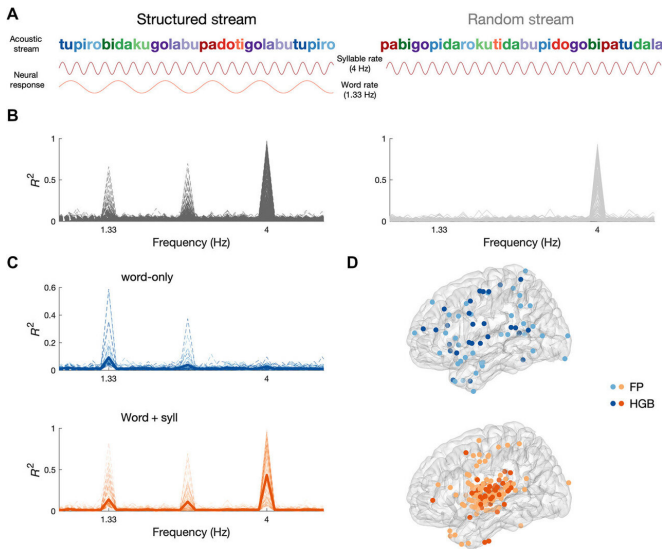


# Learning hierarchical sequence representations across human cortex and hippocampus

5 March 2021, by Thamarasee Jeewandara



Neural tracking of auditory SL. (A) Schematic depiction of the auditory SL task. The structured stream (left) contained 12 syllables [250-ms stimuli onset asynchrony (SOA), 4 Hz] in which the TPs formed four words (color-coded for visualization, 750-ms SOA, 1.33 Hz). The random stream (right) contained the same 12 syllables in a random order. The predicted neural response is shown below each syllable stream: Syllable tracking (top) was expected in both conditions, whereas word tracking (bottom) was expected only in the structured condition. (B) Phase coherence spectrum in neural data for the structured (left, black) and random (right, gray) conditions from 1898 electrodes in 17 patients. Each significant electrode is depicted with a thin line, and the average is depicted with a thick line. (C) Phase coherence spectrum in the structured condition for electrodes showing word-tracking responses, in two groups: electrodes that showed tracking responses at the word rate only (top, blue) and electrodes that showed tracking responses at both the word and syllable rate (bottom, orange). (D) Localization of word-only (top, blue) and word + syll (bottom, orange) electrodes exhibiting significant phase coherence in the field potential (FP; light blue, light orange) or the high-gamma band (HGB; dark blue, dark orange). Credit: Science Advances, doi: 10.1126/sciadv.abc4530

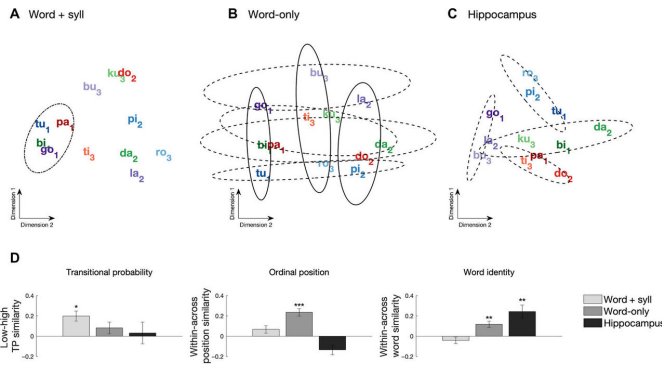
Humans experience sensory input continuously as segmented units of words and events. The ability of the brain to discover regularities is known as [statistical learning](#). This concept can be represented at multiple levels including transitional probabilities and the identity of units. In a new report now published on *Science Advances*, Simon Henin and a team of scientists at the New York University School of Medicine, Yale University and the Max Planck Institute in the U.S. and Germany recorded sequence encoding in the cortex and hippocampus of human subjects exposed to auditory and visual sequences with temporal (time-based) regularities. Using early processing, they tracked lower-level features such as syllables and learned units including words, while later processing could only track learning units. The findings showed the existence of multiple parallel computational systems in humans to assist learning across organized cortico-hippocampal units.

