

## Delta(but not theta)-band cortical entrainment involves speech-specific processing

Nicola Molinaro & Mikel Lizarazu

---

**Review timeline:**

Submission date:	29 June 2017
Editorial Decision:	05 September 2017
Revision received:	02 November 2017
Editorial Decision:	04 December 2017
Revision received:	13 December 2017
Accepted:	19 December 2017

---

Editor: Ali Mazaheri

1st Editorial Decision

05 September 2017

Dear Dr. Molinaro,

Your manuscript was reviewed by two external reviewers as well as by our Guest Section Editor, Dr. Ali Mazaheri, and ourselves. Based on these reviews, we regret to inform you that we are not able to accept your manuscript for publication in EJN in its present form. However, the research described in your manuscript is potentially of interest and therefore we invite you to resubmit a much revised version. As you will read, both reviewers share a significant concern that based on the design of the experiment it's impossible to link the findings of the manuscript with higher-order language processing, the central claim of study. We would like to give you the opportunity to make a better case for this central claim. Otherwise, you will need to substantially revise your stance and provide a more measured interpretation of your findings.

The reviewers also raise a number of other critical issues that will need to be addressed if the paper is to ultimately be successful here at EJN. In particular, reviewer 2 raises concerns about the validity of the connectivity analysis in source space, given the results weren't significant in sensor space. Here the authors should provide an explanation of why they think this is happening. We would also like to see some of the changes in power of oscillatory activity induced by the onset of stimuli in sensor space as well as source space.

Also, please attend to the following issues:

- 1) Provide the age-range of your participants
- 2) Provide an explicit statement regarding Conflict-of-interest at the end of the paper in its own section
- 3) Provide a similarly segregated data sharing statement
- 4) Provide a statement regarding the contributions from authors

Thank you for submitting your work to EJN.

Kind regards,

John Foxe & Paul Bolam  
co-Editors in Chief, EJN

Reviews:

Reviewer: 1 (Andrew Dimitrijevic, Sunnybrook Health Sciences Centre, Canada)

Comments to the Author

The manuscript "Delta-band cortical entrainment to speech involves high order language related brain regions" describes an MEG study where the authors used speech and non-speech stimuli to look at the cortical entrainment mechanisms. They found that auditory cortex follows the stimuli well for all types of stimuli (speech, 2 and 4 Hz AM white noise). However, left frontal regions were preferentially activated only the speech stimulus in the delta frequency. The authors conclude that higher order cortical areas

preferentially process linguistic stimuli.

The paper is clearly written for the most part, the content is appropriate for the readership of European Journal of Neuroscience.

There are some concerns and clarifications that need to be address before accepting the manuscript.

Choice of control stimuli: The authors chose to use 2 and 4 Hz amplitude modulated white noise. I understand the rationale for using these stimuli- they are at the frequency range they looking at. However, I believe it to be too limited. Why not simply use white noise with the identical amplitude envelope of the sentences? This comparison would be more valid. How do you know that the left frontal regions are not just sensitive to a "rich" amplitude modulation near the 2 Hz range? I.e., different modulation depths or varying AM rates. Irrespective of the linguistic aspect of the sentences. I think this is a major flaw in the design of the experiment. I am not questioning the results of the experiment (I think these are very valid) but rather I am questioning the implications of the results and interpretation. I am not convinced that speech/linguistic aspects are "special". There was a preferential activation of the left frontal activation in figure 1 for the delta frequency and when you do the subtraction this is the area becomes significant. I think the authors need to tone down the interpretation. The contrast is really speech versus 2 Hz AM (not necessarily "non-linguistic", this is a bit of stretch).

Other issues:

Top plots of figures 1 and 2. What are these lines?

Why are there no differences in the lateralization indices' in figure 3 (i.e., for the speech). Which ROIs are these based on?

What was the rationale for doing a correlation between speech and "non-linguistic" in figure 3?

Page 8: "First, we identified the frequencies that showed significant coherence" This is very vague. How was this conducted, which frequencies, how did this change per subject, per frequency, which ROIs (sensor versus source space)?

Page 11: "at the sensor level..." Figure 2 is not at sensor level.

Page 11: "positive correlations..." which ROIs?

