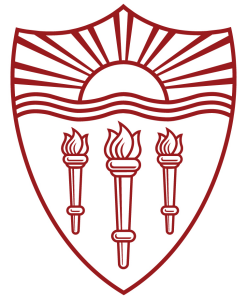
Affordance of Auditory Object Formation Improves Speech Perception in Noise

MAURY LANDER-PORTNOY & JASON ZEVIN

UNIVERSITY OF SOUTHERN CALIFORNIA



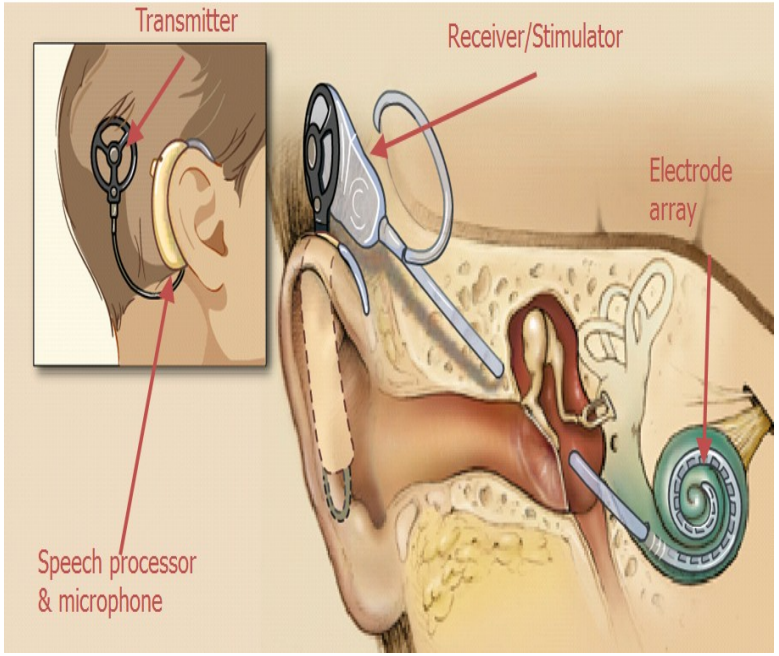
Who doesn’t love a good cocktail party?

Complex auditory environments

Lots of interfering noise

“Cocktail Party Problem” (Cherry 1957)

Who is most affected?



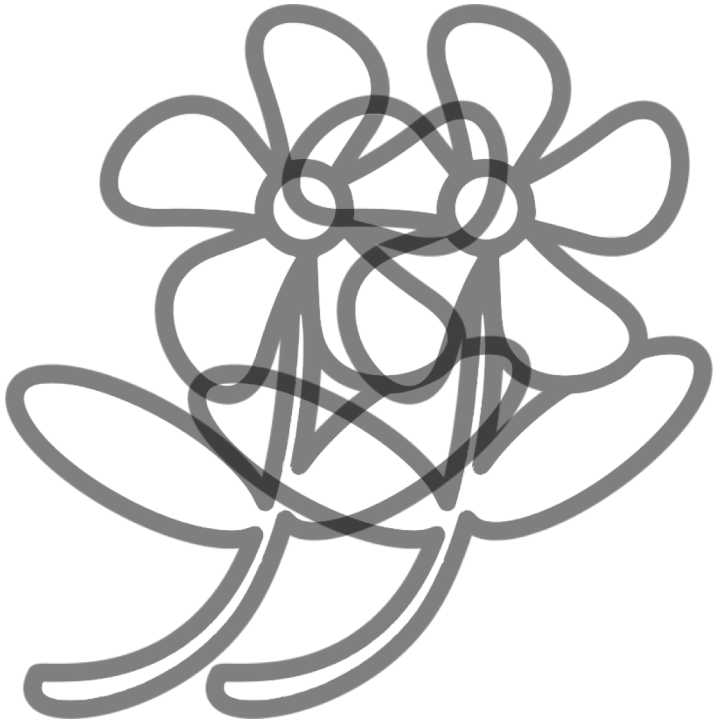
300K+ CI Users

(NIH Publication No. 00-4798)

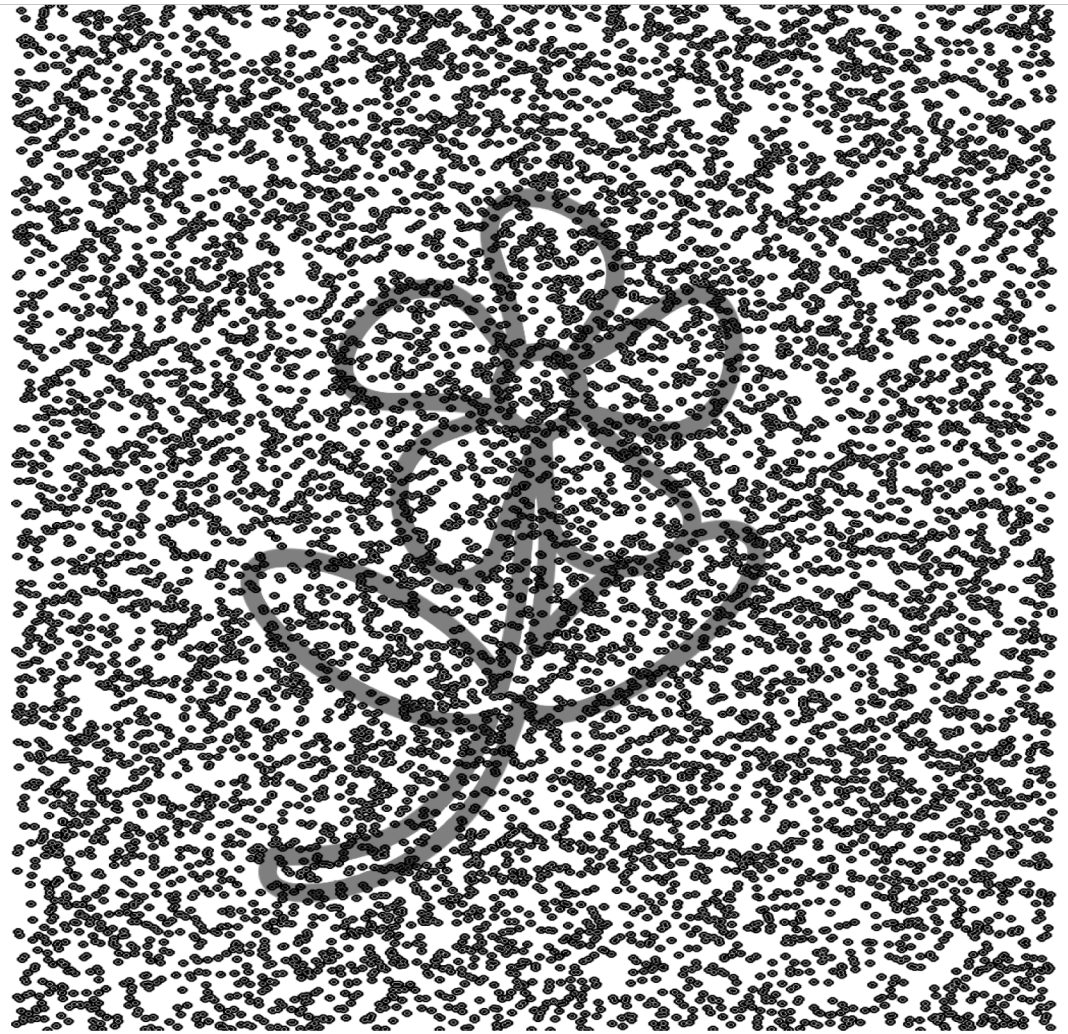
~ 25% of the US population has a hearing aid

