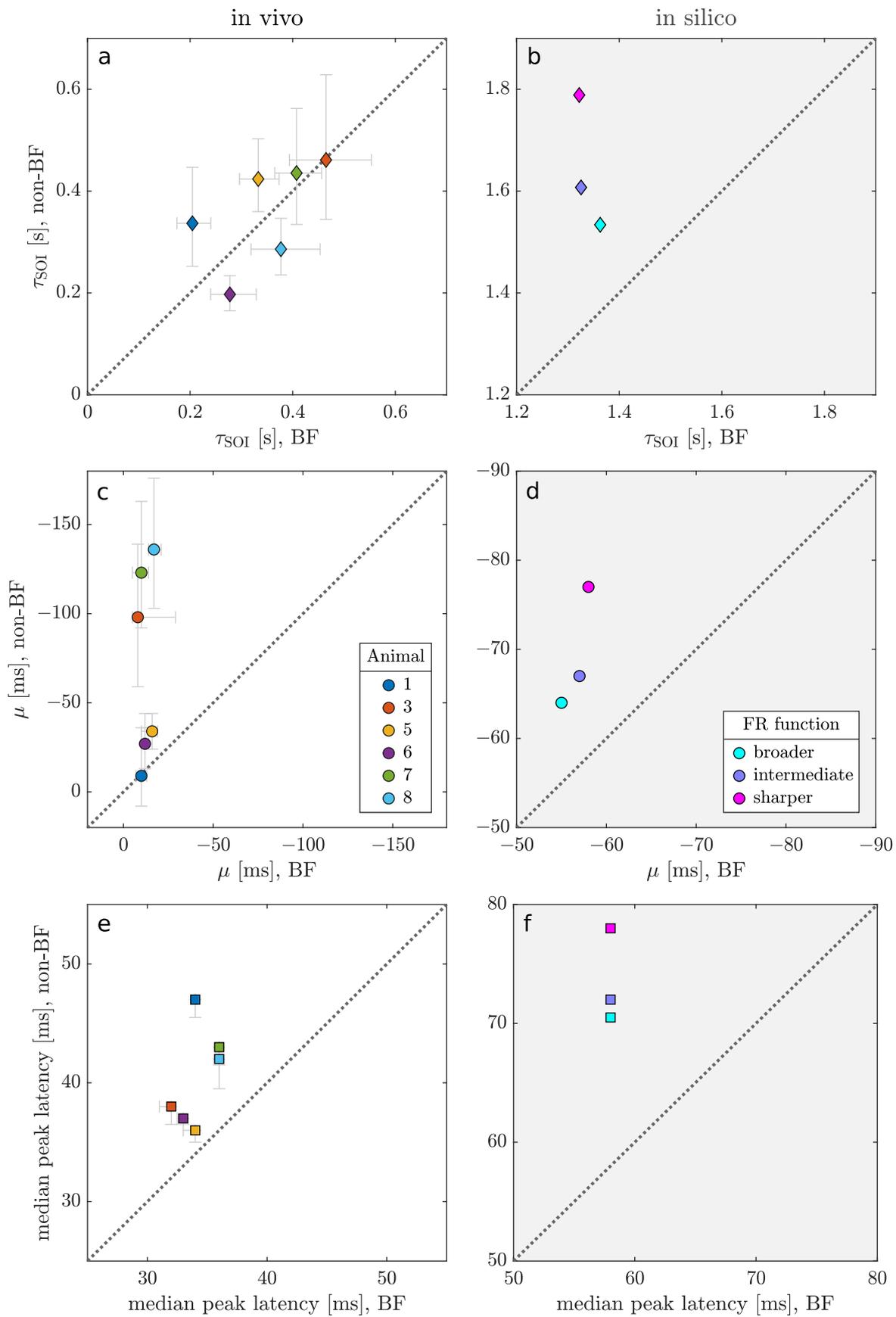


Figure 32: Comparison of in-vivo and previously introduced in-silico results. **a:** RS lifetimes in vivo. Values of τ_{SOI} for non-BF stimulation are plotted against τ_{SOI} -values for BF stimulation. Diamond markers reflect median values across the bootstrapped data and the capped grey error bars reflect the respective 95-% CIs. The dotted grey diagonal line marks equality between the conditions. **b:** RS lifetimes in silico (highlighted values from Figure 20a to c). Note that while the x- and y-axis in **a** and **b** each extend over a 700-ms range, absolute values are larger in **b** than in **a**. **c:** Latency shift in vivo. μ -values for non-BF stimulation are plotted as a function of μ -values for BF stimulation. The x- and y-axis cover the same range of μ -values and are both inverted, with increasingly negative values plotted in the right and upward direction, respectively. The circular markers reflect median values across the bootstrapped data and the capped grey error bars indicate the respective 95-% CIs. **d:** Latency shift in silico (highlighted values from Figure 20d to f). Note that the overall range of μ -values was much larger in vivo than in silico, with axes in **c** covering a range of 200 ms whereas axes in **d** only cover a range of 40 ms. **e:** Median peak latency across SOIs for non-BF stimulation, plotted against values for BF stimulation. The square markers reflect median values across the bootstrapped data and the capped grey error bars reflect the respective 95-% CIs. The dotted grey diagonal line marks equality between the conditions. **f:** Peak latency in silico (highlighted values from Figure 20g to i). Note that while the axes in **e** and **f** all extend over a 30-ms range, absolute values are larger in **f** than in **e**.

of the 95-% CIs for μ reaching into the positive range, and only up to the 8 ms mark (animal 1, BF stimulation).

For BF stimulation, μ -values are limited to a relatively small range. Median μ values extend from -8 ms (animal 3) to -17 ms (animal 8) and the grand median is -11 ms. For non-BF stimulation, the picture is much more diverse. Median μ -values range from -9 ms (animal 1) to -136 ms (animal 8). For five of the six animals (3, 5, 6, 7, and 8), the latency shift is more pronounced for non-BF than for BF stimulation. The corresponding median μ -values are above the line of equality and the respective 95-% CIs do not cross this line. For animal 1, median μ -values are equal across conditions (-10 ms at both BF and non-BF stimulation) and the 95-% CIs extend on either side of the line of equality.

As repeated in Figure 32d, simulations with the gerbil AC model predicted what the in-vivo data showed (recall Section 1.2.2 and Figures 20d to f therein):

- Peak latency decreases with increasing SOI (reflected by negative μ -values)
- μ -values for BF stimulation do not change with FR function sharpness
- The magnitude of μ increases from BF to non-BF stimulation

Moreover, simulations predicted that differences between BF and non-BF μ -values are more pronounced the sharper the FR function. This was coarsely confirmed in vivo. While there was no statistically significant correlation between FR function sharpness and the difference between BF and non-BF μ -value, the two animals with the sharpest FR functions (8 and 3) exhibited much larger differences in median BF vs. non-BF μ -values than the two animals with the broadest FR functions (1 and 5). The overall range of non-BF μ -values was much larger in vivo than in silico but this was accompanied by a larger range of FR function sharpness.

This is in agreement with the prediction that differences between BF and non-BF μ -values are more pronounced the sharper the tuning.

2.4.3 Median peak latency

Overall, peak latencies were shorter in vivo than in silico. Figure 32e summarises median peak latencies across SOIs and the associated 95-% CIs for all animals and both BF and non-BF stimulation. For BF-stimulation, latency-values are limited to a relatively small range. They extend from 32 ms (animal 3) to 36 ms (animals 7 and 8). For non-BF stimulation, the picture is more diverse. Values range from 36 ms (animal 5) to 47 ms (animal 1). In all six animals, the median peak latency is longer for non-BF than for BF stimulation. Values lie above the line of equality (grey dots) and the respective 95-% CIs do not cross this line.

As repeated in Figure 32f, simulations with the gerbil AC model predicted the increase in peak latency from BF to non-BF stimulation. Moreover, simulations predicted that the delay is more pronounced the sharper the FR function. This could not be confirmed in vivo. While the two animals with the sharpest FR functions (animals 8 and 3) exhibited larger differences in BF vs. non-BF peak-latency-values than the animal with the widest FR function, the biggest difference was observed for the animal with the second widest FR-function (animal 1).

2.4.4 Preliminary conclusions

As summarised in Figure 32, the main predictions made with the gerbil AC model were confirmed in vivo:

- RS recovery dynamics at a given measurement location are not necessarily constant. The lifetime of RS changed from BF to non-BF stimulation in four out of six animals.
- Peak latency is also not constant and varies across both SOIs and stimulus audio-frequencies.
- Peak latency is longer for non-BF than for BF stimuli.
- A decrease in SOI leads to an increase in peak latency and this increase is steeper for non-BF than for BF stimuli

Note however, that the sample size (6 animals) was rather small and should be increased to further verify results at the group level. This could also shed additional light on the predicted correlations between FR function sharpness and difference between BF and non-BF response measures. Furthermore, the FR paradigm might need to be adapted. With the current paradigm, the in-vivo FR functions only have a 1-oct resolution. Thus, FR function

sharpness could be misrepresented, especially when the ‘true BF’ at the measurement location lies halfway between two investigated audio-frequencies.

While group-level results need to be confirmed by extending the subject pool, we were able to confirm the robustness of our results at the single-subject level via a newly developed, bootstrap-based analysis pipeline. This pipeline can also be applied to regular-SOI paradigm responses recorded from other species and/or via other methods. Furthermore, it can serve as a tool to verify the robustness of group-level results. In the study summarised in Chapter 3, we were able to showcase the benefits of our analysis pipeline by applying it to MEG data recorded from the AC of 14 human subjects.

Peak latency, total duration of the evoked response, and RS recovery, were all slower in silico than in vivo. To better match in-silico and in-vivo results, the gerbil AC model’s time constants (c.f. Table 1) could be scaled down. Note, however, that optimising the model to better match in-vivo results by manually adjusting model parameter settings is a tedious process and does not allow for a systematic search of the most suitable values within the full parameter space. Therefore, an automation of the parameter-value search is desirable. In the project summarised in Chapter 4, we addressed this need and pioneered a computational approach for the systematic optimisation of parameter values in a simplified AC model. This optimisation pipeline has the potential to also be adapted for use with the gerbil AC model.

The prediction that RS lifetime is longer for non-BF than for BF stimulation was tested with mixed success. This prediction emerged from a network where τ_{rec} , the time constant of recovery from STSD, was identical across all connections targeting AC. In simulations with the single-column model, we showed that mixing τ_{rec} -values can lead to a scenario where τ_{SOI} decreases from BF to non-BF, as observed in vivo for animals 6 and 8. However, it remains to be explored how introducing different τ_{rec} values across columns and/or fields would affect the response behaviour of the gerbil AC network. As we have seen in Section 1.2.2, activation patterns across the entire network affect the RS lifetime of individual columns. Therefore, even for seemingly equivalent parameter settings across column, the response behaviour of the single-column model does not necessarily predict the response behaviour of the gerbil AC model. When τ_{rec} was constant across all connections, there was a decrease in RS lifetime τ_{SOI} as a function of distance from the BF in the single-column model (c.f. Section 1.2.1) but an increase in the gerbil AC model (c.f. Section 1.2.2). Further simulations with a range of different model parameter settings are needed to investigate how a decrease in τ_{SOI} as a function of distance from the BF might arise in the gerbil AC model. In the Discussion, we will address which changes to and expansions of the gerbil AC model could be particularly beneficial in the future, especially in the context of also tackling more complex stimulation paradigms.

Part III: Refining strategies for repetition suppression studies

Project 3: An analysis pipeline to verify the robustness of lifetime estimates

The data set we used to determine RS lifetimes for BF and non-BF stimulation in gerbil A1 came from a small group of animals ($n = 6$) and the relationship between the RS lifetimes for the two conditions was highly subject-specific. Our bootstrap-based analysis pipeline helped to confirm the robustness of the subject-specific RS lifetime estimates. Here, we demonstrate how an adapted and expanded version of our bootstrap-based analysis pipeline improved the investigation of a potential group-level asymmetry in MEG-based RS lifetimes of the left and right human AC.

Numerous hemispheric lateralisations in both structure and function have been observed across the auditory system (for a review, see [Ruthig and Schönwiesner, 2022](#); [Ocklenburg and Güntürkün, 2024](#)). This raises the question of whether lifetimes of auditory sensory memory, as reflected by RS lifetime τ_{SOI} , also differ across hemispheres. The literature reports that the amplitude of the N1/N1m response is larger in the hemisphere contralateral to the stimulated ear when stimulation is monaural (see, for example, [Hine and Debener, 2007](#); [McEvoy et al., 1997](#); [Pantev et al., 1998](#); [Reite et al., 1981](#); [Ross et al., 2005](#); [Salmelin et al., 1999](#); [Woldorff et al., 1999](#)) but with binaural stimulation, no pronounced lateralisation was observed ([Ross et al., 2005](#)). Results regarding the lateralisation of RS lifetime deduced from the N1m response amplitudes are inconclusive.

For the small group size of four subjects but a large number of investigated SOIs (11) ranging from 0.5 s to 16 s, [Lu et al. \(1992a\)](#) found similar RS lifetimes across hemispheres when stimulating the respective contralateral ear. [Rojas et al. \(1999\)](#) also delivered stimuli to the ear opposing the hemisphere where MEG recordings were obtained. They found no difference in RS lifetimes across hemispheres but while their group size was larger (25), they used only six SOIs ranging from 1 s to 6 s. Using five SOIs ranging from 0.6 s to 8 s and stimulating the left ear of eight participants, [McEvoy et al. \(1997\)](#) found no statistically significant difference in τ_{SOI} between left and right hemisphere. In contrast, [Zacharias et al. \(2012\)](#) found that mean τ_{SOI} -values across 15 subjects were about 600 ms longer in the left than in the right hemisphere when using a regular-SOI paradigm consisting of five SOIs

ranging from 0.5 s to 10 s and left-ear stimulation. However, they did not investigate the significance level of this difference. Finally, using binaural stimulation, [Cheng and Lin \(2012\)](#) found no significant difference between RS lifetimes of the left and right AC in three groups of 15 subjects each. They used five SOIs, ranging from 0.5 s to 8 s.

None of the studies introduced above investigated the robustness and confidence intervals of the subject-specific RS lifetime estimates obtained for the left and right hemisphere. Thus, while four out of five studies report that left and right AC are characterised by the same RS lifetime, conclusively answering the question of whether a difference exists requires further investigation. In the study presented below ([Dar & Härtwich et al., 2025](#)), we adapted the bootstrap-based analysis pipeline introduced for the analysis of electrophysiological recordings from gerbil AC to the analysis of MEG recordings from left and right human AC. This allowed us to compute CIs for all subject- and hemisphere-specific response measures and to base our investigation of hemispheric lateralisation of RS lifetimes at the group level on subject-specific statistics.

3.1 Materials and Methods

The following sections first introduce the subject pool, the design of the regular-SOI paradigm used to investigate RS in the human AC, and the data acquisition process. Next, pre-processing steps and ERF computation are discussed and finally, the bootstrap-based analysis pipeline applied to the data is described.

3.1.1 Subjects

Fourteen subjects, belonging mainly to the academic environment of the Leibniz Institute for Neurobiology and the Otto von Guericke University Magdeburg, participated in the study. All participants were healthy and presented with normal audiograms. The study was approved by the Ethics Committee of the Otto von Guericke University Magdeburg and subjects provided written consent to take part in MEG measurements.

3.1.2 Experimental paradigm

The regular-SOI paradigm used for the MEG measurements consisted of ten stimulus blocks, characterised by respective SOIs of 0.25 s, 0.50 s, 0.75 s, 1.00 s, 1.50 s, 2.00 s, 3.00 s, 4.00 s, 5.00 s, and 7.00 s. Stimuli had a duration of 100 ms, with a linear rise and fall time of 5 ms, respectively, and were presented at an audio-frequency of 1.5 kHz. To improve signal-to-noise ratio but keep overall measurement duration at a level manageable for the subjects, blocks with SOIs ≤ 0.750 s consisted of 120 stimuli whereas the remaining blocks consisted of 100 stimuli. Blocks were presented in a random order and the resulting full measurement session took about 45 minutes.

Stimuli were simultaneously delivered to both ears of the subject using a plastic tube connected to an earmold. The commercial *Presentation* software package (Neurobehavioral Systems Inc., Albany, CA) was used on a conventional PC to generate each tone. The onset of each stimulus block was independently initiated by the subject and required pressing a distinct button on a keypad. Breaks between subsequent blocks typically ranged from about 10 s to 30 s. The SPL of the tones was adjusted to 80 dB prior to the measurement, aiming for equal subjective levels of loudness in both ears. Two subjects required a moderate down regulation of the SPL delivered to the left ear to achieve such equality (75 dB SPL, subjects i and n). Subjects were instructed to refrain from movement during the measurement and to focus on a fixation cross displayed on a screen 1 m in front of them.

3.1.3 Data acquisition and pre-processing

A whole-head Elekta Neuromag TRIUX MEG system located in a magnetically shielded chamber was used to record data at a sampling rate of 1 kHz. The chamber was ventilated and contained a camera and loudspeakers to observe and communicate with the subject. An online filter was used during data acquisition (low-frequency cut-off: DC or 0.1 Hz; high-frequency cut-off of: 330 Hz) and the signal-space separation (SSS) method was applied to the data. Simultaneous measurements of horizontal and vertical electrooculograms were taken to identify eye movements and eye blinks.

For pre-processing, the raw data from all 102 magnetometers was loaded into the Brainstorm software (Tadel et al., 2011) and single trials containing eye blinks were rejected. For subjects with prominent regular heartbeat in the MEG signal, heartbeat artefacts were identified and corrected using the signal space projection (SSP) module included in Brainstorm. Additionally, other artefacts caused by muscle tension or technical incidents were identified via visual inspection. Subsequently, the continuous recordings from all channels were segmented into epochs ranging from -500 ms to 1500 ms relative to stimulus onset and epochs containing artefacts were discarded. Additionally, the non-adapted response to the first stimulus in each block was removed (Rosburg et al., 2010). Finally, the remaining epochs (80% to 90%, depending on SOI), were exported and band-pass filtered (high-pass 1 Hz, low-pass 30 Hz, both zero phase, Butterworth filter design of order 5) for further analysis within the Julia language (Bezanson et al., 2017).

3.1.4 Computation of SOI-specific ERFs

For each SOI respectively, ERFs were determined by computing the arithmetic average across artefact-free single trial epochs. Baseline correction was performed based on the 200-ms interval preceding stimulus onset but to verify our findings, we also ran our full analysis without baseline correction. ERFs with substantial amplitudes were only obtained for MEG

channels located over the left and right AC. Therefore, we limited our analysis to these two subsets of channels. For each subset, a *principal channel* was identified. This channel was characterised by showing, in a time window from 50 ms to 150 ms, the largest absolute N1m-peak amplitude across a majority of SOIs. For each subject, subsequent analysis steps were performed on data recorded by the principal channel above the left and right AC, respectively. As an example, Figures 33a and b show the SOI dependence of the principal-channel ERFs observed for the left and right hemisphere of subject a.

The information content of N1m peak amplitudes was assessed by computing signal-to-noise ratios (SNRs) for each subject, hemisphere, and SOI. SNR was defined as the ratio between N1m-peak amplitude and the standard deviation of the signal in the 200 ms prior to stimulus onset. We set the SNR-threshold to 1.5 after reviewing all ERFs and associated SNR-values. ERFs with an SNR below 1.5 and no clear N1m peak were excluded from further analysis. As a result, most ERFs for the 250-ms SOI were excluded and ultimately, we removed all of them from further analysis to guarantee a fair comparison across hemispheres and subjects. To keep the number of SOIs identical across hemispheres, we also excluded data from *both* hemispheres from further analysis when one hemisphere did not meet the SNR criterion for another SOI. Figures 33c and d map out the SOI-dependence of the N1m-peak amplitudes in the left and right hemisphere of subject a.

