Dai & Shinn-Cunningham, 2016:

Abstract:

* Top-down executive control modulates the sensory representation of sound in the cortex
* This representation depends also on the coding fidelity of the peripheral auditory system 🡪 which may contribute to individual differences in performance
* **Selective auditory attention paradigm**
* Measured ERPs from the scalp 🡪 which reflect cortical responses to sound
* Behavioral scores
* The stimulus conditions varied 🡪 to alter the degree to which performance might be limited due to fine stimulus details vs due to control of attentional focus
* Behavioral differences are correlated with the strength of attentional modulation of ERPs

Introduction:

* Auditory neuropathy
* Low-spontaneous rate auditory nerve fibers are more susceptible to damage from noise exposure than high-spontaneous rate fibers, which respond at hearing threshold
* AN driven by ageing and noise exposure
* A likely contributor to individual differences in the encoding of subtle spectro-temporal features of supra-threshold sound
* **If a listener cannot segregate sounds based on such features 🡪 trouble directing attention to sound of interest**
* The accuracy of how the early sensory system encodes the timing patterns of sounds that are just audible significantly affects an individual's capability to effectively communicate in difficult situations.
* The envelope following response (EFR) is an auditory evoked potential that reflects the brain's neural response to the temporal envelope of an auditory stimulus. The temporal envelope refers to the changes in intensity or amplitude of a sound wave over time
* **General cognitive ability, aging, health affect the ability to understand speech in complex settings**
* Studies on sensory coding deficits and failures of selective attention, were designed so that the features that distinguished the target source and competing speech streams differed modestly 🡪 **15 degrees separation**
* **Performance in such paradigms depends on sensory coding fidelity (detecting changes in pitch and amplitude in fine detail) and on attentional control, which modulates the neural representation of the target sound, enhancing the signal**
* Performance correlates with how strongly cortical responses to the competing sound are modulated by attentional focus 🡪 **Choi et al., 2014**
* **In addition to differences in subcortical coding fidelity**
* Tasks where peripheral coding does not limit performance 🡪 easy to segregate sounds
* Perhaps listeners with the most robust peripheral encoding of supra-threshold sounds perform better, and it does not necessarily have anything to do with central control
* Two experiments: to examine subcortical sensory coding fidelity and strength of attentional modulation of cortical responses and behavior performance
* EEG 🡪 ERPs and EFRs
  + Compared the magnitude of ERPs to the same mixture of auditory inputs when listeners attend to one stream vs the other 🡪 quantify the degree of top-down control of selective attention for individual listeners
  + EFR 🡪 degree to which the brainstem phase locks to the temporal periodicities in an input acoustic stimulus 🡪 high-frequency modulation
* One stream from the left, one from the right
* Target changed from trial to trial 🡪 visual cue to indicate target
* Tasks and auditory stimuli differed in order to try and isolate different factors contributing to individual differences
* **Experiment I**
* **Aim:** To understand how well people can focus on one specific sound when there are multiple sounds to listen to. Specifically, they wanted to see if the quality of how sounds are processed in the early stages of hearing (like in the brainstem) affects a person's ability to pay attention to a particular sound.
* **What They Did:** Participants listened to two streams of sounds coming from different directions. One stream was the target that they had to pay attention to, and the other was a distraction. The target stream had some notes that changed in pitch (deviants), and participants had to count these changes.
* **Expectations:** The researchers thought that people's ability to do the task would relate to how well their early auditory system (subcortical areas like the brainstem) processes the sounds. They also thought that how well people can control their attention (not getting distracted by the other stream) might play a role.
* **Experiment II**
* **Aim:** This experiment aimed to dive deeper into how people focus their attention on sounds by making the task more challenging. They wanted to see not just how early sound processing affects attention, but also how the brain's higher-level control of attention comes into play.
* **What They Did:** Similar to Experiment I, participants listened to sound streams, but this time there were three streams, and the task was to identify the pattern (contour) of the melody in the target stream (whether the notes were going up, down, or zigzagging). The target stream was less distinct from the distractions, making it harder to focus on.
* **Expectations:** The researchers expected that in this more challenging scenario, the ability to control attention (focusing on the target and ignoring distractions) would have a more significant impact on performance. They still thought that the quality of early sound processing would matter, but they also believed that the brain's attention control mechanisms would play a bigger role because of the increased task difficulty.
* **Why These Experiments:** By varying the task difficulty and complexity of the sound environment, the researchers aimed to uncover different factors that contribute to how well we can focus on specific sounds in noisy environments. Understanding this can help explain why some people are better than others at tasks like following a conversation in a loud room and can inform strategies or technologies to assist those who struggle with such auditory attention tasks.

**Discussion:**

**Experiment I:**

The group as a whole had less phase locking for the fully masked Stream A (97 Hz) compared to the partially masked Stream B (159 Hz).

The expected EFR (brainstem response) strength did not appear to be affected significantly by selective attention in Experiment I, which is consistent with other findings suggesting that selective auditory attention does not influence subcortical responses measured by EFRs.

**Peripheral Coding Impact**: Individual differences in performance were found to correlate with the strength of subcortical coding, as measured by the brainstem responses. Participants who could better encode the fine details of sound due to higher subcortical coding strength performed better at the task.

**Attention vs. Sensory Coding**: While selective attention did modulate cortical responses, it did not predict performance in this experiment. The ability to segregate the sound streams and focus on the target was not limited by attention but rather by the sensory coding fidelity of the auditory system.

**Coding Fidelity as Performance Predictor**: The study concluded that differences in sensory coding fidelity at the brainstem level, not the strength of attentional control, were predictive of performance when the task was primarily limited by sensory coding fidelity.

