Harvard-MIT Division of Health Sciences and Technology
HST.723: Neural Coding and Perception of Sound

Masking and Frequency Selectivity

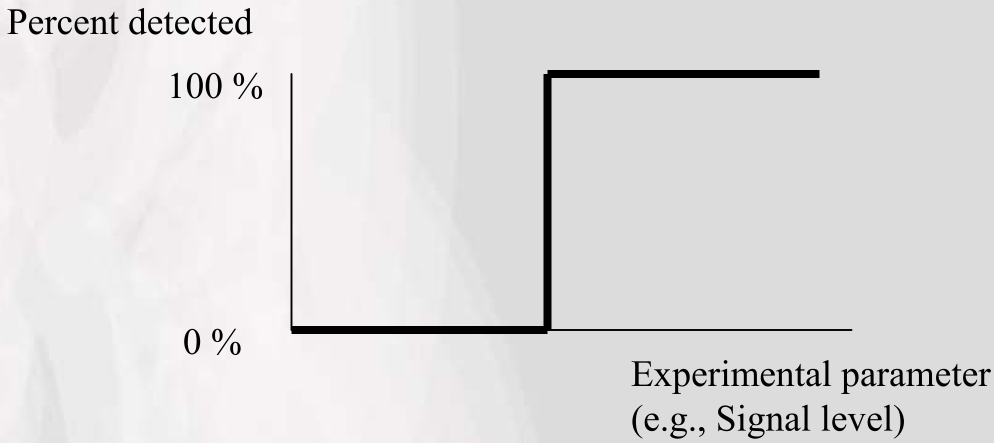
Defining Threshold

Historically two types of threshold:

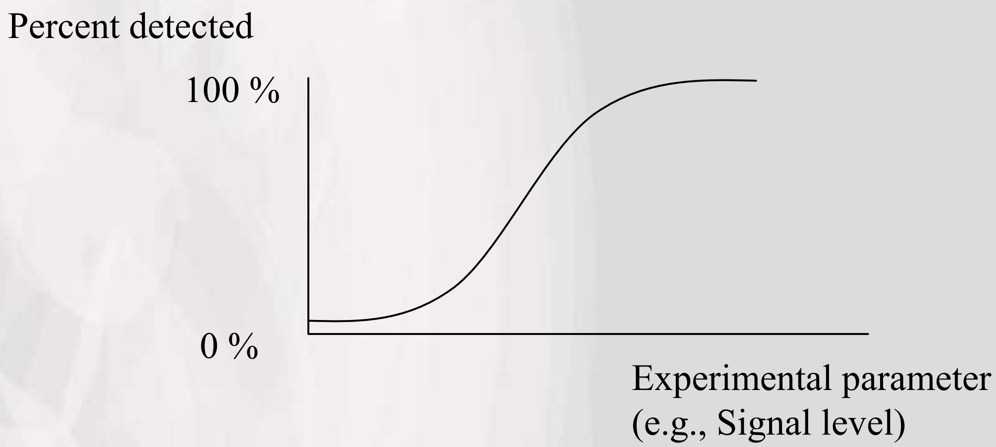
- Absolute threshold: Minimum audible signal
- Differential threshold: Minimum perceptible change, aka difference limen (DL) or just noticeable difference (jnd).

The Psychometric Function

a) Original concept of "threshold"



b) Experimental data



One-interval, two-alternative paradigms (yes-no)

Who is more sensitive?

Observer 1

	"yes"	"no"	Total
Signal	0.9	0.1	1.0
No Signal	0.9	0.1	1.0

Observer 2

	"yes"	"no"	Total
Signal	0.7	0.3	1.0
No Signal	0.1	0.9	1.0

a) Signal detection

	"Yes"	"No"
Signal	**Hits (H)**	Misses
No signal	**False alarms (F)**	Correct rejections

b) General formulation

	R2	R1
S2	**R2|S2**	R1|S2
S1	**R2|S1**	R1|S1

A good measure of sensitivity must take into account both hits and false alarms:

$$\textit{Sensitivity} = v[u(H) - u(F)]$$

Signal Detection Theory

The “internal response”, x, to a stimulus can be represented as a *random variable* (often assumed to have Gaussian (normal) distribution).

