

**Sundeeep Teki**<sup>1</sup>

**Maria Chait**<sup>3</sup>

**Deborah Williams**<sup>2</sup>

**Aiysha Siddiq**<sup>3</sup>

**Nicolas Barascud**<sup>3</sup>

**Sukhbinder Kumar**<sup>1,2</sup>

**Shihab Shamma**<sup>4</sup>

**Tim Griffiths**<sup>1,2</sup>



# **Temporal coherence and auditory object segregation in complex acoustic scenes**

- <sup>1</sup> Wellcome Trust Centre for Neuroimaging, University College London, UK
- <sup>2</sup> Newcastle Auditory Group, Newcastle University Medical School, UK
- <sup>3</sup> UCL Ear Institute, University College London, UK
- <sup>4</sup> Dept. of Electrical Engineering, University of Maryland, College Park, USA

# Auditory figure-ground segregation

## *Stimuli:*

- Studied using relatively simple signals, e.g. streaming signals

## *Mechanisms:*

- frequency selectivity
  - forward suppression
  - neural adaptation
- spatially segregated activation of neurons along the tonotopic axis corresponding to the two streams

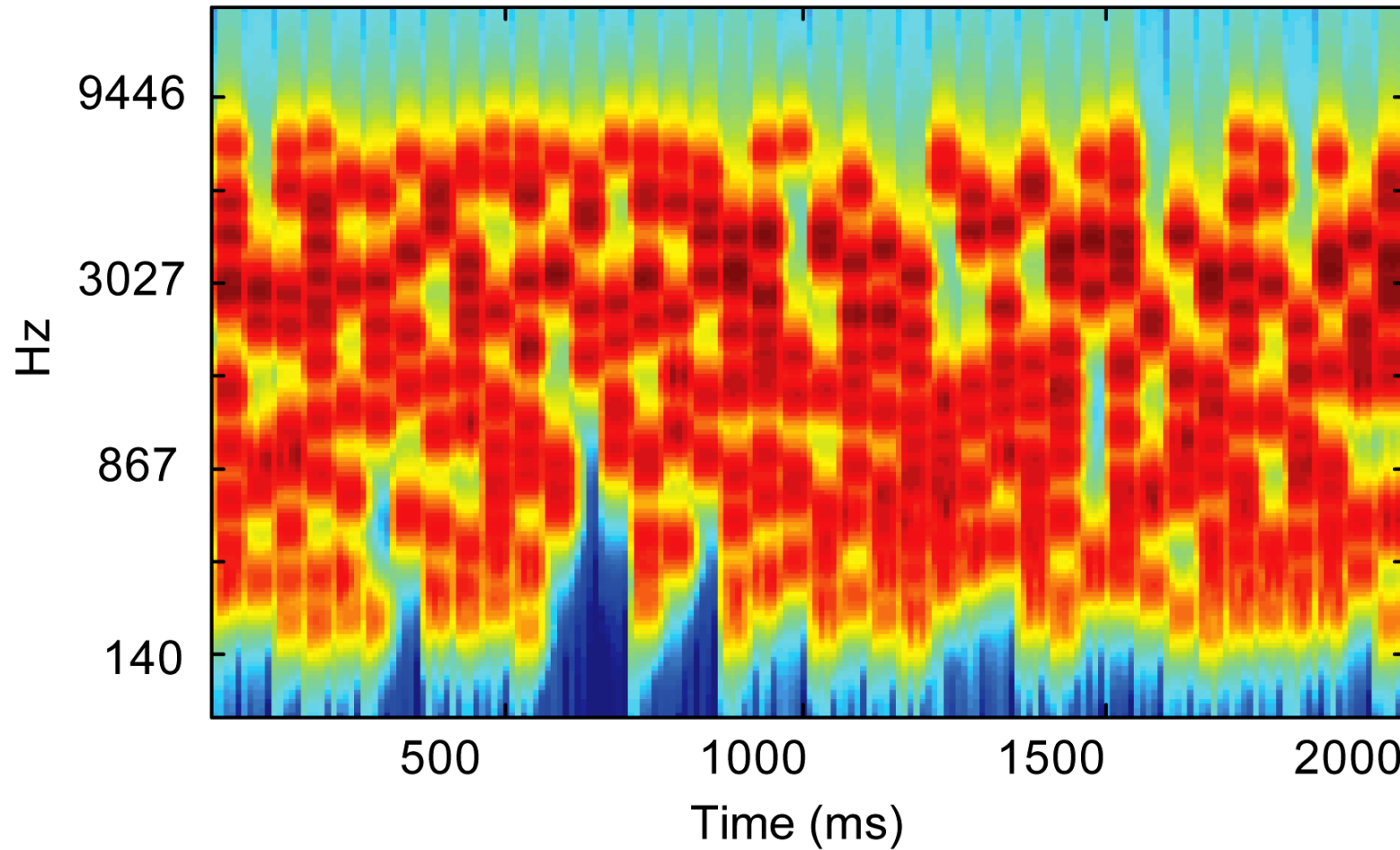
*c.f. Fishman/Steinschneider, Bee/Klump, Micheyl, Carlyon*

## *Drawbacks of streaming signals:*

- lack the rich spectrotemporal complexity of natural signals
- predictable temporal structure
- spectral components are non-overlapping and do not change with time

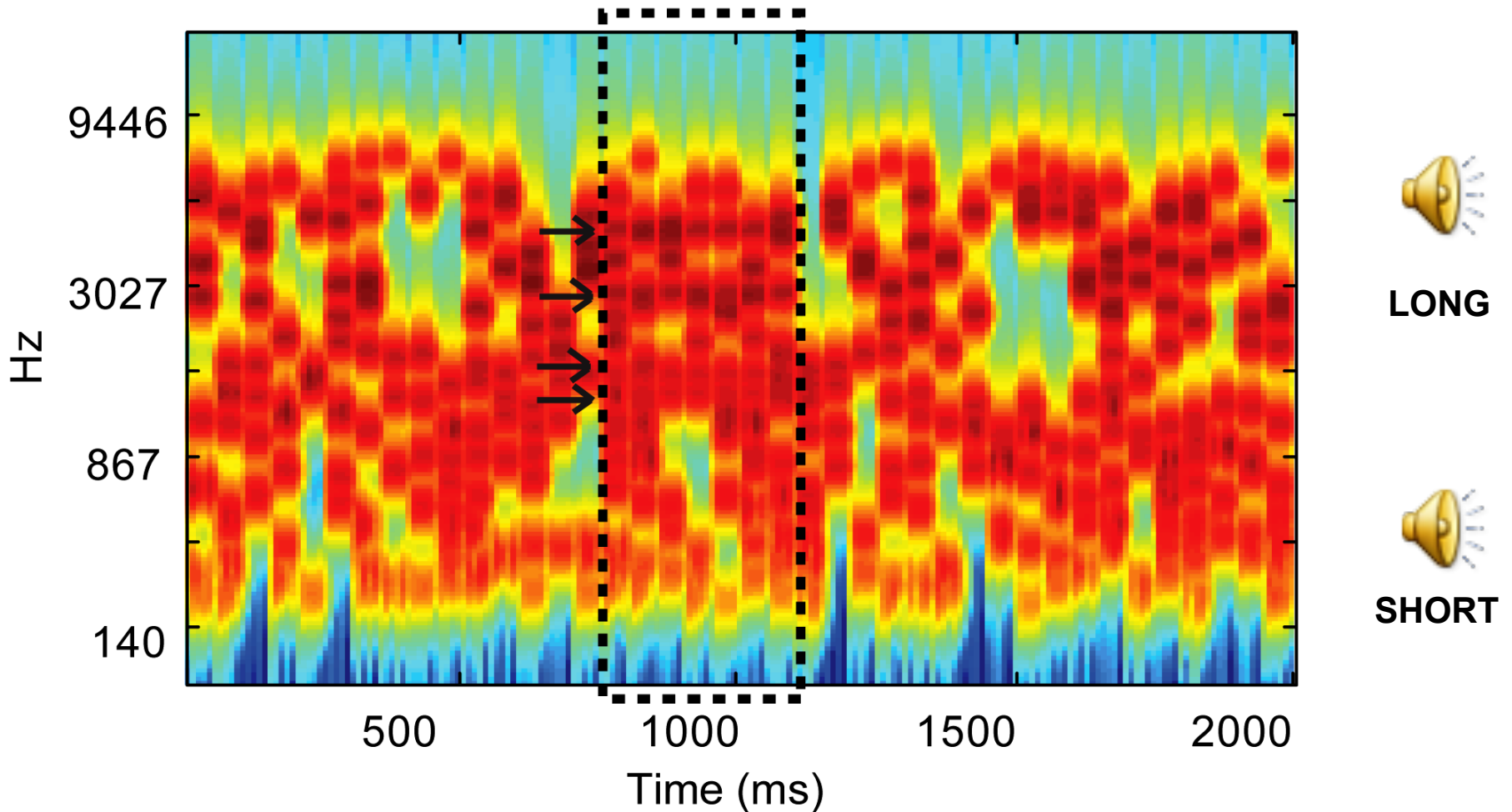
# **I. Stochastic figure-ground (SFG) stimulus**

