GOAL

Derive a computational model of *dynamic* auditory perception.

METHOD

Parameterize a Multi-band AGC-inspired model from perceptual forward masking data.

Evidence of Dynamic Audition

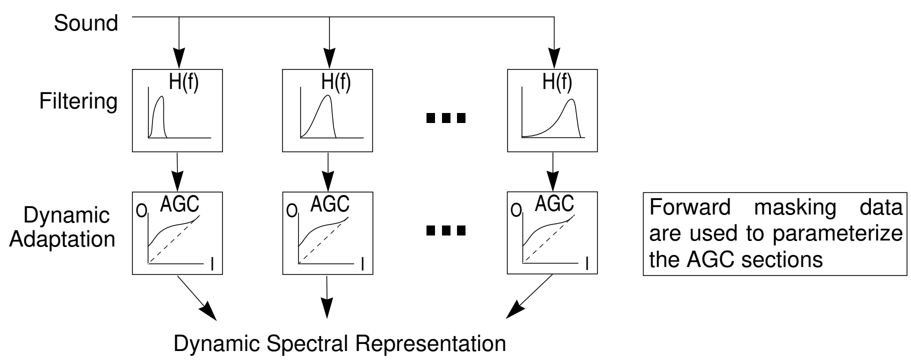
- **Physiological**

  Short-term adaptation in single auditory neurons in response to simple tones (Kiang, 1965)
  Adaptation, recovery, and forward masking in single auditory neurons (Delgutte, 1980)
  Adaptation of outer hair cell motility may include ~40ms time constant (Ashmore, 1987)
  Onsets and transitions emphasized in neural response to speech (Delgutte and Kiang, 1984)

- **Perceptual**

  Forward Masking (Plomb 1964, Jetsteadt 1982,...)
  Onsets and spectral transitions significant for speech perception (Cooper 1952, Furui 1986)

Model Overview (MB-Compression)

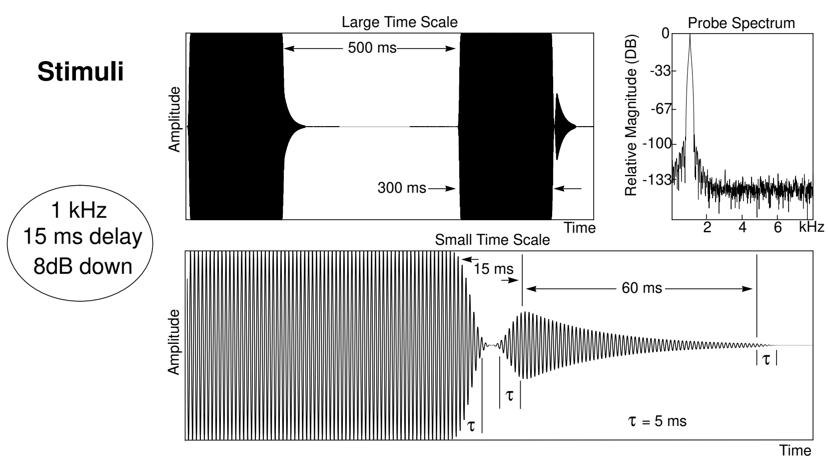


Forward Masking

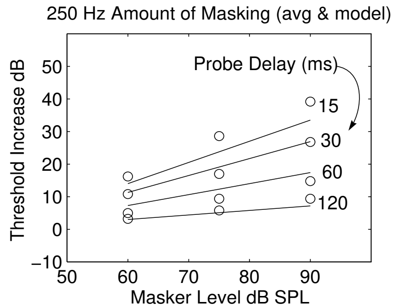
Pure-tone forward masking isolates per channel adaptation (without fine-time detail complications).
Wideband forward masking (click after noise) often used to maximize time domain resolution.
Jetsteadt et al (1982) summarizes pure-tone forward masking <u>only to 40 ms</u> probe delay.
Kidd and Feth (1982) characterize the effect of masker duration.

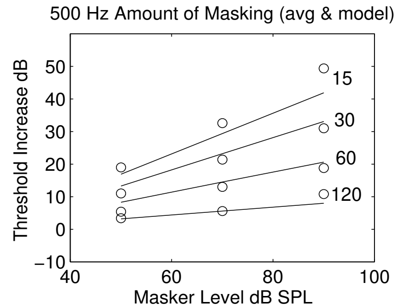
- **Extend Pure-Tone Forward Masking Data**

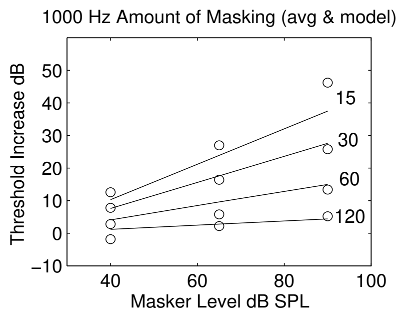
  Measurements across 5 octaves, 3 levels, and 4 delays (15, 30, 60, 120 ms).
  2-AFC, adaptive convergence to 79% correct (Levitt, 1971) for threshold.