So, two identical stimuli will *not* necessarily result in identical percepts.

Detecting a signal (or discriminating between two stimuli) relies on deciding whether the percept arose from the distribution with mean M1 or the distribution with mean M2 (both have unit variance ($\sigma^2 = 1$).

- The perceptual distance between M1 and M2, in units of standard deviations, is called d’, pronounced “d-prime”.

$$d' = z(H) - z(F),$$

where z is the inverse of the normal distribution function.

- This implies there is no “threshold”.

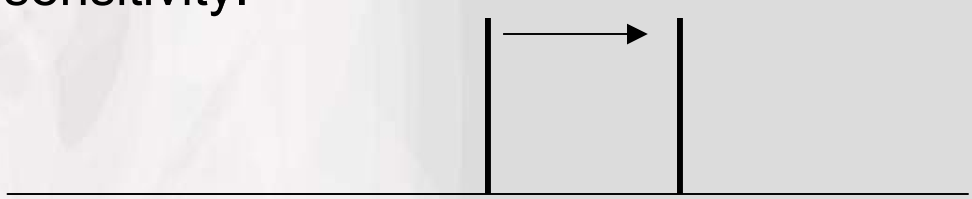
- The optimal rule is to set a criterion ‘C’:

Percept	*Response*
$x < C$	“No signal” (R1)
$x \geq C$	“Signal” (R2)

Where the criterion is set depends on:

- *a priori* probabilities of presentation
- Motivation and instructions (reward vs. punishment)

A change in the criterion does *not* mean a change in sensitivity.



Plotting various combinations of Hit rates and False Alarm rates for a given sensitivity results in a Receiver Operating Characteristic (**ROC**).

See Macmillan and Creelman (1991)

m-interval, m-alternative forced-choice experiments

One way to reduce the effects of bias is to present both types of stimulus on each trial. Generally, only one interval will contain the signal (S2), and the subject must select which interval it was.

Most popular is the 2-interval, 2-alternative forced-choice procedure (2AFC).

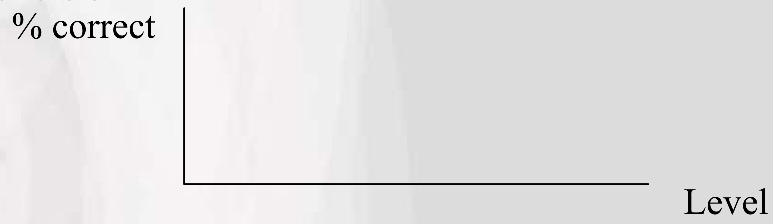
Each trial consists of either {S2 S1} or {S1 S2}, with *a priori* probability of 0.5 for each. Subjects respond '1' or '2' after each trial, depending on which interval contained S2.

Empirical results have generally shown only small biases in such experiments, meaning responses are generally symmetric. In this case, d' can be directly calculated simply from percent correct (see table).

A forced-choice paradigm does not rule out bias effects. Theoretically, it is preferable to record hits and false alarms. However, in practice most investigators only report percent correct.

