AN AUDITORY MODEL THAT EXPLAINS INDIVIDUAL DIFFERENCES OF MASKING BY SCHROEDER-PHASE COMPLEXES

PACS: 43.66.Ba, 43.66.Dc, 43.66.Mk, 43.66.Nm

NISHIMURA, Akira

Department of Media and Cultural Studies, Faculty of Informatics,
Tokyo University of Information Sciences
1200-2, Yatoh-cho, Wakaba-ku, Chiba-city, Chiba 265-8501, Japan
Tel: +81-43-236-4658
Fax: +81-43-236-4658
E-mail: akira@rsch.tuis.ac.jp

ABSTRACT

A masking model that can quantitatively account for the results of a masking experiment using positive and negative Schroeder-phase complex as a masker is introduced. The model consists of a cascade constructed of a fixed filter for outer/middle ear response, a compressive gammachirp filter bank for auditory filters, a half-wave rectifier, a direct current component adder, a leaky integrator, and a detection module. After fitting the model parameters using individual masking period patterns and notched-noise masking data, the masking model having the compressive gammachirp filter accounts well for both types of human masking data.

INTRODUCTION

Phase effects in masking experiments using harmonic complex maskers are usually associated with strong variations in the envelope of the masker. Representative examples are positive and negative Schroeder-phase complexes. A Schroeder-phase complex having a fundamental frequency f_0 in Hz is given by:

$$S(t) = \sum_{m=n_l}^{n_h} \sin(2\pi m f_0 t + sign \times phase(m)),$$

$$phase(m) = \pi m(m-1)/(n_h - n_l + 1),$$

where n_l is the lowest harmonic number and n_h is the highest harmonic number. When the $sign$ is positive, $S(t)$ represents a positive Schroeder phase complex (POS), and, when the $sign$ is negative, $S(t)$ represents a negative Schroeder phase complex (NEG). POS and NEG are time-reversed and inverted versions of each other. Both phase relationships produce very flat envelopes. However, the masked thresholds of a pure tone of frequency and phase identical to a component of the Schroeder-phase masker vary by approximately 20 dB between POS and NEG maskers [1,2].

Several studies have reported that the phase characteristics of the auditory filter create these masking differences [1,2,3,4,5]. Negative curvature of the phase characteristics of the auditory filter, that is, upward frequency glide in the impulse response of the auditory filter, has been reported to result in pronounced peaks and dips in the envelope of the POS waveform. In addition, this results in the envelope of the NEG waveform being flat.

Oxenham and Dau have investigated several types of auditory filter, including the analytic gammachirp filter, as a means to account for both the dependence of masked thresholds on the phase of the Schroeder-phase complex and the frequency selectivity of these auditory filters [6]. In the simulations using the masking model proposed by Dau et al. [7], none of the auditory filter could quantitatively account for both masking data and frequency selectivity.

The present author has proposed a masking model that adopts the analytic gammachirp filter

[8] as an auditory filter bank. The model was able to simulate the difference in masked threshold between POS and NEG obtained in previous studies [1, 2, 4]. However, a large negative value is required for c, which is a parameter of the analytic gammachirp filter, in order to obtain the appropriate negative curvature of the phase function, which causes the amplitude spectrum of the filter to have a steeper slope at high-frequency regions and a shallower slope at low frequencies compared to the shape of the auditory filter measured by the notched-noise method.

In the present study, the compressive gammachirp filter [10] is replaced by an analytic gammachirp filter as an auditory filter in the model previously proposed by the author in order to provide good agreement for both the period masking data of the Schroeder-phase masker and frequency selectivity.

MASKING EXPERIMENTS

The notched-noise method is often used to measure peripheral frequency selectivity, and masking period patterns of a short tone using a POS or NEG masker are considered to be a representation of an asymmetric phase function of the auditory filtering. The masked threshold obtained from both masking experiments should be accounted for using an auditory filter that appropriately represents the mechanism of the auditory periphery. In addition, individual differences in masking data may be accounted for by the difference in the peripheral characteristics of the auditory system, i.e. the difference in the parameter values of the appropriate auditory model.

In order to investigate these relationships, period masking data and notched-noise data were collected from the same subject. The parameters of the masking model were then fitted using these masking data.

