Sundeep Teki [1,2]
Deborah Williams [1]
Sukhbinder Kumar [1,2]
Maria Chait [3]
Aiysha Siddiq [3]
Tim Griffiths [1,2]

UCL

Auditory figure-ground segregation using a complex stochastic stimulus

[1] Newcastle Auditory Group, Newcastle University Medical School, UK
[2] Wellcome Trust Centre for Neuroimaging, University College London, UK
[3] UCL Ear Institute, University College London, UK

Outline

I. Stimulus

II. fMRI experiment (done)

III. Psychophysics (in progress)

IV. Discussion

Auditory figure-ground segregation

Listeners' ability to extract a particular sound from a background of other simultaneous sounds

Processes:

i. Grouping of simultaneous figure components from the spectral array
ii. Grouping of figure components over time
iii. Separation of grouped components from rest of the acoustic scene.

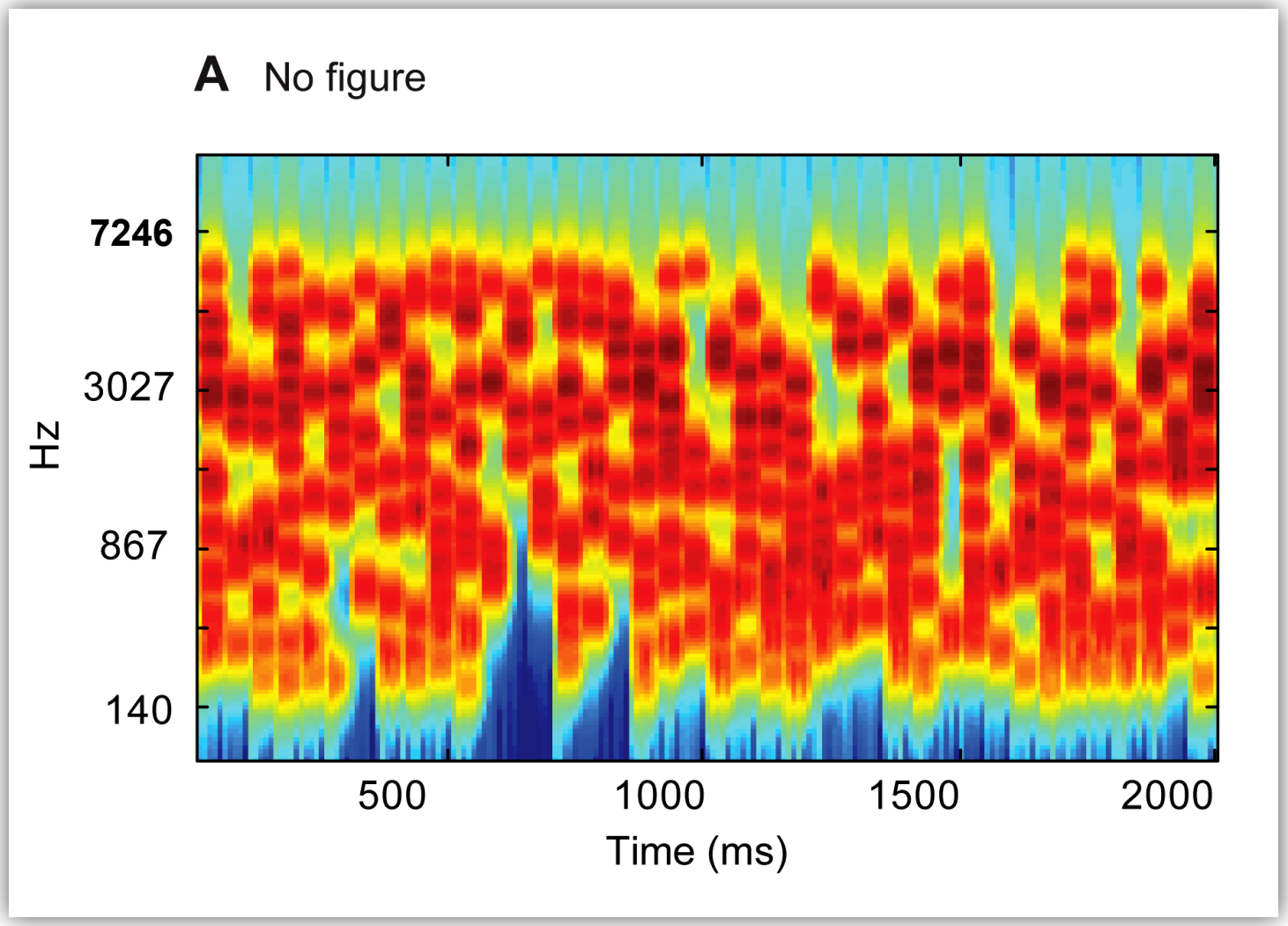
Neural Substrates:

Distributed network: auditory periphery, medial geniculate body, primary auditory cortex to non-primary auditory areas

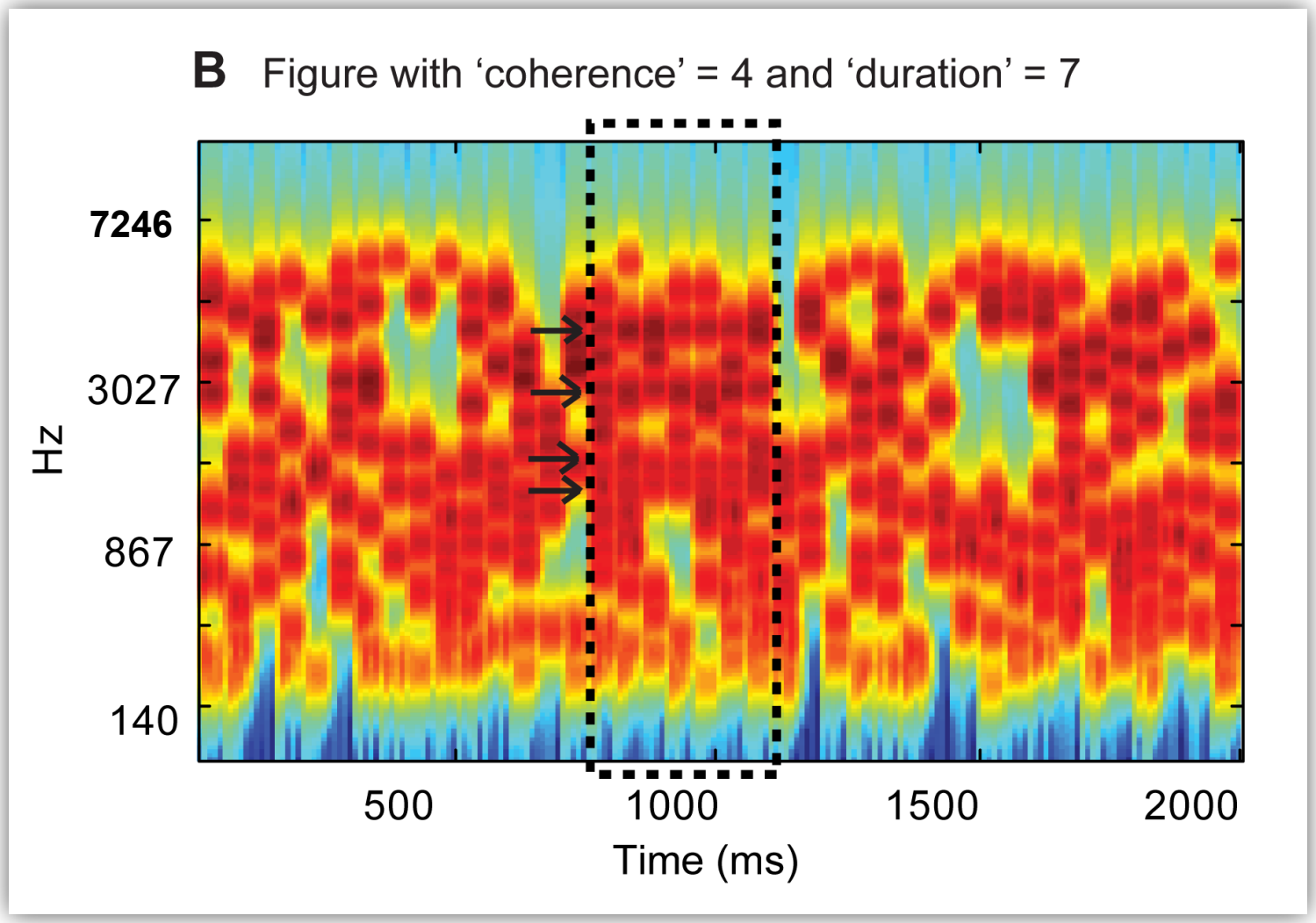
Stimuli:

Streaming stimuli:
- lack the rich spectrotemporal complexity of natural sounds.