(Fabry, 2011; Kochkin et al., 2010)

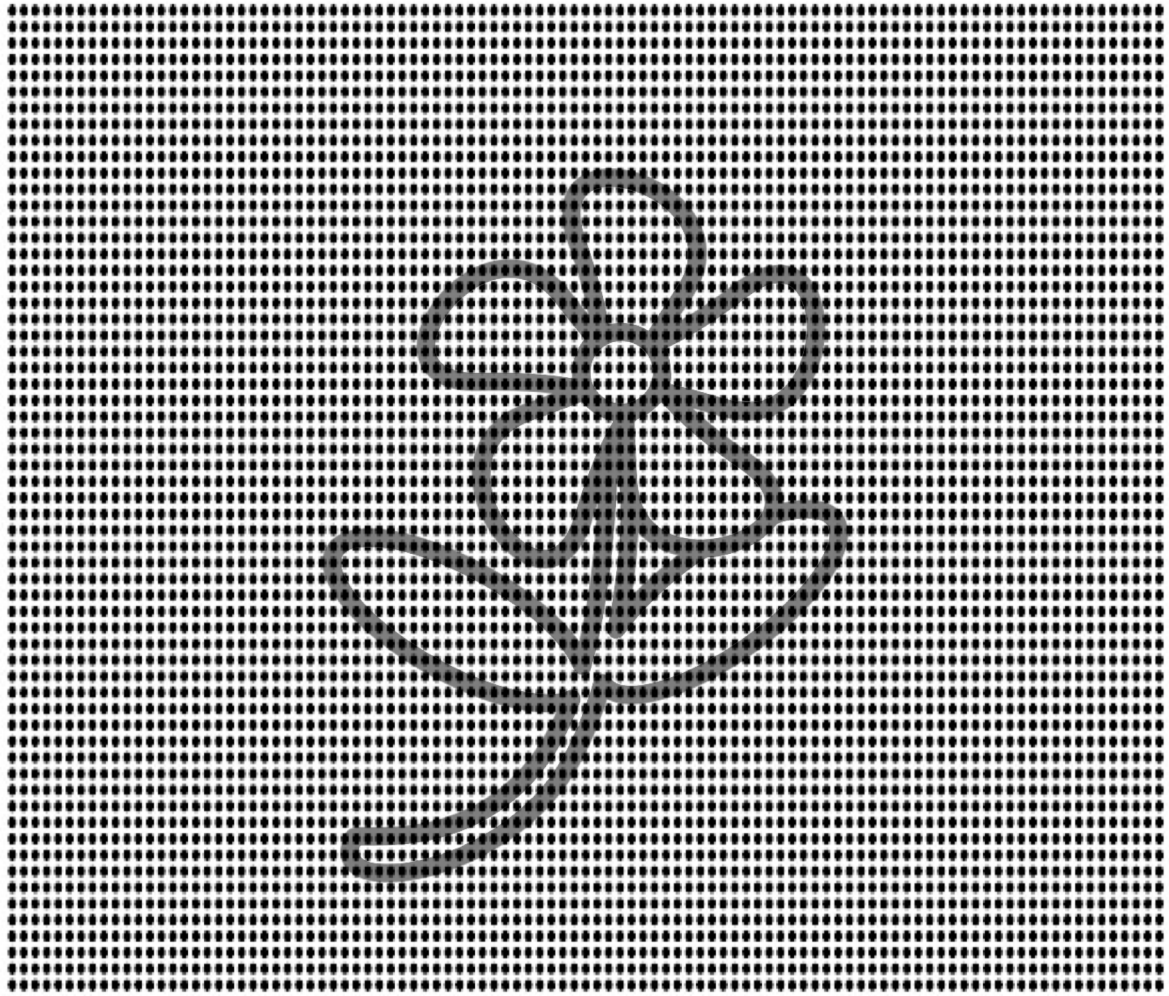
Informational Masking



Energetic Masking



Regular Energetic Masking



Behavioral Studies

Background

Initial Exploratory Experiment

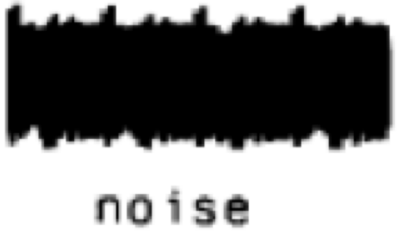
Follow-up Clarification Experiment

Question:

What role does temporal prediction play in speech perception in noise?

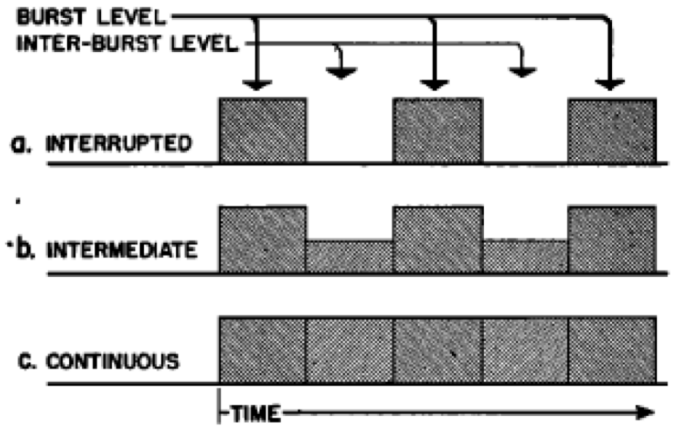
Previous Energetic Masking Work

- Steady-state masker



(Festen & Plomp, 1990)

