Editorial: Neural Tracking: Closing the Gap Between Neurophysiology and Translational Medicine

Giovanni M. Di Liberto [1,2*], Jens Hjortkjær [3] and Nima Mesgarani [4]

[1] School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland, [2] ADAPT Centre, d-real, Trinity College Institute for Neuroscience, Dublin, Ireland, [3] Hearing Systems Group, Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Ireland, [4] Electrical Engineering Department, Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, United States

Editorial on the Research Topic

Neural Tracking: Closing the Gap Between Neurophysiology and Translational Medicine

Perception involves making sense of the world around us by processing a continuous flow of multi-modal sensory information. In doing so, the human brain produces electrical activity that can be measured in a variety of scenarios and tasks to shed light on the neural basis of continuous perception. This work has shown that electrical brain activity synchronizes to particular properties of sensory inputs, a phenomenon referred to as *neural tracking* (Obleser and Kayser, 2019). Recent work demonstrated that both invasive and non-invasive electrophysiology recordings can robustly detect neural tracking (Lalor et al., 2006; Ding and Simon, 2012; Gross et al., 2013; Zion Golumbic et al., 2013), offering objective measurements to study perception in increasingly more complex tasks involving continuous real-life stimuli, such as speech and music.

The case of auditory perception is particularly remarkable. The discovery that neural signals reliably track the amplitude envelope of continuous sounds (*envelope tracking*) (Lalor et al., 2009) has led to new research directions. *In primis*, envelope tracking measurements have enabled a range of studies on auditory attention in realistic multi-talker scenarios (e.g., see COCOHA project, H2020.2.1.1.4. ID = 644732), showing that signals recorded with invasive electrocorticography (ECoG) as well as non-invasive electro- and magneto-encephalography (EEG/MEG) track attended and unattended sounds in a different manner (Ding and Simon, 2012; Zion Golumbic et al., 2013; O'Sullivan et al., 2014, 2019). This pioneering discovery led to an entire new direction for brain-computer interface research, with perspectives for novel devices such as brain-controlled hearing-aids (Eyndhoven et al., 2017; O'Sullivan et al., 2017; Ceolini et al., 2020). A parallel line of work demonstrated that multiple properties of the same stimulus are tracked simultaneously (O'Sullivan et al., 2016; Di Liberto et al., 2021a; Gillis et al., 2021). In the context of speech listening, cortical signals were shown to track progressively higher-level properties of the speech signal, from acoustical features (Lalor and Foxe, 2010; Ding et al., 2014) to linguistic units (Di Liberto et al., 2015, 2018b; Brodbeck et al., 2018; Lesenfants et al., 2019), prosody (Myers et al., 2019; Teoh et al., 2019), and semantic content (Broderick et al., 2018, 2021; Weissbart et al., 2020). As such, neural tracking measurements can offer a rich view into the hierarchical encoding of speech by providing us with distinct objective indices referring to different processing stages.

The outstanding advances in this domain have pushed scientists to explore the potentialities of studying neural tracking in translational research (Jessen et al., 2019; Dial et al., 2021; Geirnaert et al., 2021; Palana et al., 2022). Indeed, the unprecedented opportunity to assess the speech processing hierarchy as a whole (as well as for other stimuli, such as music) in a single experimental session is a very compelling reason that encourages the exploration of translational research directions. Furthermore, the possibility of using ecologically-valid tasks, such as movie or cartoon watching, opens the door to cohorts that would be difficult to assess otherwise (Di Liberto et al., 2018a; Jessen et al., 2019; Attaheri et al., 2022). Nevertheless, the feasibility for the translational applications of neural tracking metrics remains to be determined, as the theoretical and methodological challenges are yet to be uncovered.