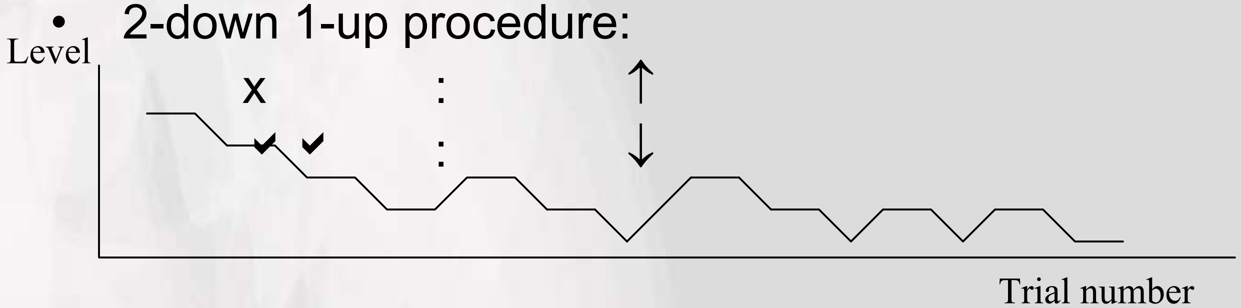
Adaptive Tracking Procedures

- Sometimes, it is necessary to find out how sensitivity (d') changes as a function of a stimulus parameter (e.g., signal level). In this case, a psychometric function can be generated by repeated measurements at a number of fixed values.



x-down y-up adaptive procedures converge on a fixed level of performance. This allows more flexible and rapid measurement of performance.

- 2-down 1-up procedure:



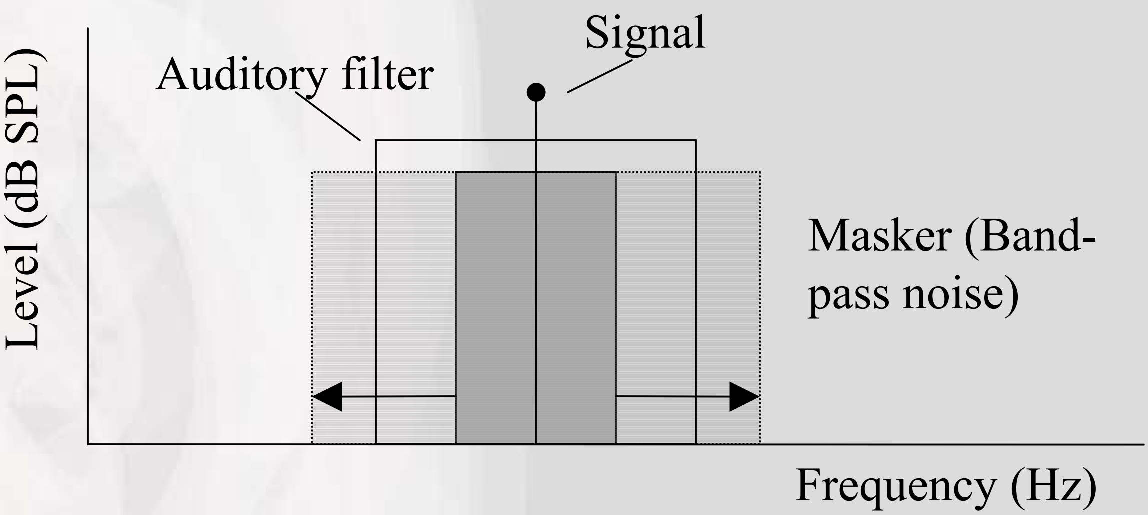
Tracks the 70.7% correct point on the psychometric function [$p(\checkmark \checkmark) = p(\checkmark)^2 = 0.5$].

Further reading:

Macmillan and Creelman (1991) "Detection Theory: A User's Guide." Cambridge Univ. Press. (Out of print)

