ARTICLE IN PRESS
PLREV:1150

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Physics of Life Reviews ••• (••••) •••–•••

PHYSICS of LIFE reviews

www.elsevier.com/locate/plrev

Comment

Computational models and neural bases of statistical learning in music and language
Comment on “Creativity, information, and consciousness: The information dynamics of thinking” by Wiggins

Tatsuya Daikoku

Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany

Received 22 August 2019; accepted 2 September 2019

Communicated by L. Perlovsky

The review by Wiggins [1] offers a comprehensive theory of the brain’s mechanism in the framework of statistical learning (SL). The theory, referred to as Information Dynamics of Thinking (IDyOT) implements memory representation, consolidation, and categorization under entropy (uncertainty) encoding, aiming to be information-efficient. The work also sheds new light on the links between SL and creativity, and opens up interdisciplinary perspectives on computational and neurophysiological underpinnings of SL. Here, I discuss how the IDyOT theory links with neural function in SL.

1. Semantic and episodic representations

The brain is innately equipped with SL machinery that encodes transitional probabilities (TPs) and entropy (i.e., uncertainty) in sequential information such as language and music. Using the acquired TP distribution, the brain predicts the high-TP events, and chunks them as semantic units [2,3]. Recent literature identifies two interdependent processes as a hallmark of SL: [1] abstraction or chunking of statistically coherent events from the sequence to generate a semantic memory and [2] integration of those units to generate episodic representation [4]. This seems to agree with the IDyOT concepts in which the chunking of several events and the re-representation as a single unit trigger the shift of hierarchy from lower (lexical and semantic) to higher (syntactic and episodic) levels. A recent study suggests that lower-order statistical and semantic units have general characteristics shared among persons, whereas higher-order statistical and episodic representation provides individualities unique to each person [5]. Thus, it is believed that the abstraction into a semantic unit provides general memory, while integration of several units provides individual, creative, and episodic representation.

The IDyOT TP- and uncertainty-based consolidation may be partially explained by the connectivity between hippocampal and neocortical functions. The hippocampus plays an important role in uncertainty encoding and episodic memory [6,7], whereas the neocortex contributes to online prediction and semantic memory [8]. The two distinct systems are not independent but are instead interdependent via the hippocampus-neocortex gateway [9]: short-term

DOI of original article: https://doi.org/10.1016/j.plrev.2018.05.001.
E-mail address: tdaikoku-tky@umin.org.

https://doi.org/10.1016/j.plrev.2019.09.001

Please cite this article in press as: Daikoku T. Computational models and neural bases of statistical learning in music and language. Phys Life Rev (2019), https://doi.org/10.1016/j.plrev.2019.09.001

memory representation in the hippocampus is consolidated, re-organized, and re-represented in the long-term space of the neocortex [10]. Thus, the consolidation process promotes quantitative and qualitative changes of memory representations [11,12], giving rise to novel and creative information [13,14]. From information-theoretical perspectives, it is hypothesized that consolidation and creativity systems are associated with certain and uncertain formations, respectively. The consolidation process allows information to be compressed as a single unit, resulting in the improvement of information efficiency (i.e., low uncertainty). The creativity, in contrast, gives rise to novel information not previously experienced, and therefore uncertain information. That is, precise evaluation of uncertainty in the hippocampus allows the brain to generate creative and uncertain information as well as consolidate certain memory. The neural findings also indicate that the hippocampus has a link to the default-mode network [15], which is important for creative and episodic future thinking [16,17]. Hence, the hippocampus-neocortex loop is likely to hold the key into IDyOT TP- and uncertainty-based consolidation, and thus the emerging creativity.

