CENTRO DE INVESTIGACIÓN EN TECNOLOGÍAS DE AUDIO
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

RESOLUCIÓN TEMPORAL

Rodrigo F. Cádiz
Octubre 2011

Resolución temporal e integración

Resolución temporal

“resolución”

La precisión mediante la cual un estímulo puede ser distinguido de otros similares en alguna dimensión

“resolución temporal

La habilidad de separar un sonido de otro en el *tiempo*

http://depts.washington.edu/sphsc461/freq_res/freq_res_files/freq_res.ppt

Resolución temporal

Medidas de resolución temporal

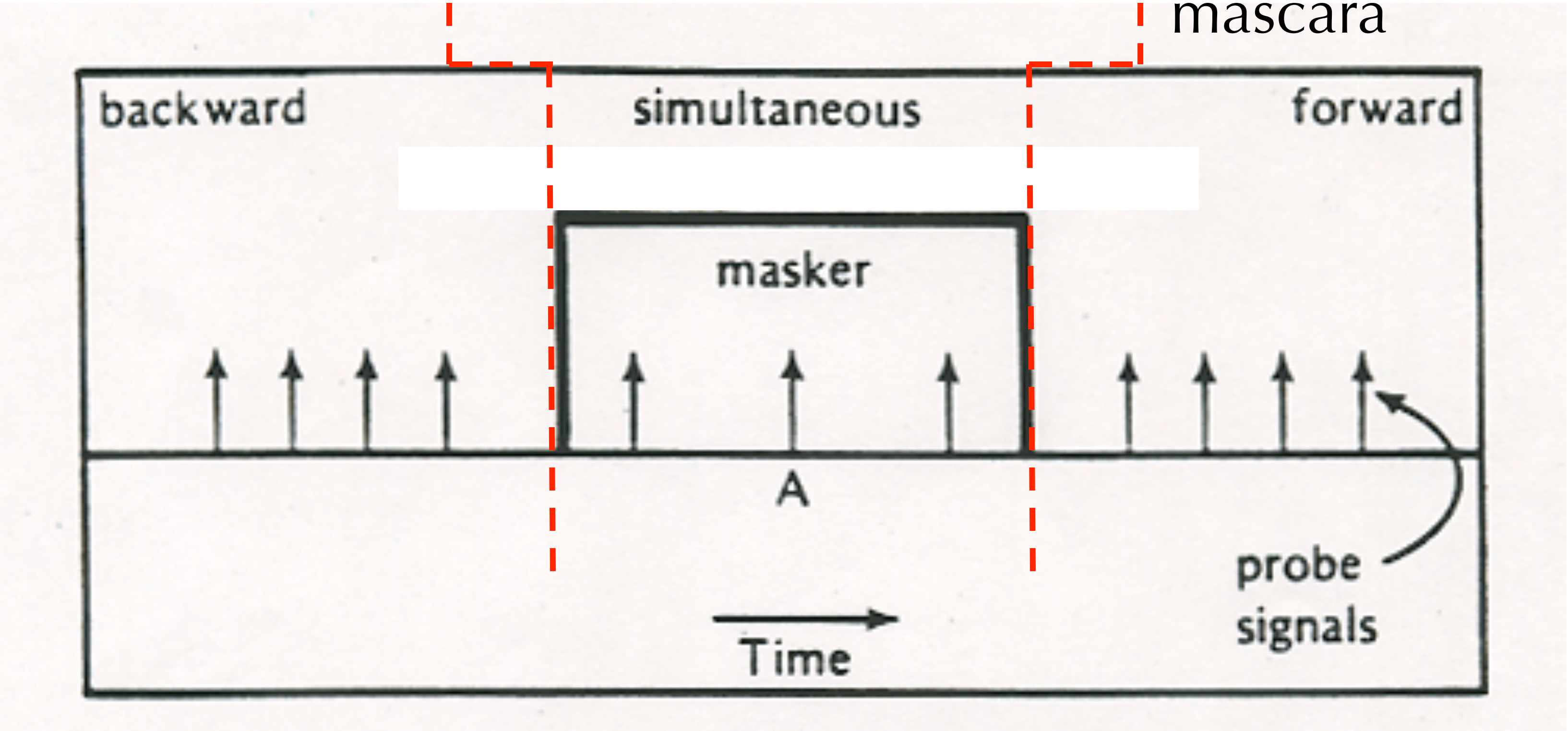
- Enmascaramiento no-simultáneo
- Detección de vacío
- Función de transferencia de modulación temporal (TMTF)
- Patrón de período de enmascaramiento

Enmascaramiento temporal

Backward masking: La señal se presenta *antes* de la máscara

Simultaneous masking: La señal se presenta *durante* la máscara

Forward masking: La señal se presenta *después* de la máscara



Ajustar el nivel de las señales para determinar el umbral

Modified from Yost (1994)

Enmascaramiento temporal

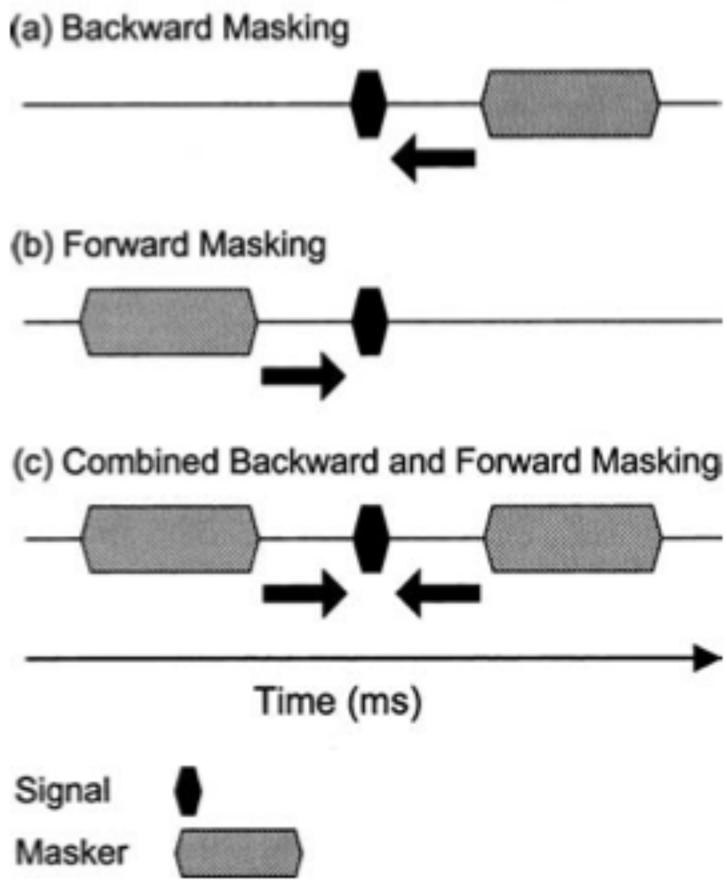
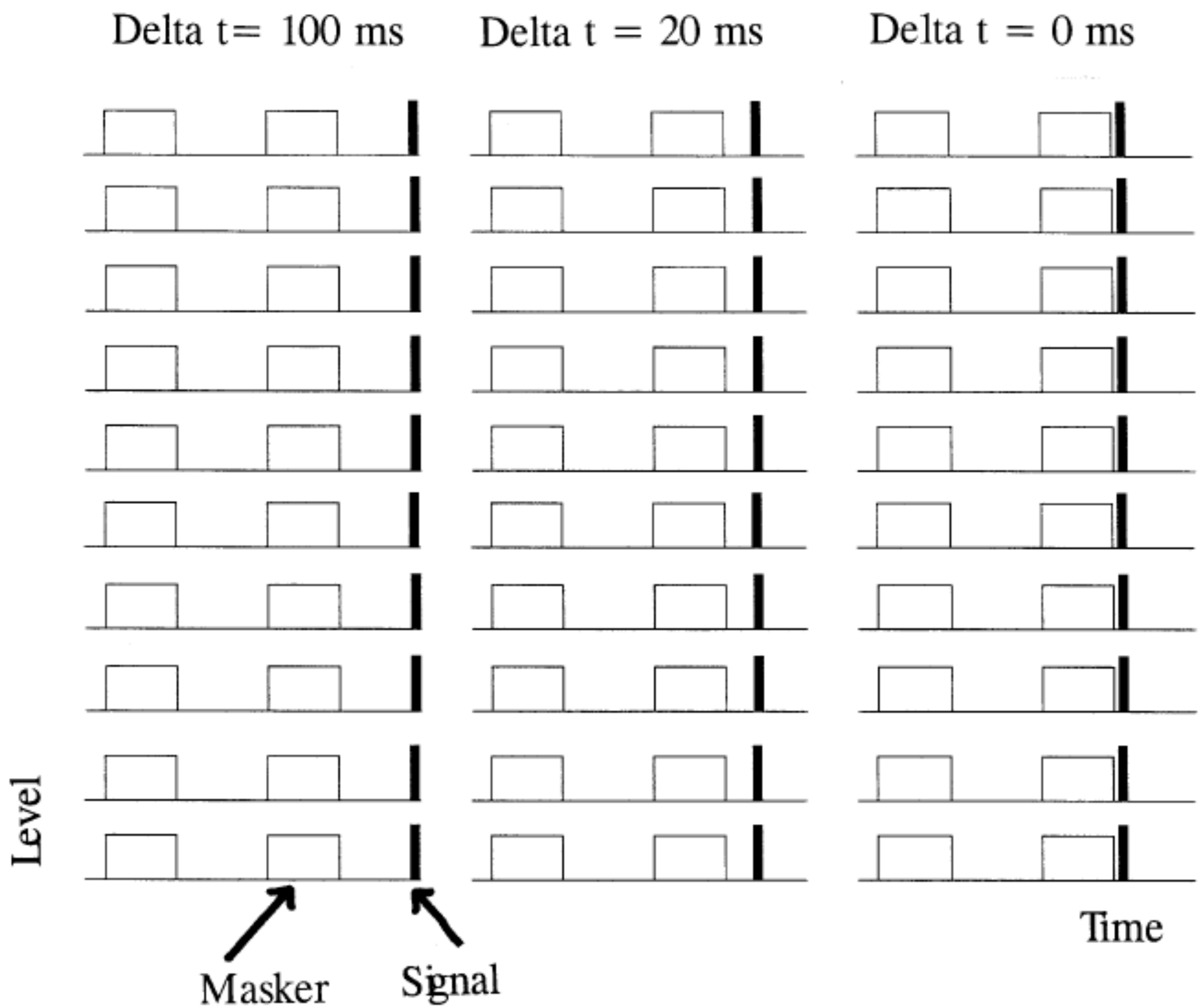
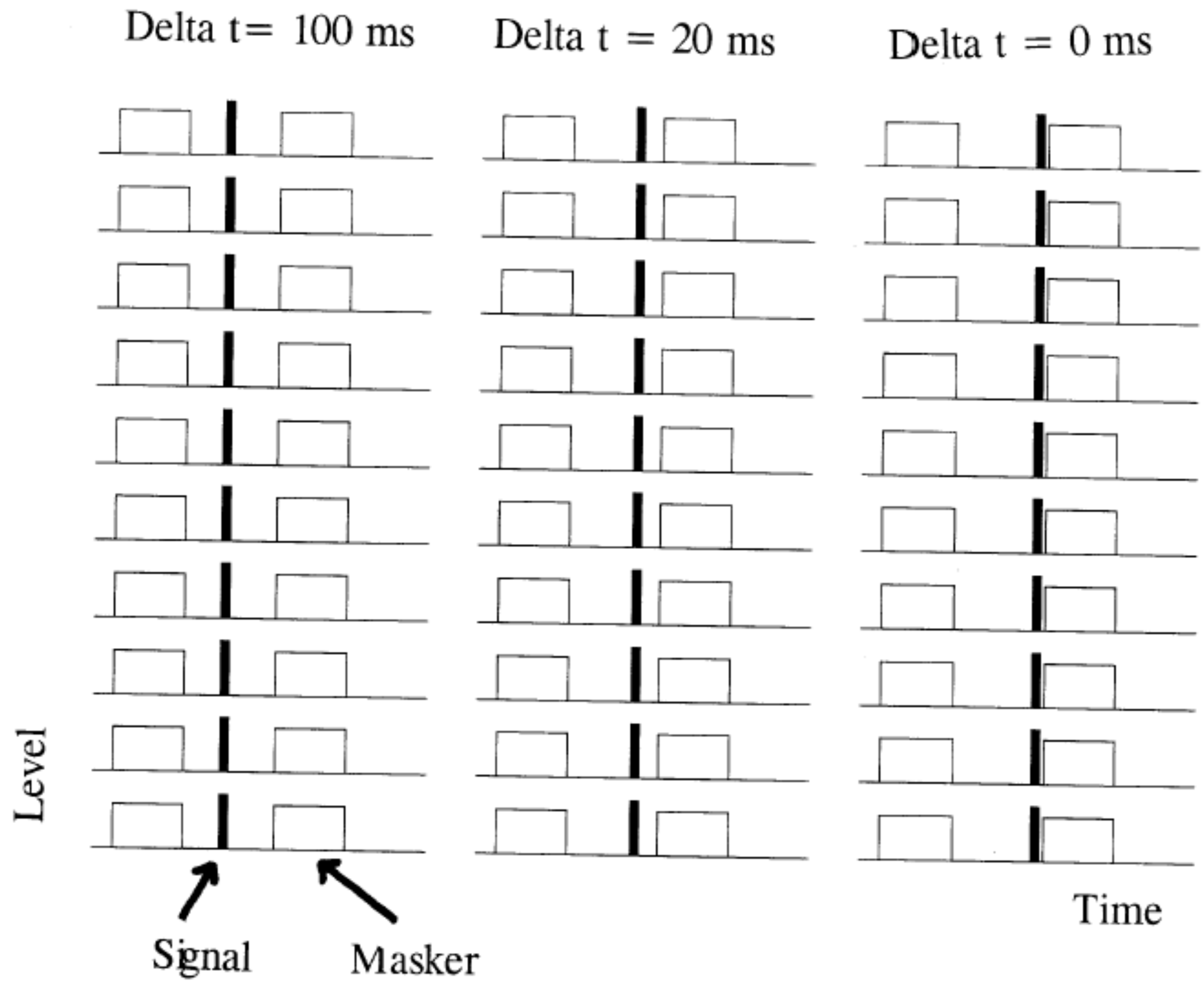


Figure 10.15 Temporal masking paradigms: (a) in backward masking the masker follows the signal, (b) in forward masking the masker precedes the signal, and (c) combined forward and backward masking. The heavy arrows show the direction of the masking effect.