3.1.5 Bootstrapping

To gauge the robustness of the obtained ERFs and N1m-peak amplitudes, we used the non-parametric bootstrap technique to generate surrogate waveforms. Recall that the power of this technique lies in its ability to yield statistical inferences – such as confidence intervals (CIs) or medians – without postulating a normal distribution or homoscedasticity for the original data (c.f. Section 2.3.2, [A bootstrap-based expansion of the standard approach](#)). Separately for each subject, hemisphere, and SOI, single trials were drawn randomly and with replacement from the respective pool of artefact-free trials. The number of drawn trials was equal to the total number of trials in each pool and surrogate ERFs were computed by arithmetically averaging the data across each newly created set of samples. This procedure was repeated 999 times, thus resulting in 1000 ERFs (original + bootstrapped) per SOI for each hemisphere in each subject.

Figures 33e and f illustrate the approach described above for the 3-s SOI. Black curves represent the original ERFs, as seen in Figures 33a and b, and the black and white markers reflect the original N1m-peak amplitudes as seen in Figures 33c and d. Light grey curves in Figures 33e and f reflect the 999 bootstrap-based ERFs, with respective N1m-peak amplitudes marked in blue for the left and pink for the right hemisphere. The variability across the single-trial responses underlying the original ERF is reflected by the grey band mapped out

by the bootstrap-based ERFs. Thus, estimates of the CIs (e.g. the 95-% CI) associated with the original N1m-peak amplitude could be deduced from the scatter of the bootstrap-based amplitude values.

To determine CIs for RS lifetime estimates, we proceeded as follows: For each subject and hemisphere, a data set reflecting the SOI-dependence of the N1m peak amplitude was created by drawing one amplitude per SOI from the respective pool of 1000 N1m-peak amplitude values, randomly and without replacement. This was repeated until the pools were depleted, thus resulting in 1000 data sets. Equation (1) was fitted separately to each of these data sets, thus resulting in 1000 exponential curves per subject and hemisphere.

For fitting, we used the method of least squares (LsqFit.jl, Julia). The fitting procedure was two-tiered: First, an arbitrary starting value of $\tau_{\text{SOI}} = 0.10$ s was used to fit Equation (1) to the data sets for each subject and hemisphere. Next, the resulting τ_{SOI} -values were arithmetically averaged. The resulting grand mean ($\tau_{\text{SOI}} = 1.30$ s) was then used as the starting value for a second round of fits. The starting value for A_{sat} was always equal to the maximum peak amplitude in the respective data set. As in previous chapters, we again used $t_0 =$ stimulus duration (100 ms), reflecting the logic that once SOI equals stimulus duration and individual tones merge into a continuous tone, no evoked response peak should be detectable beyond the first stimulus. To verify fitting results, we also recomputed all fits with t_0 as a third fitting parameter with a lower bound and starting value of 100 ms.

Figures 33g and h summarise the final results of our analysis for the left and right hemisphere of subject a. Black and white markers, respectively, reflect the original peak amplitude data and black curves represent the associated fits of Equation (1). At each SOI, the vertical distribution of blue and pink markers reflects the N1m-peak amplitudes of the bootstrap-based ERFs. Grey curves represent fits of Equation (1) to the described sets of these amplitude values. The bootstrap-based fits map out a grey band that reflects the range of fitting parameter values resulting from the variability in the single-trial responses. Thus, CIs for the τ_{SOI} - and A_{sat} -values deduced from the original ERFs could be computed. The final response recovery function for each subject and hemisphere was defined by the median values of τ_{SOI} and A_{sat} across the bootstrap-based fits.

Left Hemisphere

Right Hemisphere

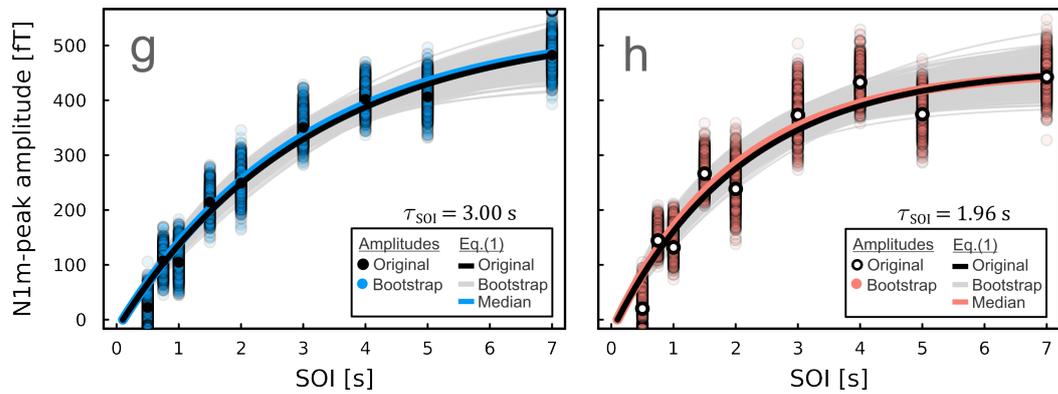
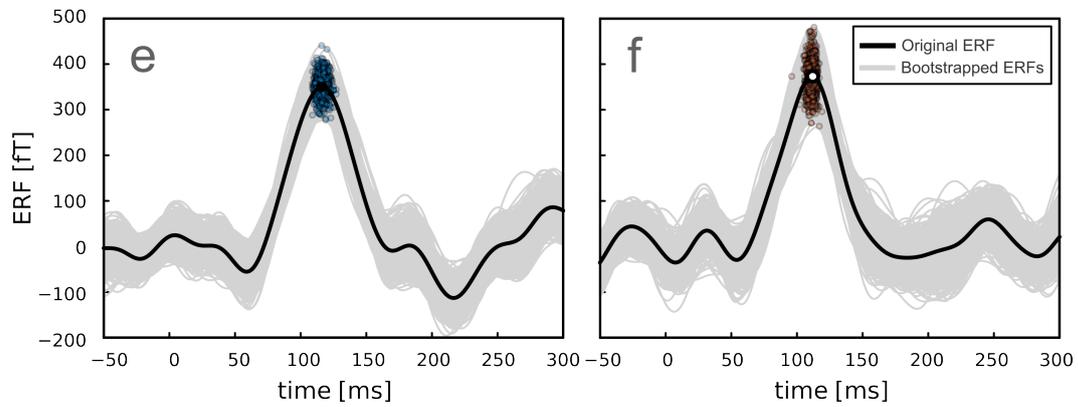
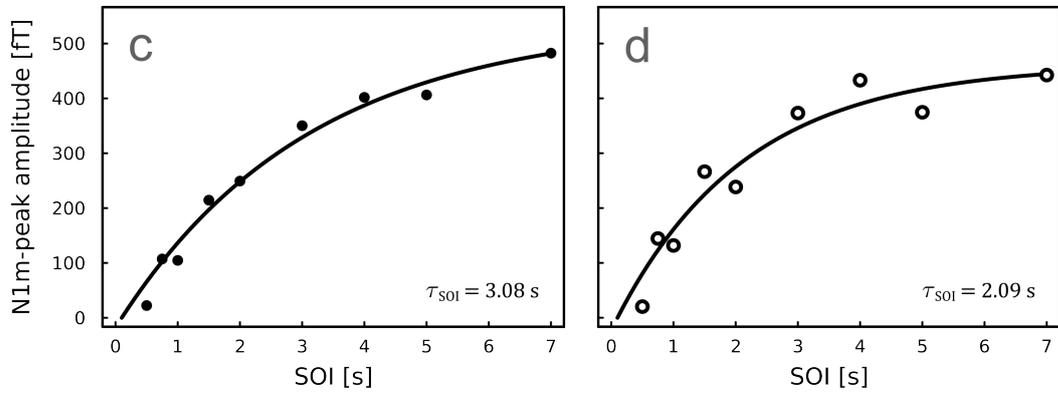
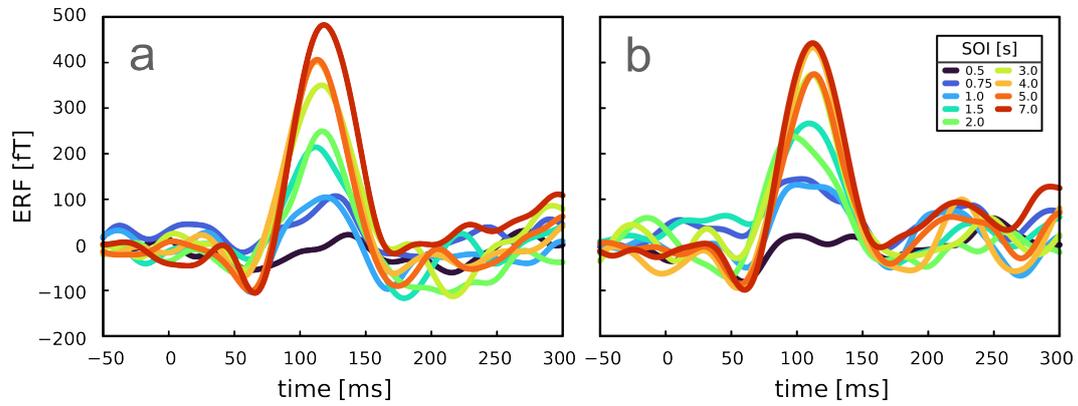


Figure 33: Procedure for determining RS lifetime τ_{SOI} using the example of a single subject. Panels on the left refer to the left hemisphere; those on the right to the right hemisphere. **a, b:** SOI-dependence of ERFs of this subject. The individual waveforms represent ERFs from SOIs between 0.5 s and 7 s, recorded by the subject- and hemisphere-specific principal MEG channel. With increasing SOI, the N1m-peak amplitude gradually increases and converges towards a saturation value at the largest SOIs. Since the magnetic field patterns of auditory ERFs have different polarities in the two hemispheres, the ERFs in **b** were multiplied by -1 for an easier comparison between the two hemispheres. Stimulus onset is at time $t = 0$ ms. **c, d:** Derivation of τ_{SOI} from the SOI-dependence of the N1m-peak amplitudes of the data shown in **a** and **b** for the left (black markers) and the right hemisphere (white markers). The curves are fits of Equation (1). **e, f:** Generation of surrogate ERFs using the bootstrap technique. ERFs for the 3-s SOI from each hemisphere are displayed as an example. The black curves represent the original ERF waveforms. The light grey waveforms represent the 999 bootstrap-based ERFs for each hemisphere, with the corresponding N1m-peak amplitudes displayed as blue (**e**) and pink (**f**) circles. **g, h:** Fits of Equation (1) to the SOI-dependence of N1m-peak amplitudes from the original (black and white circles, see **c** and **d**, respectively) and the resampled data (blue and pink circles, see **e** and **f**, respectively). The fits to the resampled data are shown as light grey curves. The blue (left hemisphere) and pink (right hemisphere) curves depict Equation (1) computed with the median values of τ_{SOI} and A_{sat} across all fits. Figure taken from and caption adapted from Dar & Härtwich et al. (2025).

3.2 Results

3.2.1 Repetition suppression lifetime and saturation amplitude across hemispheres and subjects

A comparison of median τ_{SOI} -values across hemispheres revealed longer RS lifetimes in the left hemisphere for 12 out of 14 subjects. This is illustrated in Figure 34a, where RS lifetimes for the left hemisphere are compared to RS lifetimes for the right hemisphere. The coloured markers reflect median values and the grey error bars map out the 95-% CIs associated with the respective medians. Medians of $\tau_{\text{SOI,LH}}$ covered a range from about 1.0 s to 3.0 s whereas medians of $\tau_{\text{SOI,RH}}$ ranged from about 0.4 s to 2.0 s. 95-% CIs extended from about 0.8 s to 4.9 s in the left and 0.2 s to 2.9 s in the right hemisphere. To quantify the relationship between 95-% CIs and respective medians, we computed the ratio $(q_{97.5}(x) - q_{2.5}(x))/m(x)$, where x represents one subject- and hemisphere-specific distribution of τ_{SOI} -values, $q(x)$ the respective quantiles, and $m(x)$ the median of the distribution. Overall, 95-%-CI ranges were large compared to the associated median τ_{SOI} -values, with ratios ranging from 0.34 to 1.21 across subjects and hemispheres.

We confirmed the significance of the difference between $m(\tau_{\text{SOI,LH}})$ and $m(\tau_{\text{SOI,RH}})$ via a non-parametric Wilcoxon signed rank test (estimated $z = 2.73$, p -value = 0.004). The effect size $r = z/\sqrt{n} = 0.73$ classified as large (see, for example Fritz et al., 2012). Given the substantial CIs for the τ_{SOI} -values, we verified the robustness of the finding $\tau_{\text{SOI,LH}} > \tau_{\text{SOI,RH}}$ by analysing the difference

$$\Delta\tau_{\text{SOI}} = \tau_{\text{SOI,LH}} - \tau_{\text{SOI,RH}} . \quad (29)$$

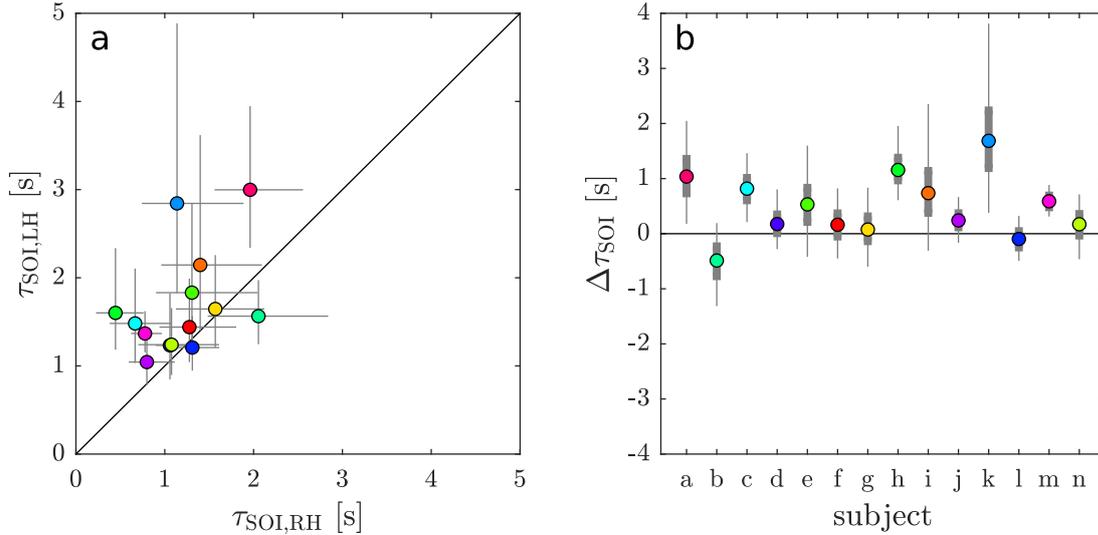


Figure 34: Subject- and hemisphere-specific results for RS lifetime τ_{SOI} . **a:** Comparison of τ_{SOI} in left and right hemisphere for all subjects of the cohort. Coloured symbols mark the median values for each subject, the corresponding error bars oriented along each axis display the 95-% CIs. The diagonal line marks equal lifetimes for the two hemispheres. A Wilcoxon signed rank test revealed a hemispheric difference in τ_{SOI} with significantly larger values for the left hemisphere (p -value = 0.004). **b:** Overview of $\Delta\tau_{\text{SOI}}$ -values across subjects. Coloured symbols (same colour code as in a) mark median values for each subject. The thicker grey error bars map out the interquartile range, and the thinner grey error bars the respective 95-% CIs. Median $\Delta\tau_{\text{SOI}}$ exceeds zero in 12 out of 14 subjects. Figure and caption adapted from Dar & Härtwich et al., 2025.

Separately for each subject, left and right hemisphere τ_{SOI} -values obtained via bootstrapping were paired randomly and without replacement to compute values of $\Delta\tau_{\text{SOI}}$. The resulting median values, interquartile ranges, and 95-% CIs of $\Delta\tau_{\text{SOI}}$ are shown in Figure 34b. In subjects where the median $\Delta\tau_{\text{SOI}}$ -value, $m(\Delta\tau_{\text{SOI}})$ was greater than zero (12 out of 14), the lower quartile, $q_1(\Delta\tau_{\text{SOI}})$ also exceeded zero in the majority of cases (8 out of 12, H1 ‘if $m(\Delta\tau_{\text{SOI}}) > 0$, $q_1(\Delta\tau_{\text{SOI}}) > 0$ ’ confirmed, $p = 0.021$). Moreover, 78% of all $\Delta\tau_{\text{SOI}}$ -values exceeded zero. These statistics confirm the robustness of our finding $\tau_{\text{SOI,LH}} > \tau_{\text{SOI,RH}}$ beyond the median values of $\Delta\tau_{\text{SOI}}$.

Across subjects, the grand median $M(\Delta\tau_{\text{SOI}})$ was equal to 0.39 s. This corresponded to 29% of the grand median $M(\tau_{\text{SOI}})$, which was equal to 1.34 s. The lower quartile of the full distribution of $\Delta\tau_{\text{SOI}}$ -values across subjects, $Q_1(\Delta\tau_{\text{SOI}})$, was also positive (0.16 s). The effect size of the hemispheric difference was thus not only large in terms of data ranks. The difference in terms of absolute values was considerable as well. As illustrated in Figure 34a, the distributions of the bootstrap-based τ_{SOI} -values computed for each subject were broad. We investigated the impact of outliers by removing data points more than 1.5 interquartile ranges (IQRs) away from the lower or upper quartile of each distribution (i.e. data point x_i from distribution x was an outlier if x_i did not lie within the range $[q_1(x) - 1.5\text{IQR}(x), q_3(x) + 1.5\text{IQR}(x)]$). The resulting changes in grand median as well as test statistics for medians and lower quartiles of $\Delta\tau_{\text{SOI}}$ were negligible.

Statistics	two-parameter fit		three-parameter fit	
	bc	no bc	bc	w/o bc
p -value $\Delta\tau_{\text{SOI}} \neq 0$	0.004	0.013	0.009	0.020
p -value $\Delta A_{\text{sat}} \neq 0$	0.426	0.391	0.463	0.426
effect size r $\Delta\tau_{\text{SOI}} \neq 0$	0.73	0.65	0.68	0.61

Table 12: Summary of p -values from the Wilcoxon signed rank test obtained for the difference between $\tau_{\text{SOI,LH}}$ and $\tau_{\text{SOI,RH}}$ and the difference between $A_{\text{sat,LH}}$ and $A_{\text{sat,RH}}$, along with effect sizes r for $\Delta\tau_{\text{SOI}} \neq 0$. Bold font highlights statistically significant results and the abbreviation ‘bc’ refers to baseline correction. Values as shown in Table 1 of Dar & Härtwich et al. (2025).

In summary, we were thus able to demonstrate that $\Delta\tau_{\text{SOI}} > 0$ is a robust finding despite the large CIs for $\tau_{\text{SOI,LH}}$ and $\tau_{\text{SOI,RH}}$. In contrast, the subject-specific difference between saturation amplitudes A_{sat} for the left and right hemisphere ($\Delta A_{\text{sat}} = A_{\text{sat,LH}} - A_{\text{sat,RH}}$) was not significant (Wilcoxon signed rank test, p -value = 0.426, for a detailed treatise, see Section 3.2 of Dar & Härtwich et al., 2025). The hemispheric difference in RS lifetime was thus not paired with a hemispheric difference in saturation amplitude. As described in [Materials and Methods](#), we repeated our analysis without baseline correction and with t_0 as a third fitting parameter to confirm our findings. Compared to the time scales observed for τ_{SOI} , deviations of the resulting t_0 -values from the previously fixed value of 100 ms were small. The median deviation across subjects and both the baseline and non-baseline corrected analysis condition was equal to 37 ms for the left hemisphere and 155 ms for the right hemisphere.

For each of the four analysis conditions (with and without baseline correction and with a two- or three-parameter fit, respectively), Table 12 summarises the p -values for the hypotheses that (1) the difference between $\tau_{\text{SOI,LH}}$ and $\tau_{\text{SOI,RH}}$ is systematic ($\Delta\tau_{\text{SOI}} \neq 0$) and (2) the difference between A_{LH} and A_{RH} is systematic ($\Delta A_{\text{sat}} \neq 0$). The hemispheric difference in τ_{SOI} was significant and the hemispheric difference in A_{sat} remained insignificant across all four conditions. This confirms our findings and suggests that, within subjects, τ_{SOI} does not systematically increase with A_{sat} .