- Periodic maskers, simple and variable noise power

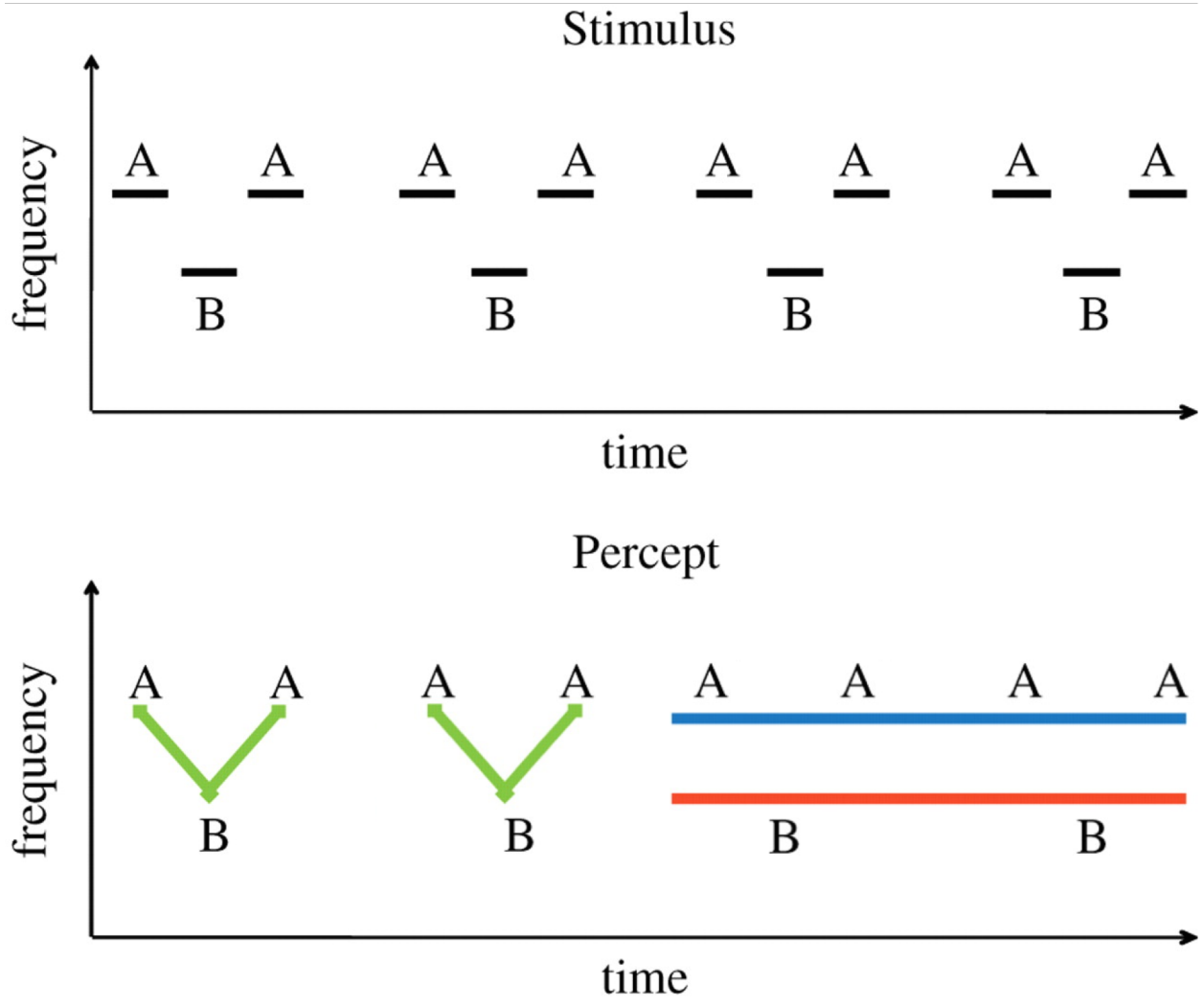


(Pollack, 1955)

- Complex structure, but no repeating rhythm

(Festen & Plomp, 1990)

Auditory scene analysis



Construct perceptual stream

Segregate streams

Attend to desired stream

Motivation for the experiments

- Noise causes decreased intelligibility
- Auditory object formation allows separation of auditory stimuli
- **Hypothesis: increased online rhythmic information will ease speech perception in noise**

Spoiler: It does

Behavioral Studies

Background

Initial Exploratory Experiment

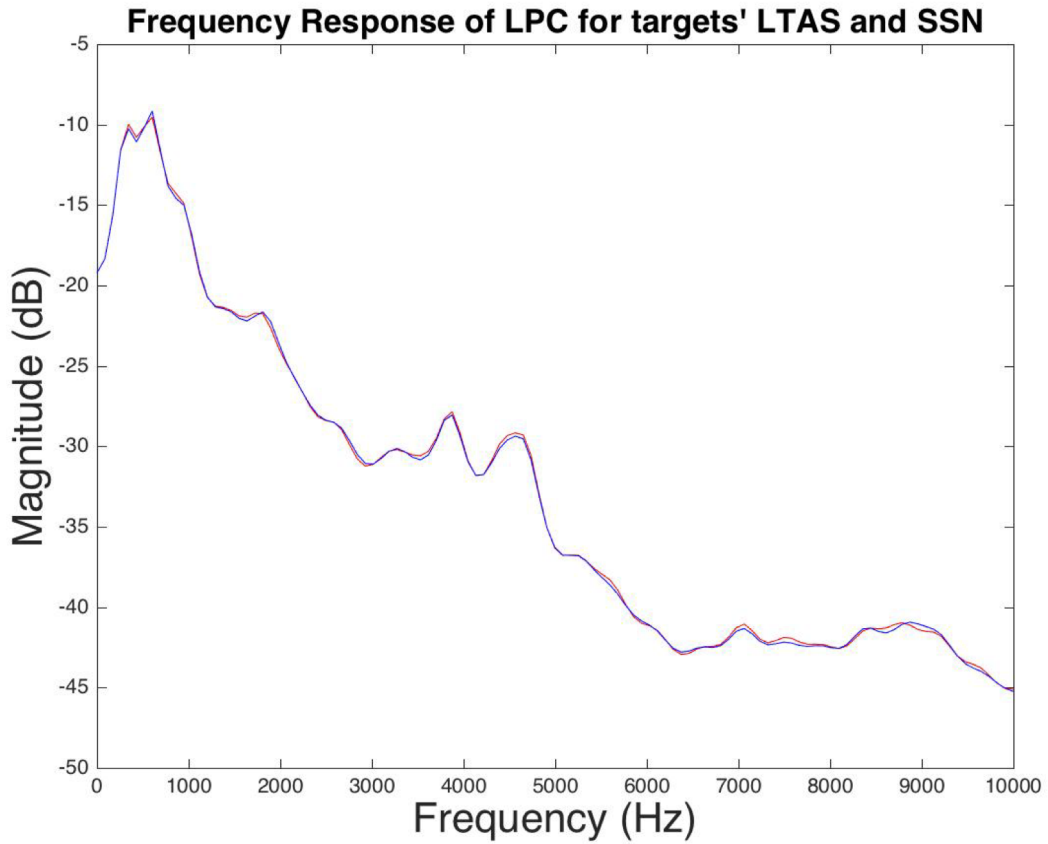
Follow-up Clarification Experiment

Question:

How do you create information from random noise?