# SFG: Figure absent



# SFG: Figure present

Figure with 'coherence' = 4 and 'duration' = 7



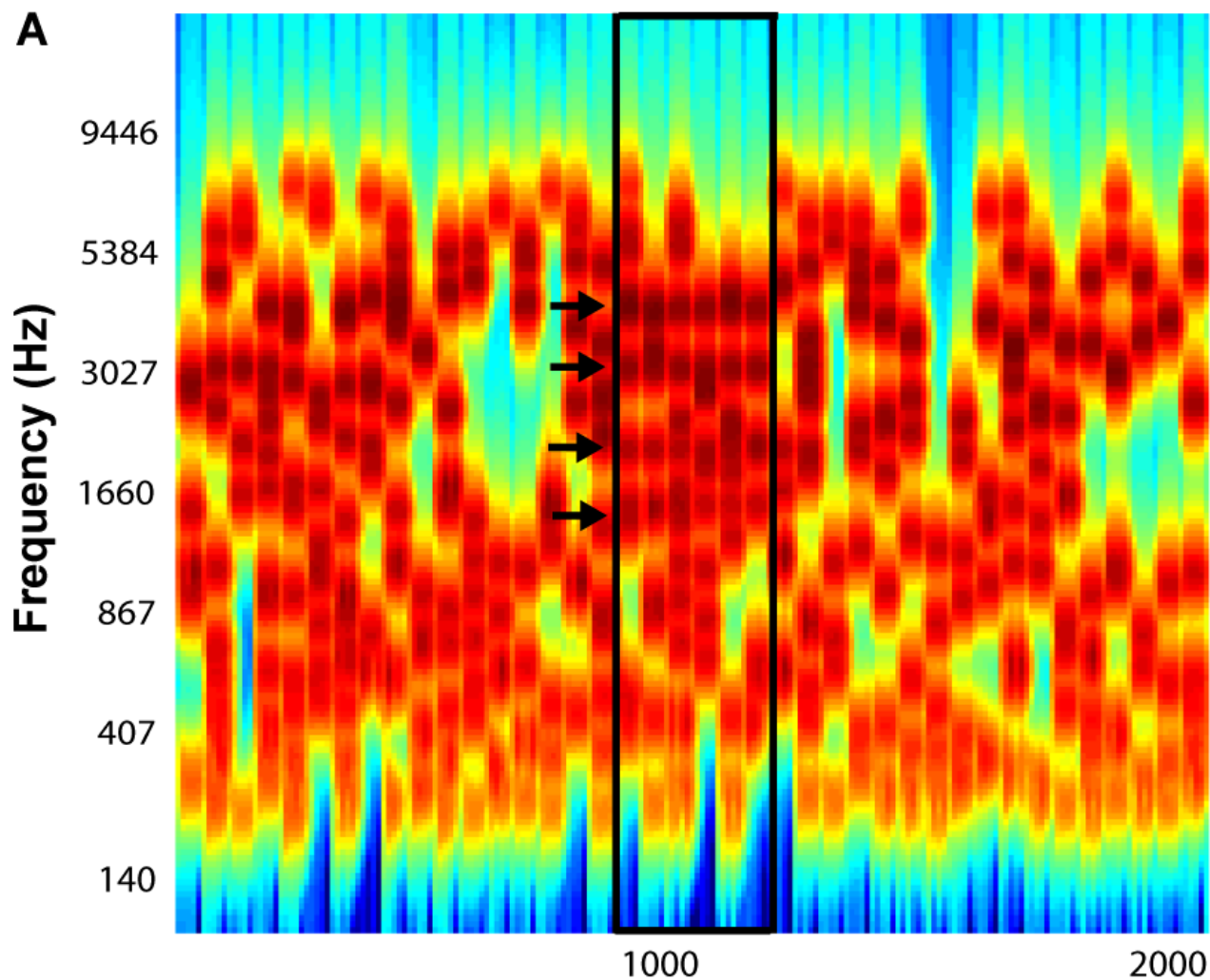
**Coherence:** 1, 2, 4, 6, 8

**Duration:** 2-7

## **II. Psychophysics**

# Expt. 1: 'Baseline' (50 ms)

Stimulus consisted of a sequence of 40 x 50ms chords (2 s long)

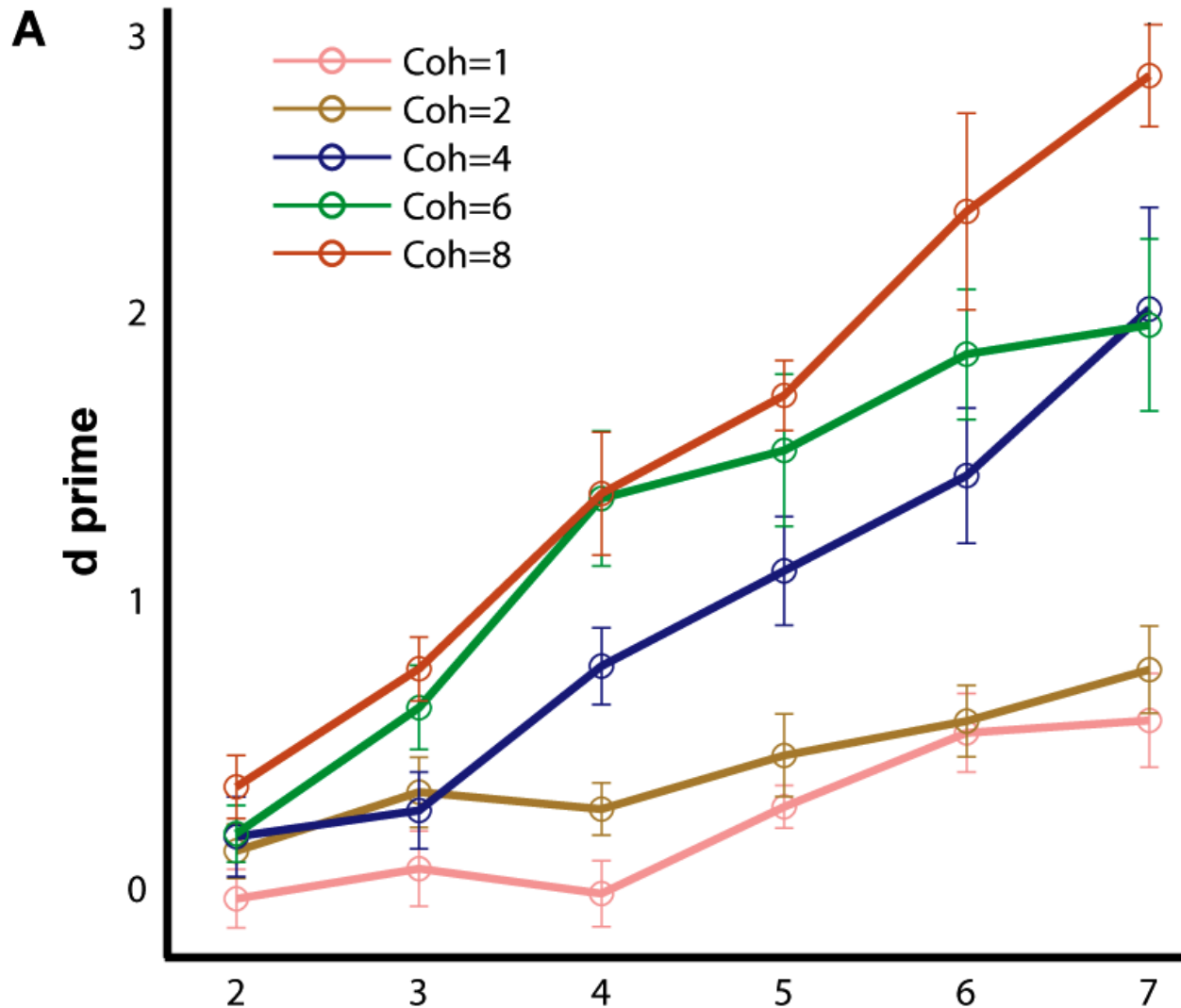


**Coherence:** [1 2 4 6 8]

**Duration:** [2:7]

# Expt. 1: Results

(n=9)

