Modeling Infant Speech Development with Spiking Neural Networks

Timothy M. Shea and Anne S. Warlaumont—Cognitive and Information Sciences, University of California, Merced

Rationale

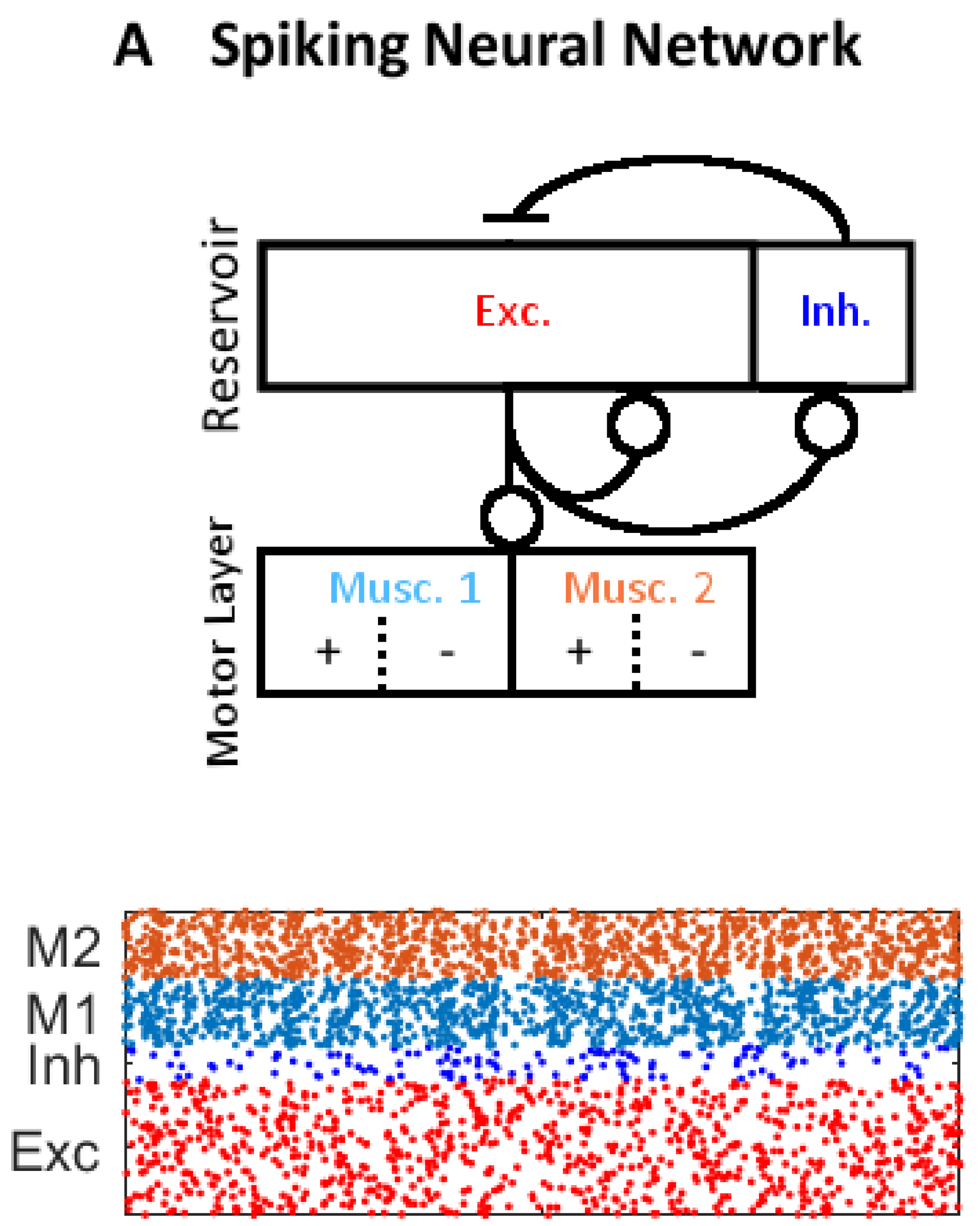
Speech involves precise coordination of numerous muscle groups. Prior work[1] has described a model capable of reproducing a major milestone in vocal learning, the development of syllabic babbling:

1. A recurrent, spiking neural network with reward-modulated spike-timing dependent plasticity (Da-STDP)[2]
2. An articulatory synthesizer simulating the human vocal tract[3]
3. A three stage model of cochlear and auditory cortex response[4]

