Development and Evaluation of a Model for Predicting the Audibility of Time-Varying Sounds in the Presence of Background Sounds*

BRIAN R. GLASBERG AND BRIAN C. J. MOORE, *AES Member*
(bg12@cam.ac.uk) (bcjm@cam.ac.uk)

Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB, England

A model for predicting the audibility of time-varying signals in background sounds is described. The model requires the calculation of time-varying excitation patterns for the signal and background, using the methods described elsewhere. A quantity called instantaneous partial loudness (IPL) is calculated from the excitation patterns. The estimates of IPL, which are updated every 1 ms, are used to calculate the short-term partial loudness (STPL) using a form of running average similar to an automatic gain control system. It is assumed that the audibility of the signal is monotonically related to the average value of the STPL over the duration of the signal. In experiment 1 thresholds were measured for detecting a 1-kHz sinusoid in four different samples each of white and pink "frozen" noise. The results were used to determine the average value of the STPL required for threshold. In experiment 2 the model was evaluated by measuring detection thresholds for nine signal types in six backgrounds (54 combinations), using a two-alternative forced-choice task. The backgrounds were chosen to be relatively steady (such as traffic noise). The correlation between the measured masked thresholds and those predicted by the model was 0.94. The root-mean-square difference between the thresholds obtained and those predicted was 3 dB. In experiment 3 psychometric functions were measured for the detection of five signals in five backgrounds (five pairs), using a two-alternative forced-choice task. Experiment 4 used the same signals and backgrounds, but psychometric functions were measured using a single-interval yes–no task. The results of experiments 3 and 4 were used to construct functions relating signal detectability d' to the average value of the STPL.

0 INTRODUCTION

There are many practical situations where it would be useful to be able to predict the audibility of signals in the presence of background sounds. Examples include the warning signals used in aircraft or trains [1] and mobile telephone ring tones, which often need to be heard in a variety of noisy situations. Previous work on predicting the audibility of such sounds has mostly been based on the power-spectrum model of masking [2], [3], which is applicable only to relatively steady signals in steady backgrounds. In this paper we describe a new model for predicting the audibility of signals, which can be applied to time-varying signals, and we illustrate the operation of the model in predicting the audibility of a variety of mobile telephone ring tones. We start with a brief review of the background to our model.

*Manuscript received 2005 February 13; revised 2005 September 8.

In an earlier paper [4] we described a model for the calculation of the loudness of steady sounds from the spectra of those sounds. The model drew on concepts and methods of many earlier researchers [5]–[10]. The 1997 model provided the basis for a new ANSI standard for the calculation of loudness [11]. The model also allowed the calculation of partial loudness, that is, the loudness of a signal in the presence of a background sound. This required separate specifications of the spectrum of the signal and the spectrum of the background. The masked threshold of the signal in the background could be predicted based on the assumption that the partial loudness at the masked threshold is a constant small amount. In fact, it was assumed that the partial loudness at masked threshold is equal to the loudness of a sound in quiet when that sound is at its absolute threshold. The loudness at absolute or masked threshold was assumed to be 0.003 sone, which is equivalent to a loudness level of 2 phons. The model was able to produce reasonably accurate predictions of the masked thresholds of a variety of steady signals in steady backgrounds, including signals composed of single sinusoidal tones, complex tones, and bands of noise.

Two aspects of the model limited its application. First the model required as input a specification of the spectrum of the sound. For most applications, such as estimating the loudness of environmental sounds arising from heating, ventilation, and air-conditioning (HVAC), this meant that spectra had to be calculated (usually in one-third-octave bands) from the time waveforms of the sounds. While one-third-octave spectra give adequate resolution for sounds with smooth broad-band spectra, they do not always give adequate resolution for sounds with discrete spectral components. A second limitation was that the model was applicable only to steady sounds. Many everyday sounds, such as speech and music, are time varying. Also, many warning or alerting signals, such as mobile telephone ring tones, are time varying.

In a later paper [12] we described a model for the calculation of the loudness (but not the partial loudness) of time-varying sounds. This model uses a waveform as its input. The stages of the model are as follows:

1) A finite impulse response (FIR) filter representing transfer through the outer and middle ear. Different filters can be used to allow for different listening situations, such as in a free field, in a diffuse field, or via headphones or a telephone.

2) Calculation of the short-term spectrum using the fast Fourier transform (FFT). To give adequate spectral resolution at low frequencies, combined with adequate temporal resolution at high frequencies, six FFTs are calculated in parallel, using longer signal segments for low frequencies and shorter segments for higher frequencies.

3) Calculation of an excitation pattern from the physical spectrum. The excitation pattern is updated at 1-ms intervals.

4) Transformation of the excitation pattern to a specific loudness pattern.

5) Determination of the area under the specific loudness pattern. This gives a value for the "instantaneous" loudness. The instantaneous loudness was taken as an intervening variable that is not accessible to conscious perception.

For sounds like speech there are two aspects to the loudness impression: the listener can judge the short-term loudness, for example, the loudness of a specific syllable; or the listener can judge the overall loudness of a relatively long segment, such as a sentence. We refer to the latter as the long-term loudness. Within the model the short-term perceived loudness is calculated from the instantaneous loudness using an averaging mechanism similar to an automatic gain control system, with an attack and release time. Then the overall loudness impression is calculated from the short-term loudness using a similar averaging mechanism, but with longer attack and release times.

In this paper we present a new model for calculating the audibility of time-varying signals in time-varying backgrounds. The model requires separate specifications of the waveforms of the signal and of the background. For applications involving listening with two ears, the waveforms of the signal and of the background are specified separately for each ear as stereo "wav" files. The model combines features of the two models described in the foregoing. The instantaneous partial loudness (IPL) for each ear is calculated from the short-term excitation patterns of the signal and background, using the method for calculating short-term excitation described by Glasberg and Moore [12] and the equations for determining partial loudness described by Moore et al. [4]. The IPL estimates for each ear are summed across ears to obtain the binaural IPL. The short-term partial loudness (STPL) is derived from the IPL in the same way as described by Glasberg and Moore [12]. We assume that the audibility of the signal is related to the mean value of the STPL, at least for reasonably long signal durations (above about 200 ms).

The model as presented here may be regarded as predicting audibility based on "energetic" masking, that is, masking that is primarily determined by the overlap of the excitation patterns of the signal and the masker in the peripheral auditory system [2]. However, it is known that under some circumstances the audibility of a signal can be much less than predicted from energetic masking. This occurs especially when the masker spectrum fluctuates markedly from moment to moment in an unpredictable way. The "extra" masking, over and above that attributable to energetic masking, is called informational masking [13]–[18]. It is presumed to arise from relatively central processes in the auditory system. The magnitude of central masking can vary markedly across individuals, whereas energetic masking varies relatively little across individuals with normal hearing [19]. While the model cannot predict the effects of informational masking, it can be useful in indicating when informational masking is occurring. This will be discussed in more detail in a later section.

1 DETAILS OF THE MODEL

In what follows we assume a sampling rate of 32 kHz and 16-bit resolution. For wav files obtained with other sampling rates, a routine for sample-rate conversion can be used. For completeness we describe all of the stages of the model here, although some of these stages are the same as in our earlier models [4], [12]. In stages 1.1 and 1.2 calculations are performed separately for the signal waveform and the background waveform. If the signal and/or background are presented binaurally and differ at the two ears, then the calculations are also performed separately for each ear.

1.1 Transfer through Outer and Middle Ear

The transfer of sound through the outer and middle ear can be modeled using fixed filters, although the filtering produced by the outer ear depends on the direction of incidence of the sound relative to the head [20]. The transfer function of the outer ear for frontal incidence used in our model is given in [4, fig. 2]. The assumed transfer function of the middle ear is given in [4 fig. 3]. Here the combined effect of the outer and middle ear is modeled by a single FIR filter with 4097 coefficients. The filter was designed using the FIR2 function in MATLAB. The transfer characteristic of this filter, for a sound with frontal incidence, is shown in [12, fig. 1]. The filter is constructed so as to have a gain of 0 dB at 1000 Hz. Other filters can be used for different directions of sound incidence. Filters can also be constructed to simulate the effects of headphone presentation, or presentation via loudspeakers, in which cases the transfer function from the headphone/loudspeaker to the eardrum needs to be taken into account. The output of the filter can be considered as representing the sound reaching the cochlea (the inner ear).

