nature biomedical engineering | PEER-REVIEW INFORMATION

Compression and amplification algorithms in hearing aids impair the selectivity of neural responses to speech

Corresponding author: Nicholas Lesica

Editorial note

This document includes relevant written communications between the manuscript's corresponding author and the editor and reviewers of the manuscript during peer review. It includes decision letters relaying any editorial points and peer-review reports, and the authors' replies to these (under 'Rebuttal' headings). The editorial decisions are signed by the manuscript's handling editor, yet the editorial team and ultimately the journal's Chief Editor share responsibility for all decisions.

Any relevant documents attached to the decision letters are referred to as **Appendix #**, and can be found appended to this document. Any information deemed confidential has been redacted or removed. Earlier versions of the manuscript are not published, yet the originally submitted version may be available as a preprint. Because of editorial edits and changes during peer review, the published title of the paper and the title mentioned in below correspondence may differ.

Correspondence

Fri 25/09/2020
Decision on Article nBME-20-1976

Dear Professor Lesica,

Thank you again for submitting to *Nature Biomedical Engineering* your manuscript, "The hearing aid dilemma: amplification, compression, and distortion of the neural code". The manuscript has been seen by 4 experts, whose reports you will find at the end of this message. You will see that the reviewers have good words for the work, and that they raise a number of technical criticisms that we hope you will be able to address. In particular, we would expect that a revised version of the manuscript provides:

* Discuss roll-over and the implications of WDRC on longer-term features of neural representations, as suggested by Rev #1.

*Textual changes for clarification, as suggested by Rev #2 and #3.

*Clarification of the impact on hearing in real-world scenarios, as suggested by Reviewer #3.

*Clarification of neuronal distribution across the arrays, as suggested by Rev #3.

*Sensitivity analysis with current data, as suggested by Rev #3.

When you are ready to resubmit your manuscript, please upload the revised files, a point-by-point rebuttal to the comments from all reviewers, the (revised, if needed) reporting summary, and a cover letter that explains the main improvements included in the revision and responds to any points highlighted in this decision.

Please follow the following recommendations:

* Clearly highlight any amendments to the text and figures to help the reviewers and editors find and understand the changes (yet keep in mind that excessive marking can hinder readability).

* If you and your co-authors disagree with a criticism, provide the arguments to the reviewer (optionally, indicate the relevant points in the cover letter).

* If a criticism or suggestion is not addressed, please indicate so in the rebuttal to the reviewer comments and explain the reason(s).

* Consider including responses to any criticisms raised by more than one reviewer at the beginning of the rebuttal, in a section addressed to all reviewers.

* The rebuttal should include the reviewer comments in point-by-point format (please note that we provide all reviewers will the reports as they appear at the end of this message).

* Provide the rebuttal to the reviewer comments and the cover letter as separate files.

We hope that you will be able to resubmit the manuscript within 25 weeks from the receipt of this message. If this is the case, you will be protected against potential scooping. Otherwise, we will be happy to consider a revised manuscript as long as the significance of the work is not compromised by work published elsewhere or accepted for publication at *Nature Biomedical Engineering*. Because of the COVID-19 pandemic, should you be unable to carry out experimental work in the near future we advise that you reply to this message with a revision plan in the form of a preliminary point-by-point rebuttal to the comments from all reviewers that also includes a response to any points highlighted in this decision. We should then be able to provide you with additional feedback.

We hope that you will find the referee reports helpful when revising the work, which we look forward to receive. Please do not hesitate to contact me should you have any questions.

Best wishes,

Michelle

__
Dr Michelle Korda
Senior Editor, *Nature Biomedical Engineering*

_____
Reviewer #1 (Report for the authors (Required)):

Results
This study demonstrates convincingly that the neural representation of simple and complex sounds in mild-to-moderate sensorineural hearing loss (the most common) is impaired by the ubiquitous algorithms employed to introduce 'wide dynamic range compression, WDRC,' into hearing aids (the most common therapy for mild-to-moderate hearing loss), compared to if these algorithms are not included. Amplification of sound per se, without WDRC restores the neural representation, including discrimination of speech sounds, to normal. WDRC actively harms this restoration.

The significance of this study

This is an outstanding manuscript, with the potential to be transformative for an entire field of therapeutics, and for the hundreds of millions of individuals living with the problem of hearing loss. Presuming they can be accessed, hearing aids remain the only option for the hearing impaired, the single biggest sensory deficit globally, affecting over 600 million people (and rising, as our populations age). No biological, pharmaceutical, or genetic therapies exist, and will not for some time. Hearing aids are the only option (other than for the tiny fraction of hearing-impaired individuals who receive cochlear implants). Despite this, the majority of people who need one don't take one, and of those that do, the majority remain under-used or not used at all. Why is this the case?

A major problem is that ALL sensorineural hearing loss (overwhelming the result of damage to the sensory hair cells of the inner ear), including so-called 'mild-to-moderate' hearing (the bracket into which most hearing-impaired listeners 'begin their journey' into hearing loss, and in which many stay, and the focus of this study), is treated as a single problem with a single solution—a hearing aid with wide-dynamic range compression (WDRC). WDRC treats both the problem of hearing loss, and the entire process of hearing per se, as an issue of inner-ear mechanics alone. To this end, current hearing aids take no account whatsoever of the existence of the auditory brain in their design and algorithmic implementation. Three decades of an increasingly sophisticated understanding of the auditory brain, its moment by moment, and longer-term functions, its rapid adaptive capacity, and its coding complexity, have contributed nothing to their design.

Distilled into a technological
solution that considers the inner ear only, then, hearing aids cannot hope to solve the problems of hearing loss as they neglect entirely to consider them, assuming only that applying a one-size-fits-all solution to a narrowly defined problem is the only therapeutic approach. This makes no sense. Considered in terms of inner-ear pathology only, there is NO 'reverse filter' or algorithm that can overcome the problem of hearing loss, and ploughing down this route guarantees failure. This leads to problems such as the increasing use of front-end algorithms to try to assuage the problems introduced by hearing-aid algorithms (e.g. WDRC for mild-to-moderate hearing loss) in the first place, or moment-by-moment denoising algorithms that introduce as many problems as they seek to resolve. This study recognises that potential solutions/effective therapies to hearing impairment lie not only in understanding the computational totality of the auditory system, including the auditory brain, but also in treating hearing loss with the current standard therapy—hearing aids—in terms of that totality. I would expect future hearing aid (and cochlear implant) algorithms to be tested experimentally in similar fashion (currently they hit the market with zero assessment in pre-clinical studies). Further, with the advent of 'hearables' that amplify only, the vast majority of indivuduals seeking their first entry to the hearing solutions market have a better solution already available, based on the outcome of this study—an amplifying device with no WDRC. This may be transformative for individuals and the hearing-aid market.

Major Issues

I really have no major 'problem' points I can raise with respect to this manuscript. Not only is it pioneering in the question it addresses, the experimental paradigm (inc. the recording of 5000! neurons using multi-electrode arrays) will define the field. The level of understanding of the broad problem and how to address it at a neural level, the sophistication of the assessments, including examining the selectivity of neural populations for identifying/discriminating tones and speech sounds, including in background noise, are outstanding.

The issue of 'roll-over' at higher sound intensities is a complex one, but widely accepted in terms of normal hearing. I would like the authors to address briefly in the Discussion what might be a potential solution to the problem of roll-over. Is it de-amplification, and how might this be achieved and impact on hearing function (since it is common to normal and impaired hearing)?

The midbrain lends itself to assessing the 'immediate encoding' aspect of hearing function. What are the implications, if any, for WDRC on other, longer-term features, of neural representations, including 'natural' dynamic-range adaptation? Can the authors comment in the Discussion?

Minor Comments

It would be worth commenting a little more directly on the choice of the midbrain for these recordings. What is the value, say,compared to brainstem or cortical?

My minor quibble is stylistic. contrast') should be avoided, and the term/s 'reduce/reduced' employed throughout.

Line 280 'missing 'than'. A thorough check of the ms for typos and grammar please, including the appropriate use of commas.