Stochastic Figure-Ground (SFG) Stimulus: Background signal



SFG: Signal with 'figure' present



Long figure

SFG: Figure

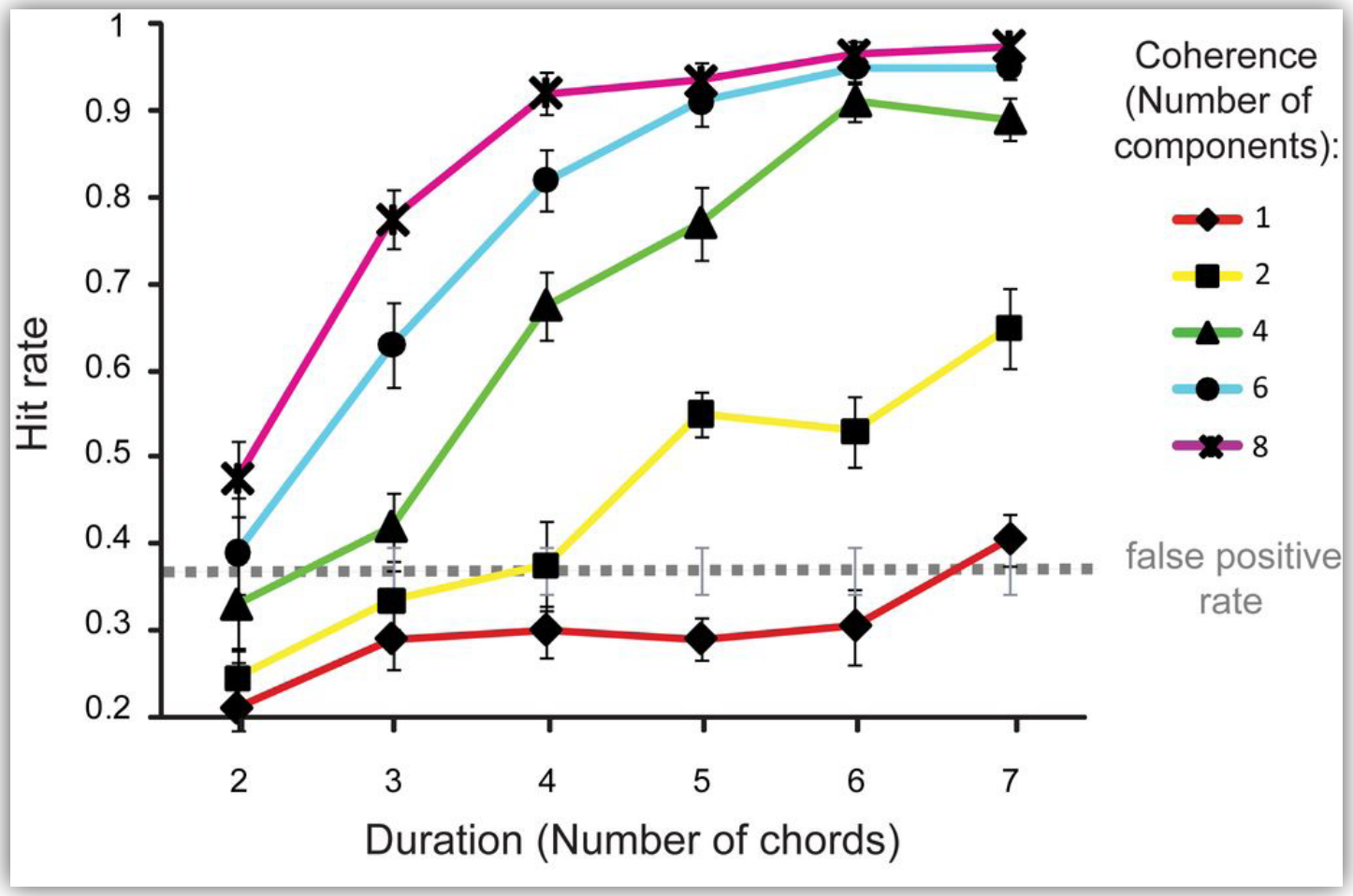
Coherence: Number of different repeating frequencies : **1,2,4,6,8**

Duration: Number of chords over which frequencies repeat : **2-7**

Features of SFG

- Figure and background signals do not differ in low-level acoustic attributes
- No spectral ‘protective’ region between figure and background
- Figure and background signals are indistinguishable at each point in time
- Figure can only be extracted by integrating over time and frequency
- Enables parametric variation of figure salience

Behaviour (n = 10)



- Listeners are remarkably sensitive to the appearance of figures
- Sensitive to parametric variations of coherence and duration

II. fMRI experiment

fMRI Experiment

Aim: Identify brain areas whose activity varies with parametric variations in coherence and duration of the figure

Stimulus:
i. Fixed coherence: 4, varying duration: 2-7 chords
ii. Fixed duration: 4, varying coherence: 1,2,4,6,8

= 9 stimulus conditions (40 repetitions each)

Paradigm:
i. Passive listening
ii. Active figure-detection

fMRI Paradigm

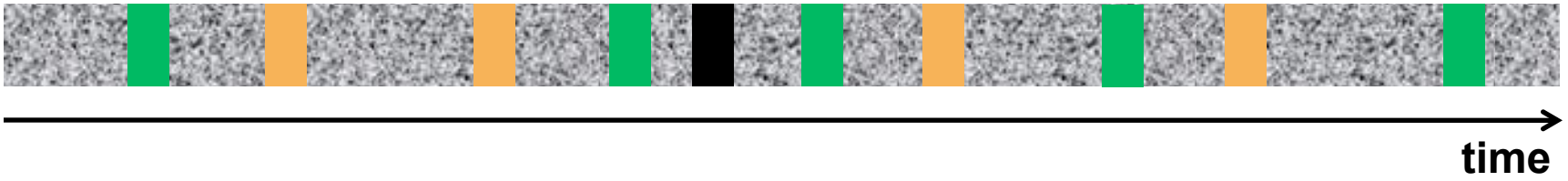


Figure (fixed coherence)

Figure (fixed duration)

Background

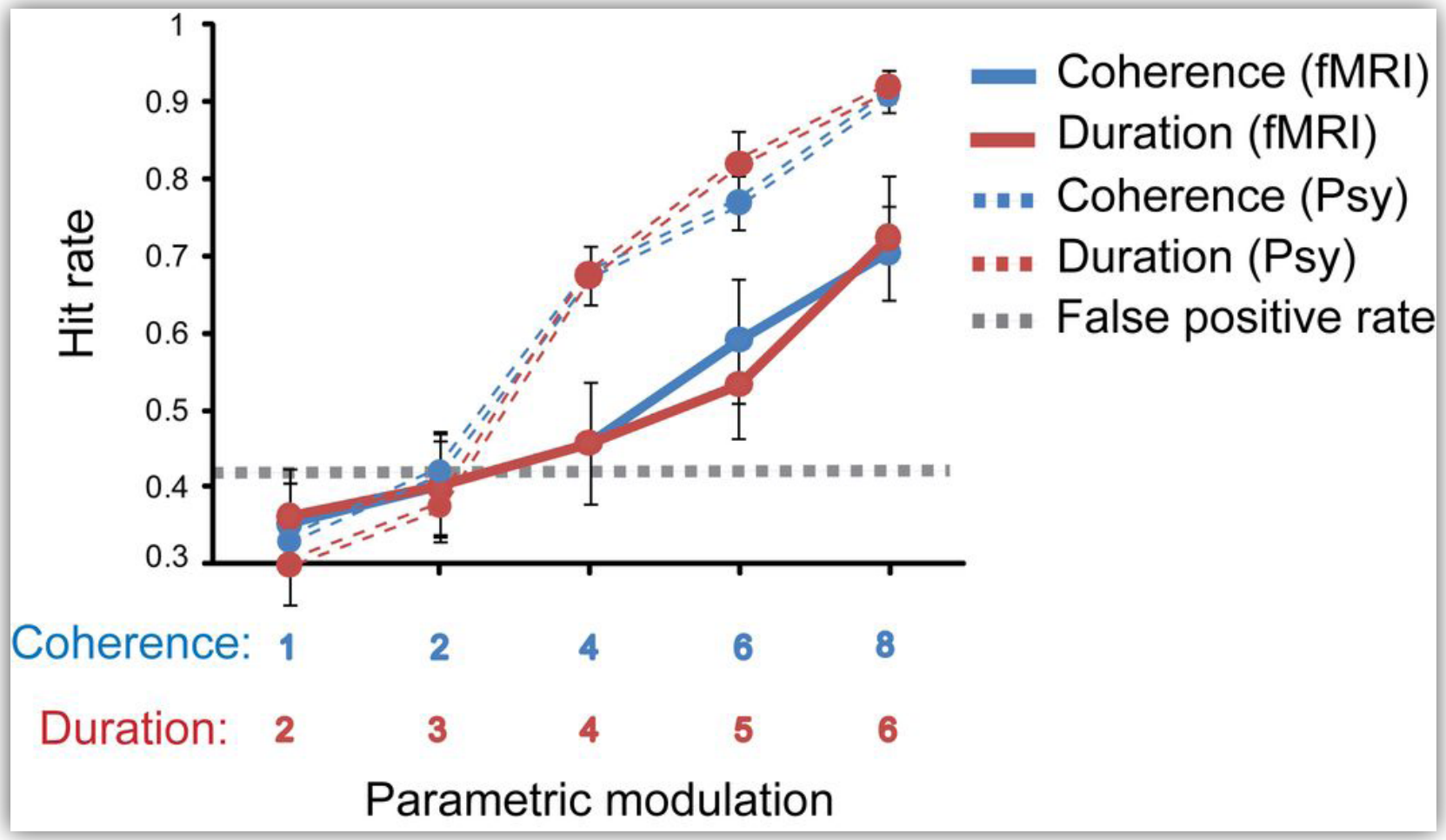
Decoy

Task: **Detect decoy stimuli (noise bursts; 10% of stimuli)**

- **Subjects not actively detecting figures**

- 3 Tesla Siemens Allegra MRI Scanner
- Continuous scanning
- 42 contiguous slices per volume
- TR: 2.52 s; TA: 2.88 s; TE: 30 ms
- Slice thickness: 2 mm with 1mm gap between slices
- In-plane resolution: 3.0 x 3.0 mm^2
- 3 scanning sessions: 510 volumes per subject

Behaviour in scanner



fMRI Analysis

- 14 subjects (normal hearing, no audiological disorders)
- Standard pre-processing with SPM8
- Whole brain analysis
- Statistical model based on General Linear Model
- Random effects design

