

HST.723J, Spring 2005 – Theme 1 Report

Madhu Shashanka
shashanka@cns.bu.edu

Introduction

The theme of this report is *frequency selectivity* and *masking*. Moore (1995) defines frequency selectivity as “the ability of the auditory system to separate (to certain extent) the components in a complex sound”. One way to demonstrate and quantify this ability is to study masking - the interference one stimulus can cause in the perception of another stimulus. This report is a discussion of the following papers: Moore (1978), Oxenham and Plack (1997), Oxenham and Shera (2003), Delgutte (1988) and Ruggero et al. (1992).

Psychophysical Studies

Moore (1978) and Oxenham and Shera (2003) deal with measuring psychophysical tuning curves. Moore (1978) studied the validity of using psychophysical tuning curves as analogues of neural tuning curves. The aim was to provide more information about the differences between simultaneous and forward masking by a detailed comparison of psychophysical tuning curves. It was observed that forward masking showed steeper slopes, particularly on the high-frequency side, and the tip bandwidths were narrower (see figure 1). This is consistent with the idea that in simultaneous masking (which shows broader tuning curves), the threshold of the probe may be influenced by lateral suppression. In forward masking, lateral suppression presumably plays no role. However, Moore pointed out some problems with using the forward-masking paradigm. One problem is that of the observer being able to distinguish between the neural activity evoked by the probe and that evoked by the masker when the probe is very close in frequency to the masker. Pitch differences between the masker and probe may act as cues which enhance the detectability of the probe relative to the case where pitch differences are minimal or absent. The result of this would be that the tuning curve derived psychophysically would be sharper than the true neural tuning curve. The second problem derives from the assumption that a given elevation in probe threshold is produced by a fixed amount of masker-evoked activity at the probe frequency. But it has been shown that as delay between the masker and the probe is increased, the level of the masker required to mask the probe increases more rapidly for maskers above the probe frequency than for maskers below this frequency. One can shorten the duration of the probe and the fall time of the masker, but only at the expense of increased “energy splatter” in the stimuli. And finally, both forward and simultaneous tuning curves may be influenced by the observer’s integration of information over an array of neurons, and by shifts in the “auditory filter” so as to optimize signal-to-masker ratios. Moore concludes that neither gives a satisfactory estimate of the shape of human neural tuning curves.

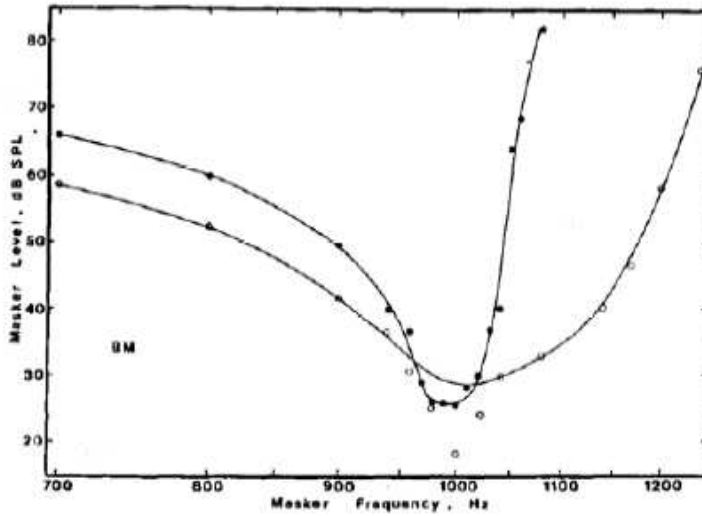


FIG. 4. Comparison of simultaneous (open circles) and forward (filled circles) masking curves for a 1-kHz probe at 10 dB SL. Subject BM.