Procedure

All stimuli were generated digitally at a sampling rate of 24 kHz and played out with 16-bit resolution using a Sound Blaster Gold 64 sound card that connected to an ISA bus of a PC/AT compatible computer. They were presented to the left earpiece of a Sennheiser HDA200 headset, whose frequency response varied within 6 dB between 200 and 2000 Hz (B&K 4153 IEC coupler with B&K 4134). The electrical signal at the input to the headphones was not corrected.

Four listeners, each with absolute threshold within 15dB, listened to the stimuli in an anechoic chamber. Thresholds were estimated using a two-down one-up adaptive procedure, with feedback being provided at the end of each two interval forced choice trial. In notched-noise experiment, the signal level was fixed and notched masker level was increased and decreased. The step size of the level increment and decrement is 4 dB at the first 3 turnpoints and 2 dB after the third turnpoint. Each run terminated after 12 turnpoints, and the estimate for that run obtained by averaging the levels at the last 8 turnpoints. All data presented here were obtained from the mean of at least six runs on separate days.

Experiment 1: Masking Period Patterns Of Schroeder-phase Complexes

Masking period patterns were measured for two maskers, POS and NEG, each consisting of equal-amplitude harmonics 2-20 of a 100-Hz fundamental. Each masking period pattern was measured at a masker level of 55 dB SPL/component. In all cases the masker duration was 250 ms, including two 20 ms raised-cosine ramps, and the signal was a 5-ms 1100-Hz sinusoid composed of two 2.5-ms raised-cosine ramps (no steady state) which was added in-phase to the 1100-Hz masker component. The signal was turned on 152, 154, 156, 158, and 160 ms after the masker (zero-voltage points) was started. These conditions were identical to those of the first experiment conducted by Carlyon and Datta [4], with the exception that they used masker levels of 39 to 69 dB SPL/component and a masker duration of 400 ms.

The results of experiment 1 are shown in Fig. 2 for each subject using data obtained from predictions of the model described in section 3. All data points are expressed relative to the level of a single masker component. Masking period patterns obtained from each subject show similar tendencies to those observed in the results reported by Carlyon and Datta [4]. The masking period patterns for NEG have a minimum at Δt = 158 ms. The masking period patterns for POS show little variation over a period compared to the NEG masker.

Experiment 2: Manipulating The Phase Of Off-frequency Components

For the first condition, the central M components were selected from the positive-phase complex

of experiment 1, and the remaining components were selected from the negative-phase complex of experiment 1. For the second condition, the central M components were selected from the negative-phase complex and the remaining components were selected from the positive-phase complex, both of experiment 1. In both conditions, thresholds were obtained for the 5-ms 1100-Hz signal at Δt = 158 ms (the minimum of the positive-phase masking period pattern) as a function of M. M was selected from 5, 7, 9 and 11. Other parameters were the same as in experiment 1, and the experimental design was based on the third experiment of Carlyon and Datta [4].

Results are shown in Fig. 3 for each subject, along with the predictions of the model simulations described in section 3. All data points are expressed relative to the level of a single masker component. The experimental data also show similar tendencies to those reported by Carlyon and Datta [4]. The threshold is still higher when the 11 harmonics between 700 and 1500 Hz are in positive phase than when all 19 harmonics are in positive phase, indicating that components more than 300-Hz distant from the 1100-Hz signal affect its threshold.

Experiment 3: Masking By Notched-noise

Masked thresholds were determined for sinusoidal probe tones fp of 1100 Hz in the presence of notched-noise maskers having variable notch widths. The probe level was fixed to 50 dB SPL, and the noise level was varied to determine the thresholds. The notches were placed both symmetrically and asymmetrically about the probe frequency. The outside edges of the masker noise were fixed at $0.8fp$ (220 and 1980 Hz). The frequencies of the edges of the notch are specified in normalized frequency units relative to the probe frequency as given by $(|f - fp|)/fp$. In the symmetric conditions, both notch edges were placed at normalized values of 0, 0.1, 0.2, 0.3 and 0.4. In the asymmetric condition one of the notch edges was set at a normalized value of 0, 0.1, 0.2 and 0.3, whereas the other was set 0.2 normalized units further away (0.2, 0.3, 0.4 and 0.5).

