

Dear reviewers,

Thank you for the careful review of our registered report and the helpful comments. We have addressed the issues raised and provide detailed responses. Please see the uploaded file. In the report, we have highlighted all changes in yellow.

Recommenders:

With regard to registered reports formatting, we have a further few items for the authors to address:

1. Version control

After each revision, it is ideal if you replace the old pdf with a new clean (no track changes) pdf on OSF to enable PCI-RR to record the version number. The tracked changes document should be uploaded to PCI-RR separately in the same page where you submit your reply to the reviewers and recommenders.

Thanks for the suggestion. We added a new clean version of our report in OSF.

2. Study design table

In the final column of the Study design table, we encourage you to think about the theoretical implications of either finding or not finding a difference in TVA activation between groups.

This section has been updated:

“The rejection of the null hypothesis would suggest that ASD individuals suffer from low level deficits in voice processing, which may eventually lead to higher order social dysfunction. The failure to reject would suggest that low level voice processing may not be impaired in ASD and that social difficulties may arise from higher-level dysfunctions. In all cases, exploratory analyses will be designed to inform the interpretation of each outcome.”

Minor:

Typo on page 3: “founding resources”

The typo has been fixed:

“funding resources”

Reviewer 1

I wonder why the authors would explore individual characteristics of only those ASD participants, who do not show any significant voice-specific responses. As no previous study specifically looked on links between brain processing of voice and individual participant’s characteristics, a dimensional approach of looking on the whole continuum of vocal brain responses in ASD could be more informative. A dimensional approach would also provide more power compared to splitting up the sample into responders and non-responders for categorical comparisons (particularly if the non-responders are expected to be in the minority).

We thank the reviewer for this insightful suggestion. As these analyses are exploratory, no specific analysis is preregistered. However, we will consider this suggestion when taking part in the exploratory phase by taking a dimensional approach. The manuscript has been modified accordingly:

“Exploratory analyses will be conducted in order to investigate links between voice brain processing and individual participants’ characteristics.”

The authors provide a high level of methodological detail in their Stage 1.

How is the alpha level of .02 justified? Is it based on any previous findings? If yes, please state this.

The .02 alpha level has been defined following the PCI-RR friendly journals requirements (namely an alpha level of .02 and power of .9 is required by the Cortex journal).

- Do the authors plan to assess hearing abilities and auditory processing abilities to ensure that different brain responses are not due to other factors than ASD.

We will assess hearing abilities using an audiogram and compare between group levels of hearing abilities in order to ensure that differences in cerebral activation are not better explained by hearing abilities.

“Participants’ hearing abilities will be assessed using an audiogram in order to discard an eventual influence of hearing abilities in between group differences in voice processing.”

- Do the authors plan to assess clinical profiles of non-autistic controls? Will they screen for other forms of neurodiversity? This is particularly important since the authors coin their control group as “typically developing”. If no such screening is planned, I recommend using the terms “non-ASD” or “non-autistic” as preferred by the community. Please pay more attention to more current vocabulary for describing ASD

We thank the reviewer for emphasizing this point. Participants from the control group will not be screen for their clinical profile. Consequently, the control group has been renamed ‘non-ASD’ group.

Please provide the description and examples of non-vocal sounds.

The methods section describing the sounds used in the voice localizer task has been expended.

“Non-vocal sounds consist of sounds from nature (e.g., wind, sea waves, ...), animals, classical music, and man-made objects (e.g., cars, clocks, ...).”

Reviewer 2

The authors hypothesize that a lower proportion of individuals with ASD will show a specific response to vocal sounds in the STS/STG (i.e., a TVA activation) than in the typically developing sample. This hypothesis is plausible and supported by previous research that has shown that individuals with ASD have difficulty processing social cues, including voices. This suggests that they may also have difficulty showing a specific response to vocal sounds in the STS/STG. It is not clear to me whether the authors intend to interpret a negative result as evidence that the effect is absent. If so, using an inferential method that is capable of drawing such a conclusion (i.e. Bayesian hypothesis or frequentist equivalence testing) would be advised.