3.2.2 Relationship between adaptation lifetime and saturation amplitude

We explored the relationship between τ_{SOI} and A_{sat} to determine whether these parameters co-vary across hemispheres and/or subjects. In Figure 35a, median τ_{SOI} -values are plotted as a function of the associated median A_{sat} -values. To investigate the correlation between these value pairs, we applied the non-parametric Kendall rank correlation test. It showed that correlation coefficients were small and that correlations were far from significant (Kendall’s tau-b = 0.14, p -value = 0.317). Correlations were also insignificant when separating the data

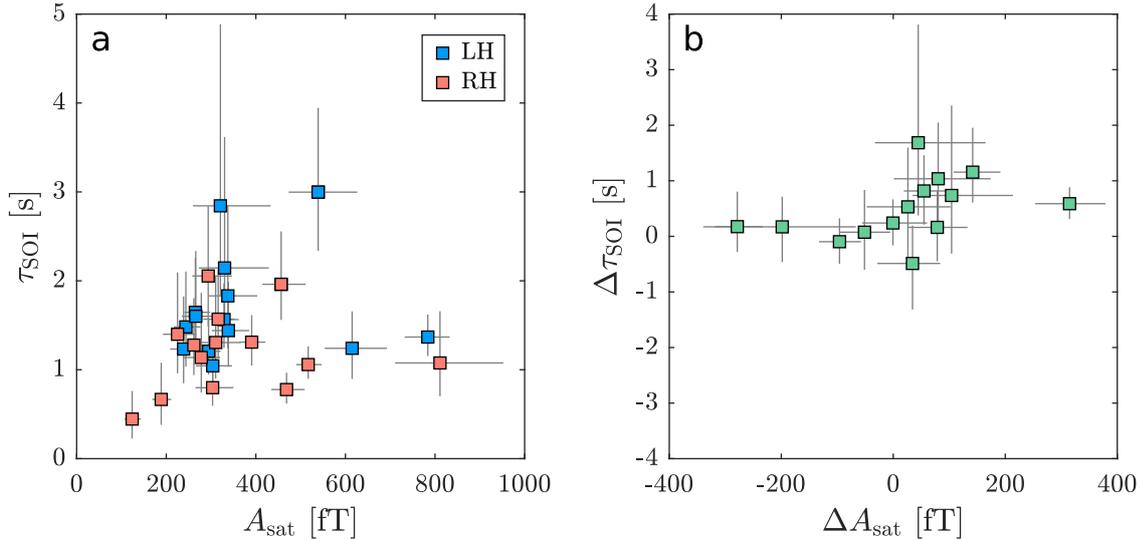


Figure 35: Relationship between RS lifetime and saturation amplitude. a: RS lifetime τ_{SOI} as a function of saturation amplitude A_{sat} . Blue and pink markers indicate the subject-specific median values for the left and right hemisphere, respectively. The error bars extend across the respective 95-% CI. There is no significant correlation between the median values for τ_{SOI} and A_{sat} (p -value = 0.317). b: Difference between RS lifetimes, $\Delta\tau_{\text{SOI}}$, as a function of the difference between saturation amplitudes, ΔA_{sat} . The green markers map out subject-specific median values and the error bars extend across the respective 95-% CI. The covariance of the median $\Delta\tau_{\text{SOI}}$ - and ΔA_{sat} -values is not statistically significant ($p = 0.062$). Figure and caption adapted from Dar & Härtwich et al. (2025).

into hemispheres (left: Kendall’s tau-b = 0.10, p -value = 0.667; right: Kendall’s tau-b = 0.14, p -value = 0.518).

To determine whether τ_{SOI} and A_{sat} co-vary across hemispheres, we assessed the relationship between $\Delta\tau_{\text{SOI}}$ and ΔA_{sat} . In Figure 35b, median values of $\Delta\tau_{\text{SOI}}$ are plotted as a function of median values of ΔA_{sat} , with error bars indicating the associated 95-% CIs. The correlation between the two measures was not significant (Kendall’s tau-b = 0.38, p -value = 0.062).

We also assessed the correlations listed above for the alternative analysis conditions previously introduced. These amounted to τ_{SOI} vs. A_{sat} for both hemispheres combined and for left and right hemisphere separately, as well as $\Delta\tau_{\text{SOI}}$ vs. ΔA_{sat} , i.e. four potential correlations \times two baseline correction options \times two fitting parameter options. The vast majority of correlations (14 out of 16) was insignificant (p -value > 0.05). Only the correlations between τ_{SOI} and A_{sat} in the right hemisphere, for a three-parameter fit *without baseline correction*, and the correlation between $\Delta\tau_{\text{SOI}}$ and ΔA_{sat} , for a three-parameter fit *with baseline correction*, had associated p -values minimally below the significance level. In summary, we thus confirmed that there is no covariation between τ_{SOI} and A_{sat} , neither across hemispheres nor across subjects.

3.2.3 Summary and outlook

We used binaural stimulation and a regular-SOI paradigm to determine RS lifetimes τ_{SOI} based on the N1m response of auditory ERFs. Our bootstrap-based analysis pipeline yielded subject- and hemisphere-specific statistical inferences and, to our knowledge, this study is the first robust demonstration of a hemispheric difference in the RS lifetime of the human AC. In the main Discussion of this thesis, I will elaborate on what might cause this difference. If it is functionally relevant, it might reflect hemispheric specialisation regarding the formation of sensory memory traces across distinct time windows. Whether the asymmetry is reduced, retained, or boosted when the auditory stimulation is task relevant and thus in the focus of selective attention remains to be explored. Moreover, given the correlation between τ_{SOI} -values and behaviourally measured lifetimes of sensory memory identified by [Lu et al. \(1992a\)](#), it needs to be established whether RS in the left or right AC better predicts behaviourally assessed lifetimes of auditory sensory memory.

Project 4: An automated approach to improve auditory cortex models

In Project 2, we have seen that the gerbil AC model was able to predict two major aspects of the general response behaviour of neural populations in gerbil A1, namely the decrease in peak latency as a function of SOI and the variation in RS lifetime as a function of stimulus audio-frequency (see Section 2.4). In order to further increase the predictive power of an AC model, not only validation but also further refinement of the model itself is required. Refinement can occur on two levels – (1) by improving the settings of the available parameters, and (2) by introducing new parameters, i.e. increasing model complexity. Here, we will focus on how to improve existing model parameter values in the context of an RS paradigm. So how could this be achieved?

Given the context of an RS paradigm, improvement refers to a better in-silico replication of the in-vivo response behaviour of the AC observed for said paradigm. For the gerbil AC, there was, for example, a mismatch between in-vivo and in-silico results for the absolute values of τ_{SOI} . This could potentially be addressed by changing parameter settings in the gerbil AC model. For the human AC, one very suitable and direct target is the detailed replication of the characteristic time course of the auditory response measured extracranially. But which model parameters should be adjusted and to what values?

Thus far, parameter settings in the May-et-al. modelling approach have been informed by a-priori knowledge about the structure and function of the AC. This knowledge only provides limited boundaries for the parameter values, which are ultimately selected on a trial-and-error basis. Such a manual approach is of course very time consuming and does not allow for a systematic search of the most suitable values within the full parameter space. Therefore, an automation of the parameter-value search is desirable.

In the project summarised in the following sections, we addressed this need and pioneered a computational approach to systematically optimise parameter values of an existing AC model. Further details of this work are reported and discussed in our recent publication (Tomana et al., 2023). The target of our optimisation were the weight values of the AC model and the optimisation criterion was the model’s ability to match a set of in-vivo auditory responses observed during an RS paradigm.

This chapter first introduces the AC model that we optimised (Section 4.1.1) and the experimental data set we used as a reference (Section 4.1.2). Next, an overview of the optimisation algorithm is provided (Section 4.1.3) and, finally, the new predictions about RS dynamics deduced from the optimised model are introduced (Section 4.2).

4.1 Materials and Methods

4.1.1 Human AC model and simulation of ERFs

The starting point for the development of our optimisation approach was the well-established human AC model developed by May et al. (2013; 2015; 2021). It can simulate extracranial recordings of auditory responses and there is thus a direct in-vivo counterpart that the in-silico signal can be compared to. The model is constructed out of the same building blocks as introduced for the gerbil AC model (columns comprising populations of excitatory and inhibitory neurons, c.f. Section 1.1.1) and the same dynamic equations govern the interactions between these network nodes (Equations (2), (3a), (3b), and (3c)). As reviewed by Moerel et al. (2014), the exact field pattern of the human AC remains elusive. Therefore, the closest available approximation was referenced: the well-studied structure of the macaque AC. The macaque AC consist of three core, eight belt, and two parabelt fields and their connection pattern is summarised in Figure 36a (for a review, see Romanski and Averbeck, 2009; Kaas and Hackett, 2000).

The AC connection pattern of the macaque maps out a set of parallel serial streams along the core-belt-parabelt hierarchy. Weight matrices of the human AC model by May et al. reflect this pattern. To reflect tonotopic organisation and produce audio-frequency-specific responses, the model consists of 16 columns per field (the same approach was adopted for the gerbil AC model, see Section 1.1.2 for a more detailed explanation). In the context of our optimisation approach, we decided to match the complexity of the AC model to the complexity of the employed RS paradigm, which only consisted of pure tones at a single audio-frequency (see Section 4.1.2 for further details). Thus, we minimised the parameter space that our optimisation algorithm had to address to one column per field. This also had the added benefit of decreasing simulation run times.

The remaining weight values were equal to those reported by May et al. (2015) and the resulting W_{ee} weight matrix is illustrated in Figure 36b. On the main diagonal (light purple), weight values were equal to 6 and all non-zero interfield-connections (green) had a weight of 0.5. Weights for W_{ie} and W_{ei} were non-zero on the leading diagonal, where values were equal to 3.5. Additionally, ie-connections from the three core areas (RT, R, A1) to thalamus (MGv), had a weight of 1. W_{ii} had no non-zero entries. Other model parameter values, such as the time constants, were also equal to those reported by May et al. (2015) and are summarised in Table 13.

Unlike the gerbil AC model, the human AC model also generates simulations of extracranial recordings of brain activity, namely data acquired via magnetoencephalography (MEG). The magnetic field detected via MEG is predominantly generated by current flow in the aligned apical dendrites of a large number of simultaneously activated pyramidal neurons in

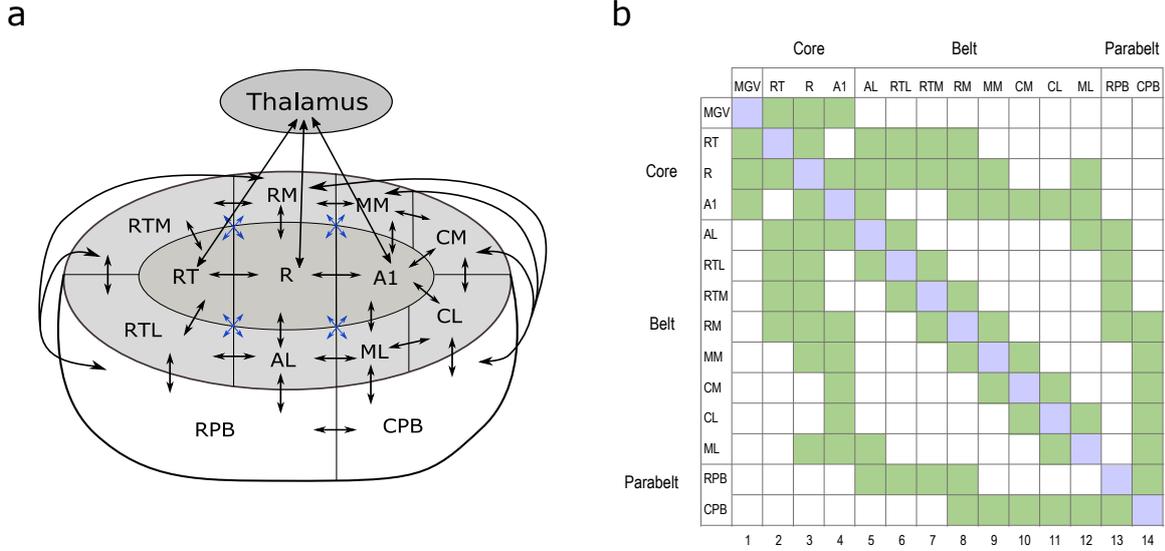


Figure 36: **a:** Schematic representation of the field structure and connectivity pattern of the macaque AC. The core area comprises three fields, the primary auditory field (A1), the rostral field (R), and the rostrotemporal field (RT), which receive parallel input from thalamus. The belt surrounds the core and consists of eight fields, the anterolateral (AL), the middle lateral (ML), the caudolateral (CL), the caudomedial (CM), the mediomedial (MM), the rostromedial (RM), the medial rostrotemporal (RTM), and the lateral rostrotemporal (RTL) field. The belt is bounded by two parabelt fields – the rostral parabelt (RPB) and the caudal parabelt (CPB). Arrows linking fields represent experimentally observed connections between fields (black: higher connection density, blue: lower connection density). Image file courtesy of Dr Aida Hajizadeh. **b:** W_{ee} matrix as used in the initial version of the simplified AC model; main diagonal (purple): $w_{i,j} = 6$; inter-field connections (green): $w_{i,j} = 0.5$.

the cortex (Hämäläinen et al., 1993). Therefore, the computation of in-silico MEG data is based on a scaled summation of the simulated synaptic currents induced via ee-connections (second term on the right-hand side of Equation (3a)):

$$R(t) = \sum_{i,j} k_{i,j} w_{i,j} q_j(t) g(u_j(t)), \quad (30)$$

where $R(t)$ is the time course of the in-silico MEG signal and i and j are column indices identifying the post- and presynaptic counterparts of each connection. The matching entries of the connection weight matrix W_{ee} are referenced by $w_{i,j}$, the factor $q_j(t)$ refers to the respective elements of the synaptic efficacy vector $\mathbf{q}(t)$, as defined in Equation (3c), and $g(u_j(t))$ is the presynaptic firing rate, as defined in Equation (2). For a given connection from j to i , the product of these three factors reflects the excitatory synaptic current induced at time t . The connection-specific factor $k_{i,j}$ scales the contribution of each synaptic current to the net in-silico MEG signal. Indices i and j range from 1 to N , the number of columns in the network, which was equal to 14 in our case.

As thalamic activation does not contribute to auditory ERFs, $k_{i,j} = 0$ for all connections targeting thalamus. Moreover, where $w_{i,j} = 0$, i.e. for connections that do not exist, $k_{i,j}$ is also set to zero. The magnitude of all non-zero k -values, i.e. for connections targeting the

description	symbol	value
firing rate threshold	θ	0.050
firing rate sensitivity	κ	$2/3$
membrane time constant	τ_m	0.030 s
time constant of synaptic depression onset	τ_{on}	thalamus: 0.020 s cortex: 0.100 s
time constant of recovery from synaptic depression	τ_{rec}	thalamus: 0.100 s cortex: 1.600 s

Table 13: Time constants and firing rate parameter values used in the simplified human AC model.

AC, was equal to 1. Their polarity was determined by connection type, specifically in terms of directionality along the core-belt-parabelt hierarchy. The apical dendrites of pyramidal neurons extend across several cortical layers and feedforward and feedback connections target these dendrites at different cortical depths. Feedforward connections predominantly form in the middle cortical layers, specifically layer IV. These are thus connections located close to the soma and synaptic activity results in dendritic current flowing away from the soma, towards the cortical surface. In contrast, feedback connections predominantly target the upper layers I and II. These are thus connections located further away from the soma and synaptic activity results in dendritic current flowing towards the soma, away from the cortical surface (see, for example, Ahlfors et al., 2015; Douglas and Martin, 2004; Kohl et al., 2022; Schroeder and Foxe, 2002).

To reflect these differences, $k_{i,j}$ was negative for feedforward connections and positive for feedback connections as well as intrafield connections (where $i = j$). Figure 37b shows an example of in-silico MEG signal time courses generated with the AC model described above. They were obtained by simulating the RS paradigm outlined in the following section (4.1.2).

4.1.2 Experimental data and paradigm

The data used to define an optimisation criterion for the human AC model was recorded during an auditory MEG study by Zacharias et al. (2012). It was approved by the Ethics Committee of the Otto von Guericke University in Magdeburg, Germany, and written informed consent was obtained from all participants. Subjects were exposed to passive listening paradigms to record auditory event-related fields (ERFs).

In the context of our optimisation project, we focussed on a data set recorded from a single subject who passively listened to five different stimulus blocks with SOIs of 0.5 s, 1.0 s, 2.5 s, and 5.0 s, respectively. The pure tone stimuli were repeated 111 times per block at an SPL of 80 dB. Each tone had an audio frequency of 1.5 kHz, a duration of 100 ms, and a linear

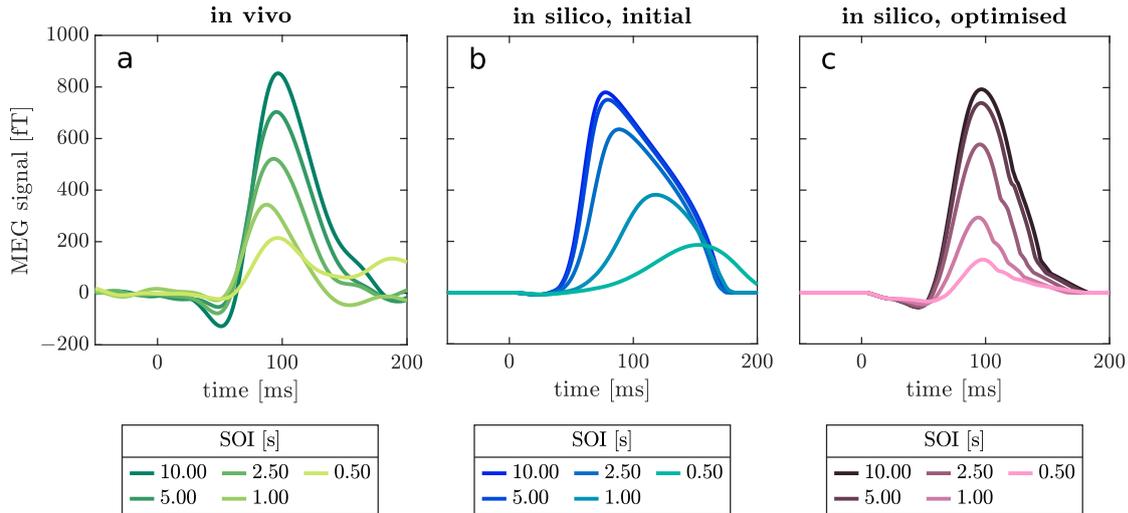


Figure 37: In-silico and in-vivo waveforms observed in response to the RS paradigm described in Section 4.1.2. **a:** MEG signals recorded from the left auditory cortex of a human subject. **b:** In-silico MEG signals produced by the initial version of the simplified human AC model. **c:** In-silico MEG signals produced by the optimised version of the simplified human AC model.

rise and fall time of 5 ms, respectively. For the MEG channel with the largest N1m signal above the left hemisphere, Figure 37a shows the artefact-corrected trial-averaged time course of the MEG signal for the five different SOIs. Repetition suppression, i.e. the characteristic decline in N1m- (and also P1m-) amplitude as a function of decreasing SOI is clearly visible. Moreover, N1m-peak latency increases with SOI and the rising slope towards the N1m peak is very homogeneous across SOIs.

The in-silico MEG data simulated for the paradigm described above successfully replicates the decrease in N1m-peak amplitude as a function of decreasing SOI. However, additional aspects of the response behaviour differ. N1m-peak latency increases with decreasing SOI, the P1m is hardly present, and the rising slope towards the N1m peak is SOI-specific. We attempted to address these discrepancies by optimising the weight values in our AC model.