WDRC is independently applied to each hearing aid, and consequently impairs cues for spatial hearing and speech-in-noise processing. This could be commented on, and is actually widely accepted in human studies.

Reviewer #2 (Report for the authors (Required)):

This is a valuable study that addresses the significant problem of why hearing aids are not more successful in ameliorating mild-to-moderate hearing loss. The results will be of great importance in improving hearing for hearing-impaired people, including many young and middle-aged people and essentially all elders. The experimental design and data analysis are excellent, albeit challenging to follow at times. My comments are intended primarily to clarify the presentation.

Line 32, "loss of audibility": Please clarify here that (I think) "audibility" refers specifically to the ability or inability to detect sounds at low levels. The next sentence ("As a result…") more or less defines audibility, but the reader still might think that audibility might refer also to, for instance, understanding of moderate-level

sounds.

Figure 4 legend: Please spell out here the abbreviations of the various spectral densities (e.g., PSD, CSD, etc).

Lines 179-182: This sentence is difficult to follow. CSD decreased with increasing frequency up to 80 Hz (i.e., covering the range of envelope frequencies), meaning that internal noise increased markedly from 8 to 80 Hz. CSD then increased, mostly in parallel with PSD, between 80 and 256 Hz (the range of voice pitch), meaning that internal noise was constant in that range. How can both of those observations be consistent with the statement that internal noise was smallest at frequencies corresponding to the envelope and voice pitch of the speech?

Lines 197 to 208: This paragraph is very difficult to follow. Lines 172 to 190 and Fig. 4B convinced me that internal noise and nuisance noise INCREASE when the relevant covariance decreases. Figure 4C plots covariance, and shows that it decreases with hearing loss and generally increases with a hearing aid. The paragraph starting at line 197, however, states that internal and nuisance noise DECREASE with hearing loss. By my understanding, it is the covariance that decreases, so noise must be INCREASING with hearing loss. To further the confusion, the titles over the panels of Fig. 4C (internal noise, nuisance noise, etc.) indicate something that is the inverse of what is actually plotted.

Figure 5D and line 231: It is counter-intuitive that tuning width decreases with hearing loss and is increased by a hearing aid. Those observation make it seem as if hearing loss is beneficial. I understand from the text that tuning width is the width of the tuning curve at a fixed sound level, and that the tuning curve is shifted upwards in the case of hearing loss. Nevertheless, I think that it would be helpful to also provide a measure of sharpness of tuning, like ERB or Q10, that is not sensitive to overall shifts of the tuning curve in the SPL dimension.

Line 243, “general loss of selectivity”: I think that you should say “general loss of frequency selectivity”. There are all kinds of “general” selectivity that are lost with hearing loss (e.g., consonant selectivity). It is only pure-tone frequency selectivity that you tested here.

Line 259, “does serve to”: This is maybe not the best word choice. “serve to” sounds like doing a good thing, whereas in this instance you are making the case that decreased spectral and temporal contrast is a bad thing.

Lines 278-282: This seems to say rather unequivocally that linear amplification solves all the problems of hearing loss and that there is no value in compression. I think that you should note here, however, that subsequent sections will show that high sound levels, possibly due to linear amplification, can lead to impaired discrimination of complex sounds (referred to later as rollover).

Line 456: It is a mystery here why the sounds in multi-talker babble were not presented diotically. Please state here that those sounds were processed with HRTFs.

Reviewer #3 (Report for the authors (Required)):

The authors provide a detailed analysis of neural coding in the inferior colliculus of a gerbil in which mild to moderate sensorineural hearing loss has been induced by noise exposure. The authors contend that compression programming intended to protect against excessive amplification actually introduces distortion that impacts hearing aid performance. They argue based on their findings, which are eloquently detailed, that simple amplification utilized in inexpensive, over the counter, devices will help the vast majority of the many people who suffer from mild to moderate hearing loss worldwide. This work has major implications for public health. The work presented is innovative, original, and impactful, and will be of interest to a wide audience. One minor comment/typo: line 280, I believe that the word ‘than’ is missing and should be the second word in the line?

Reviewer #4 (Report for the authors (Required)):

Overview:

Hearing aids help numerous people in this world on a daily basis. However, performance levels and adoption have been somewhat stagnant for decades, in which many people still do not sufficiently benefit from them and a high percentage of people who could benefit from hearing aids do not use them. One major factor is that for many, they still do not perform well in more complex and challenging environments, such as noisy conditions and multi-talker situations, which unfortunately are common situations in a person's daily life. Numerous device developments and perceptual studies in human subjects have occurred, but we still know little about the inner workings of the brain in response to hearing aids and what actual hearing aid features are or are not being properly processed in the auditory pathway; thus, we still lack a clear idea of what features need to be improved from a "brain" perspective. Surprisingly, there are very few brain studies relating to hearing aids and in characterizing hearing aid effects to guide how best to improve them. This paper by Armstrong et al. is a major contribution to the hearing aid field because it provides a detailed neurophysiological assessment of several of the challenges with current hearing aid algorithms and new insights into possible directions for improving these technologies. In particular, they recorded from numerous neurons across the inferior colliculus using large-scale invasive sensing and assessed how the neural code in response to simple and speech-like sounds differ between normal hearing and hearing loss conditions with and without hearing aid application. They were able to show that hearing thresholds/sensitivity and auditory activation in hearing loss animals can be recovered with a typical hearing aid algorithm, as well basic spectral tuning properties to simple tones. They also used well thought out and creative methods (including various types of classifiers and spectral density measures) to show that several different features of the neural code (e.g., related to intrinsic/external noise variations, general response to speech-like stimuli, and different talker variations) could be mostly recovered by a hearing aid, but that the key deficit that still remained was in the ability to sufficiently differentiate between specific speech features (i.e., consonant variations that was specifically tested in this study). A critical piece to the puzzle that the authors nicely addressed in this study, and based on previous perceptual studies in the hearing aid field, was to remove the compression effects of the hearing aid algorithm to provide a linear gain compensation, which then was able to recover the ability to differentiate specific speech features, at least in quiet conditions and for simpler conditions like two-talkers scenario. However, for more complex background noise (babble noise), the recovery was not sufficient. They showed some data to suggest this remaining deficit is due to an inherent issue caused by simply cranking sounds too loud regardless if it happening in hearing loss or in normal hearing subjects (i.e., called a rollover effect). Therefore, they were able to demonstrate, at least for mild-to-moderate hearing loss, that amplification without non-linear or suboptimal compression may be a better way to go for many individuals. Furthermore, future hearing aid algorithms need to find a way to increase perception of sound to louder levels while somehow minimizing this rollover effect.

This paper not only provides fundamental knowledge with actual large-scale brain recordings that has not yet been available in the hearing aid field, but also provides important insights into what features are compromised and what features need to be improved in future algorithms. Although the use of different classifiers and spectral density analyses are not novel in the neuroscience field, the authors were quite creative in being able to dissociate and relate the different neural code contributions to their significance for hearing performance using these methods. Furthermore, the findings open the way for future studies using these types of large-scale brain methods in animals to further understand what other sound features and neural coding components are compromised, not just for the audio part (as represented in lower brain centers like in this study), but also to do so in higher centers to incorporate multi-modal integration effects, learning and attentional contributions that all can provide a more complete real-world view of the limitations and opportunities to push the hearing aid field to the next level.

The final point in terms of broad impact is that this study is timely considering the recent surge of attention towards over-the-counter hearing devices and audio consumer technologies now being applied to hearing enhancement by mainstream mega companies. There is going to be massive interest in finding ways to simplify devices (auto and smart fitting capabilities) and enhanced hearing technologies for the mainstream public, not just for those with mild-to-moderate hearing loss but even normal hearing situations. The findings from this study provide insightful directions for the field to push forward in relation to algorithms and also in demonstrating the immense potential in using animal models in creative ways to answer critical questions not possible directly in humans.

One overall comment to the style of the paper is to commend the authors for the beautiful and well organized figures and systematic flow of the paper. It was easy to read and to understand the key points.

