Hearing Research 296 (2013) 83–95

Contents lists available at SciVerse ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Modelling detection thresholds for sounds repeated at different delays

Peter Heil [a,b,*], Jesko L. Verhey [b,c], Benedikt Zoefel [a]

[a] Leibniz Institute for Neurobiology, Brenneckestr. 6, 39118 Magdeburg, Germany
[b] Center for Behavioral Brain Sciences, 39118 Magdeburg, Germany
[c] Department of Experimental Audiology, Otto von Guericke University Magdeburg, Leipziger Str. 44, 39120 Magdeburg, Germany

ARTICLE INFO

Article history:
Received 27 June 2012
Received in revised form
3 December 2012
Accepted 5 December 2012
Available online 23 December 2012

ABSTRACT

Detection thresholds for pairs or multiple copies of sounds are better than those for a single sound, an observation commonly interpreted as indicating temporal integration by the auditory system. Detection thresholds for pairs of brief tones depend on the delay between the tones (if short) and on frequency, suggesting frequency-dependent temporal overlap of auditory-filter responses elicited by the two successive stimuli (Krumbholz and Wiegrebe, 1998). The model presented by Krumbholz and Wiegrebe did not account for all aspects of their data, despite its complexity. This study shows that a simple probabilistic model based on Neubauer and Heil (2008) predicts the increase in threshold for short temporal delays as well as the asymptotic behaviour towards longer delays. The model entails (i) a 4th-order gammatone filter with a brief impulse response and thus broad bandwidth (shorter and broader than those of a filter normally assumed), (ii) the formation of stochastic 'spikes' or 'events' whose probability of occurrence is proportional to the filter output (half-wave rectified fine-structure or amplitude envelope), raised to a power of 3, and (iii) probability summation. The same model with the same front-end filter also predicts thresholds for pairs of clicks presented in band-reject noise, measured by Hall and Lummis (1973). The model accurately predicts the magnitudes and the decay of the alternating increase and decrease of thresholds as the delay between the click varies, the small effects of click polarity, and the dependence of thresholds for pairs of clicks with unequal intensities on their temporal order. Finally, we show that this model also correctly predicts the decrease in threshold with increasing number of temporally separated brief sounds, reported in several studies. While the latter data do not constrain the characteristics of the front-end filter, they do confirm the exponent of 3 in the model. Our paper stresses the viability of the model and raises the possibility that the bandwidths of filters estimated with psychophysical techniques may depend more strongly on the experimental paradigms and stimuli than hitherto thought.

1. Introduction

In every vertebrate species studied, the threshold for the detection of a sound, when specified in terms of the required sound amplitude or level, decreases with increasing duration of the sound (for review see e.g., Fay, 1992). Similarly, detection threshold is lower for a pair of successive identical sounds than for a single sound (Hall and Lummis, 1973; Viemeister and Wakefield, 1991; Krumbholz and Wiegrebe, 1998) and decreases further with increasing number of identical sounds in each trial or observation interval (e.g. Zwislocki et al., 1962; Gerken, 1966; Carlyon et al., 1990; Gerken et al., 1990; Solecki and Gerken, 1990). This trade-off between stimulus amplitude and duration is frequently referred to as a "temporal-integration function". The name is somewhat unfortunate, as criticized by e.g. Viemeister and Wakefield (1991) and Meddis (2006), since it implies a particular neural mechanism. In fact, the most common interpretation of the trade-off has been that the auditory system acts as a temporal integrator of some quantity of the sound. The quantity often thought to be integrated by the auditory system is sound intensity (e.g., Plomp and Bouman, 1959; Dallos and Johnson, 1966; Green, 1985; Clock Eddins and Peterson, 1999; O'Connor et al., 1999; Recanzone and Sutter, 2008). Sound intensity is proportional to the square of the sound's amplitude and its temporal integral yields a measure of acoustic energy. Another widely-held view is that the integrator is leaky, rather than perfect, and that its behaviour can be

Abbreviations: ABR, auditory brainstem response; ERB, equivalent rectangular bandwidth; ERB_N, ERB of normal-hearing listeners; LIEFTS, leaky integration; event formation, temporal summation.

* Corresponding author. Leibniz Institute for Neurobiology, Brenneckestr. 6, 39118 Magdeburg, Germany. Tel.: +49 391 6263 94441; fax: +49 391 6263 95489.
E-mail addresses: peter.heil@lin-magdeburg.de (P. Heil), jesko.verhey@med.ovgu.de (J.L. Verhey), benedikt.zoefel@lin-magdeburg.de (B. Zoefel).

0378-5955/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.heares.2012.12.002

modelled analogously to the charging of a capacitor in parallel with a resistor (e.g., Plomp and Bouman, 1959; Zwislocki, 1960; Eddins and Green, 1995; Recanzone and Sutter, 2008). A comprehensive review of these ideas is provided by Verhey (2010). However, the values of the time constants required to fit experimental data with this model are often hundreds of milliseconds (for reviews see Gerken et al., 1990; O'Connor et al., 1999). Such long time constants are at variance with the observations that detection thresholds for pairs or trains of short sounds do not change once the temporal separation of these sounds exceeds a few milliseconds (Flanagan, 1961; Hall and Lummis, 1973; Gerken et al., 1990; Viemeister and Wakefield, 1991; Krumbholz and Wiegrebe, 1998). Such long time constants are also difficult to reconcile with the high temporal resolution of the auditory system, creating the "resolution-integration paradox" (deBoer, 1985; Green, 1985).

