

# Robust Neuronal Discrimination in Primary Auditory Cortex Despite Degradations of Spectro-temporal Acoustic Details: Comparison Between Guinea Pigs with Normal Hearing and Mild Age-Related Hearing Loss

Yonane Aushana, Samira Souffi, Jean-Marc Edeline, Christian Lorenzi, Chloé Huetz

## ► To cite this version:

Yonane Aushana, Samira Souffi, Jean-Marc Edeline, Christian Lorenzi, Chloé Huetz. Robust Neuronal Discrimination in Primary Auditory Cortex Despite Degradations of Spectro-temporal Acoustic Details: Comparison Between Guinea Pigs with Normal Hearing and Mild Age-Related Hearing Loss. *Journal of the Association for Research in Otolaryngology*, Springer Verlag, 2018, 19 (2), pp.163-180. 10.1007/s10162-017-0649-1 . hal-02111943

HAL Id: hal-02111943

<https://hal.archives-ouvertes.fr/hal-02111943>

Submitted on 26 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Journal of the Association for Research in Otolaryngology

## Robust neuronal discrimination in primary auditory cortex despite degradations of spectro-temporal acoustic details: comparison between guinea pigs with normal hearing and mild age-related hearing loss

--Manuscript Draft--

<b>Manuscript Number:</b>	JARO-D-17-00039R3	
<b>Full Title:</b>	Robust neuronal discrimination in primary auditory cortex despite degradations of spectro-temporal acoustic details: comparison between guinea pigs with normal hearing and mild age-related hearing loss	
<b>Article Type:</b>	Original Article	
<b>Keywords:</b>	vocoder, envelope, fine structure, electrophysiology, auditory cortex, neural discrimination performance, spike-timing	
<b>Corresponding Author:</b>	Jean-Marc EDELINE, PhD Neuro-PSI UMR CNRS 9197 ORSAY, FRANCE	
<b>Corresponding Author Secondary Information:</b>		
<b>Corresponding Author's Institution:</b>	Neuro-PSI UMR CNRS 9197	
<b>Corresponding Author's Secondary Institution:</b>		
<b>First Author:</b>	Yonane Aushana	
<b>First Author Secondary Information:</b>		
<b>Order of Authors:</b>	Yonane Aushana	
	Samira Souffi	
	Jean-Marc EDELINE, PhD	
	Christian Lorenzi	
	Chloe Huetz	
<b>Order of Authors Secondary Information:</b>		
<b>Funding Information:</b>	Agence Nationale de la Recherche (ANR-11-0001-02)	Pr Christian Lorenzi
	Agence Nationale de la Recherche (ANR-10-LABX-0087)	Pr Christian Lorenzi
<b>Abstract:</b>	<p>This study investigated to which extent the primary auditory cortex of young normal-hearing and mild hearing-impaired aged animals is able to maintain invariant representation of critical temporal-modulation features when sounds are submitted to degradations of fine spectro-temporal acoustic details. This was achieved by recording ensemble of cortical responses to conspecific vocalizations in guinea pigs with either normal hearing or mild age-related sensorineural hearing loss. The vocalizations were degraded using a tone vocoder. The neuronal responses and their discrimination capacities (estimated by mutual information) were analyzed at single recording and population levels. For normal-hearing animals, the neuronal responses decreased as a function of the number of the vocoder frequency bands, so did their discriminative capacities at the single recording level. However, small neuronal populations were found to be robust to the degradations induced by the vocoder. Similar robustness was obtained when broadband noise was added to exacerbate further the spectro-temporal distortions produced by the vocoder. A comparable pattern of robustness to degradations in fine spectro-temporal details was found for hearing-impaired animals. However, the latter showed an overall decrease in neuronal discrimination capacities between vocalizations in noisy conditions. Consistent with previous studies, these results demonstrate that the primary auditory cortex maintains robust neural representation of temporal envelope features for communication sounds under a large</p>	

range of spectro-temporal degradations.



Orsay, November 9th, 2017

Dear Editor in Chief, Dear Paul, Dear Associate Editor,

Please find on the website a revised version of our MS JARO-D-17-00039R1 that we re-submit for publication to JARO as research article.

We have taken into account the reviewers comments and have changed our MS accordingly. All the changes are in RED in the new MS for the version named "Aushana et al 09112017\_RED" and we also provide a clean version in black "Aushana et al 09112017\_B&W".

Best regards,

Yonane Aushana, Samira Souffi, Jean-Marc Edeline, Christian Lorenzi, Chloé Huetz

**Reviewer #1: The manuscript is very much improved. I am reasonably happy that they caveat their results reasonably.**

I still have a slight issue with the terminology of 'discrimination' - which is used throughout the manuscript with respect to the mutual-information analysis. My worry is that this term implies perceptual discrimination ability, and that the limitations of the stimulus - the small number and obvious differences in basic things like onset latency - make it difficult to generalise to perceptual discrimination of vocalisations. It would be nice to caveat that further.

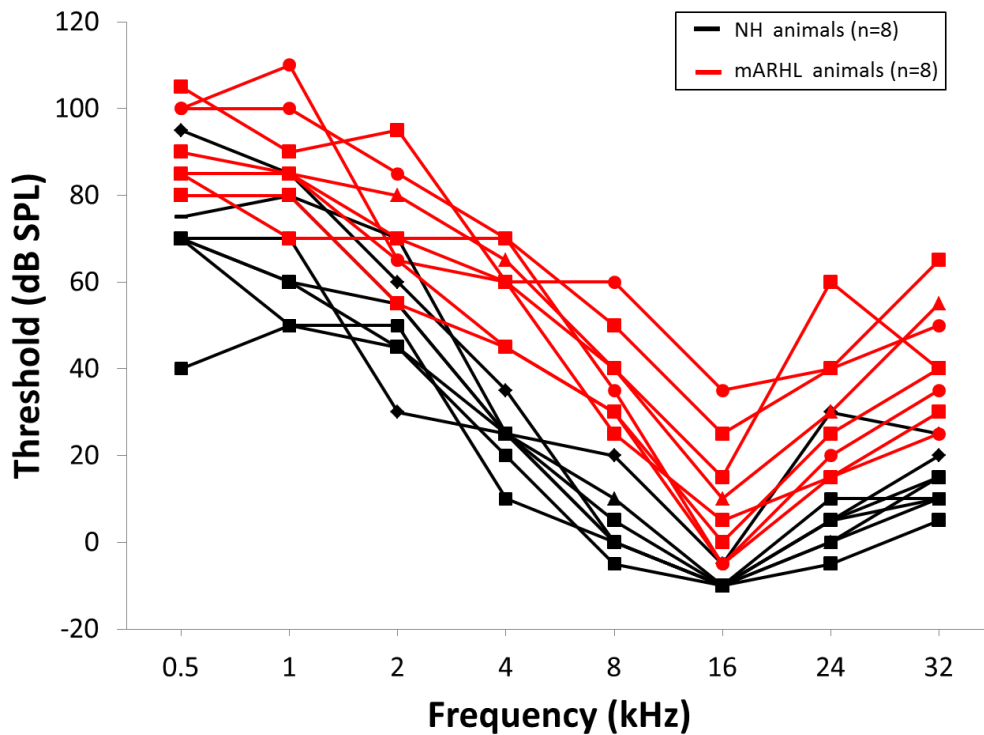
We agree with the reviewer and we clarified what should be understood by the word "discrimination". We have added a small paragraph on page 19 to clarify this point. In brief, this term was only used to point out that primary auditory cortex neurons are able to display different responses, in terms of temporal spike patterns, at presentation of set of stimuli. This term could be considered as ambiguous so we want to make it clear that our results do not suggest that guinea pigs discriminate between the original or the vocoded vocalizations.

Animal groupings are improved but still not clear.

P6L44 "Based on the ABR, two groups of animals were separated". Yet the only criteria given for splitting is age. It is clear from figure 2A that there were differences in ABR thresholds between the groups, but no indication of how ABR was a basis for separation. That could only be the case here if there were no overlap at all in the ABR thresholds of individuals in the two groups - is this the case? If so, and ABR was indeed part of the separating into groups then that criteria should be specified. I would guess that in practice they split the groups on age, and there was a group difference in the ABRs. That is OK - but they need to be clear about it. Also, this would amount to an uncontrolled confound of hearing loss age - which needs discussing.

If there IS any overlap in the ABRs then it would be interesting if there were guinea-pigs of different ages with similar audiograms.

The reviewer is correct: We have used animals of two ages and when looking at their ABR, they can be classified as normal hearing vs. aged-related hearing-loss. We show in the supplementary figure below that there is very little overlap between the two groups. We specified this point in the method section (page 6) and would like to add the figure as supplementary material.



Audiograms of the young normal-hearing (NH, black lines) and mild age-related hearing loss (mARHL, red lines) animals.

Typos

P14 L10 audiogram -> audiograms

P19 L46 "Unknown is what will be" grammar.

The typos were corrected, thank you.

**Reviewer #2: The revision addresses all of my concerns from the original submission.**

Minor concerns:

Title: The word "or" in the title should be "and", since it is a comparison.

The title was corrected, thank you.

Abstract: First sentence: The word "young" should be inserted before "normal hearing" and the word "aged" (or some variant) should be inserted before "hearing-impaired".

Introduction, page 3, line 32: "i.e., a frequency modulated (FM) carrier" is not correct in general, and especially so in this project, where the synthesized carrier is constant frequency.

These two corrections were made, thank you.

[Click here to view linked References](#)

1  
2  
3 *Robust neuronal discrimination in primary auditory cortex despite degradations of*  
4 *spectro-temporal acoustic details: comparison between guinea pigs with normal hearing*  
5 *and mild age-related hearing loss*  
6  
7  
8  
9

10 Yonane Aushana <sup>1,2</sup>, Samira Souffi <sup>1,2</sup>, Jean-Marc Edeline <sup>1,2</sup>  
11 Christian Lorenzi <sup>3</sup>, Chloé Huetz<sup>1,2</sup>  
12  
13

14 Paris-Saclay Institute of Neurosciences (Neuro-PSI)  
15 Department Cognition and Behavior  
16  
17 <sup>1</sup> CNRS UMR 9197, <sup>2</sup> Université Paris-Sud, Bâtiment 446,  
18 91405 Orsay cedex, France  
19  
20

21 <sup>3</sup> Laboratoire des systèmes Perceptifs, UMR CNRS 8248,  
22 Département d'Etudes Cognitives, Ecole normale supérieure (ENS),  
23 Paris Sciences & Lettres Research University, 75005 Paris, France  
24  
25

26 Emails:

27 Yonane Aushana: [yonane.aushana@u-psud.fr](mailto:yonane.aushana@u-psud.fr)

28 Samira Souffi: [samira.souffi@u-psud.fr](mailto:samira.souffi@u-psud.fr)

29 Jean-Marc Edeline: [jean-marc.edeline@u-psud.fr](mailto:jean-marc.edeline@u-psud.fr)

30 Christian Lorenzi: [christian.lorenzi@ens.fr](mailto:christian.lorenzi@ens.fr)

31 Chloé Huetz: [chloe.huetz@u-psud.fr](mailto:chloe.huetz@u-psud.fr)  
32  
33

34 Running title: Responses to vocoded vocalizations in AI  
35

36 Total number of pages: 22

37 Total number of figures: 6

38 Total number of words: abstract: 214; introduction: 1186; results: 1838; discussion: 2001  
39  
40

41 Correspondence to:

42 Jean-Marc Edeline

43 Department Cognition & Behavior

44 Paris-Saclay Institute of Neuroscience (Neuro-PSI)

45 UMR CNRS 9197

46 Université Paris-Sud, Bâtiment 446,

47 91405 Orsay cedex, France

48 Phone: (+33) 1 69 15 49 72

49 Fax: (+33) 1 69 15 77 26  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## Abstract

This study investigated to which extent the primary auditory cortex of young normal-hearing and mild hearing-impaired aged animals is able to maintain invariant representation of critical temporal-modulation features when sounds are submitted to degradations of fine spectro-temporal acoustic details. This was achieved by recording ensemble of cortical responses to conspecific vocalizations in guinea pigs with either normal hearing or mild age-related sensorineural hearing loss. The vocalizations were degraded using a tone vocoder. The neuronal responses and their discrimination capacities (estimated by mutual information) were analyzed at single recording and population levels. For normal-hearing animals, the neuronal responses decreased as a function of the number of the vocoder frequency bands, so did their discriminative capacities at the single recording level. However, small neuronal populations were found to be robust to the degradations induced by the vocoder. Similar robustness was obtained when broadband noise was added to exacerbate further the spectro-temporal distortions produced by the vocoder. A comparable pattern of robustness to degradations in fine spectro-temporal details was found for hearing-impaired animals. However, the latter showed an overall decrease in neuronal discrimination capacities between vocalizations in noisy conditions. Consistent with previous studies, these results demonstrate that the primary auditory cortex maintains robust neural representation of temporal envelope features for communication sounds under a large range of spectro-temporal degradations.

Key words: vocoder, envelope, fine structure, electrophysiology, auditory cortex, neural discrimination performance, spike-timing.



## Introduction

Over the last few decades, a wealth of psychophysical studies conducted with young and older normal-hearing and hearing-impaired human subjects have explored the mechanisms of *robust* speech perception, that is the ability to maintain high levels of speech recognition despite sometimes dramatic degradations of the speech signal caused by interfering sound sources such as noise, reverberation, or distortions caused by the transmission channel (e.g., the speech processor of a cochlear implant). Research efforts have mostly concentrated on the search for the spectro-temporal acoustic cues used by human subjects, and the neural codes responsible for maintaining invariant neural representations of the speech signal. A large part of this investigation was achieved using vocoders (Dudley 1939), that is signal-processing devices designed to degrade selectively specific acoustic features of the speech signal (e.g., Drullman 1995; Shannon et al. 1995; Friesen et al. 2001; Smith et al. 2002; Nelson et al. 2003; Baskent 2006; Zeng et al. 2005; Lorenzi et al. 2006; Sheft et al. 2008; for a recent review, see Shamma and Lorenzi 2013).

Vocoders decompose incoming sounds into frequency bands mimicking the spectral decomposition performed by the cochlea. The resulting bandpass filtered signals are - in turn - decomposed into a temporal-envelope component (i.e., an amplitude modulator, AM) and a temporal-fine-structure component (i.e., a frequency-modulated (FM) sinusoid). For each band, the temporal-fine-structure component is degraded (by replacing it by a pure tone or a band of noise) and then amplitude modulated by the corresponding (intact or modified) temporal-envelope component. The resulting AM carriers are finally added up and presented for discrimination or identification to human subjects.