The duration of the probe signal was 400 ms, including 10-ms raised-cosine ramps. The probe was temporally centered within the masker, which consisted of a 500-ms duration including 10-ms raised-cosine ramps. At the start of each threshold determination, a 2.73 s buffer of the notched noise was generated for use during this test. In each trial, a 500-ms portion of the buffer was chosen randomly for each of the two masker intervals within each trial.

Results are shown in Fig. 4 for each subject using data predicted by the power-spectrum model of masking described in the next section.

MASKING MODEL

The masking model previously proposed by the author [5] is revised to provide good agreement for both the period masking data of the Schroeder-phase masker and frequency selectivity. This section presents details of the model and shows how to fit the model parameters to account for masking data.

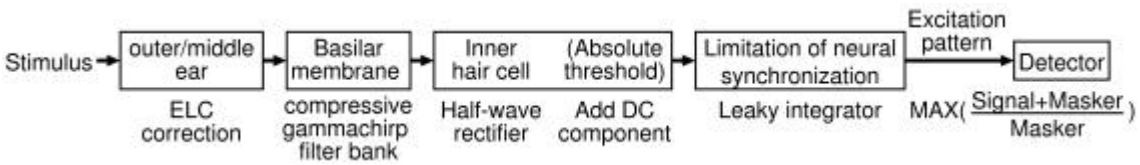

Figure 1: Schematic diagram of the masking model

Overview Of The Model

The proposed auditory model consists of a series of modules that each simulates specific functions of the auditory periphery. Figure 1 shows a schematic diagram of the model. The modules are as follows: a fixed filter for outer/middle ear response (ELC correction [9]), a compressive gammachirp filter bank for auditory filters [10], a half-wave rectifier, a direct current (DC) component adder, and a leaky integrator. A detector is connected to the output of the leaky integrator.

The compressive gammachirp filter in the present masking model is defined by five parameters, $b1$ and $c1$ in the analytic gammachirp section, and $b2$, $c2$ and $frat$ in the high-pass asymmetric function. Frequency dependent parameters were not adopted. Detail implementation of the compressive gammachirp filter in the time-domain is described by Irino and Patterson [10]. The main feature of this model is the phase response of the compressive gammachirp filters,

emphasizing peaks in the POS masker waveform and flattening the envelope of the NEG masker waveform. The half-wave rectifier represents a simple approximation of the sum of responses of the inner hair cells. The DC component adder is intended to simulate the absolute threshold of a pure tone corresponding to the center frequency of the auditory filter. The leaky integrator, having the time constant of Tc ms, simulates neural synchronization of the inner hair cells. At the detection stage, the absolute differences between the running output levels of the temporal integrator for each channel of the masker+signal waveforms and the masker alone (SMMR: Signal+Masker to Masker Ratio) are calculated. The maximum SMMR (MAXSMMR) is regarded as an index of detection of a short signal in the measurement of masking period patterns. The signal is considered to be detected in the frequency channel exhibiting the largest MAXSMMR, which simulates off-frequency listening. The center frequencies of the auditory filters were chosen in the range of $0.7fp$ and $1.3fp$.

The most remarkable feature of the present masking model for masking period pattern compared to Carlyon and Datta's model [4] is that the present model requires only a small number of parameters that can be estimated through psychoacoustic experiments, such as the parameters of a compressive gammachirp filter which can be obtained from notched-noise masking data [10].

Power-Spectrum Model

The shape of an auditory filter is generally derived using the power-spectrum model of masking. The power spectrum model assumes that threshold corresponds to a constant signal-to-masker ratio at the output of the auditory filter [9]. The compressive gammachirp filter is also adopted as an auditory filter in the power-spectrum model of masking.

Fitting The Masking Model To Masking Data

The down-simplex method was used to minimize the overall rms difference between the experimental data and the predictions of the model, that is, the difference between the masking data obtained in experiments 1 and 2 and the predictions of the present masking model, and also the difference between the masked thresholds obtained in experiment 3 and the predictions by the power-spectrum model of masking. The parameter values of the compressive gammachirp filter are identical between the proposed masking model and the power-spectrum model. All predicted values and experimental data were represented in dB. Fitting procedures were conducted for each individual subject.