Figure 1: From Moore (1978)

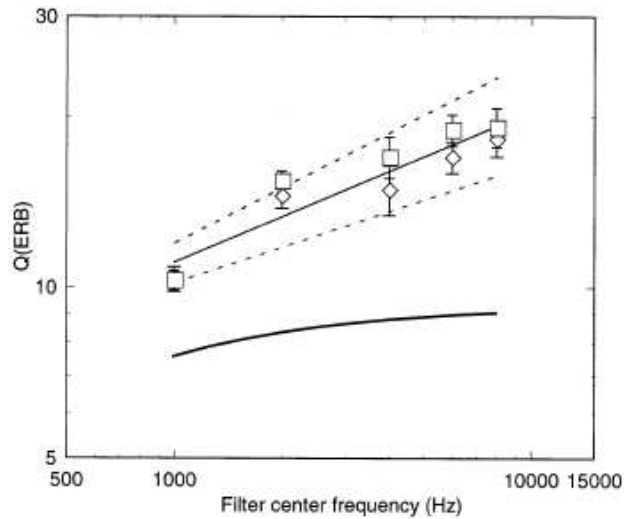


FIG. 5. Geometric mean values of Q_{ERB} from the $roex(p,w,t)$ (squares) and $roex(p,w,t,p)$ (diamonds) models. Error bars denote 1 SE of the mean. The Q_{ERB} is a dimensionless measure of tuning, defined as the center frequency divided by the ERB. The heavy solid curve shows the predicted Q_{ERB} , as defined by Glasberg and Moore (1990). The light solid line and the dashed lines represent the best-fitting power function to the pooled estimates and the 95 confidence intervals, respectively.

Figure 2: From Oxenham and Shera (2003)

Oxenham and Shera (2003) tried to estimate human cochlear tuning as accurately as possible using a variant of the notched noise method. As they rightly pointed out, while ignoring cochlear nonlinearities for the sake of simplicity has certain advantages, some important effects are dependent on these nonlinearities. If a psychoacoustic measure of cochlear tuning has to share many of the properties of neural measures of tuning, the masker and the signal should not be presented simultaneously to avoid suppression effects and the stimuli should be at a low level where tuning is generally sharpest. As Moore (1978) points out, responses of the listener will be determined by the activity over an array of neurons and the listener could shift his “auditory filter” so as to optimize the signal-to-masker ratio. However, at very low levels, the opportunities for such shifting are very small. The signal was presented at a fixed low level and the masker level was varied to measure threshold. They studied the differences between forward masking and simultaneous masking systematically as a function of frequency and found that frequency tuning increases substantially with CF between 1 and 8 kHz in both conditions. Filter shapes were derived by assuming the rounded exponential filter (figure 3) and it was found that tuning was considerably sharper than previously thought (figure 2). The data of this study is in agreement with the results from the study of Shera et al. (2002). Further studies with forward masking, using either higher signal levels or longer gaps between the masker and signal, will be required to define how tuning changes with stimulus level.

Oxenham and Plack (1997) examined the possibility of estimating basilar-membrane nonlinearity using a psychophysical technique. The method relates to the nonlinear growth of masking (GOM) observed when a masker is well below the signal frequency. BM responds linearly to tones with a frequency well below CF whereas BM response to the signal at CF is compressive. There was a discrepancy between the slope of the predicted GOM function, based on physiological measurements of BM nonlinearity, and that actually observed. The study employed a forward masking paradigm to study the on-frequency compression of the BM and used a very brief signal, presented close to the offset of the masker, in order to measure thresholds over a wide range of signal levels. For normally hearing listeners, a background noise was simultaneously presented in order to restrict “off-frequency listening”. The background noise may also have restricted “off-time listening”. Data showed that GOM corresponds to a slope of 6.25 (between signal levels 50 and 80 dB SPL - see figure 4), which contrasts strongly with previous estimates of around 2. These results do not contradict results from the physiological study of Delgutte (1988) if the signal levels are compared (those conclusions may apply for low, but not high, signal levels). Delgutte’s study concludes that the upward spread of masking was due to the effects of suppression, while this study suggests that suppression may in fact reduce the nonlinear growth of the upward spread of masking. The authors concluded that a comparison of the effects of an on-frequency forward masker well below the signal frequency provides an estimate of BM compression which is in quantitative agreement with physiological measurements. They observed that the loss of compression inferred from the results of hearing-impaired listeners was also consistent with physiological effects of damage to the cochlea and suggested that a similar measure could be used for the diagnosis of hearing-impairment. Though the presence of background noise had a large effect on masked thresholds, they concluded that the resulting estimate of BM nonlinearity was not dependent on the presence or absence of noise.

