**Synchronizing internal and external information: A commentary on Meyer, Sun & Martin (2020)**

Anastasia Klimovich-Gray[1] & Nicola Molinaro[1,2]

[1] *BCBL, Basque Center on Cognition, Brain and Language, Donostia/San Sebastian, Spain*

[2] *Ikerbasque, Basque Foundation for Science, Bilbao, Spain*

Correspondence to:

Anastasia Klimovich-Gray
BCBL, Basque center on Cognition, Brain and Language Paseo Mikeletegi, 69
20009, Donostia/San Sebastian
Spain
email: a.klimovich@bcbl.eu

Meyer and colleagues provide a compelling argument for the importance of endogenous cortical rhythms that track abstract linguistic representations, a process they refer to as *intrinsic synchronicity*. These endogenous rhythms encode probabilistic predictions about upcoming linguistic representations - at different grain sizes ranging from phonemes to syntactic phrase structures - that cannot be derived solely from the perceptual signal. Critically, while most theoretical accounts of speech perception focus on cortical entrainment to exogenous (external) acoustic cues in speech, Meyer and colleagues focus on oscillatory tracking of endogenous information which helps to integrate speech cues from the perceptual stream with higher-level representations. This addresses an important gap in the current theoretical literature, underscoring the role played by higher-level endogenous oscillatory activity in the "structured analysis" of external stimulation. This could have important implications for research on language acquisition, as nicely discussed by the authors in the final paragraph.

In their definition of *intrinsic synchronicity, the* authors indicate that this phenomenon is distinct "from the modulation of entrainment by domain-general or linguistic top-down processes" (page 4). This statement highlights two important points that, in our view, require further clarification in light of the current theoretical and experimental literature. First, is there indeed a fundamental distinction between *intrinsic synchronicity* and other processes for top-down modulation of oscillatory entrainment described elsewhere? The authors suggest that *intrinsic synchronicity* is a distinct phenomenon but, in our opinion, high-level cognitive processes must be in tune with ongoing perception to optimize perceptual sampling via hierarchical predictive inference, as previously proposed in the literature. The second, and related, question is whether *intrinsic synchronicity* is specific to the language domain? Can the same neurocognitive mechanisms be used for information processing across multiple domains? While the authors do not explicitly state that this phenomenon is language-specific, we think that their approach would benefit from being integrated with a more domain-general perspective. We believe that addressing the above questions will allow for a more parsimonious account, where cortical tracking of abstract linguistic representations works interdependently with ongoing perceptual analysis (adhering to the principles of predictive coding) to enable contextual comprehension.

The authors define *intrinsic synchronicity* as "strictly endogenous", "neither entrainment nor its top-down modulation" (page 4), implying that this mechanism is separate and not modulated by ongoing perceptual analysis. Consistent with this, Meyer and colleagues propose that *intrinsic synchronicity* and "accompanying categorical abstraction (e.g., inference of phonemic features) should be observed in sensory association cortices" (page 6), while entrainment proper should be present in the primary sensory areas. This would make it distinct from previously proposed domain-general accounts of top-down modulatory oscillatory activity such as "neural tracking" (Obleser & Kayser, 2019) or "dynamic information selection by entrainment" (Lakatos, Gross & Thut, 2019), or "active/proactive sensing" of speech components (Morillon et al., 2015; Rimmele et al., 2018), which a) work by tracking specific perceptual properties of the stimulus that signal the onset of informative segments (e.g. the speech envelope) and b) ultimately bias and optimize perceptual sampling within these segments by aligning them with periods of maximal neuronal activity. The distinction, however, becomes blurred when we consider that (a) none of these latter accounts claim that tracking or entrainment would occur exclusively for exogenous information (see direct quotes in Notes); and (b) the current experimental literature suggests that oscillatory tracking of endogenous representations (i.e. what can be considered *intrinsic synchronicity*) is indeed used to optimise perceptual sampling in the primary auditory areas. Further, this latter point is theoretically expected from predictive processing/coding accounts. Hierarchical top-down predictions need to be continuously updated by perceptual evidence (via Prediction Error PE propagation - Friston et al., 2005). These predictions must be temporally synced with incoming perceptual signals (Rimmele et al., 2018) in order to increase the gain of more informative segments (e.g. via attentional mechanisms and PE precision weighing - Friston et al., 2009; DenOuden, Kok & Lange, 2012) and optimise learning.

