

The role of transients in auditory processing

Martin Coath*, Susan L. Denham

Centre for Theoretical and Computational Neuroscience, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK

Received 11 November 2005; accepted 20 April 2006

Abstract

Enhancement of auditory transients is well documented in the auditory periphery and mid-brain, and single unit investigations have identified units with responses which may underlie this sensitivity. It is also known that transients are important in psychophysics in, for example, speech comprehension and object recognition and grouping.

In this work we use a simple phenomenological model of auditory transient extraction, based on the skewness of the distribution of energy inside a frequency dependent time window, and show that this view is consistent with electrophysiological measurements of auditory brainstem responses. In addition, we present evidence that this representation may provide a positive biological advantage in processing classes of sound that are behaviourally relevant.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Transients; Onsets; Auditory pathways; Acoustic stimulation; Perception; Statistical; Whitening; Speech

1. Introduction

A great deal of work over many years has led to the characterization of a very large number response types in the early stages of auditory processing (Trussel, 2002). This work has provided some idea, at a cellular level, of the answer to the important question: *what do cells in the early stages of the auditory pathway do?* The answer to the more general *why?* question regarding the purpose of such processing is, however, less clear.

There certainly is evidence that:

1. the auditory system performs a spectral decomposition (this can be modelled by a finite number of band pass filters) (Patterson et al., 1992), and many cells in the auditory periphery of many different types exhibit well defined characteristic frequencies (Trussel, 2002),

2. there is progressive enhancement of transients both onset and offset, throughout the ascending auditory pathway (Phillips et al., 2002; Heil, 1997a),
3. the temporal envelope within channels is emphasized at the expense of the spectral envelope which carries information about the relationship between channels (Drullman, 1995; Fu et al., 1998; Shannon et al., 1998).

Taken together these results imply that there is a role for a within-frequency-band, transient sensitive (temporal-edge sensitive) representation of stimuli within the auditory system.

It is well documented that the auditory system is sensitive to the temporal structure of the amplitude envelope, particularly rising (*onset*) transients. This has been shown both in physiological and psychophysical measurements (e.g. Phillips et al., 2002; Heil, 1997a). This sensitivity increases as measurements are made at successively higher levels in the auditory pathway. Units that detect onsets are found throughout the auditory system: in VCN (Frisina et al., 1985; Rhode and Greenberg,

* Corresponding author.

E-mail address: mcoath@plymouth.ac.uk (M. Coath).

1994), IC (Heil and Irvine, 1996; Langner and Schreiner, 1988) MGB (thalamus) (Rouiller et al., 1981, 1982), cortex (Eggermont, 2002).

Responses to falling amplitude transients (offsets) are less often reported but well attested in the literature, e.g. He et al. (1997), He (2002), Phillips et al. (2002), VanCampen et al. (1997), Kulesza et al. (2003), and Behrend et al. (2002). In particular Kulesza et al. (2003) recording in the paraolivary nucleus of the rat (a structure thought to be homologous to dorsomedial periolivary nucleus in other mammals) report that 95% of cells in this structure ($n = 116$) generated spikes wholly, or primarily, after the offset of pure tone and broadband noise stimuli with short (mean = 7.0 ms) first spike latencies. Results from recordings of auditory brainstem responses (ABR) (VanCampen et al., 1997) show that offset responses are of comparable amplitude, and exhibit similar latencies to onset responses. This is somewhat at odds with the asymmetry of onsets and offsets measured in single unit recordings (Phillips et al., 2002) and the origins of these responses at a cellular level are still obscure. However, onset responses have been shown to be locked to the maximum acceleration of pressure in electrophysiological recordings (Heil, 1997b) and this maximum acceleration is found at the very beginning of the stimulus for both sine-squared and linear onset ramps. Given this and the similar latencies measured for onsets and offsets in ABRs, it is plausible that the offset response is similarly tied to some aspect of the offset ramp.

If instances of rising energy (onsets) and falling energy (offsets) are to be identified then the question arises ‘over what time scales are these events to be judged?’ In this matter there is evidence (Wiegreb, 2001; Krumbholz et al., 2003) which suggests that parallel frequency channels may be processed on different, frequency-dependent, time scales. It seems plausible that this might apply to onset and offset judgements as well.

This of course does not represent a complete picture of sub-cortical auditory processing, but if it captures at least some essential aspects of the early auditory processing stages then it is germane to ask what the functional advantages of this processing might be. To investigate this question we implemented a simple model consisting of a cochlear filter bank followed by within-channel transient detection based on the third order moment (skewness) of energy distribution in frequency-dependent time windows. Skewness is a measure of the asymmetry of energy distribution in time, as opposed to the rate of change as might be measured by a first differential, and has the advantage that it is robust to short term variations and is relatively insensitive to initial and final values in the window.

In previous work (Coath and Denham, 2005; Coath et al., 2005) we have shown that such a representation preserves features which support a wide range of behaviourally important judgements in human speech. Here we further investigate whether the resulting representation offers any advantage to the organism with respect to processing ethological auditory stimuli.

2. Representations of auditory stimuli

2.1. Cochlear model

The first stage of the model approximates processing in the cochlea. Sounds are processed using a bank of 30 Gammatone filters (Slaney, 1994), with centre frequencies (CFs) ranging from 100 Hz to about 8000 Hz arranged evenly on an equivalent rectangular bandwidth (ERB) scale (Glasberg and Moore, 1990). The output in each frequency channel is half wave rectified as a simple model of inner hair cell (IHC) function. The output of this processing for a speech stimulus is shown in Fig. 1. For brevity this representation will be referred to as the ‘simple cochlear model’ or SCM.

2.2. Fourier representation

A widely used representation of auditory stimuli is based on a Fourier estimate of the power at each frequency present in a series of overlapping time windows. We will refer to this as the ‘fast Fourier transform’ or FFT representation. Although not a ‘biological’ model this representation is not greatly dissimilar to some published ‘cochleagrams’ which use methods similar to that

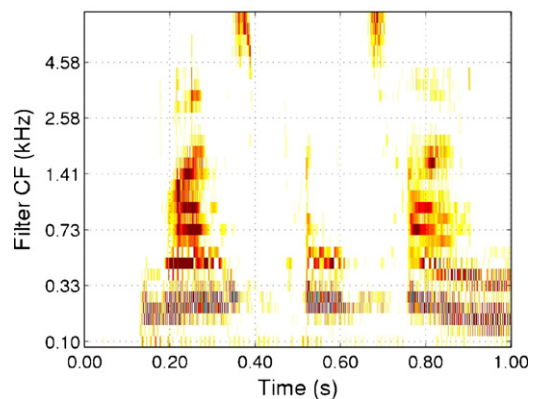


Fig. 1. Output of the cochlear model (SCM) of ‘once upon a time’ (male speaker). Speech is used here as an example stimulus to illustrate that the onsets and offsets present in a natural or ethological signal are captured by the method discussed in Section 2.3. This short sample contains examples of a wide variety of speech sounds: vowels, sibilants, plosives and so on, see Fig. 4.

Download English Version:

<https://daneshyari.com/en/article/2077156>

Download Persian Version:

<https://daneshyari.com/article/2077156>

[Daneshyari.com](https://daneshyari.com)