Viemeister and Wakefield (1991) suggested a "multiple-looks" model as an explanation for the lack of dependence of detection thresholds for pairs of short sounds on their temporal separation once it exceeds a few milliseconds. According to this model, the listener samples the sound at a high rate. The detectability during each sample or "look" is proportional to the sound's intensity. Each look provides independent information, which is stored in memory and which can be accessed and processed selectively, and information from multiple looks can be combined in a near-optimal fashion. While the model appears to be able to account for several experimental findings, including the trade-off between threshold sound level and duration, a number of uncertainties remain. They include the nature of the look, the shape of the temporal window that defines a look, the possibility that successive looks might not be independent, and the nature and characteristics of the memory used to store the information from the looks.

Heil and Neubauer (2003) suggested another alternative to long-time-constant temporal integration. According to their ideas, the stimulus generates stochastic neural detection events. The stimulus is detected when a criterion number of such events (e.g., 1) has occurred (probability summation). Thus, the higher the probability of events per unit of time, the shorter the average time required before the criterion number of events has occurred. Neither integration with long time constants nor storage of information is required. Rather, the stimulus has, at any point in time, a certain probability of exceeding threshold and being detected. Consequently, when the stimulus amplitude is fixed, the cumulative probability of exceeding threshold increases as stimulus duration increases, and performance improves. Equivalently, as stimulus duration increases, the sound amplitude required for a given performance (e.g., the threshold criterion) decreases. Meddis and Lecluyse (2011) recently formulated a similar probabilistic model of detection thresholds, which, however, differs from that of Heil and Neubauer (2003) in important aspects. For example, Meddis and Lecluyse assume the probability of detection events to be directly proportional to the (maximum) amplitude of the stimulus. In contrast, Heil and Neubauer (2003) derived that this probability was proportional to the stimulus amplitude raised to an exponent $\alpha > 1$. Initial estimates of α from different data sets and species ranged from 3 to 5 (Heil and Neubauer, 2003), but later it became clear, also from physiological studies of the auditory nerve, that the most likely value is $\alpha = 3$ (Neubauer and Heil, 2004, 2008; Heil et al., 2008, 2011). This parsimonious model has been termed LIEFTS model (for leaky integration, event formation, temporal summation; Neubauer and Heil, 2008).

Here, we primarily examine whether the LIEFTS model is also capable of accounting for the detection thresholds for pairs of brief stimuli, relative to those for a single stimulus, as a function of the temporal separation of the stimuli in a pair. Such data were reported by Hall and Lummis (1973), Viemeister and Wakefield (1991), and Krumbholz and Wiegrebe (1998). The latter study found that the range of delays over which the sharp transition from relative thresholds of about −6 dB at short delays to about −2 dB at longer delays occurred shifted to shorter delays as the carrier frequency of the brief tonal stimuli increased. This frequency-dependence of threshold-delay functions suggests the involvement of peripheral filters whose bandwidth increases with increasing centre frequency, so that the duration of their 'ringing' response decreases with increasing centre frequency. Krumbholz and Wiegrebe (1998) presented a model which included such a filter at the front end. The bandwidth of the filter was equal to the presumed equivalent rectangular bandwidth for normal-hearing subjects (ERB_N) which increases with centre frequency (Glasberg and Moore, 1990). The post-filter processing was inspired by the multiple-looks model of Viemeister and Wakefield (1991). Nevertheless, despite its complexity, the model of Krumbholz and Wiegrebe (1998) did not satisfactorily account for their data. Here, we show that a frequency-specific version of the LIEFTS model can do so. However, a filter with a bandwidth broader than ERB_N is required. We then show that the LIEFTS model with the same broad filter also correctly predicts the findings of Hall and Lummis (1973). These authors observed that relative thresholds for pairs of clicks presented in band-reject noise alternately increase and decrease as the delay varies. The LIEFTS model also correctly predicts the subtle effects of click polarity on relative thresholds and the dependence of relative thresholds for pairs of clicks with unequal intensities on their temporal order, as observed by Hall and Lummis (1973). Finally, we show that the data of Krumbholz and Wiegrebe (1998), Hall and Lummis (1973), Viemeister and Wakefield (1991), and data from several studies measuring relative thresholds for multiple copies of successive identical sounds (e.g., Carlyon et al., 1990; Gerken et al., 1990; Solecki and Gerken, 1990) confirm the exponent of 3 in the LIEFTS model which becomes apparent when the delay between the sounds exceeds the duration of the filter responses.

2. Methods

2.1. Data acquisition

The data to be modelled here are mainly from the studies of Krumbholz and Wiegrebe (1998), Hall and Lummis (1973), Carlyon et al. (1990), Gerken et al. (1990), and Solecki and Gerken (1990). All details of data acquisition can be found in these papers. We extracted the relative thresholds (in dB) from the relevant figures in the papers of Krumbholz and Wiegrebe (1998; their Figs. 3 and 5), Hall and Lummis (1973; their Figs. 4, 5 and 7–9), and Carlyon et al. (1990; their Fig. 1). The figures were exported into CorelDraw X3 (Corel Corporation) and the data points were determined with the highest accuracy possible. The remaining errors are very small and negligible compared to the errors in measuring the data. Absolute thresholds (in dB SPL) of the human listeners in the study of Gerken et al. (1990) and of the normal-hearing cats in the study of Solecki and Gerken (1990) were kindly provided to the first author by George M. Gerken.

2.2. Modelling

The frequency-specific LIEFTS model used here, as described in Section 3.1.3, was implemented in Matlab (The Mathworks Inc., Natick MA, USA). The scripts for the gammatone filter at the front-end of this model were kindly provided by Volker Hohmann (University of Oldenburg, Germany) and are described in Hohmann (2002). In order to predict relative thresholds, the output of the model (see 3.1.3.) to a single stimulus (a tone or a click) was