Theme 1: Masking and Frequency Selectivity

The main theme in this topic area is to highlight the different approaches taken by hearing scientists to characterize the human auditory system. If our auditory system behaves like a linear time-invariant system, whereby the response of the cochlear filter is completely described by the relationship between its inputs and its outputs, such characterization would be a trivial matter. However, our peripheral auditory system is non-linear and is also coupled by further up-stream processing at higher levels in the auditory system. Therefore many different approaches need to be cross-referenced to ascertain whether a given representation of the human auditory system characteristic is valid. In this report, the studies of non-linear phenomena in our auditory system, both in psychophysical and physiological experiments are highlighted. Such non-linearities include two tone rate suppression, basilar membrane (BM) compression near the characteristic frequency (CF), cochlear amplification, and the psychophysical phenomena of simultaneous and non-simultaneous masking.

Psychoacousticians try to characterize our auditory system from a top-down, non-invasive approach. Moore (1978) compared the validity of using simultaneous-masking with forward-masking tuning curves as an estimation of the shape of human neural tuning curves. Results revealed that comparing to the simultaneous-masking curves, forward-masking curves, in general, show steeper slopes particularly on the high-frequency side and the tip bandwidths are narrower (Figure 1). It was hypothesized that the threshold of the probe in simultaneous masking may be influenced by lateral suppression. This view is partially supported by later physiological experiments performed by Delgutte (1990), in which he concluded that simultaneous masking is not just suppressive, but also excitatory in nature.

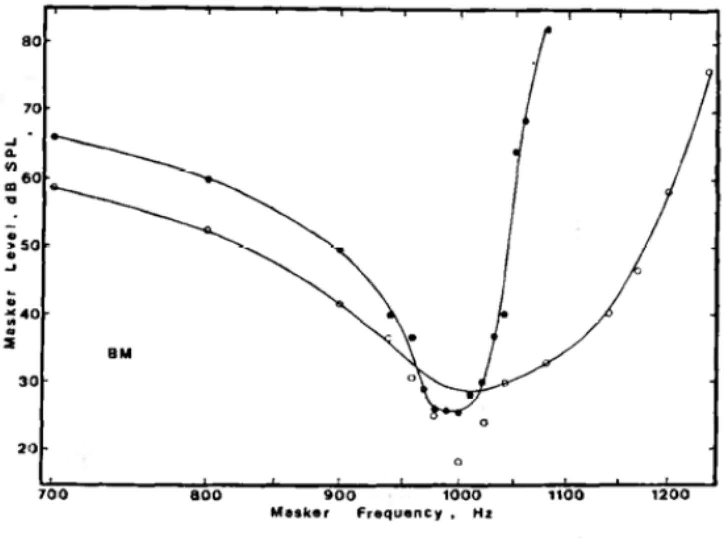


FIG. 4. Comparison of simultaneous (open circles) and forward (filled circles) masking curves for a 1-kHz probe at 10 dB SL. Subject BM.

Figure 1 Typical difference of masking curves between simultaneous and forward masking. Moore (1978).

In non-simultaneous masking, the masker does not suppress the signal (Delgutte, 1998), and thus the masker is less effective. Moore (2003) argues that psychophysical tuning curve measured in non-simultaneous masking is likely to be closely related to basilar membrane (BM) tuning curves and neural tuning curves, while the psychophysical tuning curve measured in simultaneous masking is likely to be broader than those curves. However, Moore (1978) pointed out the potential problems that are associated with the forward-masking paradigm. One problem facing the observer in a forward-masking situation is when the masker and the signal are close in frequency; the observer will find it difficult to detect the probe since it might be detected as a slight lengthening in duration of the original masker. Another major problem in comparing these psychophysical tuning curves with the single-neuron tuning curves is that the responses of the listener will be determined by the activity over an array of neurons, rather than by the activity in any single neuron (Moore, 1978). Shera *et al.* (2003), used a noise masker extending spectrally both above and below the signal frequency, to avoid effects of such “off-frequency listening” and “confusion“ between the masker and signal, and indeed, these variables have a huge bearing on the experimental results, implication of which will be discussed in the later part of this report.