Parametric Modulation

I. Effect of Duration: Fixed coherence (4); varying duration (2-7)

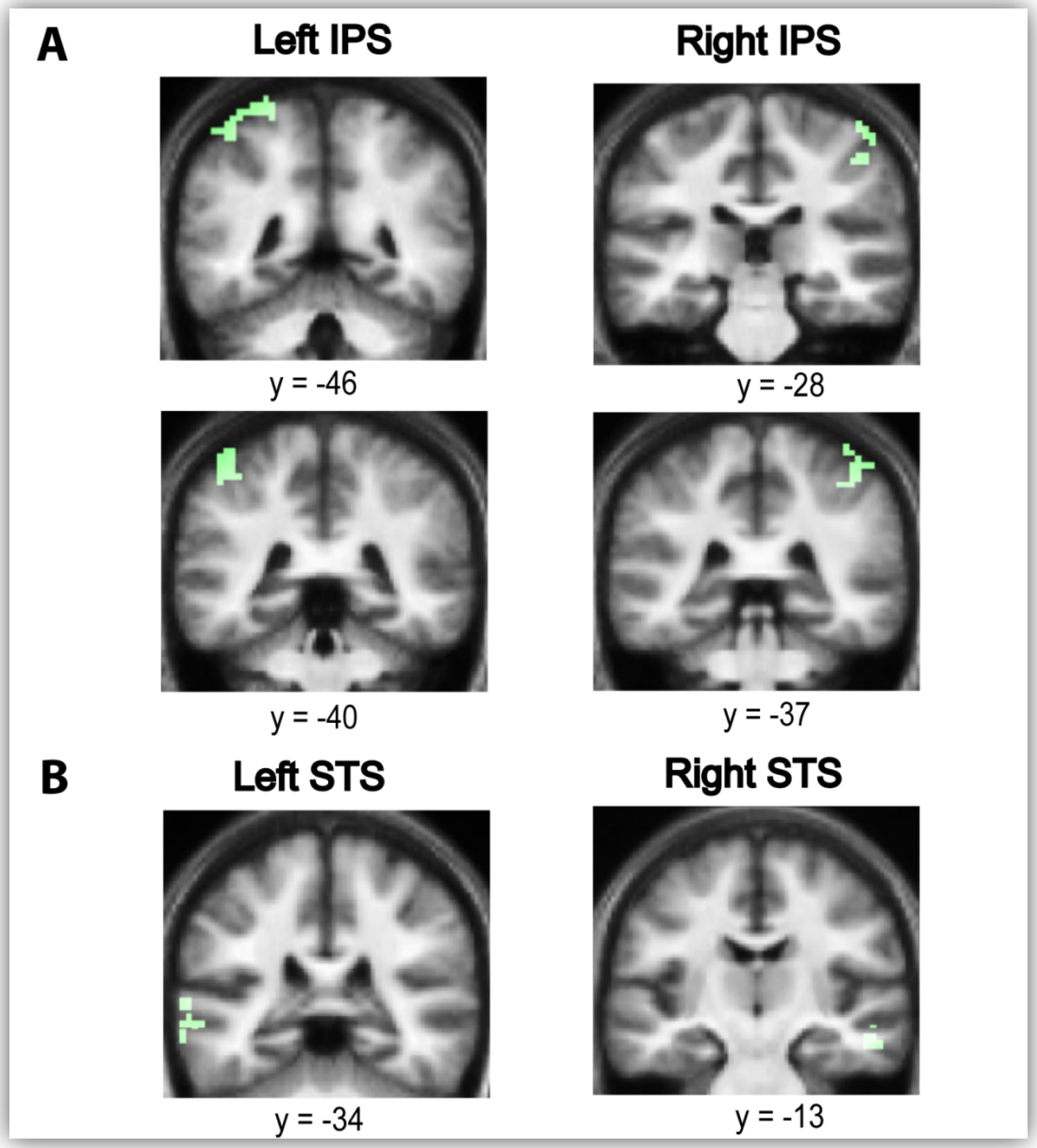
II. Effect of Coherence: Fixed duration (4); varying coherence (1,2,4,6,8)

fMRI Results

I. Effects of Duration:

Intraparietal Sulcus (IPS)	(bilateral; anterior)
Superior Temporal Sulcus (STS)	(bilateral)
Planum Temporale	(R)
Medial Geniculate Body (MGB)	(bilateral)

Effects of Duration

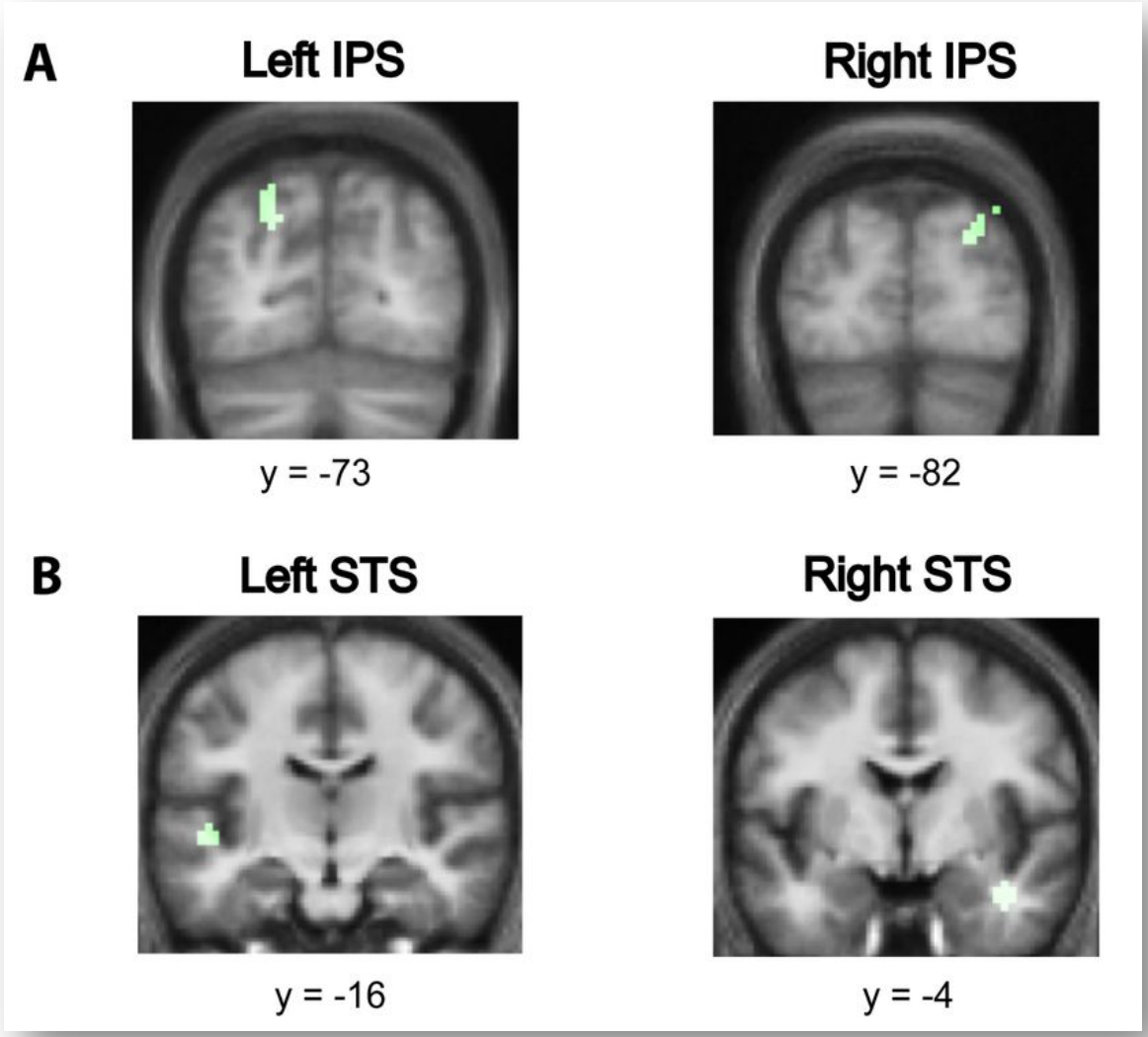


fMRI Results

II. Effects of Coherence:

Intraparietal Sulcus	(bilateral; posterior)
Superior Temporal Sulcus	(bilateral)

Effects of Coherence



What about the auditory cortex ?

- No activation in Primary Auditory Cortex (PAC) for either contrast
- Confirmed using volume of interest analysis based on PAC maps (Morosan et al., 01)
- Consistent with one previous fMRI study (Cusack, 2005)

Reasons…

- More complex and naturalistic stimulus
- Naïve subjects and short figures
- PAC recruited during active figure-ground segregation (i.e., in behavioural context) with possibly top-down modulation by IPS?

Role of STS

- STS activity modulated by changing duration and coherence of the figure

- Implicated in:
  - Analysis of spectral shape *(Warren et al., 2005)*
  - Dynamic changes in spectrum *(Overath et al., 2008)*
  - Detection of increasing changes in spectrotemporal coherence within textures *(Overath et al., 2010)*

IPS and Perceptual Organization

Role of IPS consistent with *Cusack (2005)*

- Implicated IPS in perception of two streams vs. one stream, based on the same physical streaming signal that evoked a bistable percept.
- IPS activity likely reflects top-down application of attention (shift between streams)
- Found no activation in primary auditory cortex

IPS is involved in structuring sensory input and perceptual organization

- Encoding visual object representations
- Binding of sensory features within and across different modalities
- control and shift of auditory attention

What does the IPS activity reflect?