Major Comments:

1) The findings are convincing that hearing loss leads to deficits in key sound coding features, which are then not fully recovered from a hearing aid. A critical piece was in showing most of the features tested could be sufficiently recovered when using a linear gain approach. However, one concern is that the amount of deficit seen from the hearing loss is on the order of about 5% and even with that drop, the neurons are still doing a decent/good job at detection and identification of the speech components (e.g., in Fig. 3B/D and Fig. 5F). It isn't clear if this is because the speech material is too easy to detect/discriminate (especially with only 12 consonants) compared to what actual hearing aid users in the real-world are faced with in terms of more complex and open-ended sentences in varying noise environments, and/or if it is because only a short initial portion of 150 ms is used that simplifies the task. There is clearly more dramatic deficits experienced by hearing aid users on a daily basis so it is important for the authors to more clearly link these % changes in performance and how they are clinically relevant and aligned with hearing aid users to maintain the translatability of the data to the human situation. Decoding of consonants/vowels is much easier to do, and is even possible with greatly compromised hearing with auditory implants, compared to trying to reconstruct or recite the actual words or sentences, in which the latter is more what would happen in the real-world. Making a clearer link is important to be able to justify the second half of the paper showing that the ability to recover the deficits using a linear gain is true for real-world scenarios in better understanding speech, and not just that it helps to recover "decoding" of simple speech features, like consonants.

2) The concept to use large-scale recordings across a major auditory nucleus, such as the IC, and also bilaterally, is a smart and appropriate way to get at the questions posed in this paper. One critical requirement though is that enough of the IC is spanned, at least the central nucleus of the IC, such that any interpretations of the findings are representative of the full deficits or recovery, since the assumption is that it is possible to capture the deficits via the "neural code" because all or most of the information must pass through the IC to higher perceptual centers. In an extreme case, there may be critical regions in the IC that are especially sensitive to deficits in hearing thresholds and amplification, and if those neurons are missed, then the findings do not provide a complete representation of what is going on. The authors did not provide enough information as to which regions and frequency layers were spanned by the arrays, as well as how well the neurons were
distributed across the isofrequency dimension of the IC. There are multiple studies showing that many coding properties, including timing, jitter, spectral tuning, thresholds, activation, suprathreshold activation properties, etc. can greatly change depending on location along the isofrequency layer of the IC and even in different nuclei of the IC (see Straka et al., 2014 listed below for examples and many papers on this topic). It seemed from Fig. 5B that a decent span of sites occurred across the frequency axis (though numbers of neurons per CF wasn't clear for confirming even distribution) but couldn't tell how well other regions of the IC were covered. The authors may have sufficiently spanned and checked coverage or found through other analyses that their coverage was sufficient for the purposes of their analyses. It is possible that even greater deficits might have occurred if recording from certain regions and/or some deficits maybe less recovered if looking at other
regions. The current state of the paper is already impressive in terms of demonstrating the deficits of hearing loss on selectivity of speech perception and how linear gain can recover some of those deficits. However, understanding this coverage will enable the ability to interpret the findings more comprehensively. If location of sites are not well known, which is totally understandable since doing histology and tracking all sites is very challenging, one possibility is to take the current data and perform some sensitivity analyses. For example, to take half of the more medial IC data sites and see how they perform to the more lateral sites (but having similar CFs), or other location comparisons. These findings would show that the overall findings are not so sensitive to IC location or if they are, then how to interpret the data presented. Also, further justification is needed why only 150 neurons were used for the classifications, and which neurons (i.e., what were their characteristics in terms of CF, tuning, etc. to understand what types of neurons are doing the decoding for interpreting the findings).

Straka MM, Schmitz S, Lim HH. "Response features across the auditory midbrain reveal an organization consistent with a dual lemniscal pathway." Journal of Neurophysiology, 112(4): 981-98, 2014

Minor Comments:

ABSTRACT

-Line 9: Is it true hearing aids are the "only available treatment", if including other sensory and feedback methods?

-Line 17: reword to sentence to, "To improve future hearing aids, new processing strategies must be

developed that avoid this tradeoff between neural sensitivity and selectivity."

INTRODUCTION:

-Line 27: instead of "only" maybe say "main" or something similar.

-Lines 28-29: It is a strong statement (somewhat unfair representation) to say "only a small fraction" of people…and "only a small fraction" of that use/benefit from hearing aids. More convincing would be to provide percentages or tone down the claim. Hearing aids still can do a good job and many people benefit from them. The point of the paper shouldn't be how terrible the hearing aid field is but more how it has been helpful but so much more we can do to make hearing aids better.

-Throughout the paper, the word "but" is used quite frequently and many times to start sentences. I would suggest mixing it up and even not starting sentences with the word except occasionally if needed.

-Lines 46-49: This might be a good spot to explain the concept of rollover and why it is believed to cause deficits even in normal hearing listeners. This is later talked about in the Results and then in the Discussion but it isn't clear why it is happening. Either clarify what is "rollover" and why it presumably happens here or when discussing the results (the latter may be better since Intro is already getting long and would make more sense when actually presenting the relevant data).

-Line 73: Is this study the only one doing neural recordings related to hearing aids? I would be cautious of that statement and how you define a "hearing aid study".

-Lines 89-90: Too strong of a statement. Maybe reword to: "We found that most of the distortions in the neural code that we tested that are caused by hearing loss are, in fact, corrected by a hearing aid; yet a loss of selectivity in neural responses that is.."

-Line 92: What is meant by "fundamental deficit in auditory processing"? Be more specific since it is a broad statement to make from the specific dataset.

RESULTS:

-Quite impressive of using 512 sites per animal. Yet, further info is needed why with all of the animals, the total numbers come out to be just 2000 or 3000 sites in Table S1, i.e., how was it determined which sites to keep or discard?

-In all figures, spell out "w/" and "w/o".

-Figure 3 caption, last 2 sentences: what does "each population" refer to? As with all the other figures, it is sometimes confusing to know what each point means in terms of number of animals or number of neurons or number of trials. Much of this information is in Table S1 but provide some basic info in the captions would be helpful to keep things straight.

-Figure 4C (and other similar types of plots): Are the points and 95% CIs for averages across trials per neuron or average across all trials and neurons? Please clarify also why it was done that way.

-Figure 4 caption, last sentence: Provide more explanation how spike counts was performed for (E), and what is the significance of this finding that the differential signal was evident for even this analysis compared to the more complex analysis?

-Figure 5B: What is the number of neurons/animals per CF (related to one of the major comments above). Also, what is the significance of 10 neurons used in (E) and how was that determined?

-Line 243: Add a few words to complete the sentence as, "…selectivity in neural responses to tone frequencies."

-Line 275: The use of a fixed gain of 20dB was not clearly described. Is it that 20 dB was used across all frequencies of the sound evenly?

-Figure 6 (and other figures): Please make sure to note which classifier is being presented in that particular figure.

DISCUSSION:

-There needs to be some caution with a few of the statements.

-Line 330: To say “low selectivity” of neural responses is not fully supported since the hearing loss cases the performance was still decently good. This needs to be possibly reworded based on how the authors address major comment 1 above.

-Line 332: Similarly, saying “much less distinct” for a ~5% change needs to be better worded/justified based on response to major comment 1 above.

-Lines 334, 337, 355, others: Authors say multiple times that capabilities were restored to “normal”. Not sure that statement can be made since there are other deficits that haven’t been restored that are not being captured by the specific neurons and features being analyzed in this study. Maybe better to say “comparable to normal hearing animals based on tested features” or something more along those lines.

-Lines 350: This is a good spot to explain “rollover” and why it happens, and how the findings in this study suggest it was able to explain it based on the neural code.

-Line 383: What do the authors mean by “simple” devices? Why do the findings from this study mean that now simple devices can be made and implemented, i.e., is it because the compression part is the most technologically challenging and power hungry part of a processor that doesn’t need to be included now so devices can be much cheaper and easier to build/code?

-Line 389: This statement is too strong to say “there is now sufficient…” that simple devices can provide comparable hearing to state-of-the-art devices. Please reword especially based on how "simple devices" is defined above.

MATERIALS & METHODS:

-Define RMS.

