Animal models of auditory temporal processing

Jos J. Eggermont *

Department of Physiology and Pharmacology, Department of Psychology, University of Calgary, Calgary, Alberta, Canada

A B S T R A C T

Human temporal processing relies on bottom-up as well as top-down mechanisms. Animal models thereof, in the vast majority, are only probing the bottom-up mechanisms. I will review the vast literature underlying auditory temporal processing to elucidate some basic mechanisms that underlie the majority of temporal processing findings. Some basic findings in auditory temporal processing can all be based on mechanisms determining perstimulatory adaptation of firing rate. This is based on transmitter release mechanisms in peripheral as well as central synapses. It is surprising that the adaptation and recovery time constants that define perstimulatory firing rate adaptation are not very different between auditory periphery and auditory cortex when probed with similar stimuli. It is shown that forward masking, gap and VOT detection, and temporal modulation transfer functions are all directly related to perstimulatory adaptation, whereas stimulus-specific adaptation is at least partly dependent on it. Species differences and the fact that most of the studies reviewed were done in anesthetized animals need to be taken into account when extrapolating animal findings to human perceptual studies. In addition, the accuracy of first-spike latency plays a major role in sound localization and in the brainstem mechanisms for periodicity pitch and forms the basis for understanding evoked potential studies in humans. These mechanisms are also crucial for determining neural synchrony underlying perceptual binding and some important aspects of stream segregation.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Human temporal processing relies on bottom-up as well as top-down processing. Animal models thereof, in the vast majority, are only probing the bottom-up mechanisms. I will highlight the few exceptions besides extensively reviewing the bottom-up mechanisms.

The temporal response properties of the auditory system typically refer to the way the timing of neural activity in the nervous system follows the various temporal aspects of sound. Temporal aspects of sound are, e.g., onsets and offsets, gaps, and periodicities such as amplitude modulations (AM) and carrier fine structure. The capacity of neurons in the auditory system to follow the temporal aspects of sound is limited first of all by synaptic and intrinsic neural properties (such as refractoriness following an action potential) of auditory nerve fibers (ANFs) that transmit these aspects to the central auditory nervous system. This prohibits any neuron to follow frequencies, period-by-period, of more than a few hundred Hz. This limit can be overcome by a population code that adheres to the 'volley' principle. According to this principle, neurons reliably phase-lock to the carrier fine structure but not to all periods thereof but as a population all of these individual cycles can be covered.

Some very basic mechanisms underlying temporal processing can all be based on those determining perstimulatory adaptation of firing rate. It has been shown that forward masking in ANFs is completely determined by these adaptation characteristics. Furthermore, gap detection and voice-onset-time (VOT) representation follow from forward masking. Temporal modulation transfer functions (tMTFs) are also determined by perstimulatory adaptation. The crucial mechanism is the transmitter release from the ribbon synapses in the inner hair cells, but this can be extended to what is happening in the central nervous system by taking into account intrinsic properties of central neurons and network interactions. Here I will review the phenomenology of auditory temporal processing in animals and point to mechanisms and models that aim to describe and interrelate these findings.

2. First-spike latency precision and local field potentials

One important finding (Heil and Irvine, 1997) regarding first-spike latency was that, based on the characteristics of the increase of standard deviation with latency, the precision of first-spike timing to some stimuli can actually be higher in primary auditory cortex (A1) than in the AN. Across the recorded neurons in auditory cortex (Phillips and Hall, 1990), SDs seen in responses of minimum latency were related to minimal mean latencies, and were typically in the range from 0.15 to 1.5 ms (mean = 0.88 ms). This mean SD was only slightly larger than that described for ANFs (0.73 ms: Rhode and Smith, 1986). It is,

* Department of Psychology, University of Calgary, 2500 University Drive N.W., Calgary, Alberta T2N 1N4, Canada. Tel.: +1 403 220 5214; fax: +1 403 282 8249.
  E-mail address: eggermon@ucalgary.ca.

however, inferior to that characterizing onset and chopper units in the ventral cochlear nucleus (VCN; Rhode and Smith, 1986). The onset cells had SDs that were often in the range from 0.02 to 0.05 ms (mean = 0.09 ms), and the chopper neurons typically had SDs less than 0.4 ms (mean = 0.38 ms).

It is surprising that the precision of response timing in the cortex is as good as it is. According to Rhode and Smith (1986), this probably reflects the benefits of a convergent input that may improve the response timing "signal/noise" ratio. That is, the cumulative jitter to be expected from the serial synapses in the afferent path is offset by an averaging of convergent, nearly synchronous, input response times. Phillips and Hall (1990) suggested that the cortex has little or no crucial role in the temporal representation of sounds with content in the sub-millisecond time frame, but it does indicate that the timing of the phonetically important components of spoken language such a voice-onset-time, formant transitions and prosody can be represented by spike times in the primary auditory cortex.

