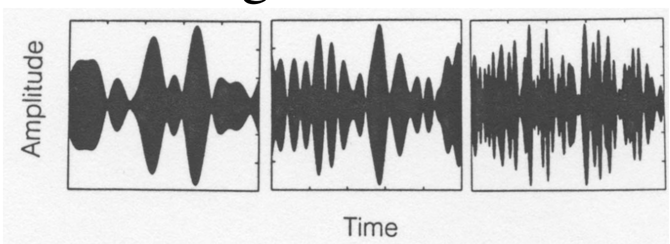
Temporal resolution

The ability to follow rapid changes in a sound over time

The bottom line

People manage to maintain good temporal resolution without compromising sensitivity by using intelligent processing.

Temporal resolution: How good is a listener at following rapid changes in a sound?

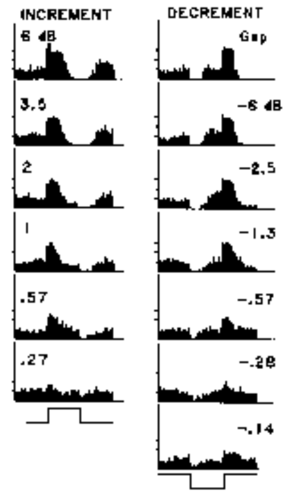


- Auditory nerve fibers do not fire at the instant at which sounds begin or end.
- Auditory nerve fibers do not fire on every cycle of sound.
- Adaptation occurs to longer duration sounds.
- Spontaneous activity occurs when no sound is present

Several characteristics of the auditory nerve response will limit the fidelity with which fluctuations in a sound can be represented.

Following rapid changes in sound

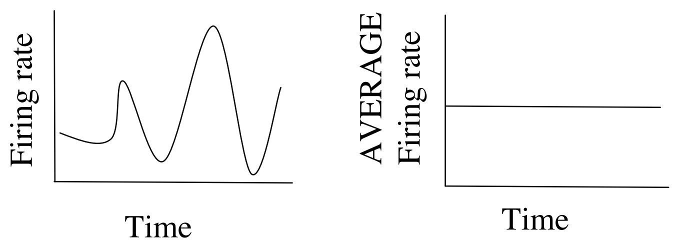
The auditory nerve response does not follow changes with perfect precision

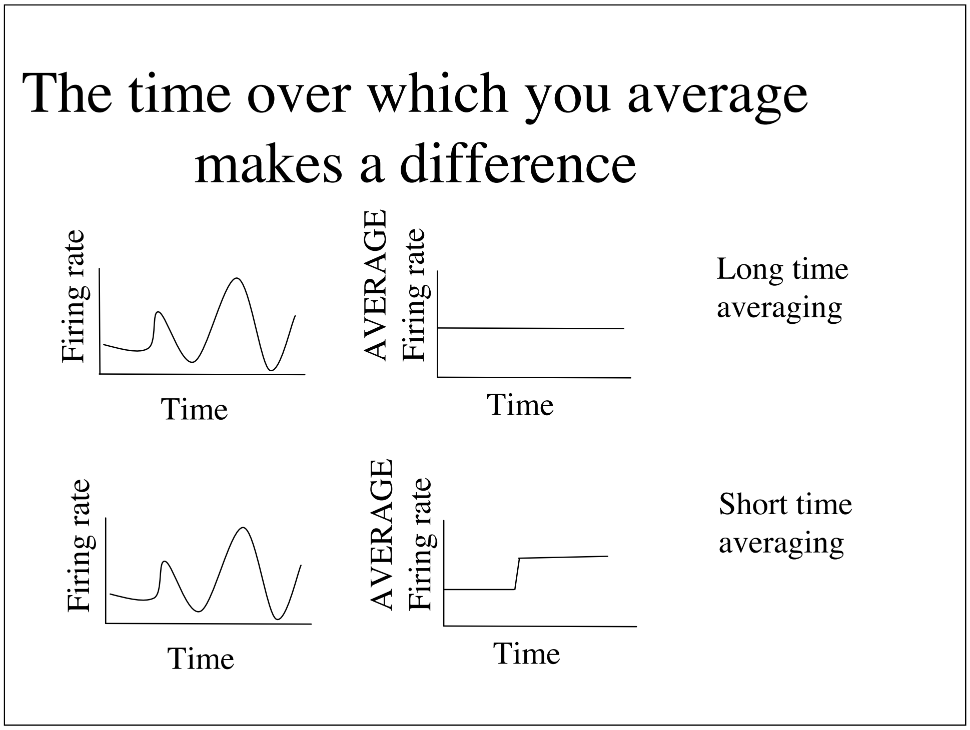


In the PST histograms on the right, the auditory nerve response does not exactly follow the clean increment or decrement in sound that is shown at the bottom of the panel.

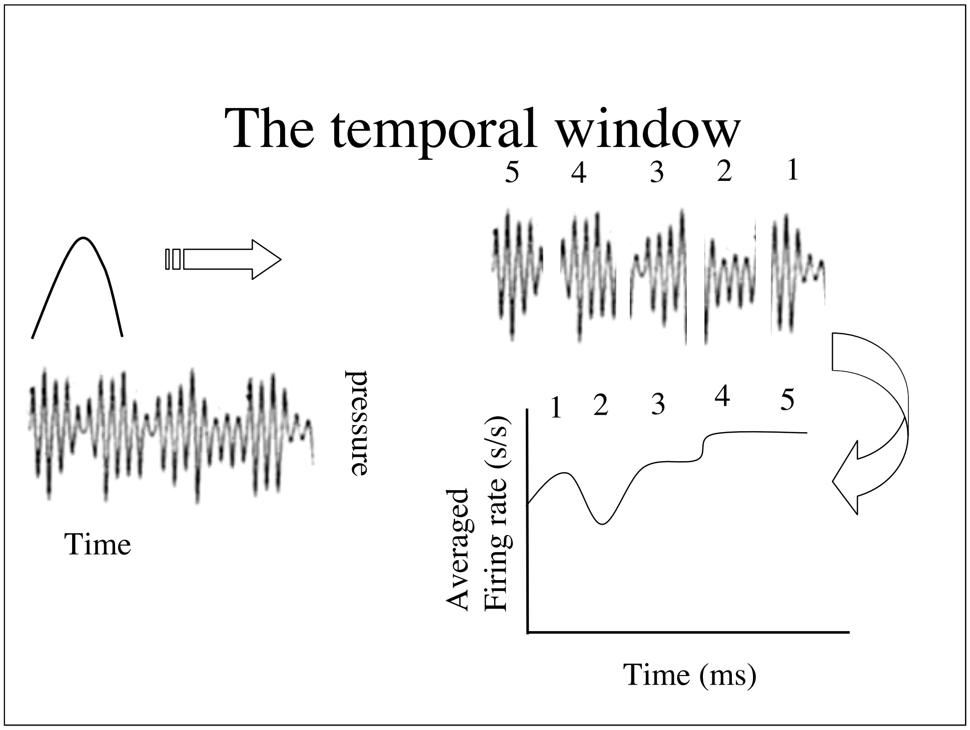
The auditory system can’t really respond fast enough to capture little fluctuations, and even if it could, a lot of those little fluctuations in the auditory nerve response would be misleading.