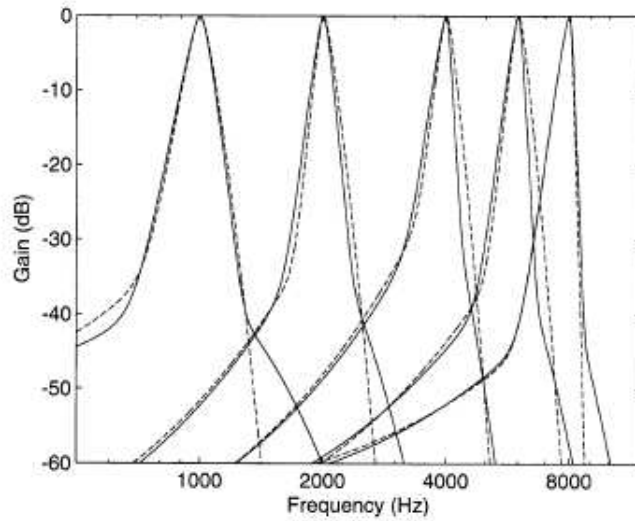


FIG. 6. Auditory filter shapes with center frequencies between 1 and 8 kHz, derived from the mean data in the forward-masking condition. Solid and dotted curves denote $\text{roex}(p,w,t)$ and $\text{roex}(p,w,t,p)$ filters, respectively.

Figure 3: From Oxenham and Shera (2003)

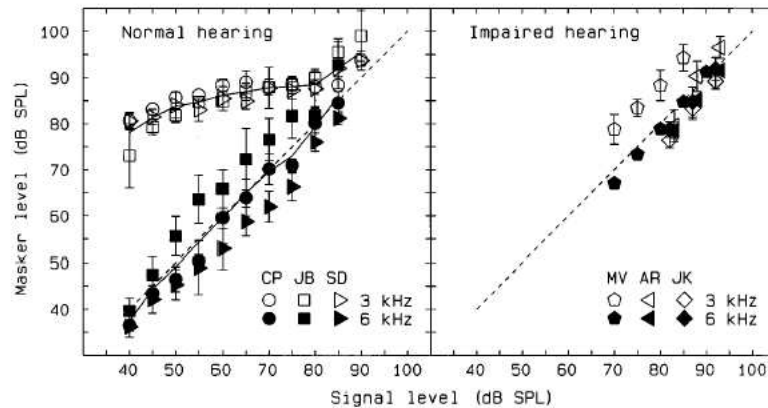


FIG. 2. The level of a masker required to mask the 6-kHz signal, as a function of signal level. Data from the three normally hearing listeners are shown in the left-hand panel and data from the three hearing-impaired listeners are shown in the right-hand panel. Error bars represent \pm one standard deviation, and are omitted if they are smaller than the respective symbol. Solid curves denote the mean thresholds of the listeners in the normally hearing group, and the dashed lines denote linear growth of masking.

Figure 4: From Oxenham and Plack (1997). Slope of this masking function is the reciprocal of the more usual GOM function.