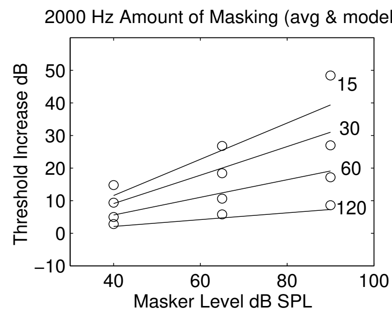


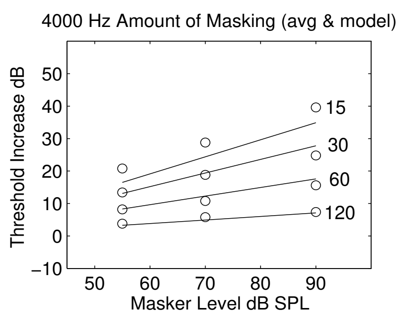
• Average Forward Masking Results (5 subjects)





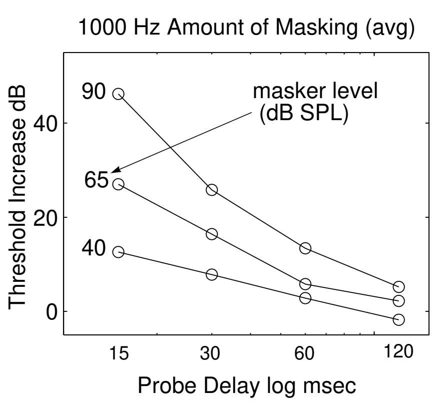


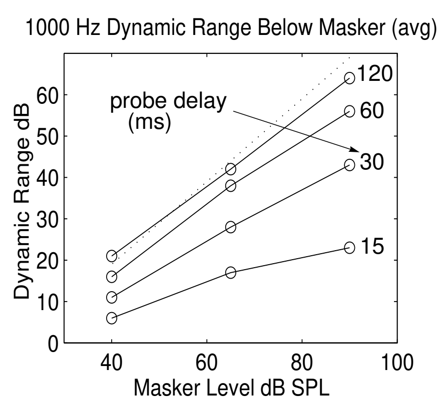




Amount of masking as a function of the masker level with contours of constant probe delay (circles), together with the model prediction of the amount of forward masking (lines), derived below.

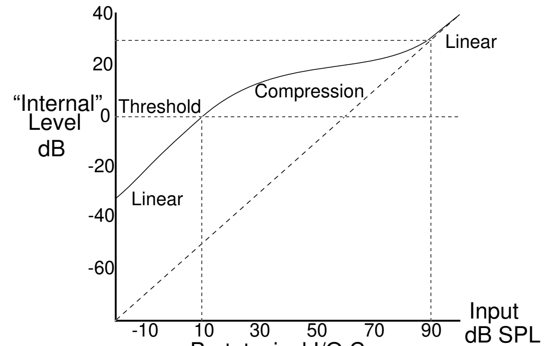
• Alternative Views of 1kHz Forward Masking Data



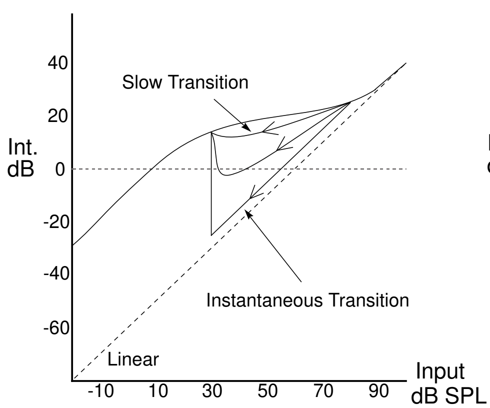


1) The rate of decay is greater for greater amounts of masking (Plomb, 1964), implying the amount of incremental adaptation should be greater for greater input changes.
2) The dynamic range below masker at short durations is still level-dependent, implying an initial estimate for the AGC I/O curve.