We thank the reviewer for this comment. Although we acknowledge that our method does not allow us to draw the conclusion that ASD and non-ASD populations does not differ in the proportion of responders to vocal sounds, we do not plan to interpret a negative result as an absence of effect.

I do have some questions about the experimental task and the participant characteristics: 1) what will the duration of the presented sounds and the sound intensity (dB) levels be? 2) At what time intervals will the sounds be presented and is this a fixed rate or jittered? 3) Will the sounds be randomly presented to participants or in a fixed order?

The methods section describing more precisely the voice localizer task has been extended to include the mentioned elements. Concerning the intensity level at which sounds will be displayed, we cannot measure it precisely using a decibel meter, as we are using the MR Confon system which operates exclusively within the scanner. However, settings from the MR Confon indicates that sounds will be displayed at an intensity level of roughly 88dB. This level was determined based on pilot sessions. This specific value was chosen because it enabled hearing the sounds above the noise from the scanner while keeping it at a comfortable level.

“The order of the blocks was determined pseudorandomly when the experiment has been designed, but the resulting order remained fixed for all participants. Within each block, stimuli are separated from each other by a delay of at most 400ms. Sounds will be displayed using the MR Confon system at an intensity level of roughly 88dB.”

4) I think it would be good to have an idea as to how well participants were engaged with the “task”. Some participants may be more attentive than others to the sounds and this may impact the results; especially if this where to be a between group difference. Would it be possible to add some sort of response task interspersed between the passive task stimuli to check if participants are (still) paying attention? Or, if that is not possible/desirable, ask a few questions after the experiment on how engaged participants were? This data then can be used to find out if differences in terms of task engagement/attention occurred between both groups for instance

The authors thank the reviewer for this insightful comment. As for reproducibility it is not desirable to change the experimental task, we added a memory task at the end of the experiment: participant will be presented with some of the sounds they heard during the voice localizer task. They will be asked to indicate, for every sound, whether they remembered hearing it during the task. This will serve as a proxy for attention during the task. Instructions for the voice localizer task has been updated accordingly. In addition, participants will be asked to indicate in a 5-points Likert scale their level of engagement with the sounds during the voice-localizer task.

“Participants are asked to close their eyes and listen carefully to the sounds. A memory task will be administered to the participants after the scanning session. More precisely, the participants will be presented with some of the sounds which were displayed during the experiment. For each sound, the participants will be asked to indicate whether they remember hearing it during the scanning session. This task allowed to monitor the attention level during the task. In addition, participants will be asked to rate their level of engagement with the sounds during the voice-localizer task on a 5-points Likert scale.”

“Behavioral data analysis

Global scores from the memory task will be compared between groups using a one-way analysis of variance (ANOVA). “

5) It is stated that participant's IQ will be evaluated. Does this mean participants will be matched on IQ as well? It may not be feasible, but if participants would not only be matched on gender and age but also on IQ that would be preferable.

Although we acknowledge that matching the participants based on IQ would be interesting, we do not plan to match participants on this measure because of two reasons. First, we are uncertain that it may be feasible, and second, we may conduct exploratory analyses that investigate how the individual specific response to vocal sounds relates to the individual's cognitive and clinical profile. This way, IQ level will be fully accounted for in exploratory analyses.

The only thing that was not clear to me in the fMRI scanning procedure is the phrase how "structural anomalies or irreparable artifacts will be replaced". Does this mean that these will not be taken into account at all or will they be replaced by other data of some sort?

Thank you for this comment which allowed us to clarify the procedure to handle artifactual data. In fact, the artifactual data will not be included in the analyses and another participant will be recruited in order to replace the data. This has been emphasized in the main text:

"Such data will be replaced (i.e., another participant will be recruited)."

We thank the reviewers for their insightful comments which enabled us to further improve the quality of our manuscript and of the resulting study.

Sincerely,

Raphaël Gautier