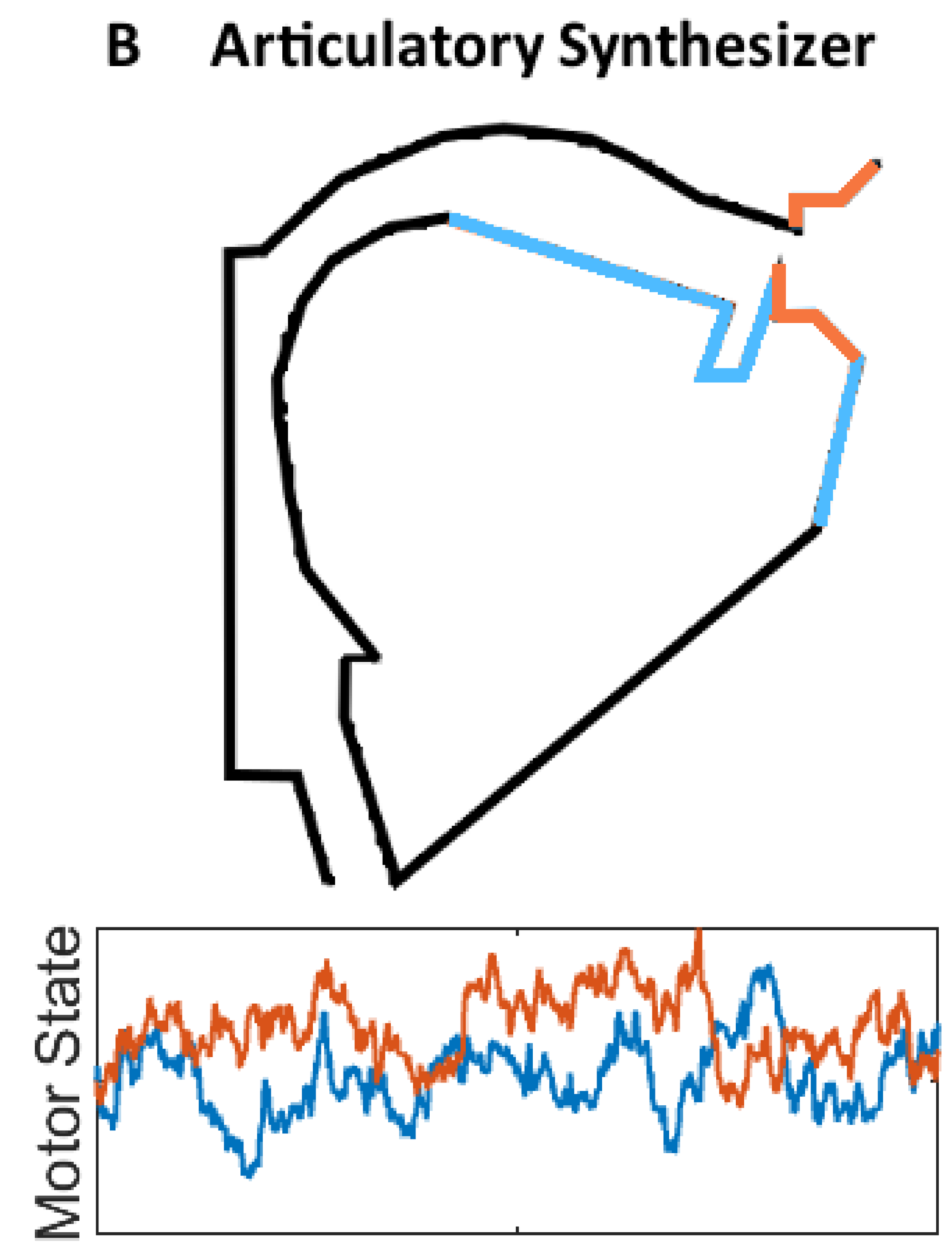
A simplification made in that work is that a single motor signal controls two muscles (1DoF), the orbicularis oris (OO) and masseter (Mass).