1.2 Calculation of Running Spectrum and Excitation Pattern

The cochlea can be characterized as containing a bank of bandpass filters whose center frequencies span the range from about 50 to 15 000 Hz [4]. The bandwidths of the filters increase with increasing center frequency. For a filter centered around 100 Hz the equivalent rectangular bandwidth for normally hearing subjects (ERB_N) is about 35 Hz, whereas at 10 000 Hz it is about 1100 Hz [21]. The filters are level-dependent, the low-frequency slopes becoming less steep with increasing level [21], [22]. The magnitude of the output of each filter in response to a given sound, plotted as a function of filter center frequency, is called the excitation pattern of that sound [23].

The excitation pattern is calculated via an initial spectral analysis using the fast Fourier transform (FFT). To obtain spectral resolution at low frequencies comparable to that of the auditory system, the analysis of relatively long (about 64-ms) segments of the input signal is required. However, for high center frequencies the use of such long segments would have the effect of limiting temporal resolution in a way that does not occur in the auditory system. For example, when a high carrier frequency is used, amplitude modulation at rates up to several hundred hertz can be detected without the resolution of spectral sidebands [24], [25]. To give adequate temporal resolution at high frequencies, an analysis using much shorter signal segments (about 2 ms) is required.

To deal with these problems, six FFTs are calculated in parallel, using signal segment durations that decrease with increasing center frequency. The six FFTs are based on Hanning-windowed segments with durations of 2, 4, 8, 16, 32, and 64 ms, all aligned at their temporal centers. The windowed segments are zero padded and all FFTs are based on 2048 sample points. All FFTs are updated at 1-ms intervals.

Each FFT is used to calculate spectral magnitudes over a specific frequency range; values outside that range are discarded. These ranges are 20 to 80 Hz, 80 to 500 Hz, 500 to 1250 Hz, 1250 to 2540 Hz, 2540 to 4050 Hz, and 4050 to 15 000 Hz, for segment durations of 64, 32, 16, 8, 4, and 2 ms, respectively. An excitation pattern is calculated from the short-term spectrum at 1-ms intervals, using exactly the same method as described in [1]. Briefly the outputs of the auditory filters are calculated for center frequencies spaced at 0.25-ERB intervals, taking into account the known variation of the auditory filter shape with the center frequency and level [21]. The excitation pattern is then defined as the output of the auditory filters as a function of center frequency.

1.3 Calculation of IPL

The next stage in the model is the calculation of the IPL for each ear, from the excitation patterns of the signal and background for that ear. The calculation of IPL from the excitation pattern is done in the same way as for the calculation of the partial loudness of steady sounds, as described in [4]. The reader is referred to that paper for details. Initially the specific partial loudness (the partial loudness per ERB_N) is calculated as a function of the center frequency (on an ERB_N number scale). The area under the resulting specific loudness pattern gives the IPL for a given ear. If two ears are being used, the instantaneous loudness is summed across ears to give the overall IPL.

1.4 Calculation of STPL

STPL, the loudness perceived at any instant, is calculated using a form of temporal integration or averaging of the IPL, which resembles the way that a control signal is generated in an automatic gain control (AGC) circuit with attack time T_a and release time T_r. We define S'_n as the running (averaged) short-term estimate of partial loudness at the time corresponding to the nth time frame (updated every 1 ms), S_n as the calculated IPL at the nth time frame, and S'_{n-1} as the running STPL at the time corresponding to frame $n-1$.

If $S_n > S'_{n-1}$ (corresponding to an attack, as the IPL at frame n is greater than the STPL at the previous frame), then

$$S'_n = \alpha_a S_n + (1 - \alpha_a) S'_{n-1} \tag{1}$$

where α_a is a constant that is related to T_a,

$$\alpha_a = 1 - e^{T_i/T_a} \tag{2}$$

with T_i being the time interval between successive values of the IPL (1 ms in this case).

If $S_n \leq S'_{n-1}$ (corresponding to a release, as the IPL is less than the STPL), then

$$S'_n = \alpha_r S_n + (1 - \alpha_r) S'_{n-1} \tag{3}$$

where α_r is a constant that is related to T_r;

$$\alpha_r = 1 - e^{T_i/T_r}. \tag{4}$$

The values of α_a and α_r are 0.045 and 0.02, respectively.

1.5 Estimation of Mean STPL Required for Threshold

The audibility of a signal in a background is assumed to be monotonically related to the value of the STPL, averaged over the duration of the signal. The first experiment, described in the next section, was performed to allow us to determine what average value for the STPL was required for the signal to be at the detection threshold as measured in perceptual experiments. It should be emphasized that there is no such thing as a threshold above which the signal is always detected and below which it is not. Rather, the probability that a signal is detected increases monotonically with increasing signal-to-background ratio [26].

However, the threshold can be defined statistically as the signal level leading to a certain percentage of correct detections in a specified task. Here the threshold is defined as the level leading to a detectability index d' of 1.16 [26].

2 EXPERIMENT 1: DETECTION THRESHOLDS FOR A 1-kHz SINUSOID IN WHITE AND PINK NOISE

We started by measuring detection thresholds for two combinations of signal and background sounds for which informational masking was likely to be minimal. Both the signal and the backgrounds were sounds that are perceived as steady. The signal was a 1-kHz sinusoid and the backgrounds were white noise and pink noise. These stimuli were also chosen because they have been used in many previous studies, and so it would be possible to check that our results were consistent with those in the literature.

To use the model to predict thresholds for the signal in these backgrounds, specific samples of the background noises have to be used as input to the model. It is known that the perceptual threshold for detecting a tone in noise depends on the specific noise sample selected, since the exact amplitudes and phases of the frequency components in the noise vary from one sample to the next [27], [28]. In our experiment we used four different samples of pink noise and four different samples of white noise. Within a block of trials the same sample of noise was used throughout, that is, the noise was "frozen." However, across blocks of trials the noise sample was varied. The noise samples used in the perceptual experiment were also used as input to the model.

2.1 Stimuli

The signal and background were generated independently from files stored in a computer, using a Tucker-Davies Technologies (TDT) 16-bit digital-to-analog converter (DD1) running at a sampling rate of 32 kHz. The signal and background were low-pass filtered at 10 kHz (Kemo VBF8/04), and their levels were controlled independently using TDT PA4 programmable attenuators. The signal and background were mixed (TDT SM3), passed through a headphone buffer (TDT HB6), and delivered to one ear via Sennheiser HD580 earphones, which have a diffuse-field response. The overall duration of each stimulus was 1000 ms, including 50-ms onset and offset ramps, shaped with a raised-cosine function. The overall level of each background noise was 70 dB SPL. The level of the signal was varied to determine the detection threshold.

2.2 Procedure

Masked thresholds were measured using an adaptive two-alternative forced-choice (2AFC) task. Within a block of trials (a run) the combination of signal and background was fixed. The background was presented in two 1000-ms bursts, separated by 500 ms, and the signal was presented randomly with either the first background burst or the second. Thresholds were measured using a three-down one-up adaptive procedure [29]. A run started with the signal easily audible. The signal level was decreased following three successive correct responses and increased following one incorrect response. This procedure tracks the point on the psychometric function (percent correct plotted as a function of signal level) corresponding to 79.4% correct, which in turn corresponds to a d' value of 1.16 [26], [30]. A change from decreasing to increasing level, and vice versa, defines a turnpoint. The step size in level was 5 dB until four reversals had occurred and was 2 dB thereafter. A run continued until 12 reversals had occurred and the threshold was estimated from the mean of the levels at the last eight reversals. At least three runs were obtained for each subject and each signal/background combination, and the final threshold was estimated by averaging across all runs obtained for that combination. Results were obtained from four normally hearing subjects.

2.3 Results

As expected from previous work [27], [28], the measured thresholds varied systematically across noise samples for any specific subject. The standard deviation of the threshold across noise samples for a given subject was typically about 3 dB. The thresholds for a given noise sample also varied somewhat across subjects, reflecting the fact that some listeners are more efficient than others [3]. For a given noise sample the standard deviation of the threshold estimates across subjects was typically about 1.2 dB. The overall mean threshold was 46.0 dB SPL for the white noise background and 50.4 dB for the pink noise background. To compare our results to those in the literature, we calculated the level per ERB_N for each noise; this was 51 dB SPL for the white noise and 55 dB SPL for the pink noise. Thus the ratio of signal level to noise level per ERB_N is about −4 to −5 dB. This corresponds reasonably well to previously reported thresholds for 1-kHz tones in noise [4], [31], [32].