Overall, these behavioral studies showed repeatedly that relatively slow (< 16 Hz) AM cues within a limited number of frequency bands (4 to 8) yield nearly perfect speech identification performance in quiet (e.g., Drullman 1995; Shannon et al. 1995; Smith et al. 2002). However, faster AM cues, FM cues and fine spectral cues were shown to be required for accurate speech identification in steady or fluctuating noise, when speech is presented against a competing speech signal or when speech is temporally interrupted (e.g., Friesen et al. 2001; Gnansia et al. 2009, 2010; Zeng et al. 2005). These findings have been interpreted as indicating that AM cues (the temporal envelope) convey important phonetic information (that is, timbre information) whereas FM and fine spectral cues (the acoustic fine structure) convey “grouping” or “segregation” cues, that is cues used to separate speech from interfering sound sources and integrate partial speech information (e.g., Gnansia et al. 2010; Apoux and Healy 2011; Apoux et al. 2013). One possibility to provide evidence for the relative roles of AM cues, FM cues and fine spectral cues in speech intelligibility can be to

1 obtain neural data in conditions where some of these cues are present (AM cues) while others are  
2 absent (FM cues). As a first attempt, the primary auditory cortex, where neurons are more sensitive  
3 to slow than to fast AM modulations (Liegeois-Chauvel et al. 2004), appears to be a structure which  
4 deserves in-depth investigation.  
5

6  
7 Very few studies have evaluated the neuronal responses to vocoded stimuli either in human  
8 or in animal. In human, magnetoencephalographic data indicate that the cortical entrainment is  
9 robust to natural and vocoded stimuli (up to 4 spectral channels) but it is largely reduced when  
10 vocoded stimuli are presented in background noise (Ding et al. 2014). In animals, only three  
11 electrophysiological studies have assessed the neural responses to vocoded speech or animal  
12 vocalizations (Nagarajan et al. 2002; Ranasinghe et al. 2012; Ter-Mikaelian et al. 2013). Altogether,  
13 these three studies found that temporal patterns of auditory cortical responses are extremely robust  
14 to reduction in the number of frequency bands, suggesting that the temporal code plays a critical  
15 role in the robust representation of communication signals cues. For instance, Nagarajan et al. (2002)  
16 found that patterns of neural discharges did not differ considerably as a function of number of  
17 frequency bands when the number of bands was varied between 2 and 16.  
18  
19

20  
21 Elderly persons often experience important difficulties understanding speech in adverse  
22 listening situations (e.g., Duquesnoy and Plomp 1980; Takahashi and Bacon 1992), sometimes in the  
23 absence of elevated audiometric thresholds (for recent reviews, see Füllgrabe et al. 2014; Schoof and  
24 Rosen 2014). One possible explanation for this is that aging per se alters supra-threshold auditory  
25 processing (for a review, see Fitzgibbons and Gordon-Salant 2010). Consistent with this idea, older  
26 listeners showed deficits in using complex AM patterns to identify speech accurately (Souza and Kitch  
27 2001; Souza and Boike 2006; Schvartz et al. 2008; Sheldon et al. 2008). However, in contrast with  
28 these studies, recent work by Schoof and Rosen (2014) and by Paraouty and Lorenzi (2017) did not  
29 find any effect of aging on AM sensitivity. Moreover, additional studies have found that sensitivity to  
30 FM cues (He et al. 2007; Grose and Mamo 2012; Wallaert et al. 2016) and binaural temporal fine  
31 structure cues decline with increasing age (Ross et al. 2007; Grose and Mamo 2010; Hopkins and  
32 Moore 2011; Moore et al. 2012a,b; Füllgrabe 2013; King et al. 2014; for a review, see Grose et al.  
33 2015). On the other hand, several studies indicate that frequency selectivity does not change with  
34 age provided that audiometric thresholds remain normal (Lutman et al. 1991; Peters et al. 1992;  
35 Hopkins and Moore 2011). These findings are consistent with neurophysiological data suggesting  
36 that aging impairs the processing of neural envelope (Frisina and Frisina 1997; Walton et al. 1998)  
37 and temporal fine structure information (Ross et al. 2007; Clinard et al. 2010) in the auditory  
38 midbrain and cortex. However, this is at odds with the higher and more precise MEG responses  
39 obtained in mild age-related hearing loss human subjects (Presacco et al. 2016; Alain et al. 2014), and  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1 also with the relatively subtle effects of aging on the spectral and temporal tuning characteristics of  
2 the cortical neurons of macaques (Engle and Recanzone 2013; Overton and Recanzone 2016).

3 Here, we investigated the cortical representation of unprocessed and vocoded conspecific  
4 vocalizations in young guinea pigs with normal hearing (NH) and older guinea pigs exhibiting mild  
5 age-related hearing loss (mARHL). Neuronal responses were recorded in the primary auditory cortex  
6 (ACx) of these guinea pigs during presentation of whistles (the alarm call for this species) left as such  
7 (unprocessed) or processed via a tone-excited vocoder using 10, 20 or 38 frequency bands. We opted  
8 for a tone vocoder rather than a noise vocoder because the latter strongly distort the signal's  
9 temporal envelope characteristics (Whitmall et al. 2007) and noise carriers introduce (non-  
10 informative) random envelope fluctuations interfering with signal envelopes (Stone et al. 2011).

11 The data showed that despite a large decrease in response strength, both the temporal reliability and  
12 the discrimination performance based on temporal spike patterns were relatively robust to the  
13 vocoding process. Importantly, this held true irrespective of age and presence of mild hearing loss.  
14 Still, discrimination performance was more affected for mARHL than for NH animals when  
15 background noise (added after the spectral degradation) was added to the vocoded stimuli. These  
16 data may provide neurophysiological bases for the difficulties in understanding speech in adverse  
17 listening situations typically reported by elderly people.  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## Materials and Methods

### Subjects

These experiments were performed under the national license A-91-557 (project 2014-25, authorization 05202.02) and using the procedures N° 32-2011 and 34-2012 validated by the Ethic committee N°59 (CEEA Paris Centre et Sud). All surgical procedures were performed in accordance with the guidelines established by the European Communities Council Directive (2010/63/EU Council Directive Decree).

Recordings were obtained in the primary ACx of 16 adult pigmented guinea pigs aged 6 months to 2 years old. Animals weighting from 450 g to 1230 g came from our own colony housed in a humidity (50-55%) and temperature (22-24° C)-controlled facility on a 12 h/12 h light/dark cycle (light on at 7:30 A.M.) with free access to food and water. None of the animals used here have been exposed to noise, except for the noise inherent to the colony room (air conditioning and the animals' own vocalizations; level estimated to 60 dB SPL).

Two to three days before each experiment, the animal's audiogram was determined by testing auditory brainstem responses (ABR) under isoflurane anaesthesia (2.5 %) as described in Gourévitch et al. (2009). The ABR was obtained by differential recordings between two subdermal electrodes (SC25-NeuroService) placed at the vertex and behind the mastoid bone. A software (RTLab, Echodia, Clermont-Ferrand, France) allowed averaging 500 responses during the presentation of nine pure-tone frequencies (between 0.5 and 32 kHz) delivered by a speaker (Knowles Electronics) placed in the animal right ear. The auditory threshold of each ABR is the lowest intensity where a small ABR wave can still be detected (usually wave III). For each frequency, the threshold was determined by gradually decreasing the sound intensity (from 80 dB down to -10 dB SPL).

Based upon ABR and the age of animals, two groups of animals were separated: half of the animals (8/16), aged 6 to 9 months were considered to have a normal audiogram which did not differ from the one published in previous studies using young guinea pigs (Gourévitch et al. 2009; Gourévitch and Edeline 2011). The other half (aged of at least 2 years old) displayed mild, but significant ( $p < 0.01$ ), age-related hearing loss which was present across the entire audible bandwidth (the "hearing spectrum") of the guinea pig (see the supplementary figure for more details).

### Surgical procedures

The animal was anesthetized by an initial injection of urethane (1.2 g/kg, i.p.) supplemented by additional doses of urethane (0.5 g/kg, i.p.) when reflex movements were observed after pinching

1 the hind paw (usually 2-4 times during the recording session). A single dose of atropine sulphate  
2 (0.06mg/kg, s.c.) was given to reduce bronchial secretions and a small dose of buprenorphine was  
3 administrated (0.05mg/kg, s.c.) as urethane has no analgic properties. After placing the animal in a  
4 stereotaxic frame, a craniotomy was performed above the left temporal cortex. A local anesthetic  
5 (Xylocain 2%) was liberally injected in the wound. The opening was 8 mm wide starting at the  
6 intersection point between parietal and temporal bones and 8-10 mm height. The dura above the  
7 auditory cortex was removed under binocular control and the cerebrospinal fluid was drained  
8 through the cysterna to prevent the occurrence of oedema. After the surgery, a pedestal in dental  
9 acrylic cement was built to allow an atraumatic fixation of the animal's head during the recording  
10 session. The stereotaxic frame supporting the animal was placed in a sound-attenuating chamber  
11 (IAC, model AC1). At the end of the recording session, a lethal dose of pentobarbital (>200 mg/kg,  
12 i.p.) was administered to the animal.  
13  
14  
15  
16  
17  
18  
19  
20  
21

## 22 **Recording procedures**

23  
24 Data were from multiunits recordings collected in the primary auditory cortex (area AI). Extracellular  
25 recordings were obtained from arrays of 16 tungsten electrodes ( $\phi$ : 33  $\mu$ m, <1 M $\Omega$ ) composed of two  
26 rows of 8 electrodes separated by 1000  $\mu$ m (350  $\mu$ m between electrodes of the same row). A silver  
27 wire, used as ground, was inserted between the temporal bone and the dura matter on the  
28 contralateral side. The location of the primary auditory cortex was estimated based on the pattern of  
29 vasculature observed in previous studies (Edeline and Weinberger 1993; Manunta and Edeline 1999;  
30 Edeline et al. 2001; Wallace et al. 2000). The raw signal was amplified 10,000 times (TDT Medusa). It  
31 was then processed by an RX5 multichannel data acquisition system (TDT). The signal collected from  
32 each electrode was filtered (610-10000 Hz) to extract multi-unit activity (MUA). The trigger level was  
33 set for each electrode to select the largest action potentials from the signal. On-line and off-line  
34 examination of the waveforms suggests that the MUA collected here was made of action potentials  
35 generated by 2 to 6 neurons at the vicinity of the electrode. For each experiment, we set the position  
36 of the electrode array in such a way that the two rows of eight electrodes sample neurons  
37 responding from low to high frequency when progressing in the rostro-caudal direction (see  
38 examples of tonotopic gradients recorded with such arrays in Figure 1 of Gaucher et al. 2012 and in  
39 Figure 6A of Occelli et al. 2016).  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

## 53 **Acoustic stimuli**

54  
55 Acoustic stimuli were generated in Matlab, transferred to a RP2.1-based sound delivery system (TDT)  
56 and sent to a Fostex speaker (FE87E). The speaker was placed at 2 cm from the guinea pig's right ear,  
57  
58  
59  
60  
61  
62  
63  
64  
65

1 a distance at which the speaker produced a flat spectrum ( $\pm 3$  dB) between 140 Hz and 36 kHz.  
2 Calibration of the speaker was made using noise and pure tones recorded by a Bruel & Kjaer  
3 microphone 4133 coupled to a preamplifier B&K 2169 and a digital recorder Marantz PMD671.  
4  
5 We first determined spectro-temporal receptive fields (STRFs) using 97 or 129 pure-tones  
6  
7 frequencies scaled with a gamma function, covering six (0.14-9 kHz or 0.28-18 kHz or 0.56-36 kHz) or  
8  
9 eight (0.14-36 kHz) octaves respectively, and presented at three levels: 85 or 75 dB, 65 or 55 dB, 45  
10 or 35 dB SPL. At a given level, each frequency was repeated eight times at a rate of 2.35 Hz in  
11  
12 pseudorandom order. The duration of these tones over half-peak amplitude was 15 ms and the total  
13  
14 duration of the tone was 50 ms, so there was no overlap between tones. A set of four conspecific  
15  
16 vocalizations was used to assess the neuronal response to communication sounds. These  
17  
18 vocalizations were recorded from animals of our colony. Pairs of animals were placed in the acoustic  
19  
20 chamber and their vocalizations were recorded by a Bruel & Kjaer microphone 4133 coupled to a  
21  
22 preamplifier B&K 2169 and a digital recorder Marantz PMD671. A large set of whistle calls was  
23  
24 loaded in the Audition software (Adobe Audition 3) and four representative examples of whistle  
25  
26 were selected. As shown in Figure 1A (bottom), despite the fact the maximal energy of the four  
27  
28 selected whistle was in the same frequency range (typically between 4 and 26 kHz), they displayed  
29  
30 marked differences in their spectrograms. In addition, their temporal envelopes clearly differed as  
31  
32 shown on their oscillograms (figure 1A top). The four selected whistles were processed by three tone  
33  
34 vocoders as previously described (Gnansia et al. 2009, 2010). The three vocoders only differed in  
35  
36 terms of the number of bands used to decompose (i.e., filter) the whistles (38, 20 or 10 bands). In the  
37  
38 following, the 38 bands are briefly presented, but the same applies for the 20-bands or the 10-bands  
39  
40 vocoders. Each digitized signal was passed through a bank of 38 *fourth-order* Gammatones  
41  
42 (Patterson 1987) filters with center frequencies uniformly spaced along a Guinea-Pig adapted ERB  
43  
44 (Equivalent Rectangular Bandwidth) scale (Sayles and Winter 2010) ranging from 50 to 35505 Hz. In  
45  
46 each frequency band, the temporal envelope was extracted using full-wave rectification and lowpass  
47  
48 filtering at 64 Hz with a zero-phase, sixth-order Butterworth filter. The resulting envelopes were used  
49  
50 to amplitude modulate sine-wave carriers with frequencies at the center frequency of the  
51  
52 Gammatone filters, and with random starting phase. Impulse responses were peak-aligned for the  
53  
54 envelope (using a group delay of 16 ms) and the temporal fine structure across frequency channels  
55  
56 (Hohmann 2002). The modulated signals were finally weighted and summed over the 38 frequency  
57  
58 bands. The weighting compensated for imperfect superposition of the bands' impulse responses at  
59  
60 the desired group delay. The weights were optimized numerically to achieve a flat frequency  
61  
62 response (Hohmann 2002). Figure 1 displays the waveforms (first row) and the spectrograms (second  
63  
64 row) of the four whistles together with the spectrograms of the 38-bands vocoded (third row), the  
65  
20-bands vocoded (fourth row) and the 10-bands vocoded (fifth row) whistles. In the following

1 figures, the unprocessed whistles will be referred as to the original, and the vocoded versions as to  
2 Voc38 (Voc20, Voc10 respectively) for the 38-bands (20-bands, 10-bands respectively) vocoded  
3 whistles. It is important to note that the broadband noise added to the stimuli was not processed  
4 through the vocoder. Broadband noise was added to the vocoded vocalizations, that is *after*  
5 vocoding. Our aim was to introduce additional spectro-temporal distortions and degrade further the  
6 acoustic differences between vocalizations. Last, in contrast to previous cortical studies that all have  
7 used noise-vocoders (Nagarajan et al. 2002; Ranasinghe et al. 2012; Ter-Mikaelian et al. 2013), a tone  
8 vocoder was used here, because noise vocoders were found to introduce random (non-informative)  
9 intrinsic temporal-envelope fluctuations distorting the signal's AM features (Stone et al. 2011;  
10 Shamma and Lorenzi 2013).

## 11 **Experimental protocol**

12 Inserting an array of 16 electrodes in the cortical tissue almost systematically induces a deformation  
13 of the cortex. A 30-minutes recovering time lap was allowed for the cortex to return to its initial  
14 shape, then the array was slowly lowered. Tests based on measures of spectro-temporal receptive  
15 fields (STRFs) were used to assess the quality of our recordings and to adjust electrodes' depth. The  
16 recording depth was 500-1000  $\mu\text{m}$ , which corresponds to layer III and the upper part of layer IV  
17 according to Wallace and Palmer (2008). However, as a result of the convex shape of the guinea-pig  
18 ACx, it is likely that all the electrodes were not exactly at the same depth.