Here we explore the effects of controlling OO and Mass muscles independently (2DoF), motivated by the questions:

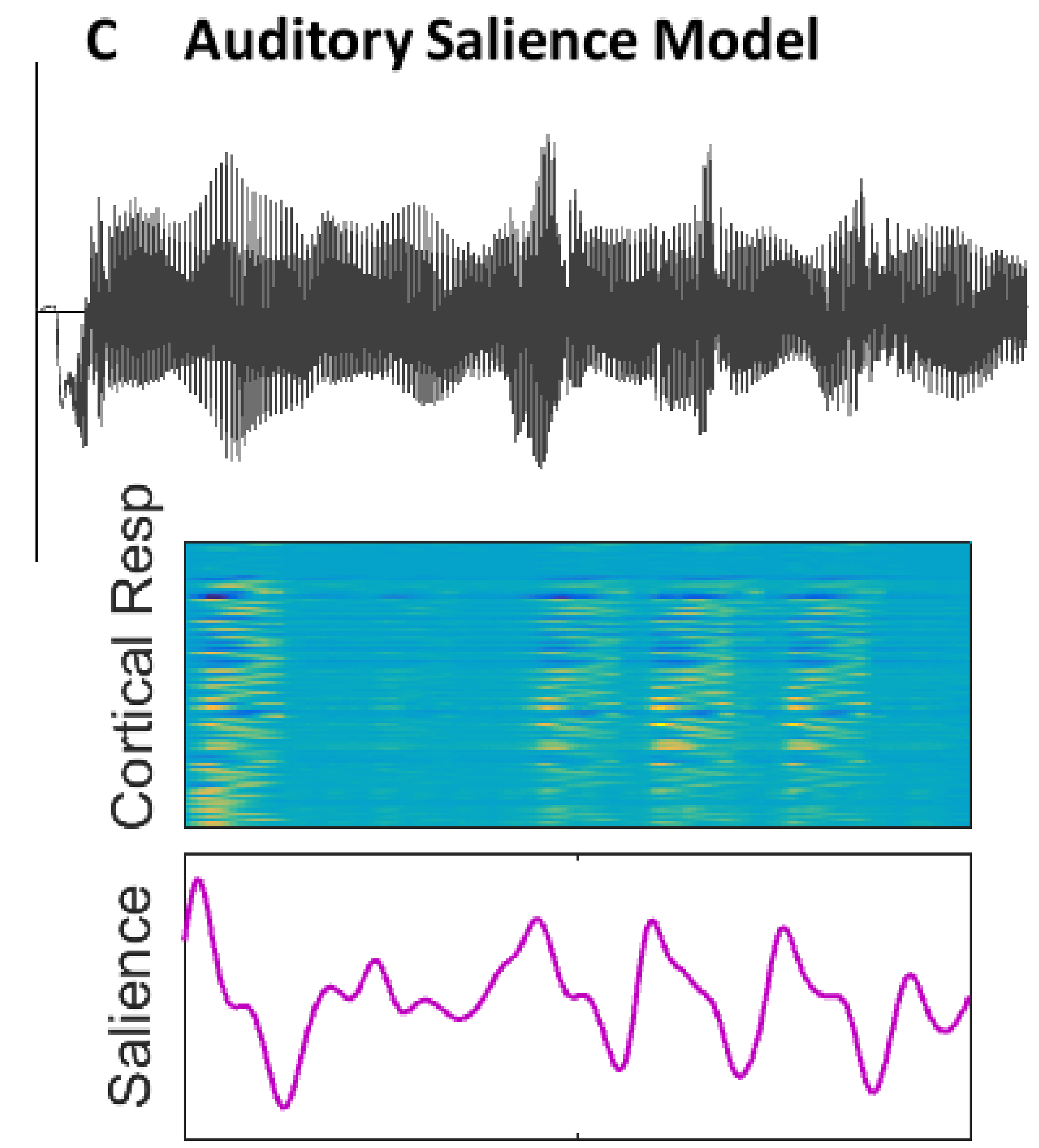
- Can the previous model be scaled up to learn multiple degrees of freedom simultaneously?
- What will we learn from this about infant speech development?



Spiking neural network using **Izhikevich neurons**[3]: 1000 reservoir (80% excitatory) with sparse, recurrent connectivity; 800 motor with excitatory **Da-STDP synapses**.