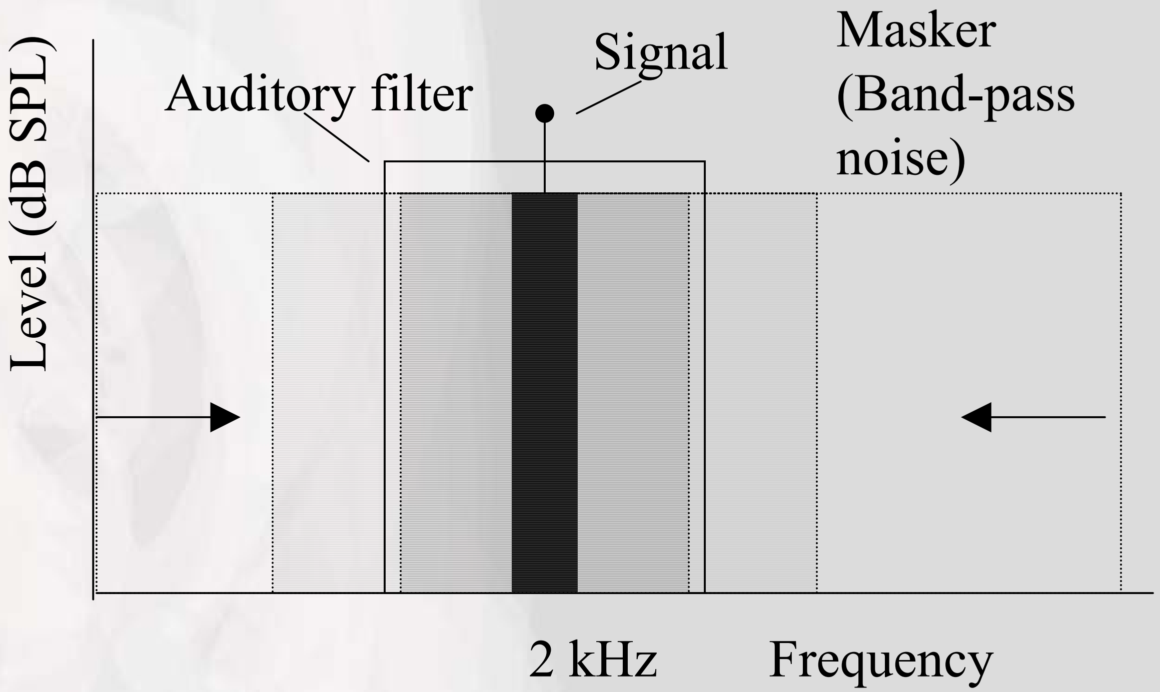
Green and Swets (1966) "Signal Detection Theory and Psychophysics," (Reprinted 1974 and 1989).

Fletcher's Band-widening Experiment (1940)



The transition point is known as the *Critical Band*. This has also been termed the *Equivalent Rectangular Bandwidth* (ERB).

Demonstration



Bandwidths:
(1) Broadband noise (0 - 20,000 Hz)
(2) 1000 Hz (1500 - 2500 Hz)
(3) 250 Hz (1875 - 2125 Hz)
(4) 10 Hz (1995 - 2005 Hz)

Spectral density* remains constant
Critical bandwidth at 2000 Hz is about 280 Hz
*Spectral density (spectrum level) is the power in a 1-Hz-wide filter.

Power Spectrum Model of Masking

A signal is detected by an increase in power at the output of the auditory filter centered at the signal frequency:

$$P_s = K\int_0^{\infty} W(f)N(f)\mathrm{d}f$$

where P_s is the power of the signal at threshold, W(f) is the filter shape, and N(f) is the masker's power spectrum. K is the detector "efficiency".

Assumptions:

- Filter is linear.
- Only one filter, centered at the signal frequency is used.
- Detection is based solely on overall power at filter output.

None of these assumptions are strictly true. However, they can often provide a reasonable first approximation.

Measures of Frequency Selectivity

Psychophysical Tuning Curves (PTCs)
Fixed signal; masker level adjusted to just mask signal.

See Moore, B.C.J (1997). An introduction to the psychology of hearing. 4th edition. Academic: London.

Advantages:
- Concept v. similar to neural tuning curves, allowing direct comparisons.

Potential problems:
- "Off-frequency listening"
- Detection of beats if using a sinusoidal masker

Notched Noise Method

Advantages:

- No influence of beats.
- Allows accurate measurement of filter “tails” (remote regions).
- Analysis can take into account off-frequency listening.

