Inferior Colliculus

*tonotopic organization*

Low freq

High freq
Structure of Inferior Colliculus

- Paracentral nuclei (somatosensory, descending auditory)
- Dorsal cortex (descending auditory)
- Central nucleus (ICC) (ascending auditory)
Laminar Structure of Central Nucleus

Disc shaped cells arranged in laminae

Stellate Cells cut across laminae

Input fibers from lateral lemniscus
Inferior Colliculus: Convergence and Reorganization

- Parallel processing pathways originating in the cochlear nucleus converge onto the central nucleus of the IC. Some of the projections are direct (monosynaptic); others are via intervening synapses in the SOC and/or the nuclei of the lateral lemniscus.
- IC performs considerable processing: its responses are more complex than those of its inputs.
Inferior Colliculus Projects to Medial Geniculate
Multiple tonotopic fields in cat auditory cortex

Reale & Brugge
Binaural & frequency organization in cat AI
Classification of binaural interactions

Simple scheme (Goldberg and Brown, 1968)
Two letter code, indicating the predominant effect of each ear on neural response, either E (excitation), I (inhibition) or O (no effect). Usually, the contralateral ear comes first.

- OE  Responds to ipsi ear only (e.g. VCN)
- EO  Responds to contra ear only (e.g. MNTB)
- IE  Ipsi ear excites, contra ear inhibits (e.g. LSO)
- EE  Binaurally excited (e.g. MSO)

Detailed scheme (Irvine, 1986)
Two letters followed by slash and third letter (e.g. EO/S). First 2 letters characterize the predominant effect of each ear alone (contralateral first). Third letter gives the type of interaction observed when both ears are stimulated: facilitation (F), suppression (S), or no interaction (O)

- EO/S  Responds to contra ear, NR to ipsi, binaural response smaller than contra.
- EE/F  Responds to each ear alone. Binaurally facilitated
- OO/F  Responds only to binaural stimulation
Discharge Rate as a function of Sound Level: Monotonic or Nonmonotonic
Parvalbumin immunoreactivity defines the auditory core
Core auditory cortex is located in the ventral bank of the lateral sulcus in NHPs (in Heschl gyrus in humans).
Schematic summary of cortical areas related to the processing of auditory, auditory plus visual, and visual stimuli. Based on 2-DG studies. (Poremba et al. 2003)
Auditory cortical subdivisions & processing streams

Macaque

Kaas & Hackett 2000
Auditory Cortical Processing Streams

PATTERN

rostral auditory cortex -> temporal & prefrontal

SPACe

caudal auditory cortex -> parietal & prefrontal

Prefrontal

Belt & parabelt

Core

Thalamus

MGm | MGv | MGd

Adapted from Rauschecker, 1998
**Temporal, parietal, & frontal cortex**
... which mediate space perception, auditory memory and other functions

**Parabelt**
Rostral and caudal divisions lateral to the belt

**Belt**
Multiple fields less precisely cochleotopic and more responsive to complex stimuli than tones.

**Core**
Cochleotopically organized fields that generally display clear responsive to pure tones

**MGB Ventral nucleus**
Principal division of the auditory thalamus

---

Human auditory cortex

- Auditory cortex
- Medial geniculate nucleus
- Inferior colliculus
- Cochlear nucleus
- Superior olivary nucleus
- Brainstem
- Cochlea
- Pons
areas 41 & 42

Primary auditory cortex
Secondary auditory cortex

Corresponds to apex of cochlea
Corresponds to base of cochlea

Sylvian Fissure

Other Cortical Areas

MGB MGB & IC MGB
Speech processing — the Wernicke-Geschwind model

**Broca's area**
*Speech production and articulation*
- stores information needed for speech production
- responsible for programming the motor cortex to move the tongue, lips and speech muscles to articulate words.

**Wernicke's area**
*Language comprehension*
- Region of angular gyrus specialized for processing human speech, and receiving input from belt/parabelt auditory cortex.
- Stores information for arranging learned vocabulary into meaningful speech
- Damage leads to **Wernicke's Aphasia** — speech may be fluent yet confused and nonsensical. Speech comprehension is also impaired.
- Damage limited to the posterior language area leads to transcortical sensory aphasia (*patient can’t understand but can repeat words*).