Averaging over time is one way the auditory system could "smooth out" the bumpy response of auditory nerve fibers



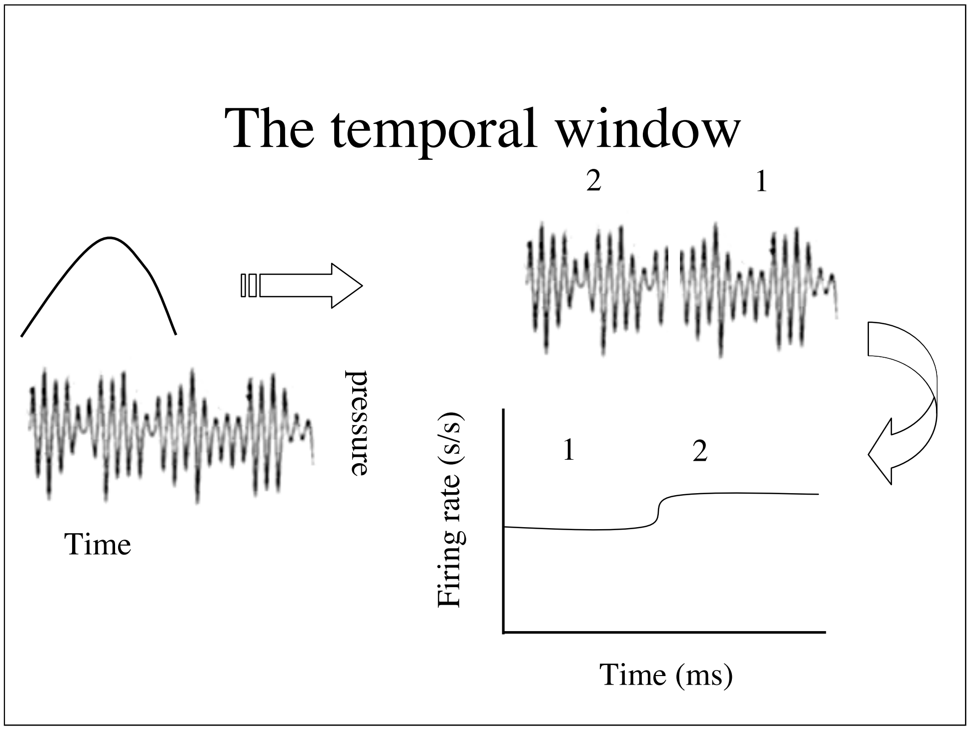


If you average over too long a time, you won't know when the intensity has changed. But if you average over too short a time, you will be fooled into thinking that an onset response or a recovery period represents a change in intensity.



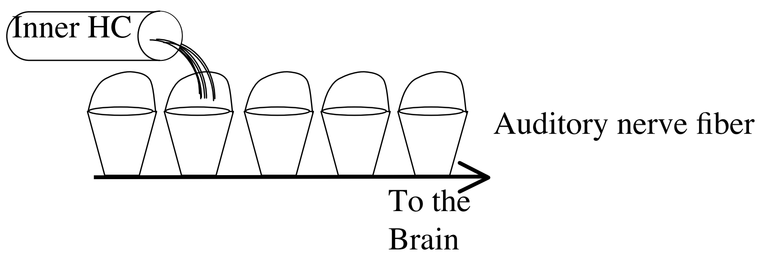
We refer to the time over which the auditory system averages firing rate the “temporal window”. As the sound passes through the temporal window, all of the response that fits into the window at one time gets averaged. I’ve shown this in terms of the time waveform of a sound, but remember it is really the firing rate of auditory nerve fibers that is getting averaged.

Notice that the average firing rate “captures” the decrease in intensity in the second “window”, but it misses some of the shorter decreases in the intensity that occur in the last 3 windows.



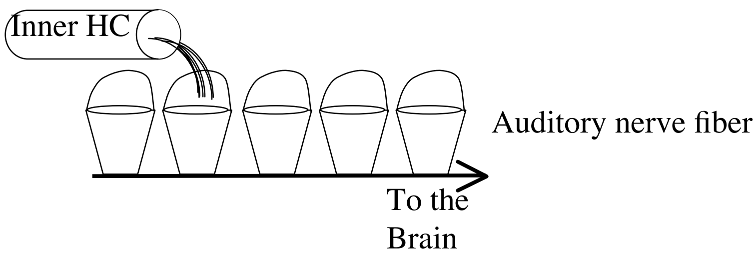
If the temporal window is longer, then we see even less of the details about the changes in the sound over time.

Hydraulic analogy: How long before the next bucket leaves for the brain?



One way to think of this: The auditory nerve is delivering the message from the cochlea to the brain in “chunks”, like the buckets in a bucket brigade. Each bucket is under the spigot for a certain period of time before it proceeds. Whatever message (“water”) collects in the bucket during that time is what the brain has to tell it what the sound amplitude was during that time period.

Hydraulic analogy: How long before the next bucket leaves for the brain?



It could go fast.

People can “add up” sound energy for

(A) 5 ms

(B) 50 ms

(C) 200 ms

(D) 1500 ms

Temporal resolution: How short are the "samples" of sound?

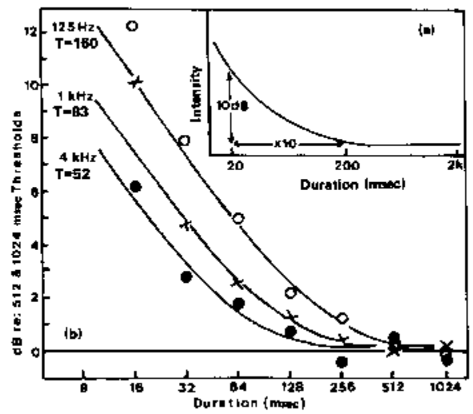


Figure 9.3 (a) Idealized temporal integration function showing that a decade change in duration is offset by an intensity change of about 10 dB up to 200–300 msec. Threshold remains essentially constant at longer durations. (b) Effect of frequency on temporal integration. (Adapted from Watson and Gengel [52], with Permission of *J. Acoust. Soc. Amer.*)

Hypothesis # 1:
We integrate
over 200-300 ms.