-Line 437: which frequencies are most relevant and what were actual CFs used in the study (would be nice to show some breakdown of types of neurons and properties, at least CF and up to authors for other properties measured.

-Lines 439-441: These sentences didn’t seem to make sense since that a high pass filter with cutoff of 500 Hz was used and then a lowpass of 300 Hz, which eliminates most of the energy. Please check the wording is correct.

-Line 466: Please provide more details about how many trials were taken for the different stimuli and then how trials were incorporated into the summary plots and classifiers.

-Lines 480-484: Would be helpful to include a bit more detail about what is attack and release times and how compression is performed including what is meant by a ratio of 1 or 2.5.

-Line 499: Provide more detail on how correlation was performed.

-Line 507: Should be “see Figures S1,S2”

-Line 529 to 548: How were the trials of data handled, so were the the spectral densities averaged across trials and how many trials?

SUPPLEMENTARY DOC:

-Figure S1: What do circles and error bars represent?

-All relevant figures: Would be helpful to indicate what is meant by *, **, ***, etc. in the captions, even though more detailed info is in Table S1.

REVIEWER: Hubert H. Lim
University of Minnesota

Thu 10/12/2020
Decision on Article NBME-20-1976A

Dear Prof Lesica,

Thank you for your revised manuscript, "The hearing aid dilemma: amplification, compression, and distortion of the neural code". Having consulted with the original reviewers (whose comments you will find at the end of this message), I am pleased to say that we shall be happy to publish the manuscript in *Nature Biomedical Engineering*, provided that the points specified in the attached instructions file are addressed.

When you are ready to submit the final version of your manuscript, please upload the files specified in the instructions file.

Also, please address all the points and queries in the marked-up reporting summary and in the extended comments (attached).

In the meantime, we will assess the main text in more detail, in particular the title and abstract, for clarity, accessibility and readability. Please expect an update to this message with additional points to address. However, you don't need to wait for the update to act on the instructions in the attached document and to submit the final files.

For primary research originally submitted after December 1, 2019, we encourage authors to take up transparent peer review. If you are eligible and opt in to transparent peer review, we will publish, as a single supplementary file, all the reviewer comments for all the versions of the manuscript, your rebuttal letters, and the editorial decision letters. **If you opt in to transparent peer review, in the attached file please tick the box 'I wish to participate in transparent peer review'; if you prefer not to, please tick 'I do NOT wish to participate in transparent peer review'.** In the interest of confidentiality, we allow redactions to the rebuttal letters and to the reviewer comments. If you are concerned about the release of confidential data, please indicate what specific information you would like to have removed; we cannot incorporate redactions for any other reasons. If any reviewers have signed their comments to authors, or if any reviewers explicitly
agree to release their name, we will include the names in the peer-review supplementary file. More information on transparent peer review is available.

Please do not hesitate to contact me should you have any questions.

Best wishes,

Michelle

Dr Michelle Korda
Senior Editor, *Nature Biomedical Engineering*

P.S. Nature Research journals encourage authors to share their step-by-step experimental protocols on a protocol-sharing platform of their choice. Nature Research's Protocol Exchange is a free-to-use and open resource for protocols; protocols deposited in Protocol Exchange are citable and can be linked from the published article. More details can be found at www.nature.com/protocolexchange/about.

Reviewer #1 (Report for the authors (Required)):

I had no specific points to address myself, and the authors' responses to specific points raised by the other reviewers are appropriate and will benefit the manuscript.

Reviewer #2 (Report for the authors (Required)):

I am satisfied with the authors' responses to me previous comments. I have just a few suggestions for clarification.

Line 54: "residual problems that persist" is redundant. I suggest deleting "residual".

Lines 386-387, "Our results suggest…": This largely begs the question. Of course the results suggest that difficulties during aided listening arise from decrease selectivity of neural responses. Neural responses were the only measure employed in this study. There might be other "difficulties during aided listening" that were not evaluated in the study. Indeed, one can't really argue that the tested neurons were "listening". I suggest that the sentence be simplified to something like: "Our results showed that sound processing algorithms like those used in conventional hearing aids resulted in decreased selectivity by neural responses, which likely could underlie difficulty during aided listening."

Lines 435-436: "decreasing the intensity of incoming sounds" does not solve the problem. One needs to decrease the noise without decreasing the signal, which is the aim of active noise cancellation, etc.

Line 462: "linear amplification without compression" is not "comparable or superior to WERC hearing aids". You need to add something like "performance of linear amplification…"

Reviewer #3 (Report for the authors (Required)):

No additional changes recommended

Reviewer #4 (Report for the authors (Required)):

The authors have sufficiently addressed my concerns. I appreciate the addition of Figure S1; it is an informative figure that addresses my initial concerns. I agree with the authors not to include the sensitivity analysis/plots and they make a valid explanation as to why it makes sense for future research and not in this paper.

I would consider revising some of the sentences below, particularly those in brackets. It is not that I disagree with the authors, but some absolute statements can be more accurately/fairly described. For example, some services are still clinically important for some patients to make sure other issues are not present that an audiologist/clinician should assess/rule out; so to claim they are not essential is a bit strong of a statement. Rather than say "compelling" maybe "growing" or "increasing", since need to ask compelling to who, all players in the field?

Overall, very nice and thorough work that will definitely attract high interest in the clinical and consumer realms and no doubt stir some discussion/debate in the field.

"Cost is a major barrier to hearing aid use, with a typical device in the US costing more than $2000 60. However, most of this cost can be attributed to associated services that are bundled with the device, e.g. testing and fitting. The hardware itself typically accounts for less than $100 (indeed, a recent study demonstrated a prototype device that provided adjustable, frequency-specific amplification costing less than $1 61). [Fortunately, neither the services nor premium features that increase cost are essential 62.] Recent clinical evaluations of over-the-counter personal sound amplification products (PSAPs) have shown that they often provide similar benefit to premium hearing aids fit by professional audiologists 63–65. Thus, there is [now compelling] physiological, psychophysical, and clinical evidence to suggest that inexpensive, self-fitting devices can provide benefit for people with mild-to-moderate hearing loss that is comparable to that provided by current state-of-the-art devices."

Signed for transparency as instructed by the journal:
Hubert Lim, University of Minnesota

Rebuttal 1

We thank the reviewers for their helpful comments. We have revised the manuscript based on their suggestions. Our point-by-point responses to each of their comments and questions are given below.

Reviewer #1 (Report for the authors (Required)):

The issue of 'roll-over' at higher sound intensities is a complex one, but widely accepted in terms of normal hearing. I would like the authors to address briefly in the Discussion what might be a potential solution to the problem of roll-over. Is it de-amplification, and how might this be achieved and impact on hearing function (since it is common to normal and impaired hearing)?

The midbrain lends itself to assessing the 'immediate encoding' aspect of hearing function. What are the implications, if any, for WDRC on other, longer-term features, of neural representations, including 'natural' dynamic-range adaptation? Can the authors comment in the Discussion?

WDRC is independently applied to each hearing aid, and consequently impairs cues for spatial hearing and speech-in-noise processing. This could be commented on, and is actually widely accepted in human studies.

We have added the following section to the Discussion to address these points:

"The mechanisms that underlie rollover are not well understood. One likely contributor is the broadening of cochlear frequency tuning with increasing sound level, which decreases the frequency selectivity of individual auditory nerve fibers and increases the spread of masking from one frequency to another [48]. But rollover is also apparent when speech is processed to contain primarily temporal cues, suggesting that there are contributions from additional factors such as increased cochlear compression at high intensities that distorts the speech envelope or reduced differential sensitivity of auditory nerve fibers at intensities that exceed their dynamic range [36]. The simplest way to avoid rollover is, of course, to decrease the intensity of incoming sounds. There are already consumer devices that seek to improve speech perception by controlling intensity through sealed in-ear headphones and active noise cancellation [49]. But for traditional open-ear hearing aids, complete control of intensity is not an option; such devices must instead employ complex sound transformations to counteract the negative effects of high intensities without necessarily changing the overall intensity itself.