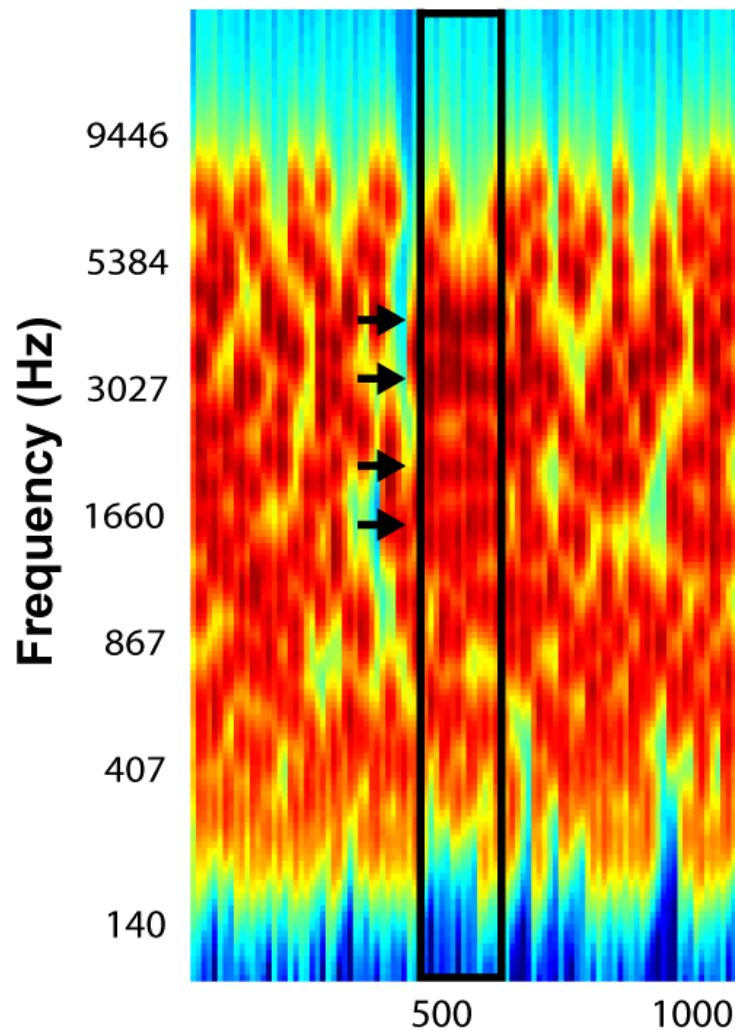




# Expt. 2: 'Baseline' (25 ms)

Stimulus consisted of a sequence of 40 x 25ms chords (1 s long)

B

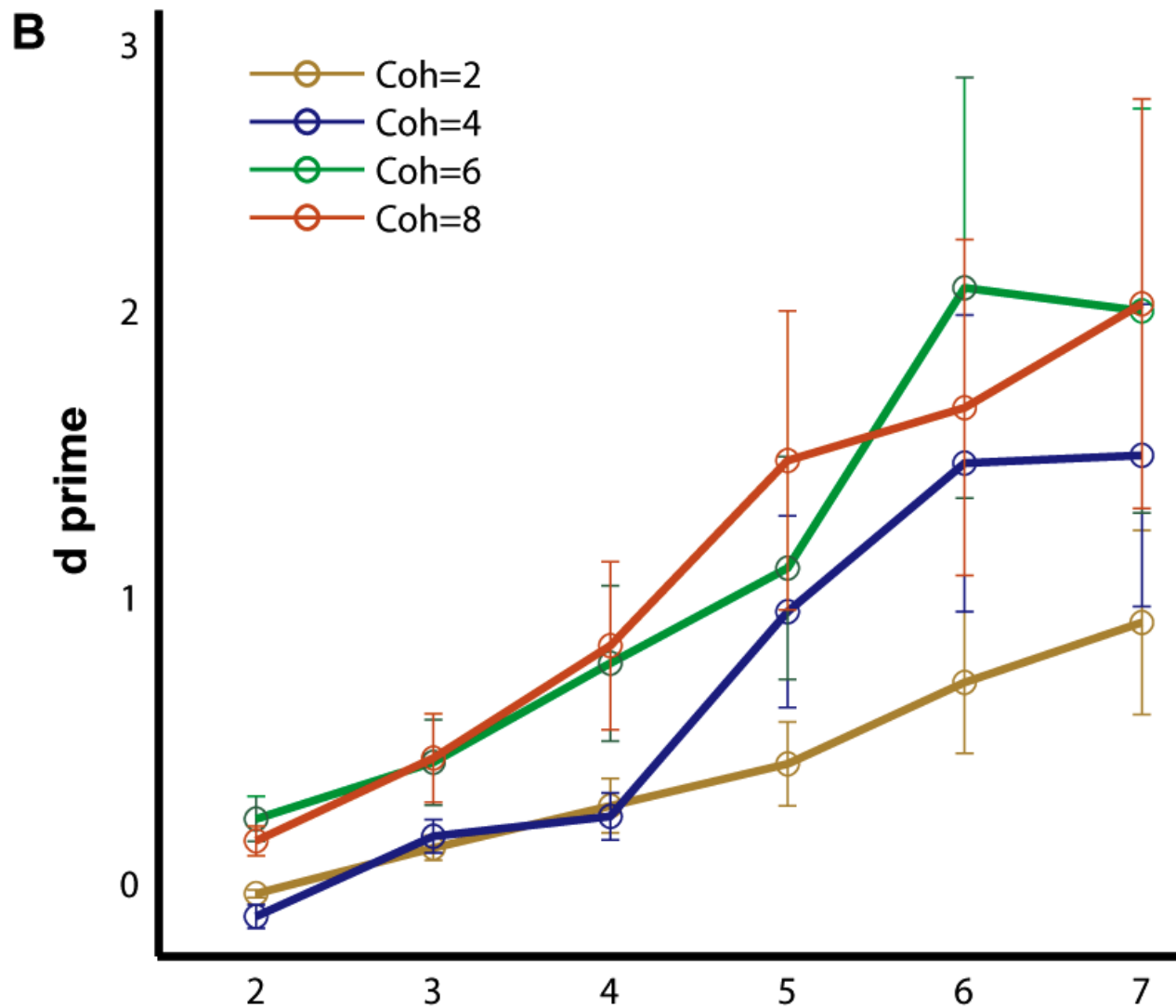


Coherence: [2 4 6 8]

Duration: [2:7]

# Expt. 2: Results

(n=8)



# Expt. 1 vs. 2

## ANOVA

- Coherence and duration as within-subject factors
- Chord length (50 ms vs. 25 ms) as between-subject factor

## Results

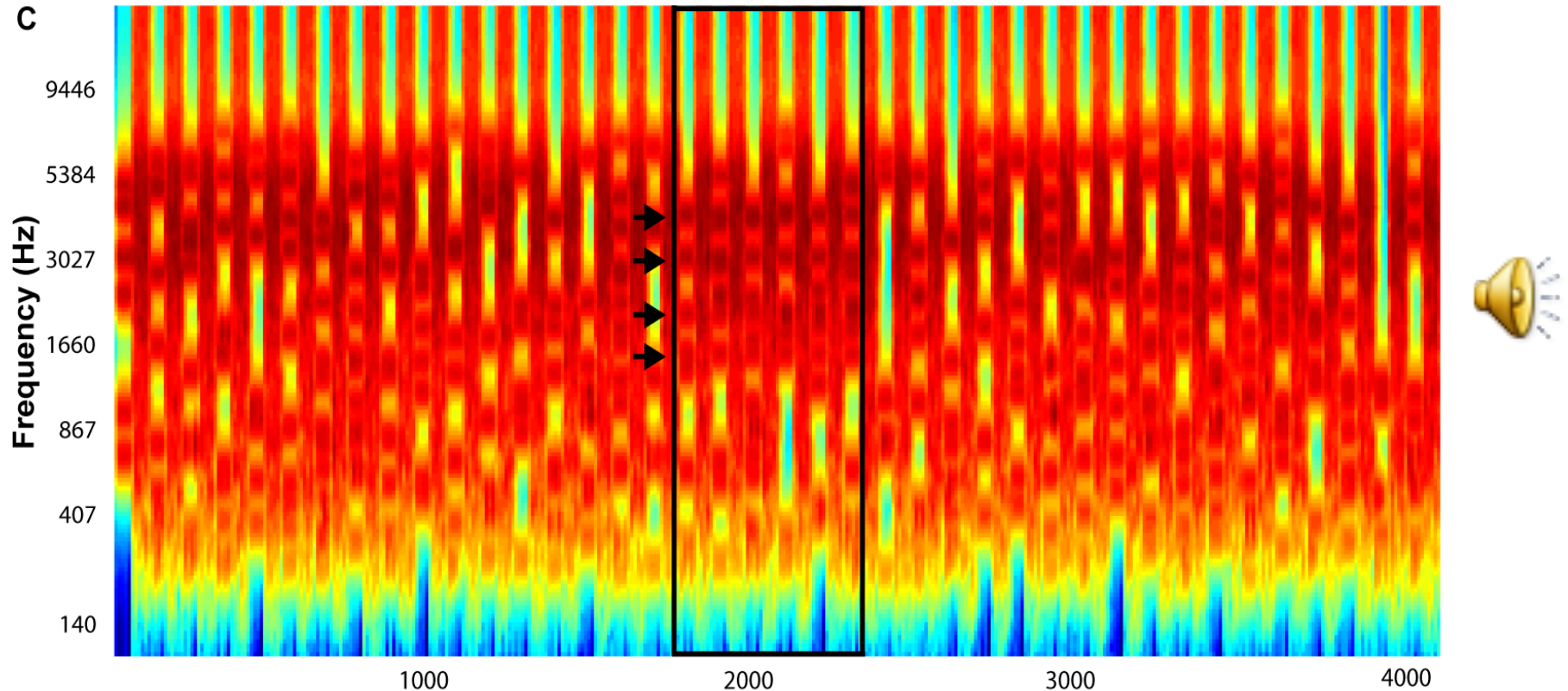
Significant effect of coherence:  $F(3, 45) = 77, p < 0.001$