Psychoacousticians would also like to characterize the compressive nature of the BM through masking experiments. The compressive non-linearity of a functional cochlea can be implied by comparing the growth of masking near the characteristic frequency (CF) with a frequency well below CF. This model is further validated by contrasting the psychoacoustical measurements of normal hearing subjects from the hearing impaired subjects, which has a characteristic loss of BM compression at CF. Oxenham *et al.* (1997) used non-simultaneous masking to obtain a behavioral measure of the BM non-linearity and contrasted it with the previous results in the literature obtained by using simultaneous masking. It was assumed that the BM responds linearly to frequencies well below the CF and compressively near the CF. Therefore, ‘the ratio of the slopes of the masking functions between a masker at the signal frequency and a masker well below the signal frequency should provide an estimate of BM compression at CF’.

Oxenham *et al.* (1997), pointed out that in previous behavioral studies using simultaneous-masking surveyed by Stelmachowicz, the slope of the growth of masking (GOM) function rarely exceeded 2. There is a huge discrepancy observed by BM motion studies in the literature, where the results point to a slope of more than 5 for levels above 40dBSPL. From Figure 2, the GOM function corresponds to a slope of 6.25. The authors argued that simultaneous masking would lead to an underestimation of BM compression because Ruggero *et al.* (1992) have shown that BM response to a tone at CF is reduced in the presence of a low frequency suppressor, and the growth of response becomes more linear, (see Figure 5[1], later in the report). It is also stated that by their use of forward masking, it eliminates the effects of suppression, and the slope of the GOM function is obtained to be more closely matched to the physiological data, at least for the signal level used (50dBSPL and higher). This paper thus argues that suppression may reduce the nonlinear growth of the upward masking, which is seemingly in contrast to Delgutte’s (1988) conclusion that upward spread of masking is primarily due to the effects of suppression. However, Oxenham *et al.* (1997) highlighted that in Delgutte’s study, signal thresholds in non-simultaneous masking conditions were rarely above 50dBSPL, and therefore, his conclusions may apply for low, but not high, signal levels.

---

[1] The paper referred to here is Ruggero, M.A., Robles, L., Rich, N. C., and Recio, A. (1992). “Basilar membrane responses to two-tone and broadband stimuli,” Philos. Trans. R. Soc. London, Ser. B **336**, 307-315. However, there is a diagram in Ruggero *et al.* (1992) that illustrates a similar observation, and since it is one of the discussion papers, the same diagram is used here for explanation.

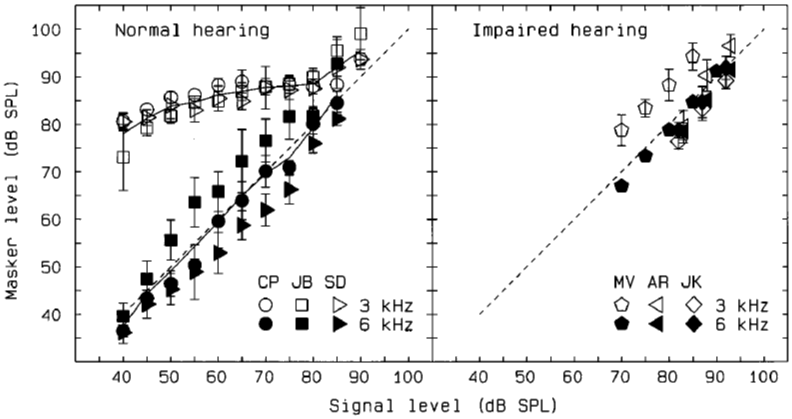


FIG. 2. The level of a masker required to mask the 6-kHz signal, as a function of signal level. Data from the three normally hearing listeners are shown in the left-hand panel and data from the three hearing-impaired listeners are shown in the right-hand panel. Error bars represent ± one standard deviation, 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 2 Masker level at threshold as a function of signal level. Slope of this masking function is the reciprocal of the more usual GOM function. Normal hearing and impaired hearing data from Oxenham *et al.* (1997).