Reviewer: 2 (Sarah Tune, University of Lübeck, Germany)

Comments to the Author

The authors present results from an MEG studies on the functional role(s) of delta and theta band neural entrainment to the envelope of speech input in the service of speech processing. To this end, the authors compared the synchronization of brain dynamics and auditory inputs for sentence materials and amplitude-modulated noise (at 2 and 7 Hz) as indexed by coherence between the two signals. The key result of the study, however, stems from the analysis of coherence at the source level. Here, the authors found that brain-speech synchronization in the delta but not the theta band involved regions in the inferior frontal region, which they interpret as suggesting that delta band entrainment 'reflects high-level processing involving language related regions'.

The question of how neural oscillations in different frequency bands are utilized in the service of complex processes such as speech processing and comprehension is a timely and important topic. As such I'm generally excited to see studies trying to disentangle and characterize the contributions of neural dynamics to speech processing.

Main concerns:

Unfortunately, I'm unsure how the results of the present study will further our understanding of the contributions of delta and theta oscillations to speech processing. This is due to two main reasons: First, I see a general disconnect between the research question asked (i.e., targeting the functional dissociation of delta and theta band "entrainment" to the speech envelope in support of speech processing) and the chosen research design and analysis. I fail to see how exactly the comparison of neural responses to linguistic vs. non-linguistic (amplitude-modulated noise) would help clarify the functional role of delta vs. theta 'entrainment' to speech processing. The neural tracking of auditory input (whether it may be speech or non-linguistic input such as amplitude-modulated noise) has been shown by many previous studies; the present study may be the first to contrast neural responses to these two input types in the same participants but I'm not sure why this should be a crucial advantage over previous studies. Moreover, the analysis itself is fairly basic, so overall I don't think that the results of the present studies provide any truly novel evidence in this respect.

The main conclusion that the authors draw, however, is based on the results of the analysis of coherence at the source level. It is the interpretation of these results that are my second main concern regarding this study. The analysis of coherence at the source level feels very post hoc given the non-conclusive results at

the sensor level. Moreover, the interpretation of the involvement of neural entrainment of delta-band oscillations as supporting 'high-level processing' given the involvement of inferior frontal regions is a severe case of drawing reverse inferences. Based on the research design and analysis, there is no way to say for certain what the involvement of regions in the inferior frontal cortices could reflect. To relate them to higher-level (language) processing is to base the interpretation of those previous results that simply appear to fit the story best. To me, the very fact that the strongest claim is based on the most ambiguous evidence in this study shows once again the disconnect between the research question the authors wish to answer and the methods they are choosing to do so.

Further comments:

Apart from the main concerns addressed above, I thought the manuscript was generally lacking clarity especially with respect to the methods description:

Methods / Results:

- p. 5: why did trials in the linguistic but not the non-linguistic block start with a 1 sec auditory tone?
- p. 6: 'sentence presentation order pseudo-randomized' -> with respect to which constraints=
- p. 6: how did the author motivate to use 2 and 7 Hz amplitude modulation?
- p. 8: given that the terminology on the literature is not always consistent, the authors should make more explicit which measure of coherence they are using
- p.8: there is no description on the preprocessing of the audio signals to derive the envelopes
- p.9: I found the description of the permutation test to be rather confusing. This was mostly due going back and forth between contrasts between conditions and the conditions themselves. It seemed like the authors were interested in the significance of specific contrasts between conditions, yet the permutation distribution was essentially built from the permutation of individual conditions.
- p. 10: the authors should motivate why they chose to investigate of lateralization of coherence values; in the current version of the manuscript this is not clear at all...

Discussion:

- the discussion provides a very in depth description of previous evidence on the possible functional role of neural entrainment in the delta band to speech processing but its relation to the present results purely speculative.

Authors' Response

02 November 2017

## RESPONSE TO REVIEWERS

Your manuscript was reviewed by two external reviewers as well as by our Guest Section Editor, Dr. Ali Mazaheri, and ourselves. Based on these reviews, we regret to inform you that we are not able to accept your manuscript for publication in EJN in its present form. However, the research described in your manuscript is potentially of interest and therefore we invite you to resubmit a much revised version. As you will read, both reviewers share a significant concern that based on the design of the experiment it's impossible to link the findings of the manuscript with higher-order language processing, the central claim of study. We would like to give you the opportunity to make a better case for this central claim. Otherwise, you will need to substantially revise your stance and provide a more measured interpretation of your findings.