In this special topic issue we have gathered contributions from scientists working in diverse disciplines who have common interests in the neural tracking phenomenon from various research domains. The current issue includes studies on speech (Alickovic, Ng, et al.) and music perception (Hausfeld et al.), selective attention (Huet et al.), and aging in healthy individuals (Mesik et al.). It also covers methodological considerations for translational research (Crosse et al.) and for measuring responses to different speech features (Bachmann et al.), as well as theoretical and practical perspectives on hearing-impairment (Alickovic, Ng, et al.), hearing-aid technology (Alickovic, Lunner, et al.), and schizophrenia (Meyer et al.). Bringing together work from a variety of research domains demonstrates the extensive width of applications for neural tracking research, while hopefully helping to build a new community of interdisciplinary research. We were very fortunate to enlist a varied and talented group of authors to contribute such a wide range of topics. Thirty-five authors contributed to the eight papers included, with a mixture of six original research articles, one review, and one hypothesis and theory. Taken together these papers present an overview of research on neural tracking from a range of perspectives, indicating a promising research framework that can greatly contribute to translational research questions, both from theoretical and applied perspectives.

As typical for new lines of work, the literature offers a diverse set of approaches and views regarding neural tracking. One issue is the apparent inconsistency in the terminology used by different research groups, leading to some confusion with terms such as neural entrainment, synchronization, and tracking. Obleser and Kayser (2019) have recently put forward an important distinction between the concepts of neural entrainment in the *narrow* and *broad sense*. In their view, neural entrainment in the *narrow sense* refers to the concept of "synchronization," whereby endogenous self-sustained neural oscillators adjust their temporal dynamics ("rhythms") to that of the sensory input (Schroeder and Lakatos, 2009). While this definition is specific to a particular neural mechanism, we use the term *neural tracking* to refer to neural entrainment in the broad sense, where the neurophysiology measurements likely reflect a combination of multiple phenomena. In fact, it is challenging (to say the least) to make any claim on the specific neural mechanisms generating such non-invasively recorded signals. Nevertheless, a somewhat agnostic view on such underlying neural mechanisms would not prevent us from making valuable theoretical and practical use of such measurements. Work using such measures has already contributed to our understanding of speech (Mesgarani et al., 2014; Di Liberto et al., 2015, 2021a; Ding et al., 2015; Brodbeck et al., 2018; Broderick et al., 2018) and music perception (Tal et al., 2017; Di Liberto et al., 2020, 2021b; Marion et al., 2021; Zuk et al., 2021), selective attention (O'Sullivan et al., 2014; Decruy et al., 2020; Fuglsang et al., 2020), multisensory integration (Crosse et al., 2016; Sullivan et al., 2021), and even abstract cognitive processes such as arithmetic (Kulasingham et al., 2021). The work in this Research Topic attempts to portray a wide set of findings while using consistent terminology.

This Research Topic is a first attempt to put together methodological, theoretical, and applied work with the common aim of projecting the study of neural tracking toward translational research. Recent reviews have discussed the neural tracking phenomenon (Obleser and Kayser, 2019; Hamilton and Huth, 2020), including specific applied research scenarios involving atypical cohorts (Palana et al., 2022). From that work, it is clear that we have only scraped the surface of a line of work with great potential, and that much more is yet to come. Neural tracking has a minimal presence in translational research at present. One challenge is that the literature portrays a complex research landscape, including many methodologies to evaluate and report the results. As for more established methodologies (e.g., ERPs), the definition of appropriate standardisations and the development of appropriate tools to more rapidly and effortlessly measure neural tracking are crucial to effectively adopting these methodologies to translational research.

One paper in this article collection contributed to this debate, presenting a set of precise guidelines on how to measure, evaluate, and report neural tracking in applied research by using one particular approach (the multivariate temporal response function—mTRF) (Crosse et al.). Others have emerged from discussions at conferences (e.g., ARO) and workshops (e.g., the Telluride Neuromorphic Engineering workshop), with special sessions revolving around neural tracking. The more specific *Cognition and Natural Sensory Processing* (CNSP) initiative, which has an educational focus, aims at bringing together researchers interested in studying and using neural tracking measurements, offering a workshop and online resources, such as standardized datasets and analysis code. Other fields such as genomics have demonstrated that resource sharing has the potential to propel research fields extensively beyond state of the art (Kaye et al., 2009; Captur et al., 2016). The benefits will be greater if resource sharing is taken as a new opportunity to answer the many open questions in our fields, rather than a separate independent niche for computational scientists. Taking inspiration from other fields could greatly help us in tackling the potential challenges that come with new opportunities.