**Experiment II:**

**Focus on Attention Control**: This experiment was designed to place greater demands on attention control by introducing a third sound stream and more complex task requirements. The hypothesis was that under these conditions, the central control of attention would play a more significant role in determining performance.

**Attention and Performance**: In contrast to Experiment I, individual differences in performance were related not only to subcortical coding strength but also to the strength of attentional modulation of cortical responses. This suggests that in more demanding listening environments, the brain's ability to control attention becomes more critical.

**Peripheral vs. Central Processing**: The findings indicated that both peripheral auditory processing and central attentional control contribute to performance, depending on the complexity of the listening task.

**Complex Stimuli and Attentional Modulation**: The study found large inter-subject differences in the ability to focus attention under more complex auditory conditions, and these differences were consistent across different tasks and stimulus configurations.

If a participant was good at modulating their attention to enhance their brain response to one stream, they were likely good at doing the same for the other stream.

Dai & Shinn-Cunningham, 2018:

* **Selective auditory attention depends on the formation of auditory objects 🡪** without properly parsing a scene, the brain cannot suppress responses to unattended sounds
* **Selective attention to auditory objects is influenced by both top-down control and bottom-up salience**
* Auditory attention operates as a form of sensory gain control 🡪 enhances the representation of attended object 🡪 these modulations of auditory-evoked responses can be decoded from EEG

The article, titled "Sensorineural hearing loss degrades behavioral and physiological measures of human spatial selective auditory attention," investigates the impacts of sensorineural hearing loss (HI) on spatial selective auditory attention. Here’s a summary:

**Objective and Hypothesis:** The study aimed to explore whether individuals with sensorineural hearing loss have difficulties in selectively attending to sounds based on their spatial location. The authors hypothesized that hearing loss could hinder the ability to focus attention, affecting the filtering out of distracting sounds, which is essential for understanding speech in noisy environments.

**Methods:** The study involved both normally hearing (NH) individuals and those with sensorineural hearing loss. Participants were asked to identify melodies presented with competing melodies from different spatial locations. The study used behavioral tasks and electroencephalography (EEG) to measure attentional control and brain responses. Participants asked to report whether the target stream was ascending, descending or zigzag (Leading or Lagging, left or right).

**Results:** The findings revealed that compared to NH listeners, HI listeners had:

* Poorer sensitivity to spatial cues.
* Lower performance in the selective attention task.
* Weaker attentional modulation of cortical responses (as seen through EEG).
* Less ability to suppress responses to unattended sounds, particularly in complex listening scenarios.
* AMI increases across time for NH listeners

Additionally, the research demonstrated that across individuals, performance on selective attention tasks correlated with both spatial cue sensitivity and the strength of attentional modulation of cortical responses.

**Conclusion:** The study concluded that sensorineural hearing loss affects spatial acuity, which in turn impacts selective auditory attention and the ability to suppress irrelevant sounds based on their location. These difficulties contribute to the challenges faced by HI individuals in social settings with background noise.

The paper underscores the intricate relationship between hearing loss, attention control, and spatial processing, and it calls for further research into technologies that might assist HI listeners based on these insights 🡪 significant correlation across participants between performance on the selective attention task, sensitivity to spatial cues (measured by ITD thresholds), and the strength of attentional modulation of cortical responses

while our ears might be able to detect very fine spatial details, our brain's ability to focus attention might not always match that precision, leading to a broader "attentional focus" when we're trying to listen to sounds from a specific location

adjacent streams were separated by 205microseconds or by 699 (small and large ITDs) 🡪 smallest possible for NH is 50μs

The article "Auditory Attention Detection via Cross-Modal Attention" by Siqi Cai, Peiwen Li, Enze Su, and Longhan Xie, focuses on enhancing the detection of auditory attention in environments where multiple speakers are present, a scenario often described as the "cocktail party problem". The research introduces a novel approach known as Cross-Modal Attention-based Auditory Attention Detection (CMAA), which utilizes both audio and EEG signals to improve the detection of which speaker an individual is focusing on.

**Key Points:**

* **Background and Motivation**: The study addresses the challenge individuals, particularly those with hearing impairments, face in distinguishing and focusing on a single speaker among multiple competing ones. While hearing aids help, they often fall short in complex auditory scenarios.
* **CMAA Model**: The proposed model employs a cross-modal attention mechanism to dynamically weigh and integrate features from both EEG and audio signals. This approach is designed to better capture the correlation between the auditory stimulus and the brain's response to it, enhancing the system's ability to detect where auditory attention is directed.
* **Experiment and Results**: The researchers conducted experiments using a publicly available dataset and compared the performance of the CMAA model against traditional linear models and state-of-the-art non-linear approaches. The CMAA model achieved notable accuracy improvements in detecting the attended speaker under various conditions, including different decision window lengths and acoustic environments.
* **Implications**: The findings suggest that incorporating a cross-modal attention mechanism can significantly enhance auditory attention detection systems. This has potential applications in developing more effective hearing aids and other auditory devices that can adapt to the user's focus in real-time, improving speech comprehension in noisy settings.
* while linear AAD models can decode where a listener's attention is directed based on auditory stimuli and brain responses, their effectiveness decreases when trying to do this rapidly, within the timeframe that aligns with natural human attention switching. 🡪 humans can switch attention within 1s.