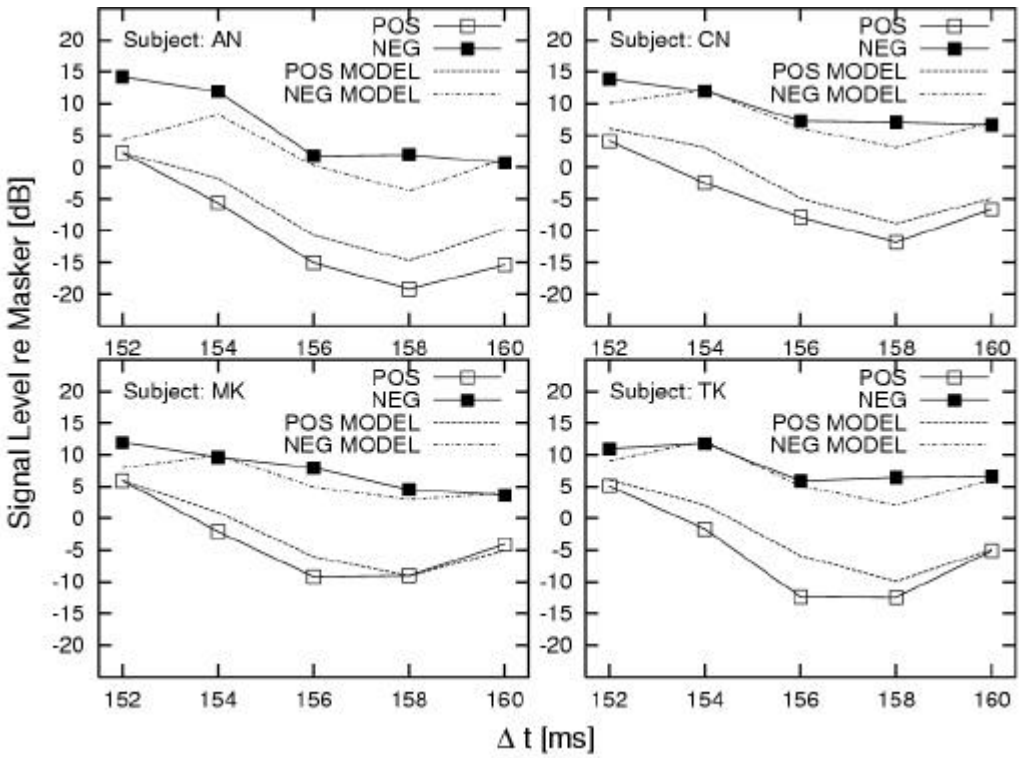

Figure 2: Results of experiment 1 and predictions by the masking model.