Spectral profile of noise vs. speech

LTAS
↓
Glimpses
↓
Patterns
↓
Trials

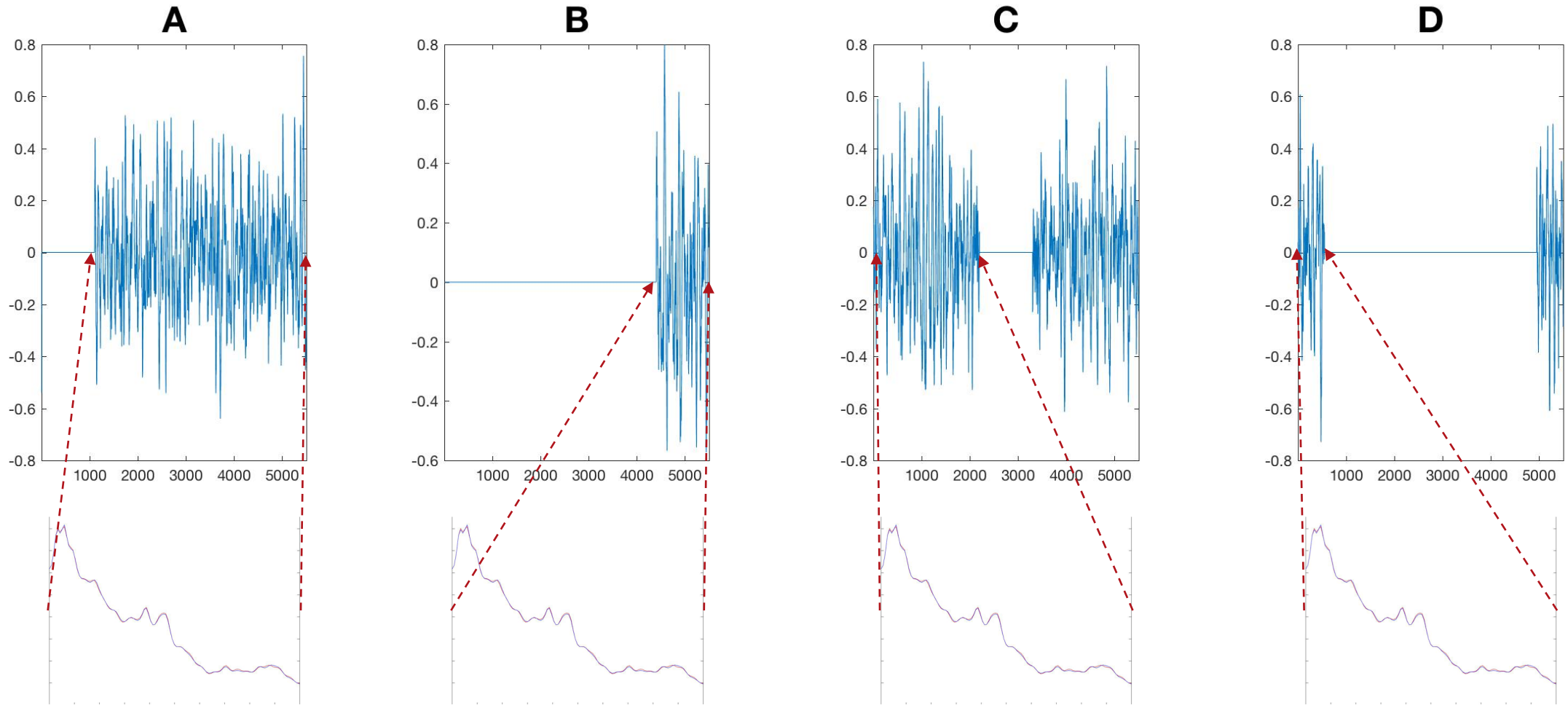


Noises' long-term average spectrum correlated to speech

Noise Patterns

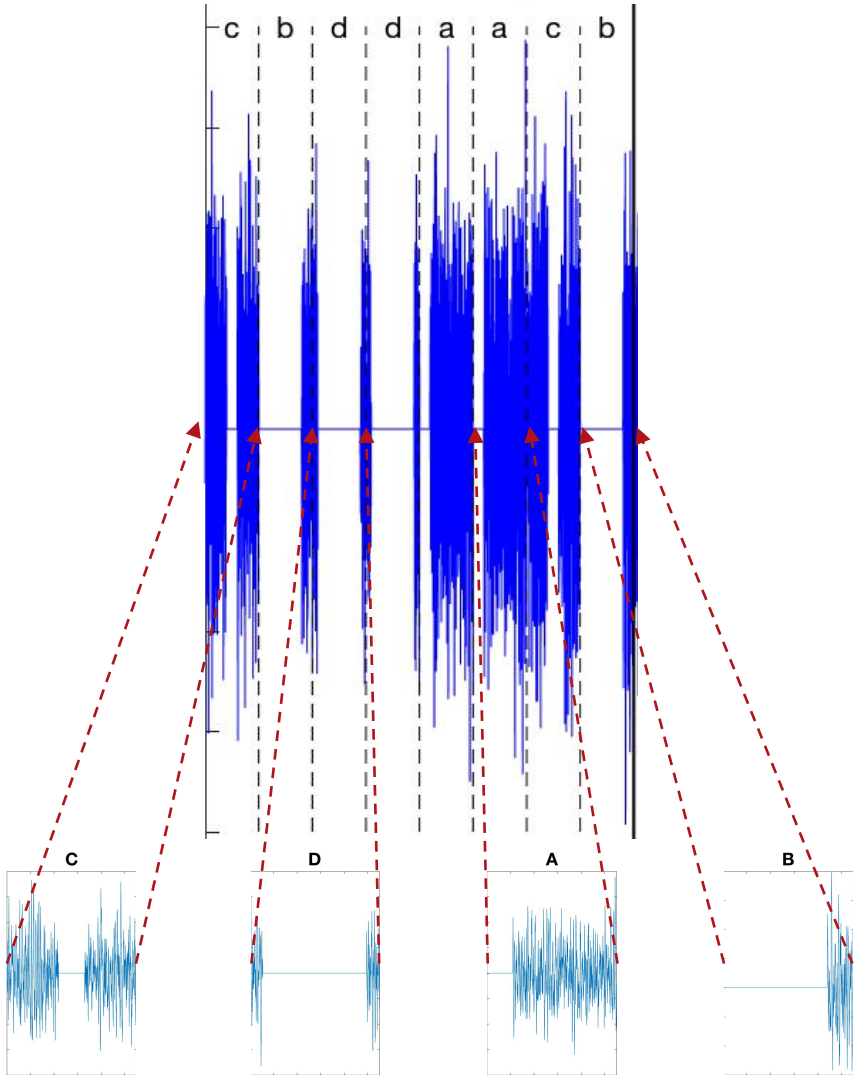
8Hz (125ms) windows containing noise and silence