The reviewers also raise a number of other critical issues that will need to be addressed if the paper is to ultimately be successful here at EJN. In particular, reviewer 2 raises concerns about the validity of the connectivity analysis in source space, given the results weren't significant in sensor space. Here the authors should provide an explanation of why they think this is happening. We would also like to see some of the changes in power of oscillatory activity induced by the onset of stimuli in sensor space as well as source space.

RESPONSE: We thank the Editor for the possibility of revising our work and allowing a resubmission. In the present version of the Manuscript:

- We added more experimental evidence supporting our claim that higher-order brain regions (such as the left IFG) get synchronized with the delta speech rhythmicity but not for theta speech-brain coupling.
- We now report an additional Experiment (with a more controlled control condition, following suggestion by Reviewer 1) in which we replicate the findings of the first experiment.

- We also report all the analyses at the sensor-level, showing that the topographical distribution of the effects is also different between conditions for the delta band (in both Experiments).
- Importantly, we investigated speech-brain coupling effects and we did not perform a connectivity analysis (connectivity based on coherence is in fact more susceptible to power modulations). This misunderstanding is likely due to the fact that we did not fully specify our methodological steps in the Methods of the previous version of the Manuscript. We now add more information in this regard to clarify the coherence analyses reported in this study.

Also, please attend to the following issues:

- 1) Provide the age-range of your participants Done
- 2) Provide an explicit statement regarding Conflict-of-interest at the end of the paper in its own section Done
- 3) Provide a similarly segregated data sharing statement Done
- 4) Provide a statement regarding the contributions from authors Done

When revising the manuscript, please embolden or underline major changes to the text so they are easily identifiable and DO NOT leave 'track change' formatting marks in your paper. If the changes made are extensive, please also provide an unmarked version.

RESPONSE: Since the changes were extensive (~ 90% changes were made to the previous version of the Manuscript), we do not highlight the modifications to the revised versions of the manuscript

Reviewer: 1

#### Comments to the Author

The manuscript "Delta-band cortical entrainment to speech involves high order language related brain regions" describes an MEG study where the authors used speech and non-speech stimuli to look at the cortical entrainment mechanisms. They found that auditory cortex follows the stimuli well for all types of stimuli (speech, 2 and 4 Hz AM white noise). However, left frontal regions were preferentially activated only the speech stimulus in the delta frequency. The authors conclude that higher order cortical areas preferentially process linguistic stimuli.

The paper is clearly written for the most part, the content is appropriate for the readership of European Journal of Neuroscience.

There are some concerns and clarifications that need to be address before accepting the manuscript.

Choice of control stimuli: The authors chose to use 2 and 4 Hz amplitude modulated white noise. I understand the rationale for using these stimuli- they are at the frequency range they looking at. However, I believe it to be too limited. Why not simply use white noise with the identical amplitude envelope of the sentences? This comparison would be more valid. How do you know that the left frontal regions are not just sensitive to a "rich" amplitude modulation near the 2 Hz range? I.e., different modulation depths or varying AM rates. Irrespective of the linguistic aspect of the sentences. I think this is a major flaw in the design of the experiment. I am not questioning the results of the experiment (I think these are very valid) but rather I am questioning the implications of the results and interpretation. I am not convinced that speech/linguistic aspects are "special". There was a preferential activation of the left frontal activation in figure 1 for the delta frequency and when you do the subtraction this is the area becomes significant. I think the authors need to tone down the interpretation. The contrast is really speech versus 2 Hz AM (not necessarily "non-linguistic", this is a bit of stretch).

RESPONSE: Following the criticism of Reviewer #1 we have added a second Experiment to the Manuscript in which we included a Rotated Speech condition (based on Blesser, 1972). We think that the consistent findings across experiments we observed across the two experiments has strengthened our claim that speech-specific processes are mainly reflected in the delta entrainment and not in theta entrainment. In this second experiment, the use of the same task for the Speech and the Rotated speech also ensured that attentional resources were equivalent across both conditions.

We have carefully chosen to use the term "speech-specific" processes and not "linguistic" processes. This terminology is more conservative but still conveys the main take-home message of the paper.

Other issues:

Top plots of figures 1 and 2. What are these lines?

