Modeling the effects of temporal context on neural responses across the cortical hierarchy

Hsiang-Yun Sherry Chien (sherry.chien@jhu.edu)
Department of Psychological & Brain Sciences, 3400 N. Charles Street, Baltimore, MD 21218 USA
Christopher J. Honey (chris.honey@jhu.edu)
Department of Psychological & Brain Sciences, 3400 N. Charles Street, Baltimore, MD 21218 USA

Abstract:

The world contains information over multiple timescales. For example, we must combine sequences of syllables to perceive a word, and sequences of words to comprehend a sentence. How does the brain process information over multiple timescales? Previous studies have demonstrated that higher-order brain regions are sensitive to temporal context on longer scales, and that they also express stable activity states over the duration of an event. We set out to model these neural phenomena using a hierarchical temporal auto-encoder (HTA). When augmented with mechanisms for a contextual reset, the HTA successfully reproduces neural phenomena. The HTA also generates a prediction regarding context construction: that low-level regions establish a new context rapidly, while higher order regions establish new context representations more gradually. We confirmed this empirical prediction by applying inter-subject pattern correlation to fMRI responses of sentences heard in different temporal contexts. Overall, we propose that a hierarchy of temporal auto-encoders is a feasible model of temporal information processing in the cortical hierarchy.

Keywords: computational modeling; hierarchical learning; sequence processing; fMRI; inter-subject correlation

Background

Information in the world is nested over multiple timescales: syllables compose words; words compose sentences; and sentences compose narratives. The human brain appears to employ a hierarchical architecture to process sequential structure across a range of scales (Hasson et al. 2015, Himberger et al. 2018). This notion of hierarchical temporal processing in the human brain is supported by two key findings:

1) *the hierarchy of temporal context*: responses in consecutive stages of cortical processing are sensitive to longer windows of prior stimulus context in higher order regions (Lerner et al. 2011).

2) *the hierarchy of neural dynamics*: neural dynamics in consecutive stages of cortical processing are slower and more stable in higher order regions (Baldassano et al. 2017, Schapiro et al. 2013).

We hypothesized that multi-scale sequence learning and perception are supported by a hierarchical computational architecture, in which each level acts as a temporal auto-encoder of its input (see also Chung et al. 2016). We formalized our hypothesis as a hierarchical temporal auto-encoder (HTA) and tested whether this model could reproduce the two key empirical findings noted above. We then compared our HTA model against an existing memory model (temporal context model, TCM) to test which computational elements are sufficient to account for hierarchical temporal processing in the human brain. Finally, we empirically tested a prediction of our HTA model regarding the timescale of context construction.

HTA and TCM

1. Our HTA model consists of stacked auto-encoder units, inspired by the TRACX2 sequence learning model (Mareschal and French 2017). Each layer of the model consists of an input unit (IN), a context unit (CNTX) and a hidden unit (HID) (Fig. 1). At each time step, the auto-encoder uses a linear-nonlinear transformation to compress CNTX+IN. The result is a lower-dimensional representation stored in the HID unit. Each level of the model then computes a local "surprise" variable, which is the absolute difference between the CNTX+IN vectors and the "reconstructed" CNTX+IN extracted from the compressed HID representation.

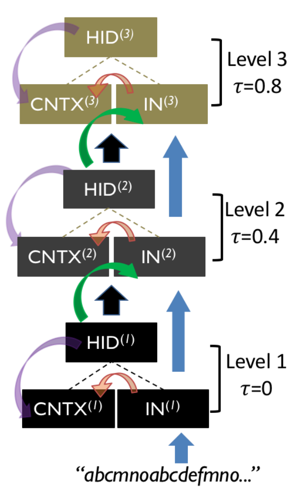


Figure 1 Hierarchical temporal auto-encoder (HTA) architecture.

HTA has two key features:

a. More local memory preserved at higher levels

Change in CNTX units are modulated by decay (governed by a *time constant*) and surprise (α, at the previous time point). At higher stages of processing, the CNTX units have longer time constants, and thus preserve more prior context.

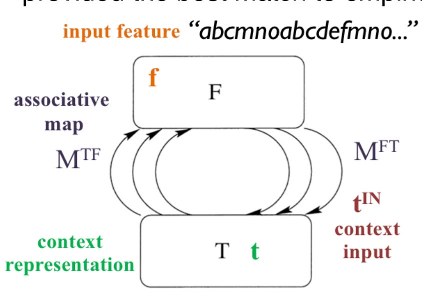
$$\Delta = \sum(output - target)^2 \qquad \alpha = \tanh(k\Delta)$$

$$CNTX_{t+1}^{i} = \left(\frac{\tau^{i}}{\alpha_t^{i}+\tau^{i}}\right) \times \boldsymbol{HID_t^{i}} + \left(\frac{\alpha_t^{i}}{\alpha_t^{i}+\tau^{i}}\right) \times \boldsymbol{IN_t^{i}} \quad \textbf{(Eq.1)}$$

b. Feedforward information modulated by surprise Information is transmitted from lower levels to higher levels of the model (i.e. $IN^{(2)}$ is a function of the units at Level 1). The nature of input to stage N+1 is modulated by *surprise* α at stage N. If α is small, the model successfully synthesized IN and CNTX, and this synthesized information is sent to higher levels.

