Representation of speech stimuli by a model of spiking neurons exhibiting gamma oscillations

Oded Ghitza

Sensory processing is associated with gamma frequency oscillations (30–80 Hz) in sensory cortices. This raises the question whether gamma oscillations can be directly involved in the representation of time-varying stimuli, in particular stimuli whose time scale is longer than a gamma cycle. Shamir *et al* (2009) addressed this issue with a dynamical model of spiking neurons and study the response to an asymmetric sawtooth input current over a range of shape parameters. (These parameters describe how fast the input current rises and falls in time.) They chose to study the response of the network to a sawtooth for a reason: sawtooth waveforms simulate short speech segments in a single frequency (cochlear) band. Hence, the study should be viewed as a first step towards developing a model capable of decoding acoustic signatures of *dyad-long speech segments*. (A dyad is the transitional region between a consonant and vowel, as in [ba], some 40-ms in duration.)

The network studied by Shamir *et al* consists of inhibitory and excitatory populations that are sufficient for generating oscillations in the gamma range. The oscillations period is about one-third of the stimulus duration. Embedded in this network is a subpopulation of excitatory cells that respond to the sawtooth stimulus and a subpopulation of cells that respond to an onset cue. The intrinsic gamma oscillations generate a temporally sparse code for the external stimuli. In this code, an excitatory cell may fire a single spike during a gamma cycle, depending on its tuning properties and on the temporal structure of the specific input; the identity of the stimulus is coded by the list of excitatory cells that fire during each cycle. A potential attractive property of the network: it is insensitive to time-scale variation of the input signal (a prerequisite for developing a speech-decoding network insensitive to phonemic variations).

The ultimate goal of the current study is to generalize the model of Shamir *et al*, to decode dyad units in the speech signal. Toward this end, we shall study the decoding capabilities of their model for test signals with a variety of shapes (e.g. asymmetric *trapezoid*, which is a better cartoon of the behavior of a speech segment in a single frequency band) and intensities. Next, we shall study the decoding capabilities of a generalized network, e.g. with 24 components, for 24 critical bands covering the speech band up to 5 kHz.

Shamir, M., Ghitza, O., Epstein, S. and Kopell, N. (2009). "Representation of time-varying stimuli by a network exhibiting oscillations on a faster time scale." PLoS Comput Biol 5(5). doi:10.1371/journal.pcbi.1000370