2. Priming, coherence, and speech entrainment

In IDyOT, the sequential information is chunked into phonemes, morphemes, then semantic words, and an episodic phrase in a stepwise fashion. According to neurophysiological studies, low-frequency oscillations in the speech motor cortex track the envelope of the speech signal [18], while high-frequency oscillations dominate in tracking the fine structure of speech [19] and bottom-up processing [20]. Furthermore, in each frequency band, SL and the chunking function control the coupling and synchronisation between phase-locked neural oscillations and speech frequencies (e.g., ∼20 Hz: phoneme, ∼4 Hz: syllable, ∼1 Hz: word and phrase) [21,22]. This has also been suggested by simulation and modelling based on cochlea function [23,24]. Another possible neurophysiological marker of IDyOT SL may be event-related potentials (ERP). Higher-order cortical representations predict plausible lower-order representations of a high-TP event in a top-down manner, and inhibit ERPs for the predictable event based on predictive coding. In the end, the SL effect manifests as the difference amplitudes of ERPs between predictable and unpredictable events [25]. Statistical abstraction to acquire semantic chunks and a statistical comparison between the acquired chunks to generate episodic syntax are both reflected in different ERPs [26]. Particularly, the N400 is known as a semantic component, and represents the SL effect [27,28]. Furthermore, the N400 and phase-locked coherence (∼1 Hz: word frequency) as semantic chunking effects can be detected in parallel [29]. Hence, the oscillations and ERPs may be candidates for unveiling IDyOT processes.

3. Categorization and non-adjacent prediction

A number of literatures indicate that SL is a domain-general and species-general learning principles, occurring for visual as well as auditory information including language and music, and in both primates and non-primates such as monkeys [30], songbirds [31,32], and rats [33]. The current SL theories, however, may not be sufficient to account for all levels of the music and linguistic knowledge acquired in the adult human brain that is likely to rely on domain-specific principles in each language and music [34,35]. The categorization [36], nonadjacent TPs [37], and the order of TPs [25] are considered the key insights into the advanced SL models in a form closer to that used for natural language and music. For example, humans learn TPs of word categories such as nouns and verbs [36]: when a verb of "drink" occurs, the brain predicts lots of subsequent words that can be drunken. The brain generalizes adjacent and non-adjacent statistical rules of that grammar and further applies those rules to novel vocabulary [38]. Hence, the brain does not have to code all received information, contributing to information efficiency, which is an important concept of the IDyOT's categorizations. I'm expected that the IDyOT will implement SL systems in a form closer to natural language and music learning

4. Future prospects: what is optimal for creativity, but not efficiency?

The brain is generally motivated to optimize prediction and minimize uncertainty for information efficiency [39]. This uncertainty resolution results in rewards. On the other hand, a certain degree of uncertainty may be necessary to maintain curiosity towards the information and learning motivation because the uncertainty allows the brain to anticipate further rewards through the resolution. That is, for sustainable "*curiosity*" and "*motivation*", humans may seek a slightly suboptimal solution if it is afforded at a significantly low uncertainty. Recent studies indicate that the

Please cite this article in press as: Daikoku T. Computational models and neural bases of statistical learning in music and language. Phys Life Rev (2019), https://doi.org/10.1016/j.plrev.2019.09.001

fluctuations of uncertainty may contribute to aesthetic appreciation of art and music [40,41], and encourage humans to create and learn new regularities [42]. The balance of uncertainty, that is, the "black box" optimization algorithm that maintains curiosity and facilitates not efficiency but creativity, will be a future prospect in SL studies. To understand it, it is important to investigate how and when the intrinsic curiosity is originated through SL. The IDyOT may imply a black box optimization by uncertainty-driven consolidation and data compression [43]. I hypothesize that this sudden improvement of information efficiency provides intrinsic rewards (i.e., the so-called "wow-effect"), triggers strong motivation to resolve highly uncertain information, and results in creative thinking possibly due to an extra capacity in memory and storage. In conclusion, the IDyOT theory proposed novel perspectives on the link between SL and creativity. I would also like to emphasize that this theory links with neural findings on SL. This suggests that the IDyOT processes can be verified by neurophysiological experiments such as phase coherence and ERP. An optimization algorithm not for efficiency but for creativity is a further key question for SL of language and music. To understand how/when creativity originates through SL processes, future study is needed to interdisciplinarily verify SL from computational and neuronal perspectives.

Acknowledgements

The present study was supported by Suntory Foundation. The funders had no role in decision to publish and preparation of the manuscript.