Physiological Studies

In Delgutte (1988), the goal was to identify the contributions of suppression and spread of excitation to tone-on-tone masking by comparing masked thresholds of auditory-nerve fibers measured with simultaneous and non-simultaneous techniques. As pointed out earlier, the difference between simultaneous and non-simultaneous thresholds gives a measure of the contribution of suppression to masking because there is no suppression when the signal and masker are not simultaneously present (see figure 5). Recordings were made from auditory-nerve fibers in anesthetized cats. It was found that excitatory masking dominates for signal frequencies near and below the masker frequency, although suppression also contributes somewhat below the masker frequency. Suppressive masking dominates for signal frequencies well above the masker. Overall, the non-simultaneous masking pattern was more sharply tuned than the simultaneous pattern. The study concluded that the upward spread of simultaneous masking was due primarily to the supralinear growth of suppression rather than to the growth of excitation. Masked thresholds were measured for signals of different frequencies rather than just measuring the threshold at the CF so that the contribution of off-frequency listening to signal detection could be accounted for. Delgutte (1990) points out that physiological studies of masking in which the signal is constrained to be at the CF of auditory neurons may considerably underestimate detectability.

The conclusions are based on the assumption that the difference between simultaneous and non-simultaneous thresholds measures the contribution of rate suppression to masking. One difficulty with this assumption is that, in the absence of a signal, the discharge rates of auditory nerve fibers during the signal interval can differ for the two masking conditions because the masker is present in the simultaneous condition, but not in the non-simultaneous condition. Hence, non-simultaneous thresholds will be somewhat higher than simultaneous thresholds when there is no suppression. When masking is both excitatory and suppressive, the contribution of suppression to masking will be systematically underestimated. A second possible bias in measuring suppression is that the responsiveness of auditory nerve fibers may differ for the two masking conditions because signal levels are higher in simultaneous than in non-simultaneous masking. The author concludes (Delgutte (1990)) that these inaccuracies are too small to invalidate the conclusions concerning the role of suppression in masking.

Ruggero et al. (1992) report their experiments to characterize mechanical two-tone suppression in the basilar membrane (chinchilla) and compare it with auditory nerve rate suppression properties. They concluded that mechanical suppression effects were demonstrated that closely resemble, qualitatively and quantitatively, most features of rate suppression in auditory nerve fibers. Suppression magnitude was maximum at probe frequencies near CF. It increased as a function of suppressor intensity and decreased with probe intensity. Suppressor tones linearized probe-intensity curves (figure 7). Rate of growth of suppression for increasing suppressor intensity was larger for below-CF suppressors than for higher frequency suppressors (figure 9). Low-frequency suppressor tones had suppression magnitudes which varied periodically (figure 8). Cochlear death was found to destroy suppression. Positive correlation was found between suppression and BM sensitivity to CF tones. They concluded that mechanical suppression could result from interactions between outer hair cells and the basilar membrane.