2.4 Determination of Average STPL at Threshold

The stimuli used in the experiment were used as input to the model. The FIR filter used in the first stage of the model was designed to take into account the frequency response of the headphone, based on measurements using a KEMAR acoustic manikin [33], averaging across results for the "large" and "small" ears supplied with KEMAR. We examined the output of the model (STPL versus time) for all signal/background combinations used in the experiment, that is, for all different samples of frozen noise used. The background level was always 70 dB SPL, the level used in the experiment. The signal level was set to the mean threshold value determined for each type of background (white noise or pink noise). The average value of the STPL was 0.014 sone for both pink and white noise backgrounds. Hence this value was taken as the STPL required for threshold ($d' = 1.16$).

In summary, under conditions where informational masking was probably minimal, detection thresholds for a sinusoidal signal in pink and white noise corresponded to a mean value for the STPL of 0.014 sone. In experiment 2 we assessed how well a model based on this value would predict detection thresholds for complex time-varying signals in a variety of background sounds.

3 DETECTION THRESHOLDS FOR TIME-VARYING SIGNALS IN BACKGROUND SOUNDS

Behavioral masked thresholds were measured for six different signal types in nine different background sounds (54 combinations). The backgrounds were chosen to be relatively steady, and not to contain very distinct musical portions, so as to minimize informational masking. However, we anticipated that informational masking might influence the results to some extent, especially when the background contained features (such as weak musical notes) that were also present in the signal.

3.1 Method

The first five of the signals were extracted from waveform files supplied by Nokia. These files contained signals of the type used for ring tones in mobile telephones; some were monophonic (containing one tone at a time) and others were polyphonic (containing multiple tones at any one time). The sixth signal was a 1000-Hz sinusoidal tone. The signals were identified by the following names:

s1 Birdsong
s2 Snowfall
s3 Running late
s4 Ring ring
s5 Kuusnepa
s6 1000-Hz tone

The maskers (backgrounds) were recorded in a variety of everyday situations (except for the pink and white noise) and were:

b1 Arcade
b2 Car
b3 Compressor
b4 Pub
b5 Supermarket
b6 Traffic
b7 Train
b8 Pink noise
b9 White noise

Only a single sample of pink noise and of white noise was used. For each signal and each background sound a 1-s segment was extracted from the longer waveform files, and these 1-s segments were used in the experiment. The masked threshold for each signal type was measured in the presence of each background type, giving 54 combinations of signal and background. Raised-cosine ramps of 50-ms duration were applied to both the signal and the background. For each signal/background combination the waveforms used were fixed (frozen). Exactly the same waveforms were used as input to the model when evaluating the predictions of the model.

The signal and background were generated and presented in the same way as for experiment 1. The procedure was also the same as for experiment 1. Results were obtained from four normally hearing subjects.

3.2 Results

For the majority of the combinations of signal and background, the standard deviation (SD) of the threshold estimates across subjects was 2 dB or less, giving a standard error (SE) of 1 dB or less. However, there were a few cases where there were relatively large differences across subjects. It appeared that one subject sometimes "found a cue" that the other subjects did not find. Cases where individual variability was large (SD > 4 dB) are shown in Table 1. It is noteworthy that all of these cases occurred when the background sound was either pink noise or white noise. Possibly, the stable nature of these backgrounds, with their smooth variation in spectrum level across frequency, made it easier to pick out spectral features of the signal in specific frequency regions, but subjects varied in which features of the signal were most salient for them. In what follows we focus on the mean thresholds across subjects.

3.3 Comparison of Obtained and Predicted Thresholds

Thresholds for each signal/background combination were predicted by adjusting the level of the signal at the input to the model until the mean STPL was equal to 0.014 sone. Table 2 gives the measured masked thresholds and compares them with the thresholds predicted by the model. Columns 1 and 2 show the signal/background combination. Column 3 shows the mean measured masked threshold in dB SPL. Column 4 shows the predicted thresholds. The final column shows the predicted threshold minus the obtained threshold. The values in the final column are generally small, indicating reasonably accurate predictions. However, there are more negative than positive differences, indicating that the thresholds obtained tend to be higher than predicted. This is consistent with the idea that the results were influenced to some extent by informational masking. There are only a few cases where the differences were positive by more than 3 dB. The largest positive differences occurred for the 1000-Hz signal in pink noise backgrounds. This happened because the specified samples of white and pink noise that were (randomly) selected for use in this experiment led to unusually low obtained thresholds (at the lower end of the range measured in experiment 1).

Table 1. Combinations of signal and background from experiment 1 for which individual differences were large (SD > 4 dB).

Signal	Background	SD (dB)
Running late	Pink noise	4.8
Running late	White noise	6.5
Ring ring	Pink noise	6.1
Kuusnepa	Pink noise	6.6
Kuusnepa	White noise	8.8

The largest negative deviation of 8.2 dB occurred for signal 1 in background 4. The predicted thresholds for this signal are lower than the thresholds obtained for all backgrounds except "car." For the remaining signals there is no consistent trend for the deviations to be negative or positive. The predictions were most accurate for the most steady backgrounds (car, compressor, train, pink noise, and white noise). Thresholds tended to be higher than predicted in the backgrounds arcade and pub, probably reflecting effects of informational masking for those backgrounds.

Table 2. Comparison of thresholds.

Signal	Background	Obtained Threshold (dB SPL)	Predicted Threshold (dB SPL)	Difference (dB)
Birdsong	Arcade	48.1	45.0	–3.1
	Car	27.3	27.4	0.1
	Compressor	35.9	35.4	–0.5
	Pub	49.0	40.8	–8.2
	Supermarket	32.9	30.0	–2.9
	Traffic	42.7	41.2	–1.5
	Train	35.1	33.8	–1.3
	Pink noise	52.7	49.1	–3.6
	White noise	50.0	46.0	–4.0
Snowfall	Arcade	41.9	42.1	0.2
	Car	27.6	27.0	–0.6
	Compressor	34.5	35.4	0.9
	Pub	43.9	40.3	–3.6
	Supermarket	27.2	30.2	3.0
	Traffic	42.4	40.4	–2.0
	Train	33.1	33.3	0.2
	Pink noise	46.3	45.8	–0.5
	White noise	41.0	44.0	3.0
Running late	Arcade	52.6	48.0	–4.6
	Car	30.6	31.0	0.4
	Compressor	38.6	39.4	0.8
	Pub	52.0	44.8	–7.2
	Supermarket	36.9	34.3	–2.6
	Traffic	47.2	45.2	–2.0
	Train	36.4	37.9	1.4
	Pink noise	50.5	49.3	–1.2
	White noise	44.2	44.8	0.6
Ring ring	Arcade	42.1	39.3	–2.8
	Car	24.4	26.4	2.0
	Compressor	35.0	33.0	–2.0
	Pub	38.1	35.6	–2.5
	Supermarket	34.2	31.0	–3.2
	Traffic	39.3	36.4	–2.9
	Train	28.7	31.4	2.7
	Pink noise	52.4	51.0	–1.4
	White noise	54.3	52.1	–2.2
Kuusnepa	Arcade	49.3	49.2	–0.1
	Car	33.3	34.0	0.7
	Compressor	44.0	42.8	–1.2
	Pub	52.8	45.5	–7.3
	Supermarket	42.6	37.8	–4.8
	Traffic	46.9	45.6	–1.3
	Train	38.8	41.3	2.6
	Pink noise	48.6	46.3	–2.3
	White noise	41.1	42.7	1.6
1000-Hz tone	Arcade	49.5	47.1	–2.4
	Car	27.8	29.4	1.6
	Compressor	44.5	43.7	–0.8
	Pub	45.9	43.6	–2.3
	Supermarket	27.0	30.2	3.2
	Traffic	47.3	46.4	–0.9
	Train	39.5	41.0	1.5
	Pink noise	45.7	49.9	4.2
	White noise	39.9	45.6	5.7

Fig. 1 shows a scatter plot of measured thresholds versus predicted thresholds. The correlation between the two quantities is 0.94. The root-mean-square difference between the thresholds obtained and those predicted is 3 dB. The dashed line in Fig. 1 indicates where the points would lie if the measured thresholds equaled the predicted thresholds.