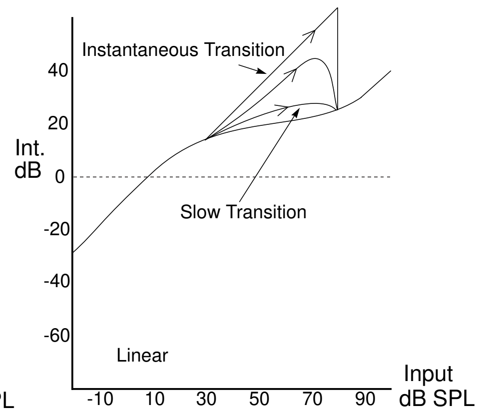
Adaptation Stage Functionality



Prototypical I/O Curve



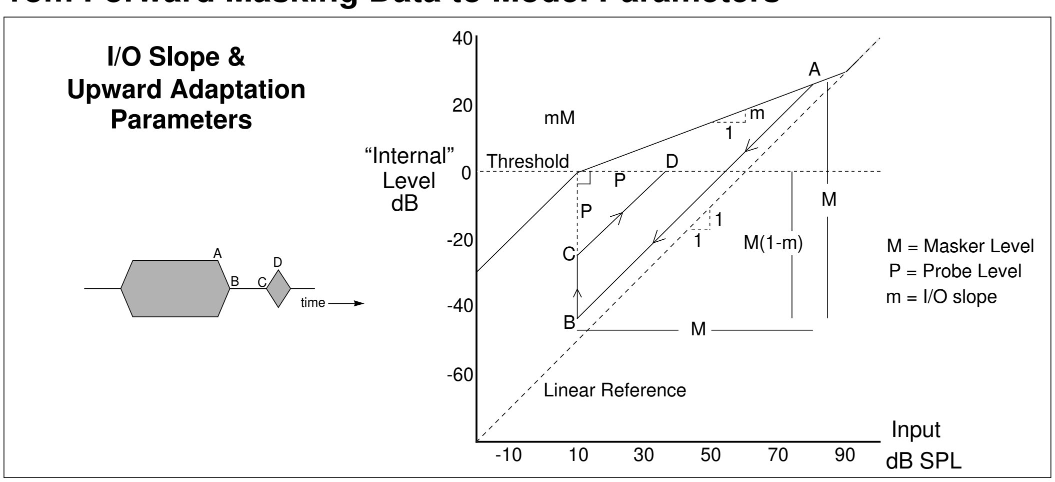
Response to Decreasing Input



Response to Increasing Input

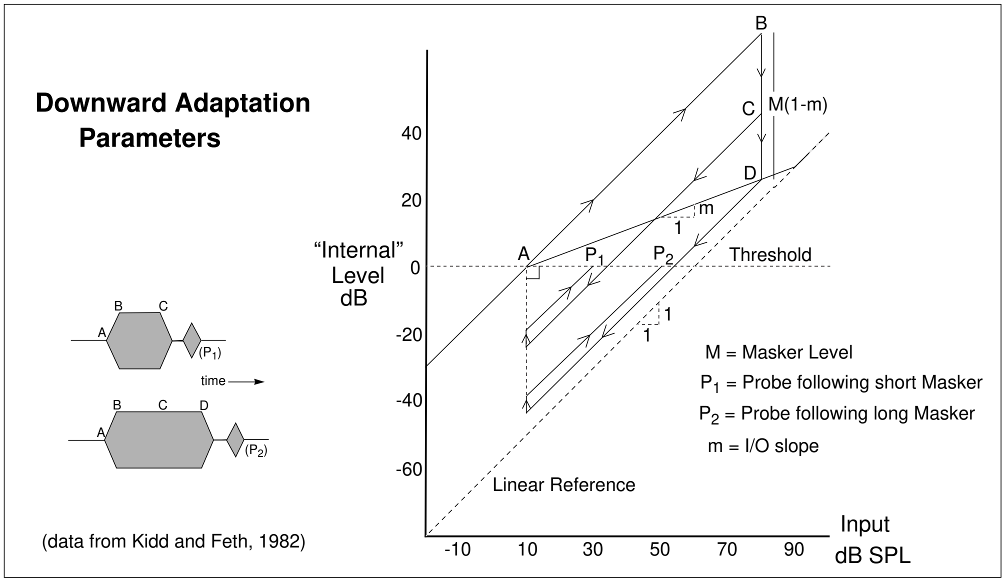
- **Statically approach the I/O target**
  - I/O shape motivated by outer hair cell motility (Johnstone, 1986).
  - Impose threshold at static audibility threshold.
  - High linear intercept at high level of equal loudness across (center) frequency.
  - Frequency-dependent loudness growth mechanism.
- ***Additive* adaptation (adjust a logarithmic offset to move toward an I/O target)**
  - Incremental auditory response is not proportional to adapting multiplier (Smith & Zwislocki, 1975).
  - Incrementally adjusting multiplicative gain does not lead to a level-dependent adaptation rate.