- Connection (arcuate fasciculus) between Wernicke & Broca’s areas enables patients with transcortical sensory aphasia to repeat words despite not understanding them.
- Damage to this connection disrupts speech, but speech sounds, articulation, and comprehension are retained (*conduction aphasia*).
Amplitude-modulated tones

- Percepts generated by AM sounds vary with modulation frequency (recall Spottiswoode’s description of the sound quality of beats)
Representation of AM auditory nerve

FIG. 1. A: superimposed waveforms of an unmodulated 1,000-Hz tone (thin line) and the same tone sinusoidally amplitude modulated (AM) (thick line) at 100% with a modulation frequency of 100 Hz, according to Equation 1. Dashed lines indicate the envelope. The amplitude is referenced to the peak amplitude of the unmodulated tone. B: idealized spectrum of the AM tone in A. At 100% modulation, the amplitude of the sidebands is half that of the carrier, i.e., a difference of 6 dB. C: average response in the form of a poststimulus time (PST) histogram of a nerve fiber to the signal shown in A (stimulus duration, 50 ms). D: spectrum of the PST histogram in C. The components at carrier frequency ($f_c$) and $f_c$ ± modulation frequency ($f_m$) indicate that there is phase-locking to the fine-structure of the stimulus waveform. The component at $f_m$ is prominently present in the response but is absent in the stimulus (B). The small circle on the ordinate indicates the average firing rate.
Transformations in representation of AM auditory nerve (dashed lines) and cochlear nucleus (solid lines)

enhancement of envelope synchronization and extended dynamic range

Emergence of bandpass tMTFs.
Transformations in representation of AM

Brainstem to midbrain

A. COCHLEAR NUCLEUS

SYNCHRONIZATION vs. $f_m$

B. INFERIOR COLICULUS

tMTF

C. RESPONSE RATE vs. $f_m$

D. rMTF

$\text{f}_m$
Transformation in representation of AM
Thalamus to cortex
Topography in representation of AM

INFERIOR COLLICULUS

Langner and Schreiner 1988,
Schreiner and Langner 1988:
Langner and Schreiner 1988 (first of two papers)

**FIG. 2.** Synchronized activity of 3 single units and 1 multiple unit (IC759) in response to 10 different modulation frequencies $f_m$. Modulation frequencies are indicated at the right of the PSTHs (binwidth: 40 ms; stimulus duration: 150 ms). Bars at the bottom left indicate the number of spikes during 100 repetitions for units IC759, IC735, and IC111 and during 30 repetitions for unit IC759. The carrier frequency $f_c$ for the amplitude modulation were at the CT of the units. They are indicated above each block of PSTHs. The arrows point to the approximate BMF of each unit.

**FIG. 1.** Distribution of best modulation frequencies (BMFs) for 536 multiple units and 226 single units in the inferior colliculus. About one-half of the recordings revealed BMFs between 30 and 100 Hz.
Fig. 3. Normalized modulation transfer functions (MTFs) for the firing rate of typical bandpass units of the ICC. Numbers above the MTFs indicate the unit number; the CF of the unit that also serves as carrier frequency of the AM stimuli and the maximal firing rate corresponding to the tip of the MTF. Single units are indicated by open circles, multiple units by closed circles.
FIG. 2. Reconstructed location of 3 frequency band laminae. The reconstruction is based on the depth recorded for each recording locations relative to the surface, the surface coordinates, and the inclination of the electrode trajectory. The resulting recording locations were projected onto appropriate coronal sections of the IC. The recorded locations had CFs of 3 ± 0.5 and 12 ± 1 kHz, respectively, indicated beside the isofrequency contours. The dashed lines correspond to contours obtained in 1 animal. The approximate location of the lateral (L) and the combined central and medial (C) subdivisions of the central nucleus of the ICC is outlined (22, 23). The numbers on the left side of each section indicate the distance from the caudal pole of the IC in micrometers.