Significant effect of duration:  $F(5, 75) = 41, p < 0.001$

No significant effect of chord length:  $F(1, 15) = 2, p = 0.174$

# Expt. 3: 'SFG/Noise'

Stimulus: SFG with 40 x 50ms chords alternating with 50ms of white noise (4 s)

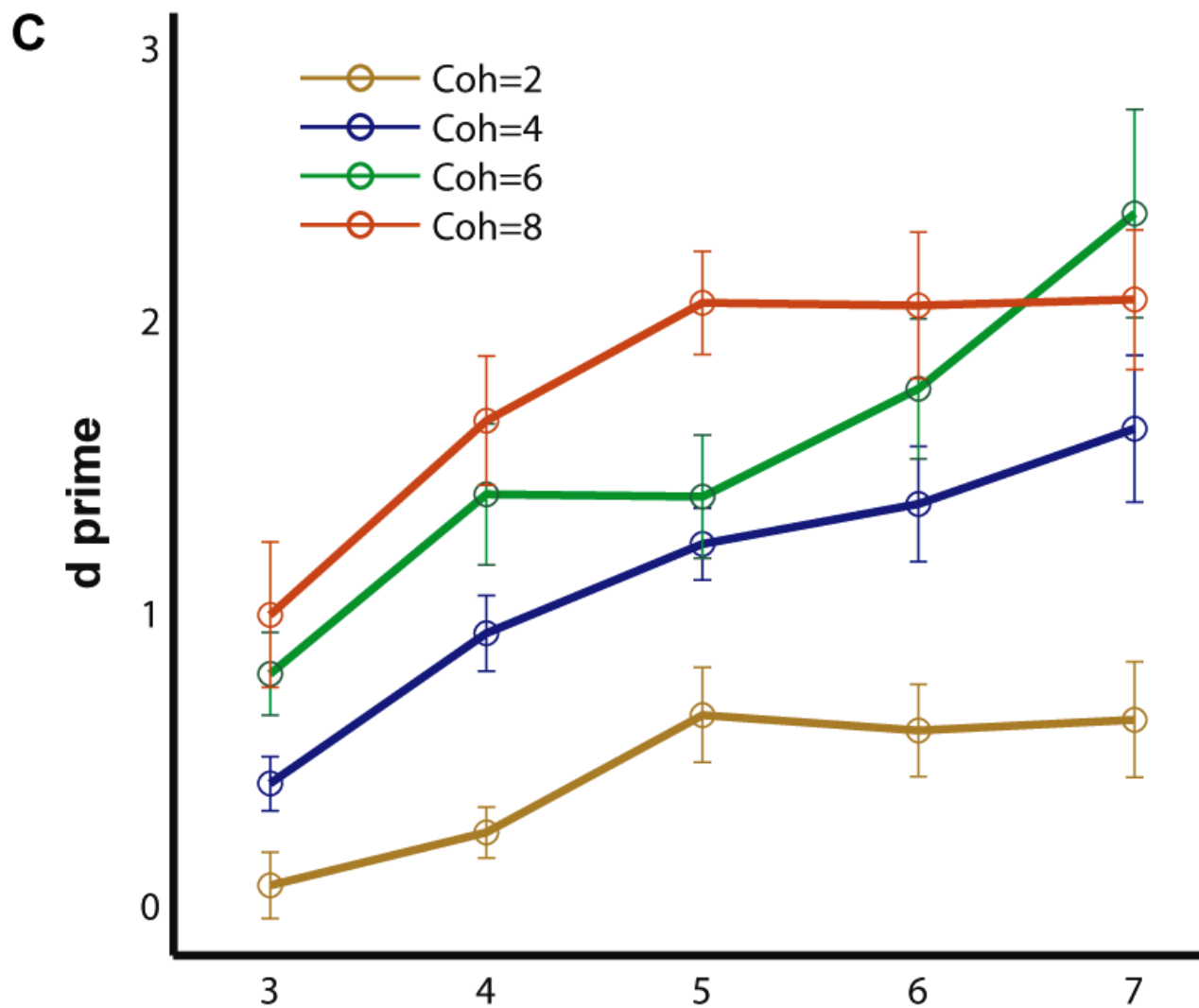


Coherence: [2 4 6 8]

Duration: [3:7]

# Expt. 3: Results

(n=10)



# Expt. 1 vs. 3

## ANOVA

- Coherence and duration as within-subject factors
- Condition (Baseline vs. SFG/Noise) as between-subject factor