Oxenham *et al.* (1997) further extended their ideas of estimating the BM non-linearity of normal hearing subjects to cochlear hearing impaired subjects. Evidence in the literature supported the hypothesis that ‘damage to the cochlea, in particular the outer hair cells (OHCs), results in a reduction in sensitivity and a loss of compression at CF’, and Figure 3 shows much less compressive masking functions compared to the normal hearing subjects. While the available data showed that limited cochlear hearing loss may not necessarily result in a uniform reduction of BM compression over the entire level range, it may instead reduce the range of levels over which “normal” compression is observed, but more data need to be obtained to validate this observation. The authors argued that ‘a reliable measure of BM compression may complement more traditional tests, such as loudness judgments and measures of frequency selectivity, in determining whether such a two-component approach can help account for the variability observed in hearing-impaired listeners with the same absolute hearing loss’.

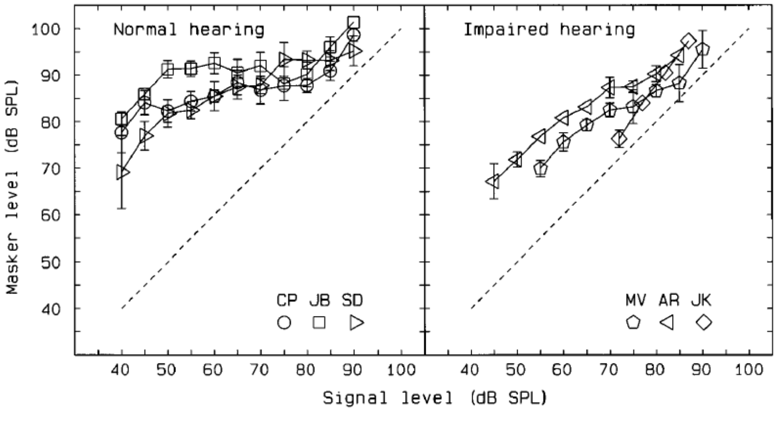


FIG. 5. Masker level at threshold as a function of signal level using a 2-kHz signal and a 1-kHz masker. Individual results from the normally hearing and hearing-impaired listeners are shown in the left- and right-hand panels, respectively. Error bars represent ± one standard deviation, and are omitted if they are smaller than the respective symbol.

Figure 3 Masker level at threshold as a function of signal level using a 2kHz signal and a 1kHz masker. Comparison between normally hearing and impaired hearing in GOM function. Oxenham *et al.* (1997).

Direct access to the cochlea using invasive techniques, such as recording of the basilar membrane motion (e.g. Ruggero *et al.*, 1992) or responses from single units of auditory nerve fibers (e.g. Delgutte, 1998), can provide us a glimpse of the non-linearity of the human peripheral auditory system, albeit by drawing inferences from the results of recording made on other mammalian species. Delgutte (1988) carried out single auditory fiber measurements to separate the contributions of two-tone rate suppression and spread of excitation to tone-on-tone masking related to psychophysical simultaneous and non-simultaneous masking phenomena. He reported that, similar to the psychoacoustical data, the non-simultaneous masking pattern is more sharply tuned than the simultaneous pattern (Figure 4). By comparing the masked thresholds of auditory nerve fibers measured with simultaneous and non-simultaneous techniques, he concluded that both suppression and spread of excitation are important for explaining psychophysical data on simultaneous masking in terms of physiological mechanisms. There is also a good correspondence between psychophysical and physiological data in simultaneous masking showing that the masked thresholds grow faster than linearly with masker level for signal frequencies above the masker, a phenomenon called "upward spread of masking". As mentioned previously in this report, Delgutte (1988) concluded that physiological masking is, in general, both excitatory and suppressive. However, excitatory masking dominates for signal frequencies near and below the masker frequency, and suppressive masking dominates for signal frequencies well above the masker, particularly with the 80-dB masker.