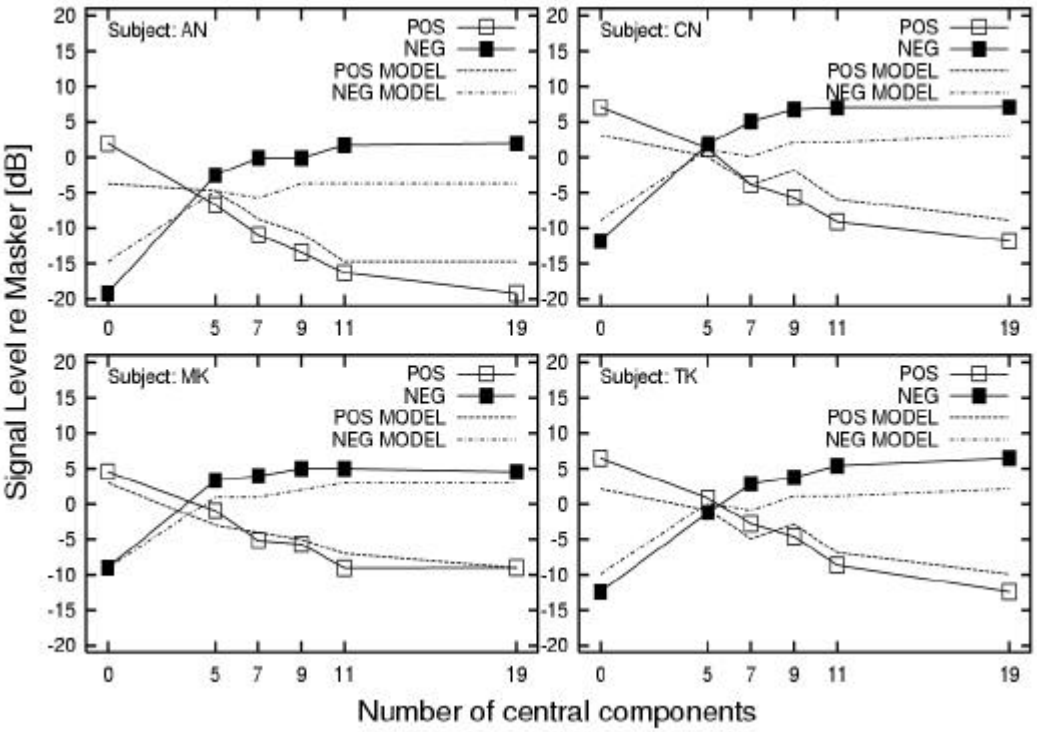

Figure 3: Results of experiment 2 and predictions by the masking model.

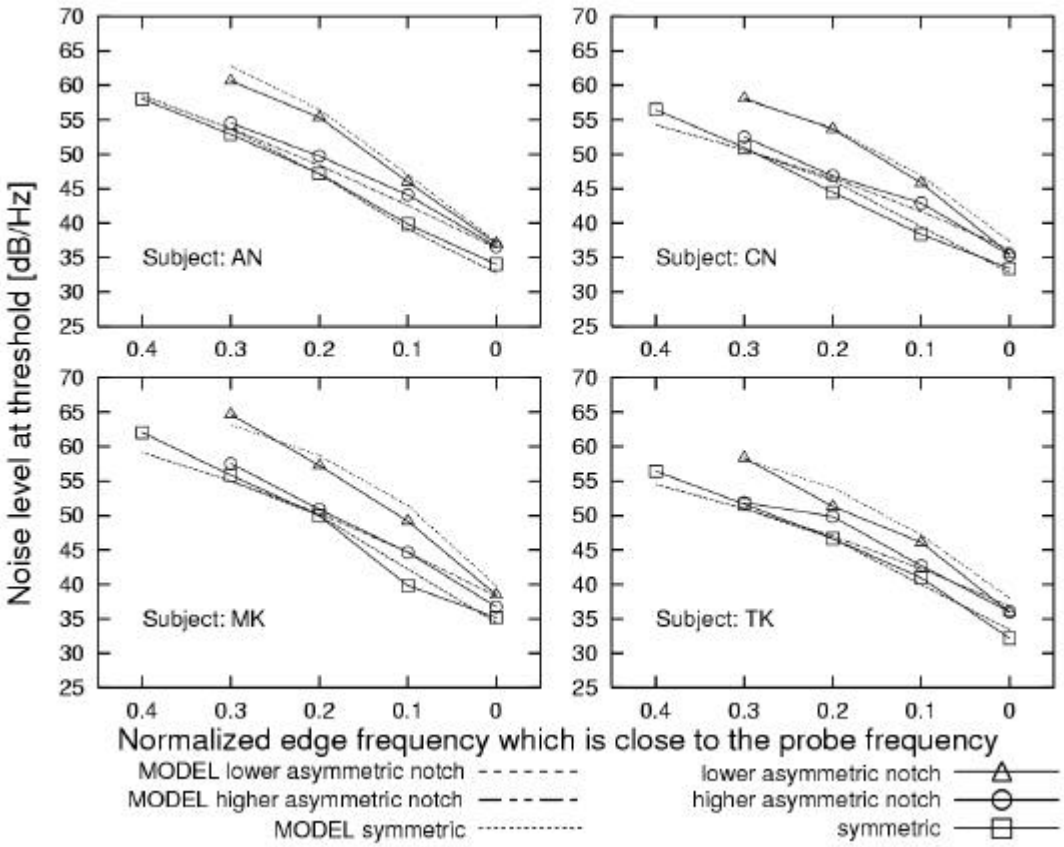

Figure 4: Results of notched-noise experiment and predictions by the power-spectrum model.