➢ ***automatic, bottom-up segregation of auditory object from stochastic background***

Teki, Chait et al., *J Neurosci (2011)*

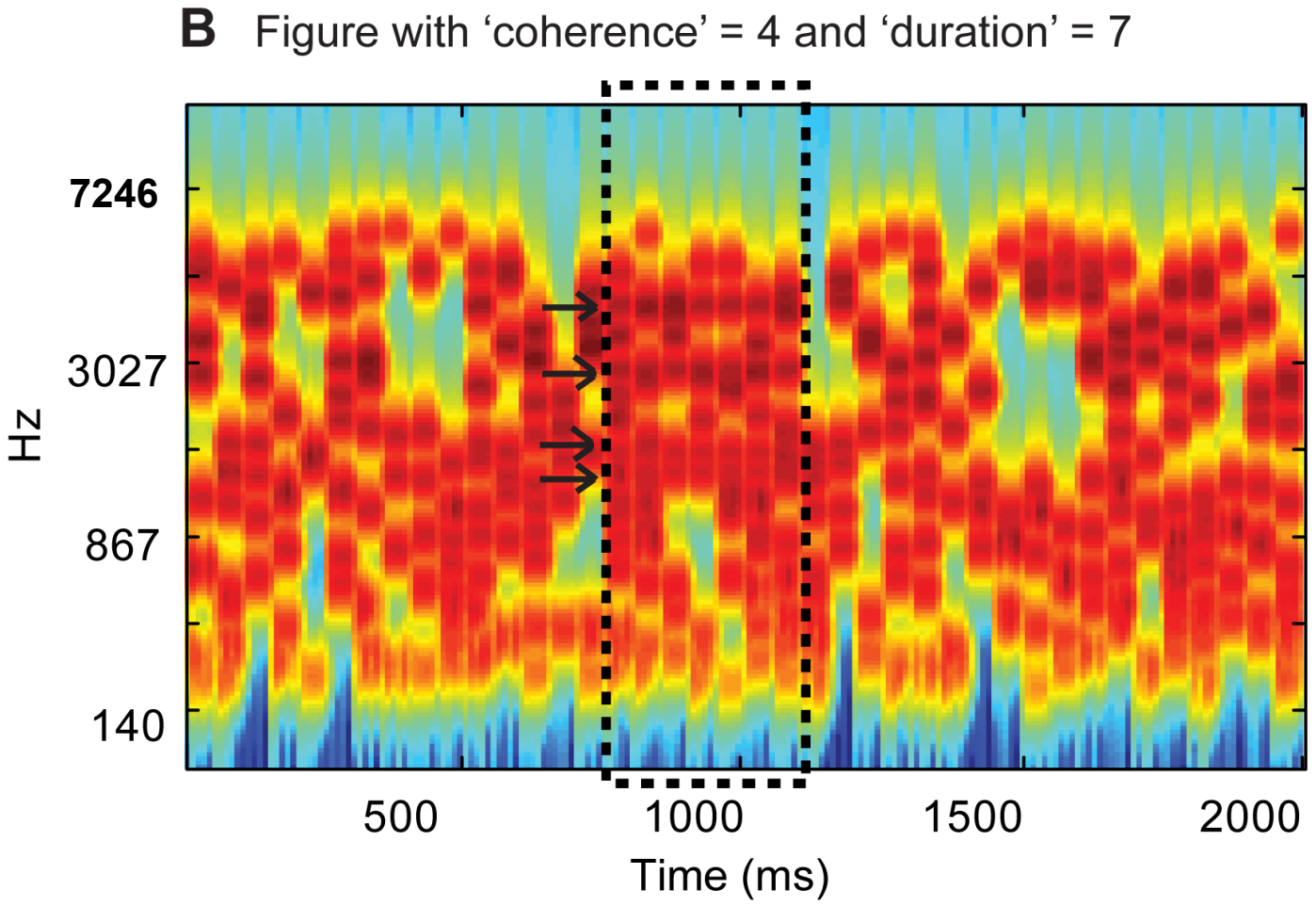
III. Psychophysics

Motivation

- Investigate neural mechanisms underlying figure-detection in the SFG stimulus
- Initial psychophysics: different stimulus conditions presented together in a block
- Present trials with specific coherence/duration in a single block and obtain **d’**

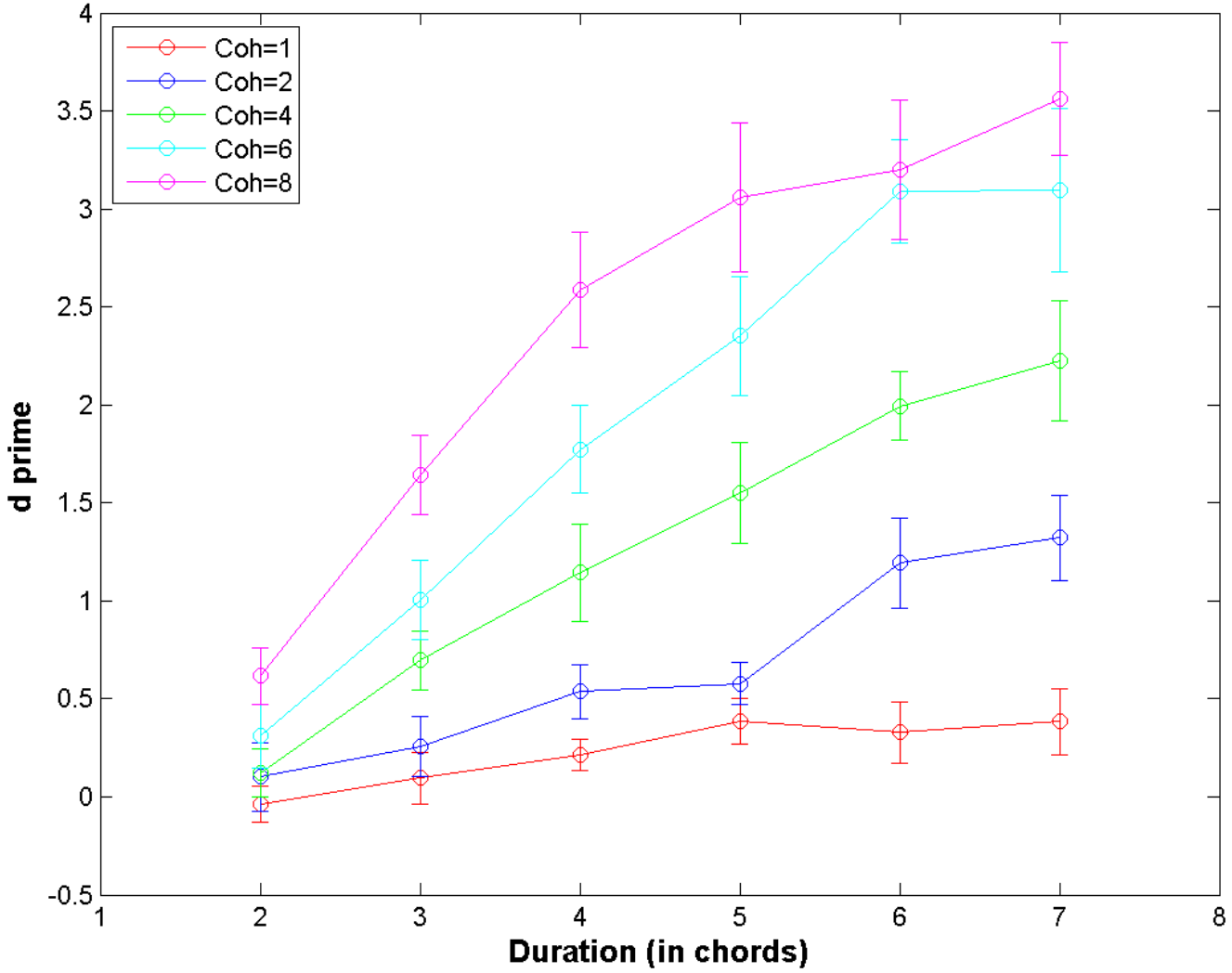
Experiment 1

Stimulus: SFG with 50ms chords



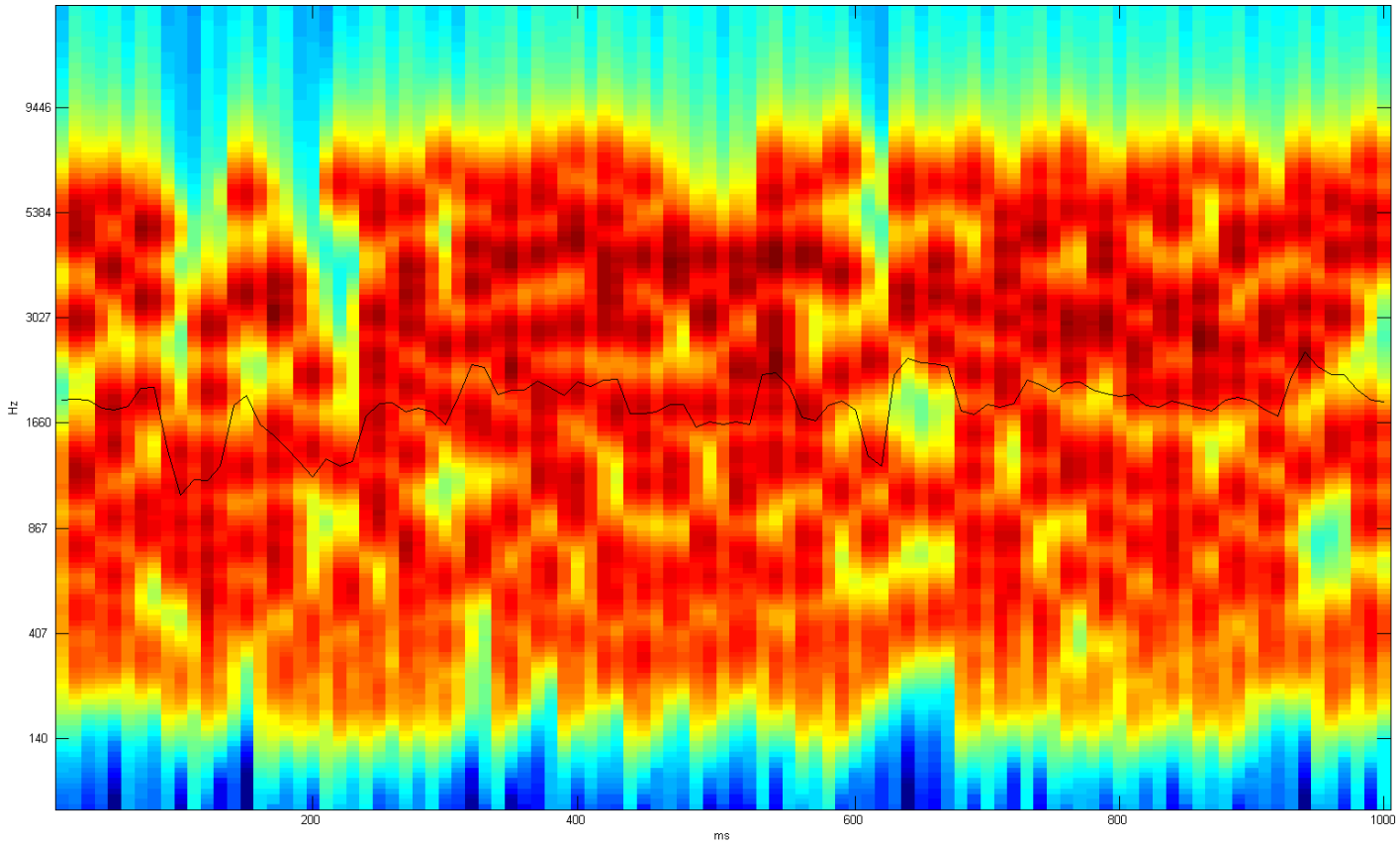
Coherence: [1 2 4 6 8] Duration: [2:7]