## Understanding the code of speech

We receive and experience continuous input from the world in digestible chunks. For example, with language, humans can acquire and extract meaningful sequences including sentences, words and phrases from [a continuous stream of sounds](#) without clear acoustic boundaries or pauses between linguistic elements. This segmentation happens incidentally and effortlessly as a core building block during development. The behavior of learning transitional probabilities between syllables or shapes in infants or adults are known as "[statistical learning](#)". However, the mechanism of the brain supporting such cognitive functions are poorly understood. It is well known for brain regions such as the hippocampus and the [inferior frontal gyrus](#) (IFG) to aid in visual and auditory statistical learning. To understand this process, Henin et al. conducted intracranial recordings from 23 human

epilepsy patients to provide mechanistic insight into the fundamental process of human learning relative to cortical areas that respond to the structure of the world. The findings highlighted [neural frequency tagging](#) (NFT) as a versatile tool to investigate incidental learning in preverbal and nonverbal patient populations.

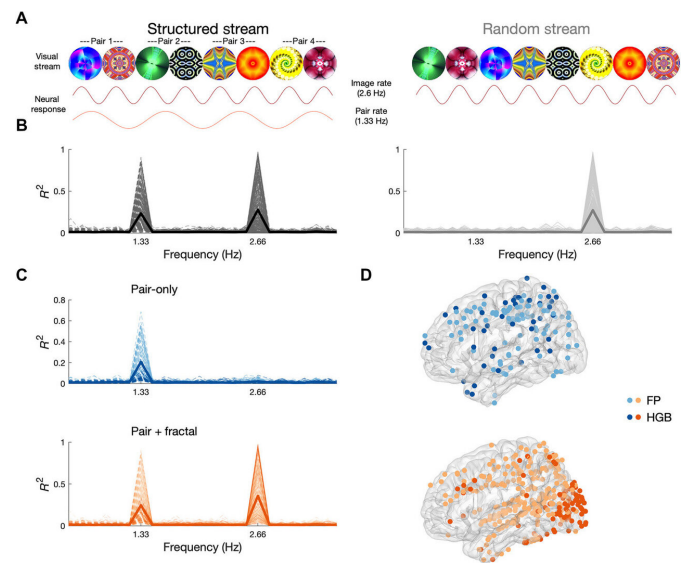
create two sequences: a structured and random word sequence. In the structured sequence, Henin et al. manipulated the transitional probabilities between syllables so that four hidden words could be embedded in sequence to create a continuous artificial language stream. They represented the underlying syllable presentation rate at 4 Hz and the word rate at 1.33 Hz. The team did not inform the participants of the structure but asked them to perform a cover task instead, where they indicated syllable repetitions randomly embedded in the auditory streams.



Pattern similarity results during auditory SL. Multidimensional scaling (MDS) of the distances between syllabic responses across electrodes showing significant (A) word + syll responses and (B) word-only responses, as well as (C) across electrodes from the hippocampus. Individual words are color-coded; subscripts represent ordinal position (e.g., “tu1pi2ro3”). Dot-dashed ellipses indicate grouping by TP, solid ellipses outline grouping by ordinal position, and dashed ellipses indicate grouping at the level of the individual words (color-coded). (D) Quantification of multivariate similarity for syllables in the auditory SL task. Left: Similarity by TP. Greater within-class similarity indicates stronger grouping of syllables with low TP (0.33) than syllables with high TP (1.0). A Friedman test indicated a main effect of electrode type on TP similarity ( $\chi^2 = 22.03$ ,  $P$

**Behavioral evidence of auditory statistical learning**

Henin et al. studied the [neural circuits](#) and computation underlying statistical learning by presenting 17 participants with auditory streams of syllables after manipulating the structure of the sequence. The team placed each syllable into the first, second and third position of a three-syllable word or a triplet in such structured streams. The resulting transitional probabilities were low and uniform without a word level of segmentation. During the auditory tasks, they generated 12 consonant-vowel syllables using [MacTalk](#) and concatenated them using [MATLAB software](#) to