The required sound transformations are likely to be highly nonlinear and identifying them through traditional engineering approaches may be difficult. But recent advances in machine learning may provide a way forward. It may be possible to train deep neural networks to learn complex sound transformations to counteract the effects of rollover in normal hearing listeners or the joint effects of rollover and hearing loss in impaired listeners. These complex transformations could also potentially

address other issues that are ignored by the WDRC algorithm in current hearing aids, such as adaptive processes that modulate neural activity based on high-order sound statistics or over long timescales [50,51]. Deep neural networks may also be able to learn sound transformations that avoid the distortions in binaural cues created by current hearing aids [53,54], enabling the design of new strategies for cooperative processing between devices."

It would be worth commenting a little more directly on the choice of the midbrain for these recordings. What is the value, say, compared to brainstem or cortical?

We have added the following section to the Introduction to address this point:

"The neural code is transformed through successive stages of processing from the auditory nerve to the auditory cortex. At the level of the auditory nerve, some of the important effects of hearing loss that underlie impaired perception are not yet manifest [26], while at the level of the thalamus and cortex, neural activity is modulated by contextual and behavioral factors (e.g. attention) that complicate the study of the general effects of hearing loss on the neural representation of acoustic features. We chose to study the neural code in the inferior colliculus (IC), the midbrain hub of the central auditory pathway that serves as an obligatory relay between the early brainstem and the thalamus. The neural activity in the IC reflects the integrated effects of processing in several peripheral pathways but is still primarily determined by the acoustic features of incoming sounds."

Reviewer #2 (Report for the authors (Required)):

Lines 179-182: This sentence is difficult to follow. CSD decreased with increasing frequency up to 80 Hz (i.e., covering the range of envelope frequencies), meaning that internal noise increased markedly from 8 to 80 Hz. CSD then increased, mostly in parallel with PSD, between 80 and 256 Hz (the range of voice pitch), meaning that internal noise was constant in that range. How can both of those observations be consistent with the statement that internal noise was smallest at frequencies corresponding to the envelope and voice pitch of the speech?

We agree that there was an inconsistency in our original description. We have updated the manuscript to say that the "difference between the $\boldsymbol{PSD}$ and the $\boldsymbol{CSD}$ increased with increasing frequency up to 80 Hz and then remained relatively constant, indicating that the internal noise was smallest (and, thus, the neural responses most reliable) at frequencies corresponding to the envelope of the speech."

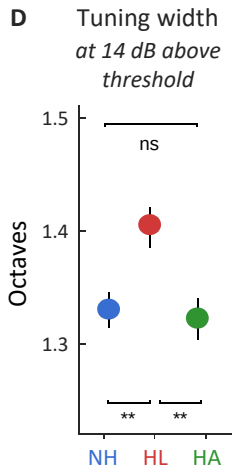
Lines 197 to 208: This paragraph is very difficult to follow. Lines 172 to 190 and Fig. 4B convinced me that internal noise and nuisance noise INCREASE when the relevant covariance decreases. Figure 4C plots covariance, and shows that it decreases with hearing loss and generally increases with a hearing aid. The paragraph starting at line 197, however, states that internal and nuisance noise DECREASE with hearing loss. By my understanding, it is the covariance that decreases, so noise must be INCREASING with hearing loss. To further the confusion, the titles over the panels of Fig. 4C (internal noise, nuisance noise, etc.) indicate something that is the inverse of what is actually plotted.

We understand the source of the confusion. The magnitudes of the actual PSD and CSDs are covariances and so, therefore, are the differences between them that we define as internal noise and nuisance noise. The plots in Figure 4C do not show the magnitudes of a single CSD but rather the magnitudes of the difference between the PSD and a CSD. The leftmost panel in Figure 4C does not show the magnitude of the *CSD* (as the reviewer seems to suggest) but rather the magnitude of the difference between the *PSD* and the *CSD* (the topmost filled area in Figure 4B). Similarly, the second leftmost panel in Figure 4C does not show the magnitude of the $\boldsymbol{CSD}_{\boldsymbol{shuff}}^{\boldsymbol{V,T}}$ but rather the magnitude of the difference between the *PSD* and the $\boldsymbol{CSD}_{\boldsymbol{shuff}}^{\boldsymbol{V,T}}$ (the second topmost filled area in Figure 4B).

To avoid this confusion, we have removed the 'covariance' label from the y-axes in Figure 4C and in any plots of the magnitudes of the response components in other figures. We have also added descriptions to the main text to explicitly state which of the filled areas in Figure 4B are shown in each of the plots in Figure 4C.

Figure 5D and line 231: It is counter-intuitive that tuning width decreases with hearing loss and is increased by a hearing aid. Those observation make it seem as if hearing loss is beneficial. I understand from the text that tuning width is the width of the tuning curve at a fixed sound level, and that the tuning curve is shifted upwards in the case of hearing loss. Nevertheless, I think that it would be helpful to also provide a measure of sharpness of tuning, like ERB or Q10, that is not sensitive to overall shifts of the tuning curve in the SPL dimension.

We have added the following panel to Figure 5 showing tuning width at the same relative intensity for each neuron (14 dB above threshold; the intensity of the tones was varied in steps of 7 dB):



The results are as expected: hearing loss increased tuning width according to this measure and the hearing aid restored it to normal.

Reviewer #3 (Report for the authors (Required)):

No specific points to address

Reviewer #4 (Report for the authors (Required)):

1) The findings are convincing that hearing loss leads to deficits in key sound coding features, which are then not fully recovered from a hearing aid. A critical piece was in showing most of the features tested could be sufficiently recovered when using a linear gain approach. However, one concern is that the amount of deficit seen from the hearing loss is on the order of about 5% and even with that drop, the neurons are still doing a decent/good job at detection and identification of the speech components (e.g., in Fig. 3B/D and Fig. 5F). It isn't clear if this is because the speech material is too easy to detect/discriminate (especially with only 12 consonants) compared to what actual hearing aid users in the real-world are faced with in terms of more complex and open-ended sentences in varying noise environments, and/or if it is because only a short initial portion of 150 ms is used that simplifies the task. There is clearly more dramatic deficits experienced by hearing aid users on a daily basis so it is important for the authors to more clearly link these % changes in performance and how they are clinically relevant and aligned with hearing aid users to maintain the translatability of the data to the human situation. Decoding of consonants/vowels is much easier to do, and is even possible with greatly compromised hearing with auditory implants, compared to trying to reconstruct or recite the actual

words or sentences, in which the latter is more what would happen in the real-world. Making a clearer link is important to be able to justify the second half of the paper showing that the ability to recover the deficits using a linear gain is true for real-world scenarios in better understanding speech, and not just that it helps to recover "decoding" of simple speech features, like consonants.

There are two related questions raised in this point: (1) a quantitative question concerning the seemingly small effect (about 5%) of hearing loss on performance, and (2) a qualitative question concerning the difference between identification of isolated phonemes and real-world speech perception.

Regarding (1)

We understand that 5% may seem small. But this is not atypical of the actual effect size in human listeners with mild-to-moderate hearing loss performing a comparable task. For example, Gordon-Salant (*Phoneme Feature Perception in Noise By Normal and Hearing Impaired Subjects*, *J Speech Hear Res*, 1985) compared performance between normal hearing listeners and those with sloping mild-to-moderate hearing loss in a consonant identification task similar to ours at different speech-to-noise ratios. She found differences ranging from 1% to 10% depending on the specific conditions tested, but for most conditions the difference was about 5%. Also note, while the average effect of hearing loss across all consonants may be about 5%, the size of the effect varies across consonants.

Regarding (2)

It is certainly the case that identification of isolated phonemes and real-world speech perception are different and care must be taken when making inferences about the latter from observations of the former. The reviewer seems to suggest that phoneme identification is easier than real-world speech perception. This may be true under some conditions with a very small phoneme set but, in general, the opposite is true; because real-world speech is redundant, it is often possible to infer the correct word even when it is not heard properly. Isolated random phonemes offer no such opportunity. Numerous studies have found that the speech-to-noise ratio thresholds for phoneme identification are much higher than those for real-world speech perception (for example, Woods et al., *PLoS One*, 2015). However, while phoneme identification may be more difficult than real-world speech perception, performance in the two tasks is highly correlated across listeners. Thus, results from consonant identification tasks are predictive of real-world speech perception.

