PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Intelligibility of filtered speech and its relation to electrophysiological markers of supra-threshold hearing deficits

Markus GARRETT[1]; Viacheslav VASILKOV[2]; Sarah VERHULST[2]

[1] Medizinische Physik and Cluster of Excellence Hearing4all, Oldenburg University, Germany

[2] Hearing Technology @ WAVES, Dept of Information Technology, Ghent University, Belgium

ABSTRACT

The origins and perceptual consequences of supra-threshold hearing deficits are not well understood. Sensory and neural deficits are likely contributors, underlining the need for an objective measure which quantifies the neural contributors and complements audiometric measures of sensory loss. The envelope-following response (EFR) is sensitive to neural fiber loss in animal models and has been proposed as a marker. However, its diagnostic sensitivity in humans and relationship to speech intelligibility are still unclear. Here we argue that the relationship between electrophysiological measures of temporal coding ability and the performance in speech-in-noise tasks crucially depends on the underlying stimulus parameters. We discuss ways in which experimental protocols can be adjusted to improve the predictive relationship between electrophysiological measures of temporal coding fidelity and behavioral measures of speech intelligibility and test these predictions in a group of young normal-hearing, elderly normal-hearing and elderly hearing-impaired participants as groups reflecting different degrees of neural or sensory hearing deficits.

Keywords: envelope following response, supra-threshold hearing deficits, speech intelligibility, filtered speech

1. INTRODUCTION

An important part of the characterization of a person's hearing status is a quantification of their ability to understand speech in the presence of background noise. Even though speech audiometry in noise gives an accurate representation of one's ability to understand speech, it does not provide information about the underlying deficits at play. Additional threshold measures such as the audiometric threshold detection and otoacoustic emission measures are routinely used to quantify deficits linked to outer-hair-cell loss. However, as speech is usually presented at supra-threshold hearing levels, these threshold metrics are not informative about a person's supra-threshold temporal coding abilities (1), which are crucial to extract different features from the speech input to ultimately make sense of it.

Animal models have shown that electrophysiological metrics such as the envelope following response (EFR) and the auditory brain stem response (ABR) are able to capture facets of temporal coding deficits. The EFR (e.g. 2) measures the phase-locked neural response of the brain to the stimulus envelope and shows, depending on the stimulus material, characteristic peaks coinciding with the modulation frequency and its harmonics in the frequency-domain response of an EEG recording using a standard electrode configuration (e.g. vertex vs. ear-lobe reference). The ABR is a transient-evoked response (3,4) elicited by a click or tone burst. However, the translation of the EFR/ABR findings to humans has proven difficult due to species-dependent differences and large inter-individual differences in human cohorts. The available literature has so far shown mixed results regarding the applicability of these electrophysiological measures for the objective quantification of supra-threshold hearing deficits in humans (5).

Especially the EFR, as a measure of temporal coding ability in the early auditory pathway, should

[1] markus.garrett@uol.de
[2] viacheslav.vasilkov@ugent.be; sarah.verhulst@ugent.be

present a strong relation to the ability to understand speech when presented in difficult listening environments (6). Nevertheless, many studies that have investigated the relationship between the EFR or ABR with measures of speech intelligibility did not show any consistent relationships (e.g. 7–9). It was therefore often concluded that these EEG measure offer limited sensitivity to supra-threshold hearing deficits (see (10) for a review on the topic) even though there are examples where empirical findings were consistent with theoretical considerations (e.g. 1,11,12).

Here we argue that the relationship between electrophysiological measures of temporal coding ability and the performance in speech-in-noise tasks crucially depends on the stimulus characteristics. We discuss ways in which experimental protocols can be adapted to improve the relationship between electrophysiological measures of temporal coding fidelity and behavioral measures of speech intelligibility.

2. RESULTS & DISCUSSION

The ability to make sense of speech, crucially relies on the extraction of speech features such as the fast fluctuations in the fine structure (TFS) or the much lower and globally varying envelope of the speech signal (ENV). As fine structure information are predominantly extracted from low frequency regions (< 1 kHz, (13)) and envelope coding takes place at higher frequency regions (14), the dominant features used for speech extraction thereby seem to strongly depend on the available frequency content of the speech signal.

As temporal coding abilities degrade with advanced age and/or exposure to noise, the relative contributions of different features of speech facilitating speech intelligibility might therefore also vary between young normal-hearing, older normal-hearing and clinical populations with sensorineural hearing loss. In this context it is also important to consider that hearing loss classically affects the basal part of the cochlear, predominantly processing the high frequency content of the speech.

Most stimulus paradigms for speech intelligibility assessment use broadband speech material to represent natural speech encountered in everyday life. On the other hand, the corresponding stimulus material used to evoke the EFR mostly consist of either a pure tone carrier at frequencies around 4 kHz or a broadband white noise carrier. These carriers are afterwards amplitude-modulated using modulation frequencies in the 80-120 Hz range and presented to the subject over earphones. The individual strength of the EFR peak is an objective measure of temporal coding and is typically compared to a speech intelligibility metric (e.g., speech reception threshold, SRT) to understand the contribution of temporal coding deficits to speech intelligibility. The interpretation of the EFR vs. SRT comparison constitutes the following problems: i) temporal coding ability of a narrow portion of the cochlea (in the case of a pure-tone EFR) is compared with the intelligibility of a broadband speech stimulus, ii) the envelope coding performance of the whole cochlea (in the case of a broadband EFR) is compared to the intelligibility of broadband speech which might depend on many different features of speech (e.g., TFS) and not only the envelope information as measured by the EFR. These comparisons therefore either lead to a mismatch of frequency content and/or mismatch of underlying mechanism at play. We therefore argue that these comparisons likely compare 'apples and pears' and that it is not surprising that measures do not show any kind of consistent relationship.