Neural tracking of visual SL. (A) Schematic depiction of the visual SL task. The structured stream (left) consisted of a continuous visual stream of eight fractals (375-ms SOA, 2.66 Hz). The TPs were adjusted to form four fractal pairs (750-ms SOA, 1.33 Hz). Note that the SOA of the fractals was elongated compared to the syllables to match the frequency of the learned units (pairs and words), given that there were two fractals per unit and three syllables. The random stream (right) contained the same fractals but in random order. The predicted neural responses are shown under each stream: Fractal tracking is expected for both streams, while pair tracking is expected for the structured stream only. (B) Phase coherence spectrum in neural data for the structured (left, black) and random (right, gray) conditions from 1606 electrodes in 12 patients. Each significant electrode is depicted with a thin line, and the average across the population is depicted with a thick line. (C) Phase coherence spectrum in the structured condition for electrodes showing pair-tracking responses, in two sets: electrodes that tracked pairs only (left, blue) and

electrodes that tracked pairs and fractals (right, orange). (D) Localization of pair-only (top, blue) and pair + fractal (bottom, orange) electrodes exhibiting significant phase coherence in the FP (light blue, light orange) or HGB (dark blue, dark orange). Credit: Science Advances, doi: 10.1126/sciadv.abc4530

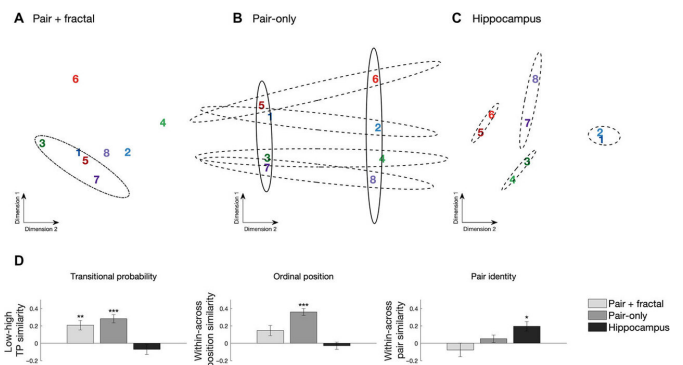
### Neural tracking of auditory statistical learning

Henin et al. obtained direct neurophysiological signals from 1898 intracranial electrodes in 17 participants to comprehensively cover the frontal parietal, occipital and temporal lobes as well as the hippocampus in both hemispheres. The participants performed a [two-alternative forced choice](#) (2AFC) task where they listened to the two audio segments presented one after the other to select the stream containing one of the hidden words. The scientists noted the responses to originate predominantly in [somatosensory/motor](#) and [temporal cortices](#). On average, they noted significantly increased word-rate coherence in the structured stream but not in the random stream, to support the sensitive and robust applications of NFT (neural frequency tagging) to assess online statistical learning. Using NFT, they tracked the representation of segmented units at [two hierarchical levels of the stream](#) and then tested the within-electrode phase coherence in the field potential [and gamma band](#) in the respective structured and random streams. Using electrocorticography, they showed the location of both words and syllable coherence to have occurred mainly in the [superior temporal gyrus](#) (STG) with smaller clusters in [the motor cortex](#) and [pars opercularis](#). In parallel, the other tuning profile reflected electrodes with significant coherence exclusively at the word rate only with locations in the inferior frontal gyrus and the [anterior temporal lobe](#) (ATL). The anatomical grouping highlighted the neuroanatomy of the auditory processing hierarchy.

### Analyzing auditory statistical learning and testing visual statistical learning.

To understand the results of neural frequency tagging (NFT), Henin et al. examined the segmentation driving the outcome, and based this on three statistical cues in the stream; including (1)

transitional probabilities, (2) ordinal position or (3) word identity to facilitate unique cognitive functions. As with auditory statistical learning tasks, the team performed visual statistical learning tasks with the patient groups, where the team formed fractals using similar sets of images as those used in [previous work](#). As before, the participants were not informed of the structure, but they performed a cover task. Henin et al. then used NFT to identify the brain areas exhibiting statistical learning in neurophysiological recordings from 1606 intracranial electrodes in 12 patients to cover the frontal, parietal temporal and occipital cortex. As with auditory statistical learning, they observed anatomical and hierarchical segregation between two temporal tuning profiles of electrodes, where one showed significant entrainment at the fractal and pair rates—mostly clustered in the [occipital and parietal cortex](#), while the other showed significant entrainment of pair only rates, in the frontal, parietal and temporal cortex.