We have updated the Discussion to include these points.

2) The concept to use large-scale recordings across a major auditory nucleus, such as the IC, and also bilaterally, is a smart and appropriate way to get at the questions posed in this paper. One critical requirement though is that enough of the IC is spanned, at least the central nucleus of the IC, such that any interpretations of the findings are representative of the full deficits or recovery, since the assumption is that it is possible to capture the deficits via the "neural code" because all or most of the information must pass through the IC to higher perceptual centers. In an extreme case, there may be critical regions in the IC that are especially sensitive to deficits in hearing thresholds and amplification, and if those neurons are missed, then the findings do not provide a complete representation of what is going on. The authors did not provide enough information as to which regions and frequency layers were spanned by the arrays, as well as how well the neurons were distributed across the isofrequency dimension of the IC. There are multiple studies showing that many coding properties, including timing, jitter, spectral tuning, thresholds, activation, suprathreshold activation properties, etc. can greatly change depending on location along the isofrequency layer of the IC and even in different nuclei of the IC (see Straka et al., 2014 listed below for examples and many papers on this topic). It seemed from Fig. 5B that a decent span of sites occurred across the frequency axis (though numbers of neurons per CF wasn't clear for confirming even distribution) but couldn't tell how well other regions of the IC were covered. The authors may have sufficiently spanned and checked coverage or found through other analyses that their coverage was sufficient for the purposes of their analyses. It is possible that even greater deficits might have occurred if recording from certain regions and/or some deficits maybe less recovered if looking at other regions. The current state of the paper is already impressive in terms of demonstrating the deficits of hearing loss on selectivity of speech perception and how linear gain can recover some of those deficits. However, understanding this coverage will enable the ability to interpret the findings more comprehensively. If location of sites are not well known, which is totally understandable since doing histology and tracking all sites is very challenging, one possibility is to take the current data and perform some sensitivity analyses. For example, to take half of the more medial IC data sites and see how they perform to the more lateral sites (but having similar CFs), or other location comparisons. These findings would show that the overall findings are not so sensitive to IC location or if they are, then how to interpret the data presented. Also, further justification is needed why only 150 neurons were used for the classifications, and which neurons (i.e., what were their characteristics in terms of CF, tuning, etc. to understand what types of neurons are doing the decoding for interpreting the findings).

There are two related questions raised in this point as well: (1) a question as to how much of the IC was covered in our recordings and (2) a question as to whether our results varied across different locations within the IC.

Regarding (1)

We have added a new supplementary figure describing the location of our recordings.

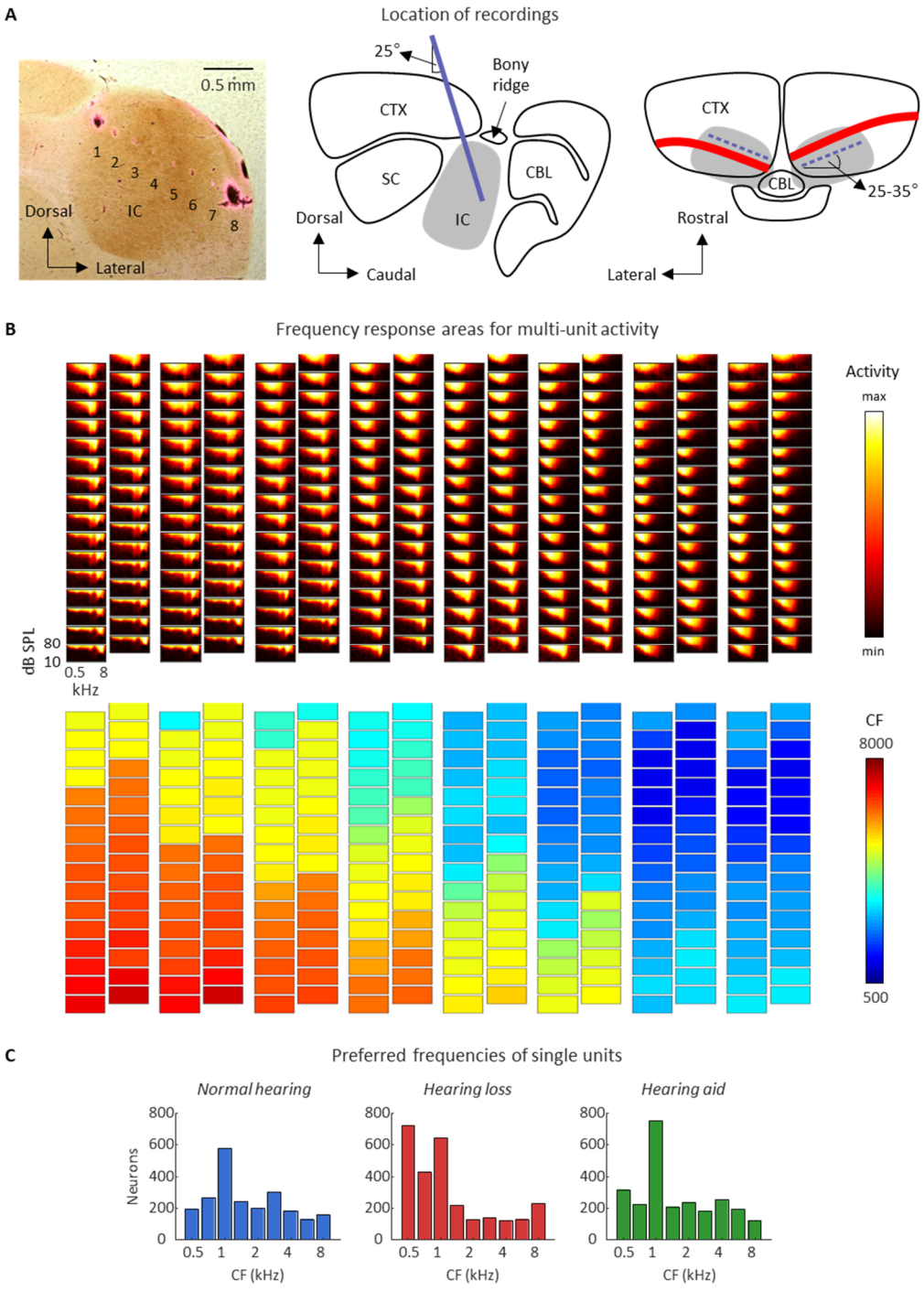


Figure S1: Location of recordings within the inferior colliculus

The geometry of our electrode arrays was designed specifically to match the layout of the speech-sensitive area in the central nucleus of the gerbil IC. The recording sites spanned a plane measuring 1.4 mm x 0.45 mm. When oriented approximately parallel to the coronal plane, one array covered the entire mediolateral extent of the central nucleus in one hemisphere and enough of its dorsoventral extent to sample from the relevant frequency layers (preferred frequencies up to ~10 kHz). (A) The left panel shows a merge of brightfield and fluorescent images of a coronal section taken for cytochrome oxidase and DiI staining, respectively (the electrode array was coated with DiI and, thus, the fluorescent areas indicate the position of each of the 8 shanks of the array within the section). The approach to the IC was constrained by the locations of a large blood vessel on the surface of the brain and a bony ridge that protrudes from the lateral wall of the skull between the cortex and midbrain, both of which varied from animal to animal and across hemispheres in the same animal. The electrode arrays were rotated by a fixed angle of 25° relative to the coronal plane about the mediolateral axis to avoid the bony ridge (see middle panel with array (blue), IC (gray), and surrounding structures) and a variable angle of 25-35° relative to the coronal plant about the dorsoventral axis to align with the blood vessel (see right panel with array shanks (blue), blood vessels (red), IC (gray), and surrounding structures). The position of the electrode arrays along the mediolateral axis was fixed but the position along the rostrocaudal axis was varied from animal to animal and across hemispheres in the same animal to avoid the blood vessel. Thus, across animals and hemispheres, the recordings sampled the full three-dimensional volume of the central nucleus. (B) MUA recorded in the inferior colliculus during the presentation of tones. The top panel shows the MUA FRAs for all 256 channels on one electrode array from an example normal hearing animal. The bottom panel indicates the center frequency (the frequency for which the mean MUA was more than 3 standard deviations above the mean MUA during silence at the lowest intensity) for each channel. (C) The distribution of center frequencies (CFs) of single units in our sample for which responses to tones were recorded for animals with normal hearing (left; n = 2249) and animals with hearing loss without (middle; n = 2959) and with (right; n = 2664) a hearing aid. The CF was defined as the frequency at which the response to a tone was significantly greater than responses recorded during silence at the lowest intensity ($p < 0.01$ for Poisson-distributed spike counts). The overrepresentation of 1 kHz is consistent with the oversized “pars lateralis” of the IC in the gerbil (Cant, Front. Neural Circuits, 2013). The distribution of CFs shifted toward lower frequencies with hearing loss, consistent with the observed effects of noise-induced hearing loss on peripheral tuning (Henry et al., J. Neurosci, 2016), but was similar to normal with the hearing aid.