AUTHOR CONTRIBUTIONS

GD wrote the first draft of the manuscript. JH and NM revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Attaheri, A., Choisdealbha, Á. N., Di Liberto, G. M., Rocha, S., Brusini, P., Mead, N., et al. (2022). Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. *Neuroimage* 247, 118698. doi: 10.1016/j.neuroimage.2021.118698

Brodbeck, C., Hong, L. E., and Simon, J. Z. (2018). Rapid transformation from auditory to linguistic representations of continuous speech. *Curr. Biol.* 28, 3976–3983.e3975. doi: 10.1016/j.cub.2018.10.042

Broderick, M. P., Anderson, A. J., Di Liberto, G. M., Crosse, M. J., and Lalor, E. C. (2018). Electrophysiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech. *Curr. Biol.* 28, 803–809. doi: 10.1016/j.cub.2018.01.080

Broderick, M. P., Di Liberto, G. M., Anderson, A. J., Rofes, A., and Lalor, E. C. (2021). Dissociable electrophysiological measures of natural language processing reveal differences in speech comprehension strategy in healthy ageing. *Sci. Rep.* 11, 4963. doi: 10.1038/s41598-021-84597-9

Captur, G., Stables, R. H., Kehoe, D., Deanfield, J., and Moon, J. C. (2016). Why democratise bioinformatics? *,MJ Innov.* 2, 166. doi: 10.1136/bmjinnov-2016-000129

Ceolini, E., Hjortkjær, J., Wong, D. D., O'Sullivan, J., Raghavan, V. S., Herrero, J., et al. (2020). Brain-informed speech separation (BISS) for enhancement of target speaker in multitalker speech perception. *NeuroImage.* 223, 117282. doi: 10.1016/j.neuroimage.2020.117282

Crosse, M. J., Di Liberto, G. M., and Lalor, E. C. (2016). Eye can hear clearly now: Inverse effectiveness in natural audiovisual speech processing relies on long-term crossmodal temporal integration. *J. Neurosci.* 36, 9888–9895. doi: 10.1523/JNEUROSCI.1396-16.2016

Decruy, L., Vanthornhout, J., and Francart, T. (2020). Hearing impairment is associated with enhanced neural tracking of the speech envelope. *Hear. Res.* 393, 107961–107961. doi: 10.1016/j.heares.2020.107961

Di Liberto, G. M., Marion, G., and Shamma, S. A. (2021a). The music of silence: Part II: Music listening induces imagery responses. *J. Neurosci.* 41, 7449. doi: 10.1523/JNEUROSCI.0184-21.2021

Di Liberto, G. M., Nie, J., Yeaton, J., Khalighinejad, B., Shamma, S. A., and Mesgarani, N. (2021b). Neural representation of linguistic feature hierarchy reflects second-language proficiency. *Neuroimage* 227, 117586. doi: 10.1016/j.neuroimage.2020.117586

Di Liberto, G. M., O'Sullivan, J. A., and Lalor, E. C. (2015). Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Curr. Biol.* 25, 2457–2465. doi: 10.1016/j.cub.2015.08.030

Di Liberto, G. M., Pelofi, C., Bianco, R., Patel, P., Mehta, A. D., Herrero, J. L., et al. (2020). Cortical encoding of melodic expectations in human temporal cortex. *eLife* 9, e51784. doi: 10.7554/eLife.51784

Di Liberto, G. M., Peter, V., Kalashnikova, M., Goswami, U., Burnham, D., and Lalor, E. C. (2018a). Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia. *NeuroImage,* 17–29, 70–79. doi: 10.1016/j.neuroimage.2018.03.072

Di Liberto, G. M., Wong, D., Melnik, G. A., and de Cheveigné, A. (2018b). Cortical responses to natural speech reflect probabilistic phonotactics. *bioRxiv.* doi: 10.1101/359828