**Core Concepts:**

1. **Cross-Modal Attention Module**: This is the central component of the CMAA model. "Cross-modal" refers to the integration and processing of information from different sensory modalities, such as auditory (sound) and visual (sight) or, in the context of this model, auditory stimuli and corresponding EEG brain responses.
2. **Top-Down and Bottom-Up Modulation**:
   * **Top-Down Modulation**: This involves cognitive processes that are driven by the brain's higher-order functions, such as attention, expectations, and previous knowledge. For example, if you're trying to listen to a specific person in a noisy room, your brain uses top-down processing to focus on sounds that match your expectations of that person's voice.
   * **Bottom-Up Modulation**: This is driven by the sensory input itself, where external stimuli capture attention due to their inherent properties, like a loud noise suddenly drawing your focus.
3. **Dynamically Assigning Weights at Run-Time**: The model adjusts its processing strategy in real-time based on the current stimulus. "Weights" refer to the importance or influence assigned to different pieces of information. By adjusting these weights on the fly, the model can more effectively focus on relevant signals and ignore distractions, mimicking the way our brain prioritizes certain sensory inputs over others depending on the context.
4. **According to the Input Stimulus**: The model's adjustments are not random but are directly influenced by the characteristics of the incoming sensory information. This means the model is responsive to the changing environment, similar to how our attention might suddenly shift to a loud sound or an unexpected movement.

**Key Matrix Concepts in CMAA:**

1. **Feature Matrices**:
   * **Audio Feature Matrix**: Imagine a table where each row represents a different moment in time during an audio recording, and each column represents a different feature of the sound at that time (like pitch, volume, etc.). This matrix captures the detailed characteristics of the auditory stimuli.
   * **EEG Feature Matrix**: Similarly, this table has rows for different time points, but the columns represent various EEG signal features captured while the subjects listened to the audio. These features might include different brainwave patterns associated with attentional focus.
2. **Weight Matrices**:
   * In machine learning, weights are numerical values used to emphasize or de-emphasize certain inputs (features) when making predictions. In CMAA, there are likely weight matrices associated with both the audio and EEG features, determining how much influence each feature should have on the model's output.

**Cross-Modal Attention Mechanism:**

* **Combining Audio and EEG Data**: The cross-modal attention mechanism is about dynamically blending the information from the audio and EEG feature matrices. It's like having two sets of ingredients (audio features and EEG features) and deciding how much of each to use in a recipe (the model's prediction) based on what you're cooking (the current listening situation).
* **Calculating Attention Weights**: The model calculates attention weights, which are values that dictate how much attention (importance) is given to each feature when making a prediction. This might involve multiplying the feature matrices by the weight matrices to emphasize certain features over others.
* **Dynamic Adjustment**: These attention weights aren't static; they change in response to the input data. If the model receives new audio and EEG data that suggest a shift in the listener's attention, the attention weights will adjust to reflect this new information.

**Mathematical Operations:**

* **Matrix Multiplication**: One of the fundamental operations in this process is matrix multiplication, where the feature matrices are multiplied by the weight matrices. This operation combines and transforms the data, factoring in the importance assigned to each feature.
* **Normalization**: After combining the features with their weights, the model might normalize the results to ensure they fall within a certain range. This step can involve adjusting the scale of the output values to maintain consistency across different data points.
* **Optimization**: The model uses optimization techniques to adjust the weight matrices so that the model's predictions closely match the known attentional focus in the training data. This might involve techniques like gradient descent, where the model iteratively adjusts the weights to minimize the difference (error) between its predictions and the actual outcomes.

**Outcome:**

* **Attention Prediction**: The outcome of all these matrix operations is a prediction of where the listener's attention is focused, based on the combined and weighted audio and EEG features. The model's success hinges on how well it can dynamically adjust the attention weights to accurately reflect the listener's focus in real-time.

The methods section of the article on the Cross-Modal Attention-based Auditory Attention Detection (CMAA) model can be summarized paragraph by paragraph as follows:

1. **EEG Data Preprocessing**:
   * EEG data from subjects is first cleaned and preprocessed. This involves filtering the raw data to remove noise and artifacts, ensuring that the subsequent analysis focuses on relevant brain signals related to auditory attention.
2. **Stimuli and Task Design**:
   * Details the auditory stimuli used in the experiments, such as the types of sounds and their presentation. It also outlines the tasks participants performed, typically involving listening to these sounds under various conditions to elicit attention-related brain responses.
3. **Feature Extraction**:
   * Explains how features are extracted from both the preprocessed EEG data and the auditory stimuli. Features might include aspects like power spectral densities from the EEG or spectral features from the audio, capturing characteristics essential for decoding attention.
4. **Cross-Modal Attention Mechanism**:
   * Introduces the core concept of the CMAA model, where the extracted features from EEG and audio data are integrated using a cross-modal attention mechanism. This mechanism dynamically adjusts to the input stimuli, modulating the features to highlight those most indicative of the subject's focus of attention.
5. **Model Training and Evaluation**:
   * Describes how the CMAA model is trained using the feature data and attention labels from the experiments. It includes details on the training process, such as the optimization algorithm used and how the model's performance is evaluated against a set of metrics.
6. **Implementation Details**:
   * Provides specifics on the technical implementation of the CMAA model, including the software and hardware used, and any particular settings or parameters critical to replicating the model's performance in other studies.
7. **Statistical Analysis**:
   * Discusses the statistical methods applied to analyze the results, detailing how the effectiveness of the CMAA model in decoding attention is quantitatively assessed, possibly including comparisons to other models or benchmarks.