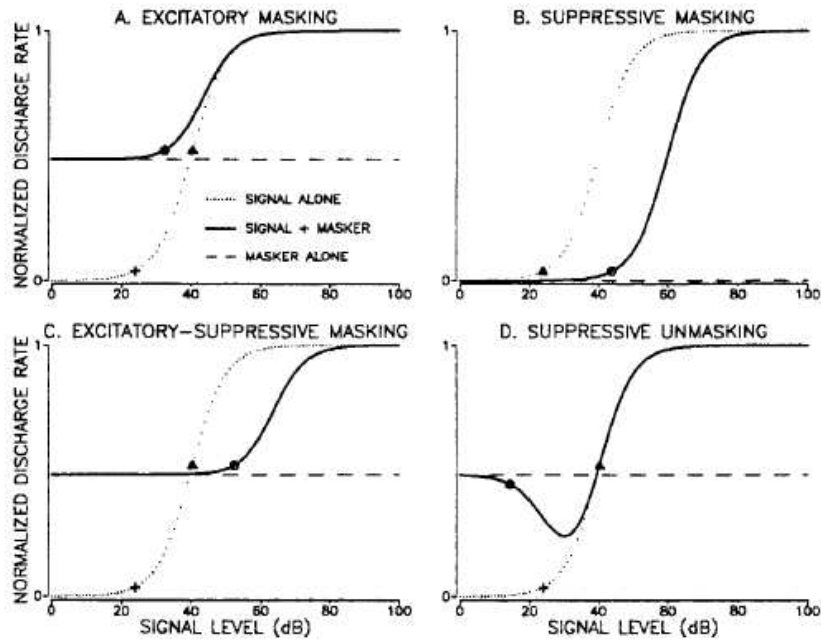


Figure 1. Schematic discharge rate of an auditory-nerve fiber as a function of the level of a tone signal, both in the presence and in the absence of different maskers. Crosses refer to thresholds in quiet, circles to simultaneous masked thresholds, and triangles to nonsimultaneous thresholds.

Figure 5: From Delgutte (1988)

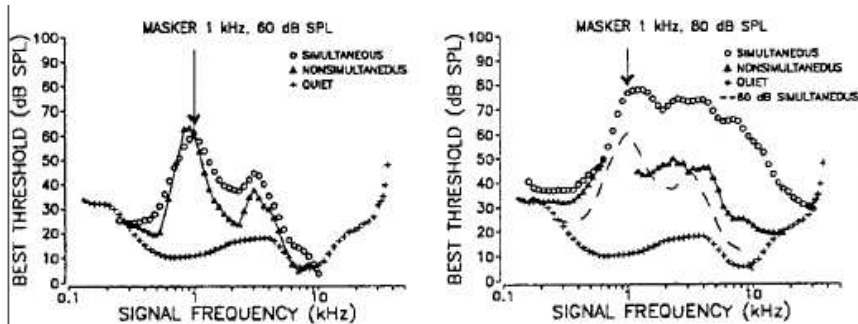


Figure 4. Best thresholds of auditory-nerve fibers as a function of signal frequency, in quiet, and for both simultaneous and nonsimultaneous 1-kHz masker at 60 dB SPL. Each data point is obtained from threshold measurements in at least 10 auditory-nerve fibers.

Figure 5. Same as Fig. 4 for an 80-dB masker. The available data did not allow reliable estimates of best nonsimultaneous thresholds for signal frequencies near 1 kHz.

Figure 6: From Delgutte (1988)

Their data is in agreement with results from Delgutte (1988). For probe frequency at CF, they report that suppression exists both for suppressor frequency above and below CF (at stimulus intensities at which the response to the suppressor is smaller than the response to the probe alone - see figure 10). Their findings that below-CF suppressors cause faster growth of suppression is consistent with the conclusions in Delgutte (1988) (figure 6).

Comments

While comparing results and conclusions, one should exercise care and caution. When comparing psychophysical data and physiological data, one should bear in mind the species differences. Oxenham and Shera (2003) mention studies on the guinea pig by Evans and colleagues which found good correspondence between auditory-nerve tuning curves and behavioral measures using simultaneous masking. This is at odds with the authors' assertion that suppression effects in simultaneous masking can severely underestimate cochlear tuning. This discrepancy can be resolved by considering that guinea pig cochlear tuning is generally poorer than that found in humans by a factor of 2 or 3. As another example, Moore (1978) mentions data of Dallos and colleagues obtained from chinchillas with kanamycin-induced hair-cell losses. Such animals may show very broad, and clearly abnormal, neural tuning curves, while showing relatively normal psychophysical tuning curves in forward masking. The possibility of species differences in more than just the frequency range of hearing needs to be examined further. One should also keep in mind differences in experiment paradigms, signal/masker conditions etc. For example, we saw earlier that the results of Oxenham and Plack (1997) might appear to be in conflict with the results of Delgutte (1990) if one does not consider the signal levels used.

The papers assigned for this theme tackle important issues like differences between simultaneous and non-simultaneous masking, measuring suppression and its role in masking, importance of off-frequency listening, nonlinearities in cochlear tuning, behavioral estimation of BM nonlinearities, relating BM properties and auditory-nerve properties and others. We have discussed their results and conclusions in the previous sections.