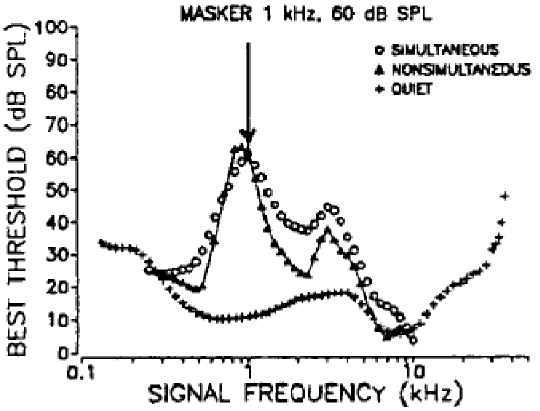


Figure 4 Best thresholds of auditory nerve fibers as a function of signal frequency, in quiet, and for both simultaneous and non-simultaneous 1kHz masker at 60 dBSPL. Delgutte (1988).

Two-tone suppression has also been studied by observing the BM motion inside the chinchilla cochlea (Ruggero *et al.*, 1992). The authors investigated whether the mechanical two-tone suppression in the BM 'permit(s) meaningful comparison with the well-known properties of rate suppression in the auditory nerve'. Their results showed that suppression magnitude grew monotonically with suppressor intensity, regardless whether the suppressor tone is at a frequency above or below CF. In general, however, suppressors at frequencies below CF caused a higher rate of growth of suppression than those with frequencies above CF (Figure 5). The authors claimed that by comparing the main features of mechanical two-tone suppression in the BM and those of its counterparts in two-tone rate suppression in the auditory nerve leads to the conclusion that, at least qualitatively, mechanical and neural two-tone suppression share identical properties. However, in their experiments, the maximum suppression magnitude in the BM for the normal chinchilla cochlea is about 34dB[2], but compared to Delgutte's (1988) data, it is still considerably lower.

---

[2] This value was not measured, but rather, it was inferred by discarding values that are considered not to be from "normal" cochleae after surgical procedures.

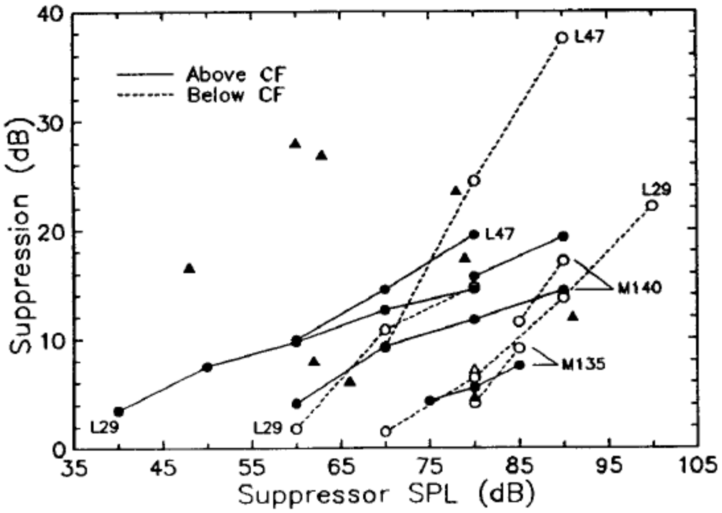


Figure 5 Suppression magnitude as a function of suppressor-tone intensity. Comparison between below-CF and above-CF compressors. Ruggero *et al.* (1992).

Hearing scientists sometimes employ system modeling and parametric variations to see whether it provides some insight into how the real system operates. Cai *et al.* (1996) developed a phenomenological model to simulate some of the non-linearities observed in the auditory nerve fiber recordings, by using functional building blocks that are based on biological cell measurements. The main goal of this model is to simulate some of the two-tone suppression results produced in the cochlea with a low-side suppressor (Figure 6). In this model, the BM displacement p(t) is represented as the sum of the CF contribution of the signal after cochlear amplification and the low-side suppressor signal. BM displacement feeds into two branches: the inner hair cell (IHC) branch and the outer hair cell (OHC) branch. The IHC branch consists of a highpass filter simulating the hydrodynamic response of the stereocilia, an asymmetric IHC transfer function, and a lowpass filter representing the basolateral membrane filter of the IHC. The OHC branch has the same basolateral membrane lowpass filter. However in the OHC branch, the operating point of the transfer function can be varied and it governs the phasic response of the suppression, since it is the saturating characteristic of this function block that produces suppression.

