

REPRESENTATION OF PITCH IN POPULATION-WIDE INTERSPIKE INTERVAL STATISTICS

Peter Cariani

Department of Otolaryngology, Harvard Medical School
Eaton Peabody Laboratory of Auditory Physiology, Mass. Eye & Ear Infirmary, 243 Charles St., Boston, MA 02114 USA.
cariani@epl.meei.harvard.edu; www.cariani.com

ABSTRACT

Temporal coding of sensory information can arise through exogenous, stimulus-locked temporal patternings of spikes or through endogenous, stimulus-triggered temporal patterns of response [1, 2]. Phase-locking to stimulus waveforms permits strategies for stimulus localization that are based on temporal cross-correlation across sensory receptors and strategies for representation of form that are based on temporal autocorrelation within sensory channels. Phase-locked neural responses can be found in a variety of sensory modalities: audition, somatoception, vision (Fig. 1), proprioception (stretch receptors), and electroception.

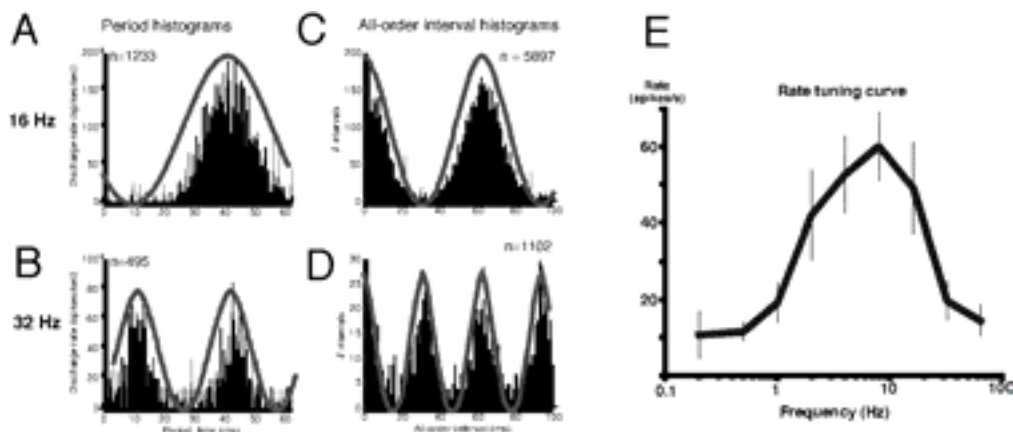


Figure 1. Phase-locking in a visual neuron. Responses of a typical lateral geniculate unit to a sinusoidal grating drifting at constant velocity. Period histograms (A,B) and all-order interval histograms (C,D) to resulting temporal modulations at 16 and 32 Hz. Spatial frequency rate tuning curve (E). Data courtesy of A. Prezybyszewski and D. Pollen [3].

In audition there is strong evidence that both localization and pitch perception utilize fine timing information provided by the phase-locking of auditory nerve fibers to acoustic stimuli. Our experimental work has focused on the early representation of pitch in the auditory pathway [4-6]. The population-wide distribution of all-order interspike intervals at the level of the auditory nerve (Fig. 2) forms an autocorrelation-like temporal representation of the stimulus whose properties explain many diverse aspects of pitch perception (e.g. missing fundamentals, level-invariance, pitch equivalence, octave similarity, pitch shifts of inharmonic complex tones) (see also [7-10]). Population-interval distributions (PIDs) also provide robust representations for those aspects of timbre that are related to stationary power spectra (e.g. formant structure and vowel quality). PID-based models of masking and harmonic resolvability that are based on the competition of interval patterns integrate this information across cochlear territories in a manner that reflects cochlear excitation patterns. Spike timing information capable of supporting PID-based representations is readily available in auditory brainstem, although this information becomes steadily sparser and less abundant as one ascends the pathway beyond the midbrain.

A central question for auditory neurophysiology concerns the means by which the central auditory system might make use of such timing information. The bandpass modulation tunings that have been observed in the auditory pathway [11] fail to account for the level-invariant nature of pitch and pitch shifts of inharmonic tones. Thus far, except for a few possible exceptions, central pitch detectors that could account for pitch equivalences between pure tones

and harmonic complexes have not been found. It is therefore difficult to envision where time-rate coding transformations and periodicity detections would exist in the ascending auditory pathway. Two possible alternative working hypotheses involve 1) covert sparse-temporal coding (based on intervals or synchrony) that is distributed across neurons, 2) top-down/bottom processing strategies in which up central auditory stations organize lower level processing and access lower level pitch representations when needed.