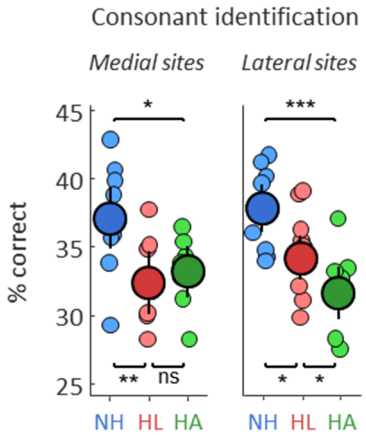
Regarding (2)

A first point to note is that we have previously published a study of the variation in response properties across isofrequency layers in the gerbil IC:

Schnupp JWH, Garcia-Lazaro JA and Lesica NA (2015) Periodotopy in the gerbil inferior colliculus: local clustering rather than a gradient map. Front. Neural Circuits 9:37.

We did not see evidence of gradients or distinct regions as in other species but rather only local clustering. However, our study considered only a limited number of response properties, mostly related to the coding of amplitude modulations. There may well be other response properties that are relevant for the effects of hearing loss and hearing aids on the coding of speech that we did not examine.

We performed the sensitivity analysis suggested by the reviewer to determine whether our results varied across an isofrequency layer. We took all neurons from our sample with a preferred frequency of 1 kHz (the most frequent preferred frequency in our recordings, as shown in the new supplementary figure above) and divided them into two groups based on whether they were recorded from the medial or lateral half of the IC. We then performed the same classification analysis for each group as we did for our full sample in Figure 3D. As shown in the figure below, the results were similar for both groups and similar to those of our original analysis of the full sample, i.e. hearing loss impaired consonant identification and the hearing aid did not restore it to normal.



Performance of a support vector machine classifier trained to identify consonants based on population single-trial responses (represented as spike count vectors with 5 ms time bins) to speech in quiet at 62 dB SPL. Populations of 25 single units were chosen at random, without replacement, from those with preferred frequencies of 1 kHz recorded in either the medial (left panel) or lateral (right panel) half of the IC (values for each population are shown along with mean ± 95% confidence intervals derived from bootstrap resampling across populations; *** indicates $p < 0.001$, ** indicates $p < 0.01$, * indicates $p < 0.05$, ns indicates not significant).

We would be happy to include these results in the manuscript if needed, but we feel this analysis is only scratching the surface of an important issue. We share the reviewer's interest in understanding how the effects of hearing loss and hearing aids on neural responses to speech vary across the population of IC neurons. The IC is highly heterogeneous -- spatial location within an isofrequency layer is just one of many dimensions along which IC neurons vary. Others include temporal dynamics (onset,

sustained, etc.), binaural preference, and, of course, preferred frequency. It is likely that the effects of hearing loss and hearing aids vary across these dimensions and understanding this variation may have practical implications for the development of treatments. But we believe this is an issue that deserves a comprehensive neuroscience study of its own (and, indeed, we plan to undertake one).

Finally, there seems to be a misunderstanding with respect to the neurons used for classification. Populations of 150 were chosen by sampling at random, without replacement, from across all animals until no further populations could be formed. No neurons were excluded (aside from those few that remained after the last full population was chosen). We have added this information to the Results and clarified it in the Methods.

-Lines 46-49: This might be a good spot to explain the concept of rollover and why it is believed to cause deficits even in normal hearing listeners. This is later talked about in the Results and then in the Discussion but it isn't clear why it is happening. Either clarify what is "rollover" and why it presumably happens here or when discussing the results (the latter may be better since Intro is already getting long and would make more sense when actually presenting the relevant data).

We have added the following to the Discussion:

"The mechanisms that underlie rollover are not well understood. One likely contributor is the broadening of cochlear frequency tuning with increasing sound level, which decreases the frequency selectivity of individual auditory nerve fibers and increases the spread of masking from one frequency to another [48]. But rollover is also apparent when speech is processed to contain primarily temporal cues, suggesting that there are contributions from additional factors such as increased cochlear compression at high intensities that distorts the speech envelope or reduced differential sensitivity of auditory nerve fibers at intensities that exceed their dynamic range [36]"

RESULTS:

-Quite impressive of using 512 sites per animal. Yet, further info is needed why with all of the animals, the total numbers come out to be just 2000 or 3000 sites in Table S1, i.e., how was it determined which sites to keep or discard?

There seems to be a misunderstanding. We did not analyze the raw recordings from each site directly. We analyzed single-unit responses extracted from the raw recordings (as described in the *Spike Sorting* section of the Methods). Our yield from each 512-channel recording was typically between 100 and 200 single-units. No neurons were excluded from the analysis.

-Figure 3 caption, last 2 sentences: what does “each population” refer to? As with all the other figures, it is sometimes confusing to know what each point means in terms of number of animals or number of neurons or number of trials. Much of this information is in Table S1 but provide some basic info in the captions would be helpful to keep things straight.

We have added clarification throughout the Results, figure legends, and Methods describing how populations were formed and the number of trials used for analysis.

- Two identical trials of the full set of syllables were presented for each condition (e.g. 62 or 82 dB SPL, with or without second talker or multi-talker noise, with or without hearing aid). All results reported are based on analysis of only the first trial, except for those relying on computation of cross spectral densities or noise correlations for which both trials were used.
- Populations were formed by sampling at random, without replacement, from neurons from across all animals for a given hearing condition until there were no longer enough neurons remaining to form another population.
- For most analyses in the Results, population size of 150 neurons was used.

-Figure 4C (and other similar types of plots): Are the points and 95% CIs for averages across trials per neuron or average across all trials and neurons? Please clarify also why it was done that way.

The points and the CIs in Figure 4C are for averages across neurons. We have added this information to all relevant figure legends.

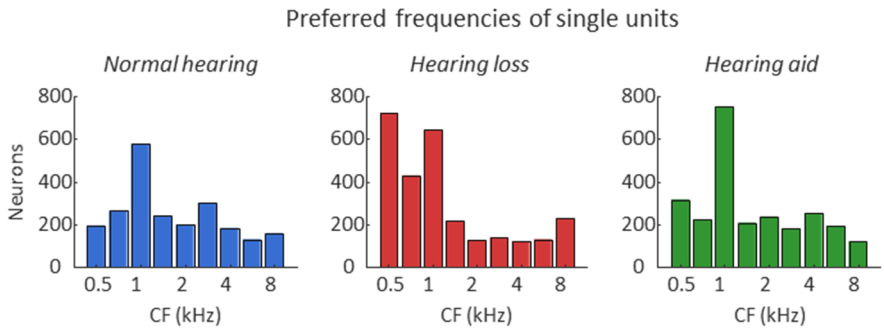
As described in the previous answer, we did not average across trials.

-Figure 4 caption, last sentence: Provide more explanation how spike counts was performed for (E), and what is the significance of this finding that the differential signal was evident for even this analysis compared to the more complex analysis?