FIG. 3. Representation of BMF's within a frequency band lamina of the ICC. A: 3-dimensional projection of the BMF distribution within the frequency band lamina of 3 ± 0.5 kHz (IC 1/26). The short axis represents the rostrocaudal spatial dimension of the mapped area of the frequency band lamina. The long axis represents the intermedial dimension. The third spatial dimension, i.e., the dorsoventral dimension, has been omitted. The elevation of the gridted surface is proportional to the BMF obtained at or near those line intersections that are marked by a dot and interpolation of the intermediate points of the grid. The grid is based on n = 43 evenly spaced measurements of the BMF (see Fig 8.4). Smoothing factor = 1. To enhance general features of the distribution, a slight smoothing operation has been employed in some of the 3-dimensional projections of a functional parameter and its location within the ICC. In principle, the degree of smoothing can vary between a smoothing factor of 1 = no smoothing and 0 = strongest smoothing (flat surface). B: contour plot of the BMF distribution in a frequency band lamina of the ICC (IC 1/26). The contour lines connect sites of neuronal responses exhibiting the same BMF (iso-BMF contours) corresponding to the spatial interpolation given by the 3-dimensional projection shown in A. The outermost contour corresponds to a BMF of 50 Hz. Each following contour represents an increase of BMF by 50 Hz.
Masking — *an upward shift in the hearing threshold of a weaker tone by a louder tone, dependent on the frequencies of the two tones.*

If the ear is exposed to two or more different tones, one tone may mask the others. Pure tones, complex sounds, narrow and broadband noise all show differences in their ability to mask other sounds. Masking of one sound can even be caused by another sound that occurs a split second before or after the masked sound.

**How can masking occur?**

- *Excitation:* Swamping of neural activity due to masker.
- *Suppression:* Reduction of response to target due to masker.
Frequency (kHz)
**Masking - Psychophysical Tuning Curves**

- present a pure tone at 10 dB
- mask it with noise of varying centre frequencies
- At what intensity will the noise prevent
detection of the pure tone?

![Graph showing frequency vs. intensity for masking](image_url)
**Masking** - *Psychophysical Tuning Curves*

![Graph showing masking effects across different frequencies and intensity levels.](image)

- **Frequency (kHz)**: 0.1, 0.2, 0.5, 1, 2, 4, 10
- **Intensity of Mask (dB SPL)**: 0-100

The graph illustrates how masking increases with higher frequency and intensity levels, highlighting the psychophysical tuning curves for auditory perception.
Masking - *Psychophysical Tuning Curves*

![Graph showing masking effects and psychophysical tuning curves.](image-url)
Some conclusions from Masking Experiments

• Pure tones close together in frequency mask each other more than tones widely separated in frequency.

• A pure tone masks tones of higher frequency more effectively than tones of lower frequency.

• The greater the intensity of the masking tone, the broader the range of frequencies it can mask.

• Masking by a narrow band of noise shows many of the same features as masking by a pure tone; again, tones of higher frequency are masked more effectively than tones having a frequency below the masking noise.

• Masking of tones by broadband ("white") noise shows an approximately linear relationship between masking and noise level (that is, increasing the noise level 10 dB raises the hearing threshold by the same amount). Broadband noise masks tones of all frequencies.
Critical bands demonstrated by masking

Results suggest that we can “tune into” the region around 2 kHz, and that only masker energy around 2 kHz affects our ability to perceive the tone.

The bandwidth of effective masking is the critical band (Fletcher, 1940).
**Masking** — *an upward shift in the hearing threshold of a weaker tone by a louder tone, dependent on the frequencies of the two tones.*

**Forward:** A tone can be masked by a sound that ends a short time (up to about 30 ms) before the tone begins. Forward masking suggests that recently stimulated cells are not as sensitive as fully-rested cells.