4 EXPERIMENT 3: PSYCHOMETRIC FUNCTIONS USING A 2AFC TASK

This experiment was conducted to check the generality of the model, and also to determine the relationship between detectability d' and the average value of the STPL. Determination of this relationship allows the prediction of signal levels that produce different degrees of detectability. For example, a d' value of 0.5 would mean that the detectability of the signal was poor and the signal would often be missed, whereas a d' value of 2 would mean that the signal was quite highly detectable. An intermediate value of d' (such as the value of 1.16 tracked in experiment 1, corresponding to threshold) would mean that the signal was moderately detectable. The analyses presented so far are all based on the assumption that when the average value of the STPL is 0.014 sone, the value of d' is equal to 1.16.

In order to determine the relationship between detectability d' and the average value of the STPL we measured psychometric functions (percent correct versus signal level) for five signals in five different backgrounds, using a 2AFC task. Each signal was paired with only one background, so there were five pairs in total. Unlike in experiment 1, some of the background sounds contained strong

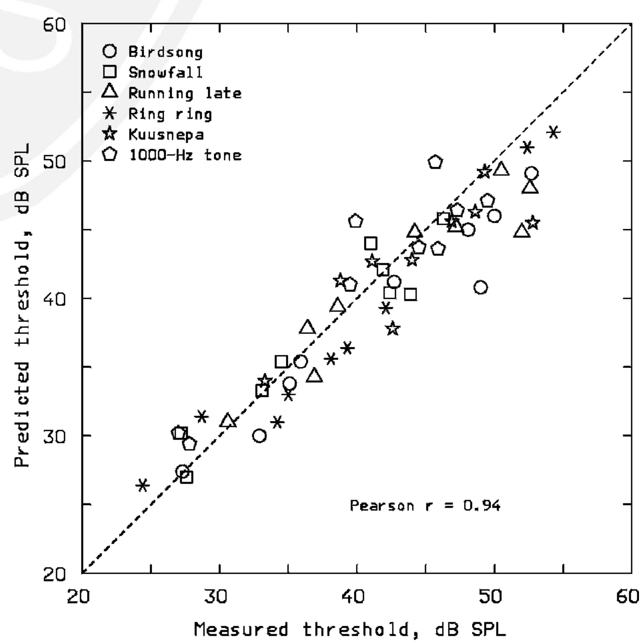

Fig. 1. Scatter plot of obtained versus predicted thresholds for experiment 2.

fluctuations and some contained music, so we expected that the results would be influenced to some extent by informational masking.

4.1 Stimuli

The signals were again of the type used for ring tones in mobile telephones. Four of the signals were also used in experiment 2. The five signal/background combinations were

1) Birdsong/arcade
2) Snowfall/band
3) Running late/café
4) Grande valse/car
5) Ring ring/bach.

The backgrounds arcade, band, and bach contained musical sounds, and the background café contained some clicking sounds. We anticipated that some informational masking might occur for these backgrounds. The timing of the stimuli was the same as for experiment 1, and the stimuli were generated and controlled in the same way as for experiment 1. Six subjects with normal hearing were tested.

4.2 Procedure

Psychometric functions were measured using a 2AFC task. Within a block of trials (a run) the combination of signal and background was fixed. The background was presented in two 1-s bursts, separated by 500 ms, and the signal was presented at random with either the first background burst or the second. The background level was fixed at 70 dB SPL throughout. The signal was initially set to a level at which detectability was high as determined in pilot experiments for each subject. Five trials at this easy level were presented to allow the subject to learn what to "listen for." Responses for these first five trials were not scored. For the remaining 50 trials in a run the signal level cycled repeatedly through five decreasing values, the highest of which was chosen to give high detectability and the lowest to give low detectability. Thus the subject received a "reminder" easily detected signal on every fifth trial to help them remember what to listen for. This helps maintain stable performance [34], [35]. Each run gave ten trials for each signal-to-background ratio. Each run (for each signal/background combination) was repeated ten times, giving 100 trials for each signal-to-background ratio.

4.3 Results

The percent correct scores were converted to the detectability index d' [36]. The psychometric functions for each signal/background combination are shown in Figs. 2–6. Within each figure each curve shows the results for one subject. The detectability index is plotted as a function of the signal-to-background ratio. In some cases the d' values were slightly negative when the signal-to-background ratio was low. In theory this could mean that a subject sys-

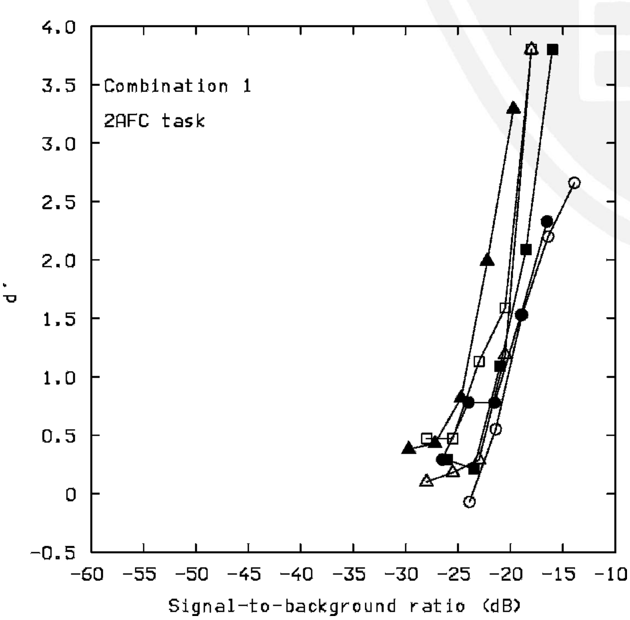

Fig. 2. Psychometric functions (d' versus signal-to-background ratio) for detection of birdsong signal in arcade background, as obtained in experiment 3 using a two-alternative forced-choice task. Each curve shows results for one subject.

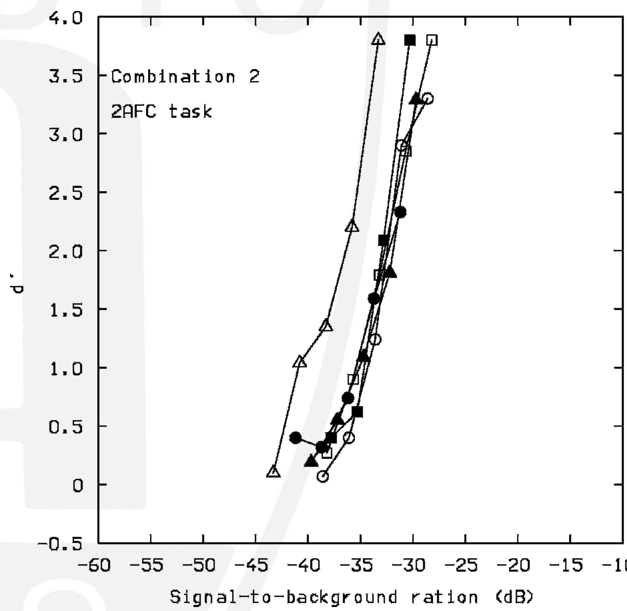

Fig. 3. As Fig. 2, but for snowfall signal in band background.

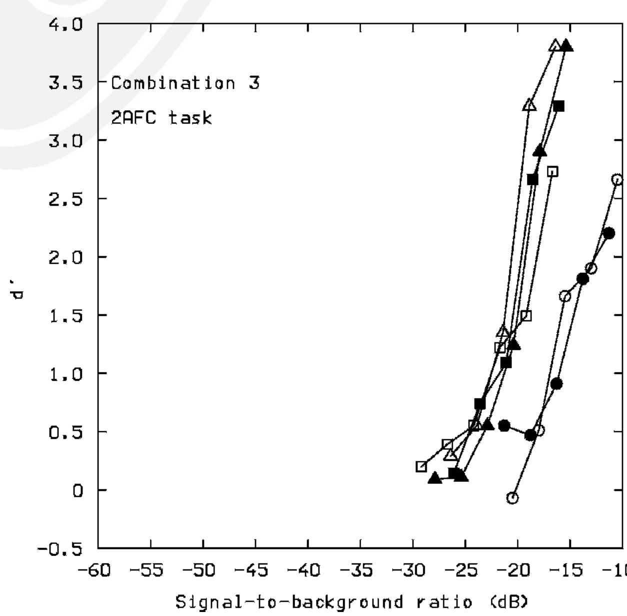

Fig. 4. As Fig. 2, but for running late signal in Bach background.

tematically gave the wrong answer (that is, identified the signal as being in the wrong interval). However, using the test described by Miller [37] none of the d' values was significantly below 0. It is therefore likely that for very low signal-to-background ratios subjects simply guessed, and negative d' values simply reflect random unlucky guesses.