LTAS
↓
Glimpses
↓
Patterns
↓
Trials



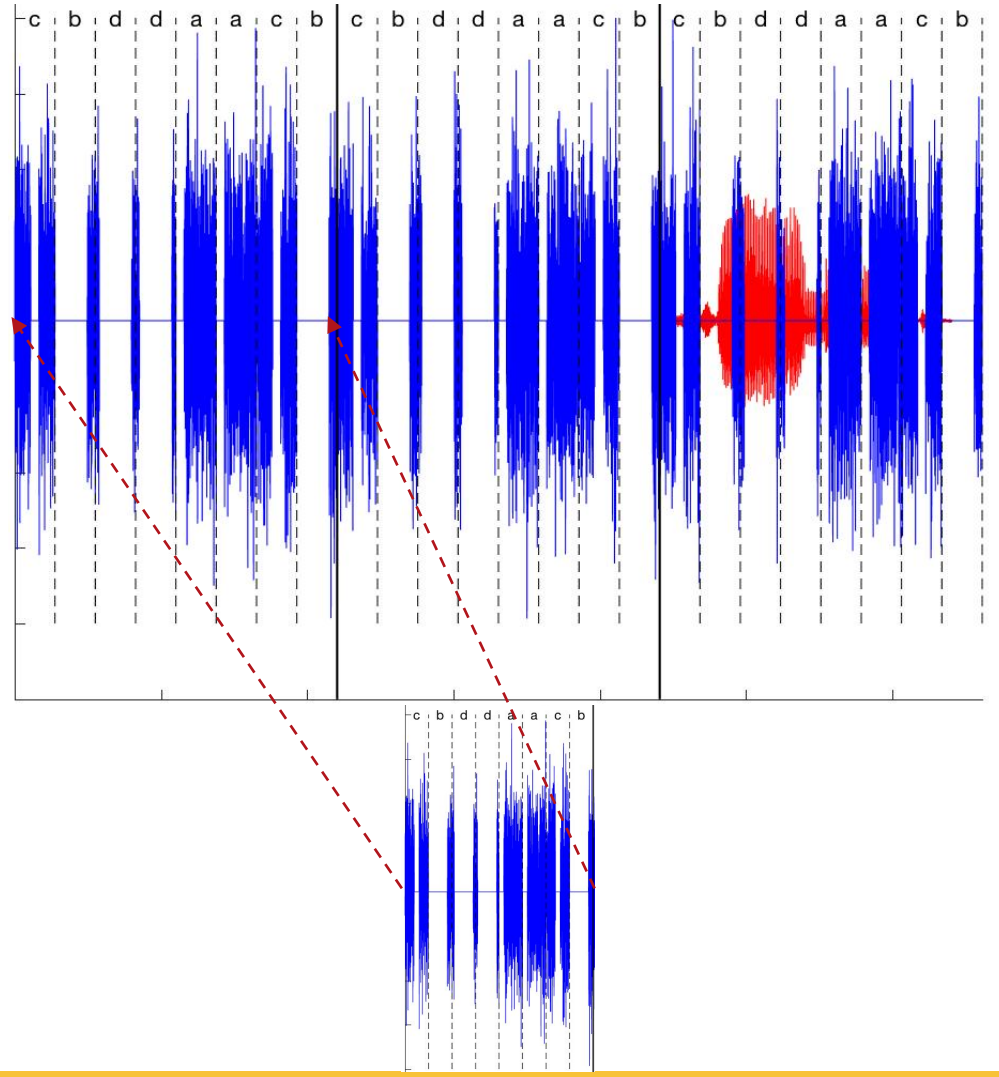
Provide temporal regularity in noise

LTAS
↓
Glimpses
↓
Patterns
↓
Trials



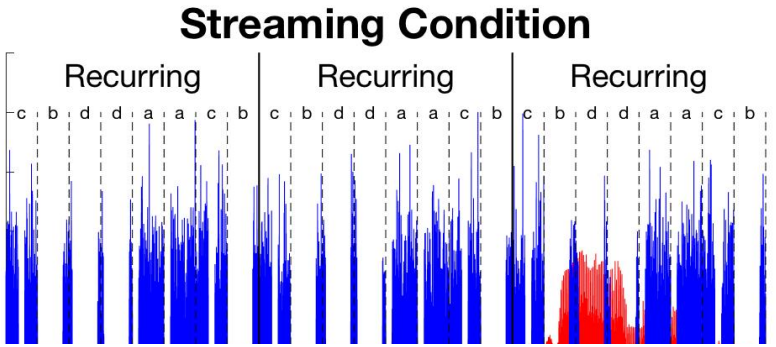
Provide temporal regularity in noise

LTAS
↓
Glimpses
↓
Patterns
↓
Trials

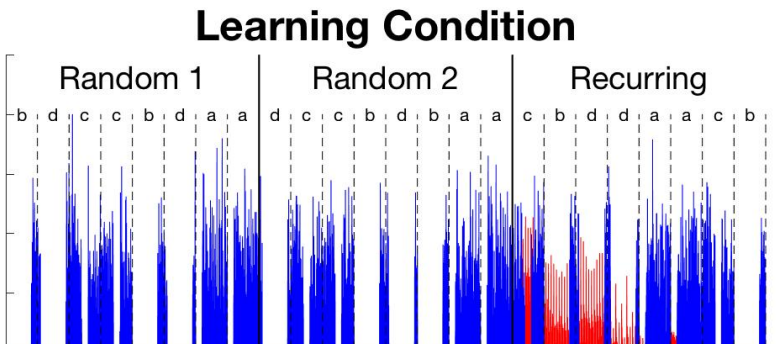


Conditions

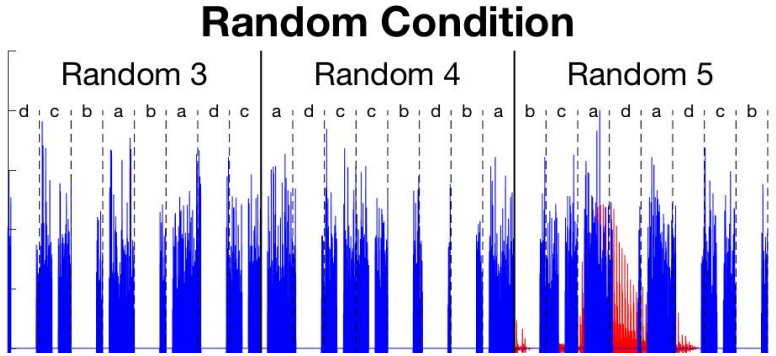
Repetitive information available for streaming



Control for SNR, glimpsing, and learning



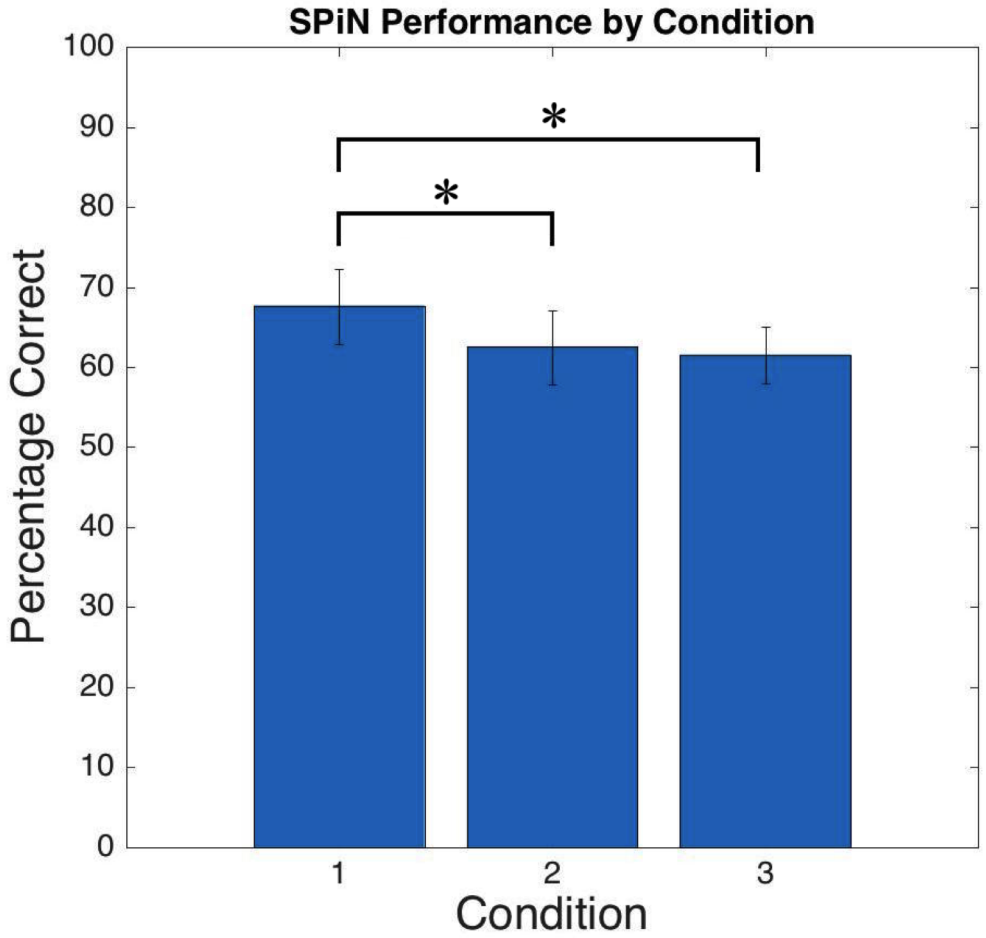
Control for SNR, baseline task performance