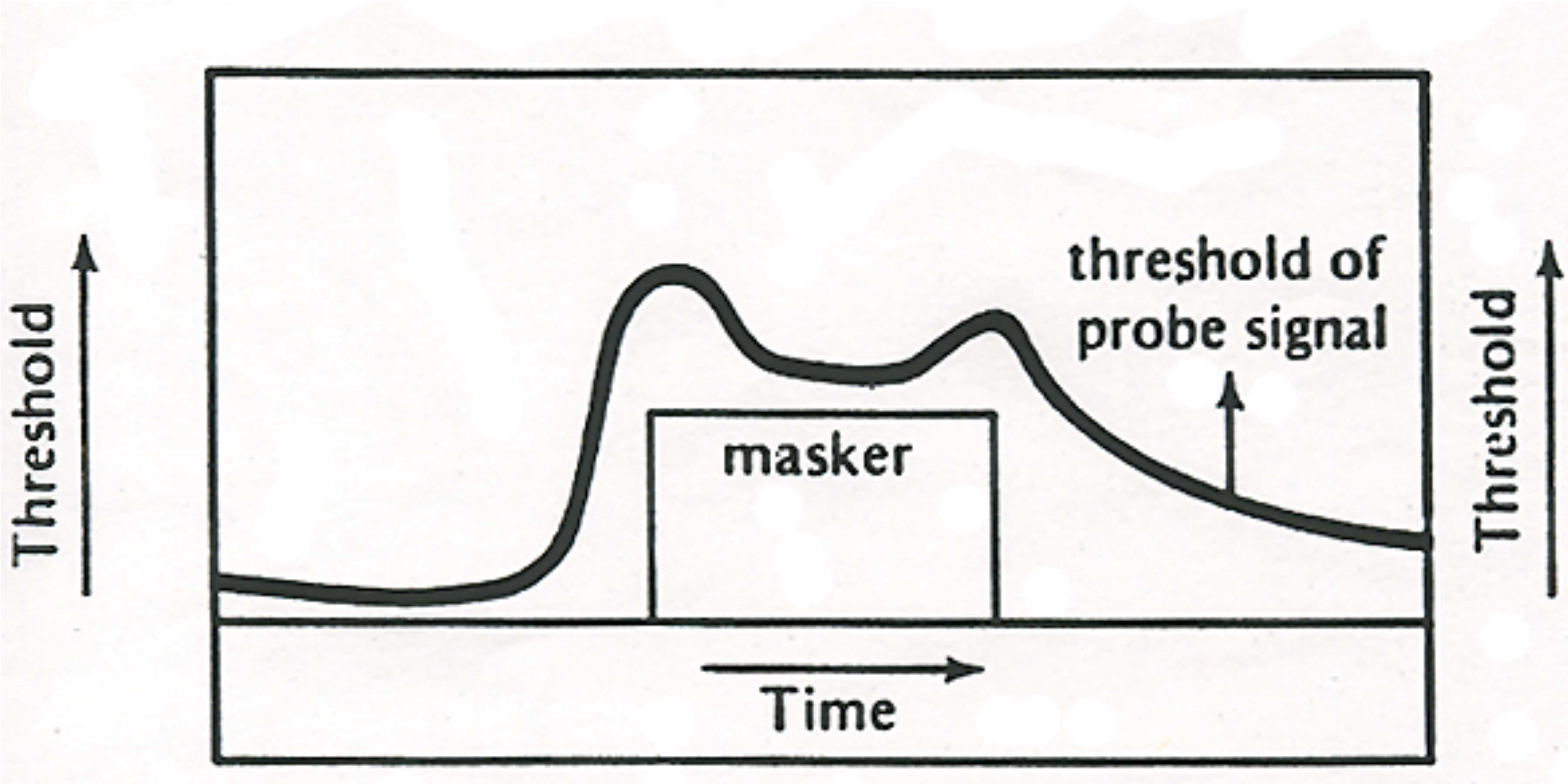
Forward masking



Backward masking



Umbral de enmascaramiento



Demostración 10   Demostración 11

From Yost (1994)

Enmascaramiento no simultáneo: efectos de tiempo

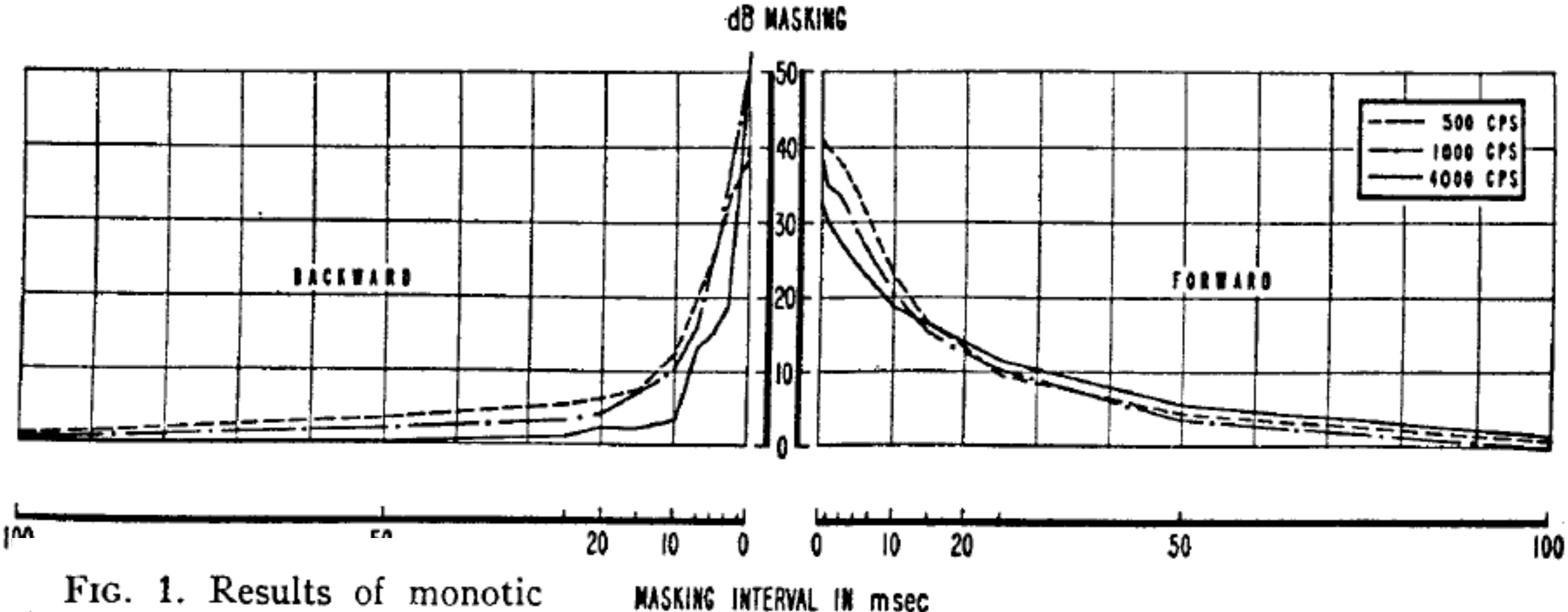


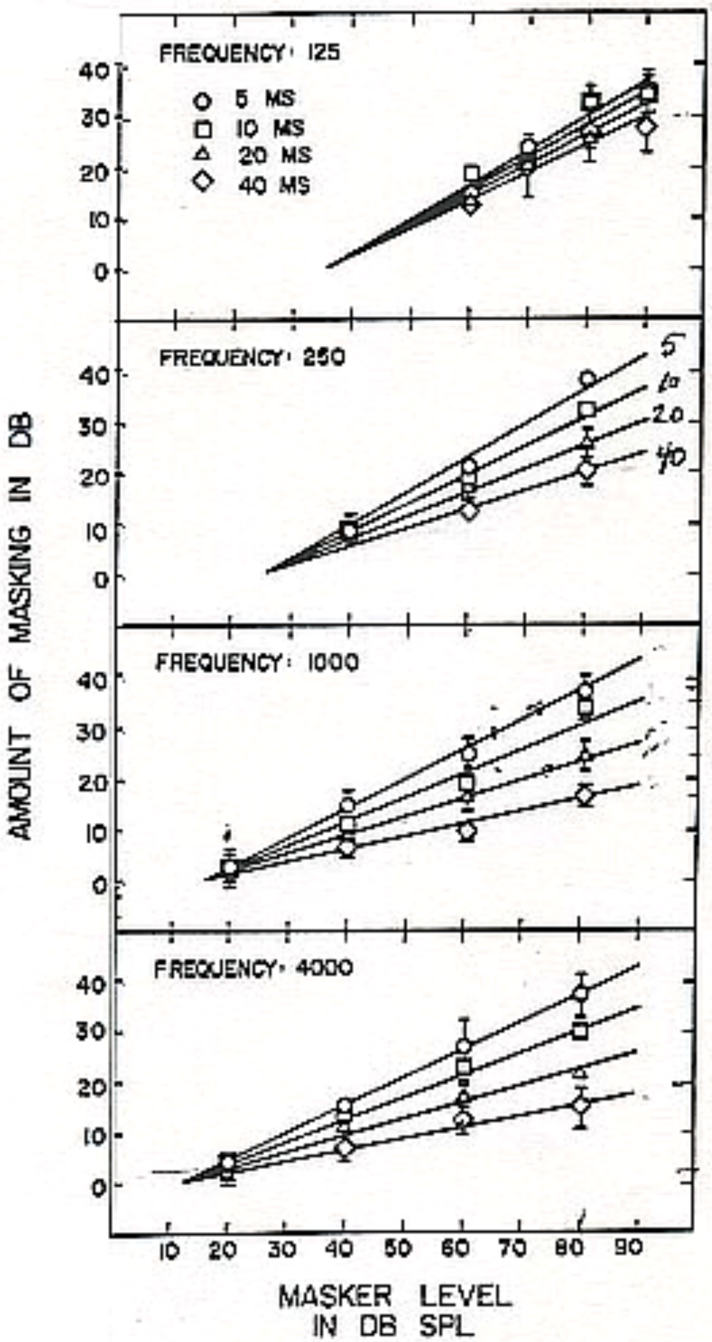
FIG. 1. Results of monotic backward and forward masking (At 0 backward masking interval the points are 50-dB masking for 1000 cps and 53.5-dB masking for 4000 cps.)

En general…

Forward masking dura ~200 ms
Backward masking dura ~50 ms

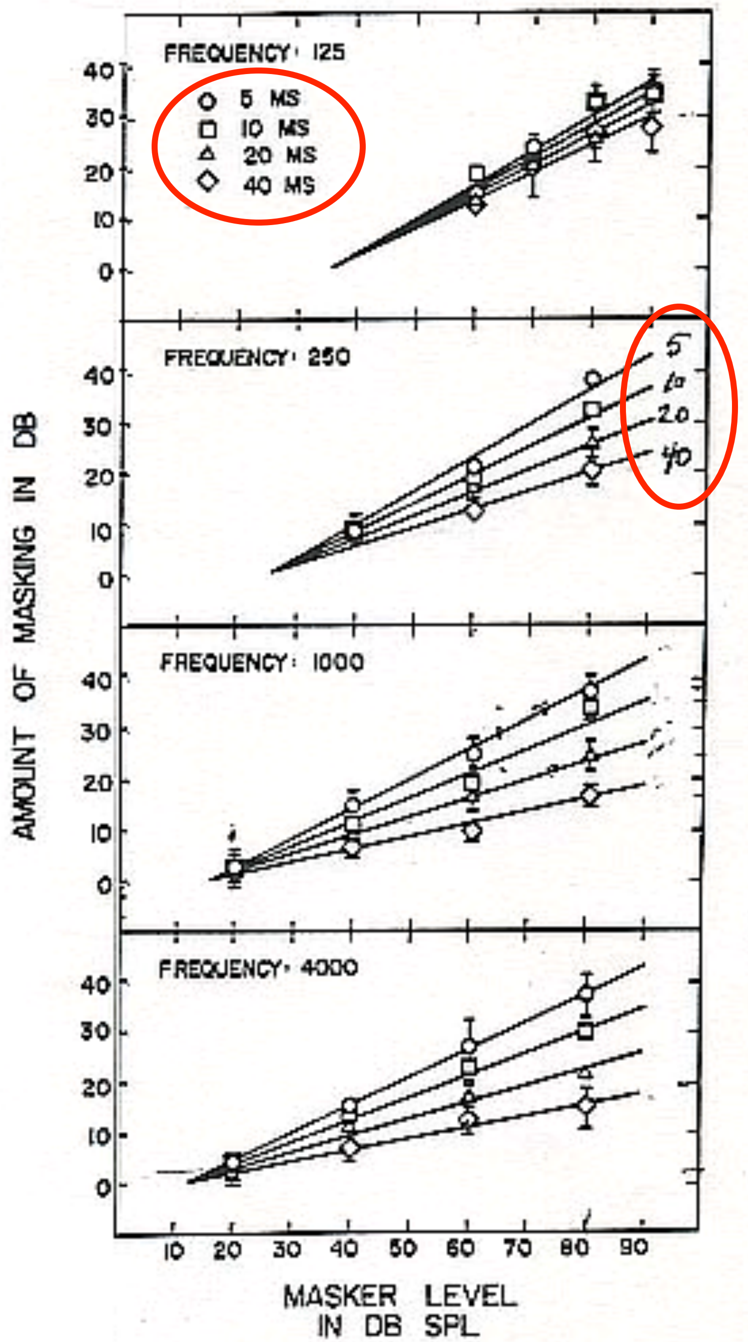
From Elliott (1962)

Crecimiento del forward masking



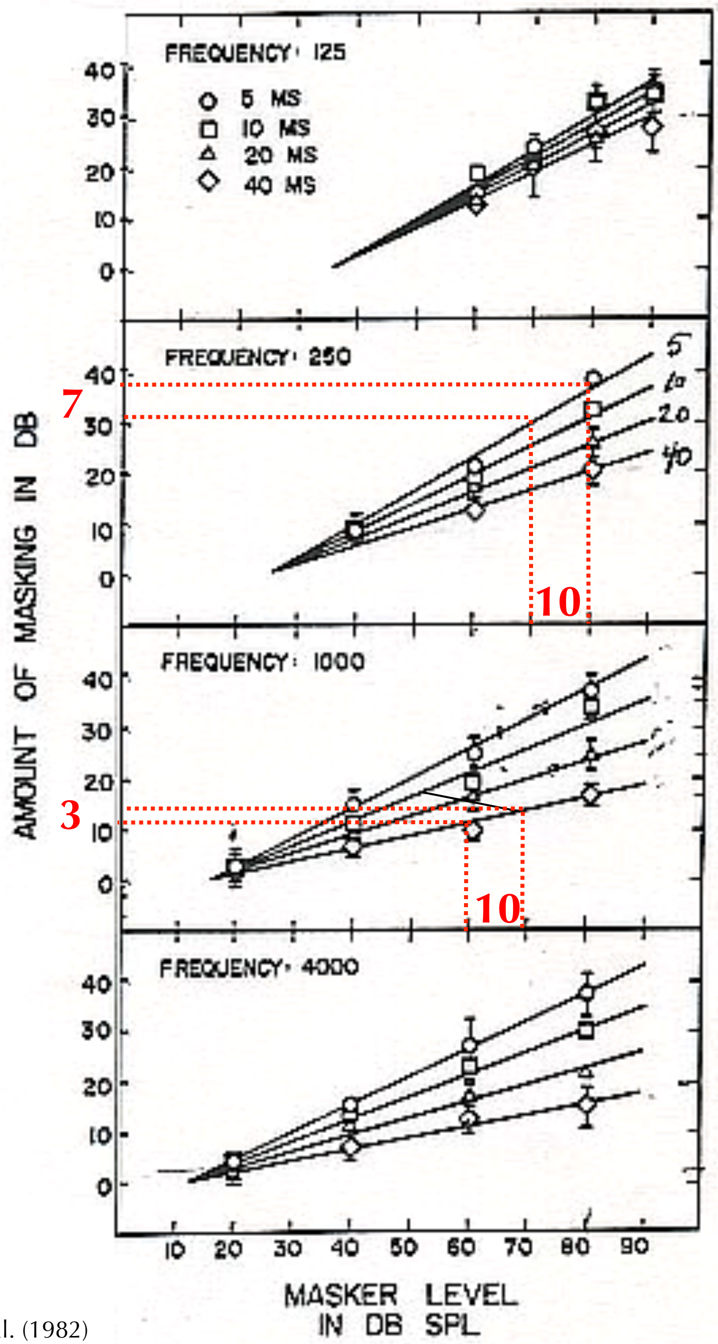
http://depts.washington.edu/sphsc461/masking/masking_files/masking.ppt