For the first signal/background pair (Fig. 2) the psychometric functions were very similar across subjects, and d' increased from about 0.5 (corresponding to 63.7% correct) to about 3 (corresponding to 98.3% correct) as the signal/background ratio increased from −25 to −17 dB (an 8-dB range). For the second signal/background pair (Fig. 3) the psychometric functions were also similar across subjects, except that one subject performed somewhat better than the other five. For the majority of subjects d' increased from about 0.5 to about 3 as the signal/background ratio increased from −38 to −30 dB. Thus although the signal levels were lower than for the first signal/background pair, the slopes of the psychometric functions were similar. For the third signal/background pair (Fig. 4) the results varied more across subjects, and the signal-to-background ratios were generally higher than for pairs 1 and 2. However, for each subject a change in signal-to-background ratio of about 8 dB was required to increase d' from about 0.5 to about 3. For the fourth signal/background pair (Fig. 5) the signal-to-background ratios were generally lower than for the first three pairs, and the slopes of the functions varied somewhat across subjects. However, on average it was again the case that a change in signal-to-background ratio of about 8 dB was required to increase d' from about 0.5 to about 3. For the final signal/background pair (Fig. 6) the positions of the psychometric functions varied markedly across subjects. Evidently some subjects were more sensitive than others to certain cues. However, the slopes of the functions did not vary markedly across subjects.

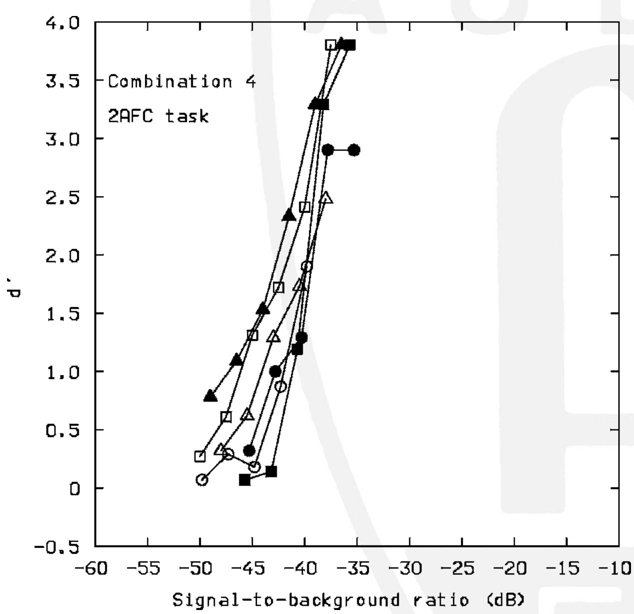

Fig. 5. As Fig. 2, but for grande valse signal in café background.

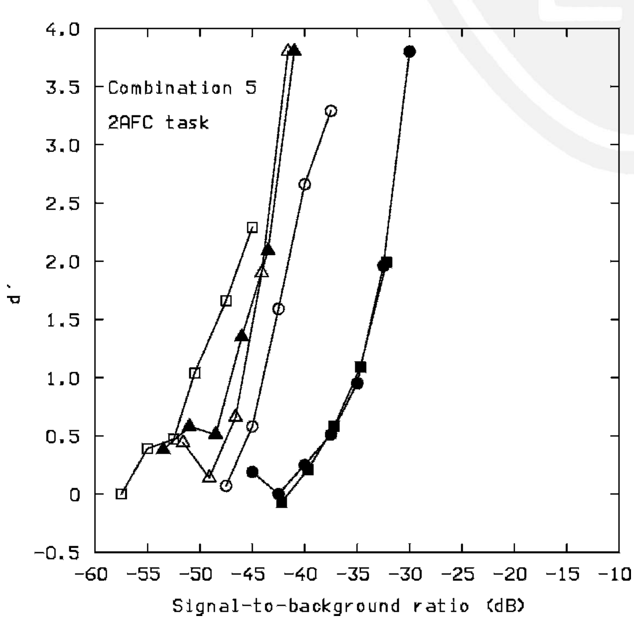

Fig. 6. As Fig. 2, but for ring ring signal in car background.

In summary, for some signal/background pairs the psychometric functions were very similar across subjects, whereas for other pairs there were large individual differences. There were also large differences across signal/background pairs in the signal-to-background ratio required to achieve a certain value of d'. However, the slopes of the psychometric functions were mostly similar across signal/background pairs and across subjects.

4.4 Relationship of d' to Mean Value of STPL

To determine the relationship between d' and the mean value of the STPL for each signal/background combination, we first determined the mean and the standard deviation across subjects of the signal-to-background ratio required to give specific values of d', namely, 0.5, 1.0, 1.5, and 2.0, using interpolation and, in a few cases, extrapolation. We denote these quantities by Mean(SBR) and SD(SBR). For each d' value we calculated the mean value of the STPL when the signal-to-background ratio was Mean(SBR) and when it was Mean(SBR) ± SD(SBR). Then for each combination of signal and background we plotted the value of d' against the mean values of the STPL determined in this way. This gave a series of functions relating d' to the mean value of STPL, one for each combination of signal and background, with corresponding error bars. The resulting functions are shown in Fig. 7.

The functions are plotted with the STPL on a logarithmic scale, as on this scale the psychometric functions for the different signal/background combinations were all approximately parallel (this was not the case when a linear scale was used). The solid symbols indicate the mean values of the STPL when the signal-to-background ratio was Mean(SBR). The error bars indicate the mean values of the STPL when the signal-to-background ratio was Mean(SBR) ± SD(SBR). The error bars are largest for combination 5 (squares), as expected from the large individual differences in the psychometric functions for that combination, as seen in Fig. 6.

If the predictions of the model were perfect, all of the curves would coincide. Clearly, they do not. Also the curves should all pass through the large open circle, which indicates a d' value of 1.16 for an STPL of 0.014. The curve for signal/background combination 2 lies close to the circle, but the curves for combinations 1 and 3 lie to the right of the circle (the signal is less detectable than predicted), whereas the curves for combinations 4 and 5 lie to the left of the circle (the signal is more detectable than predicted). The cases where the signal is less detectable than predicted may be explicable in terms of informational masking. The backgrounds for combinations 1, 2, and 3 all contained time-varying sounds with musical or clicklike segments. The better than predicted detectability of the signal for combination 4 may be related to the fact that, according to the model, the greatest contribution to the STPL came from high-frequency components in the signal, centered around 7–8 kHz. At such high frequencies the response of the Sennheiser HD580 earphone on real ears shows very large peaks and dips, but the exact position of these varies across ears, depending on the geometry of the outer ear. These peaks and dips were not present in the FIR filter that was used to simulate the response of the earphone at the input to the model. The FIR filter response was based on an average across several ears, which results in a smoothing out of the peaks and dips. It seems likely that sharp peaks in the frequency response of the earphone were present for the individual ears in the frequency range of 7–8 kHz, and these led to better detectability of the signal than predicted by the model. The better than predicted detectability of the signal for combination 5 may be related to the fact that this signal had a distinct rhythmic structure. This may enhance detectability relative to signals where no rhythmic structure is present. It should be noted again that there were considerable individual differences for combination 5, with some subjects performing more or less as predicted by the model and some performing markedly better than predicted.

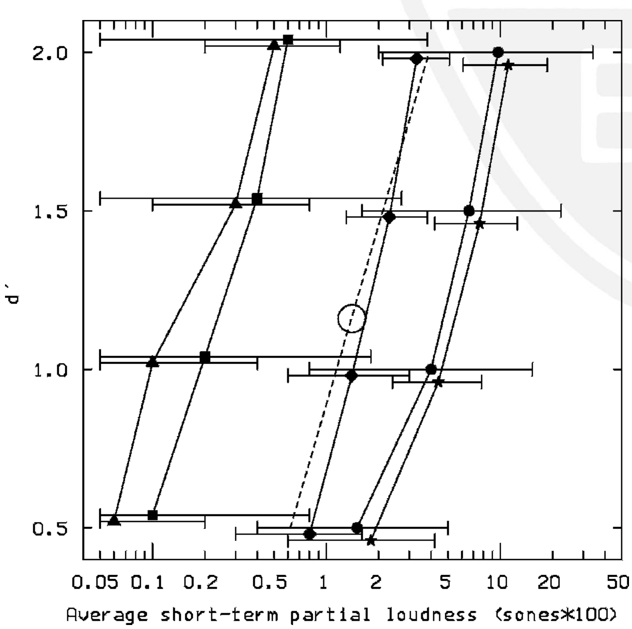

Fig. 7. Detectability index d' as a function of calculated average value of short-term partial loudness for five signal/background combinations used in experiment 3. Combinations 1–5 are indicated by stars, diamonds, circles, triangles, and squares, respectively. See text for explanation of large open circle and dashed line.