2.1. Role in sound localization

Among information-bearing features of neural responses, first-spike latency is particularly robust and (theoretically) easy to decode, especially when only one spike is elicited per stimulus. Spike timing in the auditory cortex is sensitive to sound-source location and carries as much or more spatial information than does firing rate (Brugge et al., 1996; Furukawa and Middlebrooks, 2002). The timing of the first post-stimulus spike carries a substantial amount, probably the majority, of the location-related information present in cat auditory cortex spike patterns, both in primary auditory cortex (A1; Eggermont, 1998a) and secondary auditory cortex (Furukawa and Middlebrooks, 2002). Eggermont (1998a) evaluated latencies of simultaneously recorded single-unit (SU) spikes and local field potentials (LFP) in A1 of cats for sound presented from 9 speakers arranged horizontally in the frontal half field in a semicircular array with a radius of 55 cm and the cat's head in the center. SU post-stimulus time histograms (PSTH) were made for each speaker location for a 100 ms window after noise burst onset using 1 ms bins. Differences with the minimum spike latency observed in an animal were calculated for all azimuth–intensity combinations. This relative latency showed a weakly-sigmoidal dependence on azimuth that was independent of intensity level above 40 dB SPL. Single-unit latency differences were also measured with respect to the latencies of the LFP triggers, simultaneously recorded on the same electrode. This difference was independent of stimulus intensity and showed a nearly linear dependence on sound azimuth. The mean differences across animals for both measures, however, were only significant between contralateral azimuths on one hand and frontal and ipsilateral azimuths on the other hand. Mean spike–LFP latency differences, showed a monotonic dependence on azimuth with nearly constant variance and may provide the potential for an unbiased conversion of azimuth into neural firing times. The general trend for the modal inter-spike interval (ISI) was the same as for relative spike latency: the shortest ISIs were found for contralateral azimuths (ISI usually 3 ms) and the longer ones for ipsilateral azimuths (the most frequent ISI was 4 ms, occasionally 5 ms was found). This trend was also independent of intensity level. This suggests that there is little extra information in the timing of spikes in addition to those found in the first-spike-latency histogram.

Furukawa and Middlebrooks (2002) tested the hypothesis that most of the stimulus-related information in spike patterns is carried by the latency of the first spike. Results indicated that for many units, the first spike in the response pattern carried nearly all (~89%) of the stimulus-related information. Stecker and Middlebrooks (2003a) examined the nature of spatial representation within the auditory cortex, and strongly suggested that first-spike latency is a critical information-bearing feature of cortical responses. Posterior auditory field (PAF) neurons responded with long and stimulus-sensitive latencies. In contrast, A1 neurons responded with short latencies that varied only minimally between stimulus locations. Whereas spike counts of units in A1 and PAF transmitted similar quantities of information, first-spike latencies transmitted significantly greater spatial information in PAF than A1 (Stecker et al., 2003). This suggests an important role for sound localization in PAF. Lomber and Malhotra (2008) corroborated this when they found that, after selectively cooling of the PAF, cats showed severe deficits in sound localization.

2.2. Onset representation in neural populations

Population onset, i.e., as reflected in auditory evoked potentials (AEPs), is a measure that has been used widely in studying temporal processing in humans. To appreciate this, we have to examine the relationship between the microscopic, single-unit activity and the macroscopic AEPs generated by auditory stimulation. Relevant data for this purpose are (1) locally recorded single-unit activity or multi-unit activity (MUA) and LFPs, and (2) globally measured neural activity at the scalp. This global scalp activity can sometimes be localized more or less accurately and attributed to specific structures in the brain. Scalp-recorded AEPs or magnetic fields are often onset responses. This could be because the response is generated by neurons that respond specifically to stimulus onsets. It could also be that the neuronal responses that occur at points of change in the stimulus are the only ones sufficiently synchronized to generate a field potential.

We (Eggermont et al., 2011) made multi-electrode array recordings of spike and LFP activity from A1 of 12 normal hearing, ketamine-anesthetized cats. We evaluated spectro-temporal receptive fields (STRFs) and frequency-tuning curves (FTCs) based on LFPs and spikes simultaneously recorded on the same electrode. We compared their characteristic frequency (CF) gradients and their cross-correlation distances. The CF gradient for spike-based FTCs was about twice that for 2–40 Hz-filtered LFP-based FTCs, indicating greatly reduced frequency selectivity for LFPs. For spontaneous LFP and spike activity we evaluated 1373 pair correlations for pairs with >200 spikes in 900 s per electrode. Peak correlation-coefficient space constants were similar for the 2–40 Hz filtered LFP (5.5 mm) and the 16–40 Hz LFP (7.4 mm), whereas for spike-pair correlations it was about half that, at 3.2 mm. This larger correlation distance combined with the reduced CF gradient and much broader frequency selectivity suggests that LFPs are not a substitute for spike activity in primary auditory cortex.

Kajikawa and Schroeder (2011) echoed this by comparing LFPs, current source density (CSD), and MUA in macaque auditory cortex. Estimated by frequency-tuning bandwidths, these signals' "listening areas" differed systematically in size whereby MUA < CSD < LFP. Direct measurements indicated passive spread of LFPs to sites more than a centimeter from their origins. These findings appeared to be independent of the frequency content of the LFP. These results suggest that LFPs appear as a mixture of local potentials and "volume conducted" potentials from distant sites.

To analyze this further, de Cheveigné et al. (2013) recorded LFPs and spikes from auditory cortex of guinea pigs using 16-channel electrode arrays. LFPs were processed by an independent component analysis technique that produced optimally tuned linear combinations of electrode signals. Linear combinations of LFPs were found to have sharply frequency-tuned responses, closer to spike-related frequency tuning. The existence of a sharply tuned component implies that a downstream cortical neuron (or group of neurons) capable of forming a linear combination of its inputs has access to that information contained in LFPs.

3. Perstimulatory adaptation and forward masking

3.1. Perstimulatory adaptation

In response to tone bursts of moderate sound intensity, mammalian ANFs produce a maximum firing rate at onset, followed by an adaptation