Experiment 1 (n=10)



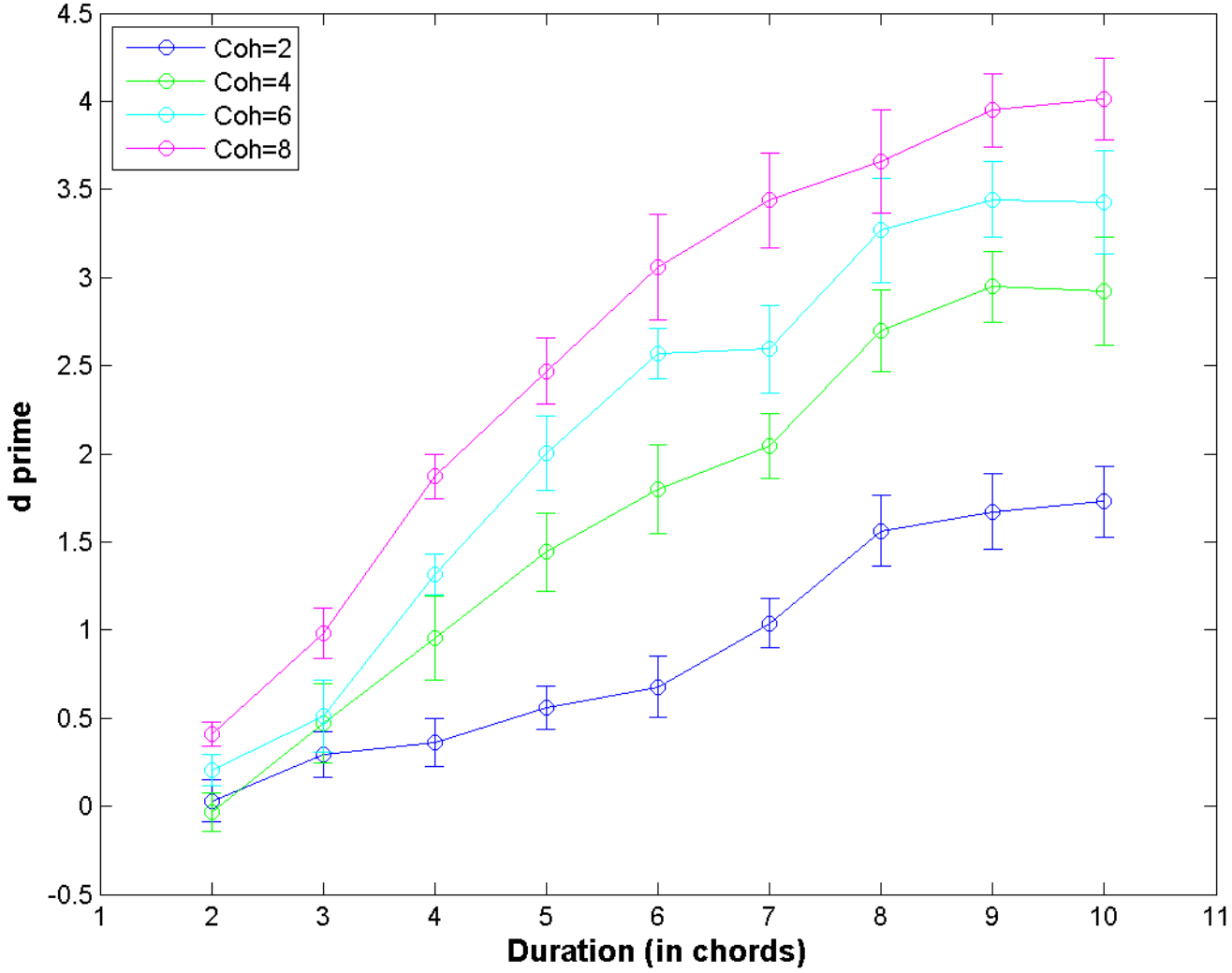
Experiment 2

Stimulus: SFG with 25ms chords; 1 s long stimulus



Coherence: [2 4 6 8] Duration: [2:10]

Experiment 2 (n=10)



Expt. 1 vs. 2

ANOVA

- Coherence and duration as within-subject factors

- Stimulus length (50ms, 25ms) as between-subject factor.

Results

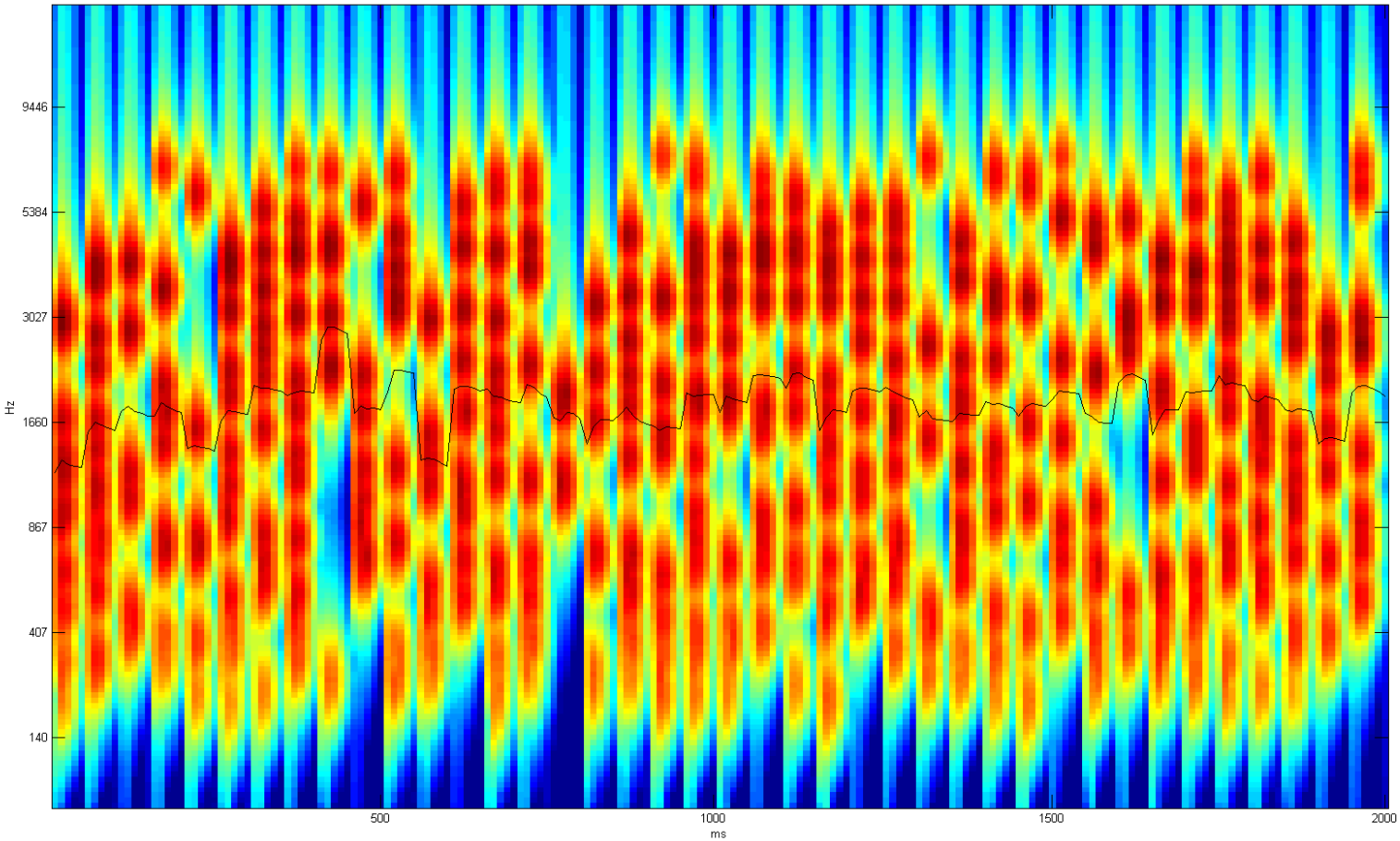
SIG. Effect of Coherence: $F(3,54) = 125$; $p < 0.001$