$$IN_t^{i+1} = \left(\boldsymbol{1-\alpha_t^{i}}\right) \times \boldsymbol{HID_t^{i}} + \boldsymbol{\alpha_t^{i}} \times \boldsymbol{IN_t^{i}} \quad \textbf{(Eq.2)}$$

2. Temporal context model (TCM) TCM is a single layer model, with an input unit (f) and a context unit (t). The model learns the association between input features and context units through updating the associative map (M^{TF}, M^{FT}) via Hebbian Learning. The updated context is a combination of the previous context and a context input (t^{IN}) from input feature, modulated by parameter $\boldsymbol{\beta}$ (Fig. 2). We modeled different processing timescales by modifying $\boldsymbol{\beta}$, the context decay rate. A manual parameter search indicated that $\boldsymbol{\beta}$ = 0.9, 0.7, 0.5 for Levels 1, 2 and 3 provided the best match to empirical effects.



$$t_i^{IN} = M_i^{FT} \boldsymbol{f_i}$$

$$\boldsymbol{t_i} = \rho_i \boldsymbol{t_{i-1}} + \beta \boldsymbol{t_i^{IN}}$$

$$M_{i+1}^{FT} = M_i^{FT} + B_i \boldsymbol{t_i} \boldsymbol{f_i}$$

$$\rho \cong \sqrt{1-\beta^2}$$

Figure 2. Temporal context model (TCM) architecture and updating rules

Model implementations and comparisons

1. Hierarchy of context dependence. First, we simulated the basic phenomenon that higher order regions are more sensitive to information presented further in the past, suggesting that higher order regions integrate information over a longer timescale (Fig. 3A, Lerner et al. (2011)). To simulate this phenomenon, we trained HTA and TCM with intact structured sequence and tested them with sequences that preserved either short or long timescales of the trained structure. We then computed the similarity (correlation) of the representations of each test item in the scrambled context against its representations in the intact context, analogous to the experimental design of Lerner et al. (2011; Fig. 3B). HTA trained with structured sequence exhibited a functional temporal hierarchy: compared to the lower levels, the higher levels are more sensitive to changes of temporal context farther in the past (Fig. 3C, *left*). When the model was trained with shuffled data (no reliable temporal structure), the hierarchical context effects were reduced (Fig. 3C, *right*). We confirmed these observations with a repeated measures ANOVA, revealing large effects of of Model Level (1,2,3) and Training Type (structured vs. random), and the interaction of Level and Training Type (all $F > 40$, all $p < 0.001$). The same overall pattern was obtained when we examined a hierarchy of TCM models with different time-constants. However, we were unable to find parameters for TCM for which the higher levels of the model showed context effects as strong as those in HTA. Finally, we tested whether the context-reset effect contributed to the ability of the HTA model to learn multi-scale temporal structure. We lesioned the model so that feedforward flow of information was no longer modulated by surprise (i.e. α=0 in Eq.2). The lesioned model exhibited reduced context dependence and increased reconstruction error (Fig. 3D). Thus, surprise-modulated context reset is important for learning multi-scale temporal structure in the HTA.

2. Community structure at distinct timescales. Schapiro et al. (2013) showed that some higher-order brain regions are sensitive to temporal community structure – they come to represent "event structure" defined by non-adjacent associations in time. Thus, we tested whether HTA and TCM could generate such event representations when exposed to state-transitions with a temporal community structure (Fig. 4A). We trained both models with sequences sampled from a transition matrix with three communities, yet with equal transition probability ($p = 0.25$) from each node to its neighbors. Both HTA and TCM learned "event representations": patterns of activity were more similar within temporal communities than between temporal communities. Furthermore, this effect was stronger at higher levels of the HTA and TCM models: community structure (block structures in Fig. 4B) was stronger at consecutive stages of the model, as within-versus-between community correlation increased from Level 1 to Level 3 (HTA: 0.41, 0.44, 0.69; TCM: 0.78, 1.05, 1.31) These results are consistent with the notion of a hierarchy of timescales in brain dynamics (Baldassano et al. 2017).

3. Summary of modeling. The HTA model reproduced two empirical phenomena of temporal processing in the human brain: hierarchical context sensitivity (Fig. 3) and representation of temporal community structure (Fig. 4). TCM provided equivalent performance in extracting temporal community structure, but was less able to generate distinct representations for sequences which differ many steps earlier (e.g. XABCDE vs. YABCDE). Our current modeling work is testing whether *surprise-driven context resetting* is the key feature that enables HTA to perform well on both tasks jointly.