Origins of frequency selectivity

See Moore (1996)

Deriving Auditory Filter Shape

See Moore (1996)

ERB as a function of center frequency

See Glasberg & Moore (1990)

Masking Patterns

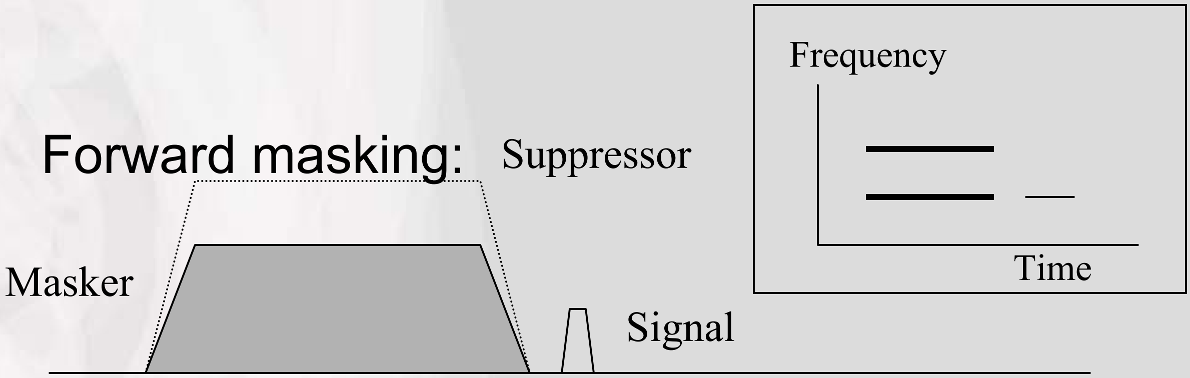
Problems with masking patterns:
- Off-frequency listening
- Detection of distortion products

See Egan and Hake (1950)

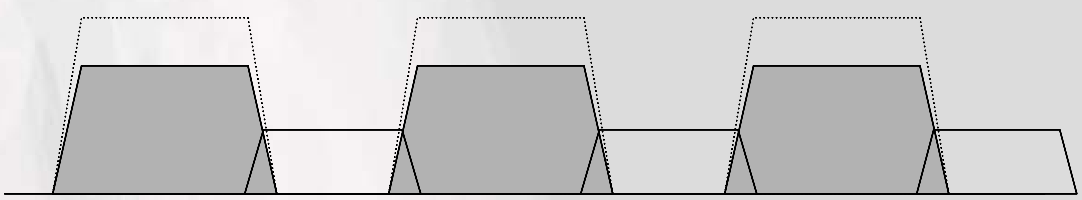
Suppression in Hearing

Houtgast pioneered the search for evidence of “lateral suppression” in psychoacoustic tasks.

- Forward masking:



- Pulsation threshold:



Using these techniques it is possible to show “two-tone suppression”. This is not possible with simultaneous masking, as the suppressor suppresses both the masker and the signal, giving zero net effect.

Example of Suppression Data

Effects of changing the suppressor frequency. Masker and probe are always at 1 kHz. (see Shannon, 1976)

Masking Patterns vs. Excitation Patterns

According to the power spectrum model, masking patterns and excitation patterns are essentially the same thing. But this is not true if masking is in part due to suppression.

See Oxenham and Plack (1998)

Failures of the Power Spectrum Model

- Temporal information (envelope or fine structure) can play an important role in detecting signals
  - Beats
  - Effects of masker modulation
  - Detection of tones in roving-level narrowband noise

- Nonlinearities, such as suppression and distortion products, are not accounted for.

- Informational masking (Neff and Green, 1987) can produce large amounts of masking without any energy around the signal frequency.

Nevertheless, the model has proven to be a very useful tool for both auditory theory and in many practical applications (perceptual audio coding; predicting speech intelligibility).