We have updated the Results and the figure legend to provide additional detail.

-Figure 5B: What is the number of neurons/animals per CF (related to one of the major comments above). Also, what is the significance of 10 neurons used in (E) and how was that determined?

We have added the following panel to Supplementary Figure 1: (see legend above)



We have added the following to the Figure 5 legend:

"A population size of 10 was used to allow for accurate classifier performance for all conditions while avoiding the 100% ceiling for any condition."

-Figure 6 (and other figures): Please make sure to note which classifier is being presented in that particular figure.

We have added the suggested information to all relevant figure legends.

DISCUSSION:

-There needs to be some caution with a few of the statements.

-Line 330: To say "low selectivity" of neural responses is not fully supported since the hearing loss cases the performance was still decently good. This needs to be possibly reworded based on how the authors address major comment 1 above.

We have changed this to:

"Our results suggest that difficulties during aided listening with mild-to-moderate hearing loss arise primarily from the decreased selectivity of neural responses. "

-Line 332: Similarly, saying "much less distinct" for a ~5% change needs to be better worded/justified based on response to major comment 1 above.

We have changed this to:

"While a hearing aid corrected many of the changes in neural response patterns that were caused by hearing loss, the average response patterns elicited by different consonants remained less distinct than with normal hearing."

-Lines 334, 337, 355, others: Authors say multiple times that capabilities were restored to "normal". Not sure that statement can be made since there are other deficits that haven't been restored that are not being captured by the specific neurons and features being analyzed in this study. Maybe better to say "comparable to normal hearing animals based on tested features" or something more along those lines.

We agree that a blanket statement about responses being returned to normal in animals with hearing loss would be inappropriate, as there are many response properties that we have not examined. But in each of the cited instances, we are explicit about exactly which response properties are returned to normal and these specific assertions are well supported by our analyses.

-Lines 350: This is a good spot to explain "rollover" and why it happens, and how the findings in this study suggest it was able to explain it based on the neural code.

We have added the following section to the Discussion:

"The mechanisms that underlie rollover are not well understood. One likely contributor is the broadening of cochlear frequency tuning with increasing sound level, which decreases the frequency selectivity of individual auditory nerve fibers and increases the spread of masking from one frequency to another [48]. But rollover is also apparent when speech is processed to contain primarily temporal cues, suggesting that there are contributions from additional factors such as increased cochlear compression at high intensities that distorts the speech envelope or reduced differential sensitivity of auditory nerve fibers at intensities that exceed their dynamic range [36].The simplest way to avoid rollover is, of course, to decrease the intensity of incoming sounds. There are already consumer devices that seek to improve speech perception by controlling intensity through sealed in-ear headphones and active noise cancellation [49]. But for traditional open-ear hearing aids, complete control of intensity is not an option; such devices must instead employ complex sound transformations to counteract the negative effects of high intensities without necessarily changing the overall intensity itself.

The required sound transformations are likely to be highly nonlinear and identifying them through traditional engineering approaches may be difficult. But recent advances in machine learning may provide a way forward. It may be possible to train deep neural networks to learn complex sound

transformations to counteract the effects of rollover in normal hearing listeners or the joint effects of rollover and hearing loss in impaired listeners. These complex transformations could also potentially address other issues that are ignored by the WDRC algorithm in current hearing aids, such as adaptive processes that modulate neural activity based on high-order sound statistics or over long timescales [50,51]. Deep neural networks may also be able to learn sound transformations that avoid the distortions in binaural cues created by current hearing aids [53,54], enabling the design of new strategies for cooperative processing between devices."

-Line 383: What do the authors mean by "simple" devices? Why do the findings from this study mean that now simple devices can be made and implemented, i.e., is it because the compression part is the most technologically challenging and power hungry part of a processor that doesn't need to be included now so devices can be much cheaper and easier to build/code?

See next answer

-Line 389: This statement is too strong to say "there is now sufficient…" that simple devices can provide comparable hearing to state-of-the-art devices. Please reword especially based on how "simple devices" is defined above.

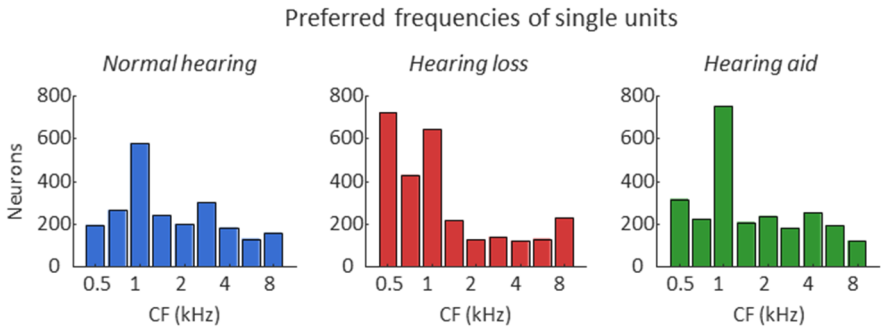
We have expanded this section of the Discussion as follows:

"Cost is a major barrier to hearing aid use, with a typical device in the US costing more than $2000 [60]. However, most of this cost can be attributed to associated services that are bundled with the device, e.g. testing and fitting. The hardware itself typically accounts for less than $100 (indeed, a recent study demonstrated a prototype device that provided adjustable, frequency-specific amplification costing less than $1 [61]). Fortunately, neither the services nor premium features that increase cost are essential [62]. Recent clinical evaluations of over-the-counter personal sound amplification products (PSAPs) have shown that they often provide similar benefit to premium hearing aids fit by professional audiologists [63–65]. Thus, there is now compelling physiological, psychophysical, and clinical evidence to suggest that inexpensive, self-fitting devices can provide benefit for people with mild-to-moderate hearing loss that is comparable to that provided by current state-of-the-art devices."

MATERIALS & METHODS:

-Line 437: which frequencies are most relevant and what were actual CFs used in the study (would be nice to show some breakdown of types of neurons and properties, at least CF and up to authors for other properties measured.

We have added the following panel to Supplementary Figure 1: (see legend above)



-Lines 439-441: These sentences didn't seem to make sense since that a high pass filter with cutoff of 500 Hz was used and then a lowpass of 300 Hz, which eliminates most of the energy. Please check the wording is correct.

It is correct. The absolute value was taken between the two filtering operations. The high pass filter -> absolute value -> low pass filter process is a simple way to extract the low-frequency envelope of a high-frequency carrier.

-Line 466: Please provide more details about how many trials were taken for the different stimuli and then how trials were incorporated into the summary plots and classifiers.

This is now clarified in the Methods:

"Two identical trials of the full set of syllables were presented for each condition (e.g. 62 or 82 dB SPL, with or without second talker or multi-talker noise, with or without hearing aid). All results reported are based on analysis of only the first trial, except for those relying on computation of cross spectral densities or noise correlations for which both trials were used."

-Line 499: Provide more detail on how correlation was performed.

We have added further detail to the legend of the relevant supplementary figure.

-Line 529 to 548: How were the trials of data handled, so were the the spectral densities averaged across trials and how many trials?

This is now clarified in the Methods:

"Two identical trials of the full set of syllables were presented for each condition (e.g. 62 or 82 dB SPL, with or without second talker or multi-talker noise, with or without hearing aid). All results reported are based on analysis of only the first trial, except for those relying on computation of cross spectral densities or noise correlations for which both trials were used."

Reviewer comment:

"Lines 435-436: “decreasing the intensity of incoming sounds” does not solve the problem. One needs to decrease the noise without decreasing the signal, which is the aim of active noise cancellation, etc."

Author response:

The full statement is: "The simplest way to avoid rollover is, of course, to decrease the intensity of incoming sounds."

Rollover is, by definition, the decrease in speech-in-noise understanding with increasing overall intensity for *fixed* speech-to-noise ratio. So decreasing the overall intensity of incoming sounds, both speech and noise, would indeed solve the problem of rollover.

Increasing the speech-to-noise ratio as the reviewer suggests would, of course, also help in general. But this section of the discussion is about how to decrease the impact of rollover per se, not how to improve perception in general.