Despite the differences across signal/background combinations, the slopes of the functions in Fig. 7 are all similar. Thus detectability d' varies in a similar way with (log) STPL. The dashed line in Fig. 7 has a slope equal to the average slope of the five curves and passes through the circle (corresponding to STPL $= 0.014$, $d' = 1.16$). The dashed line can be taken as indicating how detectability typically varies with STPL.

5 EXPERIMENT 4: PSYCHOMETRIC FUNCTIONS USING A YES–NO TASK

Experiment 3 was conducted using a 2AFC task. This task has the advantage that it is unaffected by the response criterion of the subject; the subject simply picks the interval that is most likely to contain the signal. However, in everyday life signals are not usually detected by comparing two observation intervals. Rather, the listener has to decide whether or not a signal is present at a specific time. Also, in everyday life the listener may be uncertain as to what signal may occur at any particular time (for example, their own telephone ringing versus someone else's telephone). To simulate this situation, we repeated experiment 3, but using a single-interval yes–no task rather than a 2AFC task, and with the different signal/masker combinations randomly interleaved.

5.1 Method

The subjects and signal/background combinations were the same as for experiment 3. However, on each trial only a single stimulus was presented. The task of the subject was to indicate whether or not a signal was present on that trial. For each signal/background combination the signal level for a given trial was chosen from seven levels, spaced by 2.5 dB, with the third level from the bottom chosen to correspond to the threshold ($d' = 1.16$) determined in pilot trials. For every seven trials containing a signal there were two trials without any signal (catch trials). A run started with ten practice trials, which were not scored. There were five trials containing easily detectable signals (15 dB above the estimated threshold level), one for each signal/background combination, and five catch trials. Following this, there were 45 trials that were scored, one for each signal/background combination and each signal level (35 total) plus 10 catch trials. The order of the stimuli for these 45 trials was random. Runs were repeated at least 100 times for each subject (with a different random order of the stimuli for each run) to give at least 100 trials for each signal/background combination and each signal level.

For each combination and signal level the proportion of hits (correct yes responses) was determined. The proportion of false alarms was calculated as the proportion of yes responses on trials where no signal was present. The detectability index d' was calculated from the proportion of hits and false alarms using standard methods [30].

5.2 Results

The psychometric functions for each signal/background combination are shown in Figs. 8–12. The format of the figures is the same as for Figs. 2–6. In general the functions are very similar to those obtained using the 2AFC task. However, there was a trend for the psychometric functions to be shifted slightly to the right for the single-interval yes–no task. In other words, the signal-to-background ratio had to be slightly higher in the yes–no task than in the 2AFC task to achieve the same level of detectability. This is presumably an effect of the greater level of stimulus uncertainty in the yes–no task, reflecting the fact that the different signal/background combinations were interleaved in random order. However, the effect of stimulus uncertainty is relatively small. It is noteworthy that for five of the six subjects, the proportion of false alarms was higher for combination 3 than for any other combination. This may have happened because the signal sounded rather similar to the background for that combination.

Functions relating d' to the mean value of the STPL were determined in the same way as for the 2AFC task. The resulting functions are shown in Fig. 13. They are very similar to the functions in Fig. 7, but tend to be shifted slightly to the right, reflecting the greater difficulty of the yes–no task. The dashed line can be taken as indicating how detectability typically varies with STPL.

The dashed lines in Figs. 7 and 13 are very similar in slope. Taking the mean slope of the two as representative, the relationship of d' to the average value of the STPL is

$$d' = 4.87 + 2 \log_{10} (\text{average STPL}). \qquad (5)$$

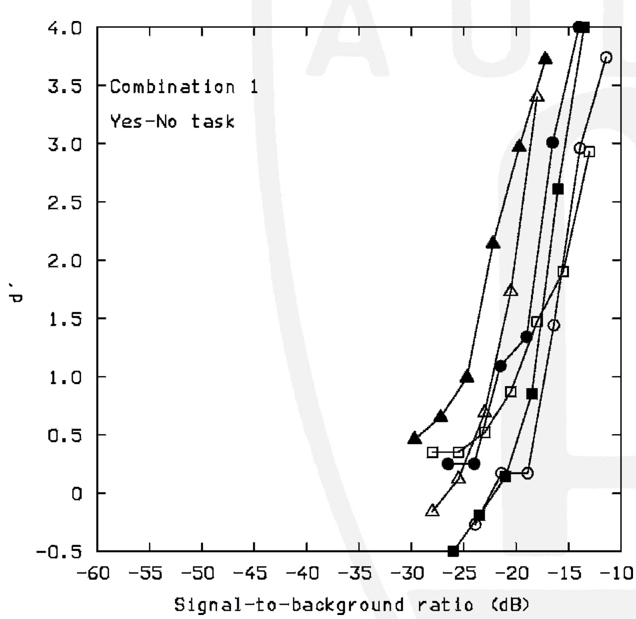

Fig. 8. As Fig. 2, but for data obtained in experiment 4 using a yes–no task.

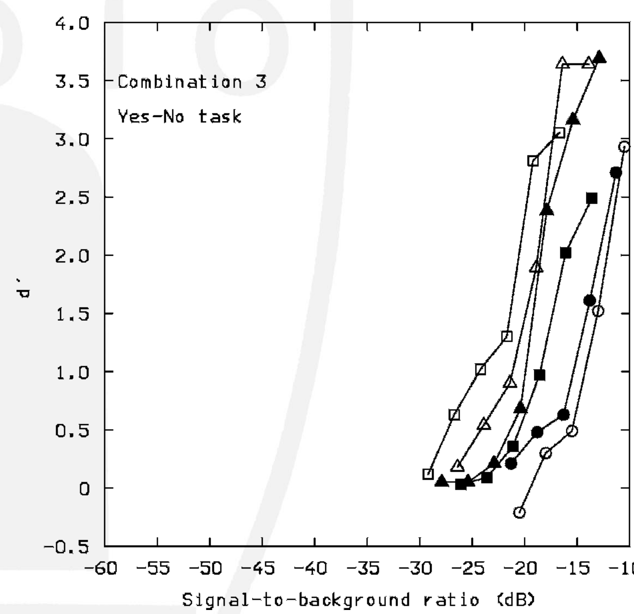

Fig. 10. As Fig. 8, but for running late signal in Bach background.

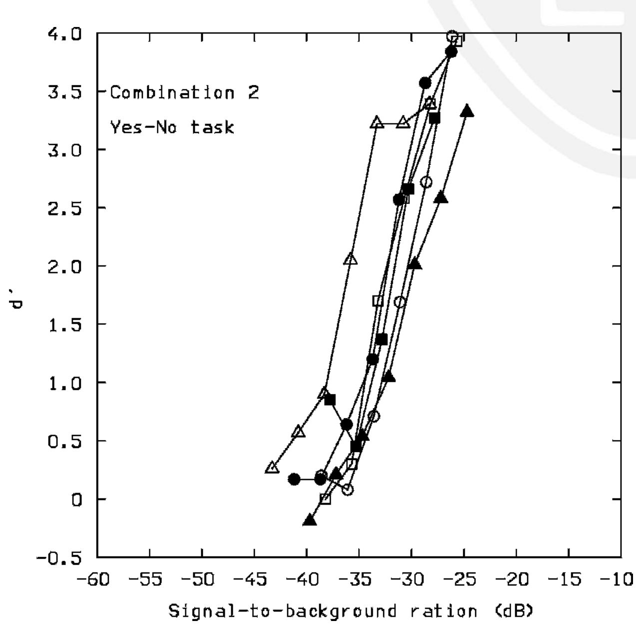

Fig. 9. As Fig. 8, but for snowfall signal in band background.