Dial, H. R., Gnanateja, G., Tessmer, R. S., Gorno-Tempini, M. L., Chandrasekaran, B., and Henry, M. L. (2021). Cortical tracking of the speech envelope in logopenic variant primary progressive aphasia. *Front. Human Neurosci.* 14, 597694. doi: 10.3389/fnhum.2020.597694

Ding, N., Chatterjee, M., and Simon, J. Z. (2014). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *NeuroImage* 88, 41–46. doi: 10.1016/j.neuroimage.2013.10.054

Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nat. Neurosci.* 19, 158–164. doi: 10.1038/nn.4186

Ding, N., and Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc. Natl. Acad. Sci. USA* 109, 11854–11859. doi: 10.1073/pnas.1205381109

Eyndhoven, S. V., Francart, T., and Bertrand, A. (2017). EEG-informed attended speaker extraction from recorded speech mixtures with application in neuro-steered hearing prostheses. *IEEE Trans. Biomed. Eng.* 64, 1045–1056. doi: 10.1109/TBME.2016.2587382

Fuglsang, S. A., Märcher-Rørsted, J., Dau, T., and Hjortkjær, J. (2020). Effects of sensorineural hearing loss on cortical synchronization to competing speech during selective attention. *J. Neurosci.* 40. 2562. doi: 10.1523/JNEUROSCI.1936-19.2020

Geirnaert, S., Vandecappelle, S., Alickovic, E., de Cheveigné, A., Lalor, E., Meyer, B. T., et al. (2021). Neuro-steered hearing devices: decoding auditory attention from the brain. *Cogn. Sci.* doi: 10.48550/arXiv.2008.04569

Gillis, M., Vanthornhout, J., Simon, J. Z., Francart, T., and Brodbeck, C. (2021). Neural markers of speech comprehension: measuring EEG tracking of linguistic speech representations controlling the speech acoustics. *J. Neurosci.* 41, 10316. doi: 10.1523/JNEUROSCI.0812-21.2021

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., et al. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol.* 11, e1001752–e1001752. doi: 10.1371/journal.pbio.1001752

Hamilton, L. S., and Huth, A. G. (2020). The revolution will not be controlled: natural stimuli in speech neuroscience. *Lang. Cogn. Neurosci.* 35, 573–582. doi: 10.1080/23273798.2018.1499946

Jessen, S., Fiedler, L., Münte, T. F., and Obleser, J. (2019). Quantifying the individual auditory and visual brain response in 7-month-old infants watching a brief cartoon movie. *NeuroImage* 202, 116060. doi: 10.1016/j.neuroimage.2019.116060

Kaye, J., Heeney, C., Hawkins, N., de Vries, J., and Boddington, P. (2009). Data sharing in genomics—re-shaping scientific practice. *Nat. Rev. Genetics* 10, 331–335. doi: 10.1038/nrg2573

Kulasingham, J. P., Joshi, N. H., Rezaeizadeh, M., and Simon, J. (2021). Cortical processing of arithmetic and simple sentences in an auditory attention task. *J. Neurosci.* 41, 8023. doi: 10.1523/JNEUROSCI.0269-21.2021

Lalor, E. C., and Foxe, J. J. (2010). Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *Eur. J. Neurosci.* 31, 189–193. doi: 10.1111/j.1460-9568.2009.07055.x

Lalor, E. C., Pearlmutter, B. A., Reilly, R. B., McDarby, G., and Foxe, J. J. (2006). The VESPA: a method for the rapid estimation of a visual evoked potential. *NeuroImage* 32, 1549–1561. doi: 10.1016/j.neuroimage.2006.05.054

Lalor, E. C., Power, A. J., Reilly, R. B., and Foxe, J. J. (2009). Resolving precise temporal processing properties of the auditory system using continuous stimuli. *J. Neurophysiol.* 102, 349–359. doi: 10.1152/jn.90896.2008