Conclusions

Psychophysical and physiological studies of masking help us to gain a better understanding of the frequency selectivity of the auditory system. However, it is not straightforward as there are many confounding factors. In this theme, I learned about the tradeoffs involved in trying to simplify the system and accurately modeling it. More importantly, I learned how nonlinearities are essential in understanding crucial properties of the peripheral auditory system.

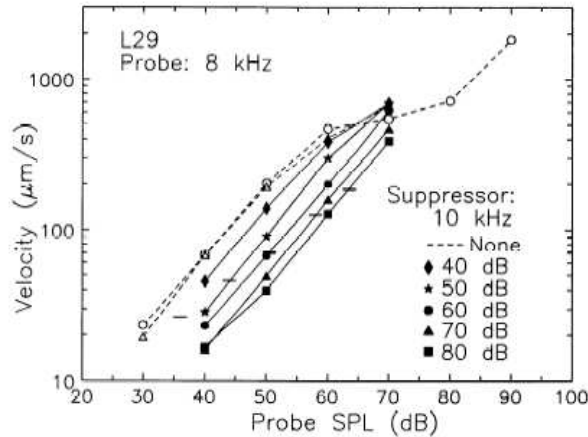


FIG. 3. Comparison of probe-tone intensity functions for a near-characteristic frequency (CF) probe tone presented alone and simultaneously with above-CF suppressor tones. Intensity functions are shown for a probe tone (8 kHz) alone (---) and in the presence of 10-kHz suppressors at several intensities (—). Probe-alone intensity functions were obtained both early in the experiment (~5–6 h preceding the suppression measurements; \circ) and interspersed with 2-tone recordings (Δ). Horizontal bars indicate responses to the suppressor tones. Data collected by means of laser velocimetry. Some of the probe and suppression velocities were measured from the waveforms of Fig. 1.

Figure 7: From Ruggero et al. (1992)

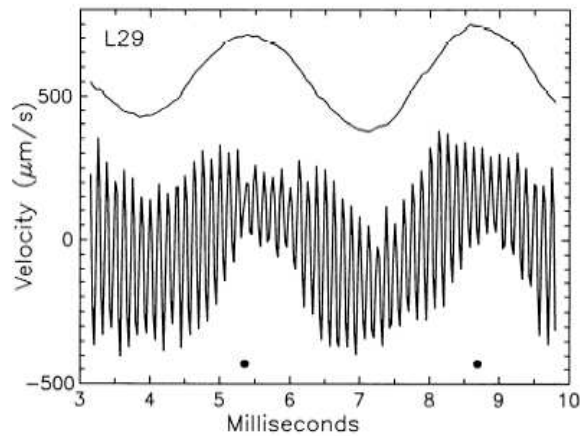


FIG. 7. Velocity response to a stimulus consisting of a characteristic frequency (CF) tone and a low-frequency suppressor tone (*bottom trace*). CF tone: 8 kHz, 60 dB SPL. Suppressor tone: 300 Hz, 80 dB SPL. Traces depict 6.67-ms segments of averaged responses to 10-ms stimuli. *Top trace*: 2 cycles of the response to the suppressor tone, extracted by low-pass filtering the response to a stimulus consisting of 2-tones: 8 kHz, 40 dB SPL and 300 Hz, 100 dB SPL. Abscissa indicates elapsed time from stimulus onset. Solid circles indicate the times at which the basilar membrane was moving with maximum velocity from scala tympani toward scala vestibuli.