19 When a clear tuning was obtained for at least 8 of the 16 electrodes, the stability of the tuning was  
20 assessed: we required that the recorded neurons displayed at least three successive similar STRFs  
21 (each lasting 6 minutes) before starting the protocol. When the stability was satisfying, the protocol  
22 started by presenting the acoustic stimuli in the following order: We first presented the 4 whistles at  
23 75 dB SPL in their natural versions, followed by the vocoded versions with 38, 20 and 10 bands. The  
24 same set of whistle in their natural and vocoded versions were then presented in the presence of  
25 continuous white noise background set at 65 dB SPL (this white noise was added after the spectral  
26 degradation produced by the vocoding, and refreshed from trial to trial). In all cases, each  
27 vocalization was repeated 20 times. Presentation of this entire stimulus set lasted 30 minutes.

## 28 **Data analysis**

### 29 *Quantification of responses to pure tones*

30 The STRFs derived from MUA were obtained by constructing post-stimulus time histograms for each  
31 frequency with 1 ms time bins. The firing rate evoked by each frequency was quantified by summing  
32 all the action potentials from the tone onset up to 100 ms after this onset. Thus, STRFs are matrices

of 100 bins in abscissa (time) multiplied by 97 or 129 bins in ordinate (frequency). All STRFs were smoothed with a uniform 5x5 bin window.

For each STRF, the Best Frequency (BF) was defined as the frequency where the highest firing rate was recorded. Peaks of significant response were automatically identified using the following procedure: A positive peak in the MU-based STRF was defined as a contour of firing rate above the average level of the baseline activity (estimated from the ten first milliseconds of STRFs at all intensity levels) plus six times the standard deviation of the baseline activity. For each electrode, to determine the  $Q_{20dB}$  and  $Q_{40dB}$ , we estimated the lowest intensity for which significant STRFs were observed. At this intensity, the BF was considered to be the Characteristic Frequency (CF). At 20 dB and 40 dB above this intensity, the low and high frequency limits of the STRFs were determined to compute the  $Q_{20dB}$  and  $Q_{40dB}$ .

The upper and lower frequency bounds of the STRF were also used to count the number of frequency bands of the vocoding stimuli that were falling within the STRF. For this analysis, only the 10-bands vocoder was considered. The possible values ranged from 0 to 4 bands falling into the neurons' STRF.

#### *Quantification of responses evoked by vocalizations*

The responses to vocalizations were quantified using two parameters: (i) The firing rate of the evoked response, which corresponds to the total number of action potentials occurring during the presentation of the stimulus, divided by its duration; (ii) the spike-timing reliability coefficient (CorrCoef) which quantifies the trial-to-trial reliability of the response. This index was computed for each vocalization: it corresponds to the normalized covariance between each pair of spike trains recorded at presentation of this vocalization and was calculated as follows:

$$CorrCoef = \frac{1}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\sigma x_i x_j}{\sigma x_i \sigma x_j}$$

where N is the number of trials and  $\sigma x_i x_j$  is the normalized covariance at zero lag between spike trains  $x_i$  and  $x_j$  where i and j are the trial numbers. Spike trains  $x_i$  and  $x_j$  were previously convolved with a 10-msec width Gaussian window. Based upon computer simulations, we have previously shown that this CorrCoef index is not a function of the neurons' firing rate (see Gaucher et al. 2013a).

As in previous studies (e.g. Gaucher and Edeline 2015), we set up a criterion of an evoked firing rate of at least 1.5 spike.sec<sup>-1</sup> and a CorrCoef of at least 0.15 to discard recordings that did not respond to any vocalization. A systematic and meticulous examination of our database has shown that recordings exhibiting only a phasic onset response generated a CorrCoef of at least 0.15, and that cells responding by tonic responses lacking of temporal organization generated at least a firing rate



of 1.5 spikes.sec<sup>-1</sup>. Only the recordings that matched this criterion for all four whistles in their original and vocoded versions were classified as “responsive” and were included in the present study.

#### *Quantification of mutual information from the responses to vocalizations.*

The method developed by Schnupp et al. (2006) was used to quantify the amount of information (Shannon 1948) contained in the responses to vocalizations obtained with natural and vocoded stimuli. This method allows quantifying how well the vocalization’s identity can be inferred from neuronal responses. Here, “neuronal responses” refers either to (i) the spike trains obtained from a small group of neurons below one electrode (for the computation of the individual Mutual Information, MI), or to (ii) a concatenation of spike trains simultaneously recorded under several electrodes (for the computation of the population MI). In both cases, the following computation steps were the same. Neuronal responses were represented using different time scales ranging from the duration of the whole response (firing rate) to a 1-ms precision (precise temporal patterns), which allows analyzing how much the spike timing contributes to the information. As this method is exhaustively described in Schnupp et al. (2006) and in Gaucher et al. (2013a), we present below the main principles only.

The method relies on a pattern-recognition algorithm that is designed to “guess which stimulus evoked a particular response pattern” by going through the following steps: From all the responses of a cortical site to the different stimuli, a single response (test pattern) is extracted and represented as a PSTH with a given bin size (different sizes were considered as discussed further below). Then, a mean response pattern is computed from the remaining responses (training set) for each stimulus class. The test pattern is then assigned to the stimulus class of the closest mean response pattern. This operation is repeated for all the responses, generating a confusion matrix where each response is assigned to a given stimulus class. From this confusion matrix, the Mutual Information (MI) is given by Shannon’s formula:

$$MI = \sum_{x,y} p(x,y) \times \log_2 \left( \frac{p(x,y)}{p(x) \times p(y)} \right)$$

where x and y are the rows and columns of the confusion matrix, or in other words, the values taken by the random variables “presented stimulus class” and “assigned stimulus class”.

In our case, we used responses to the 4 whistles and selected the first 264 ms of these responses to work on spike trains of exactly the same duration (the shortest whistle being 264-msec long). In a scenario where the responses do not carry information, the assignments of each response to a mean response pattern is equivalent to chance level (here 0.25 because we used 4 different stimuli and each stimulus was presented the same number of times) and the MI would be close to zero. In the opposite case, when responses are very different between stimulus classes and very similar within a

1 stimulus class, the confusion matrix would be diagonal and the mutual information would tend to  
2  $\log_2(4) = 2$  bits.

3 As already mentioned, this algorithm was applied with different bin sizes ranging from 264 ms to 1  
4 ms. For each cortical site, two parameters were extracted from this analysis. First, the value of MI for  
5 the largest bin size (264 ms) which represents the MI value when only the firing rate is considered  
6 ( $MI_{\text{Firing Rate}}$ ). Second, the MI value obtained for a bin of 8ms,  $MI_{\text{Patterns}}$ , which is a temporal resolution  
7 leading to the highest values of information in previous studies (Schnupp et al. 2006; Huetz et al.  
8 2009).

9 The MI estimates are subject to non-negligible positive sampling biases. Therefore, as in Schnupp et  
10 al. (2006), we estimated the expected size of this bias by calculating MI values for “shuffled” data, in  
11 which the response patterns were randomly reassigned to stimulus classes. The shuffling was  
12 repeated 100 times, resulting in 100 MI estimates of the bias ( $MI_{\text{bias}}$ ). These  $MI_{\text{bias}}$  estimates are then  
13 used as estimators for the computation of the statistical significance of the MI estimate for the real  
14 (unshuffled) datasets: the real estimate is considered as significant if its value is statistically different  
15 from the distribution of  $MI_{\text{bias}}$  shuffled estimates. Significant MI estimates were computed for MI  
16 calculated from neuronal responses under one electrode.

17 The information carried out by a group of recording was estimated by the population MI ( $MI_{\text{Population}}$ ),  
18 using the same method as for the  $MI_{\text{Patterns}}$ : responses of several simultaneous recordings were  
19 grouped together and considered as a single pattern. To assess the influence of the group size of  
20 simultaneous recordings on the information carried by that group ( $MI_{\text{Population}}$ ), the number of sites  
21 used for computing  $MI_{\text{Population}}$  varied from 2 to the maximal possible size (which is equal to 16 minus  
22 the non-responsive sites). As the number of possible combinations could be extremely large ( $C_n^k$ ,  
23 where  $k$  is the group size and  $n$  the number of responsive sites in a recording session), a threshold  
24 was fixed to spare computation time: when the number of possible combinations exceeded one  
25 hundred, 100 combinations were randomly chosen, and the mean of all combinations was taken as  
26 the  $MI_{\text{Population}}$  for this group size.

27 Since our electrode arrays can only achieve up to 16 simultaneous recordings (but often less than 12  
28 recordings responding to vocalizations), pseudo-populations (population of neurons which were not  
29 simultaneously recorded) were generated as described in the method developed by Ince et al (2013).  
30 Briefly, we sampled responsive recordings from our database to generate pseudo-populations of 2-  
31 25 recordings and compute the  $MI_{\text{population}}$  for 500 random samplings. These values of  $MI_{\text{pseudo-population}}$   
32 were computed for the original and vocoded versions of the vocalizations (38, 20 and 10).  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

### *Statistics*

To assess the significance of the multiple comparisons (vocoding process: four levels; the groups of animals: NH vs. mARHL; two conditions: quiet vs. background noise), we used an analysis of variance (ANOVA) for multiple factors to analyze the whole data set. Follow-up tests were corrected for multiple comparisons using Bonferroni corrections and were considered as significant if their p value was below 0.05. All data are presented as mean values  $\pm$  standard error (s.e.m.).

## Results

Initially, 880 stable multiunit recordings were obtained from sixteen guinea pigs tested with the original and vocoded whistles. Out of them, 581 recordings were classified as “responsive” (see Methods) and used in the subsequent analyses. As explained in the Methods, half of the animals (8/16, 6-9 months old) had normal audiograms (Figure 2A, NH, black line) which did not differ from the one published in previous studies for guinea pigs of the same age (Gourévitch et al. 2009; Gourévitch and Edeline 2011). The other half (about 2 years old) displayed mild, but significant ( $F(1,14)= 3.12$ ;  $p<0.01$ ), age-related hearing loss (Figure 2A, mARHL, red line), which was present for across the entire hearing spectrum. Analyzing the cortical STRF obtained from normal-hearing (NH) and mild age-related hearing loss (mARHL) animals revealed that the  $Q_{20dB}$  and  $Q_{40dB}$  values were lowered in the mARHL animals compared to the NH animals (Fig 2B, respectively  $F(1,157)=1.76$  and  $F(1,105)=1.37$   $p<0.05$  in both cases), indicating that hearing loss impacted on the width of frequency tuning in auditory cortex. In the following, we will first focus on the effect of vocoding on the cortical responses in NH animals ( $n=285$  recordings). Then, the influence of background noise on the processing of original and vocoded vocalizations will be presented for both the 285 recordings obtained on NH and also for a set of 296 recordings obtained in mARHL animals.

### Effect of vocoding on the responses in the ACx of normal-hearing (NH) animals

Figure 3 shows the responses of 6 simultaneous recordings at presentation of the original and vocoded whistles. Compared with the responses to the original whistles (top row), a strong decrease in response strength (for the four whistles) was observed as the number of frequency bands in the vocoder was reduced. For this set of recordings, the temporal reproducibility (as measured by the CorrCoef) also decreased (from 0.45 for the original vocalization to 0.32 for the 10-band vocalizations).

Over the set of 285 recordings obtained on these animals, vocoding reduced the evoked firing rate and the temporal reliability (Figure 4A and 4B, one way-ANOVA,  $F(3,1136)=3.36$  for firing rate and  $F(3,1136)=2.91$  for CorrCoef,  $p<0.05$  for both). Interestingly, post-hoc pairwise comparisons showed that only the 10-bands vocoded whistles elicited significantly lower responses than the original whistles (Fig 4A, Bonferroni corrected pairwise t-test,  $p=0.02$ ). To assess how the reduction in the number of frequency bands altered the neural discrimination, Mutual Information (MI) was computed either from the firing rate or from the temporal patterns (see Methods). As shown in previous studies (Schnupp et al. 2006; Gaucher et al. 2013a; Liu and Schreiner 2007; Engineer et al. 2008), temporal spike patterns allowed a much better discrimination between stimuli than the firing

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

rate did: the  $MI_{\text{Patterns}}$  values (Figure 4D) are about twice the one of  $MI_{\text{Firing Rate}}$  (Figure 4C). These two measures were impacted by the vocoding (ANOVA,  $F(3,515)=3.11$ ,  $p=0.03$  for  $MI_{\text{Firing Rate}}$ ,  $F(3,865)=5.24$ ,  $p=0.001$  for  $MI_{\text{Patterns}}$ ). However, the vocoding process only induced a small impairment of the neurons coding abilities: the MI computed on temporal patterns was only reduced by  $\sim 0.2$  bits for the vocoded 10 bands stimuli compared to the original stimuli (Fig 4D; note that a perfect discrimination between the four whistles would result in a MI equals to 2 bits).

### **Vocoding affected the cortical responses independently of the neurons' receptive field**

To determine if the vocoding process impacts similarly cortical neurons independently of their Best Frequency (BF), the recordings were split in four groups according to their BF: BF below 5 kHz ( $n=100$ ), between 5 kHz and 10 kHz ( $n=61$ ), between 10 kHz and 15 kHz ( $n=26$ ), and above 15 kHz ( $n=52$ ; note that 46 recordings had no clear BF but responded to vocalizations). A two-way ANOVA showed that there was no interaction between the BF factor and the vocoding factor for the 3 parameters quantifying the neuronal responses (the firing rate, the temporal reliability and the mutual information;  $p>0.9$  in all cases), indicating that the vocoding did not differentially impact neurons with low, middle and high frequency BF. This was not totally surprising given that our vocoded stimuli still contained energy in a wide range of frequencies. However, it could be expected that recordings whose frequency receptive field (RF) spanned none of the vocoder's frequency bands would be more impacted by the vocoding process. To investigate this issue, the recordings were classified according to the number of bands of the vocoded stimuli that fell into their RFs at 75dB. A global ANOVA indicated that all measures of neuronal activity (firing rate, temporal reliability and mutual information) were impacted by the number of bands falling into the recording's RF (number of bands factor,  $p<0.001$  for Firing Rate, CorrCoef,  $MI_{\text{Patterns}}$ ). However, there was no interaction between the "number of bands in RF" factor and the vocoding factor (two-way ANOVA interaction vocoding\*number of bands,  $p>0.7$  for Firing rate, CorrCoef and  $MI_{\text{Patterns}}$ ), indicating that the recordings were similarly impacted by the spectral degradation of the stimulus.