RESPONSE: The plots represent the Coherence spectra (corrected by the coherence spectra in the resting state condition) in some representative MEG sensors (see which ones in the Figure Captions). On the y axis, there is the Speech-MEG coherence level, on the x axis the different frequency bins we analysed.

Why are there no differences in the lateralization indices' in figure 3 (i.e., for the speech). Which ROIs are these based on?

RESPONSE: Since the lateralization indices were not very informative for the present Manuscript, we have excluded them from the present version of the Manuscript

What was the rationale for doing a correlation between speech and "non-linguistic" in figure 3?

RESPONSE: Initially, we aimed to show that the sensitivity to both speech and non-speech stimuli was similar within participant in auditory regions. This would have supported studies using non-speech stimuli in dyslexia. However, this piece of information conveys a misleading message and has been excluded in the present version of the Manuscript.

Page 8: "First, we identified the frequencies that showed significant coherence" This is very vague. How was this conducted, which frequencies, how did this change per subject, per frequency, which ROIs (sensor versus source space)?

RESPONSE: We now clarify our analysis pipeline adding more details compared to the previous submission. For the identification of the frequency of interest we reported the procedure we employed at the sensor-level (page 9):

"We used nonparametric permutation test (maximum statistic permutations, Nichols and Holmes, 2002) to identified frequency bins that showed significant coherence values at the sensor level ( $p < 0.05$ ). To do so, coherence values for each frequency bin were contrasted with resting state coherence values at the same frequency (coherence between the Env of the corresponding auditory signal and the resting state data). The sampling distribution of the maximal difference of coherence values (maximum taken across all sensors) was evaluated using the exhaustive permutation test. Frequencies for which the non-permuted maximal difference exceeded the 95th percentile of this permutation distribution were defined as frequencies of interest, and the corresponding supra-threshold sensors were identified for this frequency band. Significant frequencies were grouped in frequency bands of interest for each condition. These frequency bands were selected to compute coherence analysis in the source space."

Page 11: "at the sensor level..." Figure 2 is not at sensor level.

RESPONSE: We report now the sensor-level topoplots and statistics

Page 11: "positive correlations..." which ROIs?

RESPONSE: This passage was deleted in the present Manuscript

Reviewer: 2

Comments to the Author

The authors present results from an MEG studies on the functional role(s) of delta and theta band neural entrainment to the envelope of speech input in the service of speech processing. To this end, the authors compared the synchronization of brain dynamics and auditory inputs for sentence materials and amplitude-modulated noise (at 2 and 7 Hz) as indexed by coherence between the two signals. The key result of the study, however, stems from the analysis of coherence at the source level. Here, the authors found that brain-speech synchronization in the delta but not the theta band involved regions in the inferior frontal region, which they interpret as suggesting that delta band entrainment 'reflects high-level processing involving language related regions'.

The question of how neural oscillations in different frequency bands are utilized in the service of complex processes such as speech processing and comprehension is a timely and important topic. As such I'm generally excited to see studies trying to disentangle and characterize the contributions of neural dynamics to speech processing.

Main concerns:

Unfortunately, I'm unsure how the results of the present study will further our understanding of the contributions of delta and theta oscillations to speech processing. This is due to two main reasons: First, I see a general disconnect between the research question asked (i.e., targeting the functional dissociation of delta and theta band "entrainment" to the speech envelope in support of speech processing) and the chosen research design and analysis. I fail to see how exactly the comparison of neural responses to linguistic vs. non-linguistic (amplitude-modulated noise) would help clarify the functional role of delta vs. theta 'entrainment' to speech processing. The neural tracking of auditory input (whether it may be speech or non-linguistic input such as amplitude-modulated noise) has been shown by many previous studies; the present study may be the first to contrast neural responses to these two input types in the same participants but I'm not sure why this should be a crucial advantage over previous studies. Moreover, the analysis itself is fairly basic, so overall I don't think that the results of the present studies provide any truly novel evidence in this respect.