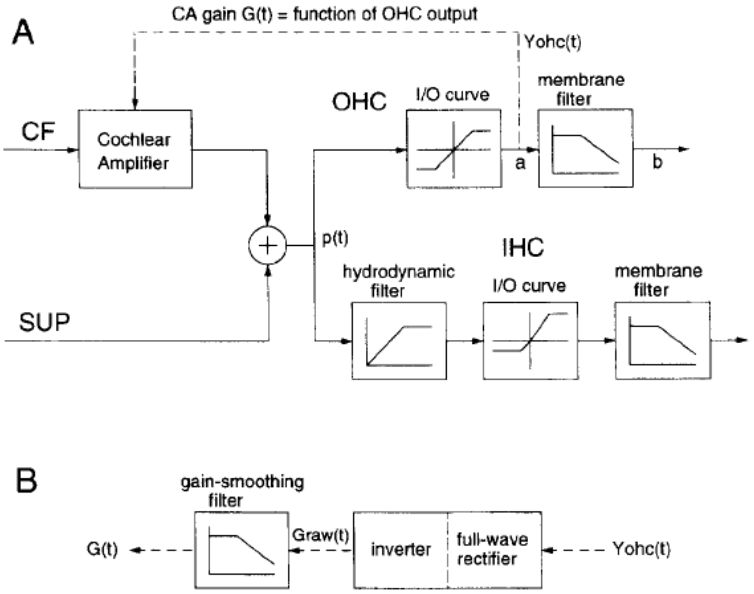


Figure 6 A–Schematic representation of phenomenological model by Cai et al. (1996). B–Gain-control stage of the model.

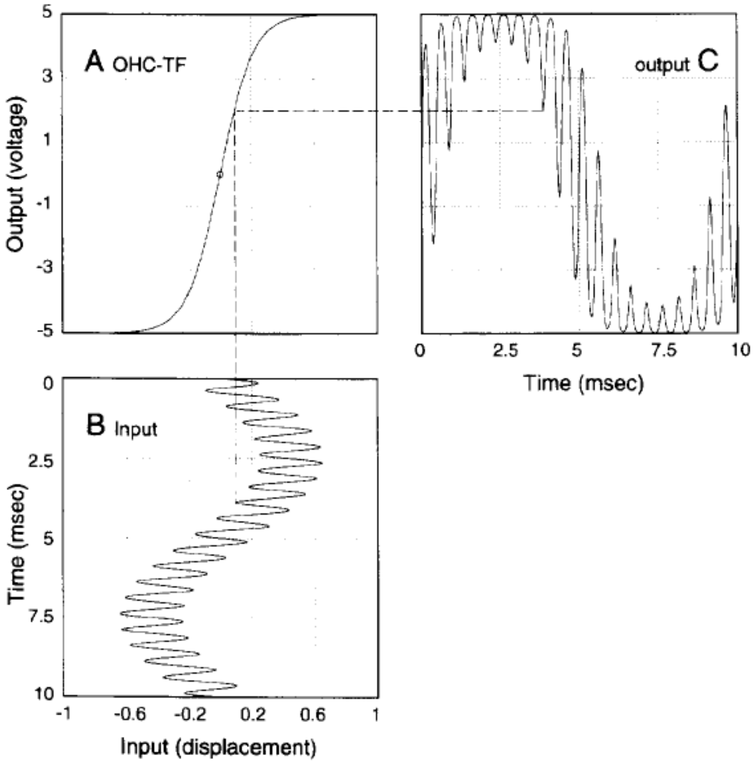


Figure 7 A schematic representation of how the saturating characteristics of a hair cell transfer function produces suppression. A: the transfer function of a hair cell. The small circle indicates the resting point of the hair cell. B: The input signal, a high-frequency tone is added to a low-frequency SUP tone. C: The output of the hair cell. Cai *et al.* (1996).