From Forward Masking Data to Model Parameters



$$P = M(1-m)\,a^{n}$$

a = upward incremental adaptation,
n = probe delay (relative to sample rate)



$$\Delta P = M(1-m)\, b^{nd} a^{nu}$$

b = downward incremental adaptation,
nu = probe delay,
nd = short masker duration

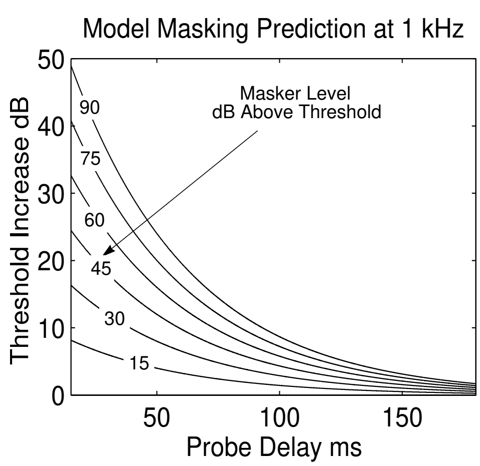
Model Parameters

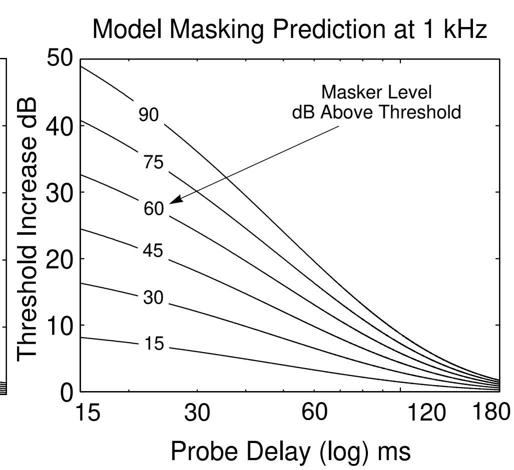
Freq. Hz	Slope m	a	b	release (ms)	attack (ms)
250	0.19	0.864	0.474	159	31
500	0.20	0.854	0.510	146	34
1000	0.26	0.816	0.543	109	45
2000	0.29	0.851	0.525	135	34
4000	0.34	0.858	0.507	139	31

Attack and release are the times for the output to settle within 2dB of I/O target after 25 dB input change (Dillon and Walker 1982). The model parameters a and b are with reference to a 100 Hz spectral sample rate (10 ms spectral sample interval).

Model Predictions

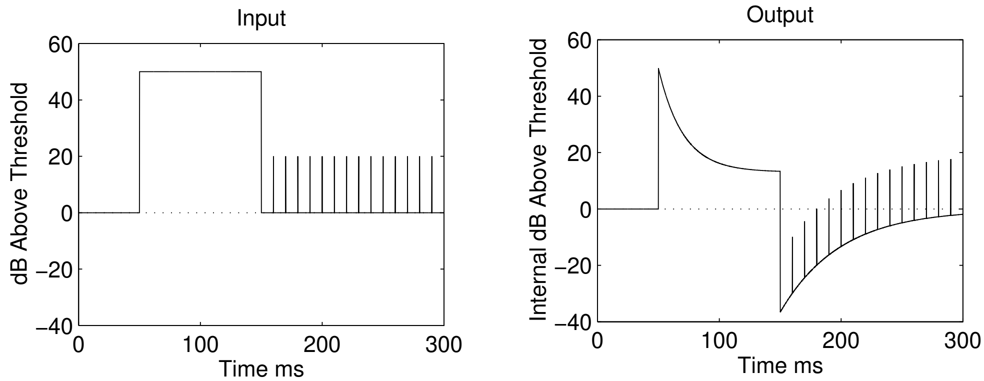
- **Rate of Decay of Forward Masking**





The model predicts a faster decay of forward masking with greater amounts of masking.

• **Specific Forward Masking Example**



Following a 100 ms tone, ideal impulse inputs to an adaptation stage are initially masked.

Bark Spectrogram Comparison

Model Applications

• **Hearing Aid Design**

Computational model of the dynamic perception of normal hearing (target performance).
Suggests an implementation strategy for dynamic hearing aids (post-log exponential adaptation).

• **Robust Speech Recognition and Coding (perception of spectral dynamics)**

Conclusion

• **Computational model using additive adaptation predicts forward masking which:**

1) Drops nearly linearly in dB with the log of the probe delay,
2) Decays faster with greater amounts of masking,
3) Increases as a fraction of the level of the masker,
4) Increases with greater masker duration, and
5) Is dependent on the frequency relation between masker and probe. (Moore summary, 1989)

• **Emphasizes onsets and transitions, and provides equal loudness mechanism.**

Future Work

• **Perceptually consistent prediction of dynamic relative loudness.**

• **Parameterization of multi-stage implementation for better model fit.**

• **Computational mechanism to model the auditory isolation of local spectral peaks.**

A First-Order Model of Dynamic Auditory Perception

Brian Strope and Abeer Alwan

Department of Electrical Engineering

University of California at Los Angeles

bps@ucla.edu

Supported in part by NIH-NIDCD Grant No. 1 R29 DC 02033-01A1.