MSc in Computational Cognitive Neuroscience

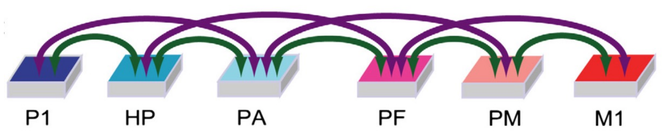
Goldsmiths
UNIVERSITY OF LONDON

Simulating intrinsically motivated vocal exploration of infants to investigate the effects of selective reward on word learning

Gabriel Dransfield – Supervisor: Max Garagnani

BACKGROUND

In the early stages of vocal development, infants produce and explore a wide variety of sounds and syllables but eventually learn to utter only spoken words heard in their environment. Currently, many explanations suggest intrinsic motivation plays a key role in driving this process.[1][2] However, at present time there is no computational model able to explain the neural mechanisms underlying the spontaneous "selection" of relevant (and "pruning" of non-relevant) word sounds. We attempt to address this gap by using an existing brain-realistic model of language areas.[3]



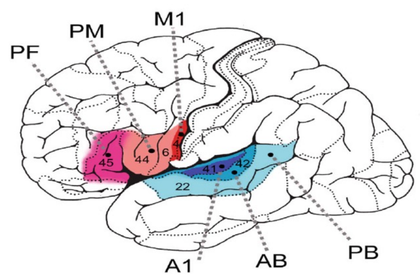
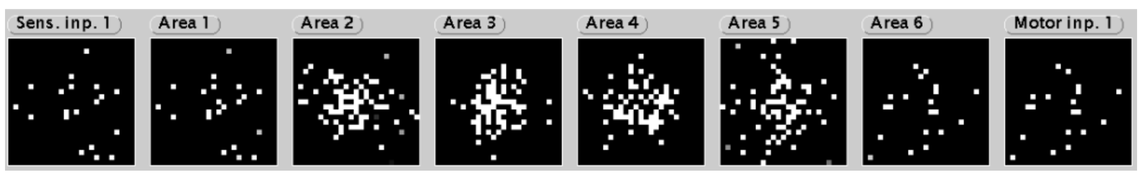


Figure on the left shows the different model areas and links. Figure on the right illustrates the analogous brain areas modelled.

RESEARCH AIM and HYPOTHESES

As a result of training (simulating "babbling"), the model [1] has been shown to exhibit the emergence of distributed "auditory-articulatory" (Cell Assembly, CA) circuits, which spontaneously ignite as a result of noise accumulation in them (see example of a CA below).



Taking such CA circuits as model-correlates of memory traces for words, we hypothesize that, if a reinforcing signal ("reward") is delivered to the network when some of its CAs spontaneously ignite (simulating a situation in which a randomly produced speech sound happens to mimic an externally perceived word), as the simulation progresses, the model should gradually develop a "bias", so that the CAs ("articulations") for rewarded "sounds" will spontaneously ignite *more often* than other, non-rewarded ones. This HP relies on a

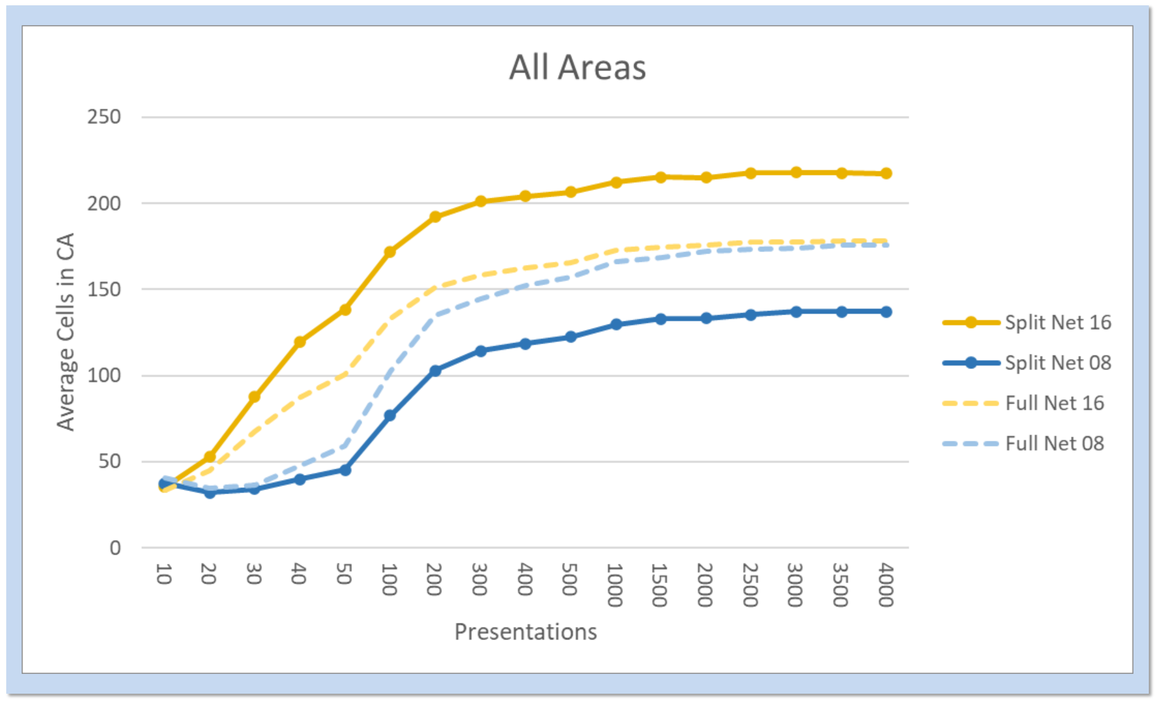
Key conjecture: an increase in CA size (# of cells) leads to a corresponding increase in CA spontaneous ignition frequency