4.1.3 Optimisation algorithm

We employed an advanced optimisation method called evolutionary algorithm (EA) to optimise weight values in our AC model. EAs explore the available parameter space in a manner that reflects mechanisms known from genetics and evolutionary theory, such as reproduction, mutation, recombination, and selection (Michalewicz, 1996). Due to their stochastic nature, EAs are less prone to getting stuck in local minima, compared to deterministic approaches like, for example, gradient based methods.

Initial population

The first step of running an EA is the creation of an initial population, made up of a number of specimens. In our case, specimens were versions of the W_{ee} and W_{ie} matrices of the human AC model. We created an initial population of 100 specimens with specimens being diversified by adding a random value from the uniform distribution $\mathcal{U}(-0.5, 0.5)$ to non-zero entries of W_{ee} and W_{ie} . Each specimen was represented by two ‘chromosomes’, one comprising the non-zero weight values of W_{ee} and one those of W_{ie} . The individual weight values represented the ‘genes’ that make up the respective chromosomes.

Fitness function

Next, the specimens were evaluated by assessing how well they were adapted to the ‘environment’. Here, the environment posed the problem of perfect replication of the experimental data introduced in Section 4.1.2. Hence, a specimen (versions of W_{ee} and W_{ie} in the human AC model) would achieve optimal adaptation to the environment if the simulation of the experimental paradigm described in Section 4.1.2 were to yield in-silico MEG waveforms that perfectly match the in-vivo data in the time window from -50 ms to 200 ms relative to stimulus onset.

To quantify the quality of the ‘solution’ that each specimen x achieves and thus rank them, a fitness function $\Phi(x)$ needs to be defined. We concatenated the in-vivo and in-silico MEG signals (c.f. Figure 37) into the respective vectors,

$$\text{MEG}^{\text{in-vivo}} = [\text{MEG}_1^{\text{in-vivo}}, \dots, \text{MEG}_S^{\text{in-vivo}}] \quad (31)$$

and

$$\text{MEG}^{\text{in-silico}} = [\text{MEG}_1^{\text{in-silico}}, \dots, \text{MEG}_S^{\text{in-silico}}], \quad (32)$$

with the subscript denoting the five different SOIs in ascending order (with $S = 5$), and quantified their alignment for specimen x via the following dot product:

$$\Phi(x) = \frac{\left\langle \text{MEG}^{\text{in-vivo}}, \frac{\text{MEG}^{\text{in-silico}}(x)}{\|\text{MEG}^{\text{in-silico}}(x)\|_2} \right\rangle}{\|\text{MEG}^{\text{in-vivo}}\|_2}. \quad (33)$$

Here, $\text{MEG}^{\text{in-silico}}(x)$ was ℓ^2 -normalised in order to focus on relative signal morphology rather than absolute amplitudes of the in-silico MEG signal. Moreover, since the maximum of the dot product in 33 is equal to

$$\left\langle \text{MEG}^{\text{in-vivo}}, \frac{\text{MEG}^{\text{in-vivo}}}{\|\text{MEG}^{\text{in-vivo}}\|_2} \right\rangle = \|\text{MEG}^{\text{in-vivo}}\|_2, \quad (34)$$

the division by the ℓ^2 -norm of MEG^{in-vivo} constrains the fitness function $\Phi(x)$ to the interval $[0, 1]$.

Selection and reproduction

When running EAs, ‘parent’ pairs are selected from the available pool of specimen in order to create ‘offspring’. This offspring possesses chromosomes that consists of a mixture of the genes from the two parents. Additionally, random mutations can be introduced. We found the ranking method (Arabas, 2001; Michalewicz, 1996; Rutkowska et al., 1997) to be a suitable choice for parent selection during our optimisation. In this method, a specimen’s probability of being selected for ‘reproduction’ depends on its rank, in terms of $\Phi(x)$, within the selection pool. The higher the rank, the greater the probability to be selected as a parent.

To create offspring, the genes in the chromosomes of the two parents (i.e., in our case, weight values for W_{ee} and W_{ie}) are recombined into new chromosomes. We used four types of arithmetic crossover to govern this reshuffle, meaning that all resulting offspring had chromosomes that were linear combinations of the respective parent chromosomes. Chromosomes of the offspring also underwent random mutations in the form of insertion, deletion, and inversion of genes (i.e. weight values). Since the chromosomes only contained weight values for connections present in the macaque AC, the associated connectivity pattern could never be contradicted by newly emerging additional connections. For full details on our approaches for selection, reproduction, and mutation, please refer to Section 4 of our publication (Tomana et al., 2023).

Succession and stopping criterion

After producing a pool of offspring, we merged it with the parent population and computed the fitness function value for each specimen. The 100 specimens with the highest fitness function values were then used as a new ‘generation’, a process called ‘succession’. Next, reproduction, mutation, and succession were performed again, thus initiating a looped procedure. Possible stopping criteria for such a procedure are (1) a predetermined number of iterations or (2) a fitness function value threshold. We opted for a total of 1000 iterations.

4.1.4 Shortening runtimes via a SOI-specific reduction of stimulus repetitions

Due to the large number of simulations performed during each of the 1000 iterations of a single EA run, it was important to minimise the runtime of the individual simulations. One important parameter that affected this runtime in the context of the RS paradigm was the number of stimuli used in each SOI block, as each additional stimulus repetition prolonged the simulation. Thus, ideally, responses should be identical from the second stimulus onward and

only two stimuli per block would be required to obtain a waveform reflecting stabilised RS. Note that the concept of RS stabilisation has been extensively explored in previous chapters, c.f. Sections 1.1.5 and 2.2.2. The same terminology is used in this section. We assessed the evolution of the in-silico response as a function of stimulus index i_{stim} in our initial AC model to determine the number of stimuli required to obtain a waveform reflecting stabilised RS. Response similarity \mathcal{S} was quantified akin to the alignment between in-vivo and in-silico MEG signals (c.f. Equation (33)), i.e. by computing normalised dot products between response pairs, with each response $R(i_{\text{stim}})$ covering a time window from -50 ms to 200 ms relative to stimulus onset:

$$\mathcal{S}(i_{\text{stim}}, k) = \frac{\left\langle R(i_{\text{stim}}), \frac{R(i_{\text{stim}+k})}{\|R(i_{\text{stim}+k})\|_2} \right\rangle}{\|R(i_{\text{stim}})\|_2}, \quad (35)$$

where k is an integer ranging from 1 to the number of stimuli presented in the stimulus block. Additionally, we compared N1m-peak amplitudes. Across 20 stimulus repetitions, we observed the following:

- For SOIs > 2.5 s, responses were fully stable from the second stimulus onward, as reflected by identical N1m-peak amplitudes (less than 0.01% difference) and identical waveform morphologies ($\mathcal{S} = 1.000$ for $i_{\text{stim}} \geq 2$).
- For the 2.5-s and 1-s SOI, N1m-peak amplitudes differed by no more than 0.05% from the third stimulus onward and responses were extremely similar in terms of waveform morphology ($\mathcal{S} > 0.999$ for $i_{\text{stim}} \geq 3$).
- For the 0.5-s SOI, N1m-peak amplitudes differed by no more than 5% from the fourth stimulus onward and responses were very similar in terms of waveform morphology ($\mathcal{S} > 0.997$ for $i_{\text{stim}} \geq 4$). Moreover, the N1m-peak of the average response across stimuli 4 to 6 and the N1m-peak of the average response across stimuli 4 to 20 differed by less than 1%.

Based on these results, we proceeded as follows: for the simulations performed during the optimisation, stimulus blocks for the 10-s and 5-s SOI consisted of two stimuli, stimulus blocks for the 2.5-s and 1-s SOI consisted of four stimuli, and the stimulus block for the 0.5-s SOI consisted of six stimuli. Evoked responses for the stabilised RS state, as depicted in Figure 37b, were reflected by the response to the second stimulus for SOIs > 2.5 s, by the average response across stimuli 3 and 4 for the 2.5-s and 1-s SOI, and the average response across stimuli 4 to 6 for the 0.5-s SOI. These waveforms were then concatenated in the vector $\text{MEG}^{\text{in-silico}}$ (Equation (32)) to evaluate the EA's fitness function (Equation (33)). The approach is summarised in Table 14.

SOI [s]	number of stimuli	MEG _[] ^{in-silico}	subscript []
0.5	6	average response across stimuli 4 to 6	1
1	4	average response across stimuli 3 and 4	2
2.5	4	average response across stimuli 3 and 4	3
5	2	response to stimulus 2	4
10	2	response to stimulus 2	5

Table 14: SOI-specific reduced number of stimuli per block and computation of the stabilised response to shorten the runtime of the simulations while properly capturing the RS phenomenon.

4.2 Results

EAs belong to the category of metaheuristics. Thus, running the full optimisation repeatedly, with the same starting point, does not lead to the same result. In order to explore the robustness of our results, we ran our full EA implementation a total of ten times.

To confirm the validity of the approach summarised in Table 14, we also assessed the evolution of the in-silico response as a function of stimulus index i_{stim} (across 20 stimuli) for the ten optimised AC models. As in the initial model, responses were always fully stable from the second stimulus onward for the 5-s and 10-s SOI ($\mathcal{S} = 1.000$, less than 0.01% difference in N1m-peak amplitude for $i_{stim} \geq 2$). For the 2.5-s and 1-s SOI, N1m-peak amplitudes differed by no more than 3% from the third stimulus onward and responses were still extremely similar in terms of waveform morphology ($\mathcal{S} > 0.999$ for $i_{stim} \geq 3$). For the 0.5-s SOI, there was one outlier (run 9) where response amplitudes showed a pronounced oscillatory pattern. For the remaining runs, N1m-peak amplitudes from the fourth stimulus onward differed by up to 15% (run 6) but less than 7% for the majority of EA runs (seven out of ten). Responses were still very similar in terms of waveform morphology ($\mathcal{S} > 0.995$ for $i_{stim} \geq 4$).

Thus, overall, our approach (summarised in Table 14) was an acceptable compromise between the runtime and the number of stimuli per block. However, the observations listed above also highlight the importance of the RS stabilisation analysis. As observed in the gerbil AC model (Section 1.1.5), changes in model connectivity can alter the RS stabilisation process and, therefore, too strong a reduction in the number of stimulus repetitions runs the risk of improperly capturing the state of stabilised RS.

4.2.1 Auditory cortex model optimisation

For each of the ten runs of the EA, the fitness function values $\Phi(x)$ of the best and worst specimen in a generation increased monotonically with the number of iterations. While no

two optimisation results were identical in terms of W_{ee} and W_{ie} (and in terms of the resulting waveforms), the algorithm always achieved a fitness function value ≥ 0.977 for the final best specimen (median: 0.981, maximum: 0.983). This is a drastic improvement given that the Φ -value for the initial model, i.e. for the waveforms depicted in Figure 37b, was equal to 0.870.

Figure 37c illustrates the in-silico MEG waveforms achieved with the optimised model for the best of the ten EA runs, i.e. the run achieving the largest Φ -value. Even without comparing the initial and final fitness function value, the pronounced improvements from b to c can easily be identified. Unlike for the initial model, where a strong delay of peak latency is observed with decreasing SOI, the in-silico waveforms produced by the optimised model show a mild delay in peak latency with increasing SOI. This is a much closer replication of the response behaviour observed experimentally. Likewise, the homogeneity of the rising slope towards the N1m-peak and the appearance of the P1m have improved. Thus, we were able to show that the response dynamics of the AC model can be modulated via the time-independent weight values for ee- and ie-connections (just as it was the case for ie-connections in the gerbil AC model, see Section 1.2).

Although optimised W_{ee} and W_{ie} were not identical across runs, we identified strong commonalities. These can be viewed as predictions regarding human AC connectivity and are discussed in detail in Sections 5 and 6 of our publication (Tomana et al., 2023). Here, we will focus on the optimised model's predictions regarding RS recovery dynamics along the core-belt-parabelt hierarchy.

4.2.2 Predictions regarding area-specific repetition suppression lifetimes

Equation (30), i.e. the net in-silico MEG signal, can be decomposed into AC-area-specific activity patterns. Columns located in the three areas have the following indices: $i_{\text{core}} = \{2, 3, 4\}$, $i_{\text{belt}} = \{5, \dots, 12\}$, and $i_{\text{parabelt}} = \{13, 14\}$ (c.f. Figure 36b). Thus, the area-specific components of the net in-silico MEG signal are:

$$R_{\text{core}}(t) = \sum_{i=2}^4 \sum_{j=1}^N k_{i,j} w_{i,j} q_j(t) g(u_j(t)), \quad (36a)$$

$$R_{\text{belt}}(t) = \sum_{i=5}^{12} \sum_{j=1}^N k_{i,j} w_{i,j} q_j(t) g(u_j(t)), \quad (36b)$$

$$R_{\text{parabelt}}(t) = \sum_{i=13}^{14} \sum_{j=1}^N k_{i,j} w_{i,j} q_j(t) g(u_j(t)). \quad (36c)$$

For the best out of the 10 runs of our EA, panels on the l.h.s. of Figure 38 illustrate the SOI-specific response behaviour of the full AC (Figure 38a, c.f. Equation (30)) as well as of the individual areas, i.e. the core (Figure 38c, c.f. Equation (36a)), the belt (Figure 38e, c.f.

area	median	τ_{SOI} [s]	
		Q_1	Q_3
AC	2.019	1.989	2.078
core	2.371	2.347	2.573
belt	1.857	1.723	1.913
parabelt	2.353	2.292	2.430

Table 15: Overview of median τ_{SOI} -values across the ten EA runs for the full AC as well as the core, the belt, and the parabelt, along with lower and upper quartiles (Q_1 and Q_3) of the respective distributions.

Equation (36b)) and the parabelt (Figure 38g, c.f. Equation (36c)).

Across all five SOIs, the largest contribution to the net ERF (orange curves) is made by the eight belt fields (pink curves). Respective response amplitudes clearly exceed those associated with the three core fields (purple curves). This observation held true for six out of the ten EA runs. For the remaining four runs, contributions of core and belt to the net ERF were roughly equal. The contribution of the two parabelt fields (green curves) to the net ERF was always negative (note the inverted y-axis in Figure 38d) and of smaller amplitude than those of both the core and the belt.

To compute net AC as well as area-specific RS lifetimes, we determined, for each SOI respectively, the peak response amplitude within the time window from 50 ms to 150 ms (relative to stimulus onset) for the corresponding in-silico signals, R , R_{core} , R_{belt} and R_{parabelt} . Next, for each run separately, we fitted Equation (1) to the data to determine RS lifetime τ_{SOI} , using $t_0 = 100$ ms. Starting points for A_{sat} were equal to the peak amplitude for the respective 10-s SOI and the starting point for τ_{SOI} was set to 1.000 s. The circular markers in the panels on the r.h.s. of Figure 38 map out N1m-peak amplitudes normalised relative to the respective A_{sat} -values as a function of SOI for all ten EA runs. The grey curves illustrate the fits of Equation (1) to the respective data sets.

Statistics regarding the fitting parameter τ_{SOI} are summarised in Table 15. Median τ_{SOI} -values for the core and parabelt exceed two seconds and are very similar ($\tau_{\text{SOI}} \approx 2.4$ s), whereas the belt recovers more quickly ($\tau_{\text{SOI}} \approx 1.9$ s). This observation also holds true at the level of the ten individual runs and at the level of the interquartile ranges (IQRs) of the respective τ_{SOI} -distributions. The IQRs for core and parabelt have a large overlap but do not intersect with the IQR of the belt. Thus, similar to the gerbil AC model, where RS lifetimes of individual A1 columns do not simply increase with response amplitude, the area with the largest response amplitudes is not equal to the area with the longest RS lifetime.

The collective response of the full AC exhibits an intermediate RS lifetime ($\tau_{\text{SOI}} \approx 2.0$ s). This is in excellent alignment with the in-vivo MEG response, where the AC of the subject was characterised by a τ_{SOI} -value of 2.006 s. In contrast, the initial model was characterised by a lower τ_{SOI} -value of 1.387 s.

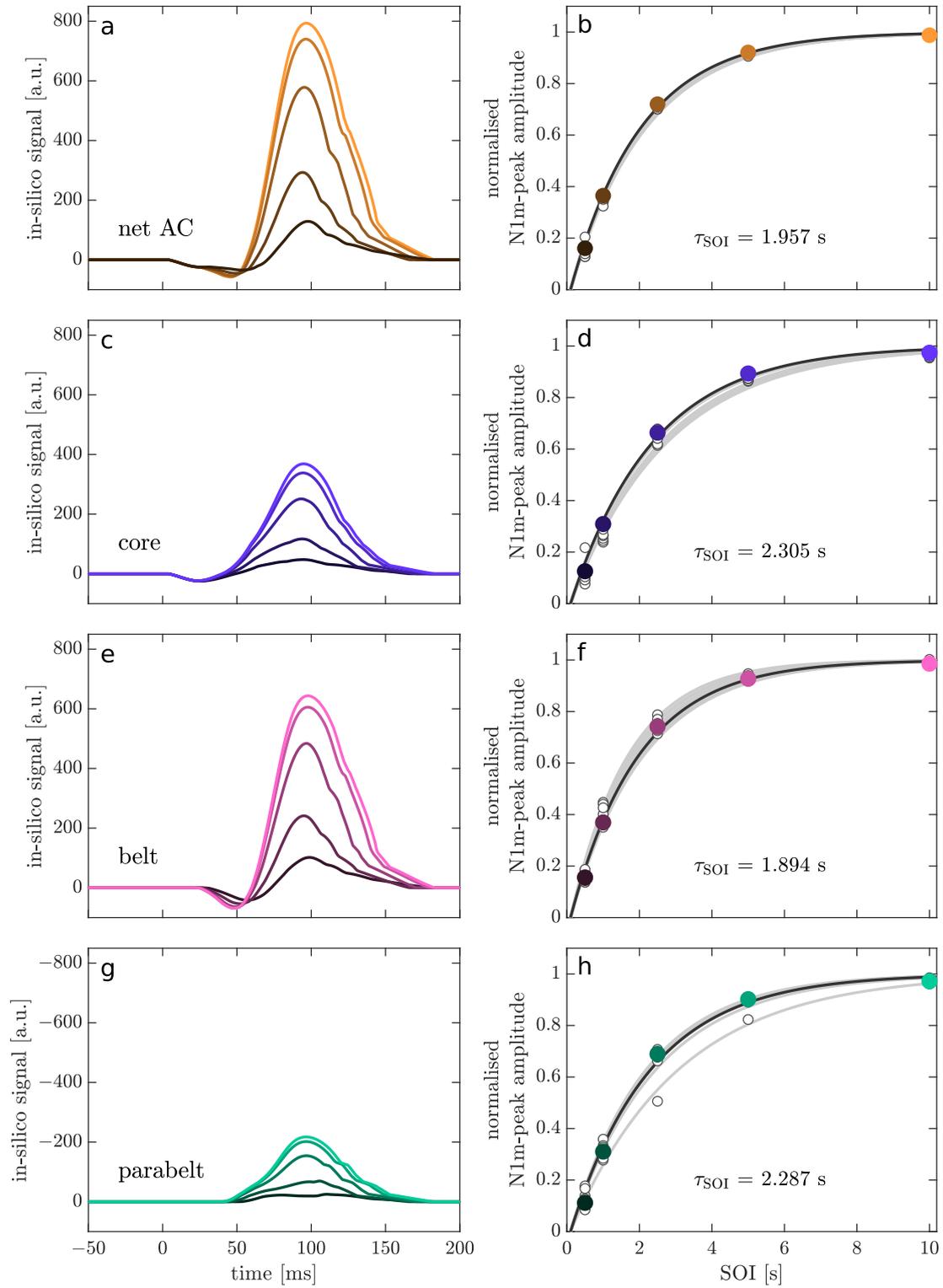


Figure 38: SOI-specific in-silico MEG signals (left) and associated N1m-peak amplitudes (right) of the full AC (a and b) as well as the core (c and d), the belt (e and f), and the parabelt area (g and h). Waveforms depicted on the left were simulated with the optimised model that achieved the highest fitness-function-value \mathcal{F} across the ten EA runs (‘best run’). The circular markers on the right map out normalised SOI-specific N1m-peak amplitudes for each of the ten runs, with amplitudes from the best run highlighted in the same colours as the adjacent waveforms. The underlying grey curves depict fits of Equation (1) to the data from each of the ten EA runs, with the fit for the best run highlighted in darker grey. Overall, the belt area has the shortest RS lifetime, with fits exhibiting the steepest rising slope. The τ_{SOI} -values listed in the panels on the right reflect the lifetime of RS deduced from the best run (i.e. the waveforms shown in the respective adjacent panels). For statistics on RS lifetime across the ten runs, see Table 15. For better comparability of the rising slope of Equation (1) across EA runs and areas, the N1m-peak amplitudes shown here were normalised with respect to the corresponding saturation amplitude A_{sat} . Note that such scaling does not affect τ_{SOI} .