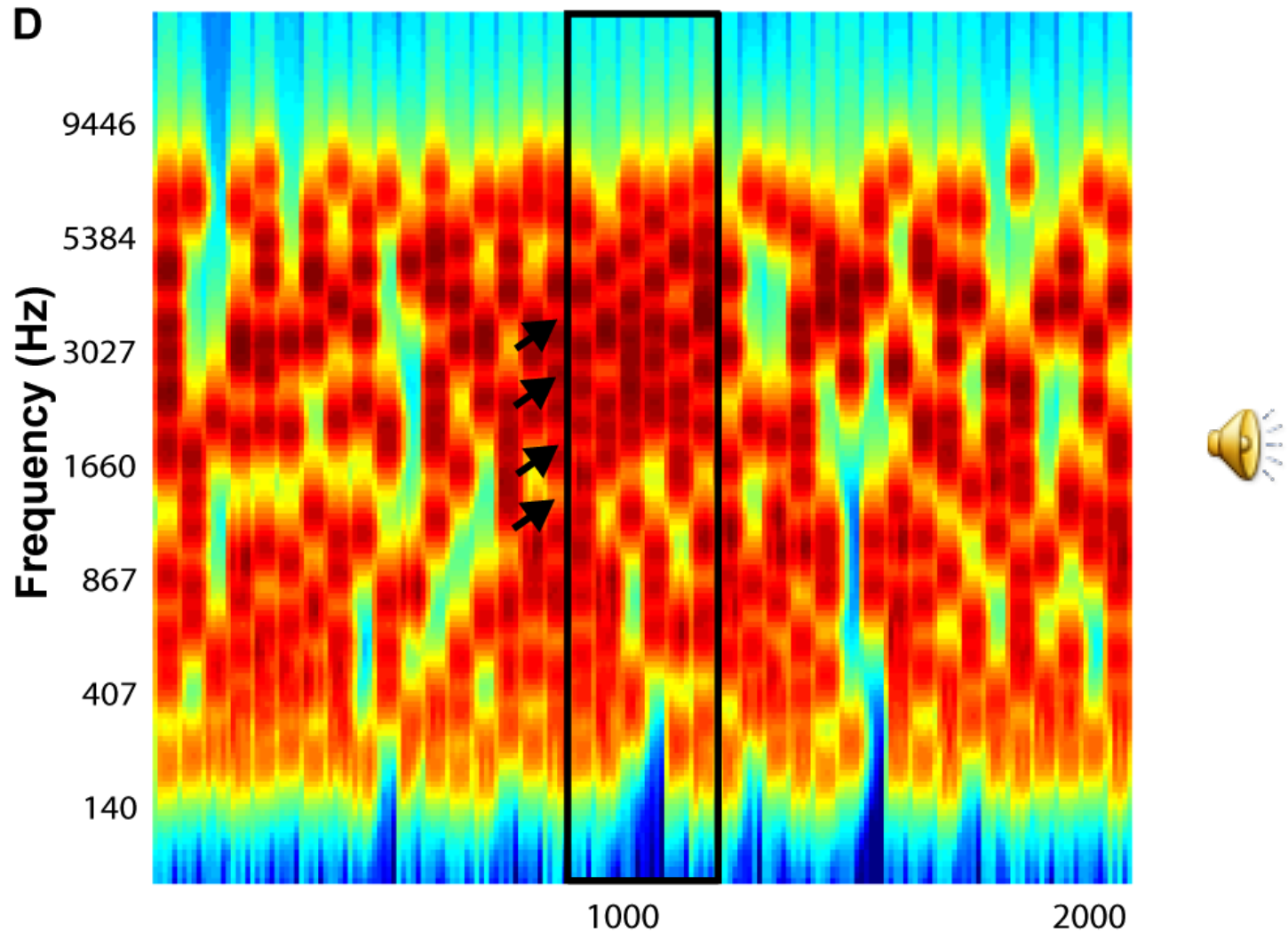
## Results

Significant effect of coherence:  $F(3, 51) = 23, p < 0.001$

Significant effect of duration:  $F(4, 68) = 29, p < 0.001$

No significant effect of condition:  $F(1, 17) = 0.004, p = 0.953$

# Expt. 4: 'Ramps'



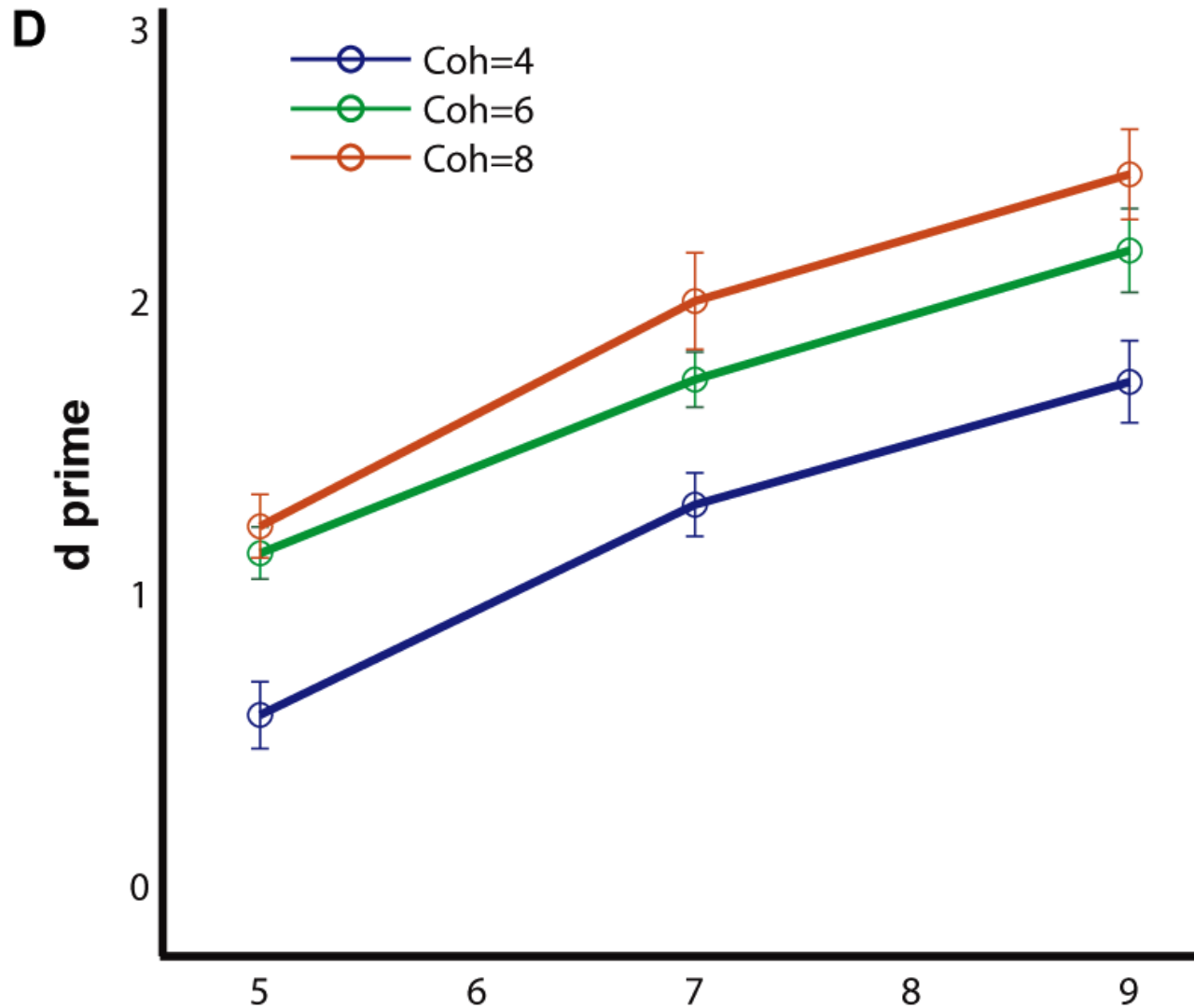
**Coherence:** [4 6 8]

**Duration:** [5 7 9]

**Ramp step:** [2 5] \* 1/24 oct.

# Results: Ramps 2

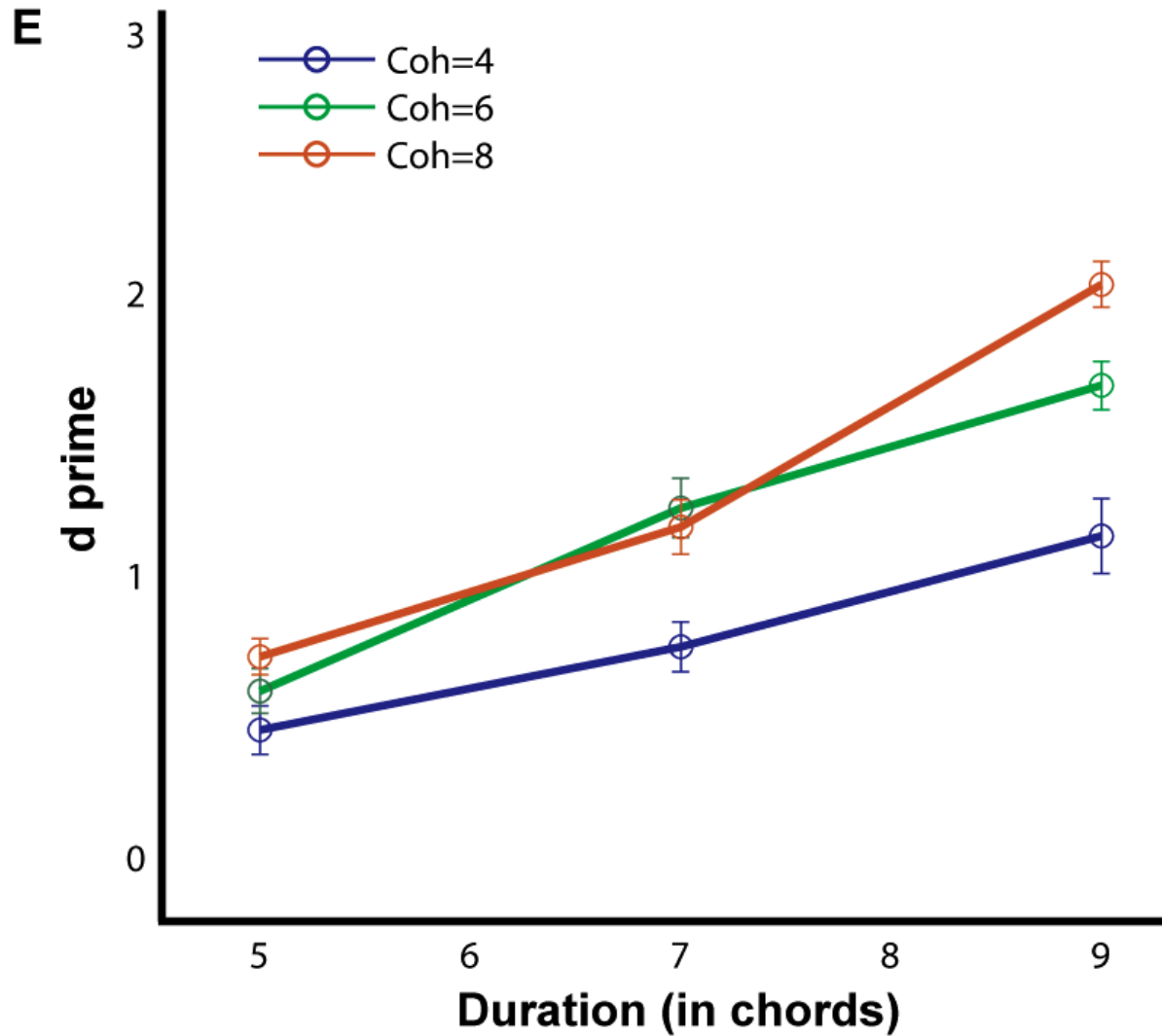
(n=10)





# Results: Ramps 5

(n=10)



# Expt. 1 vs. 4a vs. 4b

## ANOVA

- Coherence (4, 6, 8) and duration (5, 7) as within-subject factors
- Condition (Baseline vs. ramp of 2 vs. ramp of 5) as between-subject factors.