Crecimiento del forward masking



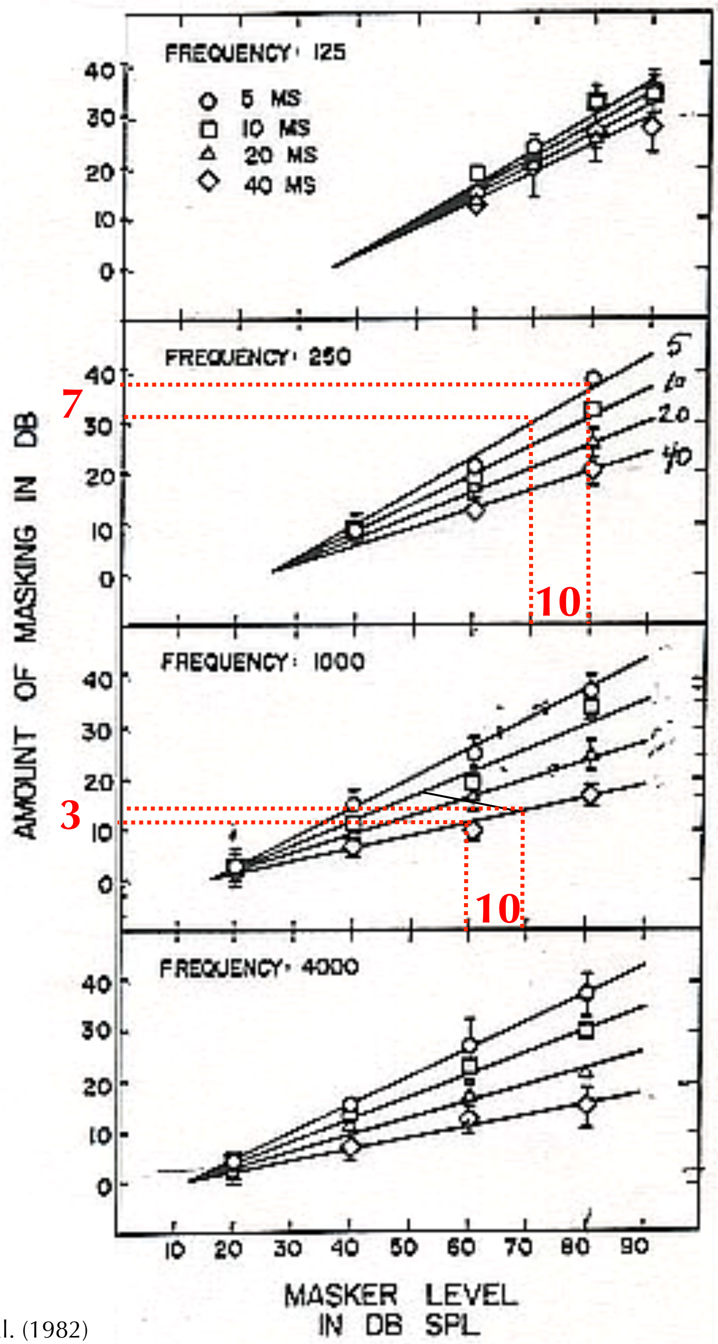
http://depts.washington.edu/sphsc461/masking/masking_files/masking.ppt

Crecimiento del forward masking



Jesteadt et al. (1982)

http://depts.washington.edu/sphsc461/masking/masking_files/masking.ppt



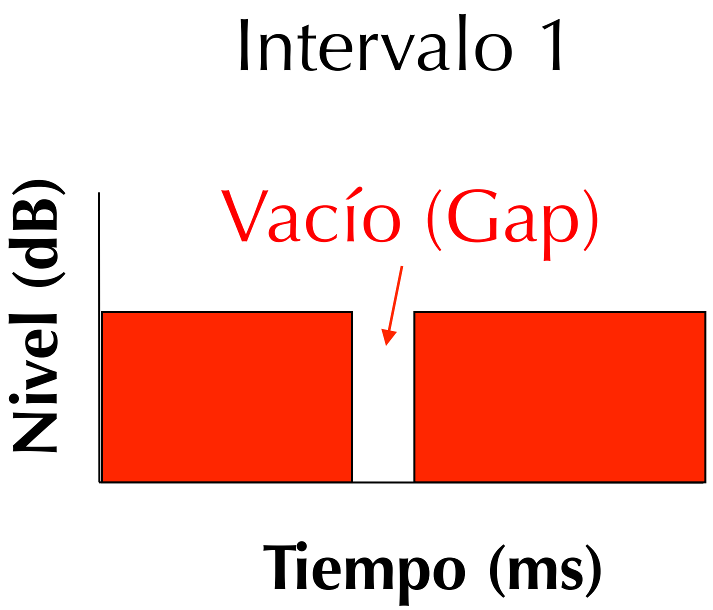
Jesteadt et al. (1982)

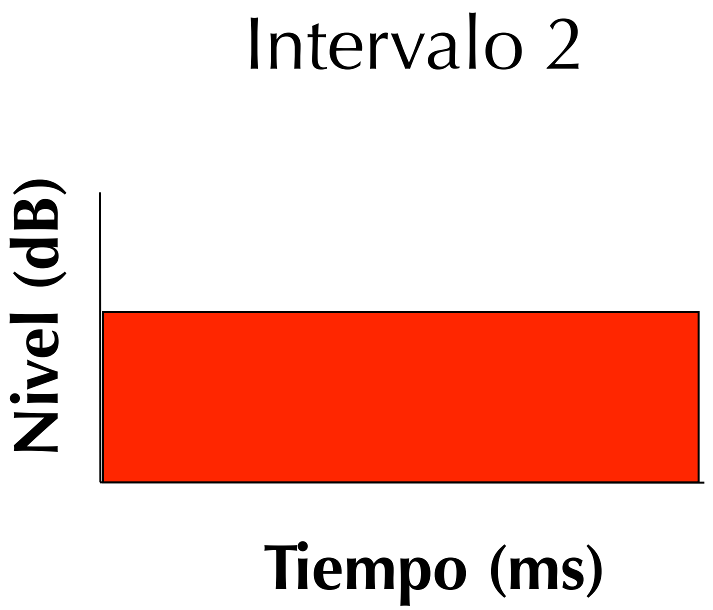
Crecimiento del forward masking

- Un incremento de 10 dB en el nivel de la máscara lleva a un incremento menor de 10 dB en la cantidad de enmascaramiento
- El crecimiento del forward masking depende del tiempo y la frecuencia

http://depts.washington.edu/sphsc461/masking/masking_files/masking.ppt

Detección de vacíos





¿Cuál intervalo tuvo un vacío?

Ajustar la duración del vacío para determinar el umbral

Umbrales de vacío

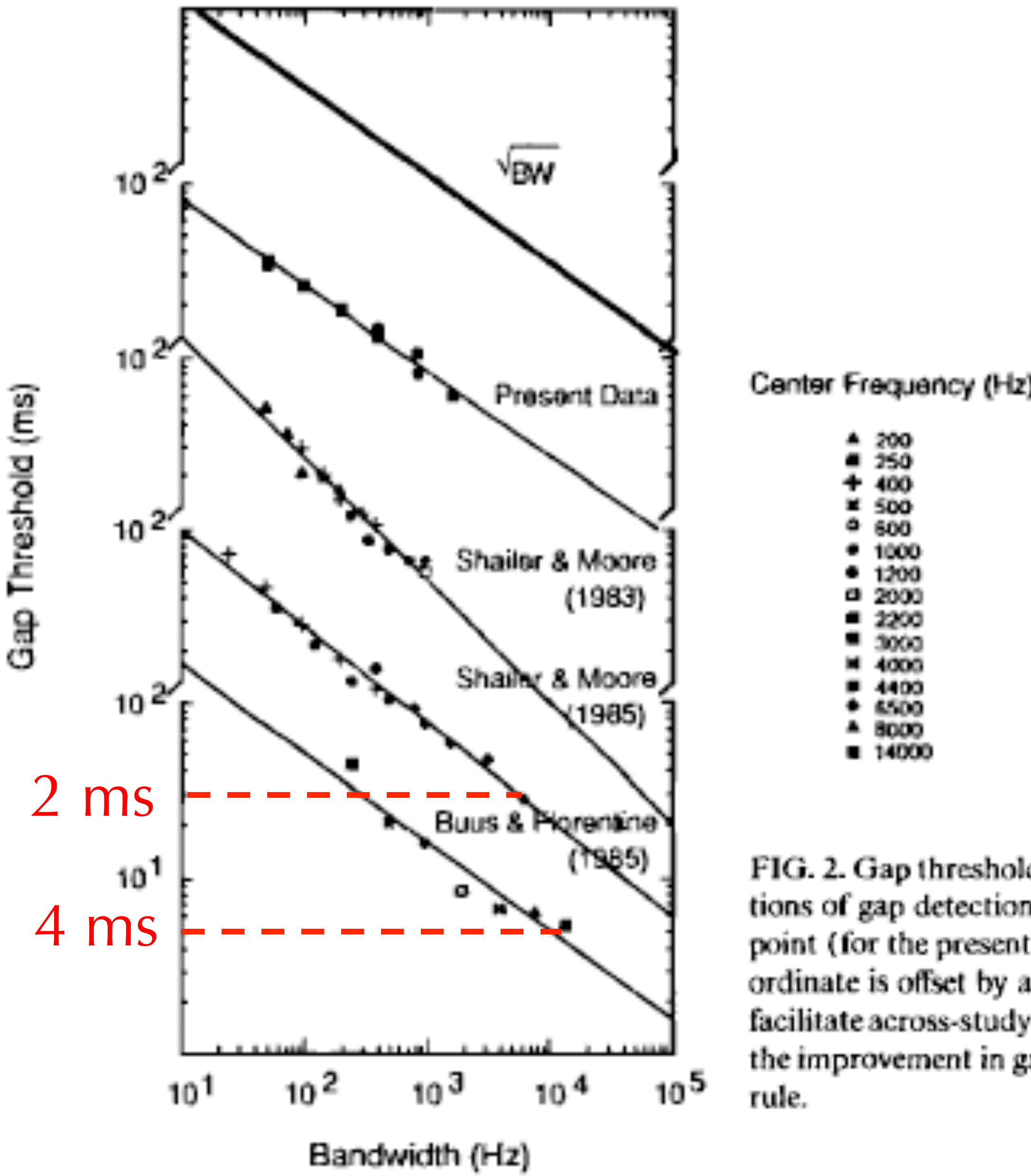


FIG. 2. Gap threshold as a function of signal bandwidth for four investigations of gap detection. Symbols indicate center frequency for each datum point (for the present data, symbols indicate upper cutoff frequency). The ordinate is offset by a factor of 10 and the data are displaced vertically to facilitate across-study comparison. The thick solid line at the top represents the improvement in gap threshold predicted by a square root of bandwidth rule.

- Detección de umbrales de vacío
  - Decrece al incrementarse el ancho de banda del ruido
  - entre ~2-4 ms para ruido ancho

Eddins et al., 1992

Detección de vacío y trastornos auditivos

TABLE I. Mean gap-detection thresholds for the normal and hearing-impaired listeners in each condition. Entries are means and standard deviations (in parentheses) in milliseconds of the final gap thresholds for each group of listeners.

	Octave-band condition		
Subjects	400–800 Hz	800–1600 Hz	2000–4000 Hz
Normal 85 dB SPL	9.17 (1.54)	6.97 (1.13)	5.09 (1.18)TL
Normal 30 dB SL	12.38 (1.37)	9.46 (0.88)	6.06 (0.49)
Impaired 30 dB SL	16.57 (2.62)	12.57 (2.07)	8.01 (2.67)

Fitzgibbons & Wightman, 1982

Detección de vacío y trastornos auditivos

TABLE I. Mean gap-detection thresholds for the normal and hearing-impaired listeners in each condition. Entries are means and standard deviations (in parentheses) in milliseconds of the final gap thresholds for each group of listeners.

	Octave-band condition		
Subjects	400–800 Hz	800–1600 Hz	2000–4000 Hz
Normal 85 dB SPL	9.17 (1.54)	6.97 (1.13)	5.09 (1.18)TL
Normal 30 dB SL	12.38 (1.37)	9.46 (0.88)	6.06 (0.49)
Impaired 30 dB SL	16.57 (2.62)	12.57 (2.07)	8.01 (2.67)

Nivel de Sensación (SL): Nivel determinado sobre un nivel base, típicamente un umbral absoluto, determinado para cada sujeto

Fitzgibbons & Wightman, 1982

Detección de vacío y trastornos auditivos

TABLE I. Mean gap-detection thresholds for the normal and hearing-impaired listeners in each condition. Entries are means and standard deviations (in parentheses) in milliseconds of the final gap thresholds for each group of listeners.

Subjects	Octave-band condition 400–800 Hz	800–1600 Hz	2000–4000 Hz
Normal 85 dB SPL	9.17 (1.54)	6.97 (1.13)	5.09 (1.18)TL
Normal 30 dB SL	12.38 (1.37)	9.46 (0.88)	6.06 (0.49)
Impaired 30 dB SL	16.57 (2.62)	12.57 (2.07)	8.01 (2.67)

Nivel de presión de sonido (SPL): Nivel de sonido estándar, independiente del sujeto

Fitzgibbons & Wightman, 1982

Nivel de sensación

<u>Se quiere un nivel de sensación level (SL) de 25 dB</u>

Sujeto A:

Umbral absoluto: 10 dB

SPL requerido: 35 dB

(25 dB SL+10 dB= 35 dB)

Sujeto B:

Umbral absoluto: 40 dB

SPL requerido: 65 dB

(25 dB SL+40 dB= 65 dB)