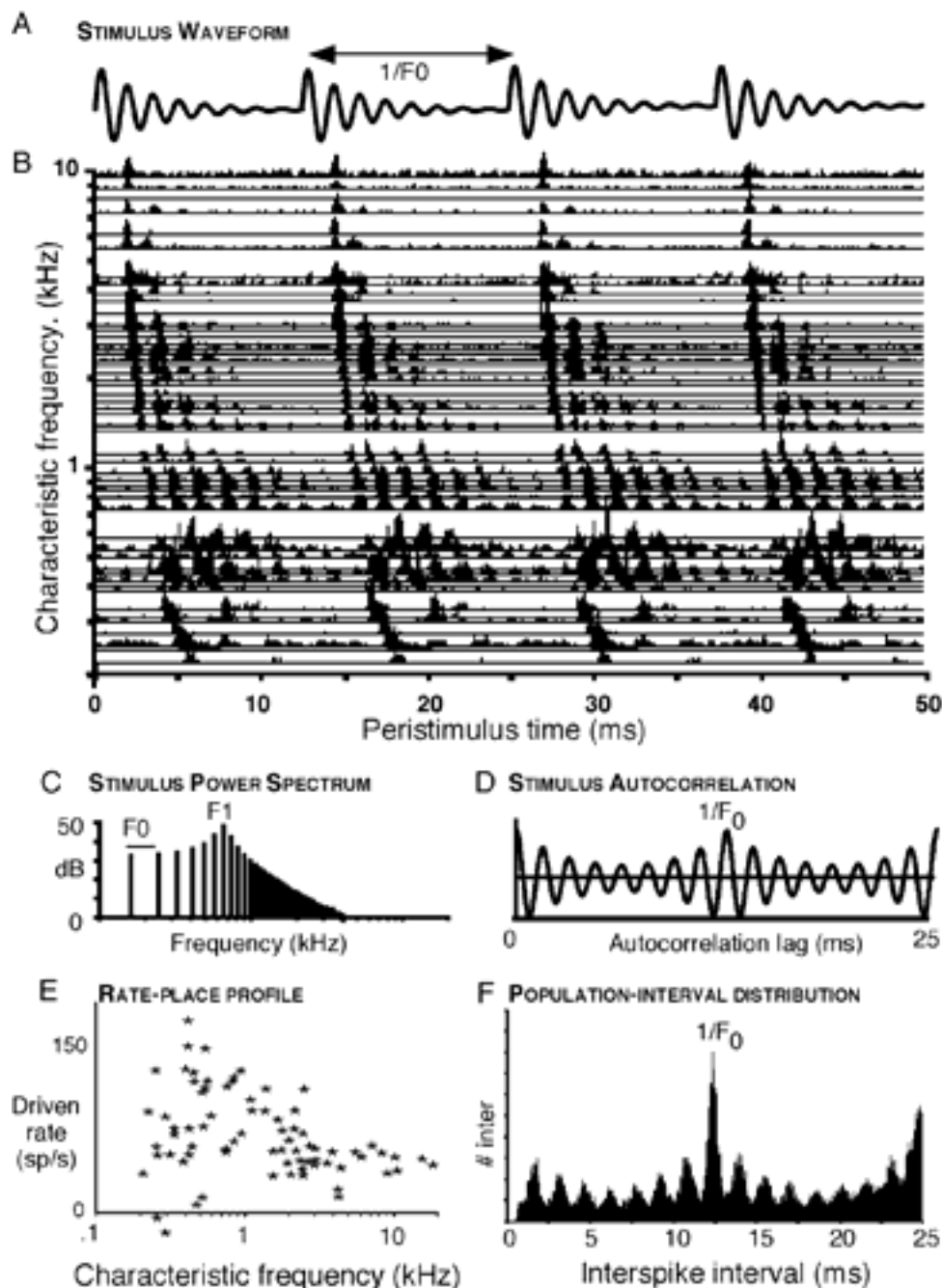


Figure 2. Population-interval representation of the low pitch of a complex tone in the auditory nerve. Auditory nerve responses to a single-formant vowel. A. Vowel waveform. A strong, low voice pitch is heard at the fundamental period of the vowel ($F_0=80$ Hz, $1/F_0=12.5$ ms). B. Peristimulus time histograms of 51 cat auditory nerve fibers to 100 stimulus presentations at 60 dB SPL. Histogram baselines indicate fiber CF's. C. Power spectrum of the vowel (0.1-10kHz, log scale). D. Driven discharge rates as a function of CF. E. Stimulus autocorrelation function. F. Population-interval distribution formed by summing all-order intervals from all fibers.

As a consequence of these theoretical difficulties, we have explored alternative, time-domain processing strategies for utilizing stimulus-driven fine timing information [12, 13]. In the spirit of Licklider's duplex time-delay neural network model [8], they are meant as heuristics that

illustrate functional principles and possibilities rather than descriptions of the input-output behavior of particular neuronal populations. A principal advantage of using temporally-coded pulse patterns rather than the rate-place patterns of classical connectionist networks is that temporally-coded signals can be liberated from their dependence on particular transmission lines and processing elements. Ultimately we seek neural networks for temporal coding and processing that are capable of mass action, multiplexing and de-multiplexing of signals, signal elaboration, broadcast of information, asynchronous-heterarchical rather than sequential-hierarchical organization, content-addressable memory, nonlocal storage of memory, compositionality, and support for multiple logical types. Thus far, only the functional capabilities of simple feedforward and recurrent timing nets have been explored.