We verified the robustness of the insights deduced from our fits by rerunning the fitting algorithm with t_0 as a third fitting parameter (starting point: 100 ms) and found that they were confirmed in that:

- The optimisation of the ee- and ie-connection weights prolongs RS lifetime, moving it closer to the in-vivo value
- RS lifetime is shorter in the belt than in the core and parabelt
- Core and parabelt have similar RS lifetimes

4.2.3 Preliminary conclusions

In the traditional view, adapting ERFs are deemed to arise from spatially discrete sources that are very localised. In contrast, our simulations predict that RS is a phenomenon that occurs across the entire AC. In our model, the mechanism that gave rise to this RS was STSD. Model parameter τ_{rec} , the time constant of recovery from STSD, was constant across all cortical areas. Yet the response behaviour of the model was area-specific, with the belt exhibiting a shorter RS lifetime than the core and parabelt. Time constant τ_{SOI} was thus not a direct reflection of model parameter τ_{rec} . Moreover, we showed that AC connectivity can affect RS recovery dynamics. While keeping the time constants of the model at fixed values, changes in connection weights alone were sufficient to alter the RS lifetime of the net AC by about 600 ms (from initial to optimised model). These results further support the prediction deduced from our simulations with the gerbil AC model: RS is a network effect resulting from the interplay of neural dynamics and connectivity patterns.

In addition to the insights gained from the response behaviour of the optimised model, the optimisation pipeline itself is also a result. We have shown that an evolutionary algorithm is an excellent tool to improve the parameter values of a computational AC model. The optimisation pipeline has the potential to be adapted for use with the gerbil AC model. In the Discussion (Part IV), we will elaborate on how this could be achieved.

Part IV: Discussion

Over the past four chapters, I have described a plethora of analysis steps and results, obtained in the context of four different projects. What ensues in this final part of the thesis is an interdisciplinary discussion and consolidation of our results and hypotheses. The projects were carried out at various levels of observation: in vivo vs. in silico, human vs. gerbil, extracranial vs. intracortical measurements. However, there was one overarching theme: the exploration of the phenomenon of RS in the AC and, more specifically, its lifetime.

Simulation results from the newly created computational model of the gerbil AC predict that the lifetime of RS does not simply reflect recovery from short-term synaptic depression (STSD). Instead, RS is a result of the interplay between STSD dynamics and network connectivity patterns. The same prediction also resulted from simulations with a simplified computational model of the human AC. When network connection strengths were altered, the lifetime of RS changed. Moreover, the gerbil AC model demonstrated that network interactions cause variations in RS lifetime as a function of stimulus audio-frequency. Electrophysiological recordings revealed audio-frequency-specific RS lifetimes in the AC of four out of six gerbils. Moreover, the RS lifetimes deduced from the intracortical recordings were shorter than lifetimes deduced from extracortical measurements of activity in the human AC.

In the first section of this discussion chapter, I review the challenges of RS lifetime quantification and discuss how we addressed them. In the subsequent section, I discuss what RS lifetime reflects. This is followed by a discussion of the role network-modulated RS might play as a memory trace facilitating the process of temporal binding. As we observed differences in RS lifetimes across species, I also discuss how these differences might arise and how they might manifest at the behavioural level. Finally, I discuss how the gerbil AC model could be refined and provide an outlook regarding future studies that could be carried out to further investigate the role network-modulated RS might play in temporal binding.

Q1: How can the challenges of repetition suppression lifetime quantification be addressed?

Paradigm selection

At the level of intracortical recordings, the literature predominantly assesses neural adaptation along midbrain and cortical stations of the auditory pathway in the context of the oddball paradigm (e.g. [Ulanovsky et al., 2004](#); [Pérez-González et al., 2005](#); [Malmierca et al., 2009](#);

Taaseh et al., 2011; Nieto-Diego and Malmierca, 2016). A sequence of stimuli, comprising two different tones each occurring with a certain probability, is presented to the subjects and the focus of the analysis lies on the average level of adaptation observed for each of the stimuli. Across sequences, the effect of temporal and spectral separation of the stimuli, as well as of their respective probability of occurrence, is investigated. As reviewed in the Introduction, responses generally adapt more strongly to the stimulus that is repeated more frequently, a phenomenon termed stimulus-specific adaptation (SSA).

RS is a fundamental aspect of the response behaviour observed for adaptation paradigms such as the oddball paradigm (Nelken, 2014; Pérez-González and Malmierca, 2014). While this paradigm creates a diverse stimulus history and does not permit the detailed study of fundamental RS dynamics, the regular-SOI paradigm creates a state of adaptation where the response stabilises across stimulus repetitions (recall, for example, the response stabilisation illustrated in Figure 27). The systematic variation of SOI across stimulus blocks permits the detailed quantification of an adaptation lifetime (τ_{SOI}) rather than just of the average level of adaptation for the ‘standard’ and ‘deviant’ stimulus. Insights on this lifetime can form the basis for the explanation of more complex phenomena, like SSA. To our knowledge, the results introduced in Chapter 2, along with the results we present in Ma & Brunk et al. (2021), are the first detailed reports of such lifetimes for individual neural populations in the AC. Note that other authors also report rates of decay or adaptation time constants based on intracranial measurements in the AC. These, however, are a reflection of the exponential decay or increase in response amplitude as a function of stimulus index (e.g. Deane et al., 2023; Ulanovsky et al., 2004) and should not be equated to the RS lifetimes reported here.

Noise sensitivity of in-vivo signals

Be it intracranial local field potentials or extracranial magnetic fields – in vivo, the noise sensitivity of the recorded brain signal poses a challenge to the quantification of RS lifetime. In the context of the RS paradigm, signal-to-noise ratios of evoked responses revealed via averaging can be very low unless a large number of single trials is averaged (see, for example, Figure 29a). Moreover, even if the number of trials is sufficient to stabilise the baseline prior to stimulus onset, questions regarding the robustness of response measures such as peak amplitude arise (Sielużycki et al., 2021). Given the large variability across the single trial data, has this measure stabilised or would it change if we were to increase the number of single trials?

For a given SOI, an increase in the number of single trials increases the measurement duration by the product of the number of additional trials and the SOI. Thus, increasing the number of trials is particularly problematic for longer SOIs, where it can quickly lead to excessive measurement durations. For extracranial recordings in awake subjects, this is a chal-

lence in terms of endurance and the increasing level of fatigue can distort the recorded data. For intracranial recordings in anaesthetised animals, maintaining a stable level of anaesthesia can be problematic and the changes in level of consciousness can distort the recorded data. Moreover, the viability of the animal is at stake.

To tackle the conflict between the need for shorter measurement durations and larger sample sizes, we developed a bootstrap-based analysis pipeline. The approach creates additional surrogate data sets based on the existing recordings by mimicking the repetition of these recordings via resampling. As a result, the number of evoked responses revealed by averaging single trials can be multiplied and the spread of these evoked responses is a measure of the CIs associated with the original evoked response. This in turn permits statistical inferences on RS lifetime, all without requiring a normal distribution and homoscedasticity for the original data. Consequently, as demonstrated in Projects 2 and 3, the significance of differences in RS lifetime observed across experimental conditions can be assessed at the single-subject level. Moreover, as demonstrated in Project 3, the robustness of results observed at the group level can be verified.

Selection of the fitting function

Traditionally, RS lifetime is quantified by fitting Equation (1), an exponentially saturating curve, to peak response amplitudes observed as a function of SOI (e.g. Lu et al., 1992a; Mäkelä et al., 1993; McEvoy et al., 1997; Sams et al., 1993). In doing so, the x-intercept t_0 can either be set to a constant value or used as a fitting parameter. For the experimental data and simulations presented here, we set t_0 equal to the stimulus duration. Lifetimes τ_{SOI} were thus always determined relative to the same reference point and could be compared directly across different conditions. Note, however, that a fixed value for t_0 is not always an appropriate choice. In our simulations, we have seen that some neural populations across the AC network activate at SOIs much longer than the stimulus duration (recall Table 9 and Figure 22). Their RS response behaviour would be more adequately summarised by a t_0 that is longer than the stimulus duration. Other simulations have revealed scenarios where t_0 is negative – for their simplified model of the human AC, Hajizadeh et al. (2022) report t_0 -values ranging from about -1.4 s to -0.4 s. In vivo, Rojas et al. (1999) used t_0 as a fitting parameter and, based on MEG data, report a significant difference in values for male and female subjects.

When t_0 is not fixed, comparisons of time constant τ_{SOI} across conditions (measurement locations, species, models, stimulus properties etc.) are less valid, because lifetime is not measured with respect to the same reference point. One approach to solve this issue could be to define RS lifetime as $t_0 + \tau_{\text{SOI}}$. In this case, however, RS lifetime no longer reflects the rising slope of the SOI-dependence of the peak amplitude. A scenario with a long t_0 and short τ_{SOI} would be reflected by the same RS lifetime as a scenario with a short t_0 and a long τ_{SOI} .

Equation (1) describes the relationship between peak amplitude and SOI via a single exponential function. It is important to note that such functions are very rare in nature (Lukichev, 2019). Hajizadeh et al. (2022) used computational modelling to investigate whether RS in the AC is the rare exception and concluded that whether Equation (1) is an appropriate description of the recovery process also depends on the number and range of sampling points. Alternatively, other fitting functions could be used, for example the stretched exponential Kohlrausch function, as recommended by Lukichev (2019). This might increase the goodness of fit, but at the cost of increasing the number of fitting parameters that need to be interpreted, both individually and in relation to each other.

Irrespective of the selected fitting function, the aim is to quantify the lifetime of RS. While the resulting values might be imperfect estimates of the time window neural populations require to recover from RS, these estimates are necessary to make comparisons of RS lifetime across conditions. In this work, Equation (1), along with a fixed t_0 , was the most parsimonious choice and provided an excellent foundation for the comparison of RS recovery dynamics across measurement locations, subjects, experimental conditions, and even species.

Q2: What does the lifetime of repetition suppression reflect?

As summarised by Gollisch and Herz (2004), potential mechanisms giving rise to RS and other forms of adaptation can be divided into two categories: (1) mechanisms that take effect on the output of neurons, such as voltage dependent conductances or hyperpolarisation, and (2) mechanisms that take effect on the input received by a neuron, such as STSD and inhibition. As reviewed in the Introduction (see ‘Mechanisms behind cortical RS’), in-vivo studies by Wehr and Zador (2003, 2005) and Ulanovsky et al. (2004) identify STSD as the most likely mechanism at the source of adaptation phenomena and in-silico studies further support this hypothesis (Loebel et al., 2007; Mill et al., 2011, 2012; David and Shamma, 2013; Yarden and Nelken, 2017; May and Tiitinen, 2013; May et al., 2015; May, 2021). Computational AC models with an implementation of STSD, but without mechanisms that take effect on the output of neurons, can replicate a multitude of adaptation-based phenomena.

STSD occurs due to a mixture of stimulus-evoked effects: the depletion of readily releasable synaptic vesicles, the inactivation of release sites, and the inactivation of presynaptic calcium channels (Fioravante and Regehr, 2011). The combined time courses of vesicle replenishment, reactivation of release sites, and reactivation of presynaptic calcium channels determine the lifetime of STSD. So is the lifetime of RS a direct reflection of the lifetime of STSD? Our intracortical measurements from gerbil A1 showed that τ_{SOI} for a given neural population can change between BF and non-BF stimuli, with respective lifetimes differing by up to 132 ms (c.f. Figure 32a). Postulating that STSD lifetime is the only factor that modulates RS lifetime, this could be explained by synapse-specific STSD lifetimes across the dendritic

tree and the activation of different synapses for BF vs. non-BF stimulation. The literature indeed reports that a wide range of time scales for STSD, spanning hundreds of milliseconds to seconds, exists for corticocortical synapses (Markram et al., 1998; Tsodyks and Markram, 1997; Varela et al., 1997). Moreover, we demonstrated in our simulations that, in the very simplified case of the single-column model, audio-frequency-specific RS lifetimes reflected the audio-frequency-specific amalgamation of the STSD recovery time constants at the afferent and recurrent synapse (Section 1.2.1).

Simulations with the gerbil AC model, however, have shown that STSD recovery dynamics cannot be directly related to RS recovery dynamics in this larger network (Section 1.2.2). Despite identical STSD lifetimes across all cortical connections, τ_{SOI} varied as a function of stimulus audio-frequency and always exceeded synaptic time constant τ_{rec} . A detailed analysis revealed that this is the result of large-scale network interactions. These in turn were affected by mild changes to the network’s connectivity pattern in the form of varying ranges of lateral inhibition. The effect of network connectivity on RS lifetime was further confirmed during our optimisation project (4), where we showed that changes in the connection weights of a simplified AC model based on macaque AC anatomy lead to changes in the ‘extracranially’ observed RS lifetime based on in-silico MEG signals.

We thus conclude that RS is a network effect. While STSD is the mechanism that gives rise to RS, the lifetime of RS reflects a plethora of interdependent network events. This hypothesis is supported by simulation results from Hajizadeh et al. (2022), who showed that changes to the serial core-belt-parabelt connection pattern of the AC affect the lifetime of RS. Furthermore, our observations are in line with results from Buonomano and Merzenich (1995). They showed that, in a network where neuronal properties are time-dependent but the model’s time constants are fixed and equal across units, the response behaviour of the individual units is highly heterogeneous. The in-silico SSA studies we reviewed in the Introduction (Insights from computational modelling) also indicate this (Loebel et al., 2007; Mill et al., 2011, 2012; Yarden and Nelken, 2017; May and Tiitinen, 2013; May et al., 2015; May, 2021).

An additional facet of the network effect that is RS could be the desynchronisation of synaptic events. In Ma & Brunk et al. (2021), we showed that, at the layer-specific level, RS is more pronounced for the measure reflecting peak current flow during sink activity than for the measure reflecting total current flow across sink activity. This also resulted in longer lifetimes of RS for the peak current flow than for the total current flow. Further analysis of the gerbil AC model responses could reveal whether this effect is also present in silico as the model is well suited for the detailed decomposition of the net response of a column into contributions from individual connections. We expect this to be a promising approach because, as demonstrated in Sections 2.4.2 and 2.4.3, the model made well-confirmed predictions regarding the timing of the net response peak as a function of SOI and stimulus audio-frequency. It could thus serve as a tool to further disentangle the different network events that contribute to RS.

RS lifetime measured extra-cortically via MEG might reflect an additional factor, namely cortical folding. As roughly expressed by the MEG equations in our simplified AC model (Equations 30, 36a, 36b, and 36c), the MEG signal is an amalgamation of individual neural responses. In vivo, the summation of these responses is not straightforward – differences in cortical folding can cause differences in the observed ERF (Shaw et al., 2013; Hajizadeh et al., 2021). Thus, two AC networks with the same connectivity patterns and STSD dynamics but different macroscopic anatomical properties might be characterised by two different MEG-based RS lifetimes.

It is well established that the folding of the AC differs across hemispheres (Heschl, 1878; Morosan et al., 2001; Rademacher et al., 2001; v. Economo and Horn, 1930) Thus, the different RS lifetimes we observed for the left and right human AC (Chapter 3) might mostly reflect hemisphere-specific cortical folding. However, they might also reflect hemisphere-specific time scales for the underlying neural mechanisms that give rise to RS and/or hemisphere specific AC connectivity that shapes RS dynamics. The three factors do not exclude each other and their role in hemisphere-specific RS lifetime needs to be explored further. If the different time scales of RS lifetime in the left vs. right hemisphere are indeed indicative of hemisphere-specific network dynamics, this difference might have a functional relevance and reflect a lateralisation in the processing of auditory information. This is discussed further in Dar & Härtwich et al. (2025).

Q3: What is the functional relevance of repetition suppression as a network effect?

We expect that the variation in the lifetime of RS as a function of stimulus audio-frequency lays the foundation for the process of temporal binding, for example the ability of individual neural populations or neurons in the AC to exhibit combination sensitivity (CS) – a selective response to a specific stimulus sequence but not its individual components or the reversed sequence. Experimentally, CS neurons have been observed in the primary auditory cortex or comparable areas of multiple species such as songbirds (Margoliash and Fortune, 1992; Lewicki and Arthur, 1996; Doupe, 1997), rats (Kilgard and Merzenich, 1999; Orduña et al., 2001), cats (McKenna et al., 1989; Brosch and Schreiner, 2000), marmosets (Bartlett and Wang, 2005; Sadagopan and Wang, 2009), and macaques (Brosch et al., 1999). However, it is unclear how this property arises.

Combination sensitivity in computational modelling

Several simulation studies have shown that a neural network where connection strengths are modulated by short-term synaptic plasticity and, if applicable, afferent input is spectrally

filtered, is a structure that can perform temporal binding and generate network nodes that exhibit forms of CS observed in vivo. [David and Shamma \(2013\)](#) demonstrated that the response behaviour of their model of an A1 neuron receiving input via a set of STSD-affected synapses (for a further review of the modelling approach, refer back to page 7) encodes stimulus envelope history for a time window of up to 290 ms .

[Lee and Buonomano \(2012\)](#) created a network model reflecting a simplified cortical column and demonstrated that this network starts to form vocalisation-sensitive neurons during an unsupervised training scheme. The model consisted of 400 excitatory and 100 inhibitory single-compartment, conductance-based leaky integrate-and-fire neurons as well as 18 input fibers, each representing an audio-frequency band. Time constants of short-term synaptic depression and facilitation (STSF) were constant across excitatory and inhibitory connections, respectively. Neurons were divided into a population reflecting cortical layer IV and a population reflecting cortical layers II and III. Layer IV neurons received tonotopically organised afferent input whereas layer II/III neurons were only driven by the propagation of activity within the network. The network was initialised with random connectivity. During the unsupervised training scheme, vocalisations – spoken digits from zero to nine – were presented and connection weights were altered. Afterwards, a subset of neurons exhibited sensitivity for one of the digits. Lee and Buonomano demonstrated that this selectivity hinges on synaptic plasticity. When removing STSD and STSF from the model, direction selectivity, e.g. a strong response to a specific vocalisation but not its time-reversed counterpart, decreased significantly.

[Goudar and Buonomano \(2015\)](#) also showed that CS neurons arise naturally from a large neural network reflecting properties of A1, including connections affected by STSD and STSF. Model parameters defining short-term plasticity were connection-type-specific but constant across cells. The model comprised 800 excitatory and 200 inhibitory single-compartment, conductance-based leaky integrate-and-fire neurons, the input layer as well as its projections were tonotopically organised, and recurrent connections in A1 were randomised but most probable for cells in close proximity. Without specific training, responses to paired tones were very diverse and included neurons exhibiting CS.

May and colleagues ([2013](#); [2015](#)) demonstrated that STSD, a serial network structure, and tonotopically organised afferent input suffice for columns in their AC model to exhibit CS to tone pairs, tone sequences, and short vocalisations. As previously reviewed (page 8), units in this modelling approach reflect columns comprising an excitatory and an inhibitory neural population and network connectivity mirrors the field structure along the core-belt-parabelt hierarchy of the macaque AC. The lifetime of STSD is constant across the network and the fine structure of the connectivity pattern is randomised. May and colleagues demonstrated that columns exhibiting CS appear without specific training or fine tuning of the model. The percentage of CS columns in the model changed as a function of several variables: the STSD

recovery time constant τ_{rec} , the balance between excitation and inhibition, the connection density in the network, the level of spectral selectivity in the network, and the duration and composition of the stimulus sequence. Moreover, temporal binding of shorter sequences mostly occurred in the core area whereas CS for longer sequences was observed higher up the serial structure, in the belt and parabelt fields. When serial processing was removed by targeting afferent input to all fields, CS almost vanished.