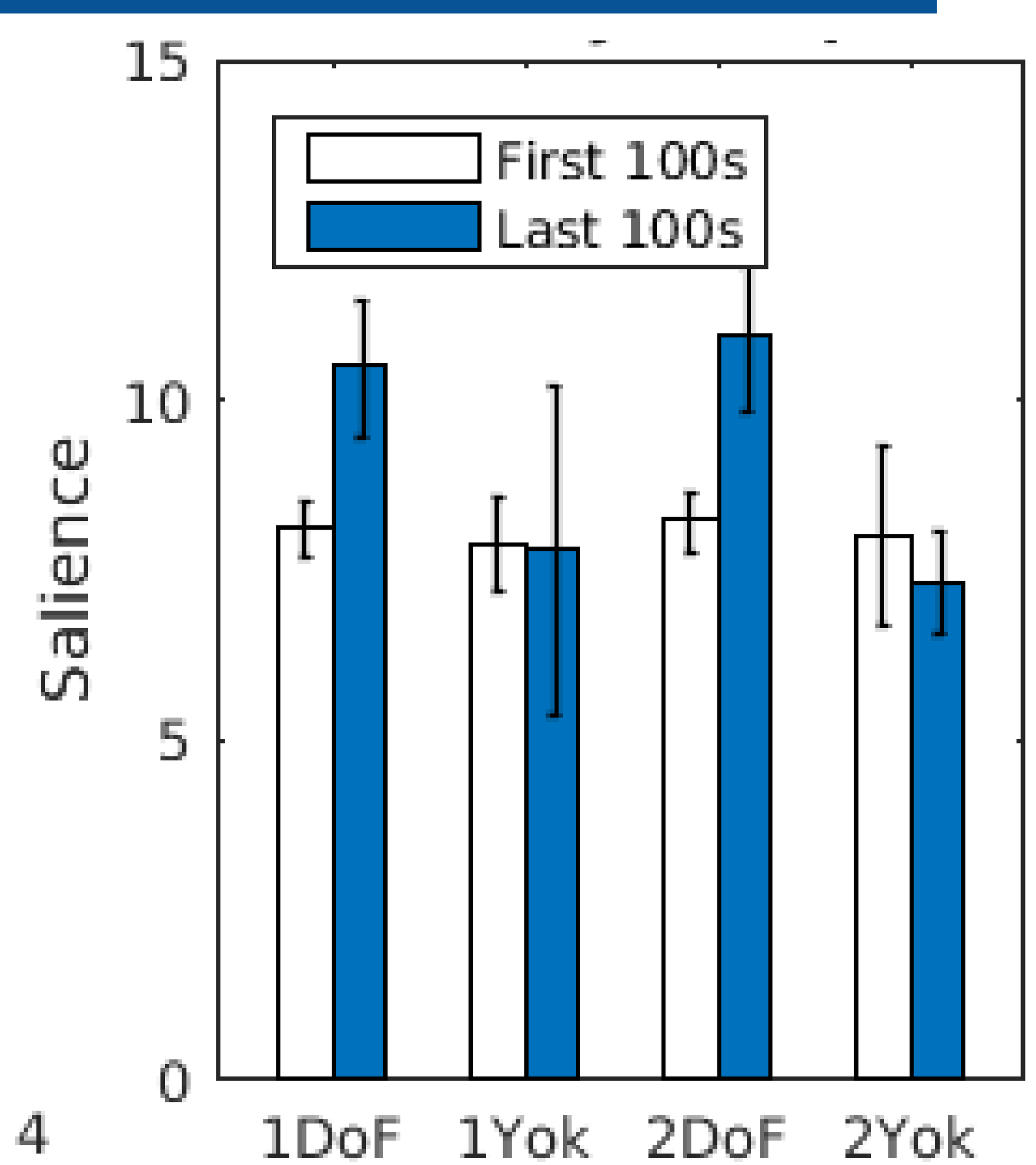


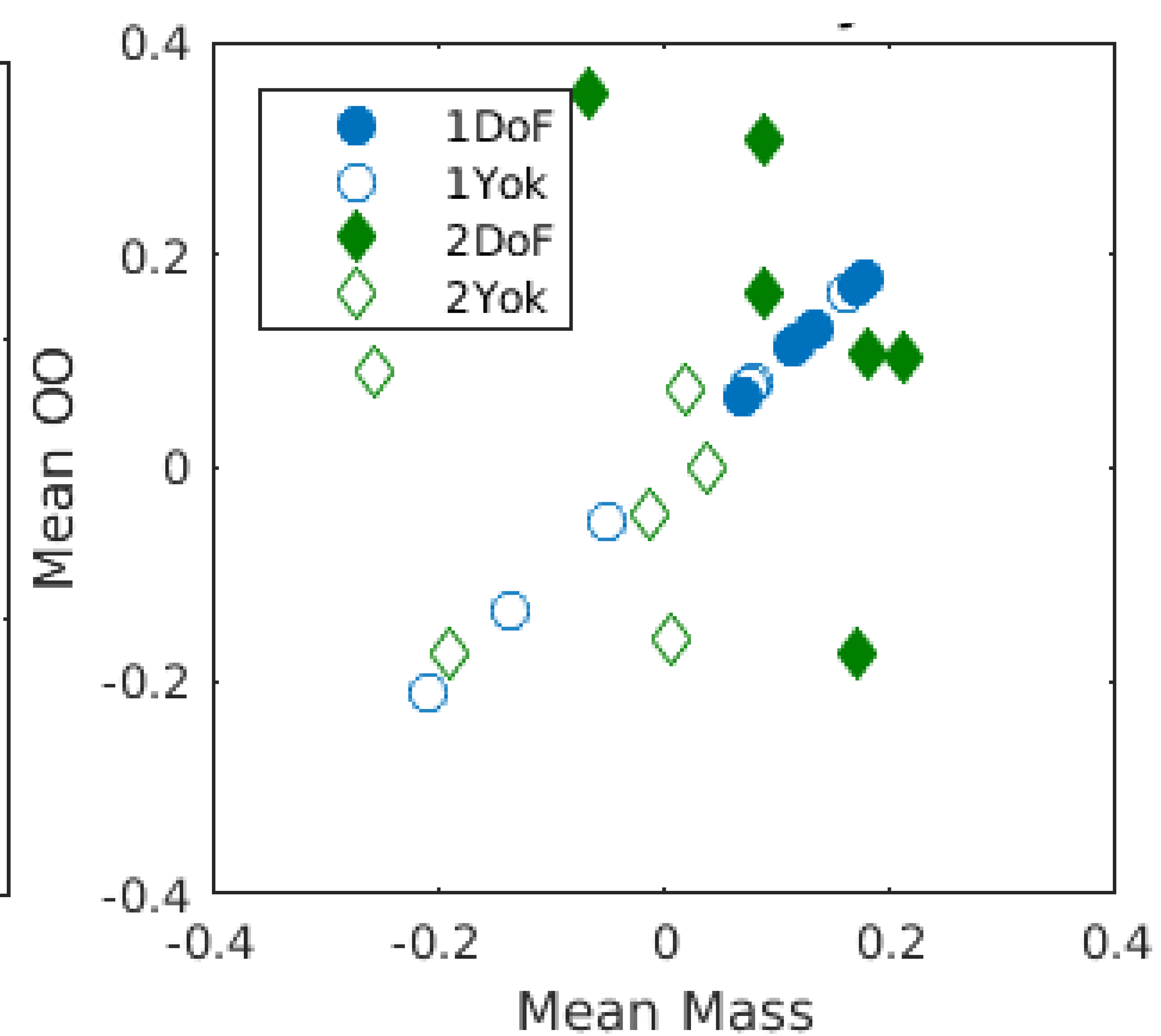
Spikes in **four motor subgroups** drive the muscles (Mass, light blue; OO, orange) of the **Praat articulatory synthesizer**.



Auditory salience model includes acoustic onset detection, cortical response filters, and an **adaptive reward threshold**.

Auditory salience of first and last 100 vocalizations for each group. **Salience increased for both 1DoF and 2DoF (p<0.01)** but not for either yoked group.





Mean position of each muscle in the last 600 trials for each run. Both **1DoF and 2DoF simulations have consistently greater muscle activity** than yokes, but there is **greater variation in 2DoF positions.**

Conclusions

We demonstrated the feasibility of learning multiple DoF simultaneously with the BabbleNN model. We also found that **individual differences in learning can occur as a result of many degrees of freedom with few task constraints.** The model requires significant modifications to support additional muscle groups (3 & 4DoF versions not presented here). This presents a challenge to extending the model control of the entire vocal tract[5].

[1] Warlaumont, AS & Finnegan, MF (2016, PLOS ONE). [2] Izhikevich, EM (2007, Cerebral Cortex). [3] Boersma, P (2001, Glot International). [4] Coath, M, Denham, SL, Smith, L, Honing, H, Hazan, A, Holonowicz, P, & Purwins, H (2007, Neural Information Processing Systems, Workshop on Music Processing in the Brain). [5] Code and sample sounds at https://github.com/tim-shea/BabbleNN.