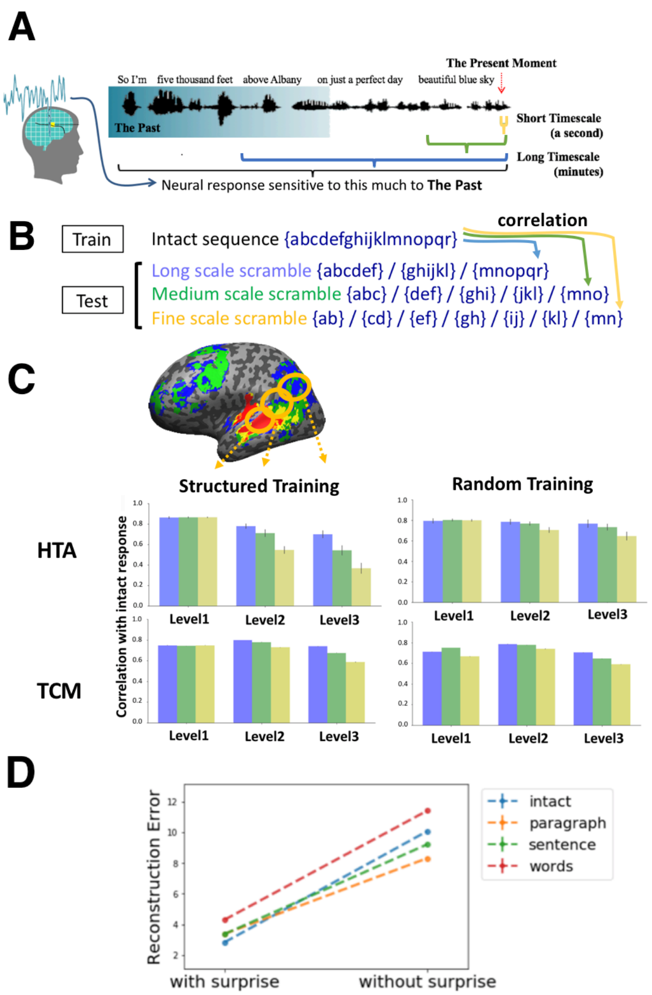


Figure 3 (A) Information from short to long timescales is integrated in hierarchical structure of the brain. Different brain regions are sensitive to different scales of temporal context. (**B**) Sequences for training and testing the models (**C**) HTA generated hierarchical context dependence under structured training. Context effects were smaller in TCM and in models trained with random sequences. (**D**) HTA without feedforward context-resetting exhibited higher reconstruction error.

Hierarchical Context Construction Effect

Model Prediction. Our HTA model predicts that, following a context reset, lower levels of the model should re-establish context more rapidly than higher levels do (Fig. 6A). We refer to this phenomenon as the *hierarchy of timescales in context construction*. To test

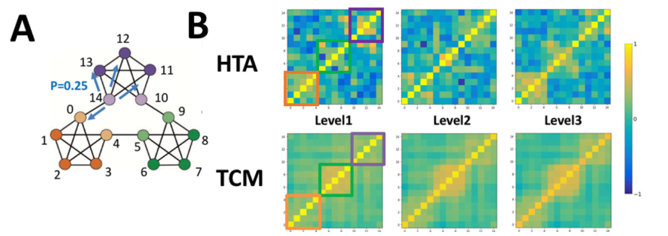


Figure 4 (**A**) Community structure with equal transition probability from one node to adjacent nodes. (**B**) Both HTA and TCM generated hierarchical community structures.

this prediction, we performed a similar analysis in fMRI data, measuring the timescale on which BOLD voxel patterns became aligned across different brains that were exposed to the same current input but had experienced a different prior context.

fMRI Analysis Methods One group of 24 subjects listened to a 9-minute auditory story in the fMRI scanner, and another group (n = 24) listened to the same story scrambled at the scale of long sentences (mean = 21.9s, s.d. = 4.3s). Whole-brain inter-subject pattern correlation (ISPC) was performed on the fMRI data to measure the neural signal alignment within each sentence (Fig. 5A). In a set of 400 ROIs (Schaefer et al. (2017)), we measured ISPC across the subjects who heard the intact and scrambled stimuli. At each moment, the correlation across subjects quantified the alignment of brain states when processing the same sentence preceded by different contexts (Fig. 5B).