Repetition suppression as a tool for temporal binding

The studies reviewed above demonstrate that CS automatically emerges in neural network models of the AC when connections are affected by synaptic plasticity and the fine structure of the connection pattern is not purely tonotopic. Moreover, the authors delineate the boundary conditions of the parameter space within which CS can be observed. Our results regarding network-modulated, audio-frequency-specific RS offer first building blocks towards an explanation of how exactly such CS arises. Figure 39 conceptually depicts an example scenario where five columns, responding as observed in the gerbil AC model, are connected to an additional column. This column exhibits CS for a distinct stimulus sequence due to the selected connection pattern.

Figure 39a summarises the response behaviour observed for A1 columns in the gerbil model. The x-axis lists intrafield column indices and the horizontal colour gradient below maps out the associated BF of each column. The green, yellow, and red bar charts illustrate how, in terms of response amplitude and RS lifetime, the columns respond to stimulus blocks of pure tones A, B, or C. Tones A, B, and C are the BF of columns 4, 8, and 12, respectively. Taller bars in the bar charts simultaneously indicate larger response amplitudes and shorter RS lifetimes. Note that the distribution of RS lifetimes across columns is the same u-shaped distribution as shown in Figures 10a to c, it was just inverted to simultaneously depict RS lifetimes (vertical axis on the l.h.s.) and response amplitudes (vertical axis on the r.h.s.). The depicted response amplitudes reflect FR functions as shown in Figure 19.

Figure 39b depicts a stimulus sequence comprising the pure tones A, B, and C. Its repetitive pattern can coarsely be likened to that of naturalistic vocalisations observed for the gerbil (Kobayasi and Riquimaroux, 2012). Figures 39c and d illustrate how the CS column responds when the last segment of the sequence is presented in isolation and as part of the sequence, respectively. Due to the connection pattern between the CS column and the other A1 columns, the firing threshold of the CS column is only reached when the full sequence is presented, whereas, due to audio-frequency-specific RS, presenting individual components of the sequence does not trigger a response. The assumption that the CS column requires parallel excitatory input from several columns to cross its firing threshold is in line with in-vivo observations from Sadagopan and Wang (2009), who reported sparse spontaneous firing for

CS neurons and concluded that this suggests a high firing threshold.

Note that Figure 39 is just one simple example to illustrate the functional potential of audio-frequency-specific RS – scenarios with a much higher level of complexity could also be thought up. At the same time, audio-frequency-specific RS is probably only one factor contributing to temporal binding in the AC. A first step towards exploring the hypothesis that audio-frequency-specific RS facilitates CS would be to implement and verify the scenario sketched out in Figure 39. Moreover, CS columns in the May et al. model as presented in 2013 and 2015 should be investigated further. First, RS lifetimes of pure-tone driven, spectrally selective columns in the network could be catalogued via the regular-SOI paradigm. Next, the input received by the previously identified CS columns could be decomposed and potentially traced back to audio-frequency-specific RS at a set of columns responsive to pure tones. In this context, it would also be interesting to classify the position of the pure-tone columns along the core-belt-parabelt hierarchy and relative to the CS column.

In vivo, the coexistence of pure-tone-driven neurons and sequence-driven neurons in A1 has been well described (e.g. Sadagopan and Wang, 2009) and, in the future, methods like optogenetic techniques (for a review, see Emiliani et al., 2022) or juxtacellular nanostimulation (Stüttgen et al., 2017) might permit the artificial alteration of the response pattern of pure-tone-driven, adapting neurons while simultaneously recording the effect this intervention might have on previously identified CS neurons. Sadagopan and Wang (2009) noted that the precision of the CS neurons they observed aligned with the audio-frequency tuning bandwidth of tone-tuned neurons and this could thus be a promising line of investigation. It could also shed light on the robustness of the neural code and the dependence of CS on input from different AC fields.

Our hypothesis suggests that audio-frequency specific RS operates as a stimulus-history-dependent network activity ‘gating mechanism’ that affects the response to future stimuli. This interpretation can be embedded in the framework of ‘active’ and ‘hidden’ internal states developed by Buonomano and colleagues (for a review, see Buonomano and Maass, 2009; Motanis et al., 2018). The active state reflects the network’s measurable response behaviour (in terms of firing rates etc.), the hidden state refers to time- and stimulus-history-dependent neuronal properties (e.g. current levels of STSD and STSF).

Buonomano and colleagues (2009) illustrate the relationship between active and internal state with the ‘pebble thrown into a pond’ analogy: "*A pebble thrown into a pond will create a spatiotemporal pattern of ripples, and the pattern produced by any subsequent pebbles will be a complex nonlinear function of the interaction of the stimulus (the pebble) with the internal state of the liquid (the pattern of ripples when the pebble makes contact).*" Thus, the hidden state serves a crucial memory function. In this context, it is important to note that the computational studies reviewed above as well as the work presented in this thesis have shown that complex response behaviour does not require a high diversity in STSD and STSF dynamics

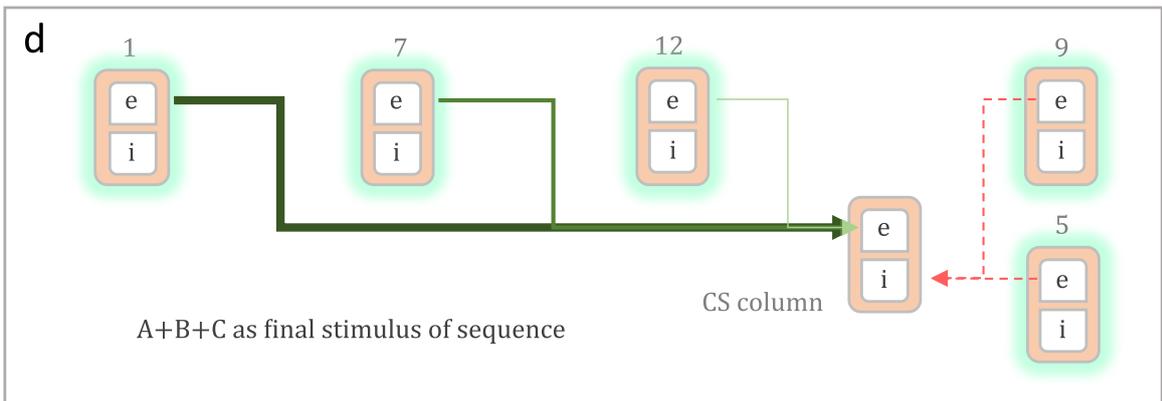
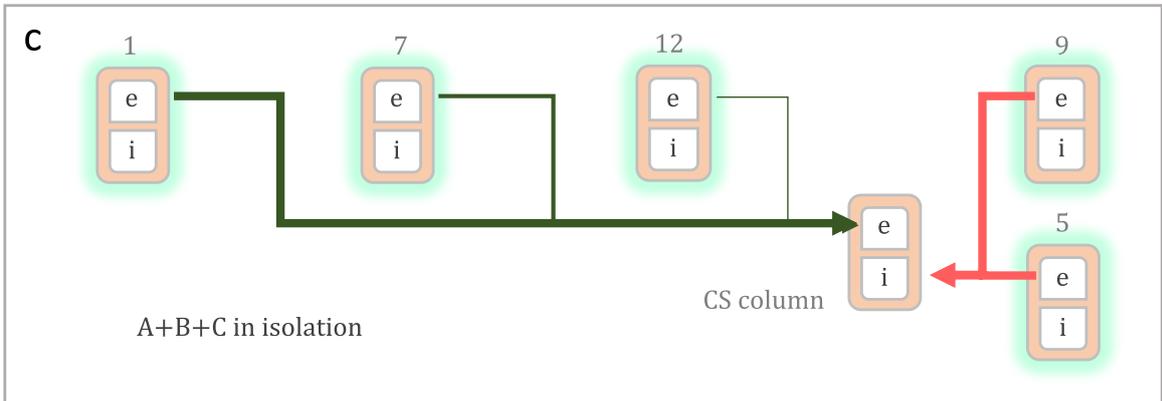
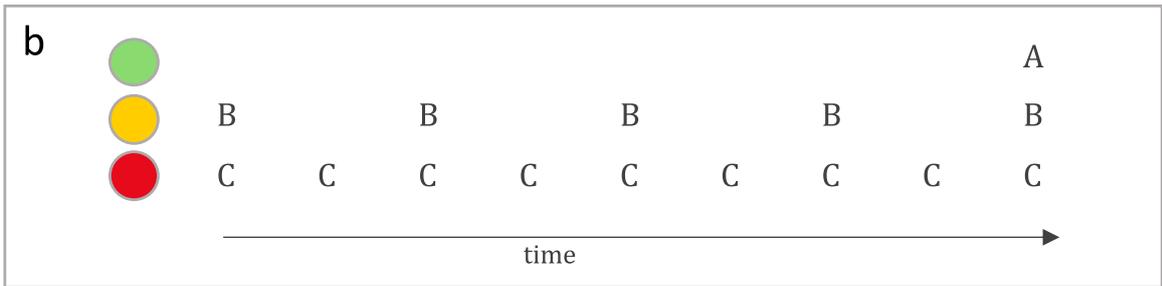
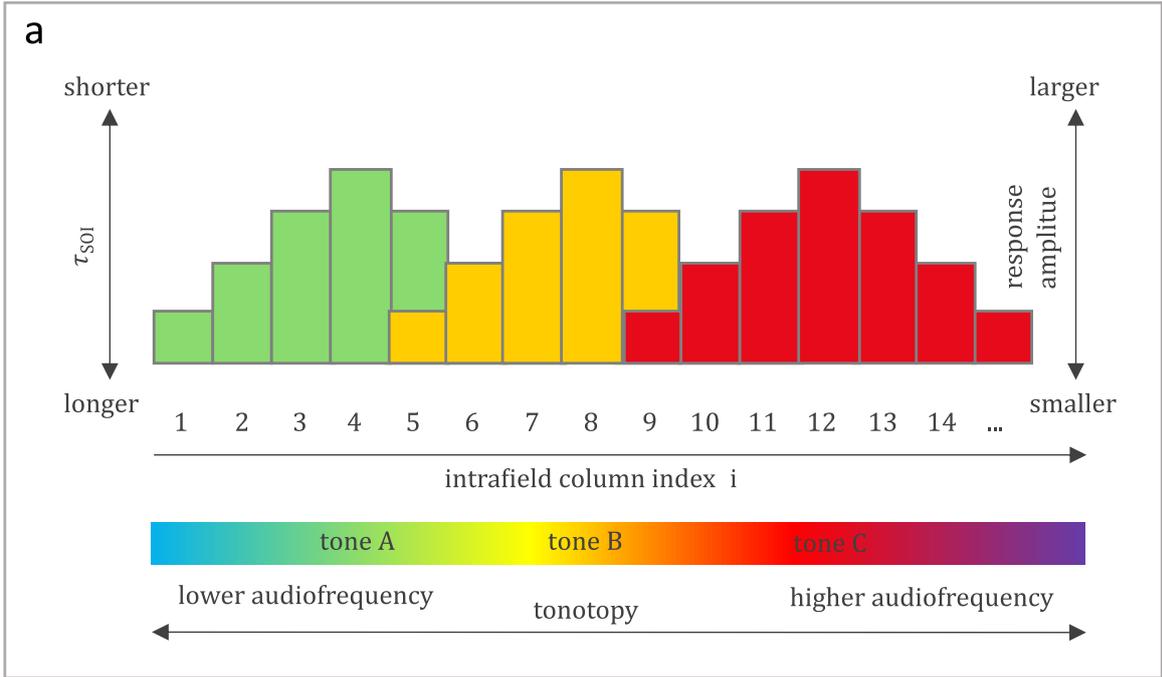


Figure 39: Audio-frequency-specific RS lifetimes as a basis for CS.

a: Response behaviour of A1 columns in the gerbil AC model (c.f. Section 1.2.2) when pure tones A (green), B (yellow), or C (red) are presented repeatedly. The horizontal axis maps out intrafield column indices and associated BFs. The vertical axes map out response amplitude (right) and RS lifetime τ_{SOI} (left). RS lifetime decreases in the upward direction.

b: Combination of tones into a sequence. The audio-frequency of tone A is the BF of Col.4. Neighbouring columns also respond but response amplitude decreases with distance from the BF column. In contrast, RS lifetime increases with this distance. Col.1 responds mildly to tone A but, due to the long RS lifetime, only when SOI is long – as is the case in the stimulus sequence. Tone B has an intermediate audio-frequency, the BF of Col.8. As for A, response amplitudes decrease and RS lifetimes increase with distance from the BF column. Col.7 responds fairly strongly to tone B, even at the intermediate SOIs present in the stimulus sequence, because RS lifetime is intermediate. In contrast, Col.5 has a longer RS lifetime for tone B and its response to it is mostly suppressed by the end of the sequence. Tone C has a higher audio-frequency, the BF of Col.12 – it responds strongly to the tone, even at the short SOIs present in the sequence (due to the short RS lifetime). Col.9 has a long RS lifetime for tone C and its mild response is suppressed at the short SOI.

c: Connection pattern for additional column that would exhibit CS for the stimulus sequence shown in b. Col.1 responds mildly to A and forms a strong ee-connection with the CS column. Col.7 responds fairly strongly to B and forms an ee-connection of intermediate strength. Col.12 responds strongly to C and forms a weak ee-connection. Col.5 responds mildly to B but forms a strong ie-connection. Col.9 responds mildly to C but forms a strong ie-connection. The joint presentation of A, B, and C with no prior stimulus history activates all five columns but inhibition dominates excitation and the CS column does not respond.

d: If A, B, and C are presented as the final stimulus of the sequence in b, the inhibitory effect of Col.5 and 9 is strongly decreased due to long-lived RS. The CS column responds. It does not respond to the previous B+C or just C because the connection weights are adjusted such that the strong contribution from Col.1 is required to activate the excitatory population. Therefore, an inversion of the sequence also does not lead to a response. Moreover, ‘just A’ is not sufficient because excitation from column 1 is counteracted by inhibition from column 5. A further decrease in the SOI for C makes the contribution of Col.12 insufficient, a further decrease in the SOI for B the contribution of Col.7 and a decrease in A the contribution of Col.1. Increasing SOI for B and C reactivates inhibition from Col.5 and 9. Thus, the CS column performs temporal binding based on audio-frequency specific RS.

but can arise through complexity in connectivity patterns and be modulated by changes in these patterns.

We propose that audio-frequency-specific RS lifetimes, produced by the network pattern of the AC, play a role in temporal binding. This in turn suggests that the upper limits of RS lifetime observed in a given species might be a reflection of the species-specific maximum time window across which the AC can perform such binding. We address this topic in the following section.

Q4: Is the lifetime of repetition suppression species-specific?

Interestingly, the τ_{SOI} -values we determined for neural populations in gerbil A1 are much shorter than the more global RS lifetimes determined for the human AC based on MEG data. In our animal study, the grand median τ_{SOI} -value (across animals and conditions) in gerbil A1 was equal to 0.357 s, with the full range spanning from 0.197 s to 0.465 s (c.f. Section 2.4.1). In contrast, in our human MEG study, the AC was characterised by a grand median τ_{SOI} -value (across subjects and hemispheres) of 1.339 s, with subject- and hemisphere-specific values ranging from 0.455 s to 3.998 s (c.f. Section 3.1).

The MEG-based values Zacharias et al. (2012) report for the human AC lie in the same range: 2.8 s for the left and 2.2 s for the right hemisphere (mean τ_{SOI} -value across 15 subjects, same stimulus duration and rise and fall times as in our RS paradigm). These results are also supported by older MEG-studies (Mäkelä et al., 1993; McEvoy et al., 1997; Sams et al., 1993), where τ -values ranged from about 0.9 s to 3 s.

For humans, the time scale of RS lifetime, based on auditory ERFs, is thus very well established and clearly exceeds the range we observed for neural populations in gerbil A1. Whether this time scale difference arises due to the different levels of observation or whether it is a true reflection of species-specific RS recovery dynamics remains to be explored. Several factors differentiate the levels of observation at which RS was studied in the gerbil and human AC. In total, there are four potential reasons that could explain the different results:

- *Measurement method*: intracranial recordings of local field potentials vs. extracranial recordings of magnetic fields
- *Mental state*: anaesthetised vs. awake subjects
- *Source of the measured response*: one neural population in A1 vs. all neural populations across the entire AC together
- *Species-specificity*: gerbil vs. human

Differences introduced by the measurement method

Relating intracranially measured LFPs reflecting neural activity to extracranially measured magnetic fields reflecting neural activity poses a challenging forward/inverse problem (for a review, see for example [Baillet, 2013](#)). Moreover, as discussed in the final paragraphs of ‘Q2: What does RS lifetime reflect?’, cortical folding affects the MEG signal. Thus, in the most extreme case, the extracranially measured response behaviour of the human AC could be characterised by longer τ_{SOI} -values even if the intracranially measured response behaviour of individual neural populations in the human AC were to perfectly align with results from gerbil A1. Simulations suggests that, in the context of the regular-SOI paradigm, the superposition of neural activity into ERFs is SOI-specific and depends on cortical folding ([Hajizadeh et al., 2022](#)). The differences we observed could thus be a reflection of the species-specific measurement method rather than species-specific RS dynamics.

Differences introduced by anaesthesia

[Deane et al. \(2020\)](#) found that the amplification of thalamocortical input in gerbil A1 via recurrent excitation is significantly increased under ketamine anaesthesia – the type of anaesthesia also used in our gerbil experiments. They concluded that this effect is most likely caused by an anaesthesia-induced *decrease* in inhibition. In [Ma & Brunk et al. \(2021\)](#), we demonstrated that pharmacological intervention *increasing* the level of inhibition and thus minimising recurrent excitation leads to longer RS lifetimes. Thus, the ketamine-induced increase in recurrent excitation might lead to faster recovery from RS in anaesthetised than in awake animals. While no literature on RS lifetimes in the awake gerbil is available, data presented in another publication by [Deane et al. \(2023\)](#) permits coarse estimates of τ_{SOI} -values for another small rodent species – the black mouse. For neural populations in A1 of *awake* mice, the electrophysiological data presented in the study suggests lifetimes of RS in the same range as we observed for the anaesthetised gerbils, with τ_{SOI} not exceeding 0.3 s. Note, however that the stimulus duration in these paradigms was much shorter than for our experiments (< 4 ms) and the results are thus only comparable to a limited extent. Further investigation in awake gerbils is required to determine whether and to what extent anaesthesia affects RS lifetime.

Differences introduced by the anatomical source of the measured response

Another explanation for the different RS lifetimes we observed in gerbils and humans could be that RS dynamics vary across auditory fields, with τ_{SOI} -values increasing when moving from the core to higher order areas, i.e. the belt and parabelt. In such a case, a value deduced from the activation of a core field would be smaller than a value deduced from the net activation of the entire AC. For the rat AC, [Nieto-Diego and Malmierca \(2016\)](#) observed that, for a

given sequence of standard tones, RS was most pronounced in the non-primary fields. This is consistent with the hypothesis of longer RS lifetimes in the belt and parabelt than in the core. Studies in human subjects also suggest different RS dynamics across AC areas (Lu et al., 1992b; Kropotov et al., 2000; Jääskeläinen et al., 2004; Opitz et al., 2005). However, the available data from the human AC is not sufficient to make clear comparative statements about the lifetime of RS along the core-belt-parabelt hierarchy.

Recent modelling work by Hajizadeh et al. (2022), who constructed a simplified computational model of the AC with one network node per AC area, predicts that RS lifetime is about 0.2 s longer in the belt and parabelt than in the core. In contrast, our recent work (see Project 4) predicts that RS lifetimes are shortest in the belt. We optimised the parameter values of our computational human AC model based on human MEG data and found that, in the optimised model, RS lifetime was about 0.5 s shorter in the belt than in core and parabelt. Further work is required to explore and verify these contradictory predictions regarding RS lifetimes in the different areas of the human AC.