### **Robust coding of vocalizations by small neuronal populations in ACx for normal-hearing animals**

As shown above, the parameters quantifying the responses of individual recordings were impacted by the stimulus spectral degradations generated by the vocoder. The next objective was to determine if this effect also impacted the coding abilities at the population level. To explore this issue, mutual information was computed based on the temporal patterns obtained from two to sixteen simultaneously recordings. Computed on a population, the MI quantifies how well the four whistles can be discriminated based on temporal patterns expressed by this population. Figure 5A

1 presents the value of the  $MI_{\text{Population}}$  as a function of the number of recordings. Each line represents a  
2 given set of simultaneous recordings, from which N recordings (N varying from 2 to 16) were selected  
3 using a bootstrap procedure to compute a mean MI value for N recordings (see Methods). During  
4 most of the recording sessions, not all 16 electrodes gave significant responses, and, as a  
5 consequence, most of the lines do not reach the right side of figure 5A (corresponding to 16  
6 simultaneous recordings). Strikingly, for the original and vocoded stimuli, the maximum value of  
7 information (2 bits) was reached by many populations of 16 simultaneous recordings. Therefore,  
8 when MI values were computed for populations of 12 individual recordings (figure 5B, pink line), the  
9 mean value was close to maximum and there was no difference between the four groups of stimuli  
10 (ANOVA,  $F(3,35)=0.9$ ,  $p=0.4$ ), indicating that even when the stimuli were strongly degraded,  
11 populations of 12 simultaneous recordings were, on average, sufficient to discriminate between the  
12 four whistles. However, when populations of only 8 simultaneous neurons were considered, the  
13 vocoding process resulted in a significant decrease of the MI carried by the neuronal responses (fig  
14 5B, orange line, ANOVA,  $F(3,81)=2.82$ ,  $p<0.05$ ).

15 As only a few simultaneous recordings were available for 12 neurons, a pseudo-population analysis  
16 was performed (see Methods). Pseudo-populations were constructed from all responsive neurons,  
17 regardless of whether or not they were simultaneously recorded (Ince et al., 2013). This analysis  
18 allowed creating populations up to 25 neurons. Figure 5C shows the values of  $MI_{\text{Pseudo-population}}$  for  
19 original (black), 38-band vocoded (blue), 20-band vocoded (green) and 10-band vocoded (red)  
20 responses as a function of the number of neurons included in the pseudo-population. It is obvious  
21 from Figure 5C that, for populations made of small numbers of neurons (less than 12), the  $MI_{\text{Pseudo-}}$   
22  $population$  decreases as a function of the vocoder's number of bands. Strikingly, when 25 neurons are  
23 included in the population (shaded grey area on Figure 5C), all curves almost reach 2 bits (the  
24 maximum possible MI value, see grey line on Figure 5D). This suggests that, even with the 10 band  
25 vocoded stimuli, 25 neurons are sufficient to perfectly discriminate the four whistles (but see  
26 Discussion for the limitations).

### 27 **Impact of mild age-related hearing loss on neural responses to vocoded stimuli in quiet and noisy** 28 **conditions**

29 The results presented above show that the discrimination between vocalizations performed by ACx  
30 neurons is relatively robust to the vocoding process: Despite a significant decrease in evoked activity,  
31 temporal reliability and mutual information (figure 4), the decrease in neural performance is modest  
32 considering how important stimulus degradations are. Moreover, a perfect discrimination between  
33 stimuli can still be achieved at the population level (figure 5). The following section now explores to  
34 which extent the above conclusions also hold for hearing-impaired animals and in noisy conditions.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
As explained at the beginning of the Results section, the ABRs for eight animals (two years old aged) revealed a mild but significant Age-Related Hearing Loss (mARHL, figure 2). The cortical neurons obtained from these animals (which displayed slightly broader frequency tuning curves) were tested with the same set of stimuli (original and vocoded whistles). The results are displayed in Figure 6A, in red lines for mARHL animals, and in black lines for NH animals. Recordings from mARHL animals showed higher firing rate and temporal reliability compared to the recordings obtained for NH animals (three-way ANOVA,  $F(1,4447)=11$  for firing rate,  $F(1,4447)=14.6$  for CorrCoef,  $p<0.001$ ). Also, recordings obtained from mARHL animals showed discriminative abilities at the level of individual recordings or neuronal populations comparable to those observed for NH animals ( $F(1,3358)=0.41$ ,  $p=0.5$  for  $MI_{Patterns}$ ,  $F(1,161)=1.86$ ,  $p=0.18$  for  $MI_{Population}$ ).

17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
To further test the robustness of cortical coding in NH and mARHL animals, a broadband background noise was added to the original and vocoded whistles (see Methods). The same parameters quantifying neural activity (firing rate, temporal reliability,  $MI_{Patterns}$  and  $MI_{Population}$ ) were measured in the noisy condition. Figure 6B shows the difference between noisy and quiet conditions for NH (grey lines) and mARHL (light red lines) animals. As expected, all points are below zero, showing that the addition of background noise reduced the magnitude of all these parameters: the firing rate and temporal reliability decreased (three-way ANOVA,  $F(1,4447)=14$  for firing rate,  $F(1,4447)=349$  for CorrCoef,  $p<0.001$ ), as well as the MI computed at the individual and population level (three-way ANOVA,  $F(1,3358)=152$  for  $MI_{Patterns}$ ,  $F(1,161)=37$  for  $MI_{Population}$ ,  $p<0.001$ ). However, for all the measures derived from the neuronal responses, there was no interaction between noise and vocoding (noise \* vocoding,  $p>0.6$  for all measures) suggesting that for NH subjects, the vocoder and background noise did not alter differentially the cortical responses. Interestingly, the effect of background noise was stronger in the recordings collected from mARHL animals compared to those collected in NH animals (red lines are often below the grey ones) both for the temporal reliability and the  $MI_{Patterns}$  (three-way ANOVA, interaction noise \* hearing loss,  $F(1,4447)=5$  for CorrCoef,  $F(1,3358)=16$  for  $MI_{Patterns}$ ,  $p<0.03$ ). It was not the case for the  $MI_{Population}$  when computed with 12-neuron populations, but when computed with 8 neurons, the interaction became significant (three-way ANOVA, interaction noise \* hearing loss,  $F(1,161)=3$ ,  $p=0.08$  for  $MI_{Population(\#neurons=12)}$ ;  $F(1,327)=7$ ,  $p<0.01$  for  $MI_{Population(\#neurons=8)}$ ). However, as in NH animals, there was no interaction between noise and vocoding for the mARHL animals (noise \* vocoding,  $p>0.6$  for all measures,  $R^2<1.6e-4$ ). Together, these last two results suggest that hearing loss amplifies the impact of noise on cortical responses to whistles whereas it has no influence on the impact of vocoding.

## Discussion

### Brief overview of the results

The responses of auditory-cortex neurons were found to be altered when tested with spectrally degraded vocalizations (in comparison with responses obtained with the original vocalizations). Nevertheless, these alterations were not drastic and were prominent only when the vocoding process was applied to 10 broad frequency bands. For normal-hearing animals, the firing rate and the spike timing reliability decreased, which in turn reduced the discrimination ability of cortical cells at the single recording level. However, when at least 12 neurons were considered, an almost perfect discrimination between the four vocalizations was reached and stayed unaffected by the vocoding process. Data collected on animals exhibiting mild age-related hearing loss showed similar patterns in quiet conditions. In the presence of background noise, the discriminative abilities of cortical cells were more reduced in mild age-related hearing loss animals than in normal-hearing animals, suggesting that the noise impacted more the cortical processing of vocoded vocalizations for the mild age-related hearing loss than for the normal hearing animals. Before discussing the functional implications of these results, several methodological points should be discussed.

### Methodological considerations

First, the data presented here are based on the responses of small groups of neurons (2-5) located at the vicinity of each recording electrodes and we did not perform spike sorting analyses to separate the different units. We consider that this do not invalidate our findings because our recordings technique and conditions allowed detecting relatively subtle effects in previous studies when manipulating either the acoustic properties of the stimuli (Ocelli et al. 2016) or pharmacological agents (Gaucher et al. 2013a; Gaucher and Edeline 2015).

Second, it could be argued that the use of anesthetics can impact the potential effect induced by the vocoding process on cortical responses. Some studies have reported important differences between anesthetized and unanesthetized animals: for example, the maximum rate of click trains that primary ACx neuron is able to follow (Rennaker et al. 2007), the percentage of bursts in the evoked responses (Massaux et al. 2003, 2004) and the evoked oscillations triggered by acoustic stimuli (Cotillon-Williams and Edeline 2003). However, other studies have reported comparable responses in awake and anesthetized rats for consonant sounds (Engineer et al. 2008) and conspecific vocalizations in guinea pig (Huetz et al. 2009).

Third, the hearing loss displayed by our mild age-related hearing loss animals (20 dB) is modest and not comparable with what was reported in a previous study (Gourévitch and Edeline 2011). Most



1 likely, this is the direct consequence of the fact that animals used in our previous study were much  
2 older (3.5 to 4 years in Gourévitch and Edeline 2011 vs. 2 years here). As a consequence, the cortical  
3 responses obtained here in our mild age-related hearing loss animals are similar to the responses  
4 obtained in normal-hearing animals, which can explain why the vocoding process impacts similarly  
5 the two groups of animals.  
6

7  
8 MI was used here as a metric to quantify the differences between neuronal responses to the four  
9 original, or the four vocoded vocalizations. It provides an objective measure to quantify to what  
10 extent neuronal responses differ between each other. The MI value can be interpreted as a way to  
11 evaluate the discrimination ability of auditory cortical neurons when tested with the original, then  
12 with the vocoded stimuli. Obviously, our use of the term “discrimination” does not refer to  
13 perceptual abilities. Only behavioral experiments, in humans or animals, can give access to  
14 perceptual abilities. To what extent the discriminative abilities of neurons recorded in a given cortical  
15 area have causal relationships with the whole perceptual ability is the subject of a vast literature, and  
16 is clearly out of the scope of the present study.  
17

18 Here, the values of MI obtained with the original vocalizations on single recordings are relatively low  
19 (below 1bit) even when computed in terms of spike patterns (the values of  $MI_{\text{firing rate}}$  were more than  
20 twice lower than the values of  $MI_{\text{patterns}}$ ). This can be explained by the fact we used four versions of  
21 the same type of vocalization which had roughly the same spectrograms (even if they were  
22 acoustically different). It is only when computing the MI values on populations of simultaneous  
23 recordings that an almost perfect discrimination could be achieved (MI value close to 2 in our case).  
24 However, our population analysis was suffering from small data sets, due to the fact that only a few  
25 numbers of simultaneous recording sets were available. To increase the number of neuronal  
26 populations, analyses were performed following the same method as in Ince et al. (2013). A large  
27 number of pseudo-populations (500) was generated from our data set (population of neurons not  
28 simultaneously recorded) and simulations were conducted to investigate the impact of vocoding on  
29 the discriminative abilities of these pseudo-populations. These analyses confirmed that when 25  
30 neurons are considered, an almost perfect discrimination is possible by ACx neurons, thus suggesting  
31 that the discriminative capabilities of cortical cells are robust to the spectral degradations produced  
32 on our stimuli set.  
33

34 One of the main limitations of our study is the small number of stimuli used to test the discriminative  
35 performance. Previous studies using larger sets of stimuli have found lower MI values (Gaucher et al.  
36 2013a, 2015; Chechik et al. 2006). Thus, one can wonder whether using a larger set of stimuli should  
37 allow us to reach similar conclusions. The parametric study performed by Ince et al. (2013) indicates  
38 that even with larger set of stimuli (up to 50), it is still possible to reach the maximal value of MI  
39 (indexing an almost perfect discrimination) with large neuronal populations. In our experimental  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

1 conditions, if a larger set of stimuli is used, the effects of the vocoding process on the cortical  
2 discrimination remain unknown.  
3  
4

### 5 **Relationship with human studies** 6

7 Vocoders have long been used to characterize the key acoustic features responsible for robust  
8 speech perception (Dudley 1939). This literature has largely documented that slow AM fluctuations  
9 (<16 Hz) in a limited number of frequency bands (4-8) is sufficient to maintain an almost perfect  
10 identification of speech in quiet (Shannon et al. 1995; Smith et al. 2002). By showing here that  
11 auditory-cortex neurons are very robust in discriminating between vocoded communication sounds,  
12 our results clearly point out that the primary auditory cortex can send relevant (i.e., invariant) cues  
13 to downstream brain areas more specialized in speech identification (Belin et al. 2000; Romanski and  
14 Averbeck 2009).  
15  
16  
17  
18  
19  
20

21 Elderly persons often experience important difficulties understanding speech in adverse  
22 listening situations sometimes in the absence of elevated audiometric thresholds (for recent reviews,  
23 see Füllgrabe et al. 2014; Schoof and Rosen 2014). Potentially, this can result from an alteration of  
24 supra-threshold auditory processing (for a review, see Fitzgibbons and Gordon-Salant 2010), which  
25 can be explained if high-threshold auditory-nerve fibers are the first to be impacted during aging  
26 (Sergeyenko et al. 2013). Initially, it was found that older listeners showed deficits in using complex  
27 AM patterns to identify speech accurately (Souza and Kitch 2001; Souza and Boike 2006; Schwartz et  
28 al. 2008; Sheldon et al. 2008). This is possibly due to reduced sensitivity to AM cues or reduced  
29 central ability to make optimal use of AM cues (e.g. Takahashi and Bacon 1992; He et al. 2008;  
30 Füllgrabe et al. 2015). However, several studies did not find any effect of aging on AM sensitivity  
31 (Schoof and Rosen 2014; Paraouty et al. 2016; Paraouty and Lorenzi 2017), suggesting that the  
32 processing of AM information is roughly comparable for younger and older listeners exhibiting  
33 similar audiometric thresholds. When elderly subjects show sensorineural hearing loss, AM  
34 sensitivity is generally improved compared to normal (Wallaert et al. 2017), presumably because of  
35 the loss in the fast-acting amplitude compression applied by outer hair cells in the cochlea (Moore et  
36 al. 1996). Thus, the potentially detrimental effects of aging on AM processing may be confounded  
37 with (and counter-balanced by) the loss of compression associated with mild hearing loss. This may  
38 explain why previous MEG studies found enhanced cortical responses to speech for older subjects  
39 with mild hearing loss compared to young and older subjects with normal hearing (Soros et al. 2009;  
40 Alain et al. 2014; Presacco et al. 2016). More specifically, the synchronization of MEG responses with  
41 the speech envelope was more precise for older subjects both in quiet and in noise (Presacco et al.  
42 2016). Other studies found higher amplitude for the waves P1 (Heschl's gyrus) and N1 (Planum  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

temporale) in hearing impaired older subjects than in normal hearing subjects (Alain et al. 2014; Soros et al. 2009). This may also explain the increase in evoked firing rate and temporal reliability observed in our cortical recordings from animals with mild aged-related hearing loss.

### **Relationship with previous electrophysiological studies at the cortical level**

As a general rule, previous cortical studies reported that the responses of cortical neurons to communication sounds are particularly robust to spectral alterations produced by vocoders, even when the vocoders only incorporated 4 or 2 bands. This was the case in the marmoset (Nagarajan et al. 2002) and gerbil (Ter-Mikaelian et al. 2013) auditory cortex at presentation of conspecific vocalizations; this was also the case in the rat auditory cortex at presentation of vocoded consonants and vowels (Ranasinghe et al. 2012). Consistent with these findings, the present study showed that the degradation produced by the vocoder had little impact on cortical responses. Even when 10-band vocoded stimuli were presented, the decrease in MI at the level of individual recordings was modest (0.2 bit on average, see Fig 4). Only one of the previous studies (Nagarajan et al. 2002) has evaluated the addition of broadband noise on the conspecific vocalizations and found little effect of masking noise up to a signal-to-noise ratio of +10 dB (but note that the vocoded vocalizations were not presented in the background noise in their study). This is consistent with other findings showing robust representation of speech in the primary auditory cortex in the presence of white/pink noise and in the presence of convolutive reverberations (Mesgarani et al. 2014).