RESPONSE: We would like to thank Reviewer #2 for the important concern about our study. In the present version of the Manuscript we tried to emphasize the differential response of delta and theta band entrainment across different auditory conditions and in two separate Experiments. We report in both Experiments consistent findings that we think contribute to the discussion about the processes supporting cortical entrainment. At the beginning of the Introduction we clarify the debate:

"it is still under debate whether cortical entrainment to speech underlies pure auditory perceptual processing or if it additionally reflects higher-order processes involved in actively parsing speech information. "

We thus changed a bit the focus of the paper (previously centred on linguistic vs. non linguistic processing) to better frame our MEG findings. Overall, we show robust evidence that theta coherence is not sensitive to the speech-specific properties of the stimuli (see Discussion section), whereas delta coherence is stronger for Speech in a larger set of brain regions beyond the perceptual sensory cortices. This result about the different response patterns of theta and delta during speech-brain coupling has never been reported in the literature and nicely integrates with available reports trying to highlight the relevant role of delta oscillations in speech perception.

The main conclusion that the authors draw, however, is based on the results of the analysis of coherence at the source level. It is the interpretation of these results that are my second main concern regarding this study. The analysis of coherence at the source level feels very post hoc given the non-conclusive results at the sensor level.

RESPONSE: In the present version of the Manuscript we also report the sensor-level statistical analysis at the sensor-level that was not included in the previous version of the paper. We show that in both Experiments differential coherence effects emerge for delta (but not the theta) band. Sensor-level analyses however could not provide evidence concerning the source origin of our effects. The source-level analyses thus are useful to show in which brain regions speech-brain delta coherence is enhanced for Speech. Importantly, it is also possible to depict the similarity between the brain maps of theta entrainment that are highly similar independently of the auditory stimulus presented.

Moreover, the interpretation of the involvement of neural entrainment of delta-band oscillations as supporting 'high-level processing' given the involvement of inferior frontal regions is a severe case of drawing reverse inferences. Based on the research design and analysis, there is no way to say for certain what the involvement of regions in the inferior frontal cortices could reflect. To relate them to higher-level (language) processing is to base the interpretation of those previous results that simply appear to fit the story best. To me, the very fact that the strongest claim is based on the most ambiguous evidence in this study shows once again the disconnect between the research question the authors wish to answer and the methods they are choosing to do so.

RESPONSE: We thank Reviewer #2 for noting this point. We have significantly toned down the “linguistic” interpretation of our coherence data. Now we leave the interpretation of our delta findings more open (more focused on high-level processing). Importantly, we stress the need for more studies focused on delta entrainment to better constrain the set of processes they reflect. Noteworthy, we assume that if speech entrainment reflects passive auditory synchronization to the acoustic regularities of speech, the coherence brain maps should be confined to the auditory regions. This was true for the theta entrainment, but not for the delta. We think that this piece of data deserves attention from the community working on speech processing.

Further comments:

Apart from the main concerns addressed above, I thought the manuscript was generally lacking clarity especially with respect to the methods description:

RESPONSE: In the present version of the Manuscript we have added more details and specifics about the methods we used in the two Experiments. Below we detail the major changes.

Methods / Results:

- p. 5: why did trials in the linguistic but not the non-linguistic block start with a 1 sec auditory tone?

RESPONSE: In Experiment 1, participants were asked to provide a response to the comprehension questions at the end of each sentence they heard. The initial auditory tone was employed to signal that a new sentence was starting.

In the AM white-noise blocks, participants were watching a silent movie, while hearing all the white noise segments. They were presented once after the other without pauses since there was no need for signalling the beginning of the next stimulus after a task (no task here).

In Experiment 2 (by using the Rotated Speech condition) we mitigated the potential confounds that might be associated with the differences between the speech and non-speech blocks. Stimuli in this experiment had similar timing and the same task.

- p. 6: 'sentence presentation order pseudo-randomized' -> with respect to which constraints

RESPONSE: In Experiment 1, the order of the sentences was randomized and described many different scenarios to avoid priming effects. However, we made sure that the content of two consecutive sentences was not too similar to avoid discourse priming effects.

- p. 6: how did the author motivate to use 2 and 7 Hz amplitude modulation?

RESPONSE: The selection of these two frequency bands was based on previous pilot data on natural speech-brain coherence in which we observed these two coherence peaks. Indeed this was observed for Speech in Experiment 1 and for both Speech and Rotated Speech in Experiment 2.

We now included this information in the Methods section.

- p. 8: given that the terminology on the literature is not always consistent, the authors should make more explicit which measure of coherence they are using

RESPONSE: We now provide these details in the Method section the functions we used for computing the Coherence values.