Detección de vacío y trastornos auditivos

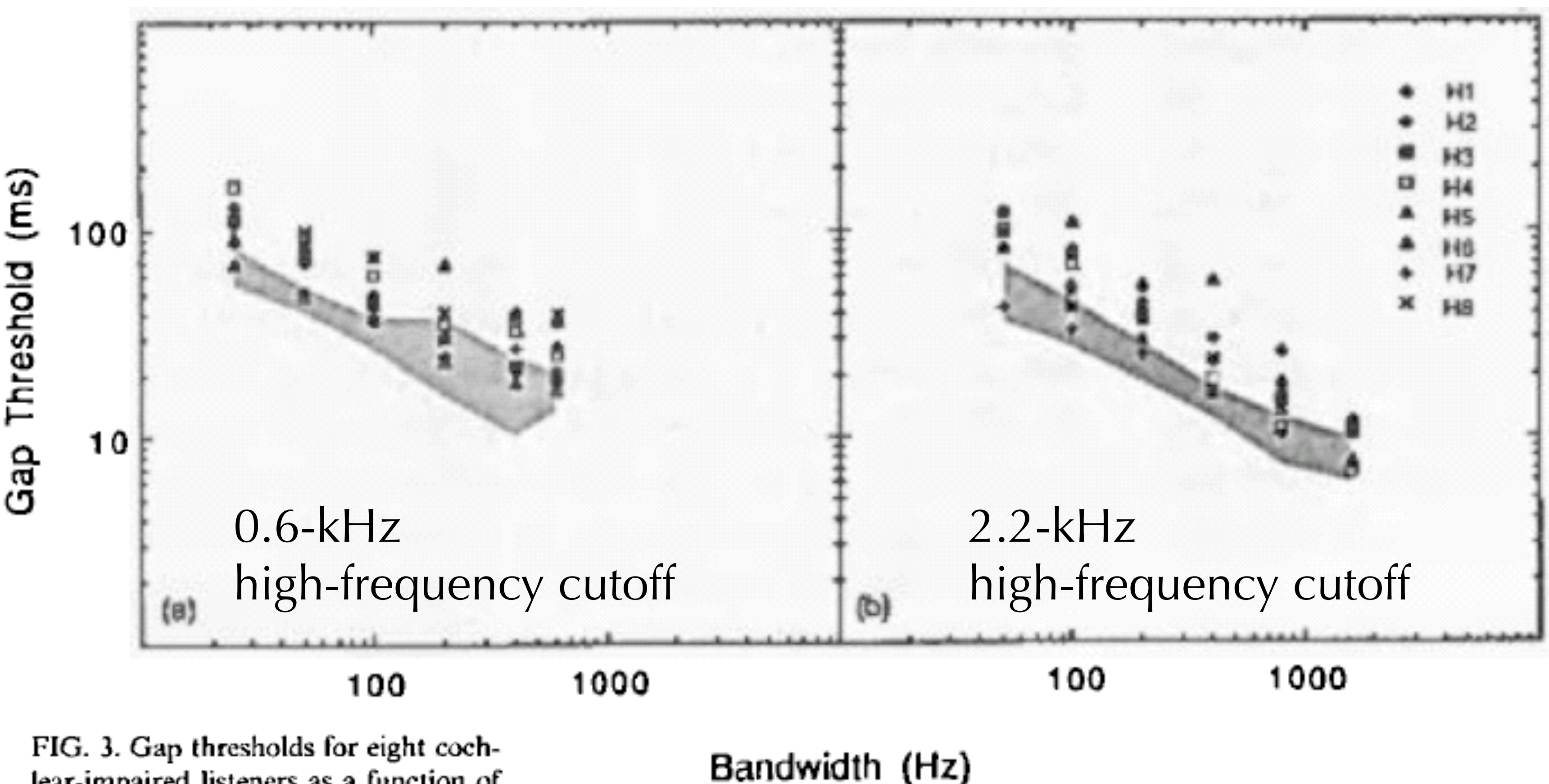
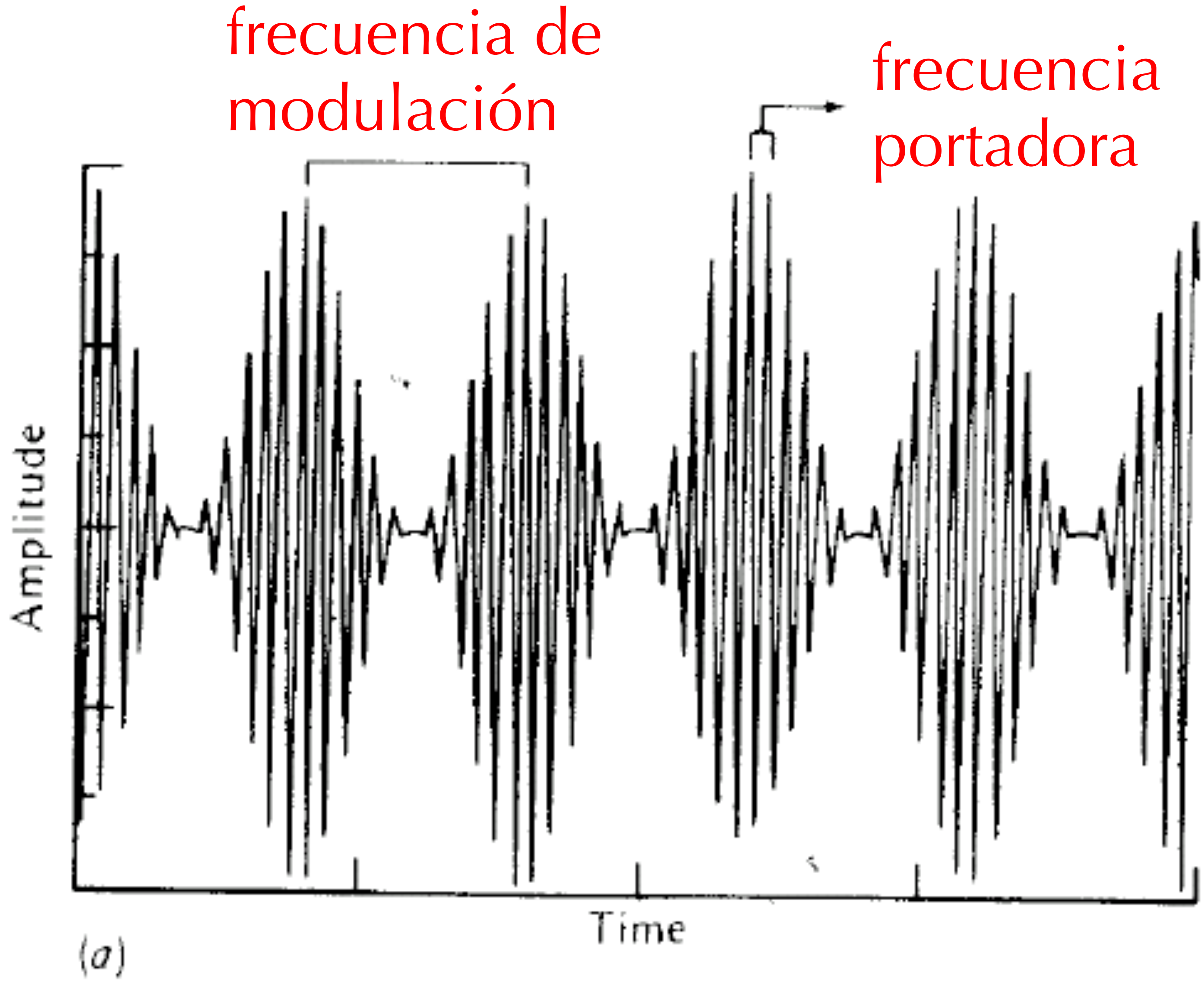


FIG. 3. Gap thresholds for eight cochlear-impaired listeners as a function of stimulus bandwidth for (a) a 0.6-kHz high-frequency cutoff and (b) a 2.2-kHz high-frequency cutoff. Shaded areas indicate the mean $\pm$ 2 s.d. of the data for normal listeners.

Grose et al., 1989

Modulación de amplitud (AM)



Detección de AM

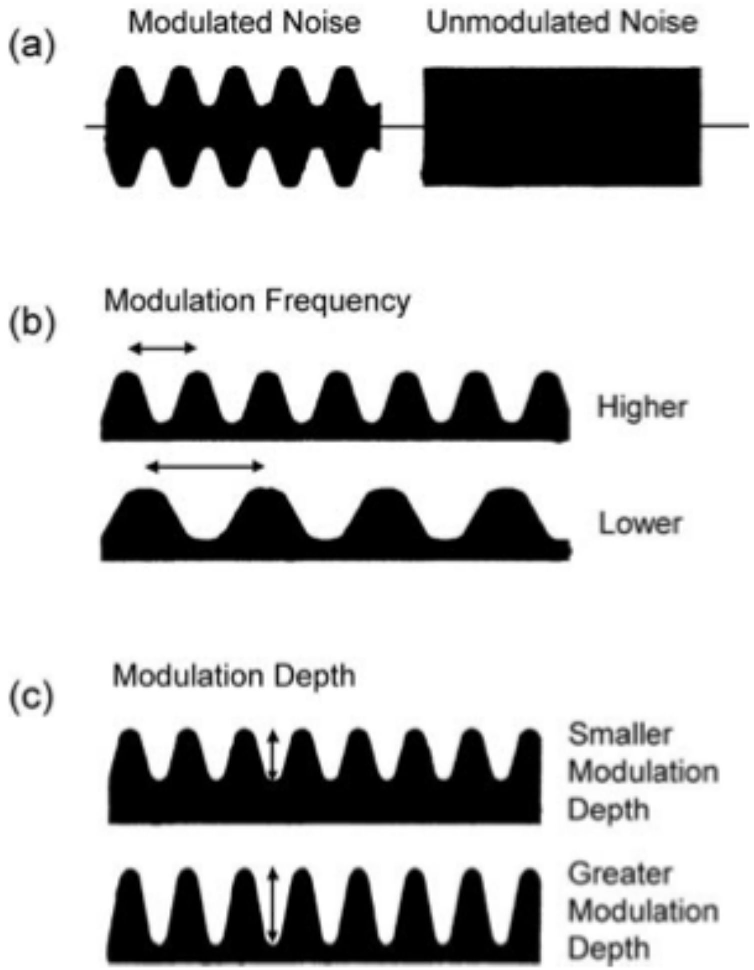
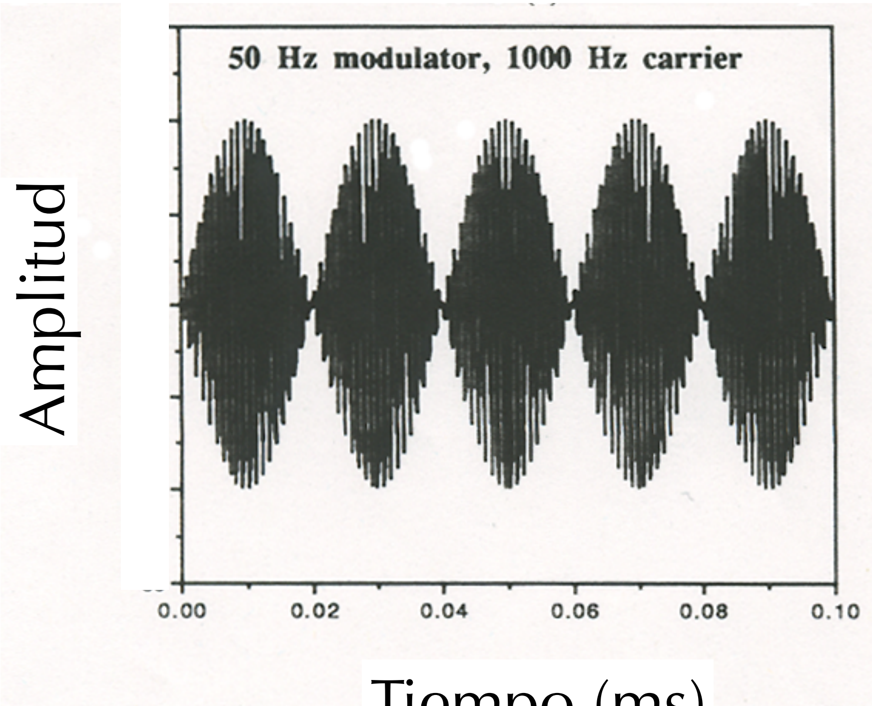


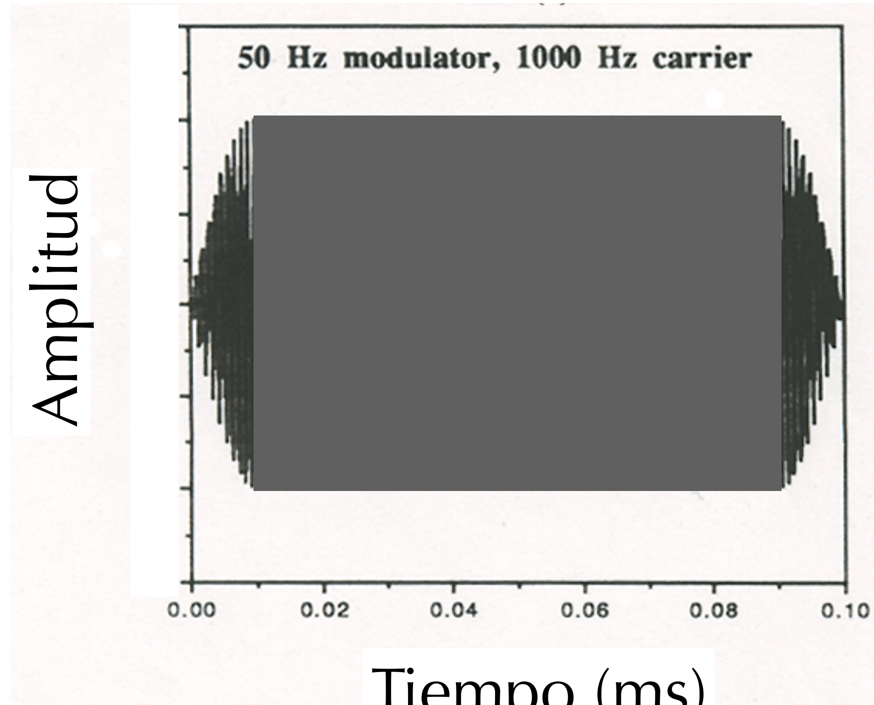
Figure 9.13 Artist's conceptualizations of (a) amplitude-modulated versus unmodulated noises, (b) different modulation rates, and (c) different modulation depths.

Detección de AM

Intervalo 1



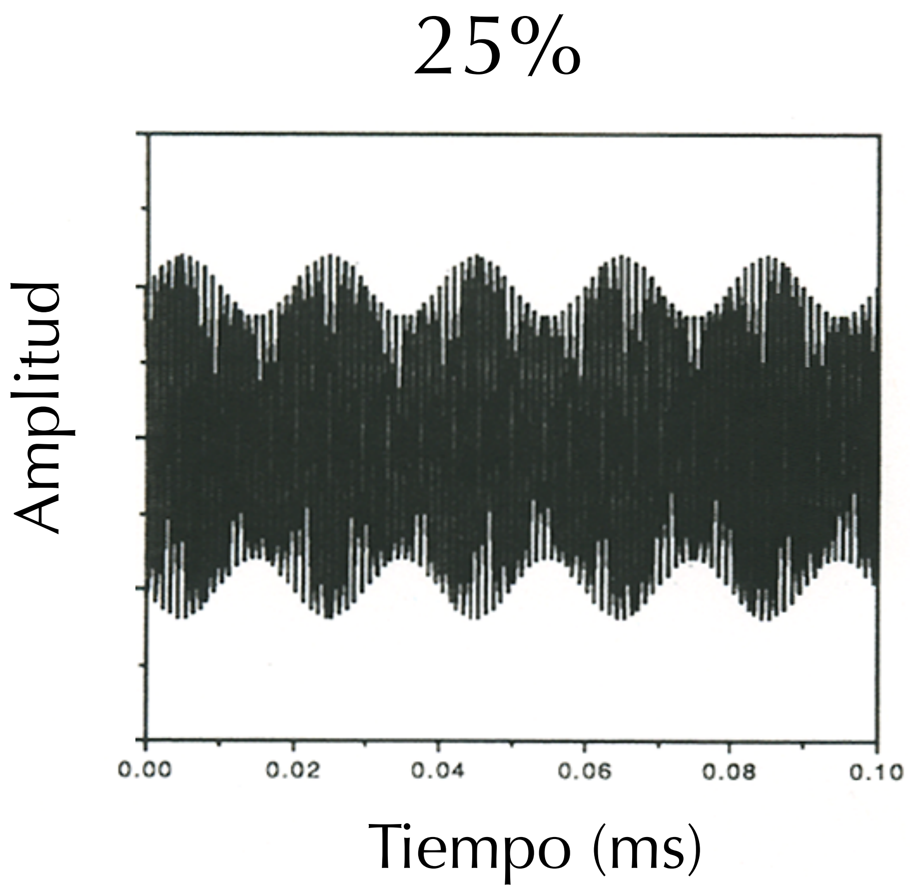
Intervalo 2

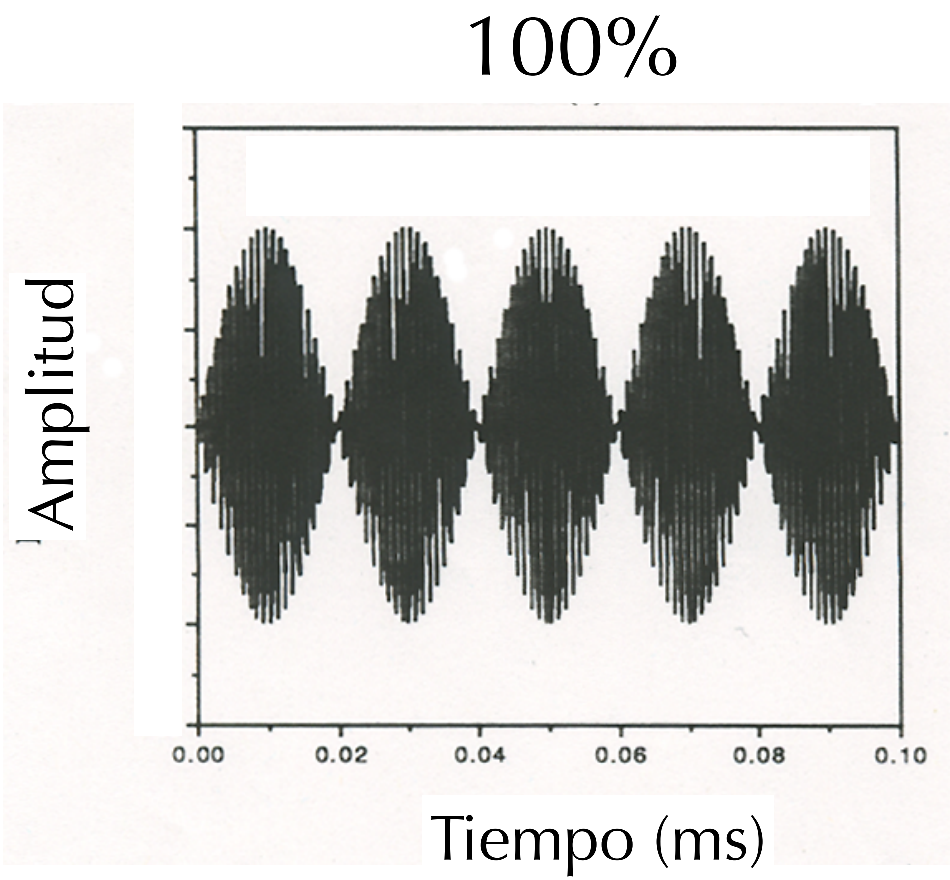


¿En cúal intervalo fue modulado el sonido?