Experiment Format

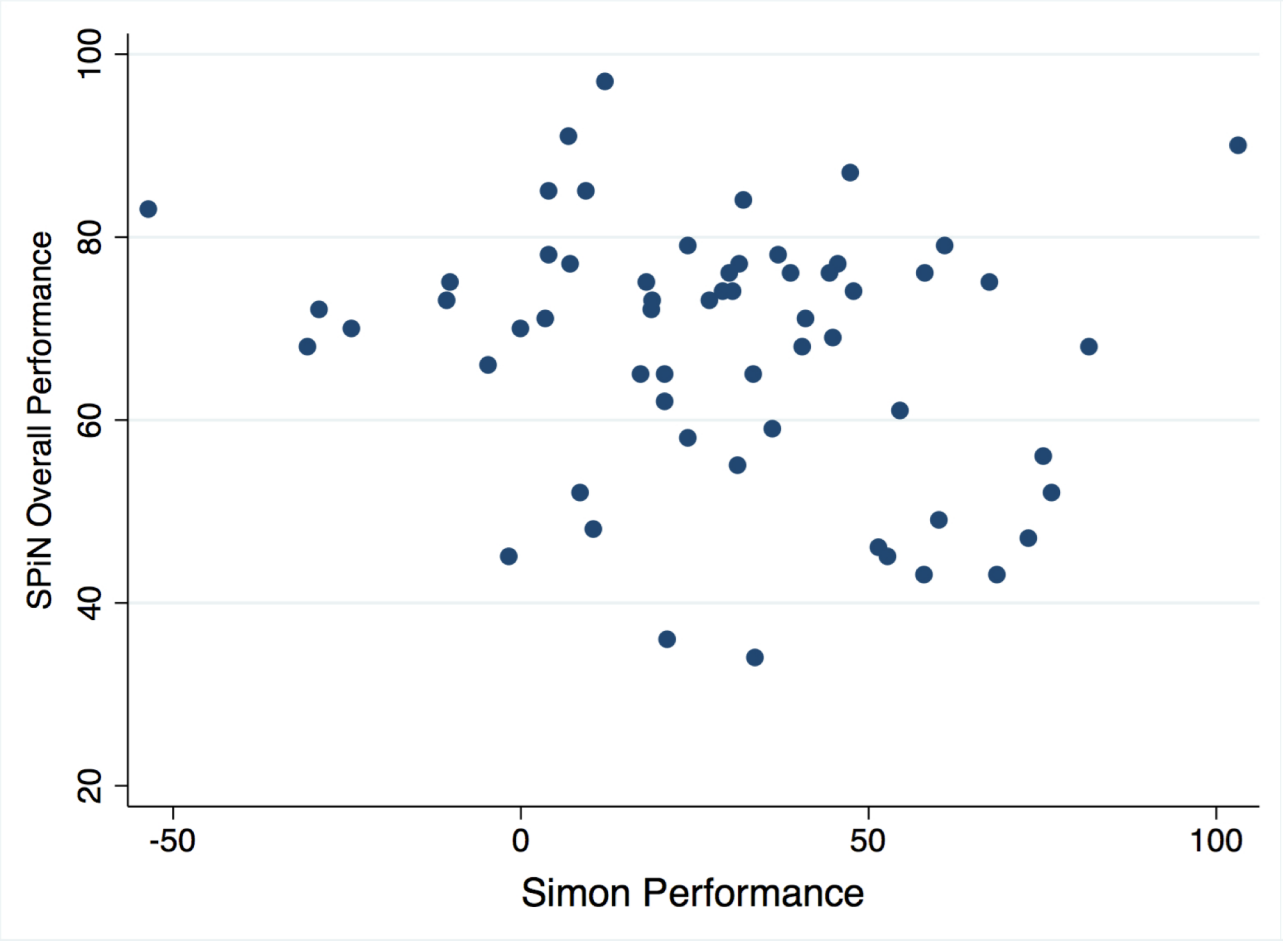
- Open-set word-recognition task
- Used 108 multisyllabic words as targets
- 60 Undergraduates at USC (n = 60)
- Executive function and working memory tasks followed

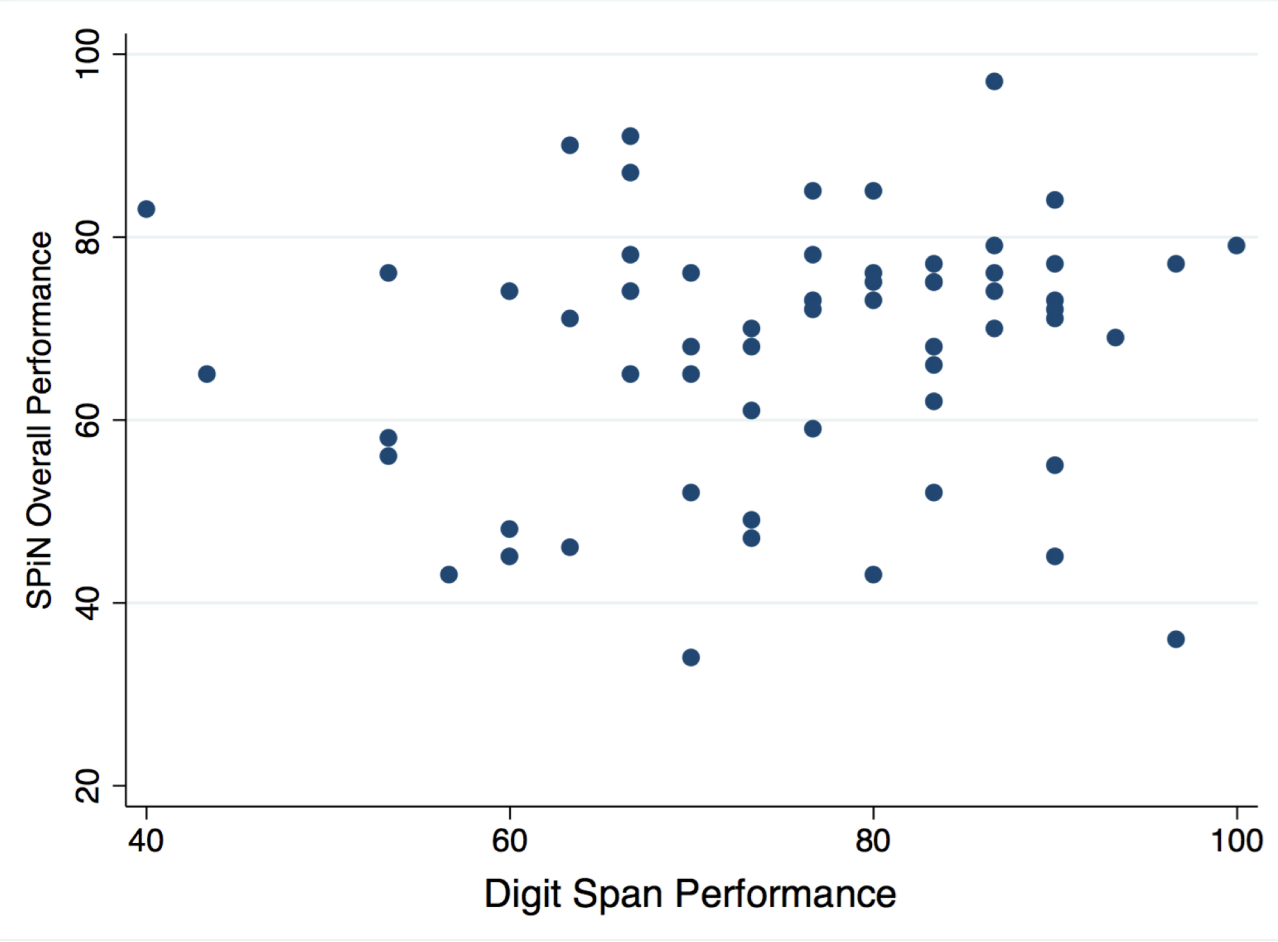
Temporal regularity significantly better