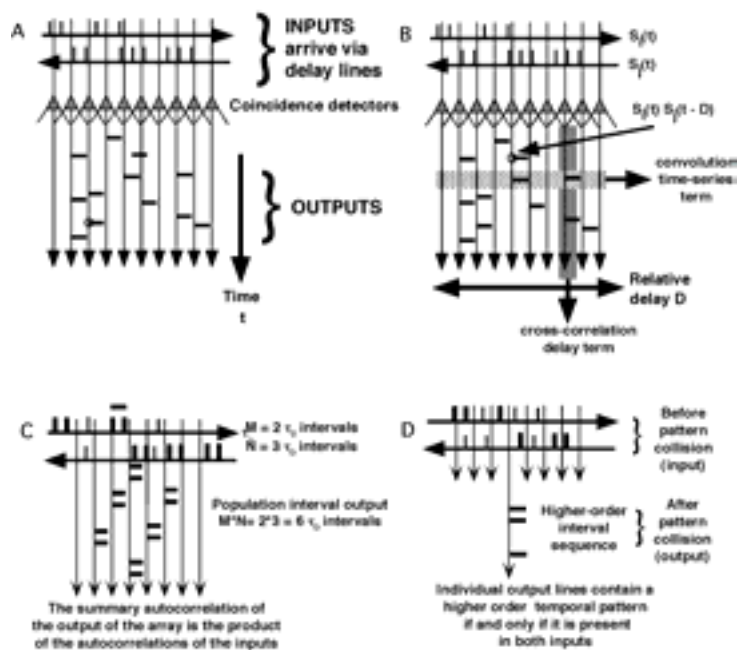


Figure 3. Feedforward timing nets. Basic computational operations.

Feedforward neural timing nets (Fig. 3) are coincidence arrays that function as temporal pattern sieves to pass intervals and interval-sequences that are shared across inputs. As a result, these simple mechanisms are potentially useful for demultiplexing embedded spike patterns. Such nets multiply the autocorrelations of their input spike trains, and thus can be used to extract intervals related to a common fundamental period (pitch) shared by two stimuli even when the stimuli have no harmonics in common (different timbres), and vice versa.

Recurrent timing nets are coincidence arrays with delay loops that build up and separate auditory objects with different fundamentals (e.g. different voices, instruments) on the basis of temporal pattern invariances. These recurrent nets suggest alternative correlation-based strategies for scene analysis that are based on temporal pattern coherence rather than on feature detection and binding.

Keywords: temporal coding, interspike interval, autocorrelation, pitch, scene analysis, neural timing nets.

References

- [1] Cariani, P. (1995) in: *Origins: Brain and Self-Organization*, (K. Pribram, Ed.), Lawrence Erlbaum, Hillsdale, NJ, pp. 208-252.
- [2] Cariani, P. (2001) Temporal coding of sensory information in the brain. *Acoust. Sci. & Tech.*, **22**: 77-84.
- [3] Przybylski, A., Gasa, J.P., Foote, W., and Pollen, D.A. (2000) Striate cortex increases contrast gain in macaque LGN neurons. *Visual Neuroscience*, **17**: 485-494.

- [4] Cariani, P. (1999) Temporal coding of periodicity pitch in the auditory system: an overview. *Neural Plasticity*, **6**: 147-172.
- [5] Cariani, P.A. and Delgutte, B. (1996) Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase-invariance, pitch circularity, and the dominance region for pitch. *J. Neurophysiol.*, **76**: 1717-1734.
- [6] Cariani, P.A. and Delgutte, B. (1996) Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J. Neurophysiol.*, **76**: 1698-1716.
- [7] Meddis, R. and Hewitt, M.J. (1991) Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification. *J. Acoust. Soc. Am.*, **89**: 2866-2882.
- [8] Licklider, J.C.R. (1951) A duplex theory of pitch perception. *Experientia*, **VII**: 128-134.
- [9] Cariani, P. (2002) Temporal codes, timing nets, and music perception. *J. New Music Res.*, **30**: 107-136.
- [10] Tramo, M.J., Cariani, P.A., Delgutte, B., and Braidá, L.D. (2001) Neurobiological foundations for the theory of harmony in western tonal music. *Ann N Y Acad Sci*, **930**: 92-116.
- [11] Langner, G. (1992) Periodicity coding in the auditory system. *Hearing Research*, **60**: 115-142.
- [12] Cariani, P. (2001) Neural timing nets. *Neural Networks*, **14**: 737-753.
- [13] Cariani, P. (2001) in: *Computational Models of Auditory Function*, (S. Greenberg and M. Slaney, Ed.), IOS Press, Amsterdam, pp. 235-249.