References

[1] Wiggins GA. Creativity, information, and consciousness: the information dynamics of thinking. Phys Life Rev 2019. https://doi.org/10.1016/j.plrev.2018.05.001. In press.
[2] Saffran J, Aslin R, Newport E. Statistical learning by 8-month-old infants. Science 1996;274:1926–8.
[3] Hasson U. The neurobiology of uncertainty: implications for statistical learning. Philos Trans R Soc B 2017;372:20160048.
[4] Thiessen ED, Kronstein AT, Hufnagle DG. The extraction and integration framework: a two-process account of statistical learning. Psychol Bull 2013;139:792–814. https://doi.org/10.1037/a0030801.
[5] Daikoku T. Musical creativity and depth of implicit knowledge: spectral and temporal individualities in improvisation. Front Comput Neurosci 2018;12:1–27. https://doi.org/10.3389/fncom.2018.00089.
[6] Moscovitch M, Cabeza R, Winocur G, Episodic Nadel L. Memory and beyond: the hippocampus and neocortex in transformation. Annu Rev Psychol 2016;67:105–34. https://doi.org/10.1146/annurev-psych-113011-143733.
[7] Harrison LM, Duggins A, Friston KJ. Encoding uncertainty in the hippocampus. Neural Netw 2006;19:535–46. https://doi.org/10.1016/j.neunet.2005.11.002.
[8] Frost R, Armstrong BC, Siegelman N, Christiansen MH. Domain generality versus modality specificity: the paradox of statistical learning. Trends Cogn Sci 2015;19:117–25. https://doi.org/10.1016/j.tics.2014.12.010.
[9] Durrant SJ, Cairney SA, Lewis PA. Overnight consolidation aids the transfer of statistical knowledge from the medial temporal lobe to the striatum. Cereb Cortex 2013;23:2467–78. https://doi.org/10.1093/cercor/bhs244.
[10] Hasselmo ME. Neuromodulation: acetylcholine and memory consolidation. Trends Cogn Sci 1999;3:351–9. https://doi.org/10.1016/S1364-6613(99)01365-0.
[11] Wagner U, Gais S, Haider H, Verleger R, Born J. Sleep inspires insight. Nature 2004;427:352. https://doi.org/10.1038/nature02223.
[12] Ellenbogen JM, Hu PT, Payne JD, Titone D, Walker MP. Human relational memory requires time and sleep. Proc Natl Acad Sci 2007;104:7723–8. https://doi.org/10.1073/pnas.0700094104.
[13] Cai DJ, Mednick SA, Harrison EM, Kanady JC, Mednick SC. REM, not incubation, improves creativity by priming associative networks. Proc Natl Acad Sci 2009;106:10130–4. https://doi.org/10.1073/pnas.0900271106.
[14] McClelland J, McNaughton B, O'Reilly R. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol Rev 1995;102:419–57.
[15] Ward AM, Schultz AP. The parahippocampal gyrus links the default-mode cortical network with the medial temporal lobe memory system. Hum Brain Mapp 2014;35:1061–73. https://doi.org/10.1002/hbm.22234.
[16] Beaty RE, Kenett YN, Christensen AP, Rosenberg MD, Benedek M, Chen Q, et al. Robust prediction of individual creative ability from brain functional connectivity. Proc Natl Acad Sci 2018:201713532. https://doi.org/10.1073/pnas.1713532115.
[17] Østby Y, Walhovd KB, Tamnes CK, Grydeland H, Westlye LT, Fjell AM. Mental time travel and default-mode network functional connectivity in the developing brain. Proc Natl Acad Sci 2012:16800–4. https://doi.org/10.1073/pnas.1210627109.
[18] Park H, Ince RAA, Schyns PG, Thut G, Gross J. Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. Curr Biol 2015;25:1649–53. https://doi.org/10.1016/j.cub.2015.04.049.
[19] Giraud AL, Poeppel D. Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci 2012;15:511–7. https://doi.org/10.1038/nn.3063.
[20] Fontolan L, Morillon B, Liegeois-Chauvel C, Giraud AL. The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. Nat Commun 2014;5:1–10. https://doi.org/10.1038/ncomms5694.

