



## Tonotopic representation of loudness in the human cortex



Andrew Thwaites<sup>a,b,\*</sup>, Josef Schlittenlacher<sup>a</sup>, Ian Nimmo-Smith<sup>b</sup>,  
William D. Marslen-Wilson<sup>a,b</sup>, Brian C.J. Moore<sup>a</sup>

<sup>a</sup> Department of Psychology, University of Cambridge, Cambridge, UK

<sup>b</sup> MRC Cognition and Brain Sciences Unit, Cambridge, UK

### ARTICLE INFO

#### Article history:

Received 28 July 2016

Received in revised form

24 November 2016

Accepted 29 November 2016

Available online 1 December 2016

#### Keywords:

Tonotopy

Magnetoencephalography

Loudness

Temporal integration

Model expression

Entrainment

### ABSTRACT

A prominent feature of the auditory system is that neurons show tuning to audio frequency; each neuron has a characteristic frequency (CF) to which it is most sensitive. Furthermore, there is an orderly mapping of CF to position, which is called tonotopic organization and which is observed at many levels of the auditory system. In a previous study (Thwaites et al., 2016) we examined cortical entrainment to two auditory transforms predicted by a model of loudness, instantaneous loudness and short-term loudness, using speech as the input signal. The model is based on the assumption that neural activity is combined across CFs (i.e. across frequency channels) before the transform to short-term loudness. However, it is also possible that short-term loudness is determined on a channel-specific basis. Here we tested these possibilities by assessing neural entrainment to the overall and channel-specific instantaneous loudness and the overall and channel-specific short-term loudness. The results showed entrainment to channel-specific instantaneous loudness at latencies of 45 and 100 ms (bilaterally, in and around Heschl's gyrus). There was entrainment to overall instantaneous loudness at 165 ms in dorso-lateral sulcus (DLS). Entrainment to overall short-term loudness occurred primarily at 275 ms, bilaterally in DLS and superior temporal sulcus. There was only weak evidence for entrainment to channel-specific short-term loudness.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### 1. Introduction

The loudness of a sound is a subjective attribute corresponding to the impression of its magnitude. Glasberg and Moore (2002) proposed that there are two aspects of loudness for time-varying sounds such as speech and music. One is the short-term loudness, which corresponds to the loudness of a short segment of sound such as a single word in speech or a single note in music. The second is the long-term loudness, which corresponds to the overall loudness of a relatively long segment of sound, such as a whole sentence or a musical phrase. Glasberg and Moore (2002) proposed a model in which transformations and processes that are assumed to occur at relatively peripheral levels in the auditory system (i.e. the outer, middle, and inner ear) are used to construct a quantity

*Abbreviations:* CF, characteristic frequency; DLS, dorso-lateral sulcus; EMEG, electro- and magneto-encephalographic; HG, Heschl's gyrus; KID, Kymata identifier  
\* Corresponding author. Department of Psychology, Downing Site, University of Cambridge, Cambridge, CB2 3EB, UK.

*E-mail addresses:* [acgt2@cam.ac.uk](mailto:acgt2@cam.ac.uk) (A. Thwaites), [js2251@cam.ac.uk](mailto:js2251@cam.ac.uk) (J. Schlittenlacher), [iannimmosmith@gmail.com](mailto:iannimmosmith@gmail.com) (I. Nimmo-Smith), [wdm10@cam.ac.uk](mailto:wdm10@cam.ac.uk) (W.D. Marslen-Wilson), [bcjm@cam.ac.uk](mailto:bcjm@cam.ac.uk) (B.C.J. Moore).

<http://dx.doi.org/10.1016/j.heares.2016.11.015>

0378-5955/© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

called “instantaneous loudness” that is not available to conscious perception, although it appears to be represented in the brain (Thwaites et al., 2016). At later stages in the auditory system the neural representation of the instantaneous loudness is assumed to be transformed into the short-term loudness and long-term loudness, via processes of temporal integration.

In a previous study (Thwaites et al., 2016), we investigated whether the time-varying instantaneous loudness or short-term loudness predicted by the loudness model of Glasberg and Moore (2002) is ‘tracked’ by cortical current, a phenomenon known as cortical entrainment. The extent of cortical entrainment was estimated from electro- and magneto-encephalographic (EMEG) measures of cortical current, recorded while normal-hearing participants listened to continuous speech. There was entrainment to instantaneous loudness bilaterally at 45 ms, 100 ms and 165 ms, in Heschl's gyrus (HG), dorso-lateral sulcus (DLS) and HG, respectively. Entrainment to short-term loudness was found in both the DLS and superior temporal sulcus at 275 ms. These results suggest that short-term loudness is derived from instantaneous loudness, and that this derivation occurs after processing in sub-cortical structures.

Missing from this account is how overall instantaneous loudness

is constructed from the information that flows from the cochlea. A prominent feature of the auditory system is that neurons show tuning to audio frequencies; each neuron has a characteristic frequency (CF) to which it is most sensitive. This is a consequence of the filtering that occurs in the cochlea, which decomposes broadband signals like speech and music into multiple narrow frequency channels. The information in each channel is transmitted, in parallel, along the afferent fibers making up the auditory nerve (Helmholtz, 1863; Fuchs, 2010; Meyer and Moser, 2010; von Bekeesy, 1949), and information encoded in these channels reaches the cortex in separate locations (Saenz and Langers, 2014). Furthermore, there is an orderly mapping of CF to position, which is called tonotopic organization and which is observed at many levels of the auditory system (Palmer, 1995).