Ajustar la cantidad de modulación para determinar el umbral

Cantidad de modulación AM





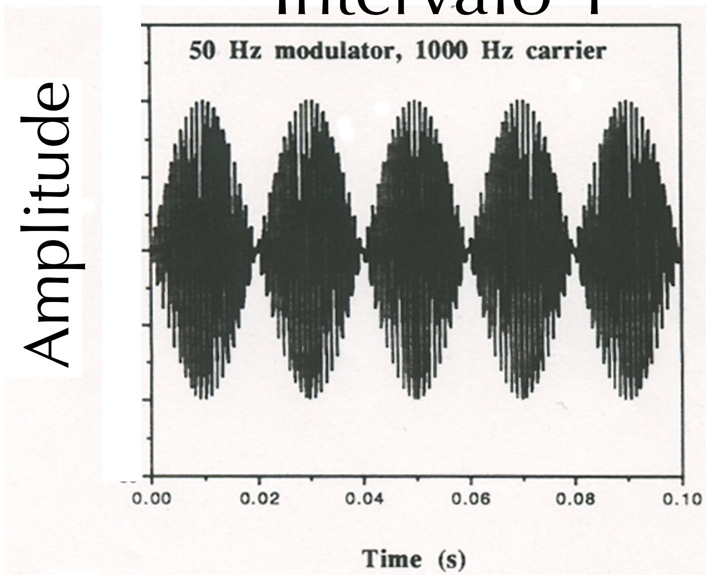
http://depts.washington.edu/sphsc461/temp_res/temp_res.ppt_files/temp_res.ppt.ppt

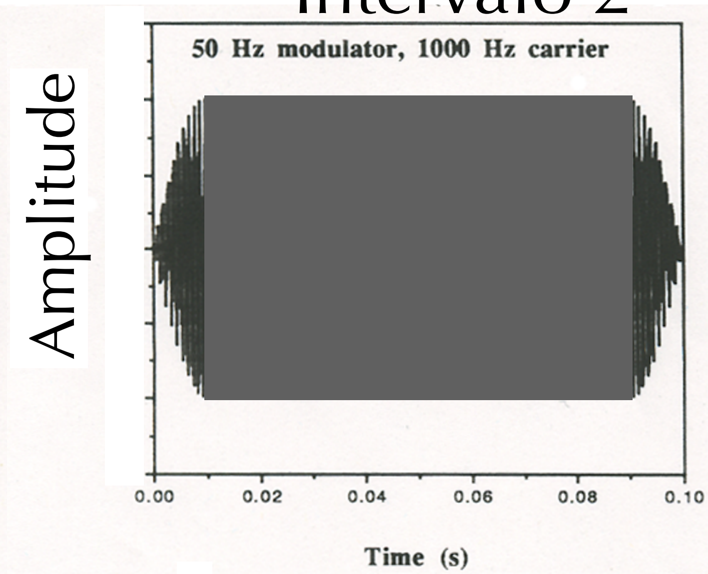
Cantidad de modulación AM

Intervalo 1

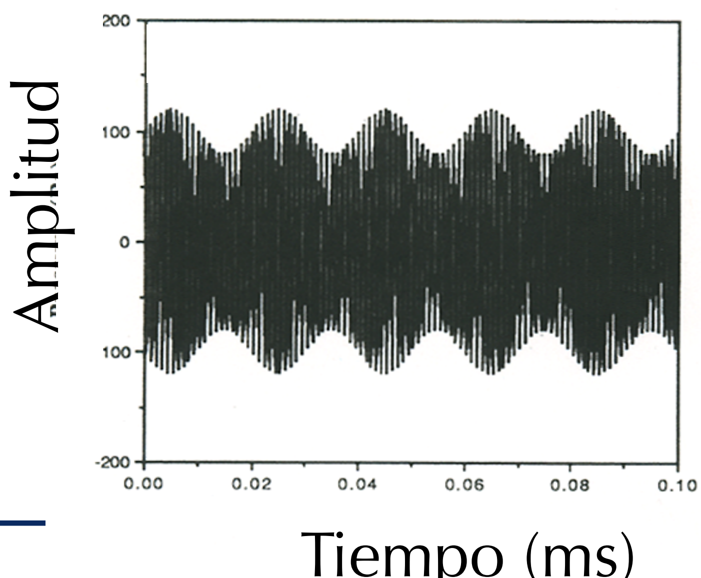
Intervalo 2

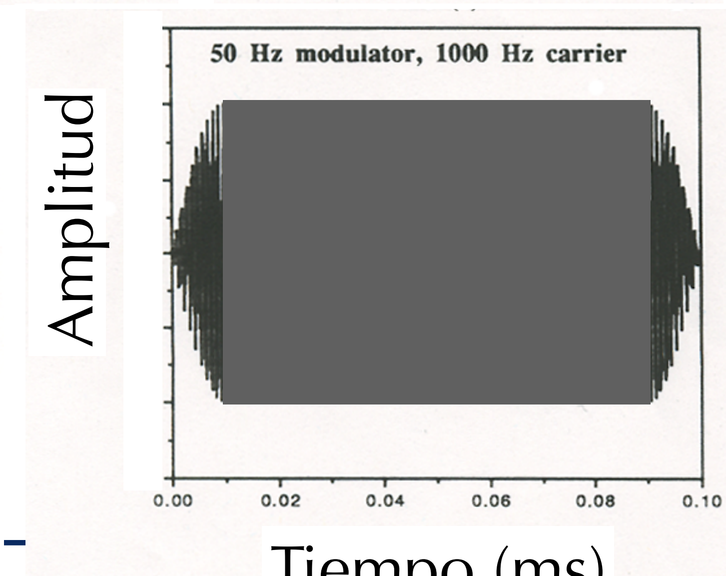
100%





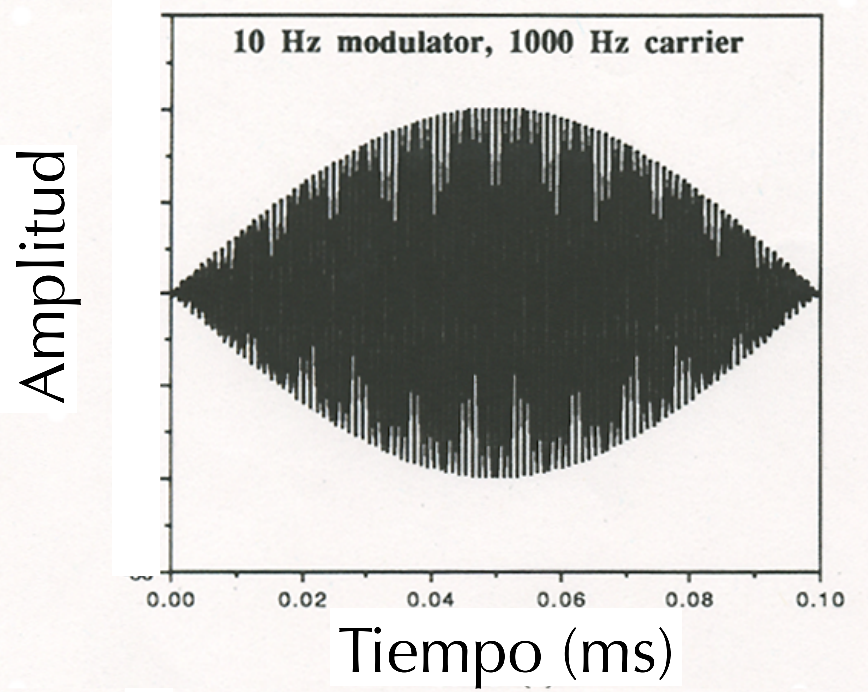
25%

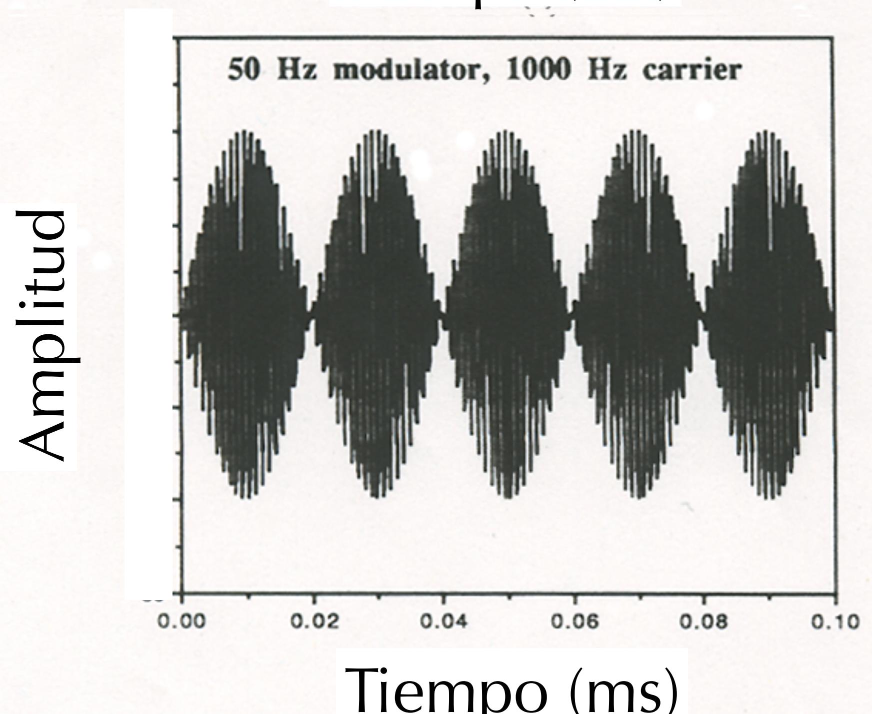


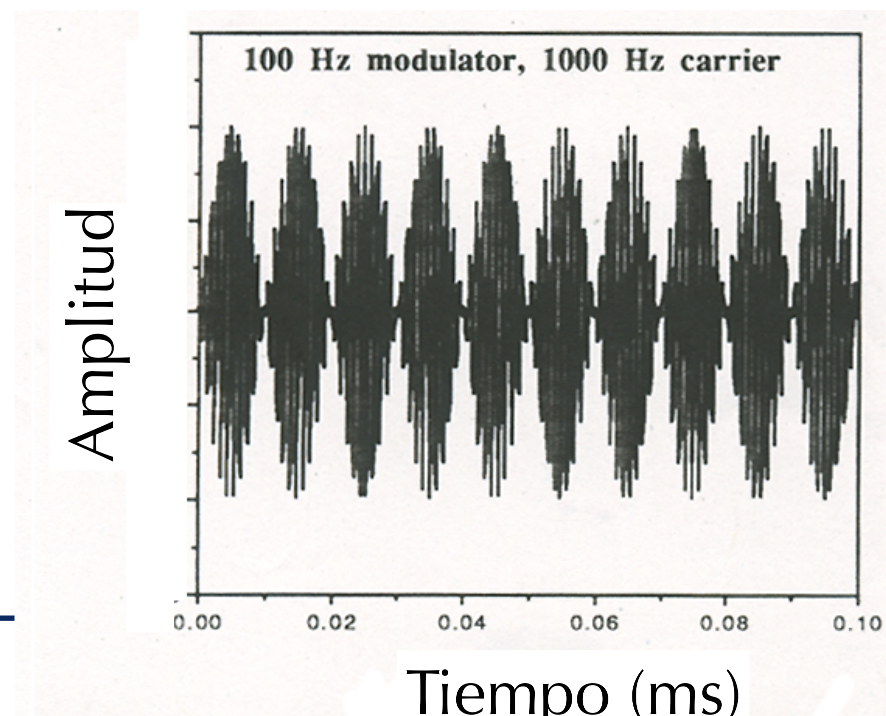


CITA

Tasa de modulación







http://depts.washington.edu/sphsc461/temp_res/temp_res.ppt_files/temp_res.ppt.ppt

Temporal Modulation Transfer Function

Detección de AM como una función de la tasa de modulación

plana

Cantidad de modulación (20 log m)

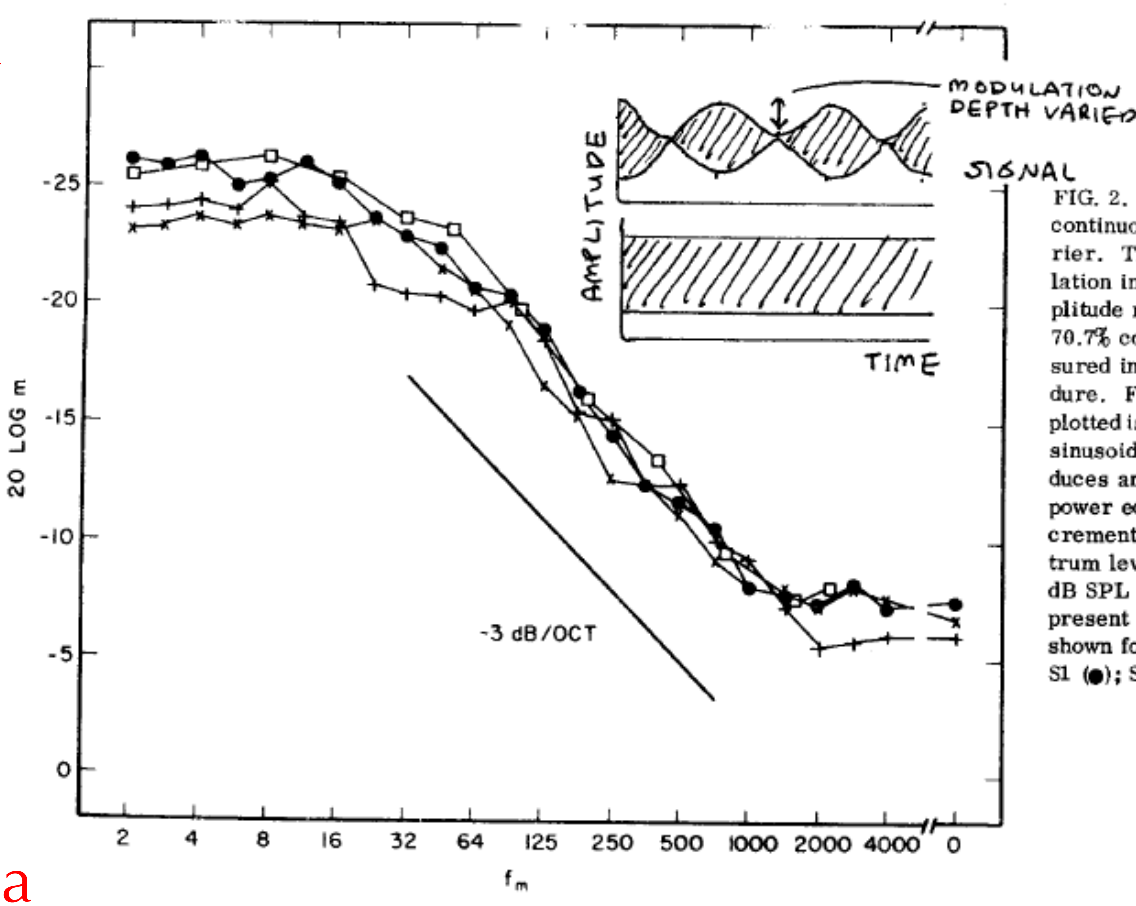


FIG. 2. TMTFs obtained with a continuous wideband noise carrier. The ordinate is the modulation index for sinusoidal amplitude modulation necessary for 70.7% correct responses as measured in a 2IFC tracking procedure. For $f_m = 0$ Hz, the value plotted is the modulation depth for sinusoidal modulation which produces an increment in average power equal to the measured increment threshold. The spectrum level of the carrier was 40 dB SPL and the modulation was present for 500 ms. Data are shown for individual observers: S1 (●); S2 (×); S3 (+); S4 (■).