Ramifications of species-specific RS lifetimes

If the time scales of RS lifetime in gerbil versus human AC are indeed species-specific, this could be a reflection of the AC dynamics being tailored to the species-specific temporal patterns of vocalisations employed to exchange information. In their detailed study of the spectrotemporal pattern of gerbil vocalisations, Kobayasi and Riquimaroux (2012) showed that gerbil “utterances” are much shorter and more repetitive than human utterances. They identified a maximum “phrase duration” of about 3 s. This means that the gerbil operates in a soundscape that requires temporal integration across much shorter time scales than human speech. For humans, Lu et al. (1992a) showed that there is a correlation between the subject-specific lifetime of RS and of a behaviourally evaluated memory trace. Thus, one way to further investigate the species-specificity of RS lifetime would be to test whether the behavioural lifetime of auditory sensory memory is shorter in gerbils than in humans.

Work by Kaernbach and Schulze (2002) suggests that this is indeed the case. They found that gerbils can distinguish a stimulus comprising of seamless repetitions of a shorter white noise stimulus (repeated noise, RN) from a stimulus made up of entirely random white noise (non-repeated noise, NRN) if the cycle length of the RN does not exceed 360 ms (observed limit after 156 days of training the behavioural response). In contrast, humans can reliably identify RN with a cycle length of several seconds (after just a few minutes of training, see Warren et al., 2001). This difference in performance might be a direct reflection of species-specific RS lifetimes, which, in turn, could reflect the brain’s evolution-based adaptation to or specialisation towards a distinct habitat.

The hypotheses introduced above – (1) differences observed due to the measurement

method, (2) differences observed due to anaesthesia, (3) differences observed due to the anatomical sources of the response, (4) differences observed due to the species-specificity of RS dynamics – do not necessarily exclude each other. Further studies are required to disentangle the potential species- and observation-level-specificity of RS lifetime. Note that species-specific RS dynamics do not necessarily suggest species-specific STSD dynamics. They could also reflect the species-specific AC field structures and connectivity patterns. As illustrated in Figures 4 vs. 36, the literature reports pronounced differences between gerbil and macaque AC. The gerbil AC consists of 8 fields, whereas the macaque AC, which is used as an approximation of the human AC, consists of 13 fields. The macaque AC connection pattern is characterised by a set of parallel serial streams along the core-belt-parabelt hierarchy, whereas the gerbil AC includes connections from core to parabelt. As discussed in ‘Q2: What does RS lifetime reflect?’, connectivity patterns shape RS dynamics and the species-specific AC networks could thus lead to species-specific RS lifetimes.

Q5: How could the gerbil auditory cortex model be refined?

In this work, we created a first version of an anatomically-informed computational model of the gerbil AC. The aim of the modelling efforts was not to create perfect replica of experimental data in terms of absolute values but rather to compare response behaviours across experimental conditions and to use the model to explore mechanisms that might give rise to the observed phenomena. While multiple time scales of synaptic depression have been reported for cortico-cortical synapses (Markram et al., 1998; Tsodyks and Markram, 1997; Varela et al., 1997), we opted for a single τ_{rec} -value (time constant of recovery from STSD) across all columns and AC fields because we wanted to focus on investigating the effect of connectivity patterns on RS dynamics. Thus, here we reported results from simulations at three distinct locations in the vast parameter space theoretically available for the gerbil AC model to illustrate how connectivity patterns can affect the lifetime of RS in the primary auditory field.

To reflect the response delay caused by sub-cortical processing, a delay term can be added to the stimulus sequence presented to the model (c.f. Hajizadeh et al., 2019, 2021, 2022). Here, we refrained from using such a term because peak latencies (and overall responses) in the animals were generally shorter than in the simulations. Therefore, future work should entail a reassessment of the gerbil AC model time constants. The RS lifetimes were also shorter in vivo than in silico. As demonstrated by May et al. (2015), RS lifetime τ_{SOI} scales with STSD lifetime τ_{rec} . Thus, the adjustment of the model’s time constant trio, τ_{m} , τ_{on} and τ_{rec} , would be a useful first target for our optimisation algorithm. While the optimisation algorithm, as presented in Chapter 4, previously addressed weight values, the time constants could similarly be reflected by a chromosome containing three genes. As the gerbil AC model

does not produce a direct equivalent of the electrophysiological response measures available in our in-vivo data set, the fitness function should not address the direct overlap between response waveforms but could instead be based on the in-vivo vs. in-silico agreement of SOI-specific peak latencies and of RS lifetime.

After this first adjustment of the response dynamics of the gerbil AC model, another optimisation process could be used to investigate whether the parameters defining the details of the network pattern can be adjusted such that the u-shaped τ_{SOI} distribution observed for the current version of the gerbil AC model (c.f. Figure 10) is inverted. This would replicate the response behaviour of animals 6 and 8 where $\tau_{\text{SOI}}(\Delta F = 0) > \tau_{\text{SOI}}(\Delta F = 2)$. The genes in the chromosome could reflect the connection-type-specific α - and σ -values (c.f. Equation (6) and Table 3) and the fitness function would be based on the condition $\tau_{\text{SOI}}(\Delta F = 0) > \tau_{\text{SOI}}(\Delta F = 0.5) > \tau_{\text{SOI}}(\Delta F = 1)$ etc. Of course, this would also require the definition of meaningful boundary conditions, such as for example $\sigma_{\text{ie}} > \sigma_{\text{ee}}$ to keep FR functions reasonably sharp. If the optimisation algorithm achieves the inversion of the τ_{SOI} distribution, this would be further proof of the strong effect of network pattern on RS lifetimes. Moreover, it would be interesting to trace back the network effects that give rise to the inverted response behaviour.

A gerbil AC model version with a stochastic network fine structure as defined in May et al. (2013; 2015) could be used to verify whether the findings reported there can be replicated in a model that reflects the AC anatomy of a different species. Unlike the AC model based on macaque anatomy, the gerbil AC model sends thalamic input to the core as well as the belt (albeit to a lesser extent). Moreover, there are direct connections from core to parabelt. Given that May et al. identified serial processing as an important condition for CS is there then, for example, still a similar overall percentage of CS columns? And does temporal binding of sequences still mostly occur in the core for shorter and in the belt and parabelt for longer sequences? As already suggested for the model described by May et al., this investigation should also entail tracking and characterisation of the network nodes that send input to the CS column.

Hajizadeh et al. (2019) reported that stochasticity in the weight matrices leads to columns that exhibit multi-peaked and/or broad tuning curves. Multi-peaked spectral tuning in the human AC has also been observed in vivo (Moerel et al., 2013) and is probably required for temporal binding of and CS to more complex stimulus patterns. In this context, the non-tonotopic organisation of the gerbil AC parabelt fields (as observed experimentally, c.f. literature referenced in Section 1.1.2) might also be of high functional relevance and should be explored in future simulations.

Closing remarks

The results presented in this doctoral thesis as well as the future studies proposed in the Discussion indicate that the triad of in-vivo experiments, simulations, and model optimisation is a powerful tool for the investigation of memory trace formation in the AC. Thus far, we have shown that RS is a network effect modulated by connectivity patterns. Further work will reveal how the brain exploits this effect in the context of temporal binding.

References

- Abbott, L. F., Varela, J., Sen, K., and Nelson, S. (1997). Synaptic depression and cortical gain control. *Science*, 275(5297):221–224.
- Ahlfors, S. P., Jones, S. R., Ahveninen, J., Hämäläinen, M. S., Belliveau, J. W., and Bar, M. (2015). Direction of magnetoencephalography sources associated with feedback and feedforward contributions in a visual object recognition task. *Neuroscience Letters*, 585:149–154.
- Anderson, L. A., Christianson, G. B., and Linden, J. F. (2009). Stimulus-specific adaptation occurs in the auditory thalamus. *Journal of Neuroscience*, 29(22):7359–7363.
- Antunes, F. M., Nelken, I., Covey, E., and Malmierca, M. S. (2010). Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *PLoS ONE*, 5(11):e14071.
- Arabas, J. (2001). *Wykłady z algorytmów ewolucyjnych*. Wydawnictwo Naukowo-Techniczne. In Polish.
- Asari, H. and Zador, A. M. (2009). Long-lasting context dependence constrains neural encoding models in rodent auditory cortex. *Journal of Neurophysiology*, 102(5):2638–2656.
- Baillet, S. (2013). *Forward and Inverse Problems of MEG/EEG*, pages 1–8. Springer New York, New York, NY.
- Bartlett, E. L. and Wang, X. (2005). Long-lasting modulation by stimulus context in primate auditory cortex. *Journal of Neurophysiology*, 94(1):83–104.
- Benda, J., Longtin, A., and Maler, L. (2005). Spike-frequency adaptation separates transient communication signals from background oscillations. *Journal of Neuroscience*, 25(9):2312–2321.
- Bertorelli, R., Adami, M., and Ongini, E. (1995). The Mongolian gerbil in experimental epilepsy. *The Italian Journal of Neurological Sciences*, 16:101–106.
- Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98.
- Brady, T. F., Robinson, M. M., Williams, J. R., and Wixted, J. T. (2023). Measuring memory is harder than you think: How to avoid problematic measurement practices in memory research. *Psychonomic Bulletin & Review*, 30(2):421–449.
- Brosch, M. and Schreiner, C. E. (1997). Time course of forward masking tuning curves in cat primary auditory cortex. *Journal of Neurophysiology*, 77(2):923–943.
- Brosch, M. and Schreiner, C. E. (2000). Sequence sensitivity of neurons in cat primary auditory cortex. *Cerebral Cortex*, 10(12):1155–1167.
- Brosch, M., Schulz, A., and Scheich, H. (1999). Processing of sound sequences in macaque auditory cortex: response enhancement. *Journal of Neurophysiology*, 82(3):1542–1559.
- Brunk, M. G., Deane, K. E., Kisse, M., Deliano, M., Vieweg, S., Ohl, F. W., Lippert, M. T., and Happel, M. F. K. (2019). Optogenetic stimulation of the VTA modulates a frequency-specific gain of thalamocortical inputs in infragranular layers of the auditory cortex. *Scientific Reports*, 9(1):1–15.

- Budinger, E., Brosch, M., Scheich, H., and Mylius, J. (2013). The subcortical auditory structures in the Mongolian gerbil: II. Frequency-related topography of the connections with cortical field AI. *Journal of Comparative Neurology*, 521(12):2772–2797.
- Budinger, E., Heil, P., Hess, A., and Scheich, H. (2006). Multisensory processing via early cortical stages: connections of the primary auditory cortical field with other sensory systems. *Neuroscience*, 143(4):1065–1083.
- Budinger, E., Heil, P., and Scheich, H. (2000a). Functional organization of auditory cortex in the Mongolian gerbil (*Meriones unguiculatus*). III. Anatomical subdivisions and cortico-cortical connections. *European Journal of Neuroscience*, 12(7):2425–2451.
- Budinger, E., Heil, P., and Scheich, H. (2000b). Functional organization of auditory cortex in the Mongolian gerbil (*Meriones unguiculatus*). IV. Connections with anatomically characterized subcortical structures. *European Journal of Neuroscience*, 12(7):2452–2474.
- Budinger, E. and Kanold, P. O. (2018). Auditory cortex circuits. In *The mammalian auditory pathways*, pages 199–233. Springer.
- Budinger, E., Laszcz, A., Lison, H., Scheich, H., and Ohl, F. W. (2008). Non-sensory cortical and subcortical connections of the primary auditory cortex in Mongolian gerbils: bottom-up and top-down processing of neuronal information via field AI. *Brain Research*, 1220:2–32.
- Budinger, E. and Scheich, H. (2009). Anatomical connections suitable for the direct processing of neuronal information of different modalities via the rodent primary auditory cortex. *Hearing Research*, 258(1-2):16–27.
- Buonomano, D. V. and Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience*, 10(2):113–125.
- Buonomano, D. V. and Merzenich, M. M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, 267(5200):1028–1030.
- Butler, R. A. (1968). Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. *The Journal of the Acoustical Society of America*, 44(4):945–950.
- Butts, D. A. and Goldman, M. S. (2006). Tuning curves, neuronal variability, and sensory coding. *PLoS Biology*, 4(4):e92.
- Chaudhuri, R. and Fiete, I. (2016). Computational principles of memory. *Nature Neuroscience*, 19(3):394–403.
- Cheng, C.-H. and Lin, Y.-Y. (2012). The effects of aging on lifetime of auditory sensory memory in humans. *Biological Psychology*, 89(2):306–312.
- Dar*, A. H., Härtwich*, N., Hajizadeh, A., Brosch, M., König, R., and May, P. J. C. (2025). Hemispheric difference of adaptation lifetime in human auditory cortex measured with MEG. *Hearing Research*, 458:109173.
- David, S. V. and Shamma, S. A. (2013). Integration over multiple timescales in primary auditory cortex. *Journal of Neuroscience*, 33(49):19154–19166.
- de la Mothe, L. A., Blumell, S., Kajikawa, Y., and Hackett, T. A. (2006). Cortical connections of the auditory cortex in marmoset monkeys: core and medial belt regions. *Journal of Comparative Neurology*, 496(1):27–71.

- Deane, K. E., Brunk, M. G., Curran, A. W., Zempeltzi, M. M., Ma, J., Lin, X., Abela, F., Aksit, S., Deliano, M., Ohl, F. W., et al. (2020). Ketamine anaesthesia induces gain enhancement via recurrent excitation in granular input layers of the auditory cortex. *The Journal of Physiology*, 598(13):2741–2755.
- Deane, K. E., García-Rosales, F., Klymentiev, R., Hechavarria, J. C., and Happel, M. F. K. (2023). Comparative analysis of primary auditory cortical responses in bats and mice to repetitive stimuli trains. *bioRxiv*.
- Deliano, M., Brunk, M. G., El-Tabbal, M., Zempeltzi, M. M., Happel, M. F. K., and Ohl, F. W. (2020). Dopaminergic neuromodulation of high gamma stimulus phase-locking in gerbil primary auditory cortex mediated by D1/D5-receptors. *European Journal of Neuroscience*, 51(5):1315–1327.
- Diedrichsen, J., Shadmehr, R., and Ivry, R. B. (2010). The coordination of movement: optimal feedback control and beyond. *Trends in Cognitive Sciences*, 14(1):31–39.
- Douglas, R. J. and Martin, K. A. C. (2004). Neuronal circuits of the neocortex. *Annual Review of Neuroscience*, 27:419–451.
- Doupe, A. J. (1997). Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *Journal of Neuroscience*, 17(3):1147–1167.
- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1 – 26.
- Emiliani, V., Entcheva, E., Hedrich, R., Hegemann, P., Konrad, K. R., Lüscher, C., Mahn, M., Pan, Z.-H., Sims, R. R., Vierock, J., et al. (2022). Optogenetics for light control of biological systems. *Nature Reviews Methods Primers*, 2(1):55.
- Finlayson, P. G. and Adam, T. J. (1997). Excitatory and inhibitory response adaptation in the superior olive complex affects binaural acoustic processing. *Hearing Research*, 103(1-2):1–18.
- Fioravante, D. and Regehr, W. G. (2011). Short-term forms of presynaptic plasticity. *Current Opinion in Neurobiology*, 21(2):269–274.
- Fortune, E. S. and Rose, G. J. (2001). Short-term synaptic plasticity as a temporal filter. *Trends in Neurosciences*, 24(7):381–385.
- Friauf, E., Fischer, A. U., and Fuhr, M. F. (2015). Synaptic plasticity in the auditory system: a review. *Cell and Tissue Research*, 361:177–213.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456):815–836.
- Fritz, C. O., Morris, P. E., and Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1):2.
- Givre, S., Schroeder, C., and Arezzo, J. (1994). Contribution of extrastriate area V4 to the surface-recorded flash VEP in the awake macaque. *Vision Research*, 34(4):415–428.
- Gleich, O. and Strutz, J. (2012). *Hearing Loss*, chapter 4: The Mongolian gerbil as a model for the analysis of peripheral and central age-dependent hearing loss, pages 67–92. InTech.
- Gollisch, T. and Herz, A. V. (2004). Input-driven components of spike-frequency adaptation can be unmasked in vivo. *Journal of Neuroscience*, 24(34):7435–7444.

- Goudar, V. and Buonomano, D. V. (2015). A model of order-selectivity based on dynamic changes in the balance of excitation and inhibition produced by short-term synaptic plasticity. *Journal of Neurophysiology*, 113(2):509–523.
- Hackett, T., Stepniewska, I., and Kaas, J. (1998). Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *Journal of Comparative Neurology*, 394(4):475–495.
- Hackett, T. A. (2011). Information flow in the auditory cortical network. *Hearing Research*, 271(1):133–146.
- Hajizadeh, A., Matysiak, A., Brechmann, A., König, R., and May, P. J. (2021). Why do humans have unique auditory event-related fields? Evidence from computational modeling and MEG experiments. *Psychophysiology*, 58(4):e13769.
- Hajizadeh, A., Matysiak, A., May, P. J. C., and König, R. (2019). Explaining event-related fields by a mechanistic model encapsulating the anatomical structure of auditory cortex. *Biological Cybernetics*, 113:321–345.
- Hajizadeh, A., Matysiak, A., Wolfrum, M., May, P. J. C., and König, R. (2022). Auditory cortex modelled as a dynamical network of oscillators: understanding event-related fields and their adaptation. *Biological Cybernetics*, 116(4):475–499.
- Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., and Lounasmaa, O. V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2):413.
- Hammond, C. (2015). *Cellular and Molecular Neurophysiology*, chapter 3: Ionic gradients, membrane potential and ionic currents. Academic Press.
- Happel, M. F., Jeschke, M., and Ohl, F. W. (2010). Spectral integration in primary auditory cortex attributable to temporally precise convergence of thalamocortical and intracortical input. *Journal of Neuroscience*, 30(33):11114–11127.
- Harding, G. (1992). The currents that flow in the somatosensory cortex during the direct cortical response. *Experimental Brain Research*, 90:29–39.
- Henschke, J. U., Noesselt, T., Scheich, H., and Budinger, E. (2015). Possible anatomical pathways for short-latency multisensory integration processes in primary sensory cortices. *Brain Structure and Function*, 220(2):955–977.
- Heschl, R. L. (1878). *Über die vordere quere Schläfenwindung des menschlichen Großhirns: aus Anlass der 25-jährigen Jubiläums-Feier der Wiener Landes-Irrenanstalt veröffentlicht.* Universitätsbibliothek.
- Hesterberg, T. (2011). Bootstrap. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3(6):497–526.
- Hine, J. and Debener, S. (2007). Late auditory evoked potentials asymmetry revisited. *Clinical Neurophysiology*, 118(6):1274–1285.
- Hopfield, J. J. and Tank, D. W. (1986). Computing with neural circuits - a model. *Science*, 233(4764):625–633.
- Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levänen, S., Lin, F.-H., May, P. J. C., Melcher, J., Stufflebeam, S., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences*, 101(17):6809–6814.