- p.8: there is no description on the preprocessing of the audio signals to derive the envelopes

RESPONSE: We now describe how the envelopes of the auditory stimuli were computed by applying the Hilbert transform to the auditory signals

p.9: I found the description of the permutation test to be rather confusing. This was mostly due going back and forth between contrasts between conditions and the conditions themselves. It seemed like the authors were interested in the significance of specific contrasts between conditions, yet the permutation distribution was essentially built from the permutation of individual conditions.

RESPONSE: As indicated previously in the Response to Reviewer #1, we used the method described by Nichols & Holmes (2002) and explain the details of the statistics in the Methods section (page 9).

- p. 10: the authors should motivate why they chose to investigate of lateralization of coherence values; in the current version of the manuscript this is not clear at all...

RESPONSE: This analysis has been omitted from the revised manuscript since it since it did not contribute to the overall claims of the present study.

Discussion:

- the discussion provides a very in depth description of previous evidence on the possible functional role of neural entrainment in the delta band to speech processing but its relation to the present results purely speculative.

RESPONSE: We agree with Reviewer #2 criticism. In the Discussion, we now soften these claims. Now, we provide a more cautious account of the present findings making it clear that future studies should be designed to test the neurocognitive components that contribute to delta speech entrainment.

2nd Editorial Decision

04 December 2017

Dear Dr. Molinaro,

Your resubmitted manuscript was reviewed by external reviewers as well as by the Section Editor, Dr. Ali Mazaheri, and ourselves. Both the reviewers are now satisfied with the manuscript in its present form. However, reviewer 2 had some minor concerns which should be easily addressable. There appears some contradictory information in the methods section and figure 1 about whether the experiment had 2 or 3 conditions. In addition, the behavioral results from the comprehension task should be included in the results section.

The reviewer does suggest that the sensor data not be included. Here we respectfully disagree with the reviewer and would prefer that you keep these data in. The topographies here are extremely informative and provide a sanity check to the source data.

Please also attend to the following issues:

1. Please ensure that you provide a text and a figure file for the Graphical Abstract (as detailed in the instructions below).
2. You will need to update or remove any reference to a manuscript not yet accepted for publication
3. Your reference list needs to be checked for missing information

If you are able to respond fully to the points raised, we would be pleased to receive a revision of your paper within 10 days.

Thank you for submitting your work to EJN.

Kind regards,

John Foxe & Paul Bolam  
co-Editors in Chief, EJN

Reviews:

Reviewer: 1 (Andrew Dimitrijevic, Sunnybrook Health Sciences Centre, Canada)

Comments to the Author

All concerns have been adequately addressed

Reviewer: 3 (Anne Keitel, University of Glasgow, UK)  
Comments to the Author

In this study, Molinaro & Lizarazu investigate delta- and theta-entrainment processes to speech and different control conditions (AM-modulated noise and rotated speech) in MEG data. The addressed question (dissociation between delta and theta entrainment) is timely and important, fits with the scope of EJN, and the used paradigm is appropriate. The manuscript is also well written, although there are quite a few grammatical errors and typos.

I have only have a few minor points that could improve the manuscript.

Minor (in no particular order):

- 1) In the methods section, Exp 1, it is stated that participants did two conditions/blocks and it sounds like the AM-modulated white noise is concurrently modulated at 2 and 7 Hz within one stimulus (similar to speech). In the results (Figure 1), however, the authors talk about three conditions, and 2- and 7-Hz are analysed separately. Here, it looks like 2- and 7-Hz modulations were applied to two different stimulus streams. Could this be clarified?
- 2) Personally, I do not see the necessity to report sensor- and source-level analyses (I would only report source-level), but it does not harm the paper to include it.
- 3) The authors did not report results from the comprehension tasks. Especially in experiment 2, it would be good to see a comparison between intelligible (normal) and unintelligible (spectrally rotated) speech. But also in experiment 1 it would be nice to see how participants performed and whether all performed equally well. Could these behavioural results be included?
- 4) 'subject' should be 'participant' throughout the manuscript.

Authors' Response

13 December 2017

Following the Editor and the Reviewers' suggestions, we have corrected the minor typos and added the behavioral results. In line with the Editor's view, we agree in maintaining the sensor-level data in the main Manuscript to show the internal validity of the present MEG findings.