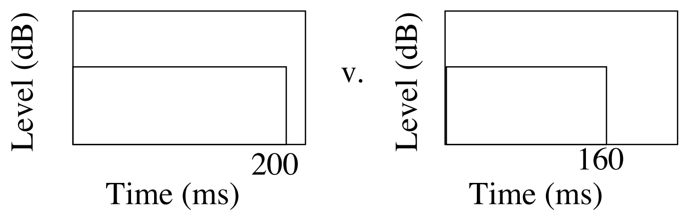
From Gelfand (1997)

We already know something that might tell us about the size of the temporal chunk: We know that absolute threshold improves with duration up to 200-300 ms. The system integrates information over that time period. So that would suggest that the "buckets collect water" for 200 or 300 ms and that maybe we could tell that there was a change in the timing if it was 50-75 ms long because that be like a 1 dB change in overall intensity.

Sensitivity-resolution tradeoff

If you extend the integration time to improve sensitivity, you lose resolution.

So how well should I be able to discriminate a change in the duration of a sound?



40 ms difference ~ 1 dB--like the jnd

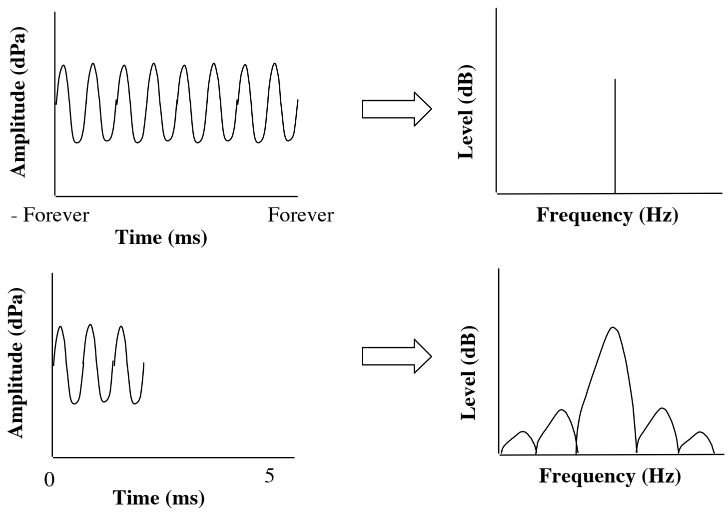
14

How to measure temporal resolution

- Duration discrimination
- Gap detection
- Amplitude modulation detection

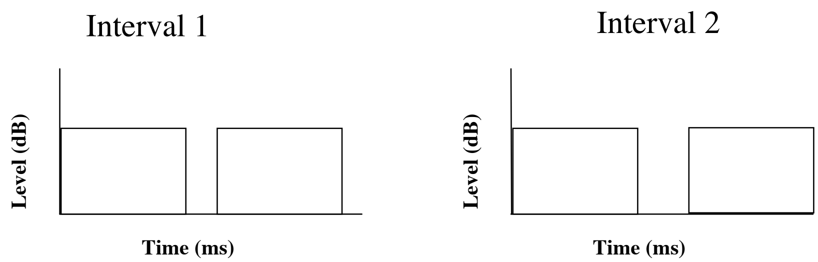
To test this hypothesis, we test people's ability to detect for short changes in a sound. These are the methods typically used.

Problem in measuring temporal resolution: "Spectral splatter"



A major problem with testing temporal resolution is that when we make temporal changes in a frequency-specific sound, we also make spectral changes. For example, a long duration tone has a single peak in its spectrum, but when we make the same sound short in duration by turning it on and off abruptly, sound energy spreads to adjacent frequencies . This phenomenon is called "spectral splatter". The problem is that a listener could tell, for example, that the duration of the sound changed by detecting the change in the spectrum, not the time waveform.

Duration discrimination



Which gap was longer?

In duration discrimination, the subject chooses the interval with the longer gap in it.

Duration discrimination

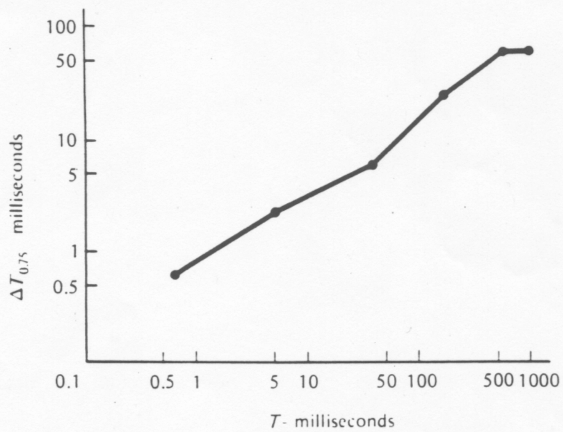


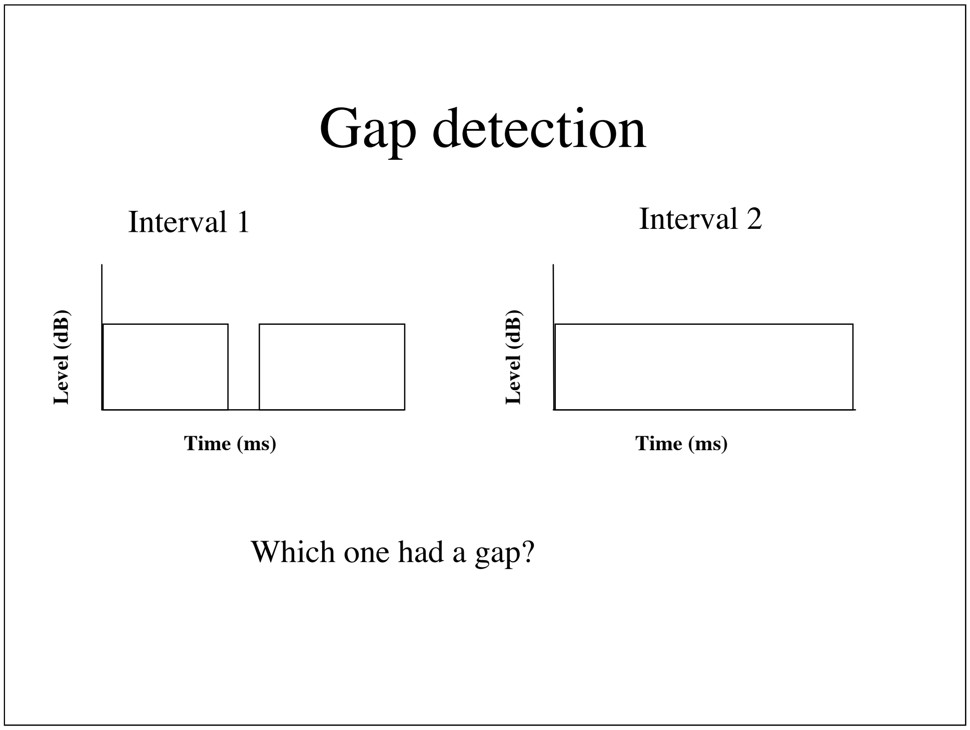
Figure 10.9 Value of $\Delta T_{0.75}$ (change in temporal separation between two markers) required for a $P(C)$ of 75 percent as a function of T (standard separation between two markers).
Adapted from Abel, 1972, by permission

- Weber’s Law? NO
- Duration discrimination can be very acute - much better than 50-75 ms.