- Jacquelin, J. (2009). Régressions et équations intégrales. <https://www.scribd.com/doc/14674814/Regressions-et-equations-integrales>, accessed 2020-02-05.
- Kaas, J. H. and Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences*, 97(22):11793–11799.
- Kaernbach, C. and Schulze, H. (2002). Auditory sensory memory for random waveforms in the Mongolian gerbil. *Neuroscience Letters*, 329(1):37–40.
- Kilgard, M. P. and Merzenich, M. M. (1999). Distributed representation of spectral and temporal information in rat primary auditory cortex. *Hearing Research*, 134(1-2):16–28.
- King, A. J., Teki, S., and Willmore, B. D. (2018). Recent advances in understanding the auditory cortex. *F1000Research*, 7:1555.
- Klein, C., von der Behrens, W., and Gaese, B. H. (2014). Stimulus-specific adaptation in field potentials and neuronal responses to frequency-modulated tones in the primary auditory cortex. *Brain Topography*, 27:599–610.
- Kobayasi, K. I. and Riquimaroux, H. (2012). Classification of vocalizations in the Mongolian gerbil, *Meriones unguiculatus*. *The Journal of the Acoustical Society of America*, 131(2):1622–1631.
- Kohashi, T., Lube, A. J., Yang, J. H., Roberts-Gaddipati, P. S., and Carlson, B. A. (2021). Pauses during communication release behavioral habituation through recovery from synaptic depression. *Current Biology*, 31(14):3145–3152.
- Kohl, C., Parviainen, T., and Jones, S. R. (2022). Neural mechanisms underlying human auditory evoked responses revealed by Human Neocortical Neurosolver. *Brain Topography*, 35(1):19–35.
- Kriegeskorte, N. and Douglas, P. K. (2018). Cognitive computational neuroscience. *Nature Neuroscience*, 21(9):1148–1160.
- Kropotov, J. D., Alho, K., Näätänen, R., Ponomarev, V. A., Kropotova, O. V., Anichkov, A. D., and Nechaev, V. B. (2000). Human auditory-cortex mechanisms of preattentive sound discrimination. *Neuroscience Letters*, 280(2):87–90.
- Kudela, P., Boatman-Reich, D., Beeman, D., and Anderson, W. S. (2018). Modeling neural adaptation in auditory cortex. *Frontiers in Neural Circuits*, 12:72.
- Kurt, S., Deutscher, A., Crook, J. M., Ohl, F. W., Budinger, E., Moeller, C. K., Scheich, H., and Schulze, H. (2008). Auditory cortical contrast enhancing by global winner-take-all inhibitory interactions. *PLoS ONE*, 3(3):e1735.
- Lee, T. P. and Buonomano, D. V. (2012). Unsupervised formation of vocalization-sensitive neurons: a cortical model based on short-term and homeostatic plasticity. *Neural Computation*, 24(10):2579–2603.
- Levy, R. B. and Reyes, A. D. (2012). Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. *Journal of Neuroscience*, 32(16):5609–5619.
- Lewicki, M. S. and Arthur, B. J. (1996). Hierarchical organization of auditory temporal context sensitivity. *Journal of Neuroscience*, 16(21):6987–6998.
- Loebel, A., Nelken, I., and Tsodyks, M. (2007). Processing of sounds by population spikes in a model of primary auditory cortex. *Frontiers in Neuroscience*, 1:15.

- Lu, Z.-L., Williamson, S. J., and Kaufman, L. (1992a). Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science*, 258(5088):1668–1670.
- Lu, Z.-L., Williamson, S. J., and Kaufman, L. (1992b). Human auditory primary and association cortex have differing lifetimes for activation traces. *Brain Research*, 572(1-2):236–241.
- Lukichev, A. (2019). Physical meaning of the stretched exponential Kohlrausch function. *Physics Letters A*, 383(24):2983–2987.
- Ma*, J., Brunk*, M., Matysiak, A., Härtwich, N., Ohl, F. W., May, P. J. C., Happel, M. F. K., König, R., and Deliano, M. (2021). Layer-specific intracortical amplification shortens the lifetime of thalamocortical repetition suppression in auditory cortex. Preprint, *Research Square*, doi.org/10.21203/rs.3.rs-1165857/v1.
- Mäkelä, J. P., Ahonen, A., Hämäläinen, M., Hari, R., Ilmoniemi, R., Kajola, M., Knuutila, J., Lounasmaa, O., McEvoy, L., Salmelin, R., et al. (1993). Functional differences between auditory cortices of the two hemispheres revealed by whole-head neuromagnetic recordings. *Human Brain Mapping*, 1(1):48–56.
- Malmierca, M. S., Cristaudo, S., Pérez-González, D., and Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *Journal of Neuroscience*, 29(17):5483–5493.
- Malmierca, M. S., Sanchez-Vives, M. V., Escera, C., and Bendixen, A. (2014). Neuronal adaptation, novelty detection and regularity encoding in audition. *Frontiers in Systems Neuroscience*, 8:111.
- Margoliash, D. and Fortune, E. S. (1992). Temporal and harmonic combination-sensitive neurons in the zebra finch’s HVC. *Journal of Neuroscience*, 12(11):4309–4326.
- Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proceedings of the National Academy of Sciences*, 95(9):5323–5328.
- May, P. J. C. (2021). The adaptation model offers a challenge for the predictive coding account of mismatch negativity. *Frontiers in Human Neuroscience*, 15:721574.
- May, P. J. C. and Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, 47(1):66–122.
- May, P. J. C. and Tiitinen, H. (2013). Temporal binding of sound emerges out of anatomical structure and synaptic dynamics of auditory cortex. *Frontiers in Computational Neuroscience*, 7:152.
- May, P. J. C., Westö, J., and Tiitinen, H. (2015). Computational modelling suggests that temporal integration results from synaptic adaptation in auditory cortex. *European Journal of Neuroscience*, 41(5):615–630.
- McEvoy, L., Levänen, S., and Loveless, N. (1997). Temporal characteristics of auditory sensory memory: neuromagnetic evidence. *Psychophysiology*, 34(3):308–316.
- McKenna, T. M., Weinberger, N. M., and Diamond, D. M. (1989). Responses of single auditory cortical neurons to tone sequences. *Brain Research*, 481(1):142–153.
- Megela, A. L. and Teyler, T. J. (1979). Habituation and the human evoked potential. *Journal of Comparative and Physiological Psychology*, 93(6):1154.
- Michalewicz, Z. (1996). *Genetic Algorithms + Data Structures = Evolution Programmes*. Springer, Berlin, Heidelberg.

- Mill, R., Coath, M., Wennekens, T., and Denham, S. L. (2011). A neurocomputational model of stimulus-specific adaptation to oddball and Markov sequences. *PLoS Computational Biology*, 7(8):e1002117.
- Mill, R., Coath, M., Wennekens, T., and Denham, S. L. (2012). Characterising stimulus-specific adaptation using a multi-layer field model. *Brain Research*, 1434:178–188.
- Mitzdorf, U. (1985). Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiological Reviews*, 65(1):37–100.
- Moeller, C. K., Kurt, S., Happel, M. F. K., and Schulze, H. (2010). Long-range effects of gabaergic inhibition in gerbil primary auditory cortex. *European Journal of Neuroscience*, 31(1):49–59.
- Moerel, M., De Martino, F., and Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience*, 8:225.
- Moerel, M., De Martino, F., Santoro, R., Ugurbil, K., Goebel, R., Yacoub, E., and Formisano, E. (2013). Processing of natural sounds: characterization of multipeak spectral tuning in human auditory cortex. *Journal of Neuroscience*, 33(29):11888–11898.
- Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., and Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage*, 13(4):684–701.
- Motanis, H., Seay, M. J., and Buonomano, D. V. (2018). Short-term synaptic plasticity as a mechanism for sensory timing. *Trends in Neurosciences*, 41(10):701–711.
- Mukamel, R. and Fried, I. (2012). Human intracranial recordings and cognitive neuroscience. *Annual Review of Psychology*, 63(1):511–537.
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, 13(2):201–233.
- Näätänen, R., Gaillard, A. W., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4):313–329.
- Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., and Paavilainen, P. (1993). Development of a memory trace for a complex sound in the human brain. *NeuroReport*, 4(5):503–506.
- Nelken, I. (2014). Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. *Biological Cybernetics*, 108:655–663.
- Newsome, W. T., Britten, K. H., and Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, 341(6237):52–54.
- Nieto-Diego, J. and Malmierca, M. S. (2016). Topographic distribution of stimulus-specific adaptation across auditory cortical fields in the anesthetized rat. *PLoS Biology*, 14(3):e1002397.
- Ocklenburg, S. and Güntürkün, O. (2024). *The lateralized brain: The neuroscience and evolution of hemispheric asymmetries*. Elsevier.
- Ohl, F. W. and Scheich, H. (1997a). Learning-induced dynamic receptive field changes in primary auditory cortex of the unanaesthetized Mongolian gerbil. *Journal of Comparative Physiology A*, 181:685–696.

- Ohl, F. W. and Scheich, H. (1997b). Learning-induced dynamic receptive field changes in primary auditory cortex of the unanaesthetized Mongolian gerbil. *Journal of Comparative Physiology A*, 181(6):685–696.
- Olshausen, B. A. and Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4):481–487.
- Opitz, B., Schröger, E., and Von Cramon, D. Y. (2005). Sensory and cognitive mechanisms for preattentive change detection in auditory cortex. *European Journal of Neuroscience*, 21(2):531–535.
- Orduña, I., Mercado III, E., Gluck, M. A., and Merzenich, M. M. (2001). Spectrotemporal sensitivities in rat auditory cortical neurons. *Hearing Research*, 160(1-2):47–57.
- Pantev, C., Ross, B., Berg, P., Elbert, T., and Rockstroh, B. (1998). Study of the human auditory cortices using a whole-head magnetometer: left vs. right hemisphere and ipsilateral vs. contralateral stimulation. *Audiology and Neurotology*, 3(2-3):183–190.
- Parras, G. G., Nieto-Diego, J., Carbajal, G. V., Valdés-Baizabal, C., Escera, C., and Malmierca, M. S. (2017). Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nature Communications*, 8(1):2148.
- Pérez-González, D. and Malmierca, M. S. (2014). Adaptation in the auditory system: an overview. *Frontiers in Integrative Neuroscience*, 8:19.
- Pérez-González, D., Malmierca, M. S., and Covey, E. (2005). Novelty detector neurons in the mammalian auditory midbrain. *European Journal of Neuroscience*, 22(11):2879–2885.
- Rademacher, J., Morosan, P., Schleicher, A., Freund, H.-J., and Zilles, K. (2001). Human primary auditory cortex in women and men. *Neuroreport*, 12(8):1561–1565.
- Radtke-Schuller, S., Schuller, G., Angenstein, F., Grosser, O. S., Goldschmidt, J., and Budinger, E. (2016). Brain atlas of the Mongolian gerbil (*Meriones unguiculatus*) in CT/MRI-aided stereotaxic coordinates. *Brain Structure and Function*, 221(1):1–272.
- Rao, R. P. and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87.
- Recanzone, G. H. (2008). Representation of con-specific vocalizations in the core and belt areas of the auditory cortex in the alert macaque monkey. *Journal of Neuroscience*, 28(49):13184–13193.
- Reite, M., Zimmerman, J. T., and Zimmerman, J. E. (1981). Magnetic auditory evoked fields: interhemispheric asymmetry. *Electroencephalography and Clinical Neurophysiology*, 51(4):388–392.
- Rojas, D. C., Teale, P., Sheeder, J., and Reite, M. (1999). Sex differences in the refractory period of the 100 ms auditory evoked magnetic field. *NeuroReport*, 10(16):3321–3325.
- Romanski, L. M. and Averbeck, B. B. (2009). The primate cortical auditory system and neural representation of conspecific vocalizations. *Annual Review of Neuroscience*, 32:315–346.
- Rosburg, T., Zimmerer, K., and Huonker, R. (2010). Short-term habituation of auditory evoked potential and neuromagnetic field components in dependence of the interstimulus interval. *Experimental Brain Research*, 205:559–570.

- Rosenbaum, R., Rubin, J., and Doiron, B. (2012). Short term synaptic depression imposes a frequency dependent filter on synaptic information transfer. *PLoS Computational Biology*, 8(6):e1002557.
- Ross, B., Herdman, A., and Pantev, C. (2005). Right hemispheric laterality of human 40 Hz auditory steady-state responses. *Cerebral Cortex*, 15(12):2029–2039.
- Ruthig, P. and Schönwiesner, M. (2022). Common principles in the lateralization of auditory cortex structure and function for vocal communication in primates and rodents. *European Journal of Neuroscience*, 55(3):827–845.
- Rutkowska, D., Piliński, M., and Rutkowski, L. (1997). *Sieci neuronowe, algorytmy genetyczne i systemy rozmyte*. Wydawnictwo Naukowe PWN. In Polish.
- Ryan, A. (1976). Hearing sensitivity of the Mongolian gerbil, *Meriones unguiculatus*. *The Journal of the Acoustical Society of America*, 59(5):1222–1226.
- Sadagopan, S. and Wang, X. (2009). Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex. *Journal of Neuroscience*, 29(36):11192–11202.
- Saldeitis, K., Happel, M. F. K., Ohl, F. W., Scheich, H., and Budinger, E. (2014). Anatomy of the auditory thalamocortical system in the Mongolian gerbil: Nuclear origins and cortical field-, layer-, and frequency-specificities. *Journal of Comparative Neurology*, 522(10):2397–2430.
- Salmasi, M., Loebel, A., Glasauer, S., and Stemmler, M. (2019). Short-term synaptic depression can increase the rate of information transfer at a release site. *PLoS Computational Biology*, 15(1):e1006666.
- Salmelin, R., Schnitzler, A., Parkkonen, L., Biermann, K., Helenius, P., Kiviniemi, K., Kuukka, K., Schmitz, F., and Freund, H.-J. (1999). Native language, gender, and functional organization of the auditory cortex. *Proceedings of the National Academy of Sciences*, 96(18):10460–10465.
- Sams, M., Hari, R., Rif, J., and Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 sec: neuromagnetic evidence. *Journal of Cognitive Neuroscience*, 5(3):363–370.
- Scheich, H. (1991). Auditory cortex: comparative aspects of maps and plasticity. *Current Opinion in Neurobiology*, 1(2):236–247.
- Scheich, H., Heil, P., and Langner, G. (1993). Functional organization of auditory cortex in the Mongolian gerbil (*Meriones unguiculatus*) II. Tonotopic 2-deoxyglucose. *European Journal of Neuroscience*, 5(7):898–914.
- Schroeder, C. E. and Foxe, J. J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Cognitive Brain Research*, 14(1):187–198.
- Schroeder, C. E., Mehta, A. D., and Givre, S. J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cerebral Cortex (New York, NY: 1991)*, 8(7):575–592.
- Shadlen, M. N. and Kiani, R. (2013). Decision making as a window on cognition. *Neuron*, 80(3):791–806.
- Shaw, M. E., Hämäläinen, M. S., and Gutschalk, A. (2013). How anatomical asymmetry of human auditory cortex can lead to a rightward bias in auditory evoked fields. *Neuroimage*, 74:22–29.

- Sielużycki, C., Matysiak, A., König, R., and Iskander, D. R. (2021). Reducing the number of MEG/EEG trials needed for the estimation of brain evoked responses: A bootstrap approach. *IEEE Transactions on Biomedical Engineering*, 68(7):2301–2312.
- Simoncelli, E. P. and Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24(1):1193–1216.
- Stüttgen, M. C., Nonkes, L. J., Geis, H. R. A., Tiesinga, P. H., and Houweling, A. R. (2017). Temporally precise control of single-neuron spiking by juxtacellular nanostimulation. *Journal of Neurophysiology*, 117(3):1363–1378.
- Szymanski, F. D., Garcia-Lazaro, J. A., and Schnupp, J. W. (2009). Current source density profiles of stimulus-specific adaptation in rat auditory cortex. *Journal of Neurophysiology*, 102(3):1483–1490.
- Taaseh, N., Yaron, A., and Nelken, I. (2011). Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS ONE*, 6(8):e23369.
- Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: A user-friendly application for MEG/EEG analysis. *Computational intelligence and neuroscience*, 2011(1):879716.
- Tchabovsky, A. V., Savinetskaya, L. E., Ovchinnikova, N. L., Safonova, A., Ilchenko, O. N., Sapozhnikova, S. R., and Vasilieva, N. A. (2019). Sociability and pair-bonding in gerbils: a comparative experimental study. *Current Zoology*, 65(4):363–373.
- Teufel, C. and Fletcher, P. C. (2016). The promises and pitfalls of applying computational models to neurological and psychiatric disorders. *Brain*, 139(10):2600–2608.
- Thomas, H. and López, V. (2003). Comparative study of inter-and intrahemispheric cortico-cortical connections in gerbil auditory cortex. *Biological Research*, 36(2):155–169.
- Thomas, H., Tillein, J., Heil, P., and Scheich, H. (1993). Functional organization of auditory cortex in the Mongolian gerbil (*Meriones unguiculatus*). I. Electrophysiological mapping of frequency representation and distinction of fields. *European Journal of Neuroscience*, 5(7):882–897.
- Tiitinen, H., May, P. J. C., Reinikainen, K., and Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*, 372(6501):90–92.
- Tomana, E., Härtwich, N., Rozmarynowski, A., König, R., May, P. J., and Sielużycki, C. (2023). Optimising a computational model of human auditory cortex with an evolutionary algorithm. *Hearing Research*, 439:108879.
- Tsodyks, M. V. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences*, 94(2):719–723.
- Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *Journal of Neuroscience*, 24(46):10440–10453.
- Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, 6(4):391–398.
- v. Economo, C. and Horn, L. (1930). Über Windungsrelief, Maße und Rindenarchitektonik der Supratemporalfläche, ihre individuellen und ihre Seitenunterschiede. *Zeitschrift für die gesamte Neurologie und Psychiatrie*, 130(1):678–757.

- Varela, J. A., Sen, K., Gibson, J., Fost, J., Abbott, L., and Nelson, S. B. (1997). A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *Journal of Neuroscience*, 17(20):7926–7940.
- Wacongne, C., Changeux, J.-P., and Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *Journal of Neuroscience*, 32(11):3665–3678.
- Wang, X.-J. (2008). Decision making in recurrent neuronal circuits. *Neuron*, 60(2):215–234.
- Warren, R. M., Bashford, J. A., Cooley, J. M., and Brubaker, B. S. (2001). Detection of acoustic repetition for very long stochastic patterns. *Perception & Psychophysics*, 63(1):175–182.
- Webster, M. A. (2012). Evolving concepts of sensory adaptation. *F1000 Biology Reports*, 4:21.
- Wehr, M. and Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature*, 426(6965):442–446.
- Wehr, M. and Zador, A. M. (2005). Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron*, 47(3):437–445.
- Westö, J., May, P. J., and Tiitinen, H. (2016). Memory stacking in hierarchical networks. *Neural Computation*, 28(2):327–353.
- Willmore, B. D. and King, A. J. (2023). Adaptation in auditory processing. *Physiological Reviews*, 103(2):1025–1058.
- Wilson, H. R. and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1):1–24.
- Woldorff, M. G., Tempelmann, C., Fell, J., Tegeler, C., Gaschler-Markefski, B., Hinrichs, H., Heinze, H.-J., and Scheich, H. (1999). Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography. *Human Brain Mapping*, 7(1):49–66.
- Wrobel, C., Dieter, A., Huet, A., Keppeler, D., Duque-Afonso, C. J., Vogl, C., Hoch, G., Jeschke, M., and Moser, T. (2018). Optogenetic stimulation of cochlear neurons activates the auditory pathway and restores auditory-driven behavior in deaf adult gerbils. *Science Translational Medicine*, 10(449):eaao0540.
- Yamins, D. L. and DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, 19(3):356–365.
- Yarden, T. S. and Nelken, I. (2017). Stimulus-specific adaptation in a recurrent network model of primary auditory cortex. *PLoS Computational Biology*, 13(3):e1005437.
- Yates, G., Cody, A., and Johnstone, B. (1983). Recovery of eighth nerve action potential thresholds after exposure to short, intense pure tones: similarities with temporary threshold shift. *Hearing Research*, 12(3):305–322.
- Zacharias, N., König, R., and Heil, P. (2012). Stimulation-history effects on the M100 revealed by its differential dependence on the stimulus onset interval. *Psychophysiology*, 49(7):909–919.

Declaration of honour

I hereby declare that I prepared this thesis without impermissible help from third parties and that none other than the aids indicated have been used.

All sources of information are clearly marked, including my own publications.

In particular I have not consciously:

- Fabricated data or rejected undesirable results,
- Misused statistical methods with the aim of drawing other conclusions than those warranted by the available data,
- Plagiarised external data or publications,
- Presented the results of other researchers in a distorted way.

I am aware that violations of copyright may lead to injunction and damage claims by the author and also to prosecution by the law enforcement authorities.

I hereby agree that the thesis may be electronically reviewed with the aim of identifying plagiarism.

This work has not yet been submitted as a doctoral thesis in the same or a similar form in Germany or in any other country. It has also not yet been published as a whole.

Magdeburg, 26 August 2024

Nina Härtwich