N = 60 Mean: 66% vs. 61% vs. 61% P-value: both <=.001

Executive Function & Working Memory tasks





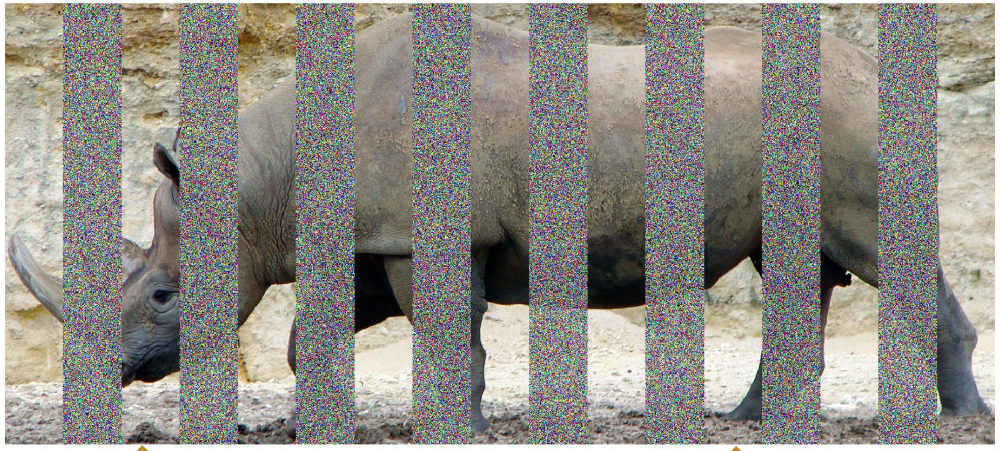
Behavioral Studies

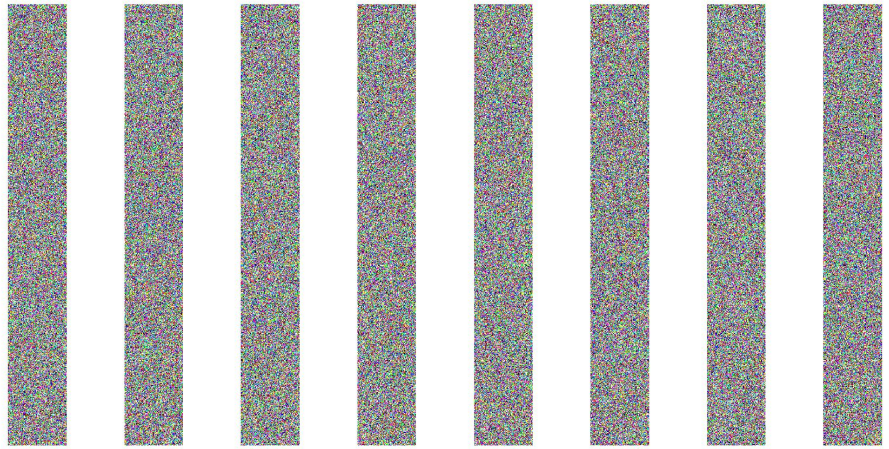
Background

Initial Exploratory Experiment

Follow-up Clarification Experiment

Is this caused by streaming?



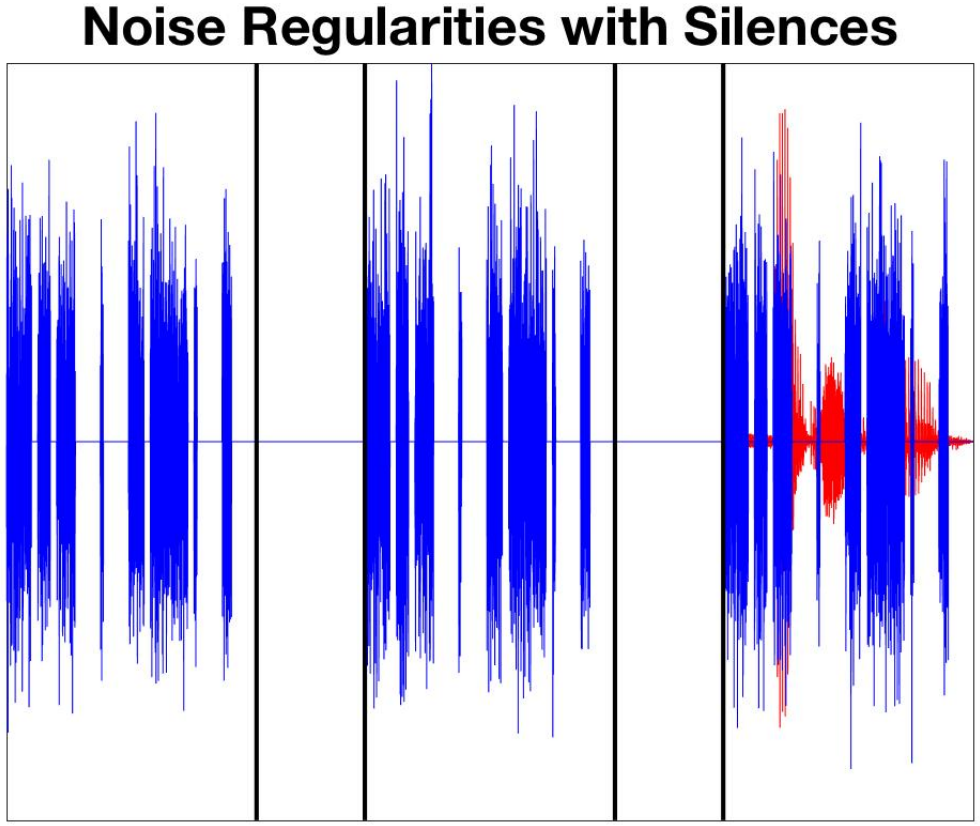


Is this caused by pure temporal prediction?

Experiment Format

- Open-set word-recognition task
- Used 108 multisyllabic words as targets
- 60 Undergraduates at USC (n = 60)
- Musical Experience Questionnaire Followed

