Fee’s elegant work uses brain cooling techniques to reveal the neural locus of the temporal dynamics of birdsong sequencing. Results support his earlier hypothesis that a chain of activation of populations of HVC neurons controls the sequencing of motor events. Each population is active for a short time (10 ms or so) and projects to RA and from there to the motor neurons, thus engaging the motor controls appropriate for that time epoch. Activation then spreads to the next HVC neuron population, and so on.

In my discussion, I will address (a) how this chain mechanism of timing control could be incorporated into the control of temporal sequencing in speech, (b) what we know about the dynamics of speech that would fail to be captured, if chaining were the sole mechanism, and (c) how this mechanism might relate to the coupled oscillatory model of speech planning (1).

In the oscillatory planning model we have developed, the lowest level of speech production control is modeled as a gestural score, composed of temporally overlapping activation intervals for a set of task-dynamic constriction (gestural) controls. In recent work (2), gestural scores for a large corpus of utterances were analyzed and it was shown that a relatively small number (~200) of distinct combinations of gestural activations (‘gestural pattern vectors’) occur with substantial frequency. Thus, it is possible to model speech as composed of sequences taken from these 200 vectors, and to envisage a chain-type model operative at this level, taking the motor control from one pattern vector to the next.

As will be discussed, at least two aspects of sequencing in speech would fail to be captured by this mechanism alone. One involves evidence for dynamics among speech units that span individual time slices, e.g., gestures. Such evidence includes speech errors, phase re-setting, and phase transitions. A second piece involves evidence for temporally localized rate changes, for example, those associated with the boundaries between prosodic events.