Another important element in the Cai *et al.* (1996) phenomenological model is their functional representation of the cochlear amplifier. The amplifier feedback was chosen to be a current sampling source at the OHC branch, even though there is compelling evidence in the literature that it is a voltage sampling feedback source. However, the authors argue that it is for simplicity, and not for any physiological reasons, that determines the functional block connection point. The feedback loop itself also consists of an inverter and a gain-smoothing filter, of which they do not have direct physiological counterparts, but were included for functional purposes. Nonetheless, the authors argue that the successful simulations of the suppression patterns support the basic assumption in the model, that the saturation of OHC transduction current produces two-tone suppression. (Figure 7). Even though this system was design to model low-side suppression, it is conceptually capable of simulating high-side suppression, but due to the spatial and phasic complexity in the region of interest for high-tone suppression, the model concentrates only on low-side suppression.

One of the latest and most exciting measurement techniques, which is now available for hearing scientists, is the use of objective, noninvasive otoacoustic experiments. Shera *et al.* (2002) used stimulus-frequency otoacoustic emissions (SFOAE) to obtain across species measurements of cochlear tuning parameters. Previous behavioral measurements suggest that the cochlear frequency selectivity (referred in their paper as parameter Q_{ERB}) of human is qualitatively different from the physiological data obtained from other mammalian species. In the SFOAE measurements, the BM group delays (N_{BM}, which is related to Q_{ERB} by a proportional factor of k, a dimensionless measure of filter shape) across species are qualitatively similar, with a constant offset, which is unlike the previous behavioral data. The authors then postulate that, if the previous behavioral data of Q_{ERB} were correct at low level cochlear tuning, it would imply that the function k must be very different for human, which is not compatible to theoretical expectations and furthermore, it adds additional discrepancies when comparing to its counterparts in other mammalian species (i.e., cat and guinea pig).

The above observations warrant a set of new behavioral measurements of cochlear tuning. Shera *et al.* (2002) used non-simultaneous masking, with low signal levels comparable to the SFOAE measurements, and a noise masker extending spectrally both above and below the signal frequency to avoid “off-frequency” listening and “confusion” between the masker and signal (c.f. to Moore, 1978 in previous discussion). The new Q_{ERB} versus frequency function obtained is substantially different from previous behavioral studies, but there is a strong quantitative agreement with the SFOAE measurements (Figure 8). The authors emphasized that the concordance of the ‘two independent measures of cochlear tuning derived from two completely different kinds of measurements interpreted using very different theoretical frameworks… is significant in itself’. The mutual agreement between physiological and behavioral measures of tuning supports the notion that, at least at the level measured, ‘human auditory frequency selectivity is determined at the level of the periphery’. The authors concluded that these revised estimates of human cochlear tuning challenge our previous view and that ‘ERB does not correspond to a constant distance along the BM’ above 1kHz. The revised measurements also suggest that the human cochlea is considerably sharper than that found in the other mammals.