SIG. Effect of Duration: $F(5,90) = 137$; $p < 0.001$

No significant effect of stimulus length: $F(1,18) = 2.866$; $p = 0.108$

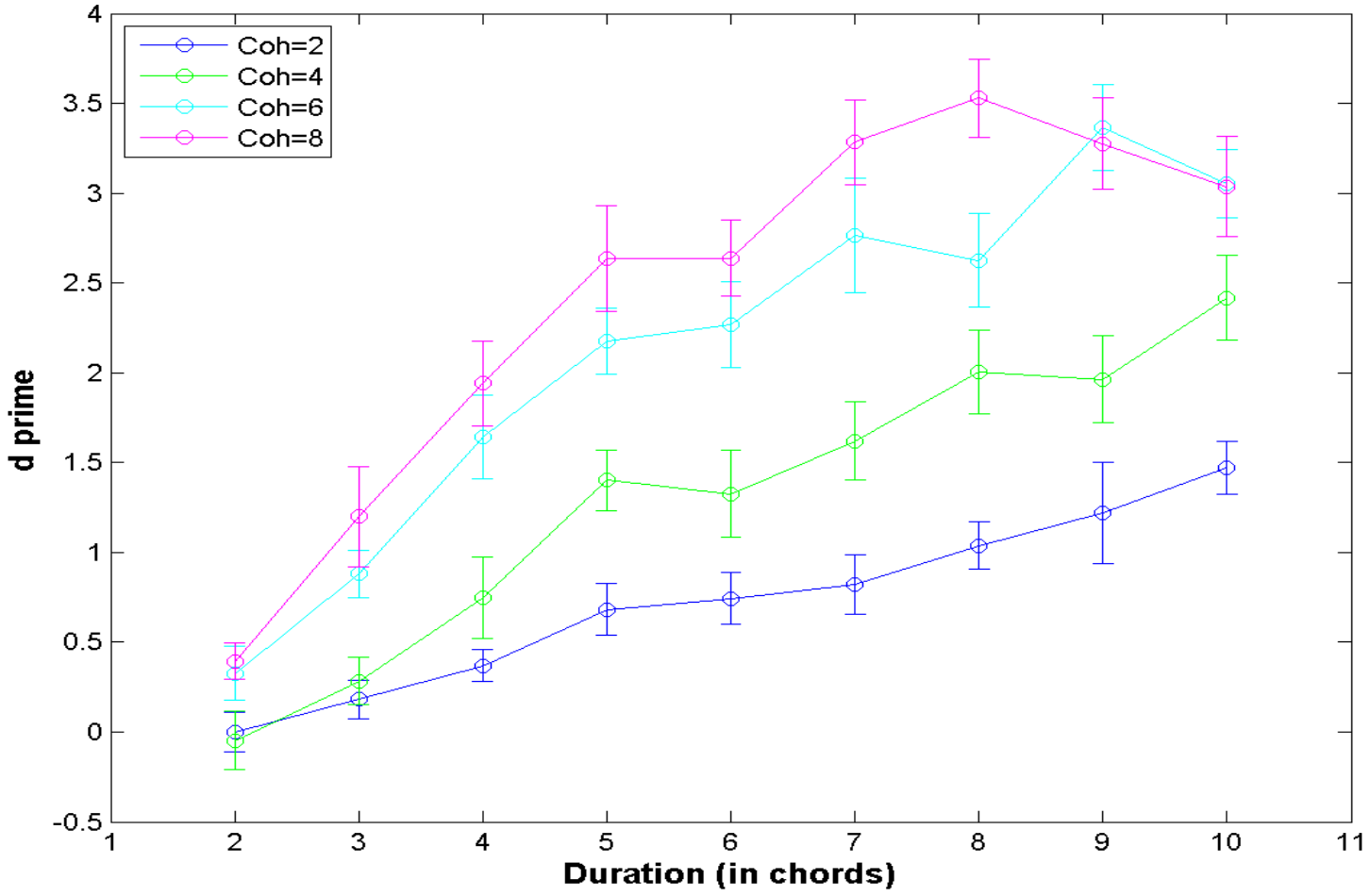
Experiment 3 (n=10)

Stimulus: SFG with 25ms chords with 25 ms silence;
2 s long stimulus



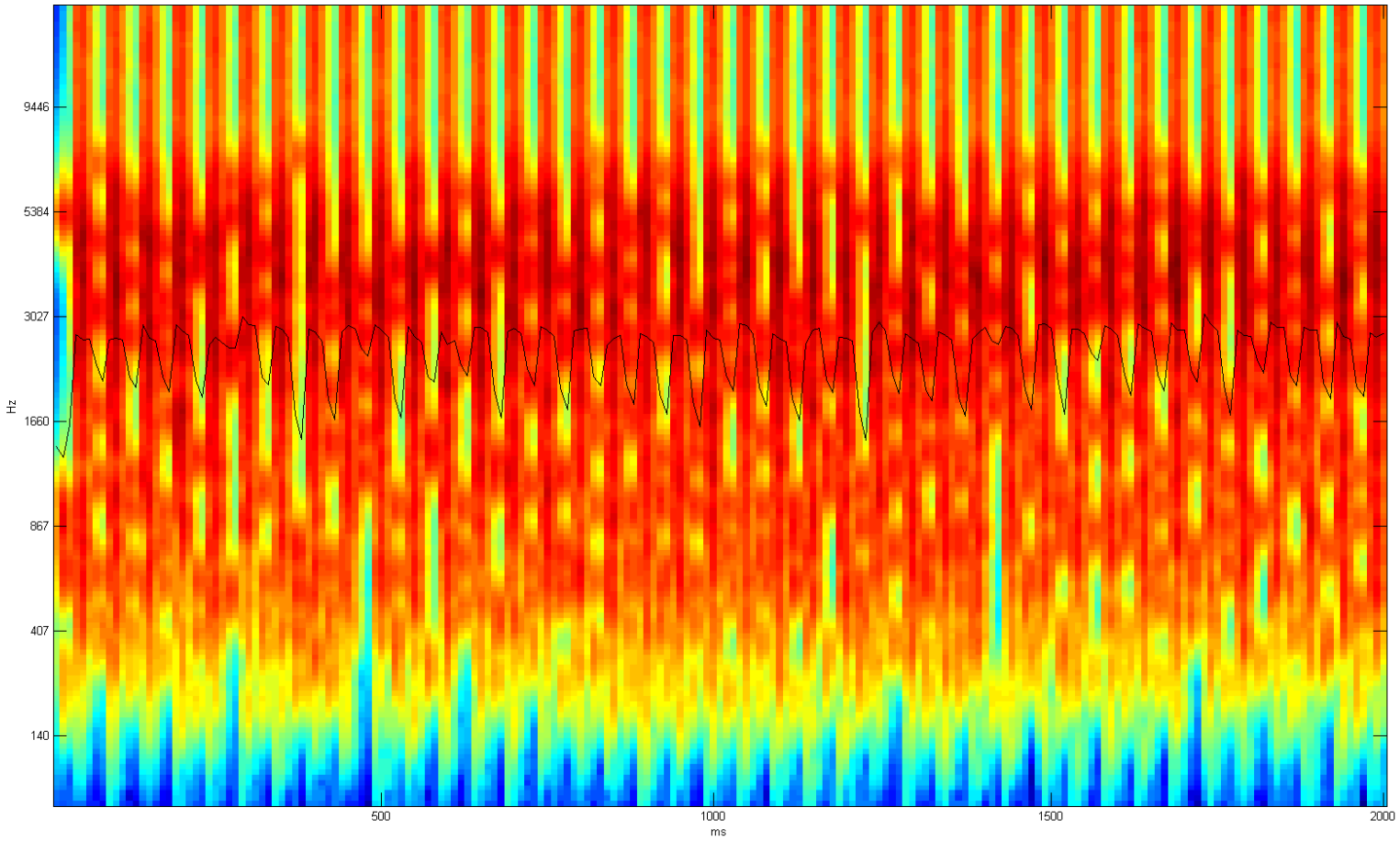
Coherence: [2 4 6 8] Duration: [2:10]

Experiment 3



Experiment 4 (n=9)