So far, none of the previous cortical studies have evaluated to what extent age-related hearing loss can potentiate the degradation produced by the vocoding process. In the present study, we found that in the case of mild age-related hearing loss (about 20 dB), this hearing loss did not contribute much to the alterations of responses produced by the vocoder (Fig 6A). However, it seems that in noisy conditions, the discrimination capabilities of cortex neurons were slightly lower in mild age-related hearing loss animals compared with normal ones (Fig 6B). Obviously this effect should be confirmed by investigating more in depth the effects of other types of background noise and testing several levels of signal-to-noise ratio.

### **What are the acoustic features triggering cortical responses to original and vocoded vocalizations?**

As shown here, populations of cortical neurons are able to discriminate between vocalizations that have been severely degraded by vocoding. What are the acoustic cues allowing for this robust discrimination? As a matter of fact, it is now well documented that auditory cortex neurons do not

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

respond to vocalizations based on their spectral content. Several studies have clearly showed that the cortical responses to natural stimuli cannot be predicted based upon the spectral content of the vocalizations (e.g. Machens et al. 2004; Laudanski et al. 2012; Bitterman et al. 2008; Chechik & Nelken 2012; review in Huetz et al. 2011; Gaucher et al. 2013b); even responses from inferior colliculus neurons cannot be predicted from spectral content (Woolley et al. 2006). In contrast, auditory cortex neurons are highly sensitive to the slope of changes in sound level (Fishbach et al. 2001; Heil 1997a,b) and as a consequence are sensitive to the temporal envelope of natural stimuli. Thus, one explanation for the present results is that despite the severe degradations in fine spectro-temporal details produced by vocoding, auditory cortex neurons can still discriminate between the 4 whistles based upon the differences in gross temporal envelope cues in each frequency band. One can envision that even if a much larger set of stimuli was used, a good cortical discrimination between vocoded stimuli is still possible provided that these stimuli differ in terms of gross temporal-envelope cues.

## Conclusions and perspectives

The present study shows that both for normal-hearing and hearing-impaired animals, auditory cortex neurons can discriminate between spectrally similar communication sounds in the absence of relevant temporal fine-structure cues (i.e. FM cues and fine spectral details). These results suggest that the critical cue for discriminating such sounds corresponds to the relatively slow temporal-envelope fluctuations (AM) of these sounds within a limited number of frequency bands. More severe conditions of stimuli presentation and degree of hearing loss should be tested in future work before generalizing these results. Also, it will be of great interest to study other cortical areas to determine to which extent the effects reported in AI impact neuronal responses in belt, parabelt and downstream cortical areas.

## Acknowledgements

This work was supported by grants from the National Research Agency (ANR Blanc2011, HearFin and ANR Blanc2014, Heart) to CL and JME. This work was also supported by ANR-11-0001-02 PSL\* and ANR-10-LABX-0087. SS was supported by the Fondation pour la Recherche Médicale (FRM grant number ECO20160736099 to [SS]). Special thanks to Fabien Lhericel and Céline Dubois for taking care of the guinea-pig colony and to Dr. Grégoire Levasseur for helping generating high quality figures.

## Figure legends

### Figure 1: Original and vocoded whistles.

A. First row shows the time-waveform of the four whistle stimuli. Their spectrogram is plotted on the second row.

B. The three rows represent the spectrogram of the vocoded version of the four whistle stimuli for each vocoder processing. The first row uses 38 bands, the second 20 bands, the third 10 bands.

### Figure 2: Brainstem and cortical thresholds for normal and mild age-related hearing loss animals.

A. Thresholds (in dB SPL) for the Auditory Brainstem Responses (ABRs) of the Normal Hearing (NH, in dark) and mild Age-Related Hearing Loss (mARHL, in red) animals. For each group, the group data are represented by the thick line with the SEM and the data from each animal are represented by the thin lighter lines.

B. Mean values ( $\pm$ sem) of the Q20dB and Q40dB of cortical neurons estimated in NH and mARHL animals. The number of recordings in each group is indicated on the bars.

### Figure 3: Rasterplots for 6-neuron recordings in response to reference and vocoded versions of each whistle.

First row show the 6-neuron recordings in response to the original whistles (different whistles on each column). For each subplot, individual neuron's APs are either in red or blue to distinguish between different neurons (1 to 6). Shaded area corresponds to spontaneous activity before stimulus presentation. Second to fourth row show the responses to the vocoded whistles (38 to 10 bands). Last row shows the waveforms of the whistles.

Mean values of Firing rate (Fr) and CorrCoef (averaged over the 6 neurons and the 4 stimuli) are presented at the end of each row.

### Figure 4: group data for normal-hearing animals.

Each subplot represents the mean (and standard error of the mean) value of a measure of neuronal activity for original whistles and vocoded versions (Vocoded 38, 20, 10 bands). A. Firing rate, B.

CorrCoef, C.  $MI_{\text{Firing Rate}}$ , D.  $MI_{\text{Patterns}}$

Stars represent a significant difference between original and vocoded (38,20 or 10), as computed with post-hoc pairwise tests corrected for multiple comparisons with the Bonferroni method.

### Figure 5: Mutual Information (MI) computed for a population of recordings.

1  
2  
3  
4  
5  
6  
A. MI as a function of the size of the population. X-axis represents the number of recordings in the population. Each curve is a given recording (maximum 16 simultaneous recordings). A MI value of 2 corresponds to perfect discrimination of whistles. Shaded colored bars (orange and pink) refer to populations of 8 or 12 simultaneous recordings (see B.).

7  
8  
9  
B. Mean values of  $MI_{Population}$  computed with 12 (pink line) or 8 (orange line) recordings for each set of stimuli (original and vocoded whistles with 38, 20, and 10 bands).

10  
11  
12  
13  
14  
15  
C.  $MI_{pseudo-population}$  as a function of the size of the population. Same axes as in A. Red, green, blue and black curves represent the different stimuli (Original, Voc38, Voc20, Voc10).  $MI_{pseudo-population}$  values for 8 (orange), 12 (pink) and 25 (grey) recordings were extracted to be plotted in D.

16  
17  
18  
19  
20  
21  
22  
D. Mean values of  $MI_{pseudo-population}$  computed with 8 (orange line), 12 (pink line), or 25 (grey line) recordings for each set of stimuli (original whistles, vocoded 38, 20, 10 bands).

### 23 **Figure 6: Effects of mild Age-Related hearing loss and background noise on vocoding**

24  
25  
26  
27  
28  
29  
30  
A. Average neural activity measure (from left to right: firing rate, CorrCoef,  $MI_{Pattern}$ ,  $MI_{Population}$ ) computed from the responses to original whistles and vocoded versions (Vocoded 38, 20, 10 bands) for normal hearing animals (NH, black lines), and mild Age-Related Hearing Loss animals (mARHL, red lines).

31  
32  
33  
34  
35  
36  
37  
38  
B. Mean differences between responses to noisy and quiet whistles. For each recording, the neural activity measure (from left to right:  $\Delta$ firing rate,  $\Delta$ CorrCoef,  $\Delta$  $MI_{Pattern}$ ,  $\Delta$  $MI_{Population}$ ) evoked in the quiet condition is subtracted from its value evoked by the noisy condition. Means of the differences are plotted in grey for NH animals, and in light red for mARHL animals.

39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
Bars represent the standard error of the mean.

## References

- 1  
2  
3  
4  
5 Alain C, Roze A, Salloum C (2014) Effects of age-related hearing loss and background noise on neuromagnetic  
6 activity from auditory cortex. *Front Syst Neurosci* 8:8.
- 7 Apoux F, Yoho SE, Youngdahl CL, Healy EW (2013) Role and relative contribution of temporal envelope and fine  
8 structure cues in sentence recognition by normal-hearing listeners. *J Acoust Soc Am* 134(3):2205-12.  
9 doi: 10.1121/1.4816413.
- 10 Apoux F, Healy EW (2011) Relative contribution of target and masker temporal fine structure to the unmasking  
11 of consonants in noise. *J Acoust Soc Am* 130(6):4044-52. doi: 10.1121/1.3652888.
- 12 Baskent D (2006) Speech recognition in normal hearing and sensorineural hearing loss as a function of the  
13 number of spectral channels. *J Acoust Soc Am* 120:2908–2925.
- 14 Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000) Voice-selective areas in human auditory cortex. *Nature*  
15 403(6767):309-12.
- 16 Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I (2008) Ultra-fine frequency tuning revealed in single  
17 neurons of human auditory cortex. *Nature* 451(7175):197-201. doi:10.1038/nature06476.
- 18 Chechik G, Nelken I (2012) Auditory abstraction from spectro-temporal features to coding auditory entities.  
19 *Proc Natl Acad Sci U S A* 109(46):18968-73. doi: 10.1073/pnas.1111242109.
- 20 Chechik G, Anderson MJ, Bar-Yosef O, Young ED, Tishby N, Nelken I (2006) Reduction of information  
21 redundancy in the ascending auditory pathway. *Neuron* 51(3):359-68.  
22 doi:10.1016/j.neuron.2006.06.030.
- 23 Clinard CG, Tremblay KL, Krishnan AR (2010) Aging alters the perception and physiological representation of  
24 frequency: evidence from human frequency-following response recordings. *Hear Res* 264(1-2):48-55.  
25 doi: 10.1016/j.heares.2009.11.010.
- 26 Cotillon-Williams N, Edeline JM (2003) Evoked oscillations in the thalamo-cortical auditory system are present  
27 in anesthetized but not in unanesthetized rats. *Journal of Neurophysiology* 89:1968–1984.
- 28 Ding N; Chatterjee M; Simon JZ (2014) Robust cortical entrainment to the speech envelope relies on the  
29 spectro-temporal fine structure. [Academic Journal]
- 30 Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. *J Acoust Soc Am*  
31 97:585-592.
- 32 Dudley H (1939) Remaking speech. *J Acoust Soc Am* 11:169–177.
- 33 Duquesnoy AJ, Plomp R (1980) Effect of reverberation and noise on the intelligibility of sentences in cases of  
34 presbycusis. *J Acoust Soc Am* 68(2):537-44.
- 35 Edeline JM, Dutilleul G, Manunta Y, Hennevin E (2001) Diversity of receptive field changes in auditory cortex  
36 during natural sleep. *Eur J Neurosci* 14:1865–1880. doi: 10.1046/j.0953-816x.2001.01821.x.
- 37 Edeline JM, Weinberger NM (1993) Receptive field plasticity in the auditory cortex during frequency  
38 discrimination training: selective retuning independent of task difficulty. *Behav Neurosci* 107:82–103.  
39 doi: 10.1037/0735-7044.107.1.82.
- 40 Engineer CT, Perez CA, Chen YH, Carraway RS, Reed AC, Shetake JA, Jakkamsetti V, Chang KQ, Kilgard MP (2008)  
41 Cortical activity patterns predict speech discrimination ability. *Nat Neurosci* 11(5):603-8. doi:  
42 10.1038/nn.2109.
- 43 Engle JR, Recanzone GH (2013) Characterizing spatial tuning functions of neurons in the auditory cortex of  
44 young and aged monkeys: a new perspective on old data. *Front Aging Neurosci* 4:36  
45 doi:10.3389/fnagi.2012.00036.
- 46 Fishbach A, Nelken I, Yeshurun Y (2001) Auditory edge detection: a neural model for physiological and  
47 psychoacoustical responses to amplitude transients. *J Neurophysiol.*;85(6):2303-23.
- 48 Fitzgibbons PJ, Gordon-Salant S (2010) Age-related differences in discrimination of temporal intervals in  
49 accented tone sequences. *Hear Res* 264(1-2):41-7. doi:10.1016/j.heares.2009.11.008.
- 50 Friesen LM, Shannon RV, Baskent D, Wang X (2001) Speech recognition in noise as a function of the number of  
51 spectral channels: Comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am* 110:1150–  
52 1163.
- 53 Frisina DR, Frisina RD (1997) Speech recognition in noise and presbycusis: relations to possible neural  
54 mechanisms. *Hear Res* 106(1-2):95-104.
- 55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