Figure 8: From Ruggero et al. (1992)

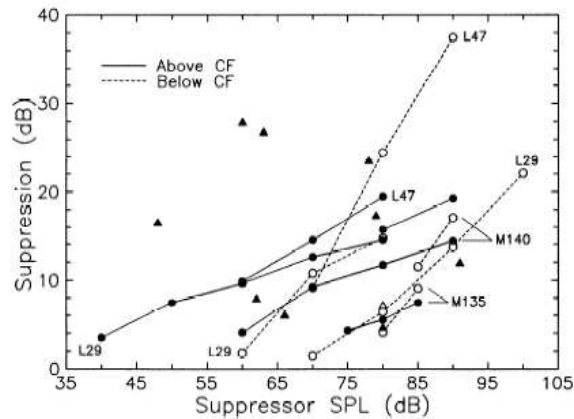


FIG. 5. Suppression magnitude as a function of suppressor-tone intensity. Single data points (triangles) and connected series of data points (circles) represent magnitudes of suppression obtained in a given experiment for a single combination of probe and suppressor frequencies, with the probe-tone frequency always near characteristic frequency (CF). Open symbols and dashed lines represent data for below-CF suppressors; filled symbols and solid lines are for above-CF suppressors. Suppression magnitude was determined as the horizontal shift of the probe-tone intensity function at a velocity near the lower extreme of the measured range. Data are from 15 animals and include both suppressors that evoked a measurable response as well as suppressors that did not. Suppression by above-CF and below-CF suppressors was studied in 14 cochleas and 5 cochleas, respectively. For the sake of clarity, 3 single data points have been displaced by 1 dB.

Figure 9: From Ruggero et al. (1992)

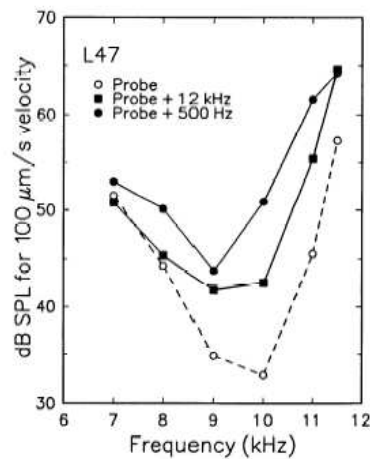


FIG. 10. Changes in basilar-membrane frequency tuning produced in cochlea L47 by below-characteristic frequency (CF: 500 Hz) and above-CF (12 kHz) suppressor tones presented at 70 dB SPL. The 500-Hz and 12-kHz suppressor tones evoked responses (66 and 50 $\mu\text{m/s}$, respectively) smaller than the isovelocity criterion (100 $\mu\text{m/s}$). The isovelocity tuning curves were derived from intensity functions for probe tones presented alone (\circ - - - \circ) and in the presence of 12-kHz or 500-Hz suppressor tones (\blacksquare - - - \blacksquare and \bullet - - - \bullet , respectively).

Figure 10: From Ruggero et al. (1992)

References

- Delgutte, B. (1988). Physiological mechanisms of masking. In Duifhuis, H., Horst, J. W., and Wit, H. P., editors, *Basic Issues in Hearing.*, pages 204–214. Academic, London.
- Delgutte, B. (1990). Physiological mechanisms of psychophysical masking: Observations from auditory nerve fibers. *J. Acoust. Soc. Am.*, 87(2):791–809.
- Moore, B. C. J. (1978). Psychophysical tuning curves measured in simultaneous and forward masking. *J. Acoust. Soc. Am.*, 63:524–532.
- Moore, B. C. J. (1995). Frequency analysis and masking. In Moore, B. C. J., editor, *Hearing*, pages 161–205. Academic, London.
- Oxenham, A. J. and Plack, C. J. (1997). A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. *J. Acoust. Soc. Am.*, 101:3666–3675.
- Oxenham, A. J. and Shera, C. A. (2003). Estimates of human cochlear tuning at low levels using forward and simultaneous masking. *J. Assoc. Res. Otolaryngol.*, 4:541–554.
- Ruggero, M. A., Robles, L., and Rich, N. C. (1992). Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression. *J. Neurophysiol.*, 68:1087–1099.
- Shera, C. A., Guinan Jr., J. J., and Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc. Natl. Acad. Sci.*, 99:3318–3323.