profunda

Frecuencia de modulación

CITA

Detección de AM y trastornos auditivos

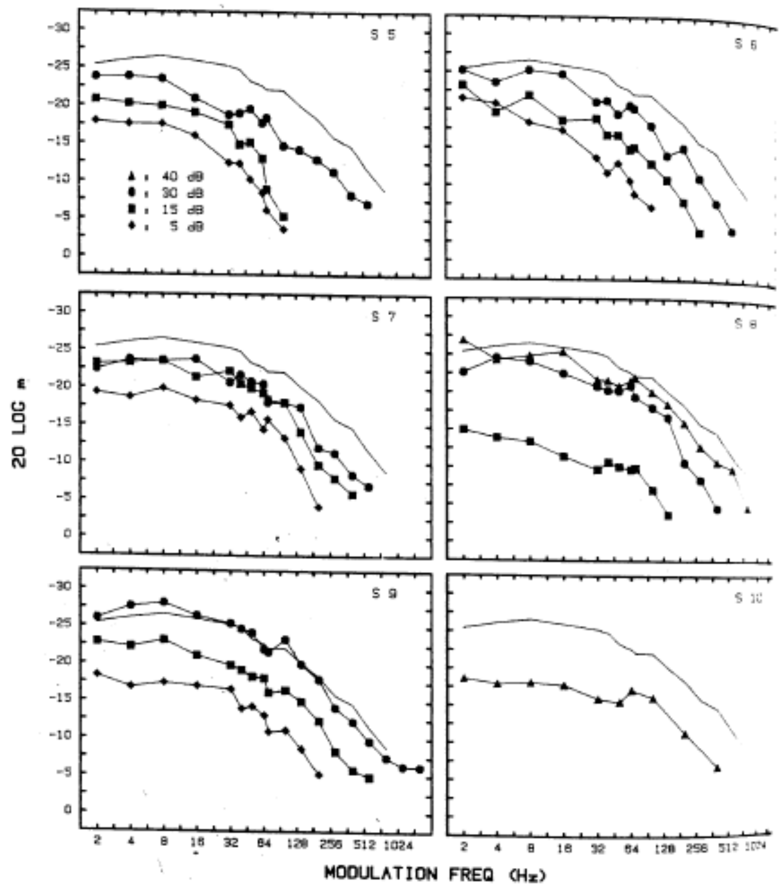


Fig. 2. TMTFs obtained with a continuous broadband noise carrier in 6 sensorineural hearing-impaired listeners. The spectrum level of the carrier is the parameter. The mean of the 4 normal-hearing listeners (fig. 1) is shown as a solid line without symbols.

NORMAL-HEARING LISTENERS
TESTED AT 30 dB SPECTRUM LEVEL

Patrón de período de enmascaramiento

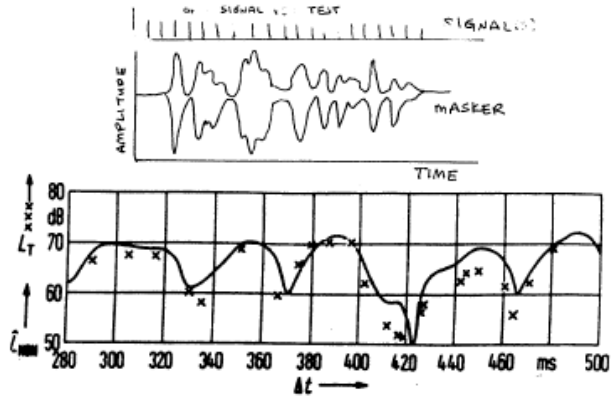


Fig. 2. Threshold level L_T (×××) of 2 ms long 3-kHz-tone impulses masked by the artificial narrow band noise centered at 2 kHz with a bandwidth of 32 Hz and with the instantaneous level L_{NBN} (solid curve) as function of the delay time Δt. The "effective" level L_{NBN} of the masking noise is 70 dB.

Tasa de modulación baja

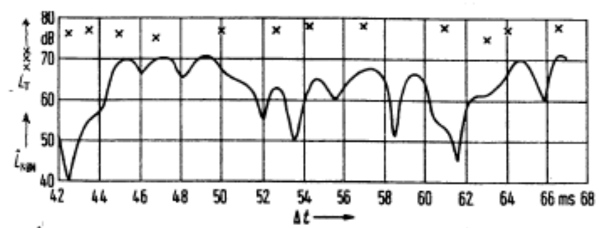


Fig. 3.
Threshold level L_T (×××) of 2 ms long 4-kHz-tone impulses masked by artificial narrow band noise centered at 4 kHz with a bandwidth of 1 kHz and with the instantaneous level L_{NBN} (solid curve) as function of the delay time Δt. The "effective" level of the masking noise is 70 dB.

Tasa de modulación alta

Patrón de período de enmascaramiento

Alta correlación

Baja correlación

The correlation coefficient r and the factor m are shown in Fig. 4 as function of the bandwidth Δf for

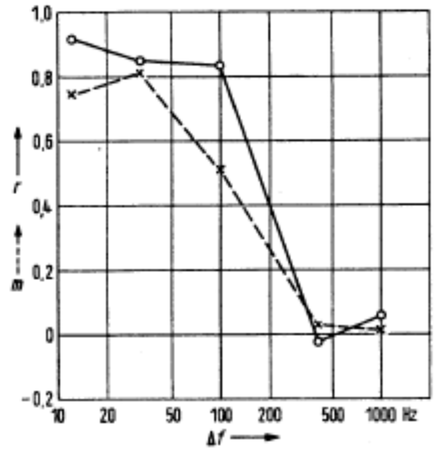


Fig. 4. Correlation coefficient r (circles and solid line) and regression factor m (crosses and dashed line) between the instantaneous level $\hat{L}_{\mathrm{NBN}}$ of the masking noise and the masked threshold L_{T} of 2-ms-tone impulses as function of the bandwidth Δf of the masking noise. Centre-frequency of the noise and tone-frequency of the impulses are each 4 kHz.

one observer. A 2-ms-tone-impulse at 4 kHz was used as signal.

Measurements with other subjects showed similar results. For small bandwidth, m and r are very close to 1. This means that L_{T} and $\hat{L}_{\mathrm{NBN}}$ change in the same way and for the same amount. At a bandwidth of 100 Hz, the correlation coefficient is still 0.85 indicating that L_{T} and $\hat{L}_{\mathrm{NBN}}$ have a similar temporal pattern but the extent of the changes of L_{T} is only half that of $\hat{L}_{\mathrm{NBN}}$ as indicated by $m = 0.5$. For $\Delta f = 400$ Hz and $\Delta f = 1$ kHz, r as well as m are small. This means that L_{T} and $\hat{L}_{\mathrm{NBN}}$ are uncorrelated and the threshold L_{T} depends on an averaged value of $\hat{L}_{\mathrm{NBN}}$ and not on the temporal structure of the envelope.

m = THE DYNAMIC RANGE OF THE THRESHOLDS RELATIVE TO THE NOISE

r = HOW CAREFULLY THE THRESHOLDS FOLLOW THE SHAPE OF THE NOISE

Enmascaramiento de modulación

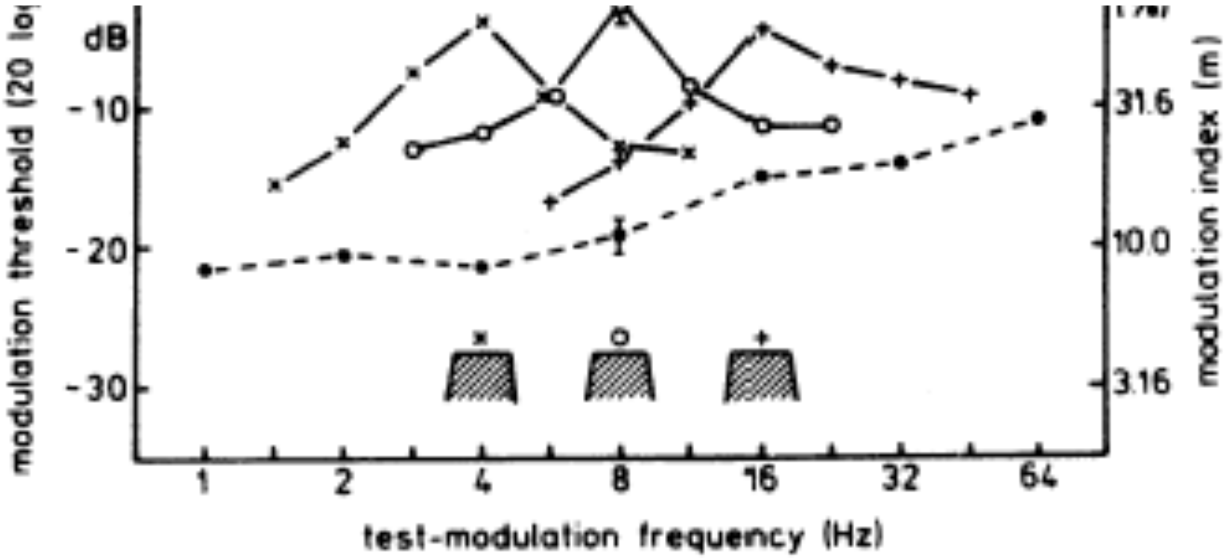


FIG. 2. The dashed curve represents the unmasked modulation-detection threshold level as a function of modulation frequency. Each of the three peaked curves represents the masked modulation-threshold patten for one of the masker-modulation bands (1/2 oct wide) indicated schematically at the bottom of the panel.

Integración temporal

Demostración 8

Integración temporal

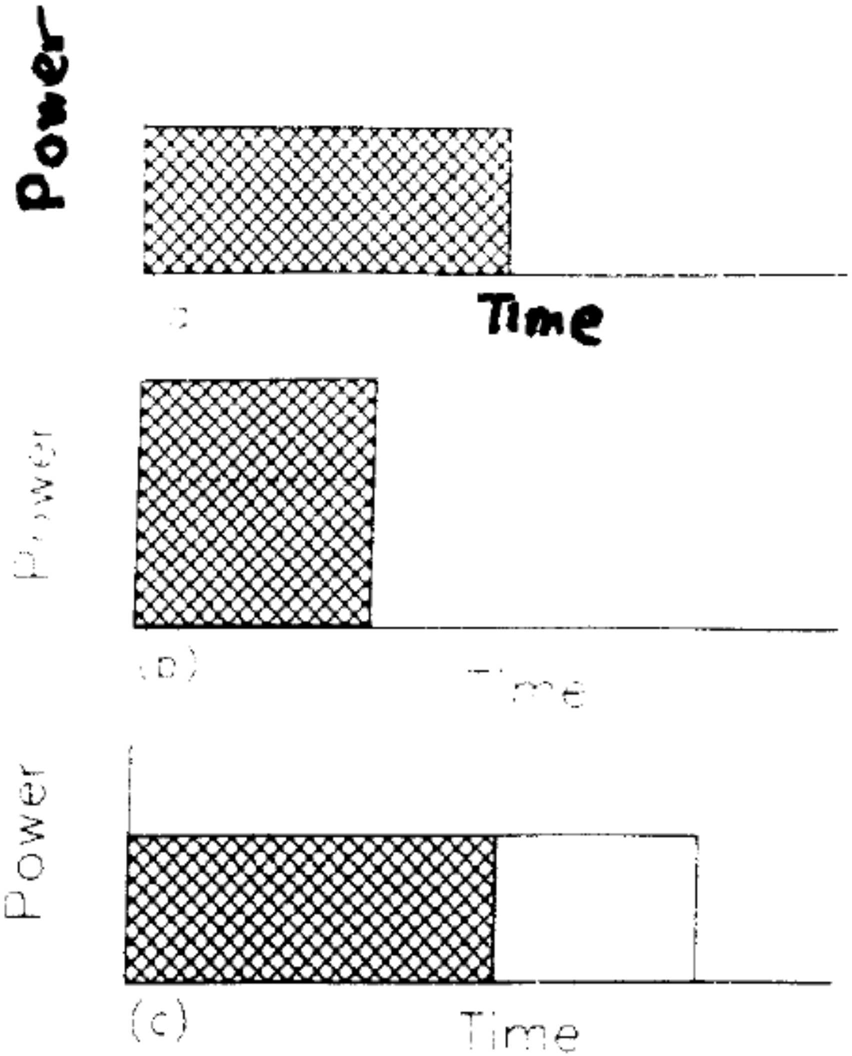


FIGURE 10.4 A schematic diagram depicting the concept of temporal integration. The hashed area in 10.4a represents the energy of a signal required to detect the presence of a tone. In 10.4b the signal is shorter than that in 10.4a; thus the power of this shorter signal must be increased over that in 10.4a to yield a just barely detectable signal. In 10.4c the signal duration is longer than the time required to achieve the necessary energy for detection; thus there is no need to change the power of the signal in order for the signal to remain just barely detectable.

Integración temporal

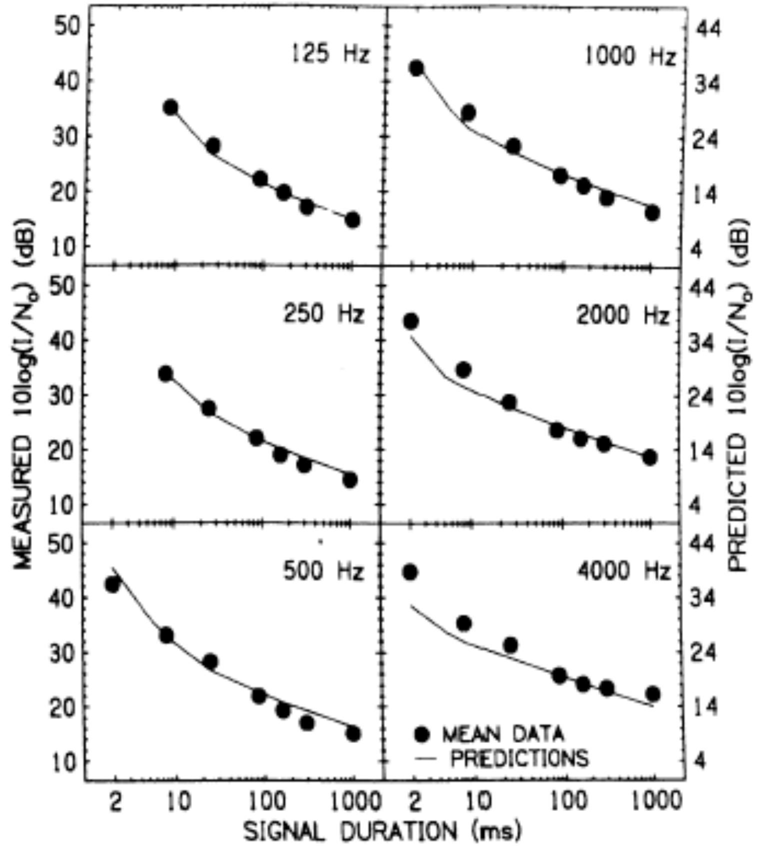


FIG. 4. Measured (filled circles) and predicted (solid curves) time-intensity trades (signal power at threshold as a function of signal duration) at six signal frequencies. The data points are referenced to the left ordinate. The right ordinate is for the model predictions only, which is shifted up by 6 dB relative to the left ordinate.