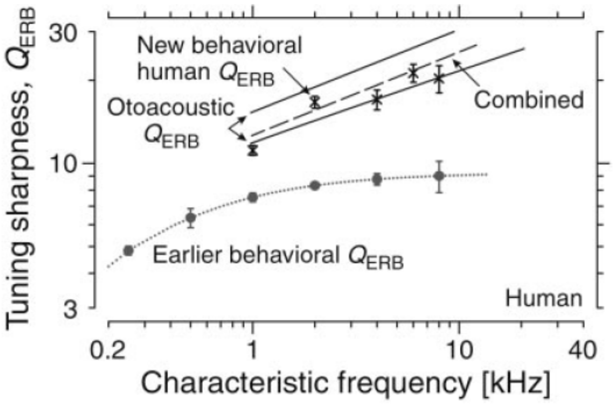


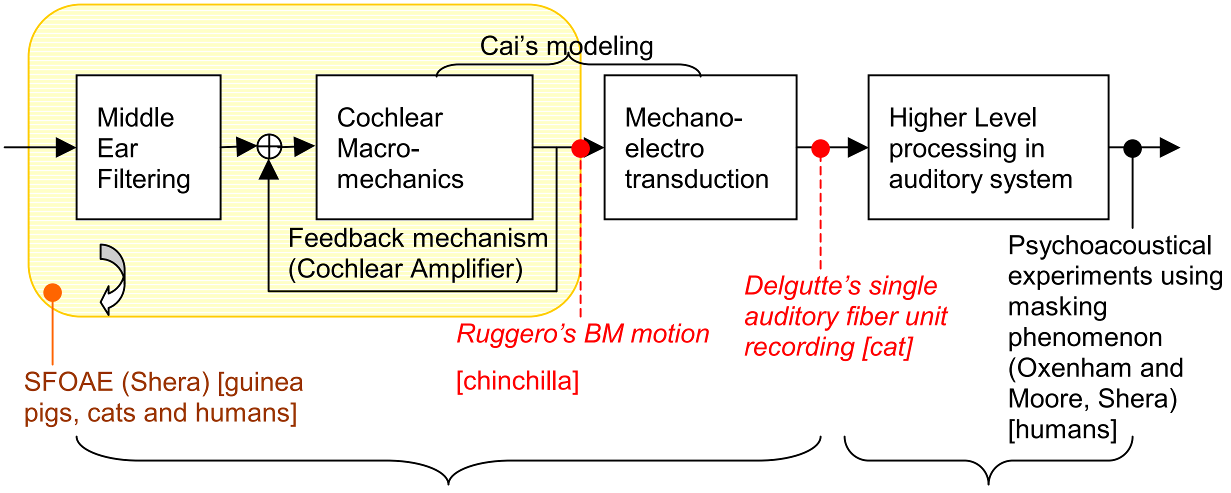
Fig. 4. New human behavioral and otoacoustic Q_{ERB}. Data points (×) give the mean behavioral and its standard error measured at 1, 2, 4, 6, and 8 kHz by using nonsimultaneous masking and other modified psychophysical procedures. The otoacoustic Q_{ERB} values from Fig. 3 are shown for comparison. The dashed line shows a power-law fit to the combined otoacoustic and behavioral values (see Table 1). The earlier behavioral Q_{ERB}, reproduced from Fig. 1, appears in gray.

Figure 8 Shera *et al.* (2002). New human behavioral and otoacoustic Q_{ERB} measurements.

Take home message from Theme 1

Hearing scientists are faced with a complex problem when characterizing the human auditory system response. Behavioral measurements often need to be compared with some benchmark measurements obtained by neural physiologists. However, these data are taken from other mammalian species. This problem highlights the impact of the last paper, Shera *et al.* (2002), especially on the ability to make measurements across species using objective and non-invasive OAE techniques. However, the exact physics behind the reflection theory behind OAE is still not fully appreciated and there is a limitation on the stimulus sound level that can be employed in this method.

The following schematic diagram attempts to summarize what I have learnt in this module. Italic font highlights measurements made from other mammalian species.



Peripheral Auditory System

Non-linearities: 2-tone suppression, non-linear active mechanism around CF

Other variables introduced:

Off-frequency listening, auditory scene analysis (Gestalt principles)

Papers discussed:

Cai Y and Geisler CD (1996). *Suppression in auditory-nerve fibers of cats using low-side suppressors. III. Model results*. Hearing Res. 96:126-140.

Delgutte B (1988). *Physiological mechanisms of masking*. In H Duifhuis, JW Horst, and HP Wit (Eds.), Basic Issues in Hearing. London: Academic, pp. 204-214.

Moore BCJ (1978). *Psychophysical tuning curves measured in simultaneous and forward masking*. J. Acoust. Soc. Am. 63:524-532.

Oxenham AJ and Plack CJ (1997). *A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing*. J. Acoust. Soc. Am. 101:3666-3675.

Ruggero MA, Robles L, and Rich NC (1992). *Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression*. J. Neurophysiol. 68:1087-1099.

Shera CA, Guinan JJ Jr and Oxenham AJ (2002). *Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements*. Proc. Natl. Acad. Sci. 99:3318-3323.

Other References:

Moore, BCJ (2003). *An Introduction to the Psychology of Hearing* (5th ED). Academic Press, UK, pp 107-121.