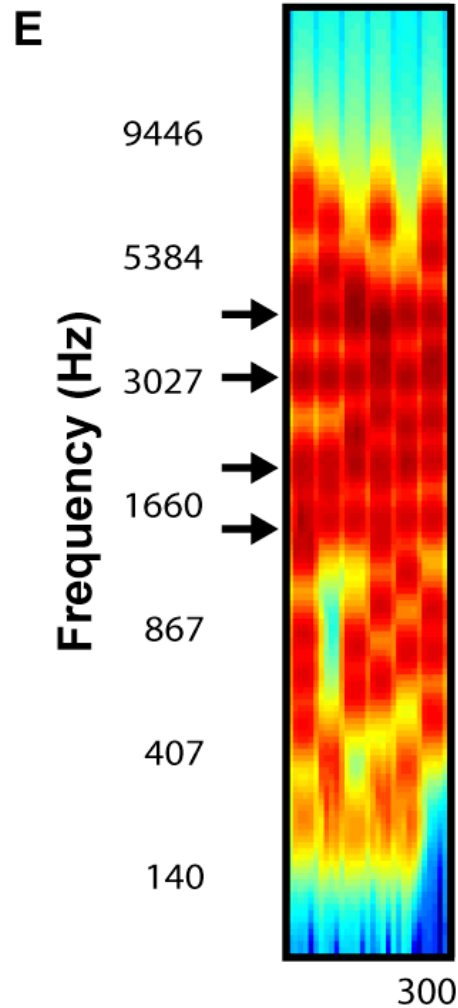
## Results

**Significant** effect of coherence:  $F(2, 50) = 25, p < 0.001$

**Significant** effect of duration:  $F(1, 25) = 110, p < 0.001$

**Significant** effect of condition:  $F(2,25) = 19, p < 0.001$

# Expt. 5: 'Isolated'



**Figure:**



**Ground:**

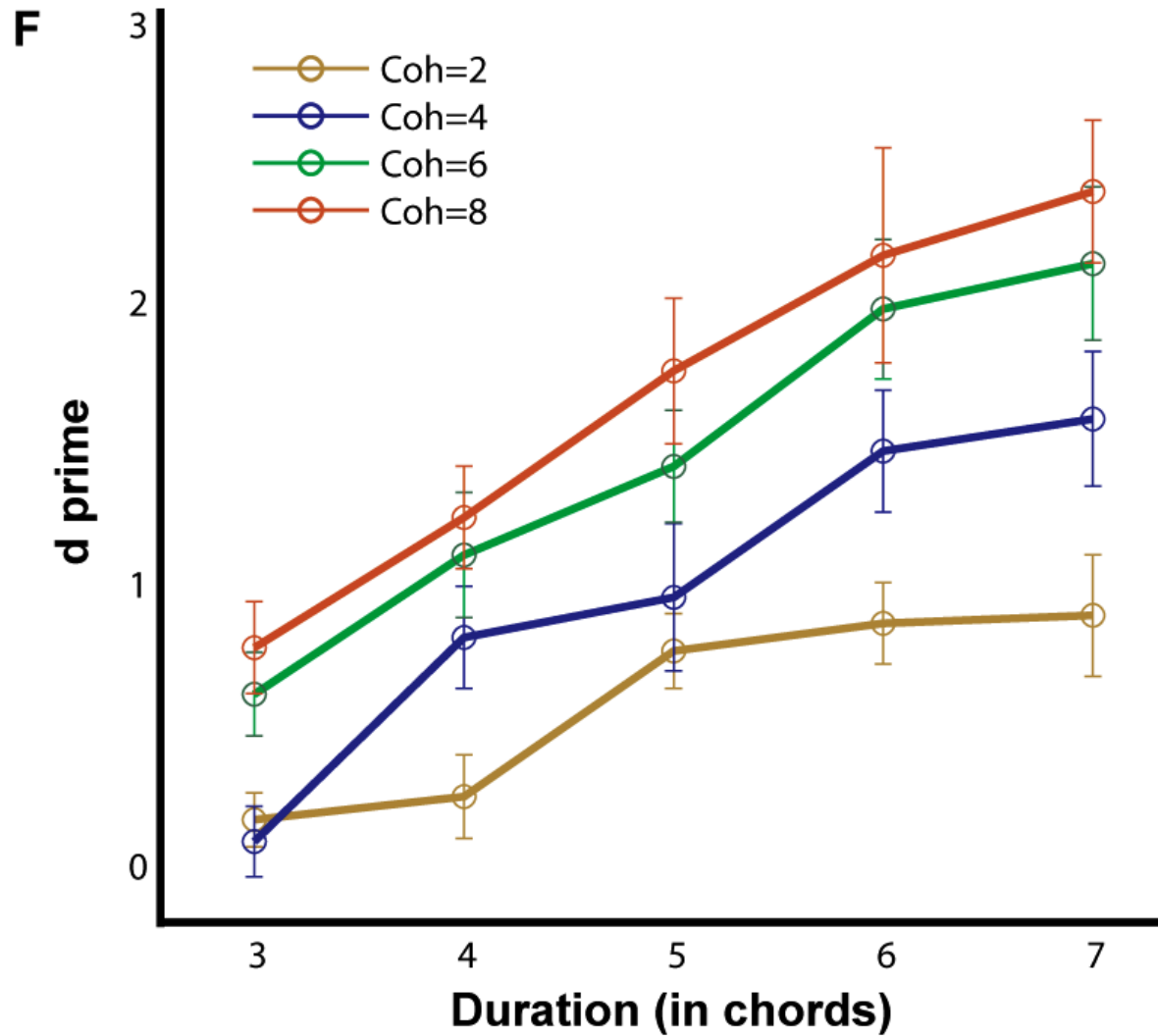


**Coherence:** [2 4 6 8]

**Duration:** [3:7]

# Expt. 5: 'Isolated'

(n=10)



# Expt. 1 vs. 5

## ANOVA

- Coherence and duration as within-subject factors
- Condition (Baseline vs. Isolated) as between-subject factor

## Results

Significant effect of coherence:  $F(3, 48) = 85, p < 0.001$

Significant effect of duration:  $F(4, 64) = 69, p < 0.001$

No significant effect of condition:  $F(1, 16) = 0.033, p = 0.859$

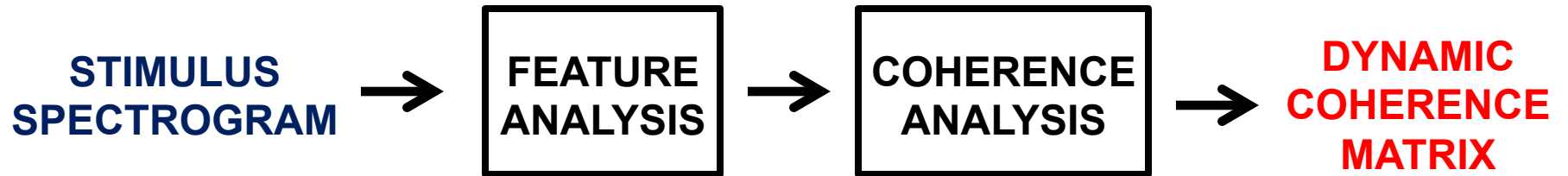
# Psychophysics summary

**Figure-detection performance in complex SFG stimulus is:**

- **Dependent on no. of repeating chords, not duration of figure** (Expt. 1 & 2)
- **Invariant to interference by white noise** (Expt. 1 & 3)
- **Sensitive to shape of target (continuous vs. ramped)** (Expt. 1 & 4)
- **Invariant to the presence of preceding background** (Expt. 1 & 5)

# **III. Temporal coherence modelling**

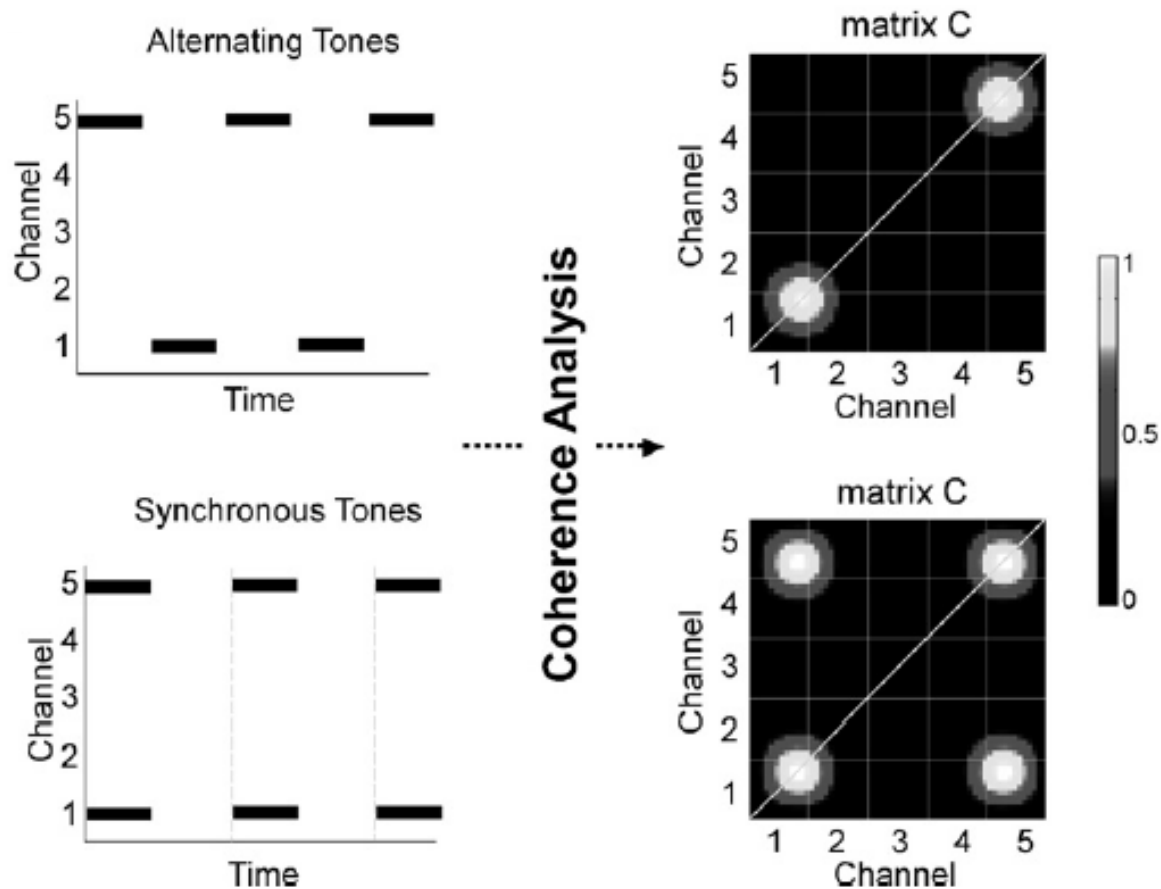
# Temporal coherence model



*Chi et al., 2005; JASA*  
*Elhilali and Shamma, 2008; JASA*  
*Elhilali et al., 2009; Neuron*  
*Shamma et al., 2011 TiNS*