The model of Glasberg and Moore (2002) is based on the assumption that neural activity is combined across CFs, i.e. across frequency channels, before the transformation to short-term loudness. It is not known where in the brain such a combination might occur, but if there is cortical entrainment to channel-specific activity, the model leads to the prediction that this entrainment should occur with a shorter latency than for short-term loudness. In this study, we aimed to test this, by assessing neural entrainment to the instantaneous loudness predicted by the loudness model for nine frequency sub-bands, hereafter called “channels”, spanning the range from low to high frequencies.

The sequence of transforms assumed in the model of Glasberg and Moore (2002) (channel-specific instantaneous loudness, followed by overall instantaneous loudness, followed by overall short-term loudness, outlined in Fig. 1) is not the only plausible sequence leading to the neural computation of overall short-term loudness. Short-term loudness may instead be derived separately for each channel, and these short-term loudness estimates may then be combined across channels to give the overall short-term loudness, as assumed in the models of Chalupper and Fastl (2002) and Moore et al. (2016). To test this alternative, we created a modified version of the model in which the short-term loudness was calculated separately for each channel and we assessed whether there was cortical entrainment to the channel-specific short-term loudness values. If neural responses are combined across CFs before the derivation of short-term loudness, then there might be cortical entrainment to the channel-specific instantaneous loudness values but not to the channel-specific short-term loudness values. If the short-term loudness is derived separately for each CF and then the short-term loudness estimates are combined across CFs, then cortical entrainment might be observed both to the channel-specific instantaneous loudness values and to the channel-specific short-term loudness values. Furthermore, the cortical entrainment to the channel-specific short-term loudness values should occur earlier in time than the cortical entrainment to the overall short-term loudness values.

In addition to standard graphic representations, an interactive representation of this study’s results can be viewed on the online Kymata Atlas (<http://kymata.org>). For easy reference, each model in this paper (referred to as a ‘function’ in Kymata) is assigned a *Kymata ID* [KID].

## 2. Defining candidate models

To measure cortical entrainment, some constraints must be imposed on the models/transforms that can be tested. Any model that takes a time-varying signal as input and produces a time-varying signal as output can be used, with function  $f()$  characterizing the mechanism by which the information (in this case the acoustic waveform) is transformed before it produces cortical entrainment. Thus, if both input  $x_1, \dots, x_m$  and output  $y_1, \dots, y_n$  are of

duration  $t$ , the model takes the form:

$$f(x_1, x_2, x_3, \dots, x_t) = (y_1, y_2, y_3, \dots, y_t), \quad (1)$$

where  $f()$  is bounded by a set of formal requirements (Davis et al., 1994) and a requirement that  $y_i$  cannot be dependent on any  $x_k$  where  $k > i$  (this last requirement avoids hypothesizing a non-causal  $f()$  where a region can express an output before it has the appropriate input).

Previously, three ‘auditory magnitude’ models were tested using the same dataset as the one used here: instantaneous loudness, short-term loudness, and the Hilbert envelope (KIDs: QRLFE, B3PU3 and ZDSQ9, respectively). The instantaneous loudness model (Moore et al., 1997; Glasberg and Moore, 2002) and short-term loudness model (Glasberg and Moore, 2002) represent successive transformations that approximate physiological processing in the peripheral and central auditory system. The third model, the Hilbert envelope model, was uninformed by physiological processing, and was included as a naïve comparison model.

The loudness model of Glasberg and Moore (2002) is based on a series of stages that mimic processes that are known to occur in the auditory system. These stages are: (1) A linear filter to account for the transfer of sound from the source (e.g. a headphone or loudspeaker) to the tympanic membrane; (2) A linear filter to account for the transfer of sound pressure from the tympanic membrane through the middle ear to pressure difference across the basilar membrane within the cochlea; (3) An array of level-dependent bandpass filters, resembling the filters that exist within the cochlea (Glasberg and Moore, 1990). These filters are often called the “auditory filters” and they provide the basis for the tonotopic organization of the auditory system; (4) A compressive nonlinearity, which is applied to the output of each auditory filter, resembling the compression that occurs in the cochlea (Robles and Ruggero, 2001).

The model of Glasberg and Moore (2002) is based on the assumption that the compressed outputs of the filters are combined to give a quantity proportional to instantaneous loudness. The moment-by-moment estimates of instantaneous loudness are then smoothed over time to give the short-term loudness. However, here we also consider a model based on the assumption that the smoothing over time occurs on a channel-specific basis, and that the overall short-term loudness is derived by summing the short-term specific loudness values across channels. The estimates of overall instantaneous loudness and channel-specific instantaneous loudness were updated every 1 ms. Examples of the structure of these models (and the models’ predicted activity) are shown in Fig. 1.

We chose to use nine channels, each 4 Cams wide on the  $ERB_N$ -number scale (Glasberg and Moore, 2002; Moore, 2012). The  $ERB_N$ -number scale is a perceptually relevant transformation of the frequency scale that corresponds approximately to a scale of distance along the basilar membrane. The equation relating  $ERB_N$ -number in Cams to frequency,  $f$ , in Hz is (Glasberg and Moore, 1990):

$$ERB_N\text{-number} = 21.4 \log_{10}(0.00437f + 1) \quad (2)$$

The normal auditory system uses channels whose bandwidths are approximately 1-Cam wide (Moore, 2012). However, cortical entrainment to the activity in 1-Cam wide channels might be very weak, because each channel contains only a small amount of energy relative to the overall energy. We used 4-Cam wide channels so as to give reasonable frequency specificity while avoiding channels that were so narrow that they would contain little energy and therefore lead to very “noisy” cortical representations. A second reason for not using 1-Cam wide channels is that the

Download English Version:

<https://daneshyari.com/en/article/5739385>

Download Persian Version:

<https://daneshyari.com/article/5739385>

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