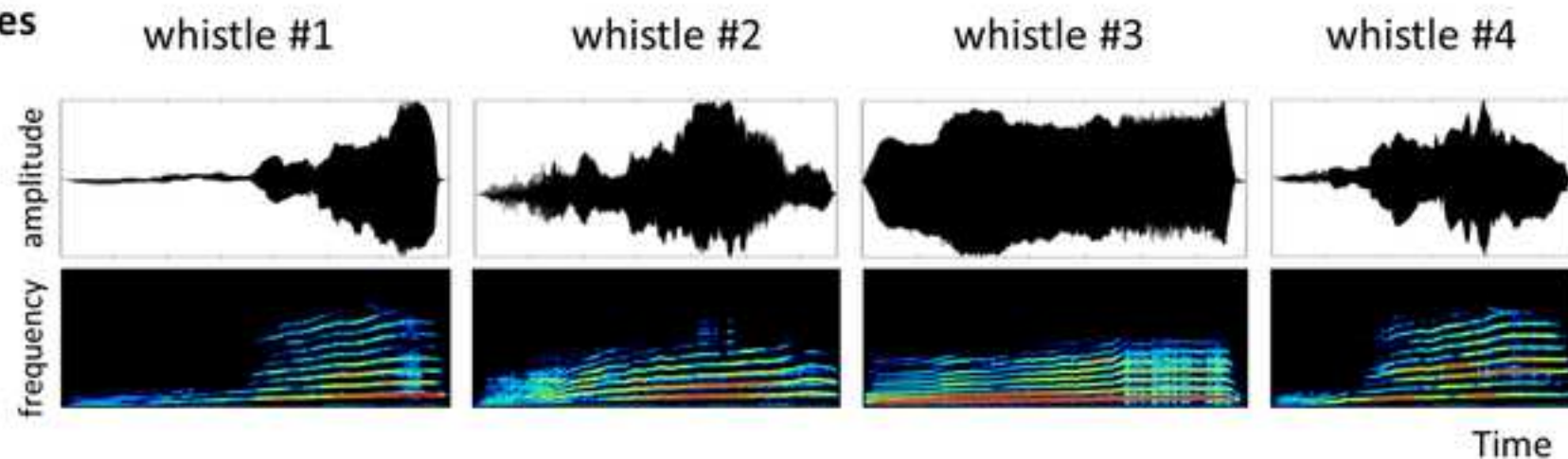
- 1 Füllgrabe C, Moore BC, Stone MA (2015) Age-group differences in speech identification despite matched  
2 audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front*  
3 *Aging Neurosci* 6:347. doi: 10.3389/fnagi.2014.00347.
- 4 Füllgrabe C, Moore BC (2014) Effects of age and hearing loss on stream segregation based on interaural time  
5 differences. *J Acoust Soc Am* 136(2):EL185-91. doi: 10.1121/1.4890201.
- 6 Füllgrabe C (2013) Age-dependent changes in temporal-fine-structure processing in the absence of peripheral  
7 hearing loss. *Am J Audiol* 22(2):313-5. doi: 10.1044/1059-0889(2013/12-0070).
- 8 Gaucher Q, Edeline JM (2015) Stimulus-specific effects of noradrenaline in auditory cortex: Implications for the  
9 discrimination of communication sounds. *Journal of Physiology (London)* 593.4, 1003-1020.
- 10 Gaucher Q, Huetz C, Gourévitch B, Edeline J-M (2013a) Cortical inhibition reduces information redundancy at  
11 presentation of communication sounds in the primary auditory cortex. *Journal of Neuroscience* 33(26):  
12 10713-10728.
- 13 Gaucher Q, Huetz C, Gourévitch B, Laudanski J, Occelli F, Edeline JM (2013b) How do auditory cortex neurons  
14 represent communication sounds ? *Hear Res* 305:102-112.
- 15 Gaucher Q, Edeline JM, Gourévitch B (2012) How different are the local field potentials and spiking activities ?  
16 Insights from multi-electrodes arrays. *J Physiol Paris* 106(3-4):93-103. doi:  
17 10.1016/j.jphysparis.2011.09.006.
- 18 Gnansia D, Pressnitzer D, Pean V, Meyer B, Lorenzi C (2010) Intelligibility of interrupted and interleaved speech  
19 for normal-hearing listeners and cochlear implantees. *Hear Res* 6108.
- 20 Gnansia D, Pean V, Meyer B, Lorenzi C (2009) Effects of spectral smearing and temporal fine structure  
21 degradation on speech masking release. *J Acoust Soc Am* 125:4023-4033.
- 22 Gourévitch B, Edeline JM (2011) Age-related changes in the guinea pig auditory cortex: relationship with  
23 peripheral changes and comparison with tone-induced hearing loss. *Eur J Neurosci* 34:1953–1965. doi:  
24 10.1111/j.1460-9568.2011.07905.x.
- 25 Gourévitch B, Doisy T, Avillac M, Edeline JM (2009) Follow-up of latency and threshold shifts of auditory  
26 brainstem responses after single and interrupted acoustic trauma in guinea pig. *Brain Res* 1304:66–79.  
27 doi: 10.1016/j.brainres.2009.09.041.
- 28 Grose JH, Mamo SK, Buss E, Hall JW (2015). Temporal Processing Deficits in Middle Age. *Am J Audiol* 24(2):91-3.  
29 doi: 10.1044/2015\_AJA-14-0053.
- 30 Grose JH, Mamo SK (2012) Frequency modulation detection as a measure of temporal processing: age-related  
31 monaural and binaural effects. *Hear Res* 294(1-2):49-54. doi:10.1016/j.heares.2012.09.007.
- 32 Grose JH, Mamo SK (2010) Processing of temporal fine structure as a function of age. *Ear Hear* (6):755-60. doi:  
33 10.1097/AUD.0b013e3181e627e7.
- 34 He NJ, Mills JH, Ahlstrom JB, Dubno JR (2008) Age-related differences in the temporal modulation transfer  
35 function with pure-tone carriers. *J Acoust Soc Am* 124(6):3841-9. doi: 10.1121/1.2998779.
- 36 He NJ, Mills JH, Dubno JR (2007) Frequency modulation detection: Effects of age, psychophysical method, and  
37 modulation waveform. *J Acoust Soc Am* 122:467–477.
- 38 Heil P (1997b) Auditory cortical onset responses revisited. II. Response strength. *J Neurophysiol.*;77(5):2642-  
39 60.
- 40 Heil P (1997a) Auditory cortical onset responses revisited. I. First-spike timing. *J Neurophysiol.*;77(5):2616-41.  
41
- 42 Hohmann V (2002) Frequency analysis and synthesis using a Gammatone filterbank. *Acust Acta Acust* 88: 433-  
43 442.
- 44 Hopkins K, Moore BC (2011) The effects of age and cochlear hearing loss on temporal fine structure sensitivity,  
45 frequency selectivity, and speech reception in noise. *J Acoust Soc Am* 130(1):334-49. doi:  
46 10.1121/1.3585848.
- 47 Huetz C, Gourévitch B, Edeline J-M (2011) Neural codes in the thalamocortical auditory system: from artificial  
48 stimuli to communication sounds. *Hear Res* 271:147-158.
- 49 Huetz C, Philibert B, Edeline JM (2009) A spike-timing code for discriminating conspecific vocalizations in the  
50 thalamocortical system of anesthetized and awake guinea pigs. *J Neurosci* 29:334–350. doi:  
51 10.1523/JNEUROSCI.3269-08.2009.
- 52 Ince RAA, Panzeri S, Kayser C (2013) Neural Codes Formed by Small and Temporally Precise Populations in  
53 Auditory Cortex. *J Neurosci.* 33(46):18277-87. doi: 10.1523/JNEUROSCI.2631-13.2013.
- 54 King A, Hopkins K, Plack CJ (2014) The effects of age and hearing loss on interaural phase difference  
55 discrimination. *J Acoust Soc Am* 135:342–351.
- 56 Laudanski J, Edeline JM, Huetz C (2012) Differences between spectro-temporal receptive fields derived from  
57 artificial and natural stimuli in the auditory cortex. *PLoS One* ;7(11):e50539. doi:  
58 10.1371/journal.pone.0050539.
- 59  
60  
61  
62  
63  
64  
65



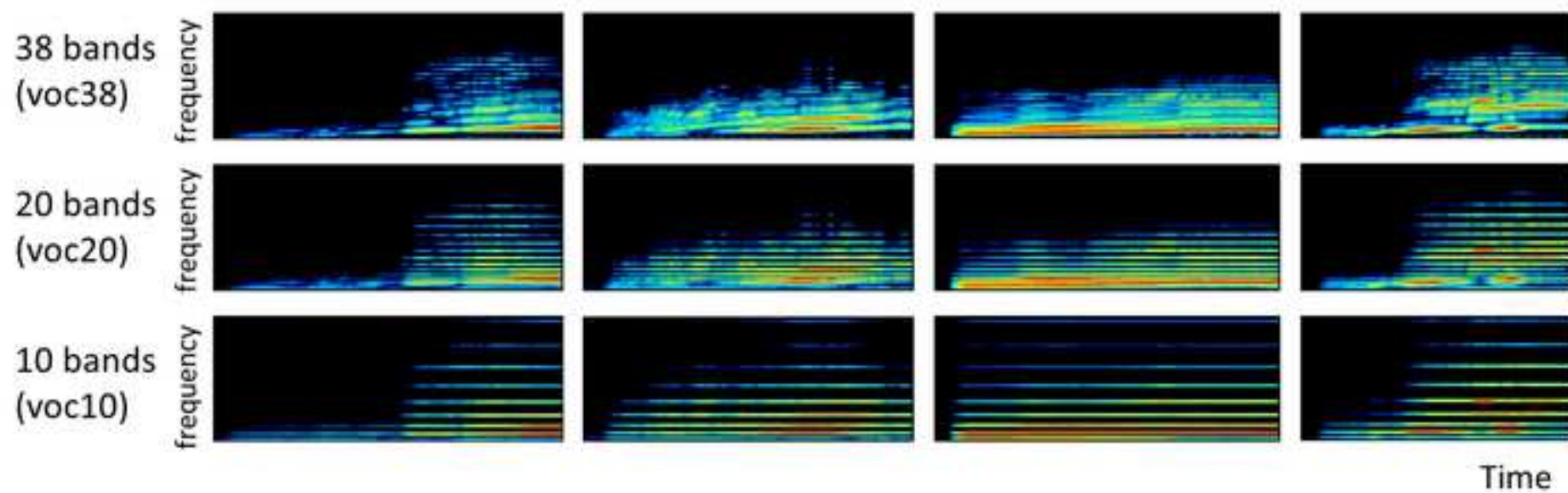
- 1 Liégeois-Chauvel C, Lorenzi C, Trébuchon A, Régis J, Chauvel P(2004) Temporal envelope processing in the  
2 human left and right auditory cortices. *Cereb Cortex* 14(7):731-40. doi:[10.1093/cercor/bhh033](https://doi.org/10.1093/cercor/bhh033).
- 3 Liu RC, Schreiner CE (2007) Auditory cortical detection and discrimination correlates with communicative  
4 significance. *PLoS Biol.* 5, e173.
- 5 Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BCJ (2006) Speech perception problems of the hearing impaired  
6 reflect inability to use temporal fine structure. *Proc Natl Acad Sci USA* 103:18866-18869.
- 7 Lutman ME, Gatehouse S, Worthington AG (1991) Frequency resolution as a function of hearing threshold and  
8 age. *J Acoust Soc Am* 89:320–328.
- 9 Machens CK, Wehr MS, Zador AM (2004) Linearity of cortical receptive fields measured with natural sounds.  
10 *Journal of Neuroscience* 24:1089–1100.
- 11 Manunta Y, Edeline JM (1999) Effects of noradrenaline on frequency tuning of auditory cortex neurons during  
12 wakefulness and slow-wave sleep. *Eur J Neurosci* 11:2134–2150. doi: [10.1046/j.1460-](https://doi.org/10.1046/j.1460-9568.1999.00633.x)  
13 [9568.1999.00633.x](https://doi.org/10.1046/j.1460-9568.1999.00633.x).
- 14 Massaux A, Dutrieux G, Cotillon-Williams N, Manunta Y, Edeline JM (2004) Auditory thalamus bursts in  
15 anesthetized and non-anesthetized states: contribution to functional properties. *Journal of*  
16 *Neurophysiology*, 91, 2117-2134.
- 17 Massaux A, Edeline JM (2003) Bursts in the Medial Geniculate Body: a comparison between anesthetized and  
18 unanesthetized states in guinea-pig. *Experimental Brain Research*, 153, 573-578.
- 19 Mesgarani N, David SV, Fritz JB, Shamma SA (2014) Mechanisms of noise robust representation of speech in  
20 primary auditory cortex. *Proc Natl Acad Sci USA* 111(18):6792-7. doi: [10.1073/pnas.1318017111](https://doi.org/10.1073/pnas.1318017111).
- 21 Moore BC, Glasberg BR, Stoev M, Füllgrabe C, Hopkins K (2012a) The influence of age and high-frequency  
22 hearing loss on sensitivity to temporal fine structure at low frequencies (L). *J Acoust Soc Am*  
23 131(2):1003-6. doi: [10.1121/1.3672808](https://doi.org/10.1121/1.3672808)
- 24 Moore BC, Vickers DA, Mehta A (2012b) The effects of age on temporal fine structure sensitivity in monaural  
25 and binaural conditions. *Int J Audiol* 51(10):715-21.
- 26 Moore BC, Hafter ER, Glasberg BR (1996) The probe-signal method and auditory-filter shape: results from  
27 normal- and hearing-impaired subjects. *J Acoust Soc Am* 99(1):542-52.
- 28 Nagarajan SS, Cheung SW, Bedenbaugh P, Beitel RE, Schreiner CE, Merzenich MM (2002) Representation of  
29 spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. *J*  
30 *Neurophysiol* 87(4):1723-37.
- 31 Nelson PB, Jin SH, Carney AE, Nelson DA (2003) Understanding speech in modulated interference : cochlear  
32 implant users and normal-hearing listeners. *J Acoust Soc Am* 115:2286-2294.
- 33 Occelli F, Suied C, Pressnitzer D, Edeline JM, Gourévitch B (2016) A neural substrate for rapid timbre  
34 recognition ? Neural and behavioral discrimination of very brief acoustic vowels. *Cerebral Cortex* 1-14.
- 35 Overton JA, Recanzone GH (2016) Effects of aging on the response of single neurons to amplitude-modulated  
36 noise in primary auditory cortex of rhesus macaque. *J Neurophysiol* 115(6):2911-23. doi:  
37 [10.1152/jn.01098.2015](https://doi.org/10.1152/jn.01098.2015).
- 38 Paraouty N, Lorenzi C (2017) Using individual differences to assess modulation-processing mechanisms and age  
39 effects. *Hear Res* 344:38-49. doi: [10.1016/j.heares.2016.10.024](https://doi.org/10.1016/j.heares.2016.10.024).
- 40 Paraouty N, Ewert SD, Wallaert N, Lorenzi C (2016) Interactions between amplitude modulation and frequency  
41 modulation processing: Effects of age and hearing loss. *J Acoust Soc Am* 140(1):121. doi:  
42 [10.1121/1.4955078](https://doi.org/10.1121/1.4955078).
- 43 Patterson RD (1987) A pulse ribbon model of monaural phase perception. *J Acoust Soc Am* 82(5):1560-86.
- 44 Peters RW, Moore BCJ (1992) Auditory filters and aging: filters when auditory thresholds are normal, in  
45 *Auditory Physiology and Perception*, eds Cazals Y., Demany L., Horner K., editors. (Oxford: Pergamon; ),  
46 179–185.
- 47 Presacco A, Simon JZ, Anderson S (2016) Evidence of degraded representation of speech in noise, in the aging  
48 midbrain and cortex. *J Neurophysiol*. doi: [10.1152/jn.00372.2016](https://doi.org/10.1152/jn.00372.2016).
- 49 Ranasinghe KG, Vrana WA, Matney CJ, Kilgard MP (2012) Neural Mechanisms Supporting Robust Discrimination  
50 of Spectrally and Temporally Degraded Speech. *J Assoc Res Otolaryngol* 13(4): 527–542.
- 51 Rennaker RL, Carey HL, Anderson SE, Sloan AM, Kilgard MP (2007) Anesthesia suppresses nonsynchronous  
52 responses to repetitive broadband stimuli. *Neuroscience* 145(1):357-69.
- 53 Romanski LM, Averbeck BB (2009) The primate cortical auditory system and neural representation of conspecific  
54 vocalizations. *Annu Rev Neurosci* 32:315-46. doi: [10.1146/annurev.neuro.051508.135431](https://doi.org/10.1146/annurev.neuro.051508.135431).
- 55 Ross B, Fujioka T, Tremblay KL, Picton TW (2007) Aging in binaural hearing begins in mid-life: evidence from  
56 cortical auditory-evoked responses to changes in interaural phase. *J Neurosci* 27:11172–11178. doi:  
57 [10.1523/JNEUROSCI.1813-07.2007](https://doi.org/10.1523/JNEUROSCI.1813-07.2007).
- 58  
59  
60  
61  
62  
63  
64  
65