**Backward:** A tone can be masked by a noise that begins up to 10 milliseconds later, although the amount of masking decreases as the time interval increases (Elliott, 1962). Backward masking apparently occurs at higher centers of processing where the later-occurring stimulus of greater intensity overtakes and interferes with the weaker stimulus.

Basilar membrane tuning produces asymmetry; a tail extends toward the high-frequency end. Thus it is easier to mask a tone of higher frequency than one of lower frequency.

As the intensity of the masking tone increases, a greater part of its tail has amplitude sufficient to mask tones of higher frequency. This "upward spread" of masking tends to reduce the perception of the high-frequency signals that are so important in the intelligibility of speech.
Auditory Scene Analysis
(sound source determination)

• Sounds from multiple sources arrive at the auditory periphery as one complex sound field. Individual sources are then resolved from spectro-temporal patterns.

• A single source can stimulate a wide region of the cochlea, with multiple sources simultaneously stimulating the same peripheral frequency channels.

• Determining the source of a sound depends on analytic processing to segregate spectral components according to source.

• An auditory stream refers to the perception of a series of sounds as the output of a single sound source. Separation of the auditory streams is affected by time and frequency separation.
Waves from environmental sound sources mingle before reaching our ears.
Auditory system disentangles the sources.
Yost calls this “Sound Source Determination”
Mechanisms

Simultaneous

Harmonicity,
Onset/offset,
Co-modulation

Sequential
Auditory ‘streaming’

Level

Clarinet
Voice

Freq
Deviations from harmonicity promote segregation

*If a harmonic is mistuned by > 2-3%, it stands out perceptually*

(Moore et al., 1985, 1986; Hartmann et al., 1990)

---

**stimulus**

---

**percept**

1 sound  
*pitch = 200 Hz*

2 sounds  
*Harmonic (pitch = 200 Hz) + Pure tone (618 Hz)*
Onset asynchronies promote perceptual segregation

**stimulus**

**percept**

1 sound

2 sounds

(A four-tone cluster)

- Tone 1 = 800 Hz
- Tone 2 = 700 Hz
- Tone 3 = 900 Hz
- Tone 4 = 800 Hz
Co-modulation promotes perceptual segregation

Coherent FM promotes the fusion of harmonics
Darwin et al. (1994)

**stimulus**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**percept**

2 sounds
Auditory streaming

repeating triplet of tones.

When all the tones frequencies are similar, you can hear a single stream of sound, with a galloping rhythm.

But when the "A" and "B" tones are far apart, you hear 2 streams and the gallop disappears.
Cues for sound segregation

Spectral separation
• Sometimes two or more sounds having different spectral components can be resolved by frequency selectivity.

Harmonicity
• A sound consisting of harmonics is usually perceived as a single pitch equal to the fundamental frequency even if the fundamental is absent in the sound spectrum.
• Fusion of spectral components into an image or entity is described as a pitch.

Profile analysis & Timbre
• Sound spectra are characterized by their intensity variation as a function of frequency.
• The peripheral auditory system is sensitive to changes in spectral profile despite variation in overall sound level.
• Timbre allows discrimination of one sound from another when both have the same pitch, loudness, and duration. Different spectra give rise to different timbres.

Spectral Modulation
• Changing a frequency modulation or modulating a complex set of harmonics facilitates segregation.

Temporal separation & Asynchrony
• Multiple sound sources are often asynchronous (onset or offset) & temporally complex, aiding segregation.

Spatial separation (e.g., cocktail party effect)
• IID and ITD are cues for sound source determination, particularly with competing signals.
• Increasing separation between masker and signal improves discrimination. e.g., Two individuals speaking simultaneously at various locations.

Temporal modulation
• Coherence of change among spectral components may help the auditory system to identify those components as originating from a single source (can be AM or FM).
• A coherently modulated cue band offers relief from masking (comodulation masking release).
• A modulated masker can impair detection of a modulated signal (modulation detection interference).

Informational masking
• The inverse relationship between pattern complexity and target resolution is called informational masking.
• Variation in target location in a pattern, total number of elements in the pattern, and stimulus uncertainty affect the information content of a sequence.