Pattern similarity results during visual SL. MDS of the distances between responses to individual fractals across (A) pair-only, (B) pair + fractal, and (C) hippocampal electrodes. Pairs are color-coded; odd numbers refer to the first position, and even numbers refer to the second position. Dot-dashed ellipses outline grouping by TP/ordinal position in pair + fractal electrodes. Solid ellipses outline grouping by TP/ordinal position in pair-only electrodes. Dashed ellipses indicate grouping by pair in pair-only and hippocampal electrodes. (D) Comparison of multivariate pattern similarity for fractals in the visual SL task. Left: Within versus between similarity for low versus high TP. Greater within-class similarity indicates stronger grouping of fractals with a low TP (0.33) over fractals with a high TP (1.0). A Friedman test indicated a main effect of electrode type on TP similarity

( $\eta^2 = 19.3$ ,  $P$

**Outlook**

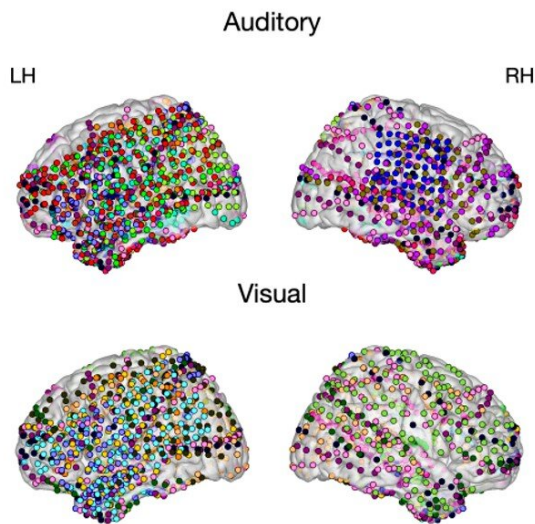
In this way, Simon Henin and colleagues used intracranial recordings in humans to describe how the brain tracks and learns structure within sensory information. The statistical learning process accompanied rapid changes in neural representations reflected in two functionally and anatomically distinct brain responses. These distinct responses revealed an anatomical hierarchy, which they mapped into early sensory processing stages in the superior temporal gyrus and occipital cortex. The team also mapped late, [amodal processing](#) stages in the [inferior frontal gyrus](#) and anterior temporal lobe. The patients extracted and represented nested structures within sensory streams in the brain [in as little as two minutes](#), even when they were not aware of the process.

time. The neural frequency tagging (NFT) technique provided an exciting opportunity to characterize learning trajectories across clinical and healthy populations across sensory modalities, to track the acquisition of knowledge across the lifespan from newborns to the elderly. By combining NFT with representational similarity analysis (RSA), the team provided a powerful toolkit to reveal how the brain engaged in statistical learning across multiple levels of organization within the human brain.

**More information:** Henin S. et al. Learning hierarchical sequence representations across human cortex and hippocampus, *Science Advances*, 10.1126/sciadv.abc4530

Kuhl P. K. et al. Early language acquisition: Cracking the speech code, *Nature Reviews Neuroscience*, [doi.org/10.1038/nrn1533](https://doi.org/10.1038/nrn1533)

Saffran J. R. et al. Statistical learning by 8-month-old infants. *Science*, 10.1126/science.274.5294.1926



© 2021 Science X Network

ECoG electrode coverage for participants in the auditory SL task (Top) and visual SL task (Bottom) from the left hemisphere (LH) and right hemisphere (RH). Each participant’s electrode coverage is shown with a different color. Credit: Science Advances, doi: 10.1126/sciadv.abc4530

The work [agreed with previous studies](#) to demonstrate how the cortical hierarchy integrated information across seemingly longer windows of

APA citation: Learning hierarchical sequence representations across human cortex and hippocampus (2021, March 5) retrieved 16 July 2022 from <https://medicalxpress.com/news/2021-03-hierarchical-sequence-representations-human-cortex.html>

*This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.*