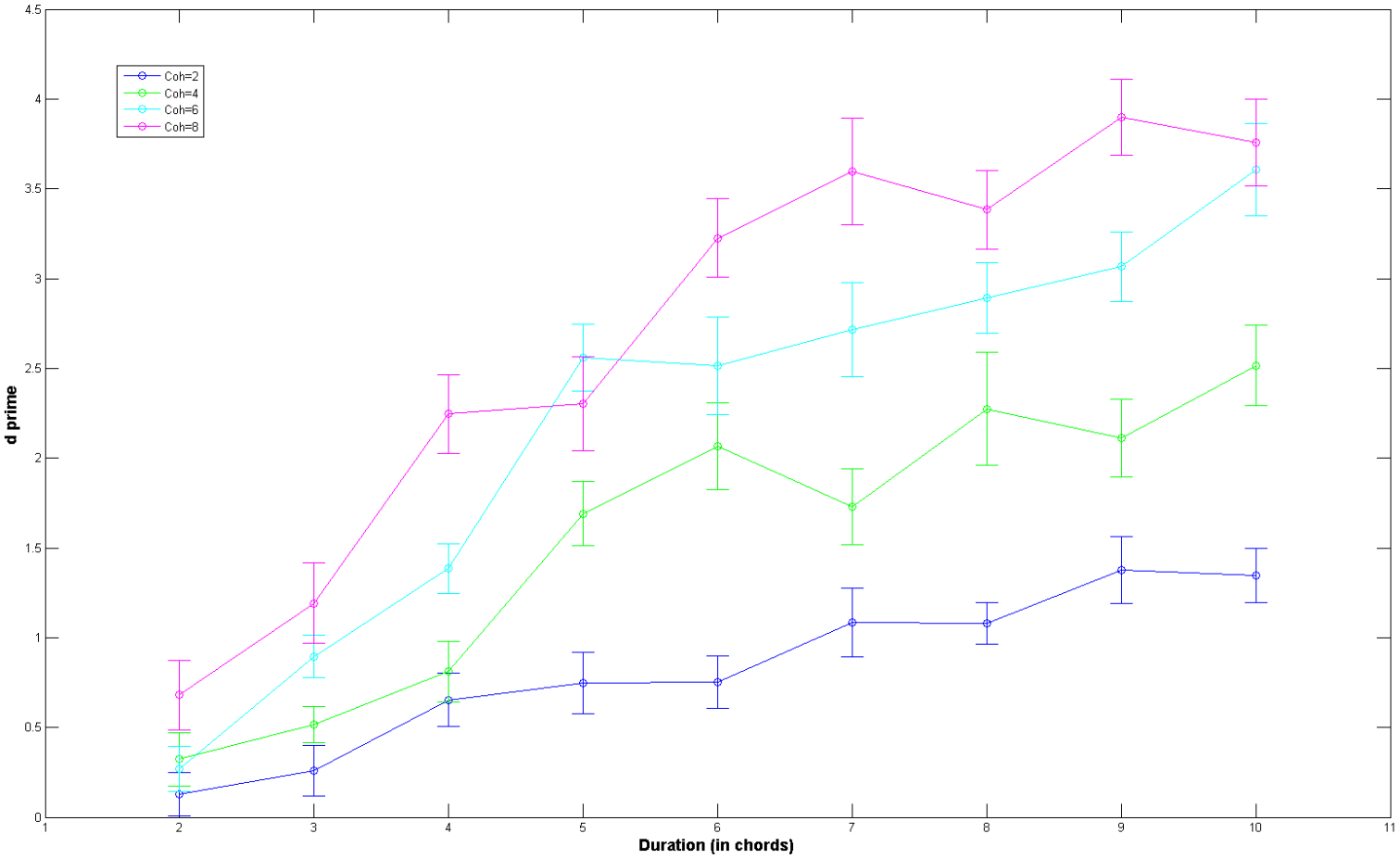
Stimulus: SFG with 25ms chords with 25 ms white noise;
2 s long stimulus



Coherence: [2 4 6 8] Duration: [2:10]

Experiment 4

**Stimulus: SFG with 25ms chords with 25 ms white noise;
2 s long stimulus**



Expt. 2 vs. 3 vs. 4

ANOVA

- Coherence and duration as within-subject factors

- Condition (no-gap, silence, noise) as between-subject factor.

Results

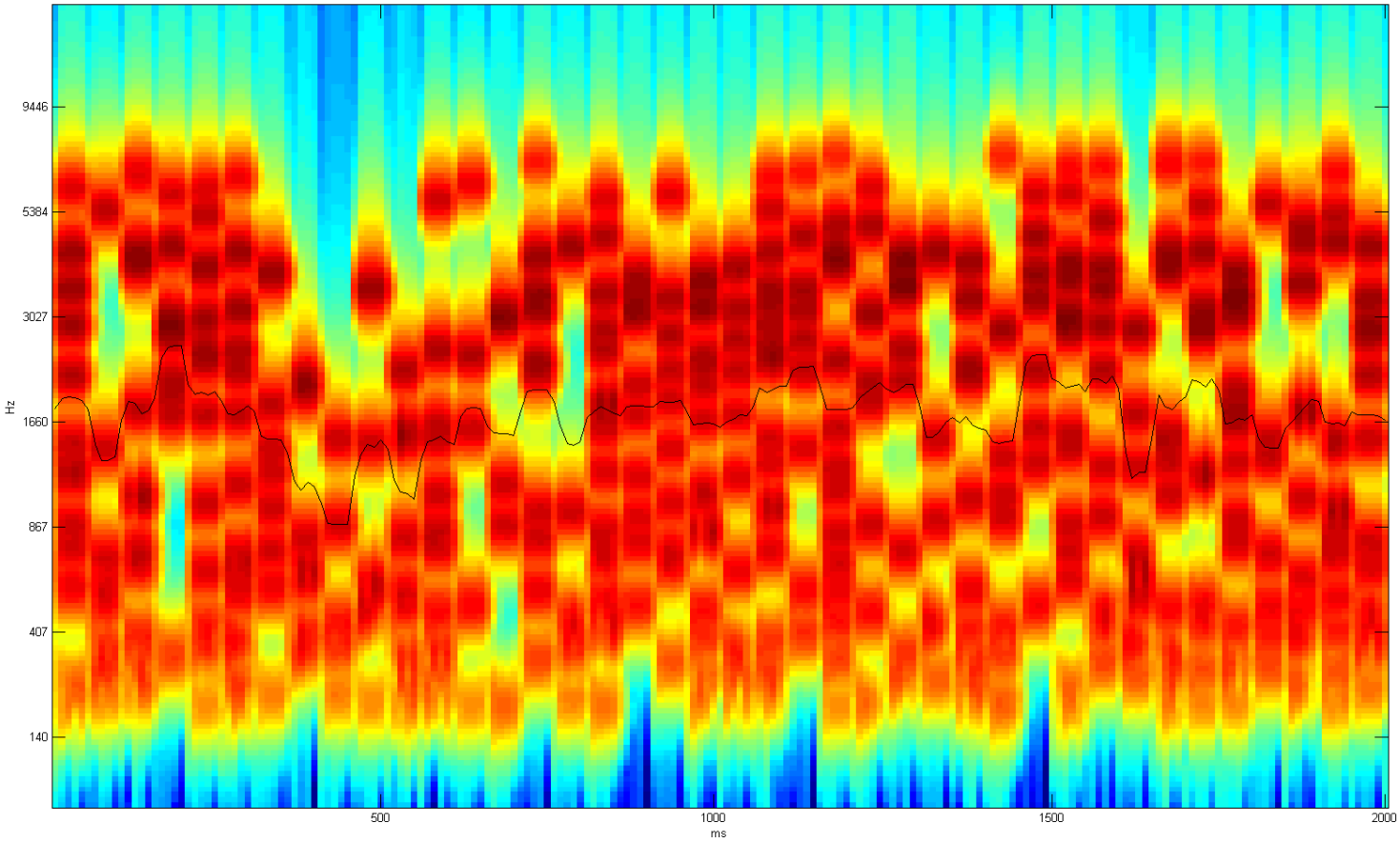
SIG. Effect of coherence: $F(3, 78) = 349$, $p < 0.001$

SIG. Effect of duration: $F(8, 208) = 241$, $p < 0.001$

No significant effect of condition: $F(2, 26) = 1.15$, $p = 0.332$

Experiment 5

Stimulus: SFG with 50ms chords with figure: positive frequency ramps (ramps within critical band)

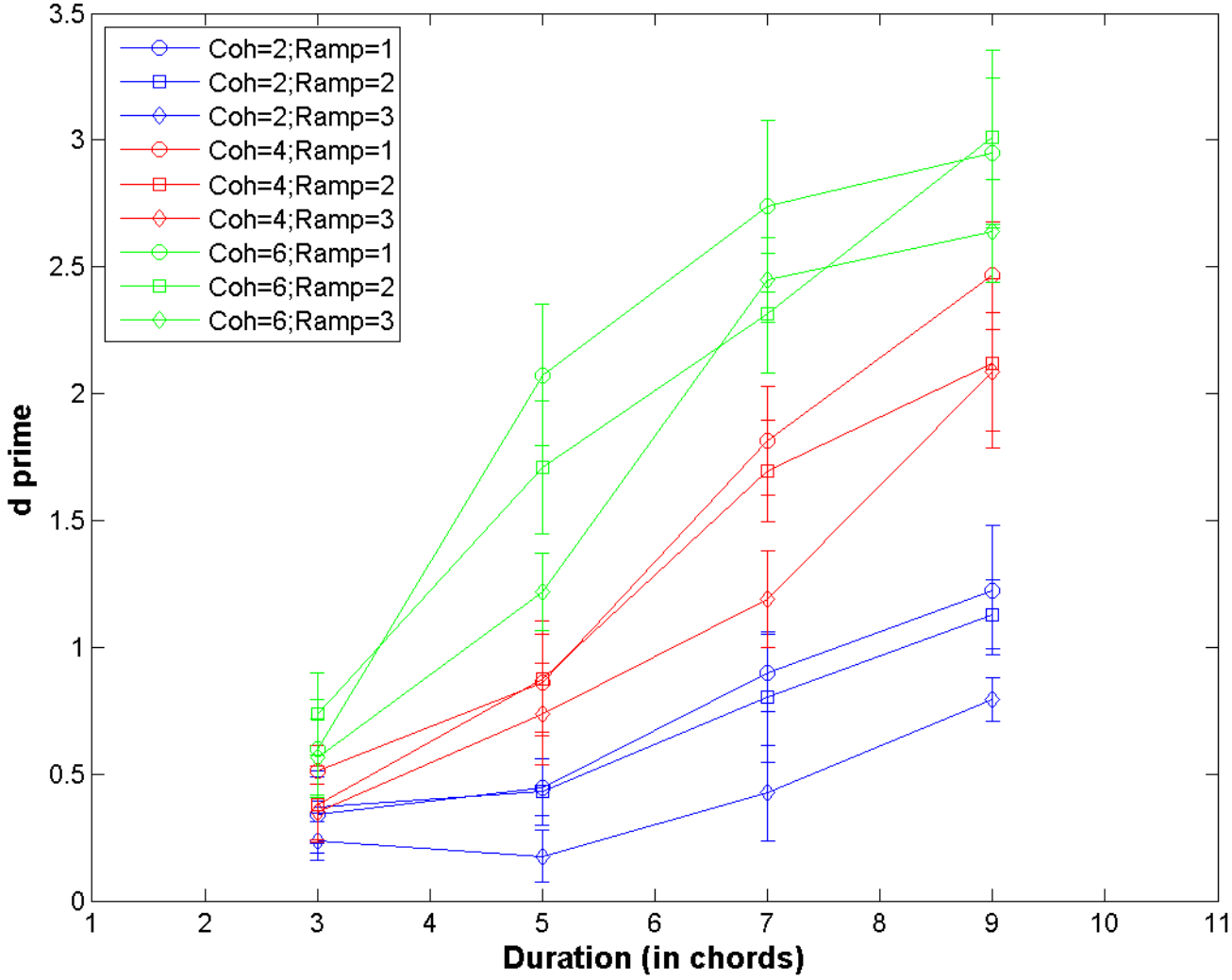


Coherence: [2 4 6]