From Yost (1994)

Duration discrimination thresholds go up as the duration goes up, but not at a constant rate as Weber’s Law predicts.

Notice that people can discriminate very short differences in duration, much shorter than their performance in temporal integration would lead us to believe.



In gap detection, the listener chooses the interval with the gap in it.

Gap detection

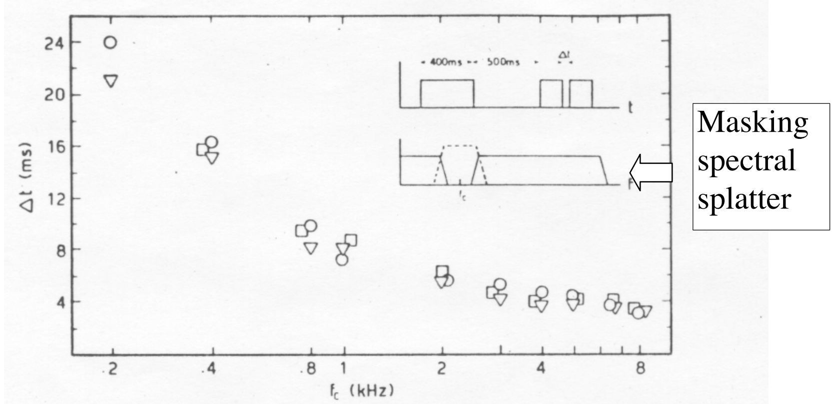
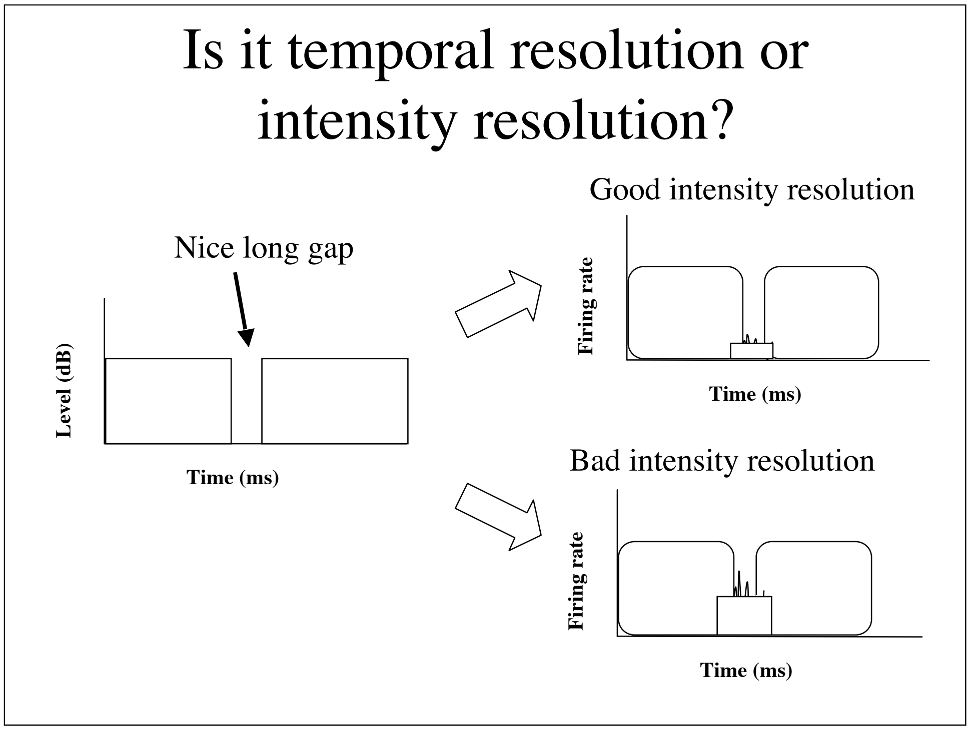


FIG. 1. Gap thresholds for the three subjects in experiment I plotted as a function of the center frequency f_c of the signal, which had a spectrum level of 25 dB. The 0.2-kHz data are from experiment II. The inset shows schematized stimulus configurations for time and frequency. The notched background masker was continuous, so the signal is shown dotted in the spectral characteristics. Details are given in the text. (MS, ○; EL, ▽; GW, □).

From Moore (1997)

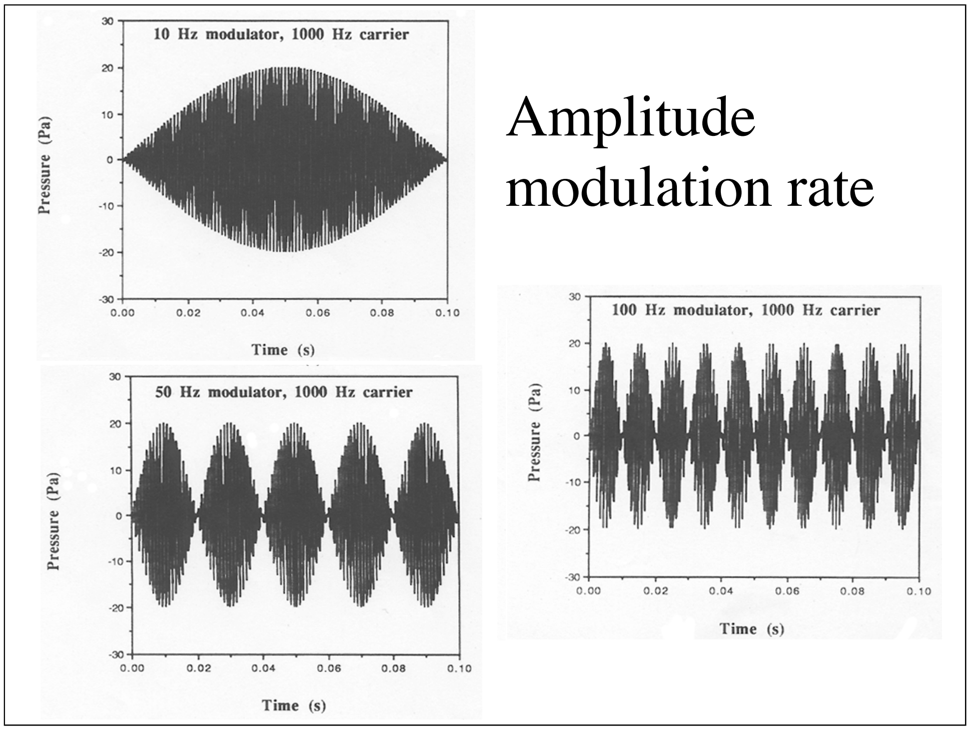
Moore et al. measured gap detection threshold in noise bands centered at various frequencies. To mask spectral splatter they played the noise band in a continuous background of notched noise. Gap detection threshold improved from about 22 ms at 200 Hz to about 3 ms at 8000 Hz-- certainly much better than 50-75 ms.