Results We found that, when processing the same sentence preceded by different contexts, the neural responses became more aligned (higher ISPC) in the later part of the sentence. The low alignment at the beginning of the sentence is due to hemodynamic carry-over from the preceding sentence. However, on top of this hemodynamic effect we observe variability in the time to align, consistent with a hierarchy of context reconstruction (Fig. 6B). As predicted by the HTA model, higher-order regions such as inferior parietal lobule (IPL) and posterior superior temporal gyrus (STG) became aligned more slowly, compared with lower order regions such as primary auditory cortex (A1+) and middle STG (Fig. 6B)

Conclusion

We tested the ability of a hierarchical temporal auto-encoder (HTA) model to reproduce features of temporal processing in the human brain. The HTA successfully reproduced key empirical phenomena of hierarchical temporal integration in the human brain. Comparisons with the TCM model suggest that *surprise-modulated context reset* may enable the HTA model to better balance the trade-offs between the learning of strict ordered relationships (local sequence structure) and the extraction of slow temporal associations (temporal community structure). Finally, we confirmed a prediction of the HTA model using fMRI responses recorded from participants listening to spoken narratives: following a sharp event boundary, early brain regions re-establish context representations more rapidly than higher-order regions.

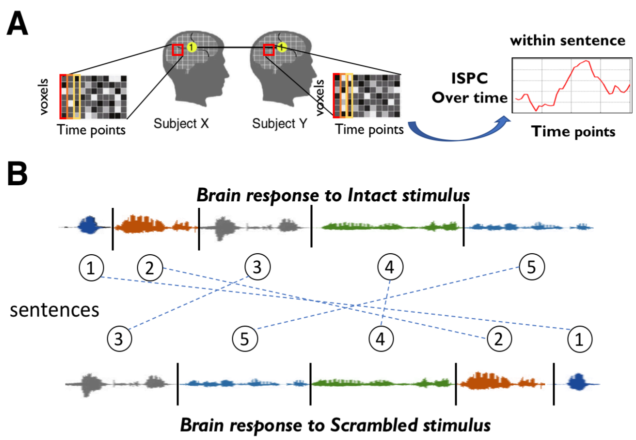


Figure 5 Inter-subject pattern correlation (ISPC) analysis **(A)** ROI-wise ISPC was calculated at each time point. **(B)** ISPC was calculated across subjects exposed to intact and scrambled stimuli. Thus, we examined the moment-by-moment neural signal alignment when hearing the same sentences in different contexts.

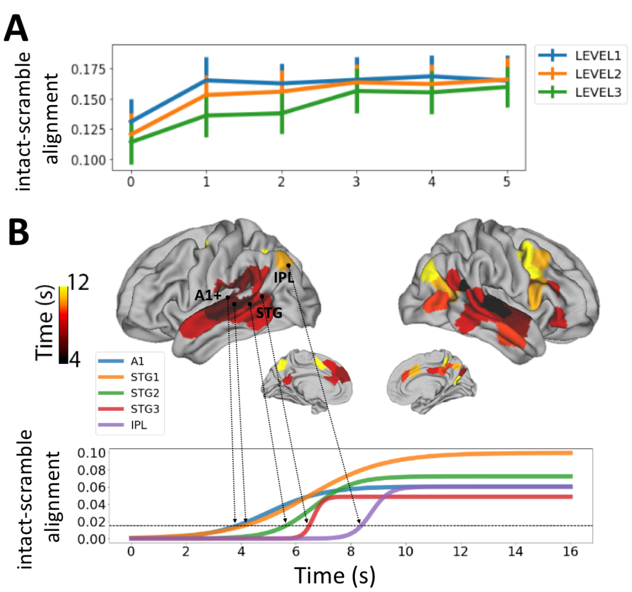


Figure 6 (A) Higher levels of the HTA model take longer to synchronize, when comparing responses to the same input preceded by different contexts. **(B)** fMRI inter-subject pattern correlation results confirm the model prediction: higher-order regions synchronize more slowly across subjects exposed to different contexts.

Acknowledgments

We acknowledge the support of the Alfred P. Sloan Foundation (Fellowship to CH). We acknowledge the use of fMRI data collected under Princeton IRB #5516.

References

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. Neuron, 95(3), 709-721.

Chung, J., Ahn, S., & Bengio, Y. (2016). Hierarchical multiscale recurrent neural networks. arXiv preprint arXiv:1609.01704.

Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. Trends in cognitive sciences, 19(6), 304-313.

Himberger, K. D., Chien, H. Y., & Honey, C. J. (2018). Principles of Temporal Processing Across the Cortical Hierarchy. Neuroscience. *In Press*

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. Journal of Mathematical Psychology, 46(3), 269-299.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. Journal of Neuroscience, 31(8), 2906-2915.

Mareschal, D., & French, R. M. (2017). TRACX2: a connectionist autoencoder using graded chunks to model infant visual statistical learning. Phil. Trans. R. Soc. B, 372(1711), 20160057.

Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. Nature neuroscience, 16(4), 486.

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., ... & Yeo, B. T. (2017). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. Cerebral Cortex, 1-20.