**Methods:**

* Speaker A and Speaker B: These are the two sources of auditory stimuli, representing the two different speakers in the auditory attention task.
* H-LP (High-pass Linear Processing): This process is applied to the audio signals from both speakers to enhance the speech envelope. It helps to make the characteristics of the speech more prominent and easier for the model to analyze.
* EEG Signals: The blue lines represent the brainwave signals collected from a person while listening to the two speakers.
* CSP (Common Spatial Patterns): This method is applied to the EEG signals to extract features that can differentiate between the listener's attention to Speaker A vs. Speaker B. The CSP essentially acts as a filter to enhance the brainwave signals that are most relevant for determining attention.
* Cross-modal Transformers (E→A and E→B): These components take the EEG features as input and dynamically interact with the enhanced audio signals of Speaker A and Speaker B. They perform feature extraction and attentional modulation to find the best combination of features representing the attention towards each speaker.
* Similarity Matrix: After the cross-modal transformer processes the data, a similarity matrix is calculated for each speaker. This involves computing the correlation between the EEG signals and the audio features for Speaker A and Speaker B. The correlation is likely measured using cosine similarity, which quantifies how similar the EEG patterns are to the audio patterns for each speaker.
* FC (Fully Connected Layer): This is a type of neural network layer that takes the similarity scores and combines them, possibly adding additional complexity and decision-making ability to the model.
* Binary Decision: Finally, the model makes a decision about which speaker the person is attending to. This decision is binary, meaning it's one or the other—either Speaker A or Speaker B. The decision is made based on which similarity score is higher, indicating a stronger correlation between the EEG signals and the audio features of that speaker.

Fiebelkorn et al., 2019:

* **Environmental sampling** 🡪 brain directs its limited resources to first select and then boost the processing of behaviourally relevant stimuli
  + Selective attention 🡪 preferential sensory processing
  + Selective attention + exploratory movements 🡪 environmental sampling
  + Prioritization of stimuli
  + Enhancement of neural processing of relevant stimuli
  + Movements orient the sensory organs toward relevant stimuli
* **Rhythmic neural activity 🡪**  shapes both prioritization of stimuli and exploratory movement
  + theta rhythms 🡪 in the attention network
  + resolve potential functional conflicts 🡪 brain's ability to manage competing demands for attention and resources when processing sensory information
  + even during sustained attention 🡪 alternating periods of enhanced or diminished perceptual sensitivity
  + diminished sensitivity 🡪 opportunity to shift attention 🡪 **every 250ms**
* **Saccades and sensory processing**
  + Coupled to eye position during overt sampling
  + Uncoupled during covert
  + Sensory and motor systems co-evolved and are functionally integrated
  + Saccades determine the targets for preferential sensor sampling **and** a sensory analysis of stimulus properties determined the end points for saccades
* **Anatomical overlap**
  + FEF, superior colliculus, lateral intraparietal area (LIP, cortical), thalamus 🡪 at the nexus of sensory and motor functions, playing a role in both
  + Part of the attentional network
* **Premotor theory of attention** 🡪 covert spatial attention 🡪 weaker activation of the **same** neural population that typically guides saccades and primarily reflects saccadic preparation 🡪 in FEF and superior colliculus
  + Weakly stimulating a frontal region that contributes to the generation of saccades 🡪 behavioral and neural effects that mimicked covert spatial attention
  + **But**  frontal cortex has functionally dissociable cell types, present in other regions of the attention network
  + These cells might make it possible to separately engage sensory processing or exploratory eye movement, but the presence of some neurons with both sensory and motor responses provides a neural basis for integrating of sensory and motor processing
  + Manipulations in monkeys and humans proved that spatial attention and saccades have separate control mechanisms
* **Spatial attention**
  + Discontinuous
  + Fluctuates 🡪 3-8Hz cycles
  + *A spotlight* 🡪 scanning the visual field and pausing to illuminate potentially relevant stimuli
  + EEG study 🡪 brief light flashes 🡪 phase of fronto-central theta rhythms, just prior to stimulus onset, was predictive of successful visual-target detection
  + if a theta oscillation has a frequency of 6 Hz, one complete cycle would last approximately 166.7 milliseconds. At any given time, a phase of 0 degrees would correspond to the beginning of the cycle, while a phase of 180 degrees would occur halfway through the cycle.
  + MEG study 🡪 alternating, equiprobable targets 🡪 observed increases in gamma-band activity (30–90 Hz) that were synchronized with the phase of theta rhythm (30-90Hz, related to better visual-target detection) 🡪 attention-related boosts in visual processing alternated between visual hemifields
* **The experiment: ECoG**
  + when theta rhythms were in certain phases, people tended to perform better on the tasks. This relationship was strongest in areas of the brain involved in attention, such as the frontal and parietal cortices
  + Interestingly, they also found that the phase of theta rhythms was related to how excitable the brain was in these areas. When theta rhythms were in certain phases, there was more activity in these brain regions, which is a sign of increased excitability
  + Theta rhythmic-coupling is not context-specific🡪 a fundamental property of spatial attention, operating regardless of the specific task requirements
  + Coupling of theta phase and higher-frequency bands linked with frontal and parietal cortices 🡪 the **attention network**
* **LFP and single-unit-activity SUA**
  + **Macaques**
  + Targeted FEF and LIP 🡪 theta rhythms in both were predictive of behavioral performance
  + Theta-rhythmic sampling evolutionarily preserved across at least 2 primate species 🡪 a fundamental property of spatial attention
* **Phase-amplitude coupling:**
  + Modulation of high-frequency band power by the phase of a lower frequency band 🡪 **mechanism for temporally coordinating cognitive functions**
* **The mechanisms: proposal**
  + Gamma band increase 🡪 in LIP 🡪 associated with visual-sensory neurons
  + Beta band increase 🡪 FEF 🡪 associated with visual-movement neurons
  + ***First attentional state*** *🡪 enhancement of sensory processing + suppression of attentional shifts*
  + ***Second attentional state 🡪*** *diminished perceptual sensitivity 🡪 LIP-specific alphaband activity 🡪 alpha band linked to attenuation of sensory processing + with a temporary release from beta-related motor suppression 🡪 increase the likelihood of attentional shifts*
* **Reweighing** 🡪 dynamic interactions between cortical and subcortical hubs of the attentional network
  + Regulated connectivity with higher-order cortex 🡪 FEF and LIP during the enhanced sensory state
  + LIP regulates subcortical hubs during the diminished sensory state
  + **Reweighing of functional connections between regions of the attentional network 🡪 some promoting sampling, others shifting**
* **Microsaccades**
  + The "rhythmic theory of attention" predicts more exploratory eye movements during specific phases of theta rhythms.
  + Microsaccades occur more frequently during these theta-dependent periods of reduced perceptual sensitivity, whether directed away from or toward the cued location.
  + Microsaccades away from the cued location suggest attentional shift, while those toward it indicate re-selection of the cued location for further sampling.
  + Theta rhythms are linked not only to microsaccades but also to larger saccadic eye movements used in overt exploration.
  + Studies show that sensory and motor aspects of environmental sampling are synchronized with theta rhythms.
  + Saccades tend to happen during periods of diminished perceptual sensitivity, suggesting alternating cycles of attention engagement and disengagement.
  + Saccades occur during times when perception is already less sharp, without altering the timing of perception rhythms.
* **Cognitive flexibility**
  + Theta-rhythmic sampling resolved functional conflicts by temporally isolating sensory and motor processes during environmental sampling
  + Sensory (sampling) and motor (shifting) linked to opposite phases of the same theta rhythm 🡪 organizes activity into two alternating states
  + 1st state 🡪 sensory sampling at a behaviorally relevant location + supressing attentional shifts 🡪 increased gamma synchronization between attentional hubs and visual cortices
  + 2nd state 🡪 attenuation of sensory processing at presently attended location 🡪 shift not necessary, but attentional network is primed should such shift be warranted
  + Depends on stimulus properties and behavioral goals 🡪 **a priority map**
  + They predicted a decrease in perceptual sensitivity that co-occurs with an increase in functional connectivity between the attention network and regions typically associated with the initiation of covert or overt attentional shifts 🡪 i.e. superior colliculus