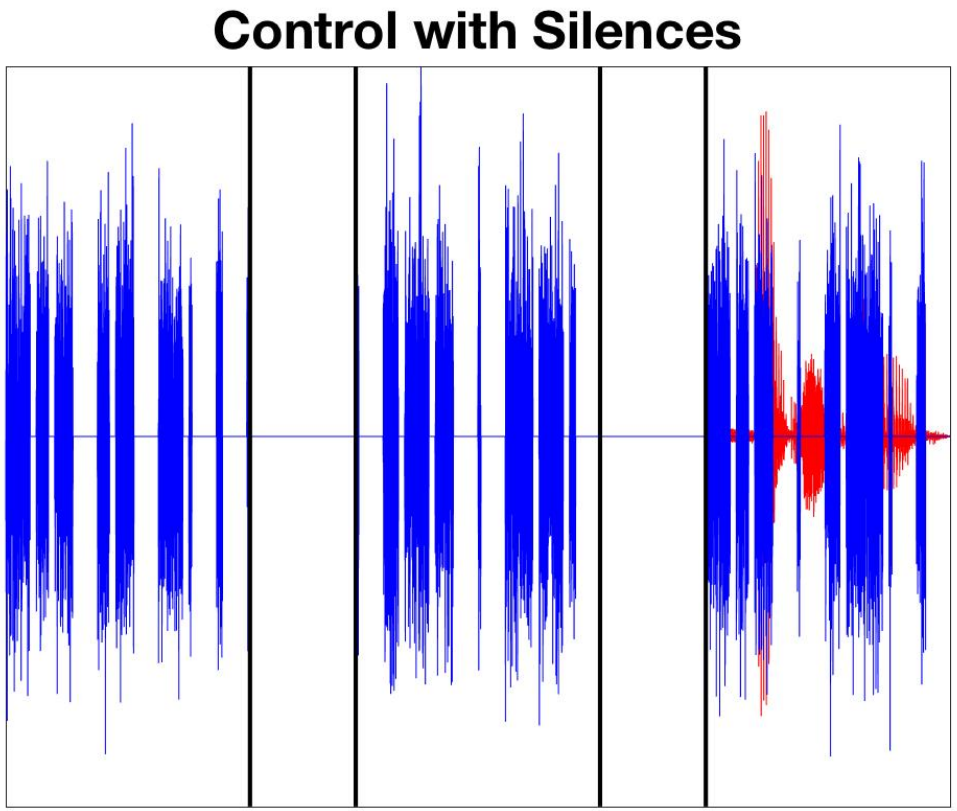
Insert silences to defeat streaming

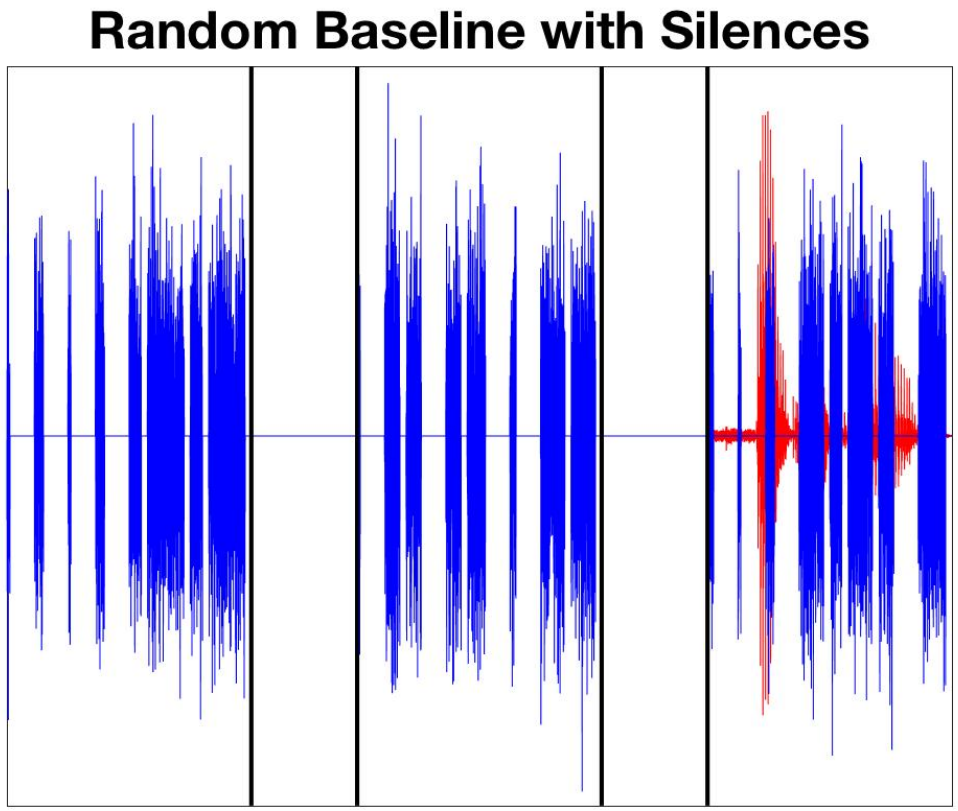


Bregman showed streaming “builds up”

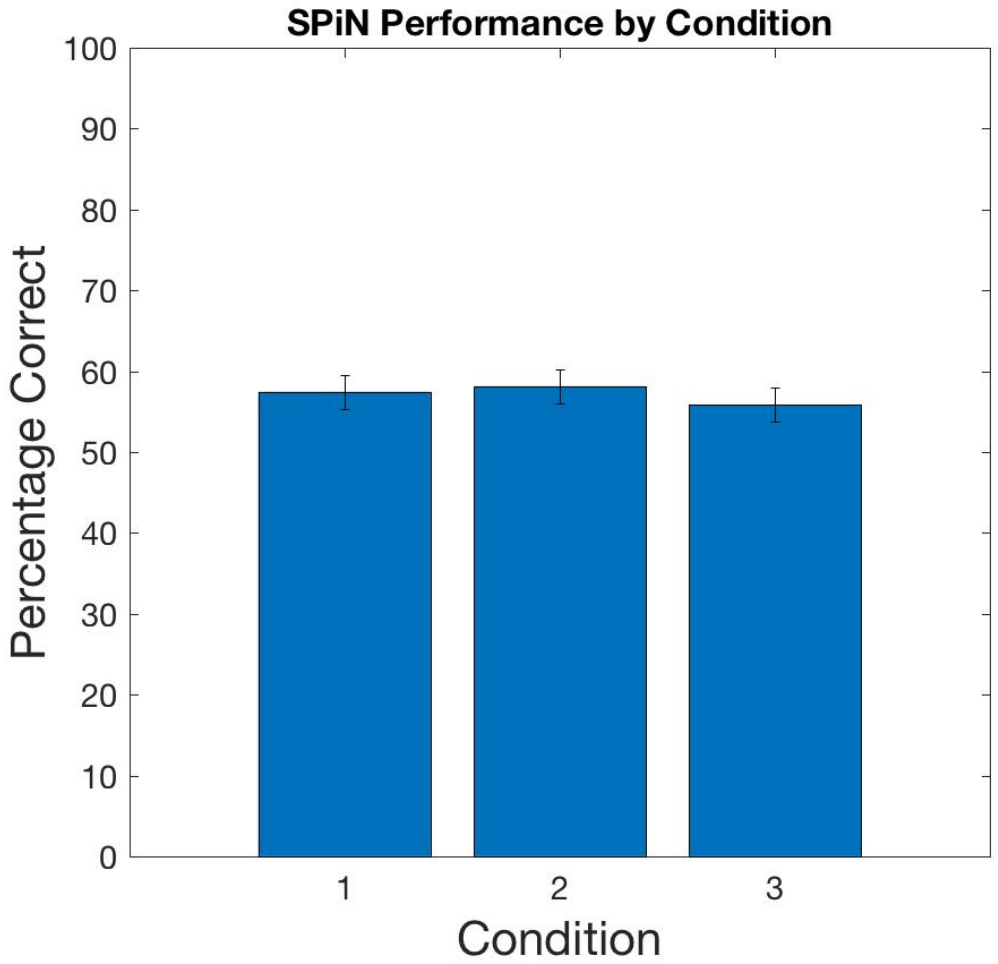
We inserted silences between patterns

Again compared with control and baseline





No significant difference



N=60 Mean: 57.4% vs. 58.1% vs. 55.8% P-value: .681, .269, and .129

We have evidence for streaming, now what?

Frequency manipulations
- Constructing noise “melody”

Manipulation of auditory attention
- Streaming highly dependent on attention

Testing of pattern learning vs. streaming
- Streaming patterns AAA, AAA, AAA vs. XXX, YYY, ZZZ

THANK YOU!

THANK YOU TO THE ZEVINLAB RAS AND THE PARTICIPANTS WHOSE DATA MADE THIS WORK POSSIBLE