From Dai & Wright (1996)

Integración temporal y trastornos auditivos

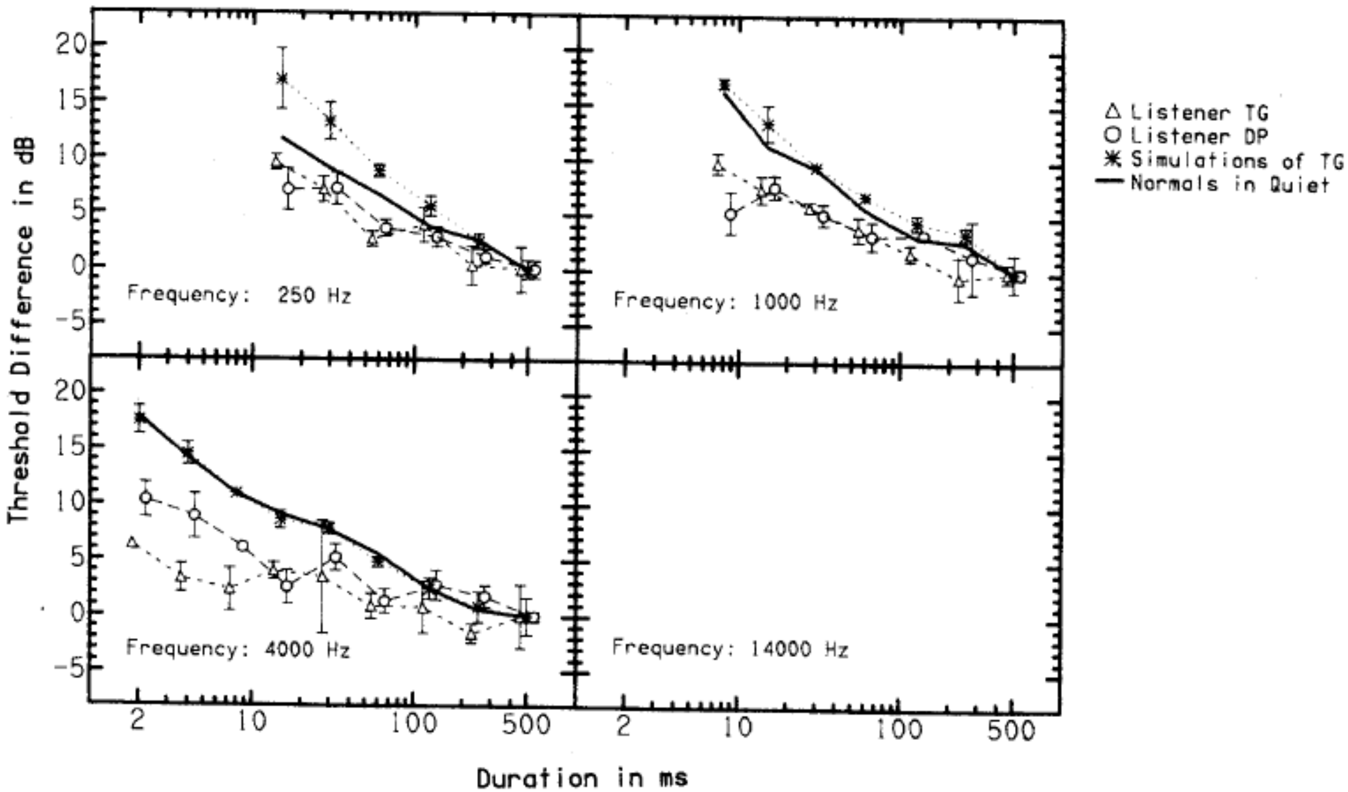


FIG. 3. The threshold difference in dB, normalized *re*: 500 ms, for listeners TG (△) and DP (○) with flat audiograms and the average data from two listeners with simulations (∗) of TG's audiogram. (The average data are shown for the two simulations because results for the individual listeners are similar.) Some symbols are shifted along the abscissa to avoid overlap. The vertical bars show plus and minus 1 standard error. For symbols without bars, plus and minus 1 standard error is less than the size of the symbol. For comparison, the average data obtained in quiet from the five normal listeners from Fig. 2 are shown (solid line).

Paradoja resolución-integración

Suppose I have two pails, one with a large hole at its base and one with a small hole. I quickly dump a quantity of water into each of the pails and measure the flow at the output holes. For the large-holed pail, the output flow as a function of time will closely mimic the input flow: Everything leaks out quickly. For the small-holed pail, water will be accumulated, or integrated, and the flow will persist well beyond the initial dumping. Now I successively dump two quantities of water, with a short interval between, into each of the pails. The flow from the large-holed pail will again mimic the input flow: There will be two relatively discrete bursts of flow. The flow from the small-holed pail, however, will not mimic the input flow because the water cannot flow out fast enough. The discrete bursts of the input will appear as a relatively continuous flow from the small hole; the bursts are smeared together. Thus, the large-holed pail shows good temporal resolution, and the small-holed pail shows good temporal integration.

Auditory models typically employ some type of "leaky integration," essentially, a more sophisticated version of the pail analogy. A leaky integrator can be characterized by its time constant.[5] To account for the resolution data, a 3- to 5-ms time constant (large hole) is required; the temporal integration data require a 200- to 400-ms time constant (small hole). This enormous discrepancy, about two orders of magnitude, is at the heart of the resolution-integration paradox.

LARGE-HOLED PAIL
HAS GOOD
TEMPORAL RESOLUTION

SMALL-HOLED PAIL
HAS GOOD
TEMPORAL INTEGRATION

Hipótesis de multiples-vistas

From Viemeister & Wakefield (1991)

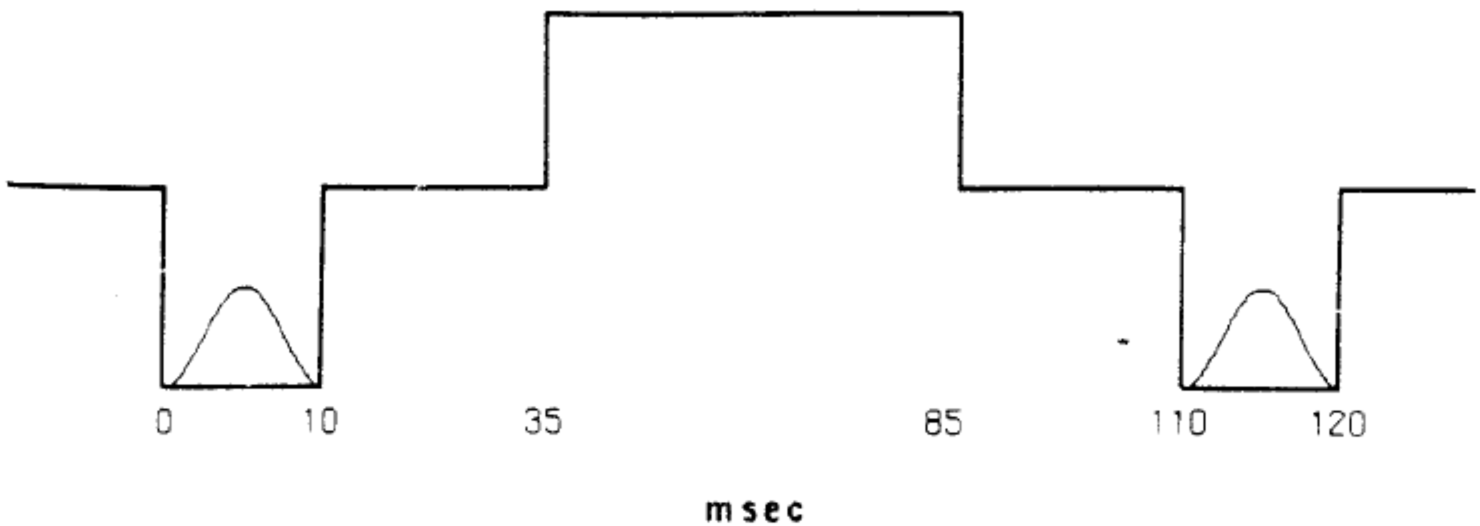
Hipótesis de multiples-vistas

From Viemeister & Wakefield (1991)

Hipótesis de multiples-vistas

From Viemeister & Wakefield (1991)

Hipótesis de multiples-vistas

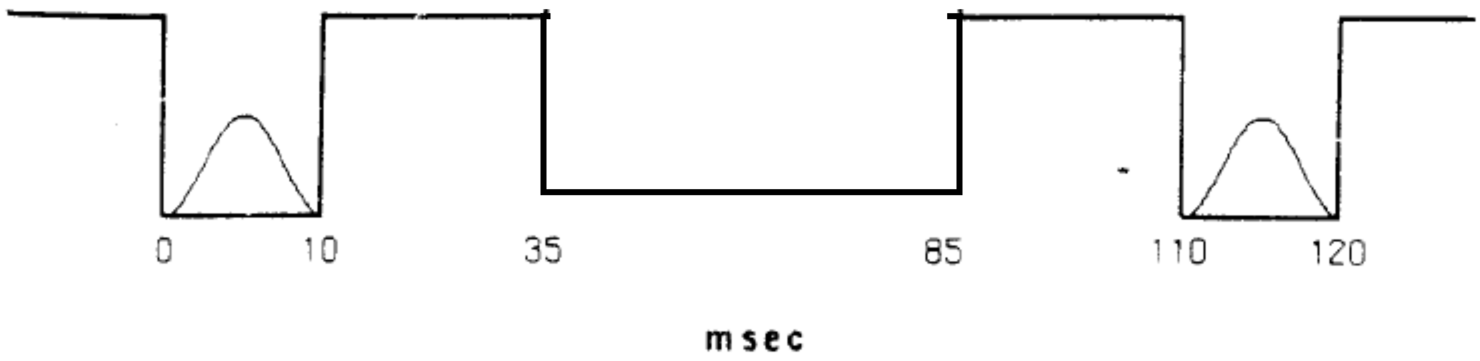


From Viemeister & Wakefield (1991)

Hipótesis de multiples-vistas

From Viemeister & Wakefield (1991)

Hipótesis de multiples-vistas



From Viemeister & Wakefield (1991)

Hipótesis de multiples-vistas

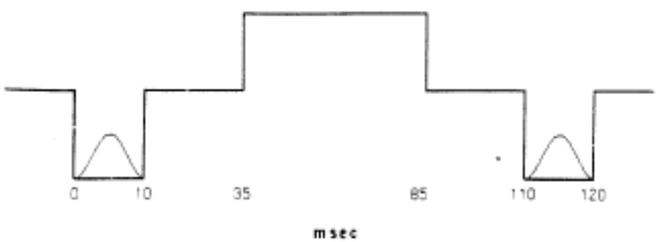


FIG. 2. Envelopes of the noise (heavy lines) and signal (light) during an observation interval. In this example the noise is incremented by 6 dB between the gaps and the signal is presented during both gaps.

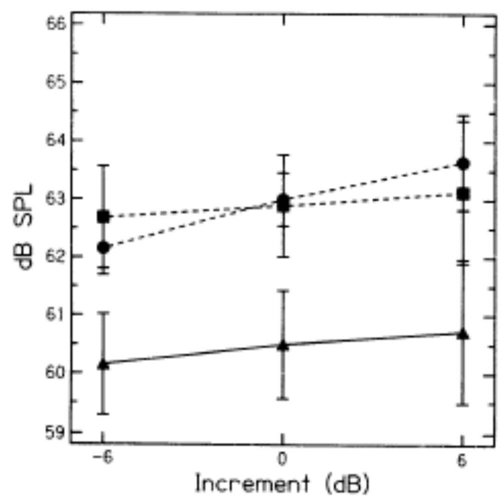


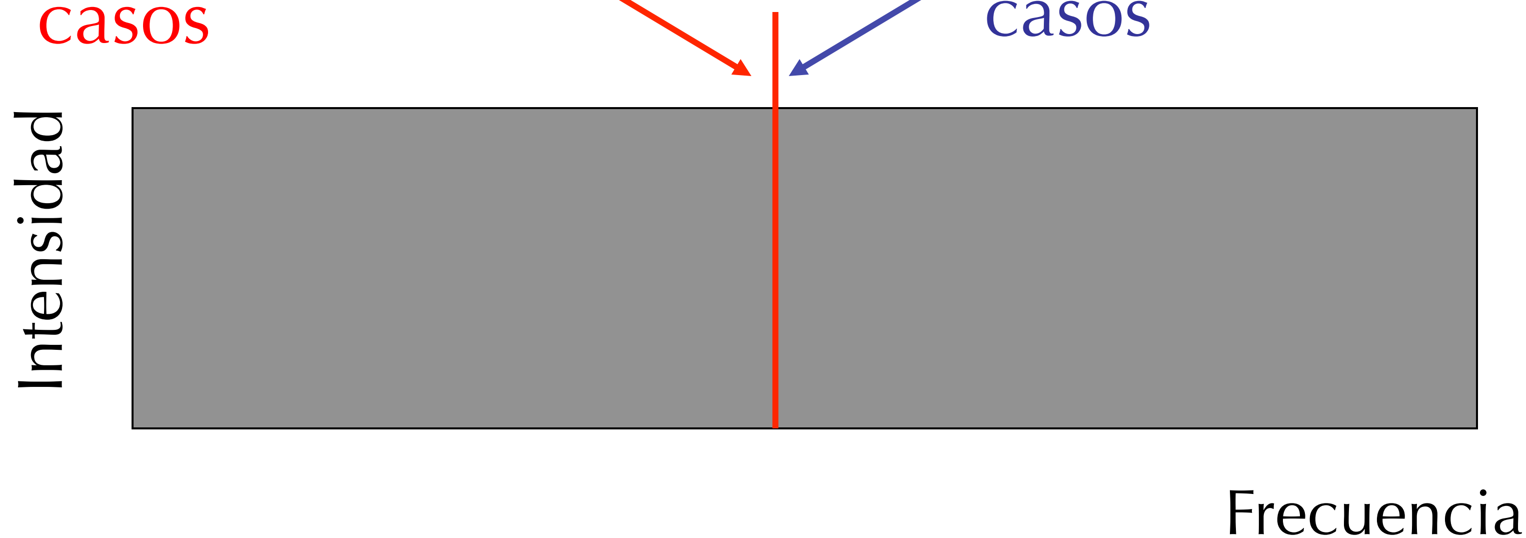
FIG. 3. Average thresholds as a function of the relative level of the intervening noise. Squares: first pulse only; circles: second pulse only; triangles: pulse pair. In quiet, the average thresholds are 26.6 dB for one pulse and 25.2 dB for the pulse pair.