But even if the gap is long enough that we can easily detect it, it will be harder to detect a gap if the auditory system doesn't respond to the change in intensity very much. In other words, intensity resolution will also influence how well we detect gaps. How can we separate temporal and intensity resolution?

Amplitude modulation detection

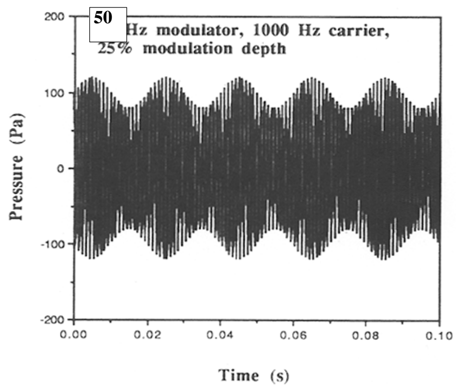
By how much do I have to modulate the amplitude of the sound for the listener to tell that it is amplitude modulated, at different rates of modulation?



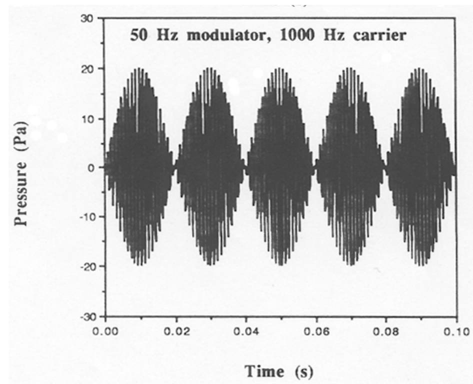
By AM rate is meant how often the amplitude of the sound goes from minimum to maximum to minimum in a second. These graphs represent a 1000-Hz carrier tone modulated at 3 different rates.

Modulation depth

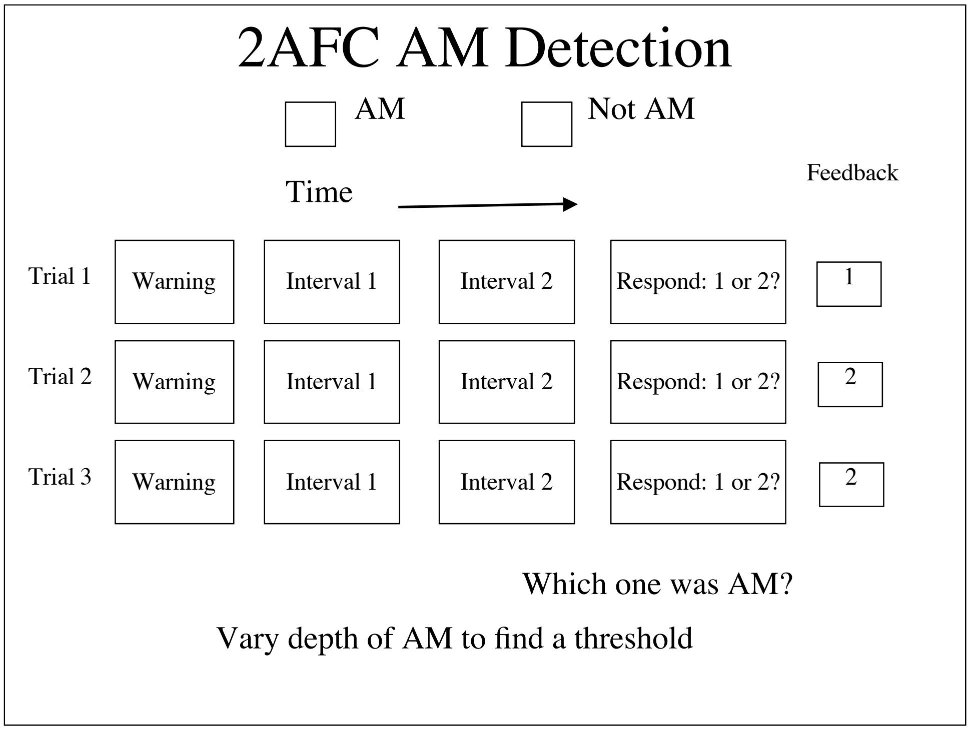
25%



100%



The minimum amplitude doesn't have to be zero. The modulation depth is the percentage by which the amplitude changes between the maximum and the minimum.



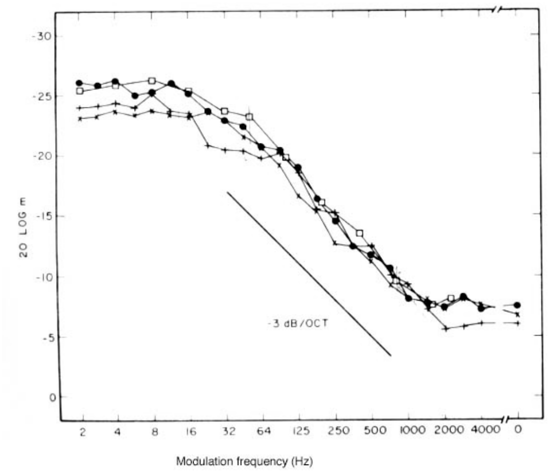
The task in AM detection is illustrated here. The listener gets a warning followed by two interval indicators. During one interval, randomly chosen, an AM sound is presented; in the other interval the same sound but unmodulated is presented. The listener chooses the interval containing the AM sound and receives feedback. The modulation depth is varied to find the threshold for AM detection, the amount of modulation required to hear the modulation some proportion of the time.

Modulation depth,
20 log m

m (%)	20 log m (dB)
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	-[illegible]
[illegible]	[illegible]

Modulation depth is usually expressed in deciBels, calculated as 20 times the log of the percent modulation.

AM detection as a function of modulation rate



The temporal modulation transfer function (TMTF)

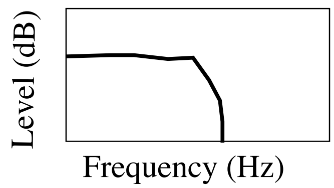
From Viemeister (1979)

The plot of AM detection thresholds as a function of modulation rate is called the temporal modulation transfer function. Threshold is plotted as modulation depth in dB with low (good) values at the top. The carrier in this case was a broadband noise. Each curve in the graph represents the data of one listener.

Notice that for modulation rates below 50-60 Hz, thresholds are about -23 to -26 dB, about 5% modulation depth. Above 50-60 Hz thresholds get worse. At 500 Hz modulation rate, something like 50% modulation is needed before we hear it.