[21] Assaneo MF, Ripollés P, Orpella J, Lin WM, de Diego-Balaguer R, Poeppel D. Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. Nat Neurosci 2019;22:627–32. https://doi.org/10.1038/s41593-019-0353-z.
[22] Ding N, Melloni L, Zhang H, Tian X, Poeppel D. Cortical tracking of hierarchical linguistic structures in connected speech. Nat Neurosci 2016;19:158–64. https://doi.org/10.1038/nn.4186.
[23] Ding N, Patel AD, Chen L, Butler H, Luo C, Poeppel D. Temporal modulations in speech and music. Neurosci Biobehav Rev 2017;81:181–7. https://doi.org/10.1016/j.neubiorev.2017.02.011.
[24] Leong V, Acoustic-Emergent Goswami U. Phonology in the amplitude envelope of child-directed speech. PLoS ONE 2015;10:e0144411. https://doi.org/10.1371/journal.pone.0144411.
[25] Daikoku T. Neurophysiological markers of statistical learning in music and language: hierarchy, entropy, and uncertainty. Brain Sci 2018:8. https://doi.org/10.3390/brainsci8060114.
[26] Daikoku T, Yatomi Y, Yumoto M. Statistical learning of an auditory sequence and reorganization of acquired knowledge: a time course of word segmentation and ordering. Neuropsychologia 2017:95. https://doi.org/10.1016/j.neuropsychologia.2016.12.006.
[27] François C, Chobert J, Besson M, Schön D. Music training for the development of speech segmentation. Cereb Cortex 2013;23:2038–43. https://doi.org/10.1093/cercor/bhs180.
[28] Sanders LD, Newport EL, Neville HJ. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. Nat Neurosci 2002;5:700–3. https://doi.org/10.1038/nn873.
[29] Batterink LJ, Paller KA. Online neural monitoring of statistical learning. Cortex 2017. https://doi.org/10.1016/j.cortex.2017.02.004.
[30] Saffran JR, Hauser M, Seibel R, Kapfhamer J, Tsao F, Cushman F. Grammatical pattern learning by human infants and cotton-top tamarin monkeys. Cognition 2008;107:479–500. https://doi.org/10.1016/j.cognition.2007.10.010.
[31] Lu K, Vicario DS. Statistical learning of recurring sound patterns encodes auditory objects in songbird forebrain. Proc Natl Acad Sci USA 2014;111:14553–8. https://doi.org/10.1073/pnas.1412109111.
[32] Lu K, Vicario DS. Familiar but unexpected: effects of sound context statistics on auditory responses in the songbird forebrain. J Neurosci 2017;37:12006–17. https://doi.org/10.1523/JNEUROSCI.5722-12.2017.
[33] Toro JM, Sinnett S, Soto-Faraco S. Speech segmentation by statistical learning depends on attention. Cognition 2005;97:25–34. https://doi.org/10.1016/j.cognition.2005.01.006.
[34] Jackendoff R, Lerdahl F. The capacity for music: what is it, and what's special about it? Cognition 2006;100:33–72. https://doi.org/10.1016/j.cognition.2005.11.005.
[35] Hauser MD, Chomsky N, Fitch WT. The faculty of language: what is it, who has it, and how did it evolve? Science 2002;298:1569–79. https://doi.org/10.1126/science.298.5598.1569.
[36] Jones MN, Mewhort DJK. Representing word meaning and order information in a composite holographic lexicon. Psychol Rev 2007;114:1–37. https://doi.org/10.1037/0033-295X.114.1.1.
[37] Frost RLA, Monaghan P. Simultaneous segmentation and generalisation of non-adjacent dependencies from continuous speech. Cognition 2016. https://doi.org/10.1016/j.cognition.2015.11.010.
[38] Gomez RL, Gerken LA. Articial grammar learning by one-year-olds leads to specic and abstract knowledge. Cognition 1999;70:109–36.
[39] Friston K. The free-energy principle: a unified brain theory? Nat Rev Neurosci 2010;11:127–38. https://doi.org/10.1038/nrn2787.
[40] Daikoku T. Depth and the uncertainty of statistical knowledge on musical creativity fluctuate over a composer's lifetime. Front Comput Neurosci 2019;13:27. https://doi.org/10.3389/fncom.2019.00027.
[41] Koelsch S. Brain correlates of music-evoked emotions. Nat Rev Neurosci 2014;15:170–80. https://doi.org/10.1038/nrn3666.
[42] Schmidhuber J. Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. Connect Sci 2006;18:173–87. https://doi.org/10.1080/09540090600768658.
[43] Wiggins GA. Consolidation as re-representation: revising the meaning of memory. Front Psychol 2019:1–22.