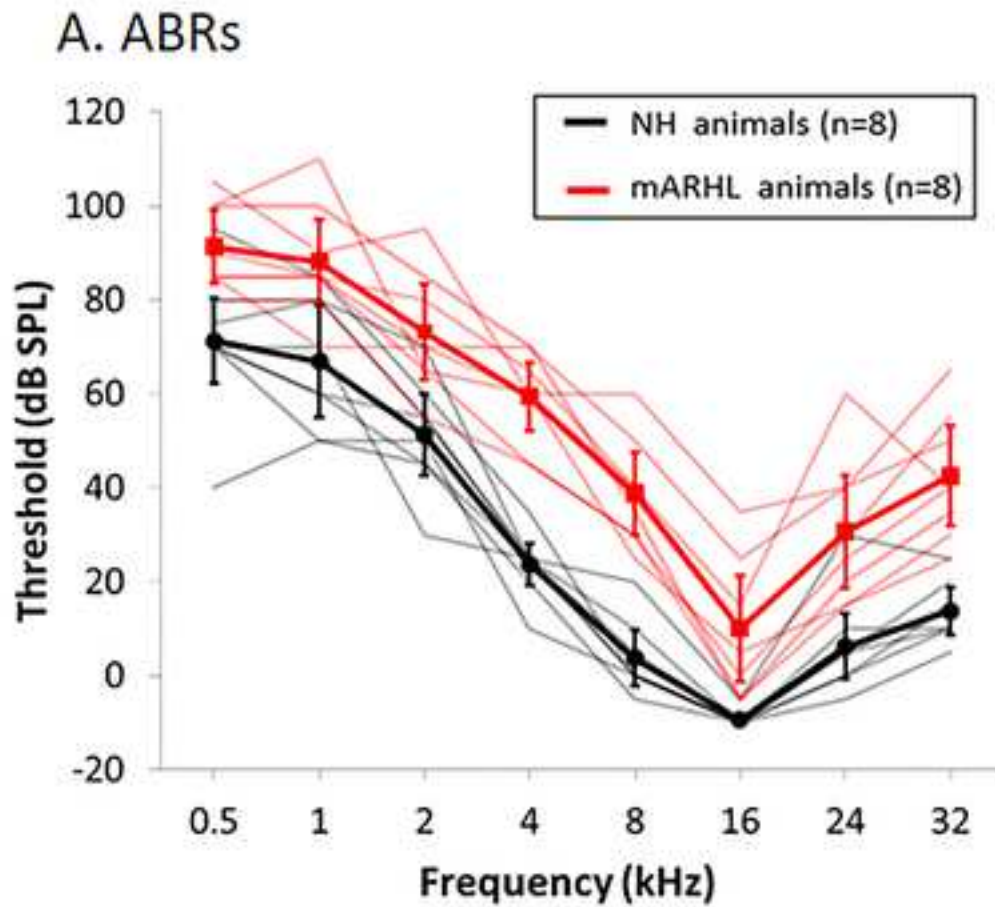
- 1 Sayles M, Winter IM (2010) Equivalent-rectangular bandwidth of single units in the anaesthetized guinea-pig  
2 ventral cochlear nucleus. *Hear Res* 262(1-2):26-33. doi: 10.1016/j.heares.2010.01.015.
- 3 Schnupp JW, Hall TM, Kokelaar RF, Ahmed B (2006) Plasticity of temporal pattern codes for vocalization stimuli  
4 in primary auditory cortex. *J Neurosci* 26:4785–4795. doi: 10.1523/JNEUROSCI.4330-05.2006.
- 5 Schoof T, Rosen S (2014) The role of auditory and cognitive factors in understanding speech in noise by normal-  
6 hearing older listeners. *Front Aging Neurosci* 6:307. doi: 10.3389/fnagi.2014.00307.
- 7 Schvartz KC, Chatterjee M, Gordon-Salant S (2008) Recognition of spectrally degraded phonemes by younger,  
8 middle-aged, and older normal-hearing listeners. *J Acoust Soc Am* 124(6):3972-88. doi:  
9 10.1121/1.2997434.
- 10 Sergeyenko Y, Lall K, Liberman MC, Kujawa SG. Age-related cochlear synaptopathy: an early-onset contributor  
11 to auditory functional decline. *J Neurosci* 33(34):13686-94. doi: 10.1523/JNEUROSCI.1783-13.2013.
- 12 Shamma S, Lorenzi C (2013) On the balance of envelope and temporal fine structure in the encoding of speech  
13 in the early auditory system. *J Acoust Soc Am* 133(5):2818-33. doi: 10.1121/1.4795783.
- 14 Shannon CE (1948) A Mathematical Theory of Communication. *Bell Syst. Tech. J.* 27, 379–423.
- 15 Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues.  
16 *Science* 270:303-304
- 17 Sheft S, Ardoint M, Lorenzi C (2008) Speech identification based on temporal fine structure cues. *J. Acoust. Soc.*  
18 *Am* 124:562-575.
- 19 Sheldon S, Pichora-Fuller MK, Schneider BA (2008) Effect of age, presentation method, and learning on  
20 identification of noise-vocoded words. *J Acoust Soc Am* 123(1):476-88. doi: 10.1121/1.2805676.
- 21 Smith ZM, Delgutte B, Oxenham AJ (2002) Chimaeric sounds reveal dichotomies in auditory perception. *Nature*  
22 416:87-90.
- 23 Sörös P, Teismann IK, Manemann E, Lütkenhöner B (2009) Auditory temporal processing in healthy aging: a  
24 magnetoencephalographic study. *BMC Neurosci* 10:34. doi: 10.1186/1471-2202-10-34.
- 25 Souza PE, Boike KT (2006) Combining temporal-envelope cues across channels: effects of age and hearing loss. *J*  
26 *Speech Lang Hear Res* 49(1):138-49.
- 27 Souza PE, Kitch V (2001) The contribution of amplitude envelope cues to sentence identification in young and  
28 aged listeners. *Ear Hear* 22(2):112-9.
- 29 Stone MA, Füllgrabe C, Mackinnon RC, Moore BCJ (2011). The importance for speech intelligibility of random  
30 fluctuations in "steady" background noise. *J. Acoust. Soc. Am.* 130, 2874-2881.
- 31 Takahashi GA, Bacon SP (1992) Modulation detection, modulation masking, and speech understanding in noise  
32 in the elderly. *J Speech Hear Res* 35(6):1410-21.
- 33 Ter-Mikaelian M, Semple MN, Sanes DH (2013) Effects of spectral and temporal disruption on cortical encoding  
34 of gerbil vocalizations. *J Neurophysiol* 110(5):1190-204. doi: 10.1152/jn.00645.2012.
- 35 Wallace MN, Palmer AR (2008) Laminar differences in the response properties of cells in the primary auditory  
36 cortex. *Exp Brain Res* 184:179–191
- 37 Wallace MN, Rutkowski RG, Palmer AR (2000) Identification and localisation of auditory areas in guinea pig  
38 cortex. *Exp Brain Res* 132:445–456. doi: 10.1007/s002210000362.
- 39 Wallaert N, Moore BC, Ewert SD, Lorenzi C (2017) Sensorineural hearing loss enhances auditory sensitivity and  
40 temporal integration for amplitude modulation. *J Acoust Soc Am.*;141(2):971. doi: 10.1121/1.4976080.
- 41 Wallaert N, Moore BC, Lorenzi C (2016) Comparing the effects of age on amplitude modulation and frequency  
42 modulation detection. *J Acoust Soc Am* 139(6):3088. doi: 10.1121/1.4953019.
- 43 Walton JP, Frisina RD, O'Neill WE (1998) Age-Related Alteration in Processing of Temporal Sound Features in  
44 the Auditory Midbrain of the CBA Mouse. *J Neurosci* 18:2764–2776.
- 45 Whitmal NA III, Poissant SF, Freyman RL, Helfer KS (2007) Speech intelligibility in cochlear implant simulations:  
46 Effects of carrier type, interfering noise, and subject experience.
- 47 Woolley SM, Gill PR, Theunissen FE (2006) Stimulus-dependent auditory tuning results in synchronous  
48 population coding of vocalizations in the songbird midbrain. *J. Neurosci.* 26, 2499–2512.
- 49 Zeng FG, Nie K, Stickney GS, Kong YY, Vongphoe M, Bhargava A, Wei C, Cao K (2005) Speech recognition with  
50 amplitude and frequency modulations. *Proc Natl Acad Sci* 102:2293-2298.
- 51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