The main aim of this project, thus, is to show that the brain-constrained model exhibits an emergent reward-seeking behaviour that replicates the gradual speech sound selection process seen in infants, i.e., that it autonomously develops a tendency to produce more frequently "words" that match those "heard" in its surroundings.

METHODS

- **Phase 1: Simulating "babbling"** -- The network was trained by repeatedly presenting (4000 times each) twelve pairs of patterns to the model's primary "auditory" and "motor" areas, each pair modelling speech-induced acoustic-articulatory neural activity. This resulted in the spontaneous formation of 12 distinct CA circuits, binding sensory and motor information, distributed across all areas of the model.
- **Phase 2: Investigating the effects of "intrinsic reward" –** Phase 1 was repeated under different conditions: namely, during training, a "reward" (i.e., a learning rate 200% that of normal) was used only for half (six) of the twelve patterns.
- **Phase 3 (PENDING): Simulating the emergence of a reward-driven behaviour –** Here, Phase 2 above will be replaced by a a 2-stage training, consisting of **(i)** a period analogous to Phase 1 (though shorter), followed by **(ii)** a period during which the network will be allowed to run without any input stimulation, giving rise to spontaneous CA ignitions. The reward will be applied only when any of six arbitrarily chosen CAs ignites. (This will require identifying a method of detecting the spontaneous ignition of a CA circuit).

PRELIMINARY RESULTS



- Using the same reward for all 12 pattern pairs produces similarly sized CAs (regardless of learning rate value); see **dashed lines**
- Applying the reward only to half of the patterns pairs results in the two sets of CAs exhibiting different sizes; see **solid lines**
- The preliminary results suggest the presence of "competition" between the emerging CA circuits, with size increases in one group resulting, at the same time, in a decrease of similar amplitude in the other group, on average.

NEXT STEPS & (EXPECTED) FINAL RESULTS

- The next phase of testing will involve implementing **Phase 3** (see METHODS above). We expect to observe the same "divergence" in CA sizes as in Phase 2, with rewarded CAs showing larger sizes than non-rewarded ones, although the fact that reward is more "rare" (only present during spontaneous CA ignition) may require longer simulation times.
- In addition, we will also test the "**Key conjecture**", namely, that significant increases in CA size leads to corresponding increases in spontaneous CA ignition frequency.

[1] Triche, A., Maida, A.S. and Kumar, A. (2022) 'Exploration in neo-Hebbian Reinforcement Learning: Computational Approaches to the exploration–exploitation balance with bio-inspired neural networks', Neural Networks, 151, pp. 16–33. doi:10.1016/j.neunet.2022.03.021. [2] Oudeyer, P.-Y. (2018) Computational theories of curiosity-driven learning [Preprint]. doi:10.31234/osf.io/3p8f6. [3] Garagnani, M. and Pulvermüller, F. (2013) 'Neuronal correlates of decisions to speak and act: Spontaneous emergence and dynamic topographies in a computational model of frontal and temporal areas', Brain and Language, 127(1), pp. 75–85. doi:10.1016/j.bandl.2013.02.001.