# Temporal coherence & Streaming



# Temporal coherence & SFG

## Hypotheses:

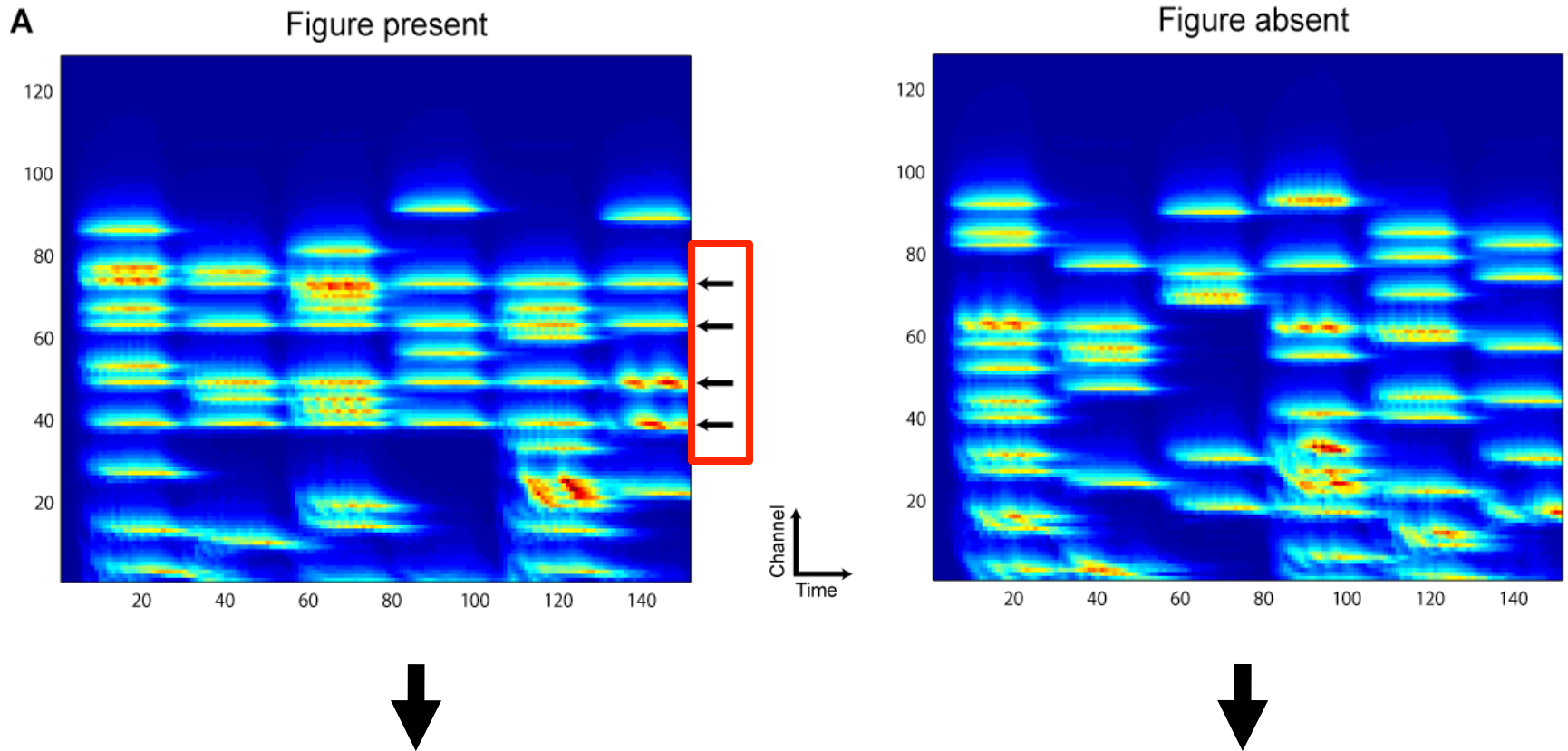
Channels with repeating frequency components would be temporally coherent; and these components may be grouped together and perceived as a single object.

## Parameters of the model:

<b>Temporal modulation:</b>	<b>20 Hz</b>	(tuned to chord repetition period of 50 ms)
<b>Spectral resolution:</b>	<b>8 cyc/oct.</b>	(corresponding to BW in streaming)

# Modelling expt 5 ('isolated')

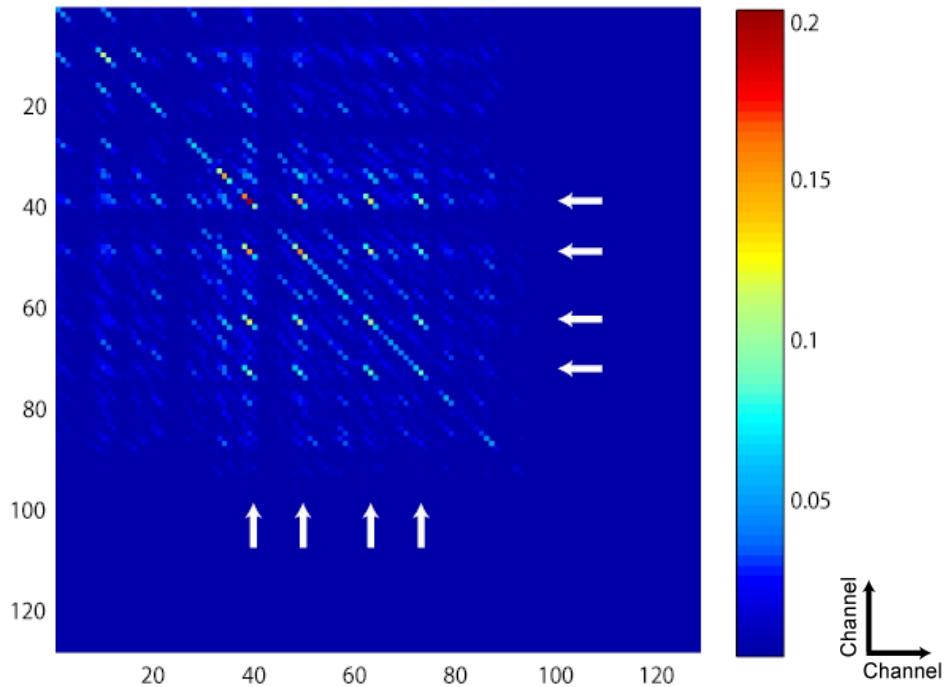
**I. Input:** Different examples of figure and ground stimuli for each (coh, dur) x 1000



# Temporal coherence model



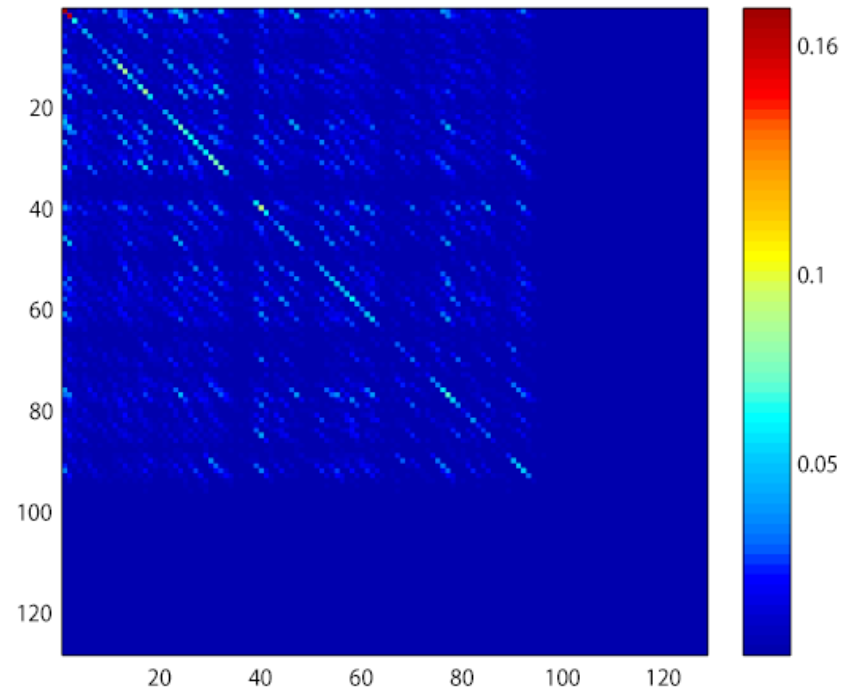
**Coherence matrix**  
(figure present)