Menon & Udin, 2010: Insula

* The insula, especially the anterior part (the hidden cortex region)
* Part of the salience network
* Assists target brain regions in the generation of appropriate behavioral responses to salient stimuli
* **Mechanisms (4):**
  + Detection of salient events
  + helps the brain switch its focus and memory resources to pay attention to and process information related to significant or salient events
  + interaction of anterior and posterior insula regions to modulate the motor response/reaction to salient stimulus
  + coupling with the anterior cingulate cortex 🡪 facilitates rapid access to the motor system
* point of the article, is to provide us with a simple model to understand the different insula functions
* **with its core function being marking salient stimuli in time and space, for additional processing**
* the insula is situated at the interface of cognitive, homeostatic and affective systems of the human brain
* key to understanding the functions of any specific brain region lies in understanding how its connectivity differs from the pattern of connections of other functionally related brain areas
* Critchley et al. (2001) found that patients with pure autonomic failure (an idiopathic disorder in which peripheral denervation disrupts autonomic responses) show **reduced activation in the right insula** 🡪 strong link between perception of one’s bodily state and experience of emotions
* A role in representation of bodily urges 🡪 in smokers with damaged insula reduced
* AI active ini tasks involving negative and positive feelings
* Involved in **deception** 🡪 high-level cognitive process 🡪 breach of a promise can be predicted by brain activity patterns in AI and ACC + frontal gyrus
* Empathy
* **Structural connectivity of the insula:**
  + while the AI is a part of both networks, its roles within each network may differ. In the SN, it helps prioritize salient stimuli, while in the CEN, it may contribute to cognitive control processes
  + the AI and ACC, within the SN, initiate control signals that regulate the activation of the CEN and deactivation of the DMN during cognitive processing
  + coordinating the activity of DMN and CEN
* Combined EEG and fMRI data analysis reveals insights into the role of the salience network (SN) in attentional control (Crottaz-Herbette and Menon 2006)
  + A schematic model suggests four stages of spatio-temporal dynamics in attentional control:
    - Stage 1: Primary sensory areas detect a deviant stimulus (MMN).
    - Stage 2: The MMN signal is transmitted to the AI and ACC.
    - Stage 3: The AI and ACC generate a top-down control signal (N2b/P3a) transmitted to other brain regions.
    - Stage 4: Neocortical regions respond to the attentional shift (P3b-evoked potential).
    - Stage 5: The ACC facilitates response selection and motor response.
  + Within this framework, the AI plays a prominent role in detecting salient stimuli, while the ACC modulates responses in sensory, motor, and association cortices
* When the anterior cingulate cortex (ACC) is dysfunctional, the anterior insula (AI) may compensate by triggering alternate cognitive control signals through other lateral cortical regions such as the ventrolateral prefrontal cortex (VLPFC) and dorsolateral prefrontal cortex (DLPFC)

Zatorre, 1999:

* PET scan used
* Hypothesis: similar neural systems for attending to spectral and spatial features of sounds
* Tones varying randomly in frequency and location 🡪 response to either low or high-frequency, or to left or right 🡪 ignoring the other stimulus that is not targeted
* IF modulation occurs depending on the features attended 🡪 changes in cerebral activity should be seen
* Test if right hemispheric cortical regions will be also relevant
* CBF increases in network in right hemisphere cortical regions 🡪 for both conditions
* **Auditory attention operates at a level at which features have been integrated to a unitary representation**
* Selective attention guided by both location and frequency cues (Dai, 1991)
* Auditory objects, **not individual features**, are subject to selective attention
  + Attending to stimulus features which are integrated 🡪 similar patterns of cerebral activity 🡪 features analyzed preattentively
  + Features may still be independently registered at an early processing stage
  + Woods, 1994: different ERPs distributions for attending one feature vs the other
    - **But** features appear to integrate within 130ms
  + **Enhancement of ERP amplitude of attended stimulus** 🡪 Hillyard, 1993
* sometimes stimuli processing does not need extra effort 🡪 information extracted fully, no further modulation of sensory regions needed
* **previous studies** provide evidence for right parietal and frontal regions involved in selective attention
* **Methods:**
  + They set a high t-value threshold (3.5) to be very cautious and avoid mistaking random fluctuations in blood flow for real activation.
  + This means they were very cautious and only considered brain regions with a substantial and consistent increase in blood flow (low variability) during listening tasks as truly active
  + This ensured that only brain regions with substantial increases in blood flow compared to resting were considered active during the listening tasks.
  + **Higher t-value means stronger evidence:** A t-value greater than a certain threshold (like 3.5 in this study) indicates a stronger likelihood that the observed difference in blood flow between the listening task and resting baseline is not due to chance
  + **If average CBF during the task is slightly higher than rest, but data is more scattered in each condition** 🡪small difference in averages might be misleading. The high variability could be due to factors unrelated to the listening task, like individual differences in brain anatomy or how relaxed someone was during the scan
* **Results**
  + They found very similar brain activation patterns across all listening tasks.
  + These included areas expected for hearing (auditory cortex), movement planning (motor areas), and balance (cerebellum).
  + Interestingly, they also found increased activity in specific regions of the right hemisphere (superior parietal, premotor, and mid-dorsolateral frontal) during all listening tasks.
  + These right-hemisphere regions seem to be consistently involved in all listening tasks, regardless of whether you focus on sound location or pitch.
  + There was one region in the right inferior frontal cortex that seemed to be more active only during the "attend left" task, suggesting it might be involved more specifically in tasks requiring leftward attention.
  + No significant modulation of CBF within the auditory cortices, as a function of attentional manipulation
  + The study found that specific brain regions, particularly in the right hemisphere, are generally activated during listening tasks, regardless of whether you focus on sound location or pitch.
  + There might be some minor differences in activity depending on the specific task (direction of attention or sound pitch), but these seem less prominent.
* **Discussion**
* **Right Hemisphere Network for Attention:**
  + The study supports the idea of a specialized network in the right hemisphere of the brain that's activated during auditory attention tasks.
  + This network includes regions in the parietal, frontal, and temporal lobes.
  + These regions likely work together to focus attention on specific sounds (spatial location or pitch) based on integrated sensory information.
* **Other Involved Brain Areas:**
  + Areas related to motor response (movement planning), general auditory processing, and a part of the left thalamus were also activated during the tasks.
  + The thalamus activity might be related to a general alertness mechanism needed for any task, not specific to auditory attention.
* **Similar Brain Activity for Different Attention Tasks:**
  + Regardless of whether participants focused on sound location (left/right) or pitch (high/low), the brain activity patterns were very similar.
  + This suggests the network works similarly for different auditory attention tasks.
* **Possible Role of Frontal Areas:**
  + There might be some subtle differences in activity in the lower frontal cortex depending on whether attention is directed to the left or right side.
  + However, overall, there wasn't strong evidence that the network activity is heavily influenced by the specific side of attention (left vs right).
* **Comparison to Other Studies:**
  + This study aligns with previous research suggesting the right hemisphere plays a key role in attention across different senses (sight, sound, touch).
  + The right superior parietal cortex likely receives input from various sensory areas to create a combined representation for attention selection.

It is entirely possible, of course, that differential neuronal events do take place within the auditory cortices as a function of attention, but that these occur on a time scale to which PET is insensitive.

Others reported increases in primary AC, associated with attentional manipulations (Alho, 1999) 🡪 maybe due to nature of attentional processes elicited by the task itself, or rate of stimulus 🡪 with slow rates, automatic and complete feature processing, fast rates 🡪 early selection reflected by increased CBD in AC

Faster presentation rates forcing the brain to prioritize selecting the relevant sound feature first, leading to more activity in the auditory cortex.

Attending to competing sounds presented simultaneously (dichotic listening) might require more active processing in the auditory cortex compared to listening to sounds presented one at a time.

Attentional selection likely happens on these already extracted auditory "streams" based on factors like location or pitch.

Choi, 2014:

**Methods:**

* **Participants:** 18 volunteers with normal hearing participated.
* **Stimuli:** Three melodies (left, right, center) were played simultaneously on each trial. The center melody was always ignored. The other two melodies differed in pitch range (same vs different) and had specific note timings to allow researchers to distinguish brain activity for each.
* **Task:** Participants focused on the left or right melody based on an audio cue (attend left/right) and identified the contour (ascending, descending, zigzagging) of that melody after the sounds stopped. In some trials (passive), participants were instructed to ignore all melodies.
* **EEG Recording:** Brain activity was measured using EEG during the main experiment. Activity was recorded from 32 scalp locations at a high sampling rate.
* **Data Analysis:**
  + Only correctly answered trials from the "different pitch" condition were analyzed for EEG (not enough correct responses for "same pitch").
  + EEG recordings were segmented into epochs around the time of melody presentation and baseline-corrected.
  + Trials with very large voltage spikes were excluded.
  + Researchers focused on the N1 component of the ERP (electrical response in the brain to sound) as a marker of attention.
  + N1 amplitude (strength) in frontal-central areas (associated with auditory processing) was calculated for each melody (attended vs ignored) and attention condition (attend leading vs lagging melody).
  + N1 amplitude was normalized for each participant by their strongest N1 response (likely to the center melody).
  + Amplification/suppression effects were calculated to quantify how attention influenced the N1 response for the attended vs ignored melodies.