To overcome these mismatches and to strengthen the relationship between these metrics, we suggest the use of a frequency-specific carrier signal within a frequency region know to use envelope information (e.g. 4 kHz) and compare these to the speech intelligibility performance of a filtered speech signal (e.g. > 1.6 kHz) (15). We argue that the underlying mechanisms of the metrics computed in the afore mentioned way will be comparable and therefore yield much stronger relationships compared to classically employed stimulus paradigms.

To explore the variables in the relationship between EFRs and speech intelligibility, we conducted a study which investigated the match and mismatch between EFR and speech intelligibility measures in a cohort of young normal-hearing, elderly normal-hearing and elderly hearing-impaired participants. Our results confirm that the relationship between both metrics grows stronger as the frequency content of both stimuli match the regions of suspected envelope coding, showing the highest correlations for high-pass filtered speech with a 120 Hz amplitude-modulated 4 kHz pure tone carrier.

3. CONCLUSIONS

Our results support the underlying assumptions that the mixed results found in literature between objective supra-threshold electrophysiological measures of temporal coding and the behavioral measures of speech-in-noise tasks (e.g. 7–9) are likely due to the mismatch of underlying frequency content and/or coding mechanisms at play. We provide evidence that the manipulation of both speech and AM-stimuli towards better matching the frequency content and coding mechanism will yield significant relationships between the metrics in agreement with theoretical considerations. A detailed account of the results of our work will be presented in our talk at the ICA 2019.

ACKNOWLEDGEMENTS

The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the German Research Foundation (DFG Cluster of Excellence EXC 1077/1 "Hearing4all": MG, SV; PP1608: VV) and the European Research Council (ERC) under the Horizon 2020 Research and Innovation Programme (grant agreement No 678120 RobSpear, VV, SV).”

REFERENCES

1. Bharadwaj HM, Masud S, Mehraei G, Verhulst S, Shinn-Cunningham BG. Individual differences reveal correlates of hidden hearing deficits. J Neurosci. 2015 Feb 4;35(5):2161–72.
2. Dolphin WF, Mountain DC. The envelope following response: scalp potentials elicited in the Mongolian gerbil using sinusoidally AM acoustic signals. Hear Res. 1992 Feb;58(1):70–8.
3. Sohmer H, Feinmesser M. Cochlear action potentials recorded from the external ear in man. Ann Otol Rhinol Laryngol. 1967 Jun;76(2):427–35.
4. Jewett DL, Williston JS. Auditory-evoked far fields averaged from the scalp of humans. Brain. 1971;94(4):681–96.
5. Bramhall N, Beach EF, Epp B, Le Prell CG, Lopez-Poveda EA, Plack CJ, et al. The search for noise-induced cochlear synaptopathy in humans: Mission impossible? Hear Res. 2019 Mar 9;377:88–103.
6. Bharadwaj HM, Verhulst S, Shaheen L, Liberman MC, Shinn-Cunningham BG. Cochlear neuropathy and the coding of supra-threshold sound. Front Syst Neurosci. 2014;8:26.
7. Guest H, Munro KJ, Prendergast G, Millman RE, Plack CJ. Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. Hear Res. 2018;364:142–51.
8. Valderrama JT, Beach EF, Yeend I, Sharma M, Van Dun B, Dillon H. Effects of lifetime noise exposure on the middle-age human auditory brainstem response, tinnitus and speech-in-noise intelligibility. Hear Res. 2018 Aug;365:36–48.
9. Johannesen PT, Buzo BC, Lopez-Poveda EA. Evidence for age-related cochlear synaptopathy in humans unconnected to speech-in-noise intelligibility deficits. Hear Res. 2019 Mar 15;374:35–48.
10. Le Prell CG. Effects of noise exposure on auditory brainstem response and speech-in-noise tasks: a review of the literature. Int J Audiol. 2019 Feb;58(sup1):S3–32.
11. Mehraei G, Hickox AE, Bharadwaj HM, Goldberg H, Verhulst S, Liberman MC, et al. Auditory Brainstem Response Latency in Noise as a Marker of Cochlear Synaptopathy. J Neurosci. 2016 Mar

30;36(13):3755–64.
12. Liberman MC, Kujawa SG. Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. Hear Res. 2017 Jun;349:138–47.
13. Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BCJ. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. Proc Natl Acad Sci USA. 2006 Dec 5;103(49):18866–9.
14. Henry KS, Kale S, Heinz MG. Distorted Tonotopic Coding of Temporal Envelope and Fine Structure with Noise-Induced Hearing Loss. J Neurosci. 2016 Feb 17;36(7):2227–37.
15. Verhulst S, Warzybok A. Contributions of Low- and High-Frequency Sensorineural Hearing Deficits to Speech Intelligibility in Noise. bioRxiv. 2019 Feb 25;358127.