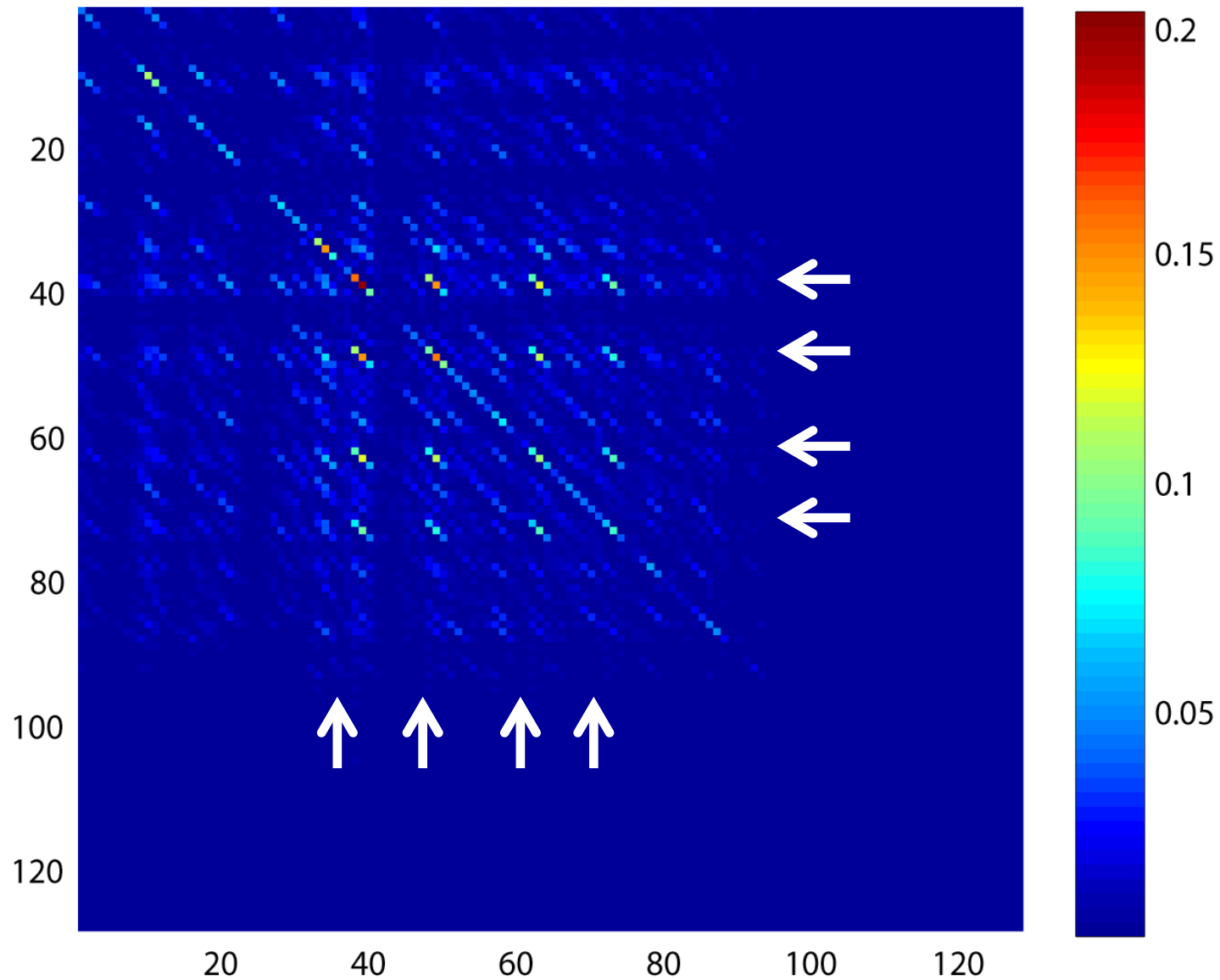


▪ **Measure:** Maximum cross-correlation value for each stimulus



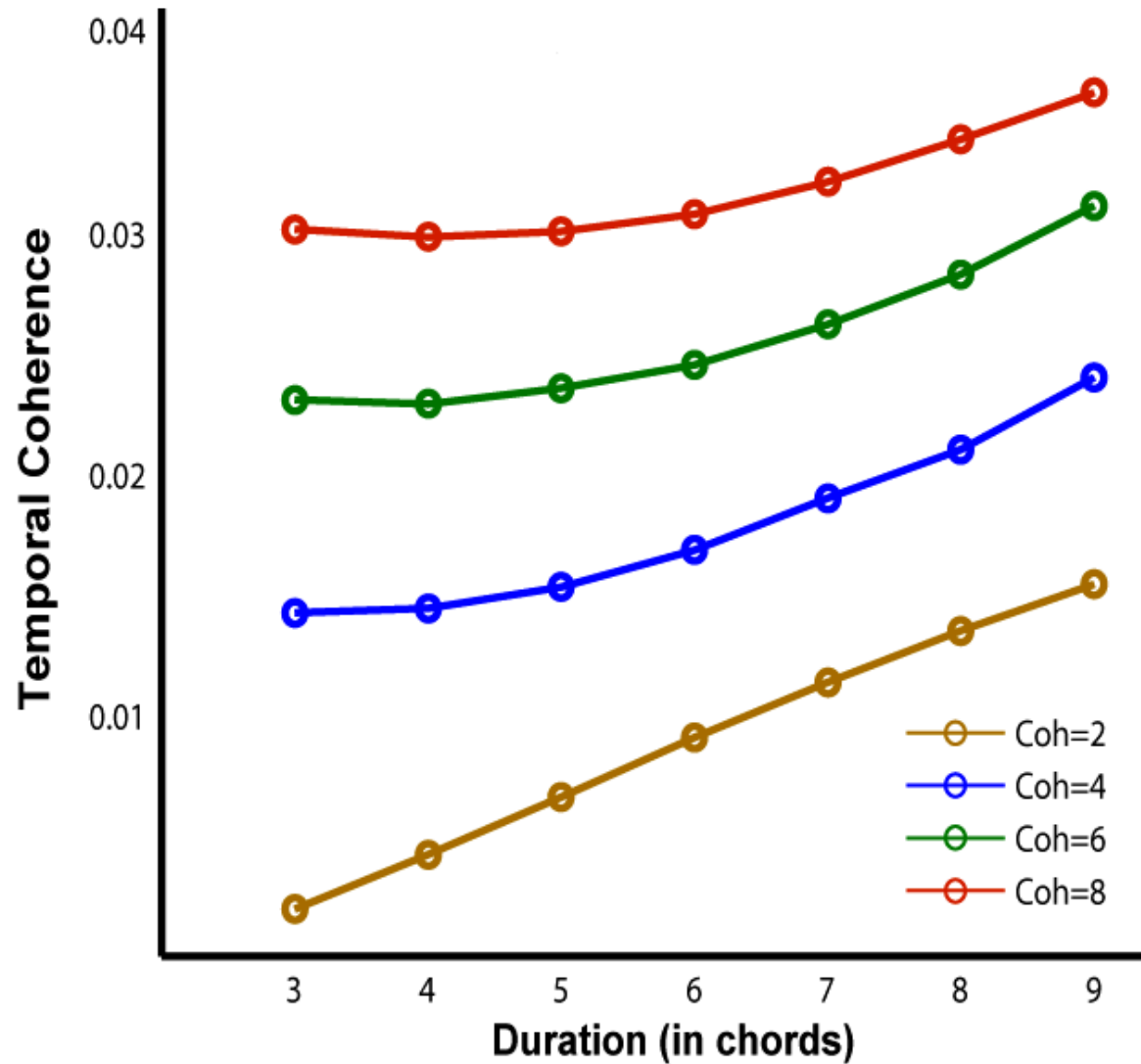
**Coherence matrix**  
(figure absent)





Channels containing repeating figure components show strong cross-correlation (temp. coherence) and this may contribute to the pop-out of the figure that is composed of these channels.

▪ **Output:**    **Average cross-correlation<sub>(figure)</sub> - Average cross-correlation<sub>(ground)</sub>**



# Modelling summary

- Temporal coherence model can explain figure-detection in complex SFG stimulus (works for each psychophysics experiment).
- Model performs better than humans at very short durations of the figure.

# Summary

## **SFG stimulus:**

Represents a complex acoustic scene and allows parametric stimulus control

## **Psychophysics:**

- Listeners can segregate figure from ongoing background very well
- Adaptation does not prove to be critical for segregation in SFG stimuli

## **Temporal coherence model:**

Auditory segregation in complex acoustic scenes may be based on computation of cross-channel temporal coherence.



# Acknowledgments



**Deborah, Sukhbinder and Tim**  
**Newcastle Auditory Group**



**Maria, Aiysha and Nicolas**  
**UCL Ear Institute**

**Shihab Shamma**  
**University of Maryland, College Park**