So the auditory system acts like a low-pass filter when it comes to temporal fluctuations in sounds-- with a cutoff frequency (3-dB down point) around 50-60 Hz. We can follow fluctuations up that modulation rate fine, but at higher rates we don’t follow as well.

What sort of filter has a response that looks like this?



(A) low-pass

(B) high-pass

(C) bandpass

(D) band reject

The TMTF is like a low-pass filter. That means that we can't hear

(A) slow amplitude modulations

(B) high frequencies

(C) low frequencies

(D) fast amplitude modulations

29

TMTF at different carrier frequencies

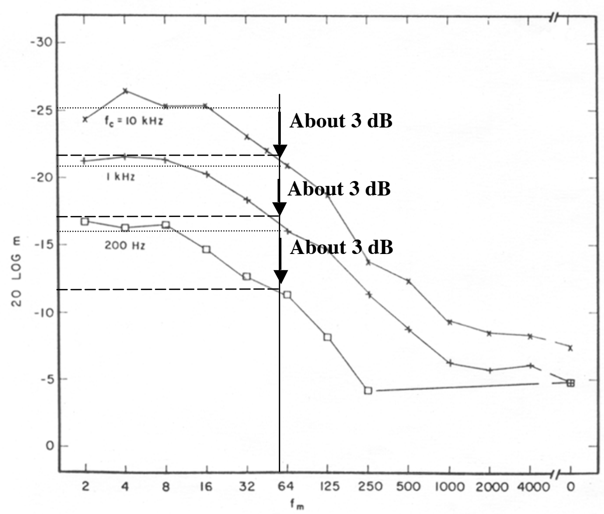


FIG. 4. TMTFs for different spectral regions of modulation. Broadband modulated noise was bandlimited after modulation to a frequency region centered at f_c.

From Viemeister (1979)

If you use a band of noise centered at some frequency as the carrier, sensitivity is better at higher frequencies, but the shape of the TMTF (the 3 dB down point) is the same at all frequencies. That would suggest that temporal resolution does not depend on frequency.

Conclusions from TMTF

- People are very good at AM detection up to 50-60 Hz modulation rate (and intensity resolution effects are controlled)
- 50-60 Hz = 17-20 ms/cycle of modulation
- 17-20 ms < 40 ms
- Somehow the auditory system is getting around the sensitivity-resolution tradeoff

The auditory system can follow amplitude modulation well up to about

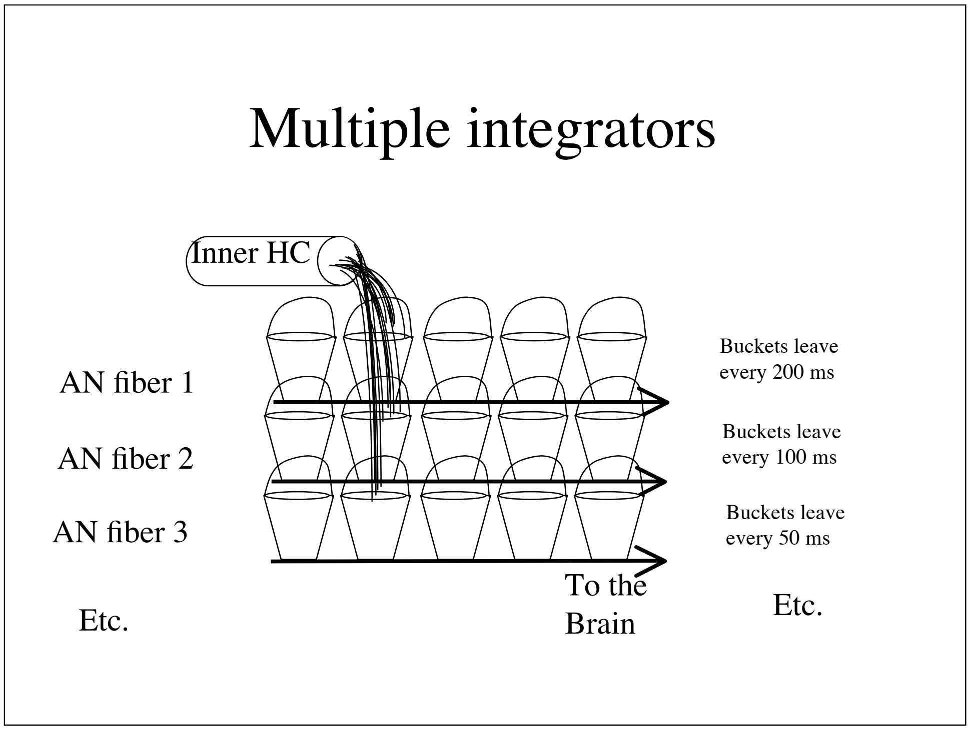
(A) 50-60 Hz
(B) 120 Hz
(C) 4 Hz
(D) 2000 Hz

32

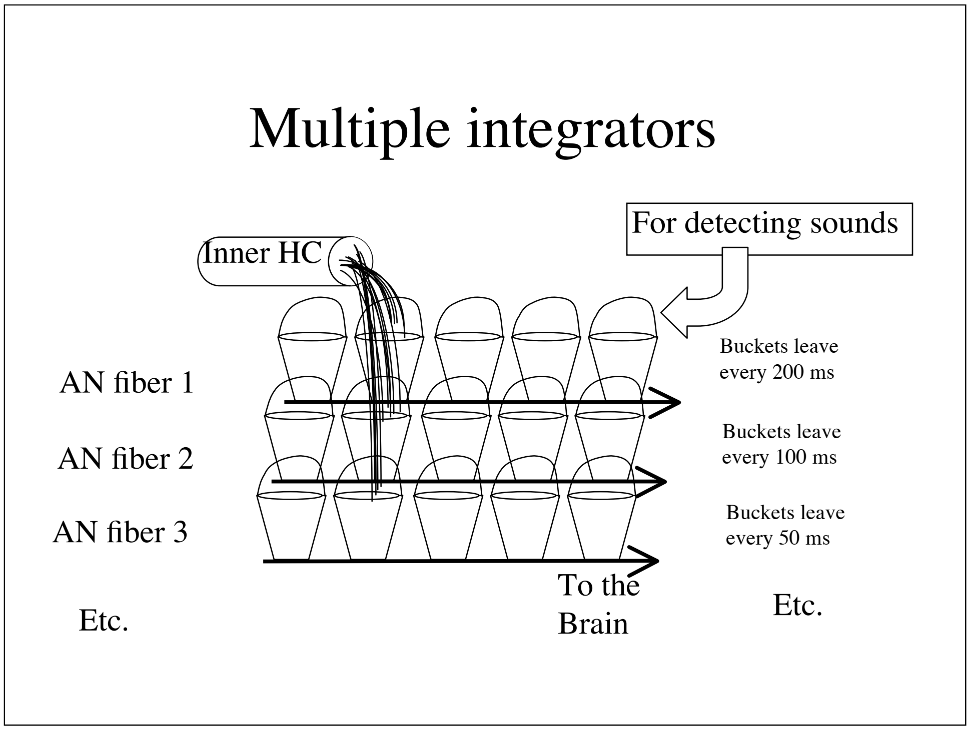
So how can we detect such short changes in a sound and still be able to integrate sound energy over 200-300 ms?