The two cases Meyer and colleagues present as good examples of *intrinsic synchronicity* in language, arguably suggest that endogenous representations are, in fact, used as top-down guides to perceptual sampling. In both examples, endogenous linguistic expectations were found to bias ambiguous perceptual signal sampling. In the Kösem et al. (2016) study, top-down bias induced alignment between the amplitude of high frequency oscillations (beta and gamma) and the onsets of preferentially perceived words when the auditory stream was perceptually ambiguous (bistable auditory percept). Kösem and colleagues argued that higher frequency effects reflected top-down facilitation of perceptual segmentation by boosting (or maybe biasing) processing of consciously “preferred” acoustic segments and improving feedforward information transfer. Meyer et al. (2016) showed that delta phase predicted phrasal boundary grouping in syntactically ambiguous sentences. Moreover, when acoustic (prosody) cues were inconsistent with a more expected phrasal boundary, coherence between speech pitch tracks and the delta band was reduced. This can also be thought of as top-down modulation of sampling via delta entrainment to speech prosody, i.e. assigning less perceptual weight to what would otherwise be perceptually meaningful segments (the authors call this reduction of “auditory processing efficacy”).

Recent work by Donhauser and Baillet (2019) showed that increases in phoneme Entropy (uncertainty about a contextually predicted phoneme) predicted gain increases in ongoing theta-band oscillations in the primary auditory cortex (AC) during sentence processing, while phoneme Surprisal (quantifying required prediction-updating) modulated delta responses in associative AC (as well as pAC). Theta-by-Entropy modulation can be considered a form of *intrinsic synchronicity* since phoneme Entropy is derived by listeners from their internal generative model of incoming speech. However, differing from Meyer and colleagues proposed definition of *intrinsic synchronicity*, it modulates processing in primary auditory areas, not just the neighbouring associative areas, suggesting a much more direct link with perceptual analysis. Donhauser and colleagues argued that perceptual sampling related to the theta band is optimised by internal predictions that increase gain of perceptual signals during periods of uncertainty, while delta modulation magnifies prediction-errors required for prediction updating. Work by Bourguignon et al. (2019) further showed that auditory areas track both overt and internally “synthesised” speech (endogenous representations of coarse-grained auditory features) at <1Hz frequencies, when the audio signal is absent but has to be derived from visual lip reading. This top-down internal feature synthesis, the authors argue, is a mechanism for facilitating speech parsing and audio-visual cue integration via top-down predictions propagated from the visual to the auditory domain.

To summarise our argument so far, there is consistent evidence that cortical oscillations track endogenous contextual predictions during language processing, but none of this evidence suggests that this tracking is not modulating or modulated by ongoing perceptual analysis. In fact, it implies the opposite: that endogenous information modulates perceptual experience to optimise incoming signal sampling and facilitate contextual comprehension. By spelling out the link between *intrinsic synchronicity* and perceptual sampling optimization, this phenomenon can be better integrated with other proposed prediction-oriented mechanisms (e.g. Morillon et al., 2015; Rimmele et al., 2018) beyond the domain of language.

This brings us to the second question – can the concept of *intrinsic synchronicity* extend beyond the language domain and, if so, are same/similar neuro-cognitive mechanisms at play? Examples of perceptual processing optimisation via endogenous predictive information tracking exist in both vision and audition (Barczaka et al., 2018; Morillion & Baillet, 2017; Breska & Deouell, 2017). For instance, endogenous temporal predictions from the sensorimotor cortex have been shown to top-down modulate auditory processing. Specifically, Morillion & Baillet (2017) have shown that when trained to expect task-relevant auditory information at a specific rate (beeps, at 1.5 Hz delta), the sensorimotor cortex predictively entrained to this frequency even if subsequent auditory information arrived at a different rate (3 Hz). This continuous predictive entrainment was endogenous and memory-driven since the unique auditory cue at delta range was not present during the task. Critically, sensorimotor cortex delta oscillations modulated beta (18-24 Hz) amplitude, which in turn top-down modulated (directional connectivity analysis) activity in the auditory cortex, suggesting that temporal predictions enhanced ongoing perceptual analysis (see Park et al., 2015 for a similar argument in regard to speech processing). Entrainment to predictable but not perceptually available cues has also been shown in music, even if the overt auditory stimulus was expected but not physically present (Tal et al., 2017). Furthermore, intracortical recordings from nonhuman primates (Barczak et al., 2018) have shown delta entrainment in the primary auditory cortex (A1) to contextual (i.e. endogenous) statistical patterns present in a continuous auditory stream with no clear perceptual onset-offset markers. Such evidence suggests that oscillatory tracking (especially at lower frequencies) of predictable information, derived through inference rather than overt perceptual cues, guides perceptual optimisation and is not specific to human language processing but is a more basic, domain-general cognitive mechanism (c.f Rimmele et al., 2018; Obleser & Kayser, 2019; Kayser et al., 2015).

