

We thank the reviewer for the constructive feedback to improve the paper. We incorporated reviewers' comments to our manuscript the best we could.

I personally like this work very much. It describes a well designed study about encoding specific sounds and pauses made in speech in order to obtain text documents that might be then classified with much less computing power than raw audio recordings.

The authors put much effort into testing various kinds of encoding, finally obtaining results that are comparable with BERT-S, which is impressive. Also the application of the results is in an area of study that is very important, namely cognitive disorder detection (though possibly could be also applied elsewhere, like for detecting specific kinds of emotional states).

What I don't exactly like in the paper is the layout of its chapters. For instance, the comparison with BERT-S could well be placed into the section about experiments, rather than the discussion. Also I would love to see a separate section about related work, not necessarily as a part of introduction. However, I admit that it is not a very major obstacle in reading the text, and thus I leave the decision about changing the layout to the authors.

We thank the reviewer for their positive comments on the paper. Per reviewer's suggestion, we moved the results under the result section with the section titled "Literature Comparison" for the paper to be more coherent.

In terms of changing the layout, considering other reviewers' comments, to stay with the current layout. We would consider using the suggested layout in our future works. We added more text throughout the introduction section and other sections for more clarity.

Minor remarks: line 139 at the end, "constrative" should probably be replaced with "contrastive", line 324 at the end, "which this combination", it looks like something is missing here.

We thank the reviewer for pointing out these problems. Both have been resolved.