Table 1 shows parameter values of the model for each subject after fitting. Figures 2, 3, and 4 show the masking data and the predictions of the model for each subject. These figures show good agreement between the predictions of the models and the experimental data for individual subjects.

DISCUSSION

Irino and Patterson [10] showed that the compressive gammachirp provides reasonable fit to the level-dependent impulse response of auditory-nerve fiber in cats [11], including their instantaneous frequency trajectories. However, direct measurement of the phase characteristics of the auditory periphery is not possible in humans. The dependence of the phase of the complex maskers on a short signal indirectly provides useful insight into the phase characteristics of the auditory filter. In addition, constraints of both amplitude and phase characteristics make estimation of the auditory filter more realistic.

The present simulations of the models may be affected by the amplitude and phase characteristics of the headphones used in the experiments. The amplitude response of the headphones affects estimations of the shape of auditory filters [9]. One possible way to estimate the effect of the headphones is insertion of a module that simulates headphones just before the fixed filter for outer/middle ear response in the model. Therefore, the impulse response of the headphones, which was measured by generating a TSP-signal in the IEC coupler (B&K 4153, B&K 4134), was convolved with the input signal. Then, simulations of both the power-spectrum model and the present masking model were conducted once more while maintaining the same parameter values obtained from the fitting.

The results of this simulation show 3-dB differences in the maximum compared with the results of the previous simulation. However, the tendency of the experimental data was maintained in the predictions of the model. Therefore, the effects of the amplitude and phase characteristics of the headphones on the present results are believed to be relatively small.

Table 1: Parameter values after fitting for each subject.

Subject	compressive gammachirp					leaky integrator	detection	overall
	b_1	c_1	b_2	c_2	f_{rat}	Tc[ms]	MAXSMMR [dB]	rms error [dB]
AN	1.09	−3.90	2.69	0.67	0.63	1.05	9.56	3.34
CN	1.01	−3.50	1.51	1.43	2.95	1.19	11.5	2.56
MK	0.92	−3.46	1.46	2.74	4.38	1.32	9.41	1.97
TK	1.00	−3.70	3.52	1.45	3.10	1.21	10.5	2.34

BIBLIOGRAPHICAL REFERENCES

[1] 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., 80, 1631-1637 (1986).

[2] A. Kohlrausch and A. Sander, "Phase effects in masking related to dispersion in the inner ear II. Masking period patterns of short targets," J. Acoust. Soc. Am., 97,1817-1829 (1995).

[3] R. P. Carlyon and A. J. Datta, "Excitation produced by Schroeder-phase complexes - Evidence for fast-acting compression in the auditory system," J. Acoust. Soc. Am., 101, 3636-3647 (1997).

[4] R. P. Carlyon and A. J. Datta, "Masking period patterns of Schroeder-phase complexes: Effects of level, number of components, and phase of flanking components," J. Acoust. Soc. Am., 101, 3648-3657 (1997).

[5] Akira Nishimura, "An auditory model that explains masking by Schroeder-phase complexes," Acoustical Science and Technology, 22, 439-441 (2001).

[6] A. J. Oxenham and T. Dau, "Reconciling frequency selectivity and phase effects in masking," J. Acoust. Soc. Am., 110, 1525-1538 (2001).

[7] T. Dau, B. Kollmeier, and A. Kohlrausch, "Modeling auditory processing of amplitude modulation I. Detection and masking with narrowband carriers," J. Acoust. Soc. Am., 102, 2892-2905 (1997).

[8] T. Irino and R. D. Patterson, "A time-domain, level-dependent auditory filter: The gammachirp," J. Acoust. Soc. Am., 101, 412-419 (1997).

[9] B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," Hearing Research, 47, 103-138 (1990).

[10] T. Irino and R. D. Patterson, "A compressive gammachirp auditory filter for both physiological and psychophysical data," J. Acoust. Soc. Am., 109, 2008-2022 (2001).

[11] L. H. Carney, M. J. McDuffy, and I. Shekhter, "Frequency glides in the impulse response of auditory-nerve fibers," J. Acoust. Soc. Am., 105, 2384-2391 (1999).