In this commentary, we outlined how the concept of *intrinsic synchronicity* can be more directly and explicitly integrated with the process of perceptual sampling optimisation and providing examples from current literature. The merits of this integration are twofold. It provides a holistic view of how endogenous linguistic representations and exogenous perceptual cues continuously interact to optimise contextual comprehension. Second, it aligns the concept of *intrinsic synchronicity* with broader domain-general predictive processing/coding frameworks, inviting exploration of the mechanisms for cross-domain information integration.

**Acknowledgements**

AK was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 798971. NM was supported by the Spanish Ministry of Science, Innovation and Universities (grant RTI2018-096311-B-I00), the Agencia Estatal de Investigación (AEI), the Fondo Europeo de Desarrollo Regional (FEDER). The authors acknowledge financial support from the "Severo Ochoa" Programme for Center/Unit of Excellence in R&D (SEV-2015-490) and by the Basque Government through the BERC 2018-2021 program.

**Notes**

Lakatos, Gross & Thut, 2019 "Entrainment by inputs conveying internally generated information streams like connected memories or 'trains of thought' is also conceivable. While we acknowledge that the existence of this form of internal entrainment is highly speculative, there is some evidence that it does occur." (p.R892)

Obleser & Kayser, 2019 "However results from clever stimulus manipulations show that entrainment can also be shaped by high-level linguistic processes and may not require such low-level acoustic regularities, a case in which the neural signal may be oscillatory but the stimulus not." (p.922)

Rimmele et al., 2018 "Temporal predictions (periodic, aperiodic) operate upon endogenous constraints by predictively aligning neuronal excitability in time to facilitate the processing of anticipated events. Top-down influences correspond to an anticipatory phase reset (originating from higher-level processes; not directly driven by low-level stimulus features) of ongoing oscillations in those neuronal subpopulations involved in the processing of the expected event (e.g., beat or syllable)". (p.875)

**References**

Barczak, A., O'Connell, M. N., McGinnis, T., Ross, D., Mowery, T., Falchier, A., & Lakatos, P. (2018). Top-down, contextual entrainment of neuronal oscillations in the auditory thalamocortical circuit. *Proceedings of the National Academy of Sciences*, *115*(32), E7605-E7614.

Bourguignon, M., Baart, M., Kapnoula, E. C., & Molinaro, N. (2019). Lip-reading enables the brain to synthesize auditory features of unknown silent speech. *Journal of Neuroscience*.

Den Ouden, H. E., Kok, P., & De Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*, 548.

Donhauser, P. W., & Baillet, S. (2019). Two Distinct Neural Timescales for Predictive Speech Processing. *Neuron*.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815-836.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211-1221.

Kayser, S. J., Ince, R. A., Gross, J., & Kayser, C. (2015). Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *Journal of Neuroscience*, *35*(44), 14691-14701.

Kösem, A., Basirat, A., Azizi, L., & Van Wassenhove, V. (2016). High-frequency neural activity predicts word parsing in ambiguous speech streams. *Journal of Neurophysiology*, *116*(6), 2497-2512.

Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment. *Current Biology*, *29*(18), R890-R905.

Meyer, L., Henry, M. J., Gaston, P., Schmuck, N., & Friederici, A. D. (2016). Linguistic bias modulates interpretation of speech via neural delta-band oscillations. *Cerebral Cortex*, *27*(9), 4293-4302.

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences*, *114*(42), E8913-E8921.

Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences*, *1337*(1), 26.

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*.

Park, H., Ince, R. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology*, *25*(12), 1649-1653.

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences*, *22*(10), 870-882.

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., & Golumbic, E. Z. (2017). Neural entrainment to the beat: the “missing-pulse” phenomenon. *Journal of Neuroscience*, *37*(26), 6331-6341.