From Viemeister & Wakefield (1991)

Método de señal y prueba

Duración objetivo (esperada) presentada en la mayoría de los casos

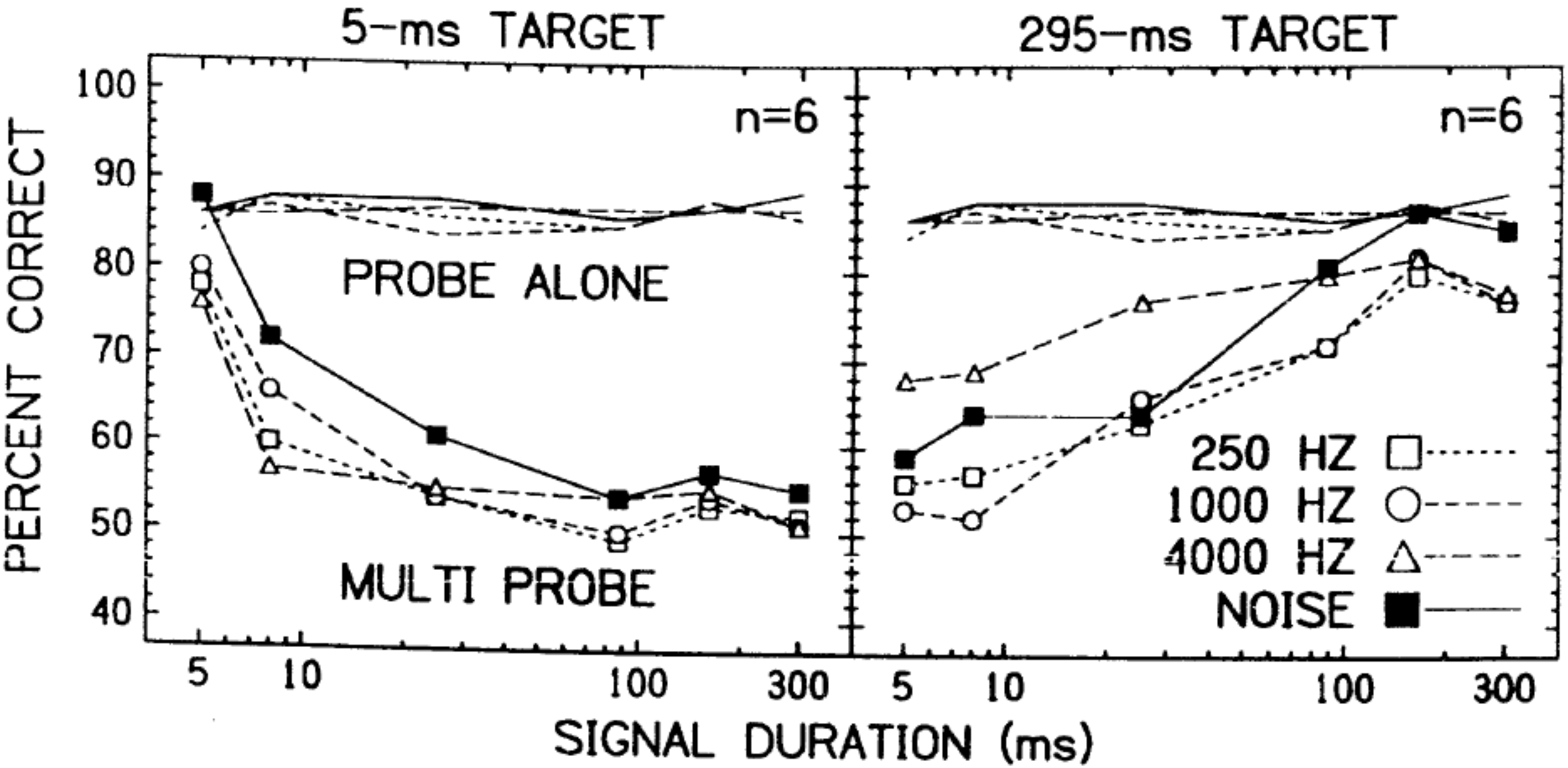
Duración de prueba (inesperada) se presenta en pocos casos



Nivel de señal: se establece para cada señal de forma que el auditor perciba en forma correcta el ~90% cuando la señal estaba presente

Medición: % de detecciones correctas para cada señal

Expectación y detección de señal en base a la duración



From Dai & Wright (1995)

Discriminación de duración

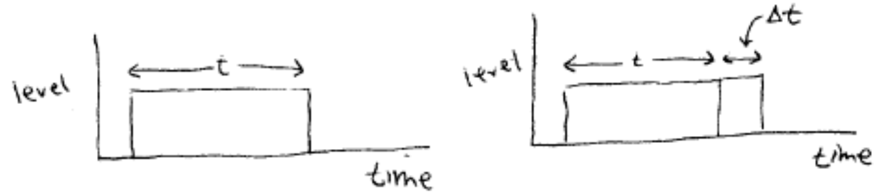
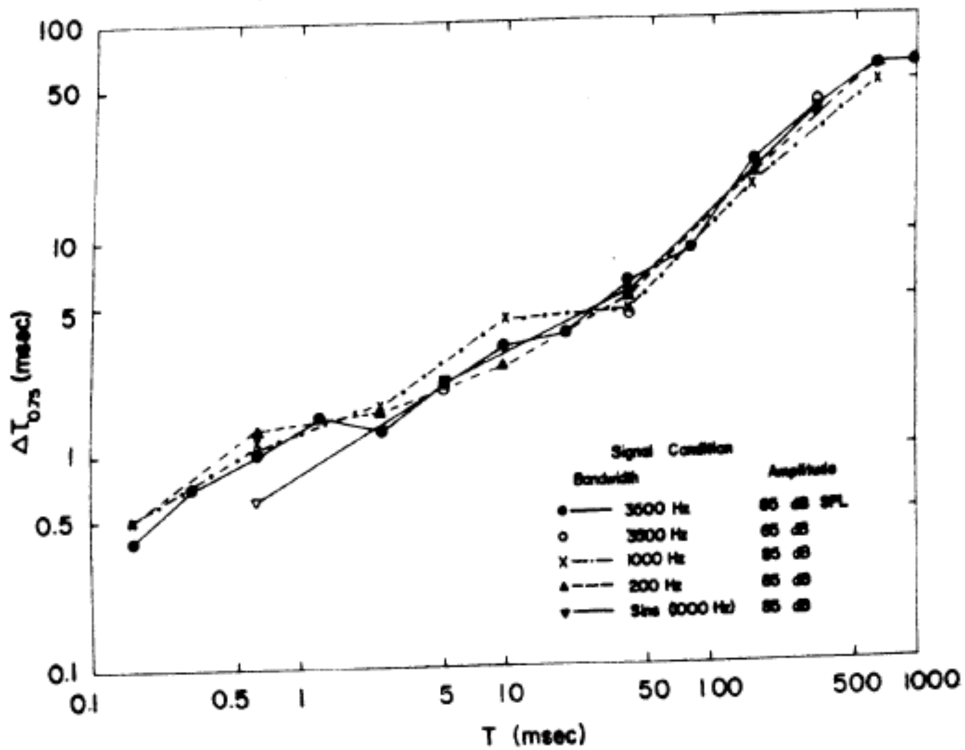
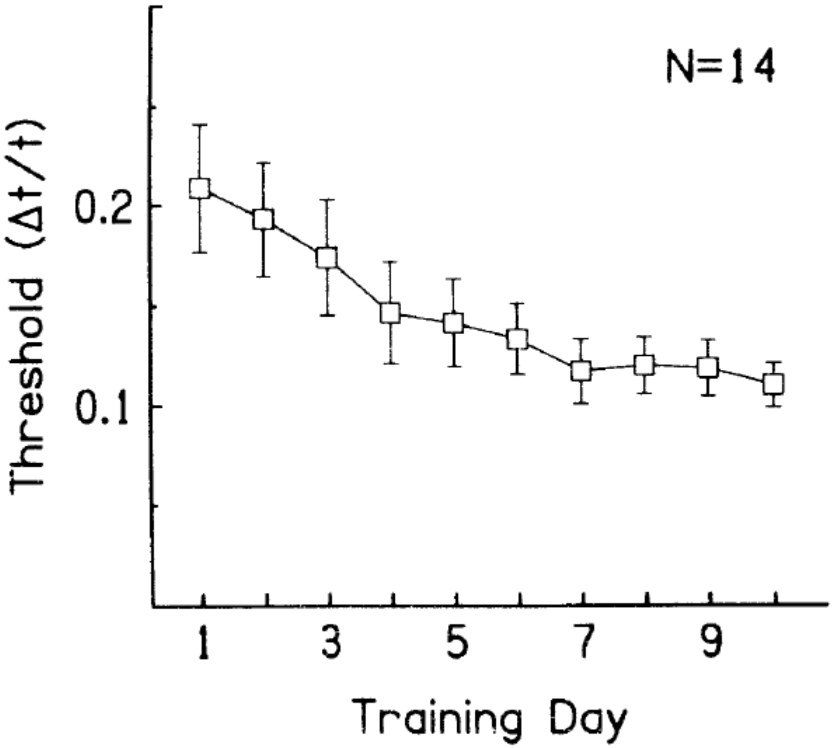
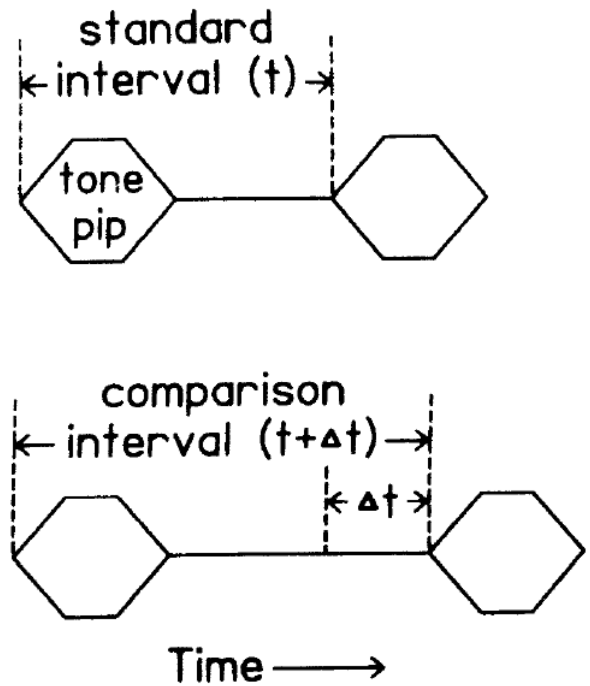


Fig. 1. Estimates of ΔT for 75% discrimination as a function of T. The parameters are signal bandwidth, waveform, and amplitude.



Aprendizaje en la discriminación de la duración





From Wright et al. (1997)

Aprendizaje en la discriminación de la duración

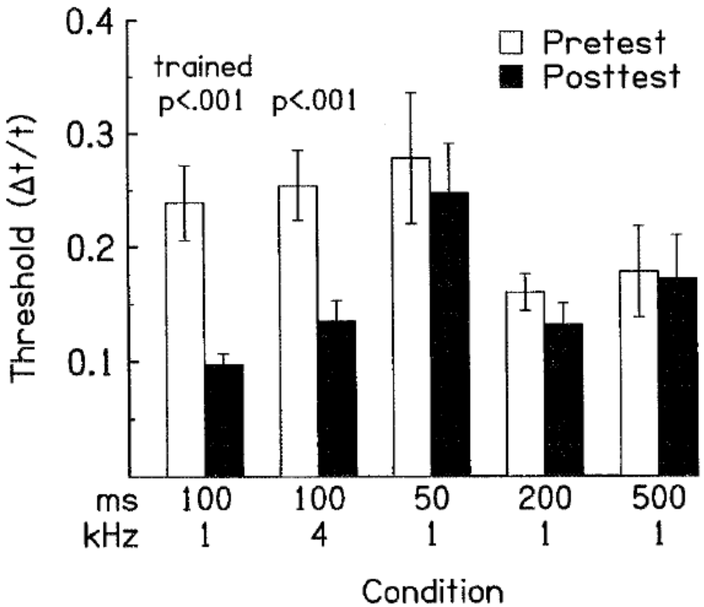
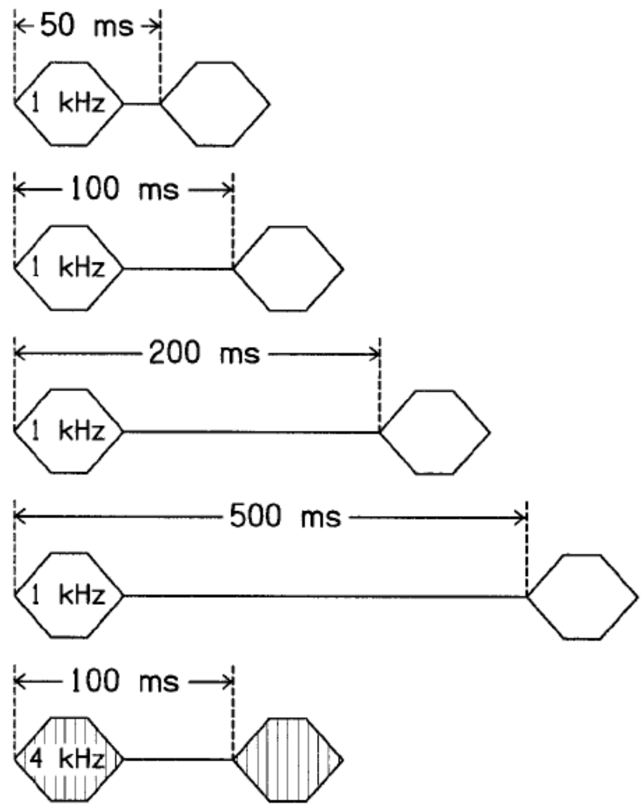


Figure 3. Generalization results. The mean threshold values in the five conditions measured before (*white bars*) and after (*black bars*) the training phase. The difference between the pre- and post-test thresholds in the 100 msec at 1 kHz condition represents the learning resulting from the training in that condition. The corresponding decrease in the 100 msec at 4 kHz condition reflects the generalization of that learning to an untrained frequency. The lack of significant differences between the pre- and post-tests in the remaining conditions indicates that the learning did not generalize to untrained temporal intervals. Threshold is expressed as the Δ millisecond value at threshold divided by the standard duration. The error bars represent ±1 SEM across subjects. Only the data of subjects who individually demonstrated learning during the training phase are included. The data of 11 of the 14 subjects tested are represented in three conditions: 100 and 200 msec at 1 kHz and 100 msec at 4 kHz. The data of five of the six subjects tested are represented in the other two conditions: 50 and 500 msec at 1 kHz.



From Wright et al. (1997)

Discriminación de la duración y trastornos auditivos

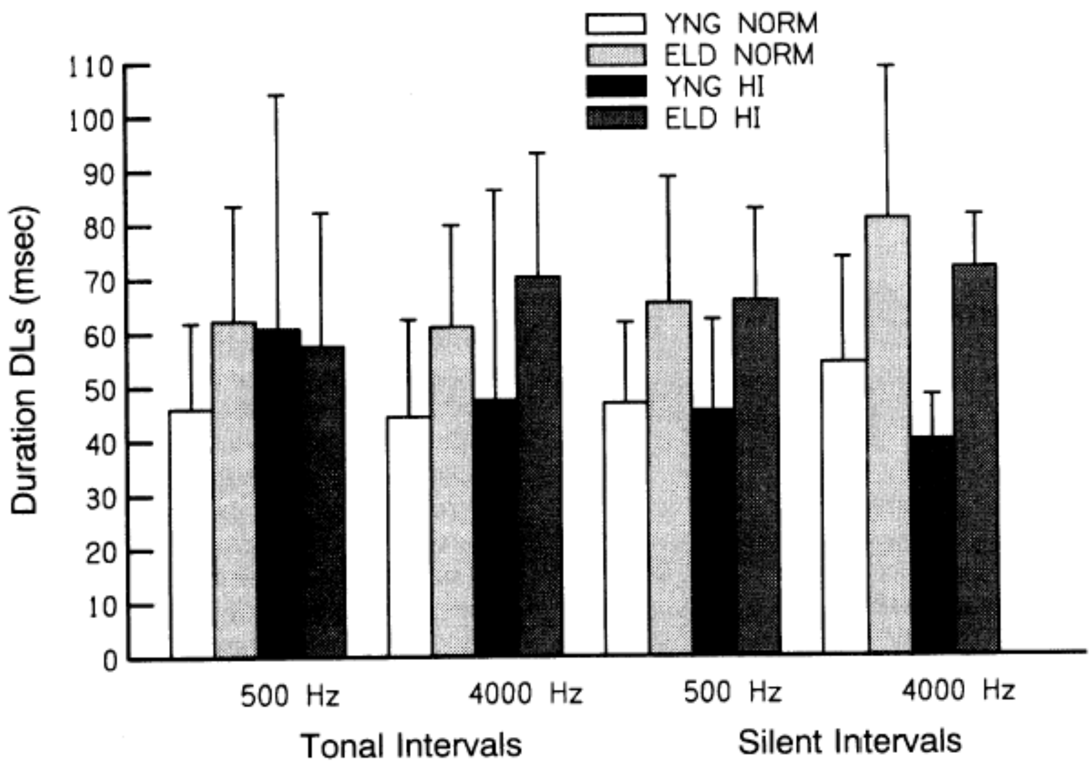


FIGURE 1. Mean duration discrimination performance of the four subject groups for tonal and silent-interval signals at the 250 msec reference duration in two frequency regions. Error bars represent one standard deviation.