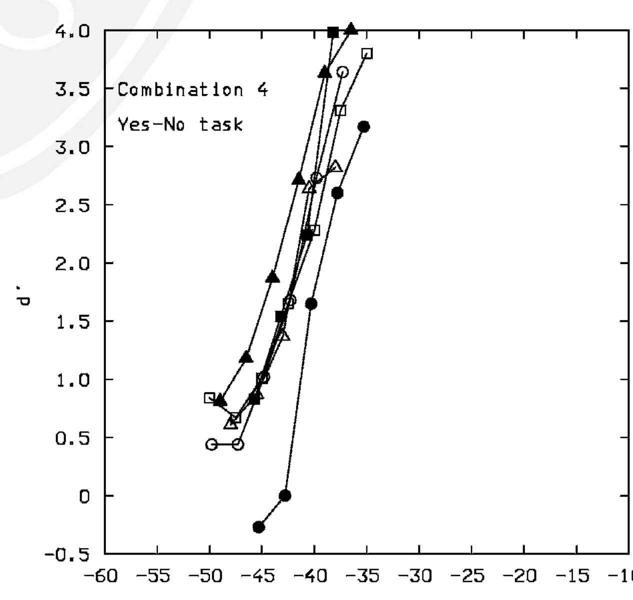

Fig. 11. As Fig. 8, but for "grande valse" signal in café background.

Using this equation, the STPL values generated by the model can be used to estimate the detectability of a given signal in a given background.

6 LIMITATIONS OF MODEL

The model does not contain a realistic front-end auditory filter bank. Under some conditions the phases of the components in a masker can have a considerable effect on the audibility of a signal [38], [39]. The model does not predict these effects correctly since it is based on the short-term power spectrum, and therefore does not explicitly take phase into account. However, the effects of phase for real-life background sounds are typically small. The model also does not take into account binaural effects in masking.

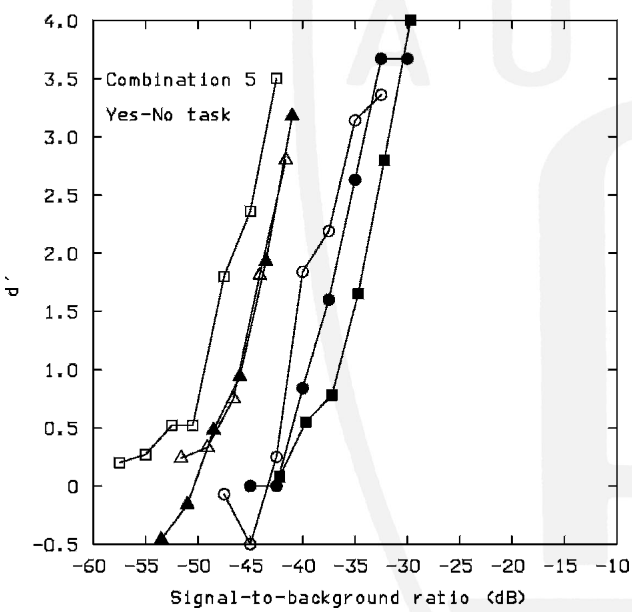

Fig. 12. As Fig. 8, but for ring ring signal in car background.

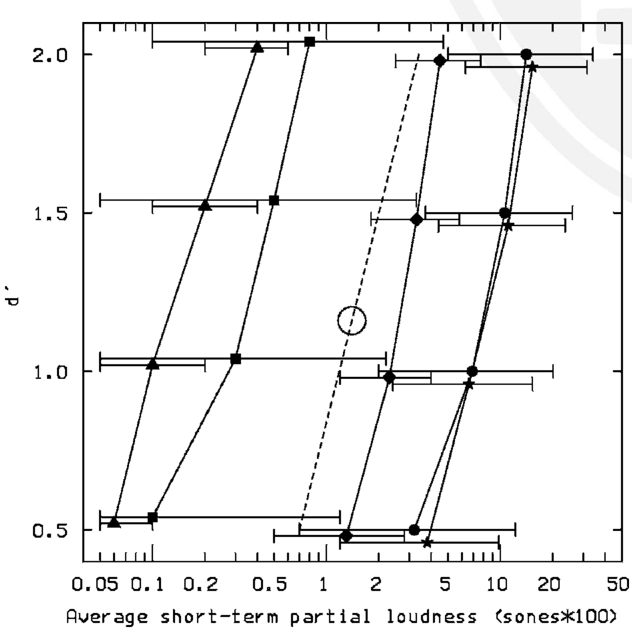

Fig. 13. As Fig. 7, but for data obtained in experiment 4 using a yes–no task.

When the interaural phase and/or level of a signal differ from those of a masker, the threshold of the signal is usually lower than when the signal and the masker have the same interaural level and phase [40]. Finally the model does not take into account the fact that the audibility of a signal may be improved when the masker contains amplitude fluctuations that are correlated in different frequency regions (comodulation masking release) [41], [42].

7 IMPLEMENTATION OF MODEL

The model requires specification of the conditions of presentation. These affect the FIR filter used to simulate the effect of transmission through the outer and middle ear. The options at present are free-field (frontal incidence), diffuse-field, or flat response at the eardrum. The last option is appropriate for earphones designed to have a flat response at the eardrum (such as Etymõtic Research ER2 insert earphones). For earphones designed to have a diffuse-field response (such as Etymõtic Research ER4 or ER6, Sennheiser HD414 and HD580, and many others) the diffuse-field option can be used as an approximation, although it is better to estimate the true response at the eardrum. For presentation via loudspeakers the transfer function from the loudspeakers to each eardrum needs to be taken into account and combined with the assumed transfer function through the middle ear.

The model also requires specification of a reference level. For example, it can be "told" that a full-scale 16-bit sinusoid corresponds to a free-field level (that is, the level measured at the point corresponding to the center of the listener's head, the listener having been removed from the sound field) of, say, 100 dB.

8 CONCLUSIONS

We have described a model that can be used to predict the audibility of complex time-varying signals in complex background sounds. The model is based on calculation of the average value of the short-term partial loudness (STPL), and it combines features of two earlier loudness models [4], [12]. Based on measurements of detection thresholds for a 1000-Hz sinusoid in different samples of pink and white noise (experiment 1) it was determined that when a signal is at threshold, corresponding to a detectability index d' of 1.16, the average value of the STPL is 0.014. In experiment 2 the model was evaluated by measuring detection thresholds for nine signal types in six backgrounds (54 combinations), using a two-alternative forced-choice task. The backgrounds were chosen to be relatively steady (such as traffic noise). The correlation between the measured masked thresholds and those predicted by the model was 0.94. The root-mean-square difference between the thresholds obtained and those predicted was 3 dB. For some combinations of signal and background, the measured thresholds were markedly higher than predicted, perhaps reflecting the effects of informational masking.

In experiment 3 psychometric functions were measured for the detection of five signals in five backgrounds (five pairs) using a two-alternative forced-choice task. The signal/background combination was fixed within a block of trials. The results showed marked individual differences for some signal/background combinations, but the slopes of the psychometric functions (d' versus signal-to-background ratio) were similar across subjects and across combinations. Experiment 4 used the same signals and backgrounds, but psychometric functions were measured using a single-interval yes–no task and with the different combinations randomly interleaved within a block of trials. The results were similar to those of experiment 1; the increase in signal uncertainty resulted in only a small worsening in performance. The results of experiments 3 and 4 were used to construct functions relating signal detectability, d', to the average value of the STPL.

9 ACKNOWLEDGMENT

This work was supported by the Medical Research Council (UK) and by Nokia Research Center (Finland). The authors thank Nick Zacharov and Ville-Veikko Mattila for helpful discussions and suggestions.

10 REFERENCES

[1] R. D. Patterson, "Auditory Warning Sounds in the Work Environment," *Phil. Trans. R. Soc. Lond. B,* vol. 327, pp. 485–492 (1990).

[2] H. Fletcher, "Auditory Patterns," *Rev. Mod. Phys.,* vol. 12, pp. 47–65 (1940).

[3] R. D. Patterson and B. C. J. Moore, "Auditory Filters and Excitation Patterns as Representations of Frequency Resolution," in B. C. J. Moore, Ed., *Frequency Selectivity in Hearing* (Academic Press, London, 1986), pp. 123–177.

[4] B. C. J. Moore, B. R. Glasberg, and T. Baer, "A Model for the Prediction of Thresholds, Loudness, and Partial Loudness," *J. Audio Eng. Soc.,* vol. 45, pp. 224–240 (1997 Apr.).