Lesenfants, D., Vanthornhout, J., Verschueren, E., and Francart, T. (2019). Data-driven spatial filtering for improved measurement of cortical tracking of multiple representations of speech. *J. Neural Eng.* 16, 066017. doi: 10.1101/551218

Marion, G., Di Liberto, G. M., and Shamma, S. A. (2021). The music of silence, Art I: responses to musical imagery accurately encode melodic expectations and acoustics. *J. Neurosci.* 41, 7435–7448. doi: 10.1523/JNEUROSCI.0183-21.2021

Mesgarani, N., Cheung, C., Johnson, K., and Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science* 343, 1006–1010. doi: 10.1126/science.1245994

Myers, B. R., Lense, M. D., and Gordon, R. L. (2019). Pushing the envelope: developments in neural entrainment to speech and the

biological underpinnings of prosody perception brain. *Science* 9, 70. doi: 10.3390/brainsci9030070

Obleser, J. and Kayser, C. (2019). *Trends in Cognitive Sciences*. Vol. 23 913-926. Amsterdam: Elsevier Ltd.

O'Sullivan, A. E., Crosse, M. J., Di Liberto, G. M., and Lalor, E. C. (2016). Visual cortical entrainment to motion and categorical speech features during silent lipreading frontiers in human. *Neuroscience* 10, 679–679. doi: 10.3389/fnhum.2016.00679

O'Sullivan, J., Chen, Z., Herrero, J., McKhann, G. M., Sheth, S. A., Mehta, A. D., et al. (2017). Neural decoding of attentional selection in multi-speaker environments without access to clean sources. *J. Neural Eng.* 14, 056001. doi: 10.1088/1741-2552/aa7ab4

O'Sullivan, J., Herrero, J., Smith, E., Schevon, C., McKhann, G. M., Sheth, S. A., et al. (2019). Hierarchical encoding of attended auditory objects in multi-talker speech perception. *Neuron.* 104, 1195–1209.e1193. doi: 10.1016/j.neuron.2019.09.007

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., et al. (2014). Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cereb. Cortex* 25, 1697–1706. doi: 10.1093/cercor/bht355

Palana, J., Schwartz, S., and Tager-Flusberg, H. (2022). Evaluating the use of cortical entrainment to measure atypical speech processing: a systematic review. *Neurosci. Biobehav. Rev.* 133, 104506. doi: 10.1016/j.neubiorev.2021.12.029

Schroeder, C. E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 32, 9–18. doi: 10.1016/j.tins.2008.09.012

Sullivan, A. E., Crosse, M. J., Liberto, G. M. D., de Cheveigné A., and Lalor, E. C. (2021). Neurophysiological indices of audiovisual speech processing reveal a hierarchy of multisensory integration effects. *J. Neurosci.* 41, 4991. doi: 10.1523/JNEUROSCI.0906-20. 2021

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., et al. (2017). Neural entrainment to the beat: The "missing-pulse" phenomenon. *J. Neurosci.* 37, 6331–6341. doi: 10.1523/JNEUROSCI.2500-16. 2017

Teoh, E. S., Cappelloni, M. S., and Lalor, E. C. (2019). Prosodic pitch processing is represented in delta-band EEG and is dissociable from the cortical tracking of other acoustic and phonetic features. *Eur. J. Neurosci.* 50, 3831–3842. doi: 10.1111/ejn.14510

Weissbart, H., Kandylaki, K. D., and Reichenbach, T. (2020). Cortical tracking of surprisal during continuous speech comprehension. *J. Cogn. Neurosci.* 32, 155–166. doi: 10.1162/jocn_a_01467

Zion Golumbic, E. M. Z., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "Cocktail Party". *Neuron* 77, 980–991. doi: 10.1016/j.neuron.2012.12.037

Zuk, N. J., Murphy, J. W., Reilly, R. B., and Lalor, E. C. (2021). Envelope reconstruction of speech and music highlights stronger tracking of speech at low frequencies. *PLOS Comput. Biol.* 17, e1009358. doi: 10.1371/journal.pcbi.100 9358