Two theories of temporal resolution-temporal integration discrepancy

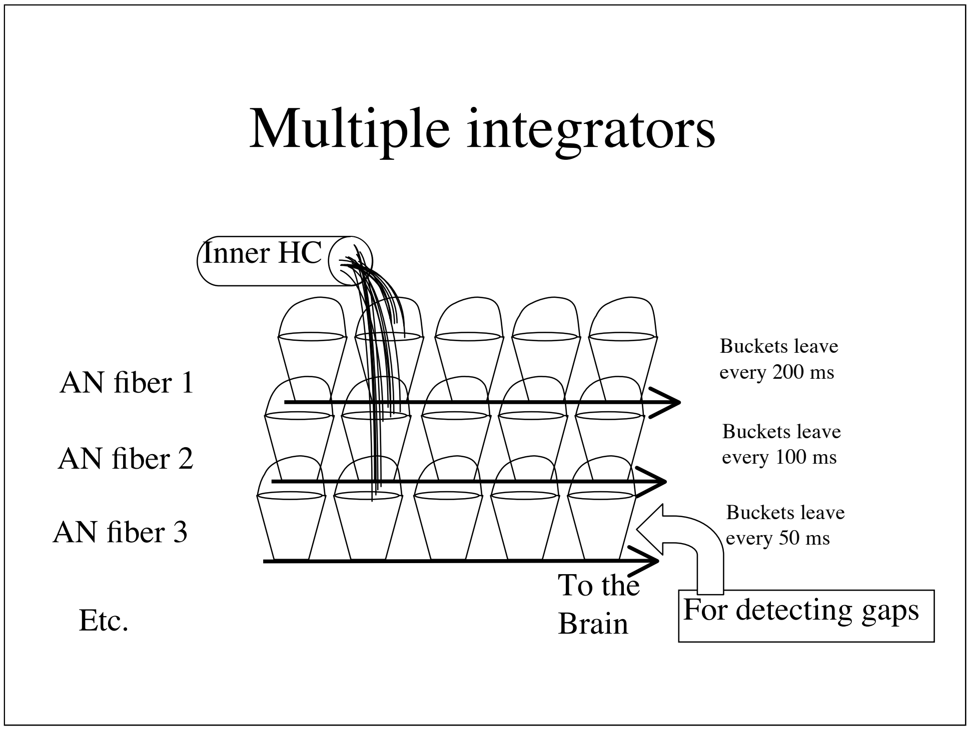
- Multiple integrators
- Multiple looks



The multiple integrator model says that different nerve fibers have different integration times, some long, some short.



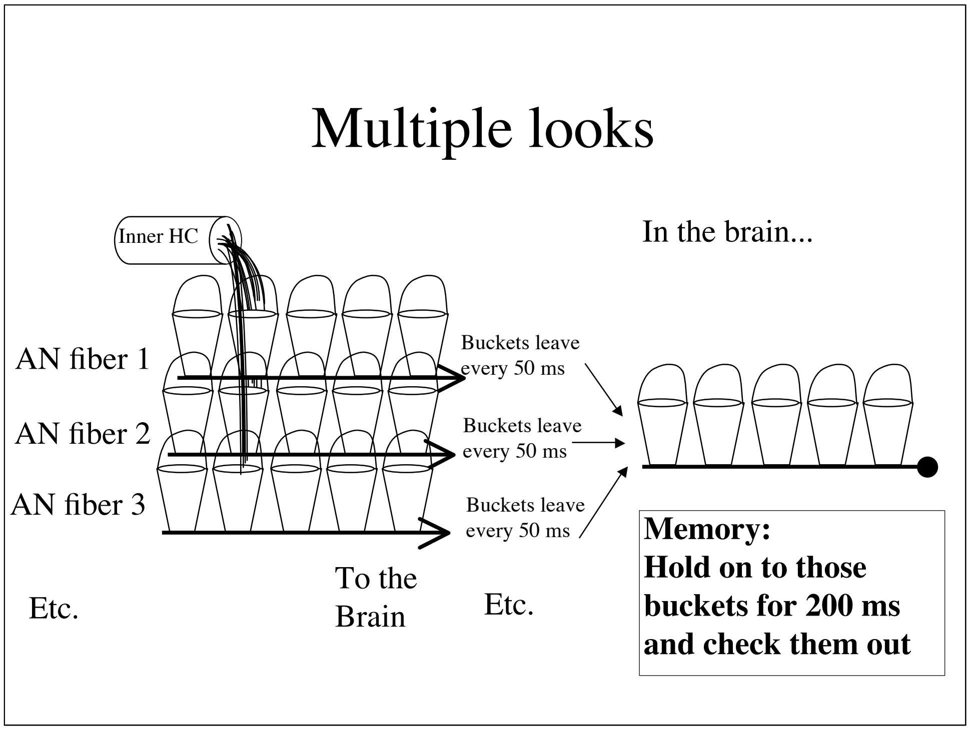
If you are trying to detect a sound, you "listen" to the fibers with long integration times, so you can add up as much sound energy as possible.



If you want to detect a short gap, you "listen" to the fibers with short integration times, so you can get as much detail as possible about how the intensity changed over time.

AN fibers don’t have different integration times

But of course the integrators could be somewhere else in the brain.



The multiple looks model holds that all of the nerve fibers have similar integration times, but that at some point in the brain, a kind of memory operates. The outputs of the nerve fibers (the buckets) are stored for 200 or 300 ms, while the listener scans their contents. If you are trying to detect a sound, you add up the contents of all the buckets. If you are trying to detect a gap, you look in the bucket where the gap was supposed to have occurred to so if there is less of a response in it.The buckets in our analogy are referred to as “looks”.

Multiple looks theory says

(A) we have good temporal resolution because we use memory to integrate sound "energy"

(B) we have good temporal resolution because we have some neurons that have good temporal resolution and some neurons that don't.

Multiple integrators theory says

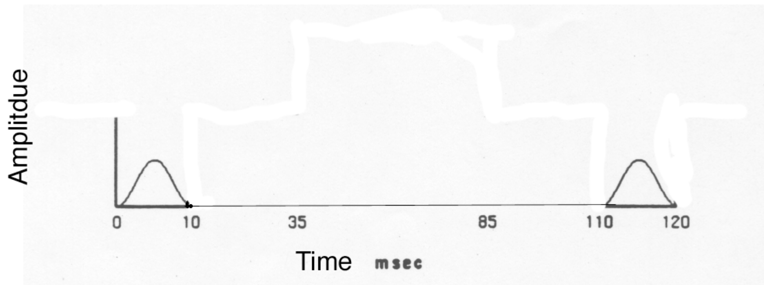
(A) we have good temporal resolution because we use memory to integrate sound "energy"

(B) we have good temporal resolution because we have some neurons that have good temporal resolution and some neurons that don't.