### A. Original whistles

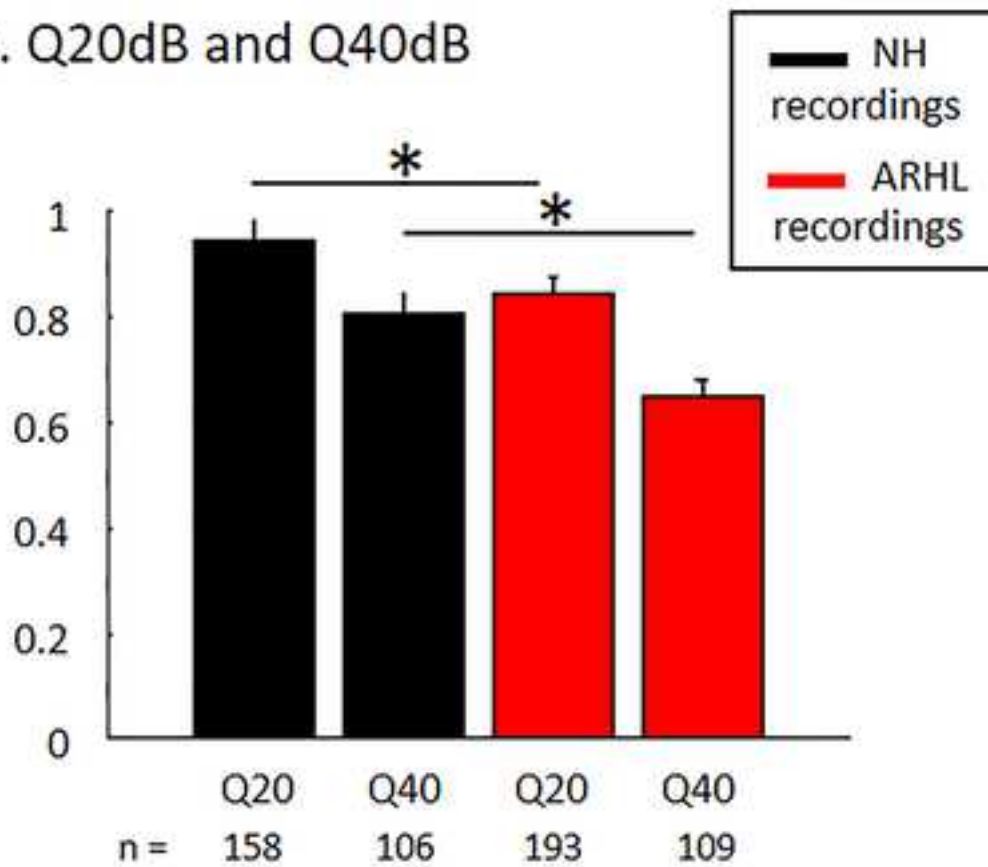


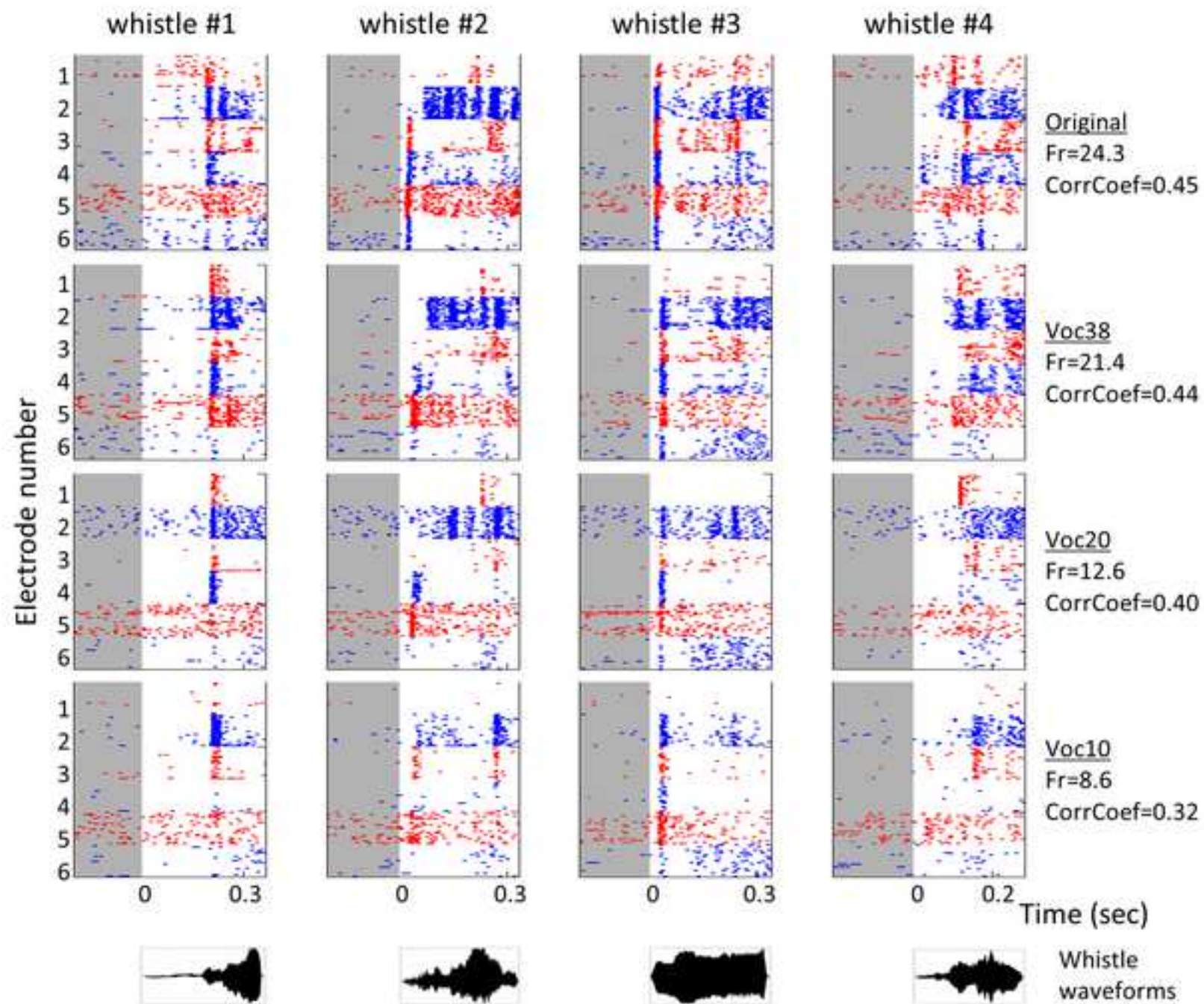
### B. Vcoded whistles

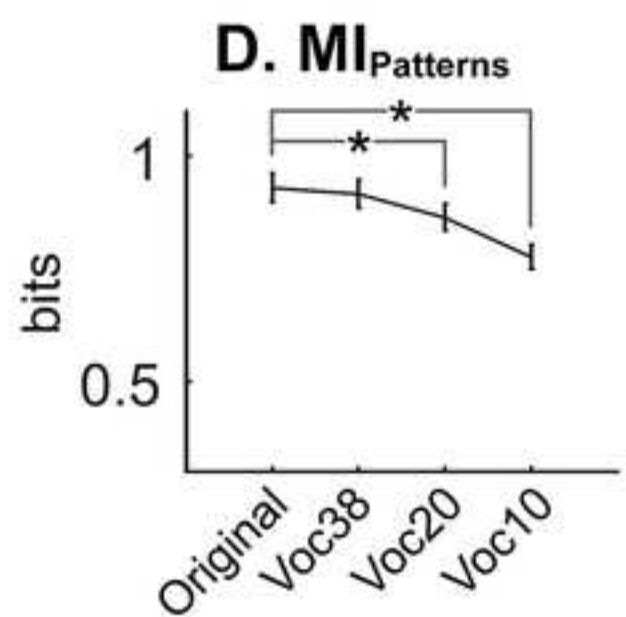
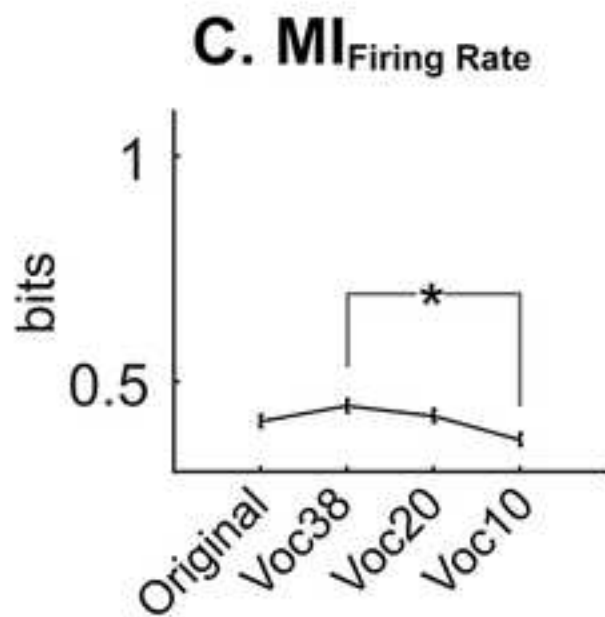
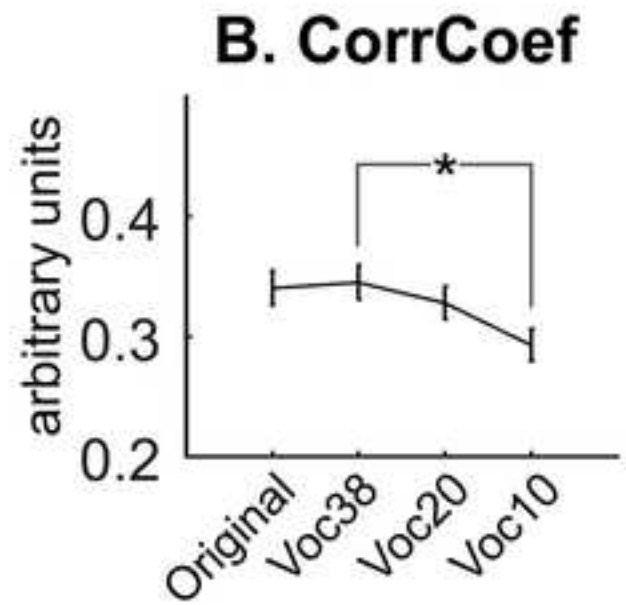
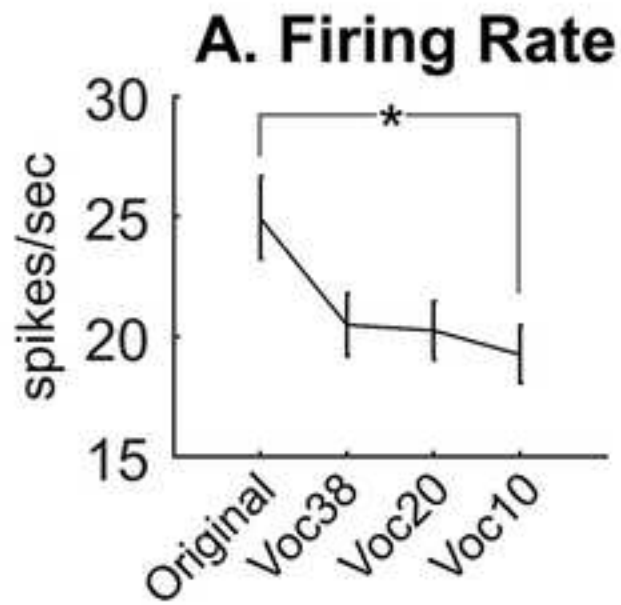




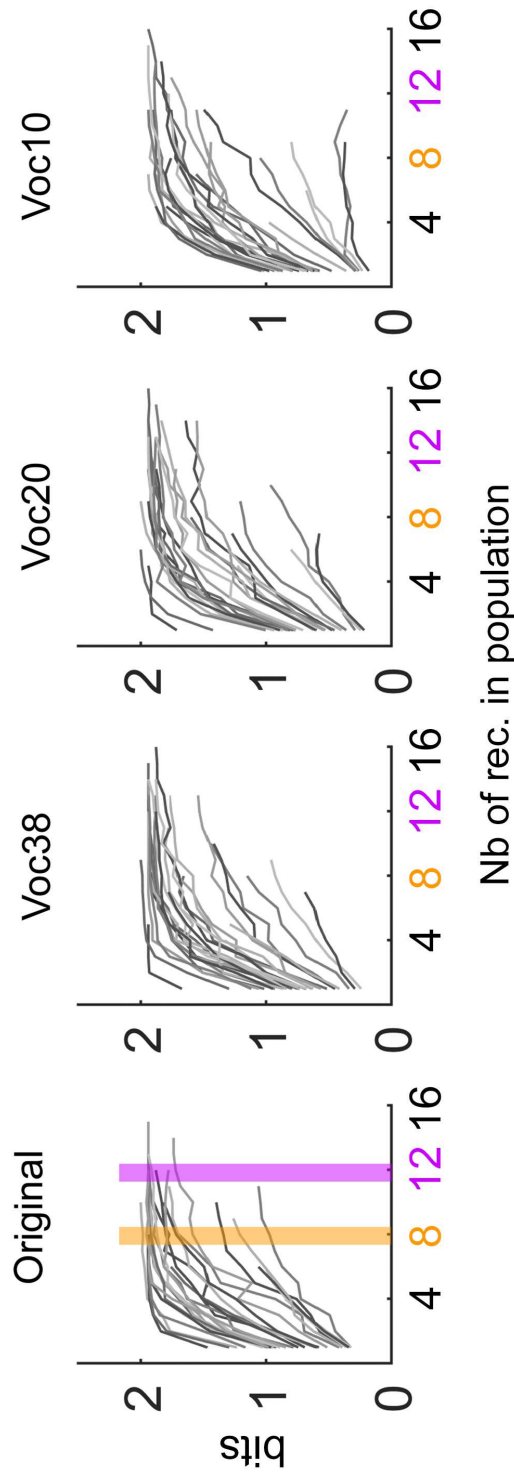
### B. Q20dB and Q40dB



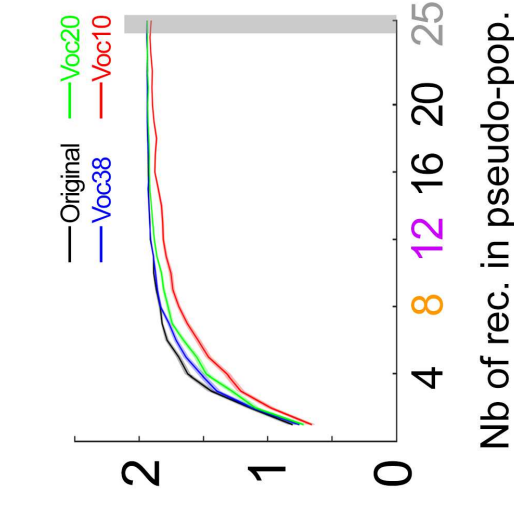




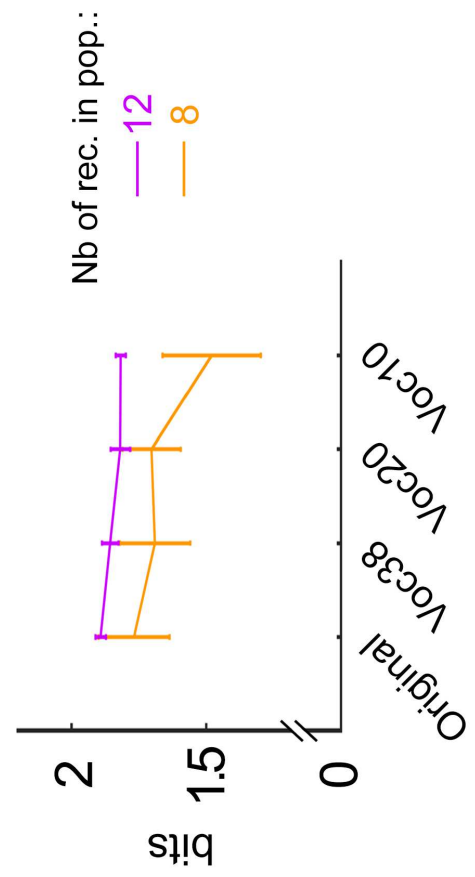
### A. Effect of population size on $MI_{Population}$



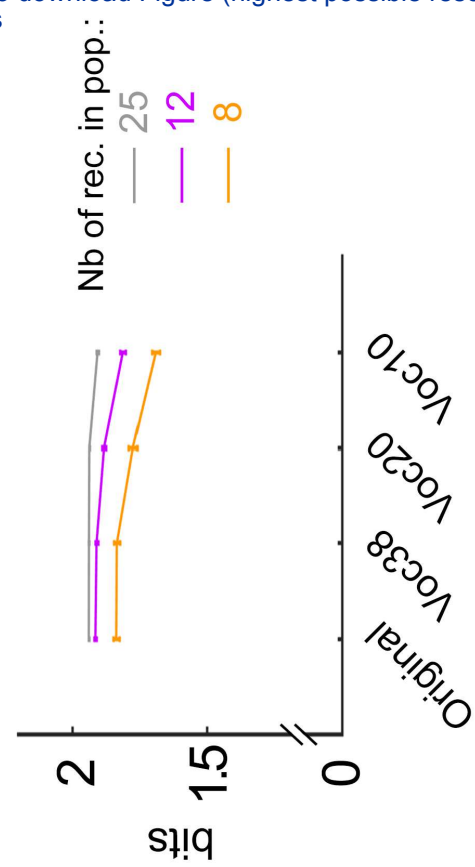
### C. $MI_{pseudo-population}$



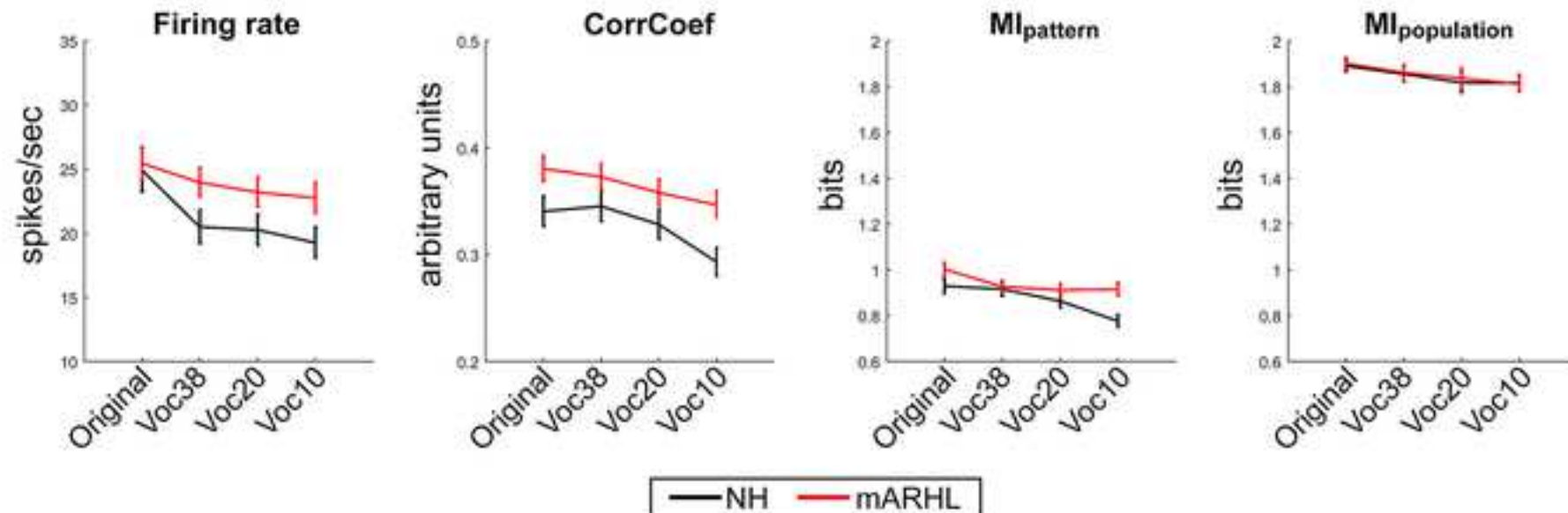
### B. Mean $MI_{Population}$ values



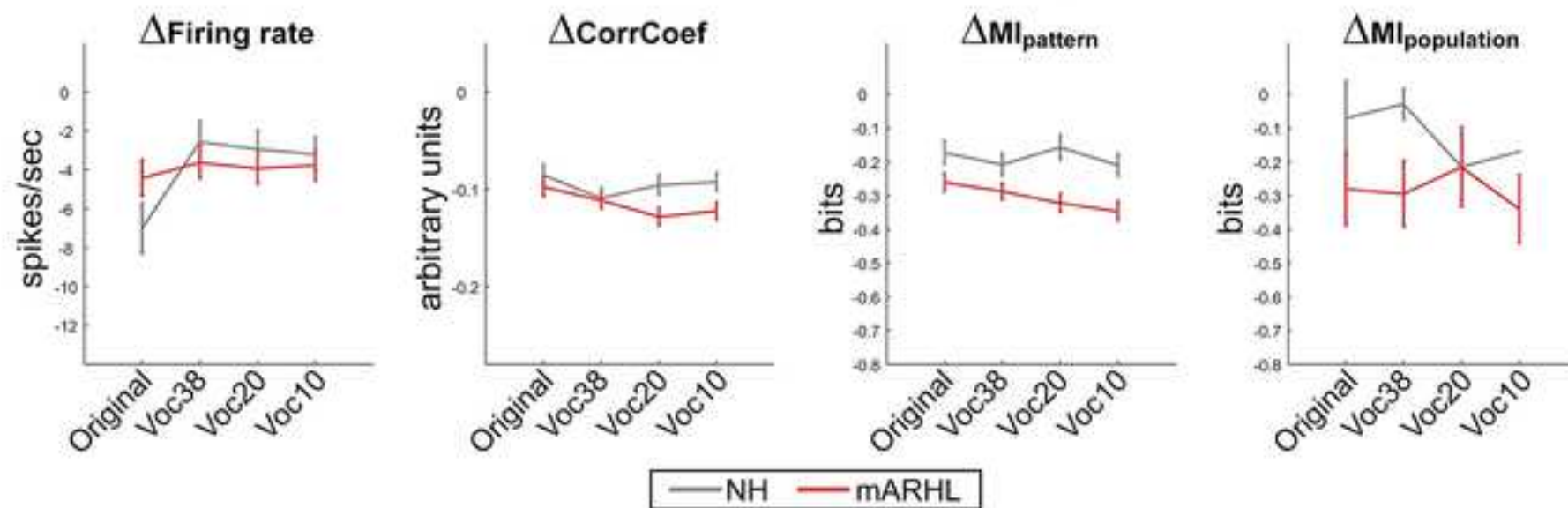
### D. Mean $MI_{pseudo-population}$



## A. Neuronal activity measures for mARHL vs NH animals



## B. Effect of background noise on neuronal activity measures







Orsay, September 22th, 2017

Dear Editor in Chief, Dear Paul, Dear Associate Editor,

The authors declare that no competing interests exist.

Best regards,

Yonane Aushana, Samira Souffi, Jean-Marc Edeline, Christian Lorenzi, Chloé Huetz