**Main Findings:**

* Participants performed better identifying melodies when the melodies had different pitches compared to when they had the same pitch.
* Focusing attention on a specific melody (attend condition) amplified the brain's response (N1) to that melody compared to passively listening or focusing on the other melody (lag effect). This amplification effect was stronger for the different-pitch condition.
* There were large individual differences in attention ability, with some participants showing much stronger amplification than others.
* The strength of the amplification effect (lag amplification) correlated with performance, especially for the difficult same-pitch condition. This suggests that people who showed a stronger neural amplification when focusing on a melody also performed better at identifying that melody.
* Other factors besides neural amplification, like how often participants accidentally responded during passive trials (false alarm rate), also influenced performance. When considering both neural amplification and false alarm rate, the researchers got a better prediction of performance, especially for the difficult same-pitch condition.

**Additional Notes:**

* The researchers controlled for learning and fatigue effects by analyzing performance across all trials.
* Attention did not significantly affect the brain's response to the initial part of the leading melody, regardless of whether participants were focusing on it or not.

**Bialas, 2023:**

* Encoding of elevation through monoaural spectral cues remains unclear
* Previous studies linear representation of elevation in AC
  + Changes in sound elevation lead to linear adjustment in the level of activity within certain cortical areas
  + As the sound decreased, rate of neural firing increases
  + The rate is a cue that the brain uses to determine elevation
  + Proven in fMRI studies 🡪 greater blood-oxygen level-dependent responses
* Methods:
  + Seated on a chair in an anechoic chamber
  + Ear molds to manipulate adaptation of neural responses to the varying elevations of sounds
  + Analyzed the ERPs elicited by sounds at different elevations
* Results:
  + EEG captured cortical responses related to changes in elevation
  + Responses increased with decreased sound elevation
  + Significant clusters of elevation-specific responses were revealed
  + Population-rate code hypothesis 🡪 information about sound features is encoded in the overall firing rates of a group of neurons within a population, rather than the specific timing or pattern of firing of individual neurons
    - Collective activity across a population of AC neurons can represent more complex auditory spatial information
  + Differences in elevation of sound could be decoded from EEG data 🡪 predictive of subject’s ability to localize sound elevation
* Discussion:
  + Population-rate code in the AC
  + Elevation represented at around 400ms in the cortex, after sound onset
  + This encoding allows for a detailed representation of elevation
  + Robust representation mechanism
  + Sounds from lower elevations elicited stronger and more distinct responses
  + Decoding techniques applied to EEG data demonstrated that elevation of a sound source could be predicted from the patterns of brain activity
  + By integrating signals across a population of neurons, the brain can derive meaningful spatial information from relatively simple changes in firing rates
  + This mechanism may be more robust to noise and variability in neural responses 🡪 relies on average output of many neurons rather than on the precise timing of spikes from individual neurons

Codes for Sound-Source Location in Nontonotopic Auditory Cortex by Middlebrooks et al. (1998):

1. **Research Objective**: The study aims to determine how the auditory cortex encodes the location of sound sources, focusing on two hypotheses: topographical coding and distributed coding.
2. **Experiment Setup**: Researchers used adult cats as subjects, conducting experiments in anechoic chambers with controlled noise bursts directed from various azimuths. They specifically examined neural responses in two cortical areas, the anterior ectosylvian sulcus (AES) and area A2.
3. **Methodology**: The research involved varying sound pressure levels to investigate their effects on the spatial tuning of neurons. Neural responses were recorded through electrodes, and the data were analyzed using artificial neural networks to infer the locations of sound sources based on the patterns of neural spikes.
4. **Key Findings**:
   * Neurons in the AES and A2 areas displayed broad spatial tuning, responding to sound sources across large azimuth ranges, often over 180°.
   * Such broad responses did not support the topographical code hypothesis, which posits a mapped spatial representation in the auditory cortex.
   * Instead, the findings were more consistent with a distributed code hypothesis, where sound locations are encoded by a collective response of neurons across a wide range of azimuths.
5. **Conclusions**:
   * The study suggests that sound localization in the auditory cortex, particularly in non-tonotopic areas, involves more complex and distributed neural interactions rather than a straightforward mapped representation.
   * This implies a panoramic approach to sound localization where multiple neurons collectively contribute to determining the position of a sound source.

Middlebrooks, 1984:

1. **Study Objective**: The study aimed to investigate how the cat's superior colliculus encodes auditory space, with an emphasis on understanding the spatial representation of sound sources.
2. **Key Findings**:
   * Most auditory units in the superior colliculus have sharply defined receptive fields, responding maximally to sounds from specific horizontal and vertical locations within these fields.
   * The receptive fields form two distinct classes based on their location and size, with a third class of units responding to sounds from any location.
   * Systematic shifts in the locations of these "best areas" (areas within the receptive fields where responses are maximal) are observed as a function of unit position, creating a continuous and structured map of auditory space.
     1. where auditory stimuli elicit the strongest or maximal response from the neuron
     2. as you examine different locations within the superior colliculus, the "best area" for sound localization systematically moves across the auditory space. This arrangement allows for a structured map of auditory space, where different locations in the superior colliculus correspond to different spatial locations of sound sources
3. **Methodological Approach**:
   * The researchers used free field auditory stimuli to examine the responses of single units and clusters of units in the intermediate and deep layers of the superior colliculus of 29 cats.
   * Detailed attention was given to maintaining consistent experimental conditions to reliably map the spatial selectivity of the auditory responses.
4. **Significance of Findings**:
   * The results demonstrate a systematic mapping of auditory space in the superior colliculus that is analogous to the mapping of visual space, indicating a parallel processing of sensory information.
   * This mapping challenges the notion of a simple topographical representation of sound space, suggesting instead a more complex, distributed coding mechanism.
5. **Implications for Understanding Auditory Processing**:
   * The study highlights the complexity of auditory spatial processing in the superior colliculus, an area traditionally associated more with visual and somatosensory processing.
   * It suggests that the auditory system uses a combination of spatial cues derived from acoustic signals to form a detailed and dynamic representation of the environment, which is crucial for orienting responses to auditory stimuli.
6. **First Class - Sharply Delimited Fields**:
   * These receptive fields are highly specific and tightly confined to particular areas. Units with these fields respond maximally to sounds originating from precise locations, making them crucial for fine-grained localization of sound sources.
7. **Second Class - Broad or Panoramic Fields**:
   * These fields are larger and less precisely defined compared to the first class. Units with these fields respond to sounds from a broader range of locations. This characteristic might be advantageous for detecting sounds that do not come from specific, pinpointed locations but rather from broader areas.
8. **Third Class - Omnidirectional Units**:
   * The units in this class respond to sounds from any location, showing no spatial selectivity in their auditory responses. This class forms a contrast with the other two, as these units can be activated by sounds coming from any direction. This feature is essential for overall auditory awareness and readiness to respond to sound stimuli from any possible direction.

**Distributed Coding**: Instead of each point corresponding to a specific location in the brain, distributed coding uses patterns of neural activity across many neurons to represent information. In the context of auditory spatial localization in the superior colliculus, the study suggests that sound locations are not represented by distinct, localized points in the brain. Instead, sound locations are encoded by the pattern of activity across a group of neurons, with each neuron potentially responding to multiple spatial cues but in varying degrees.

**Evidence for Distributed Coding from the Study**

* **Broad Tuning of Receptive Fields**: Many neurons in the superior colliculus respond to sounds from a broad range of locations, which is indicative of a distributed coding mechanism. This broad tuning means that rather than having a precise point-to-point correspondence, the superior colliculus encodes sound location through a more flexible and integrative approach.
* **Systematic Shifts and Overlaps**: The systematic shifts in the "best areas" and the overlaps in the receptive fields across neurons suggest that the encoding of space is more about the interplay and integration of responses from multiple neurons rather than from a single neuron responding solely to a specific location. This approach allows the system to be more resilient to noise and more adaptive to varying acoustic environments.
* **Parallel to Visual Mapping**: While the auditory mapping mirrors some aspects of visual space mapping in terms of having structured organization (systematic shifts), it does not conform to a strict topographical order as seen in purely visual processes. This suggests a model where auditory spatial information is processed in a complementary but more distributed fashion relative to the more topographically organized visual inputs.

**Integration in the Midbrain**

* **Topographic Representation**: In the midbrain, such as in the superior colliculus and inferior colliculus, neurons create a topographic map of auditory space. This means that neurons responding to sounds from specific, narrowly defined regions of space are organized in a way that physically corresponds to the spatial map. Studies like those by Middlebrooks and Knudsen (1984) and others have shown that these neurons are organized topographically, each tuned to different spatial cues, facilitating precise localization of sound sources based on where neurons fire in this map.

**Cortical Processing**

* **Broad Tuning in Cortex**: Unlike the midbrain, cortical neurons (those in the brain's cortex) typically do not respond selectively to a narrow location. Instead, their response covers a broad area of space, especially within the auditory field opposite to their hemisphere (contralateral hemifield). This implies that while midbrain neurons pinpoint sound location, cortical neurons adjust their responses over larger spatial areas, suggesting a more integrated and less precise role in sound localization at this level of processing.

**Opponent-Channel Code**

* **Two Neural Populations**: The paragraph mentions the opponent-channel code model, where two populations of neurons are each broadly tuned to one side of the auditory space (left or right, relative to the body's midline). This model suggests that sound location is encoded by the relative activity levels of these two populations.
* **Greatest Change at Midline**: According to this model, the perceptual resolution—how distinctly the brain can determine sound location—is highest around the midline. This is where the tuning curves of the two neural populations intersect, leading to the greatest change in neural response when the sound source moves around this area.

**Discussion**

* **Evoked Responses Encode Sound Elevation**: The experiments demonstrated that EEG can be used to assess the cortical processing of sound elevation. This finding is consistent with previous fMRI studies that identified specific regions in the auditory cortex tuned to sound elevation. The adaptation design of the study helped separate neural activity related to sound onset from that related to elevation, revealing distinct cortical responses that could be accurately decoded and were predictive of individual localization performance .
* **A Monotonic Population-Rate Code for Elevation**: The study supports the hypothesis that the auditory cortex represents sound elevation in a population response that decreases monotonically with increasing elevation. The experiments showed that ERP amplitude increased with the separation of adapter and probe, especially when the adapter was at the highest elevation, which aligns with the expected pattern under a population-rate code hypothesis .
* **Latency of Cortical Elevation Processing**: The latency for the largest elevation-specific response was found to be between 200ms and 250ms after probe onset, which is similar to previous studies using MEG and EEG. The study's second experiment showed a prolonged elevation-specific response, lasting between 200ms and 800ms, suggesting a relationship with the experimental task's demand to maintain the perception of the probe for performance optimization .
* **The Cortical Representation of Sound Direction**: The study discusses how the auditory cortex might represent sound azimuth as the difference between the activity rates in two opponent neural channels, each tuned to the contralateral hemifield. It is suggested that sounds at low elevations cause larger responses than sounds at high elevations, and this phenomenon is consistent with a rate code for sound localization .