A test of the multiple looks theory: Viemeister & Wakefield (1991)

Set up a situation in which the two theories predict different outcomes...

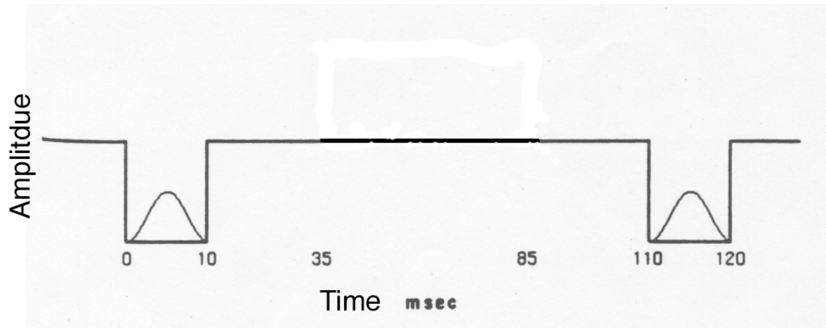
Viemeister & Wakefield (1991)



It would be useful to integrate the 2 tone pips to improve detection, and both theories say you could do that.

Viemeister & Wakefield asked listeners to detect two short tone pips-- when each was presented alone and when both were presented in sequence. You would expect people to be able to detect the two tones better than they do either tone alone by any theory of temporal processing.

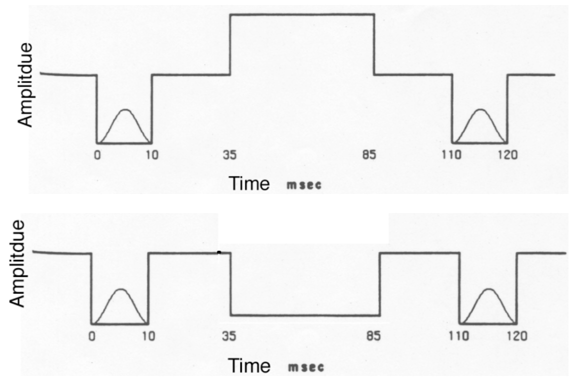
Viemeister & Wakefield (1991)



But if you put noise on between the tone pips, you can't integrate them without integrating in the noise. If you're taking short looks, you can use the looks with the tone pips, but ignore the looks in between.

Viemeister and Wakefield also presented the tones when there was a noise background that stopped just when the tones were presented, but was on in the time between the tones. The multiple integrator theory says that you would not be able to hear two tones better than one in this situation because to add the energy of the two tones you need an integrator that integrates for 120 ms, but that integrator will also have to add in the noise between the tones, making it harder to hear them. Multiple looks theory says that it shouldn't matter if there is noise between the tones.

Viemeister & Wakefield (1991)



Multiple integrator "performance" will get worse if the noise goes up more, and better if the noise goes down some, but multiple looks are not affected by what happens between the tone pips.

The top panel shows another condition in which the noise between the tones goes up in intensity by 6 dB, the bottom panel shows the case where the noise goes down by 6 dB between the tones. Multiple integrator theory says that the changes in the noise level will affect detection of the two tones, but multiple looks theory says it won't matter.

Viemeister & Wakefield (1991): Results

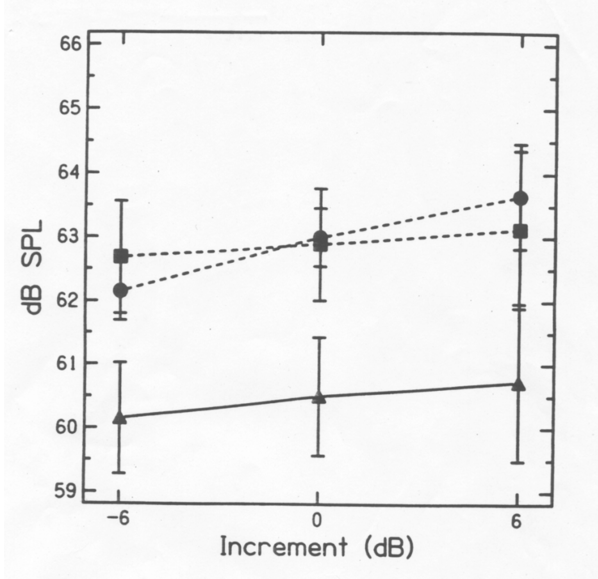


FIG. 3. Average thresholds as a function of the relative level of the intervening noise. Squares: first pulse only; circles: second pulse only; triangles: pulse pair. In quiet, the average thresholds are 26.6 dB for one pulse and 25.2 dB for the pulse pair.

This graph shows the results of the experiment. The dashed lines show the thresholds when just one of the tone (1st or 2nd) is presented; the solid curve shows thresholds when the tones are presented in sequence. On the x-axis is the level of the noise between the tones in the three conditions. Thresholds were better when there were two tones instead of one, and the level of the noise doesn't make much difference. This result supports the multiple looks theory and argues against the multiple integrators theory,

The results of Viemeister & Wakefield are most consistent with

(A) multiple looks theory

(B) multiple integrators theory

47

Conclusions

- People can detect very short duration changes in sound, such as 2-3 ms long interruptions.
- People can integrate sound energy over 200-300 ms to improve sound detection.
- The auditory system gets around the sensitivity-resolution tradeoff by using short-term integration and intelligent central processing.

Text sources

- Gelfand, S.A. (1998) Hearing: An introduction to psychological and physiological acoustics. New York: Marcel Dekker.
- Moore, B.C.J. (1997) An introduction to the psychology of hearing. (4th Edition) San Diego: Academic Press.
- Viemeister, N.F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *J. Acoust. Soc. Am., 66*, 1564-1380.
- Viemeister, N.F. & Wakefield, G. (1991) Temporal integration and multiple looks. *J. Acoust. Soc. Am., 90*, 858-865.
- Yost, W.A. (1994) Fundamentals of hearing: an introduction. San Diego: Academic Press.

Text sources

- Gelfand, S.A. (1998) Hearing: An introduction to psychological and physiological acoustics. New York: Marcel Dekker.
- Moore, B.C.J. (1997) An introduction to the psychology of hearing. (4th Edition) San Diego: Academic Press.
- Viemeister, N.F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *J. Acoust. Soc. Am., 66*, 1564-1380.
- Viemeister, N.F. & Wakefield, G. (1991) Temporal integration and multiple looks. *J. Acoust. Soc. Am., 90*, 858-865.
- Yost, W.A. (1994) Fundamentals of hearing: an introduction. San Diego: Academic Press.