[5] H. Fletcher and W. A. Munson, "Loudness, Its Definition, Measurement and Calculation," *J. Acoust. Soc. Am.,* vol. 5, pp. 82–108 (1933).

[6] W. A. Munson and M. B. Gardner, "Loudness Patterns—A New Approach," *J. Acoust. Soc. Am.,* vol. 22, pp. 177–190 (1950).

[7] S. S. Stevens, "Calculation of the Loudness of Complex Noise," *J. Acoust. Soc. Am.,* vol. 28, pp. 807–832 (1956).

[8] S. S. Stevens, "Procedure for Calculating Loudness: Mark VI," *J. Acoust. Soc. Am.,* vol. 33, pp. 1577–1583 (1961).

[9] E. Zwicker, "Über psychologische und methodische Grundlagen der Lautheit," *Acustica,* vol. 8, pp. 237–258 (1958).

[10] E. Zwicker and B. Scharf, "A Model of Loudness Summation," *Psych. Rev.,* vol. 72, pp. 3–26 (1965).

[11] ANSI, S3.4-2005, "Procedure for the Computation of Loudness of Steady Sounds," American National Standards Institute, New York (2005).

[12] B. R. Glasberg and B. C. J. Moore, "A Model of Loudness Applicable to Time-Varying Sounds," *J. Audio Eng. Soc.,* vol. 50, pp. 331–342 (2002 May).

[13] M. F. Spiegel, M. C. Picardi, and D. M. Green, "Signal and Masker Uncertainty in Intensity Discrimination," *J. Acoust. Soc. Am.,* vol. 70, pp. 1015–1019 (1981).

[14] M. F. Spiegel and D. M. Green, "Signal and Masker Uncertainty with Noise Maskers of Varying Duration, Bandwidth, and Centre Frequency," *J. Acoust. Soc. Am.,* vol. 71, pp. 1204–1210 (1982).

[15] D. L. Neff and B. P. Callaghan, "Simultaneous Masking by Small Numbers of Sinusoids under Conditions of Uncertainty," in W. A. Yost and C. S. Watson, Ed., *Auditory Processing of Complex Sounds* (Lawrence Erlbaum Assoc., Hillsdale, NJ, 1987).

[16] D. L. Neff and D. M. Green, "Masking Produced by Spectral Uncertainty with Multi- Component Maskers," *Percept. Psychophys.,* vol. 41, pp. 409–415 (1987).

[17] D. L. Neff and B. P. Callaghan, "Effective Properties of Multicomponent Simultaneous Maskers under Conditions of Uncertainty," *J. Acoust. Soc. Am.,* vol. 83, pp. 1833–1838 (1988).

[18] D. L. Neff, T. M. Dethlefs, and W. Jesteadt, "Informational Masking for Multicomponent Maskers with Spectral Gaps," *J. Acoust. Soc. Am.,* vol. 94, pp. 3112–3126 (1993).

[19] B. C. J. Moore, "Distribution of Auditory-Filter Bandwidths at 2 kHz in Young Normal Listeners," *J. Acoust. Soc. Am.,* vol. 81, pp. 1633–1635 (1987).

[20] E. A. G. Shaw, "Transformation of Sound Pressure Level from the Free Field to the Eardrum in the Horizontal Plane," *J. Acoust. Soc. Am.,* vol. 56, pp. 1848–1861 (1974).

[21] B. R. Glasberg and B. C. J. Moore, "Derivation of Auditory Filter Shapes from Notched-Noise Data," *Hear. Res.,* vol. 47, pp. 103–138 (1990).

[22] B. C. J. Moore and B. R. Glasberg, "Formulae Describing Frequency Selectivity as a Function of Frequency and Level and Their Use in Calculating Excitation Patterns," *Hear. Res.,* vol. 28, pp. 209–225 (1987).

[23] B. C. J. Moore and B. R. Glasberg, "Suggested Formulae for Calculating Auditory-Filter Bandwidths and Excitation Patterns," *J. Acoust. Soc. Am.,* vol. 74, pp. 750–753 (1983).

[24] A. Kohlrausch, R. Fassel, and T. Dau, "The Influence of Carrier Level and Frequency on Modulation and Beat-Detection Thresholds for Sinusoidal Carriers," *J. Acoust. Soc. Am.,* vol. 108, pp. 723–734 (2000).

[25] B. C. J. Moore and B. R. Glasberg, "Temporal Modulation Transfer Functions Obtained Using Sinusoidal Carriers with Normally Hearing and Hearing-Impaired Listeners," *J. Acoust. Soc. Am.,* vol. 110, pp. 1067–1073 (2001).

[26] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Krieger, New York, 1974).

[27] A. Ahumada and J. Lovell, "Stimulus Features in Signal Detection," *J. Acoust. Soc. Am.,* vol. 49, pp. 1751–1756 (1971).

[28] T. E. Hanna and D. E. Robinson, "Phase Effects for a Sine Wave Masked by Reproducible Noise," *J. Acoust. Soc. Am.,* vol. 77, pp. 1129–1140 (1985).

[29] H. Levitt, "Transformed Up–Down Methods in Psychoacoustics," *J. Acoust. Soc. Am.,* vol. 49, pp. 467–477 (1971).

[30] N. A. Macmillan and C. D. Creelman, *Detection Theory: A User's Guide* (Cambridge University Press, Cambridge, UK, 1991).

[31] B. Scharf, "Critical Bands," in J. V. Tobias, Ed., *Foundations of Modern Auditory Theory* (Academic Press, New York, 1970), pp. 157–202.

[32] S. Fidell, R. Horonjeff, S. Teffeteller, and D. M. Green, "Effective Masking Bandwidths at Low Frequencies," *J. Acoust. Soc. Am.,* vol. 73, pp. 628–638 (1983).

[33] M. D. Burkhard and R. M. Sachs, "Anthropometric Manikin for Acoustic Research," *J. Acoust. Soc. Am.,* vol. 58, pp. 214–222 (1975).

[34] M. M. Taylor, S. M. Forbes, and C. D. Creelman, "Pest Reduces Bias in Forced-Choice Psychophysics," *J. Acoust. Soc. Am.,* vol. 74, pp. 1367–1374 (1983).

[35] B. C. J. Moore and A. Sek, "Detection of Combined Frequency and Amplitude Modulation," *J. Acoust. Soc. Am.,* vol. 92, pp. 3119–3131 (1992).

[36] M. J. Hacker and R. Ratcliff, "A Revised Table of d' for M-Alternative Forced Choice," *Percept. Psychophys.,* vol. 26, pp. 168–170 (1979).

[37] J. Miller, "The Sampling Distribution of d'," *Percept. Psychophys.,* vol. 58, pp. 65–72 (1996).

[38] B. K. Smith, U. K. Sieben, A. Kohlrausch, and M. R. Schroeder, "Phase Effects in Masking Related to Dispersion in the Inner Ear," *J. Acoust. Soc. Am.,* vol. 80, pp. 1631–1637 (1986).

[39] B. C. J. Moore, "Interference Effects and Phase Sensitivity in Hearing," *Phil. Trans. R. Soc. Lond. A,* vol. 360, pp. 833–858 (2002).

[40] I. J. Hirsh, "Binaural Summation and Interaural Inhibition as a Function of the Level of the Masking Noise," *J. Acoust. Soc. Am.,* vol. 56, pp. 205–213 (1948).

[41] J. W. Hall, M. P. Haggard, and M. A. Fernandes, "Detection in Noise by Spectro-Temporal Pattern Analysis," *J. Acoust. Soc. Am.,* vol. 76, pp. 50–56 (1984).

[42] B. C. J. Moore, "Across-Channel Processes in Auditory Masking," *J. Acoust. Soc. Jpn. (E),* vol. 13, pp. 25–37 (1992).

THE AUTHORS

B. R. Glasberg

Brian R. Glasberg received B.Sc. and Ph.D. degrees in applied chemistry from Salford University in 1968 and 1972, respectively.

He then worked as a chemist refining precious metals. He spent some time as a researcher of process control before joining the laboratory of Brian C. J. Moore at the University of Cambridge, UK, initially as a research associate and then as a senior research associate.

Dr. Glasberg's research focuses on the perception of sound in both normally hearing and hearing-impaired people. He also works on the development and evaluation of hearing aids, especially digital hearing aids. He is a member of the Acoustical Society of America and has published 107 research papers and book chapters.

The biography of Brian C. J. Moore was published in the 2005 July/August issue of the *Journal.*