Duration: [3 5 7 9]

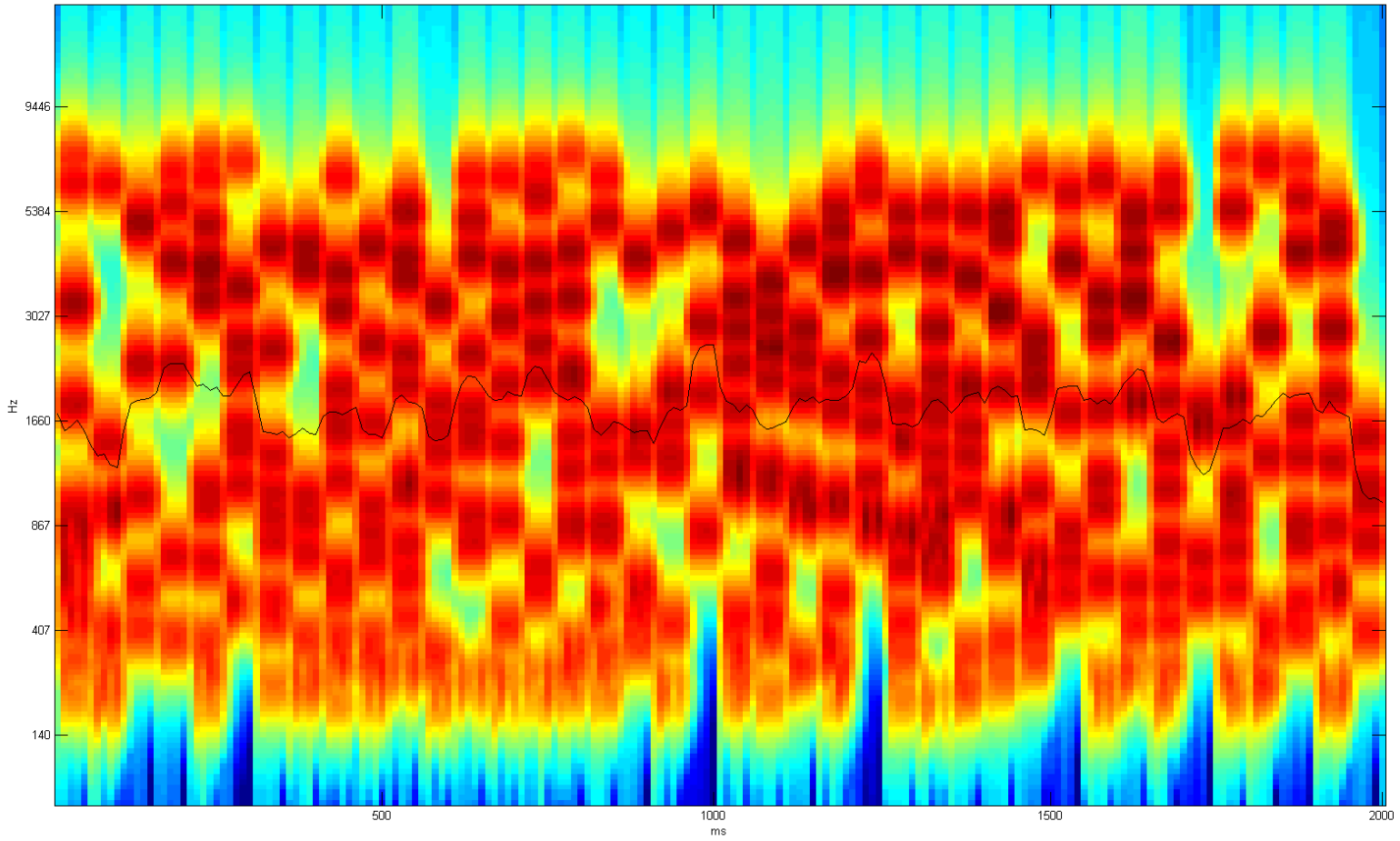
Ramp steps: [1 2 3]

Experiment 5 (n=10)



Experiment 6

Stimulus: SFG with 50ms chords with figure: negative frequency ramps (ramps within critical band)

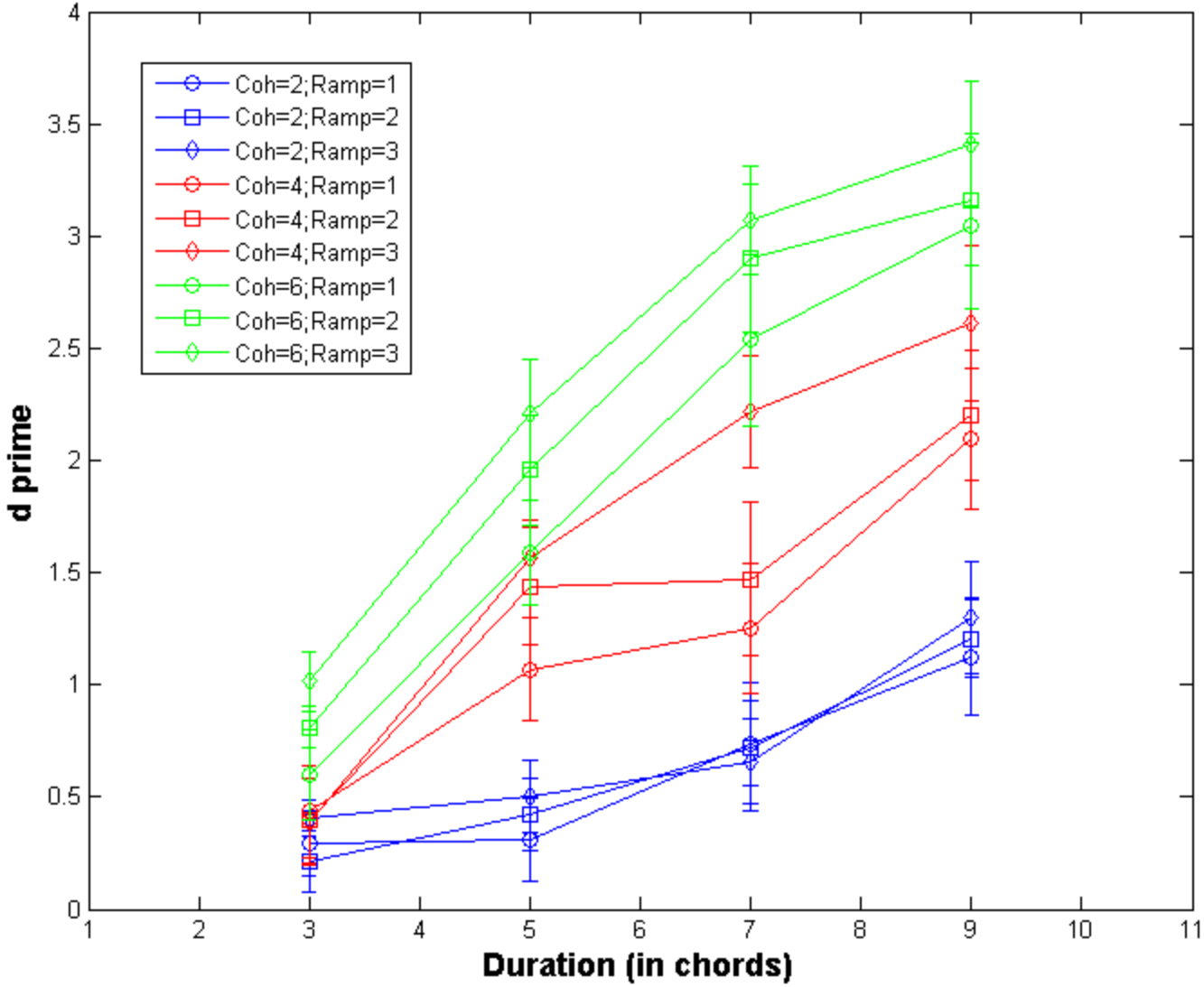


Coherence: [2 4 6]

Duration: [3 5 7 9]

Ramp steps: [-1 -2 -3]

Experiment 6 (n=10)



Expt. 1 vs. 5 vs. 6

ANOVA

- Coherence and duration as within-subject factors

- Condition (no-gap, positive and negative ramps) as between-subject factor.

Results

SIG. Effect of coherence: $F(2, 134) = 98$, $p < 0.001$

SIG. Effect of duration: $F(2, 134) = 31$, $p < 0.001$

No significant effect of condition: $F(2,67) = 2$, $p = 0.140$

Expt. 5 vs. 6

ANOVA

- Coherence, duration and ramp step as within-subject factors

- Condition (positive and negative ramps) as between-subject factor.

Results

SIG. Effect of coherence: $F(2,36) = 376$; $p < 0.001$

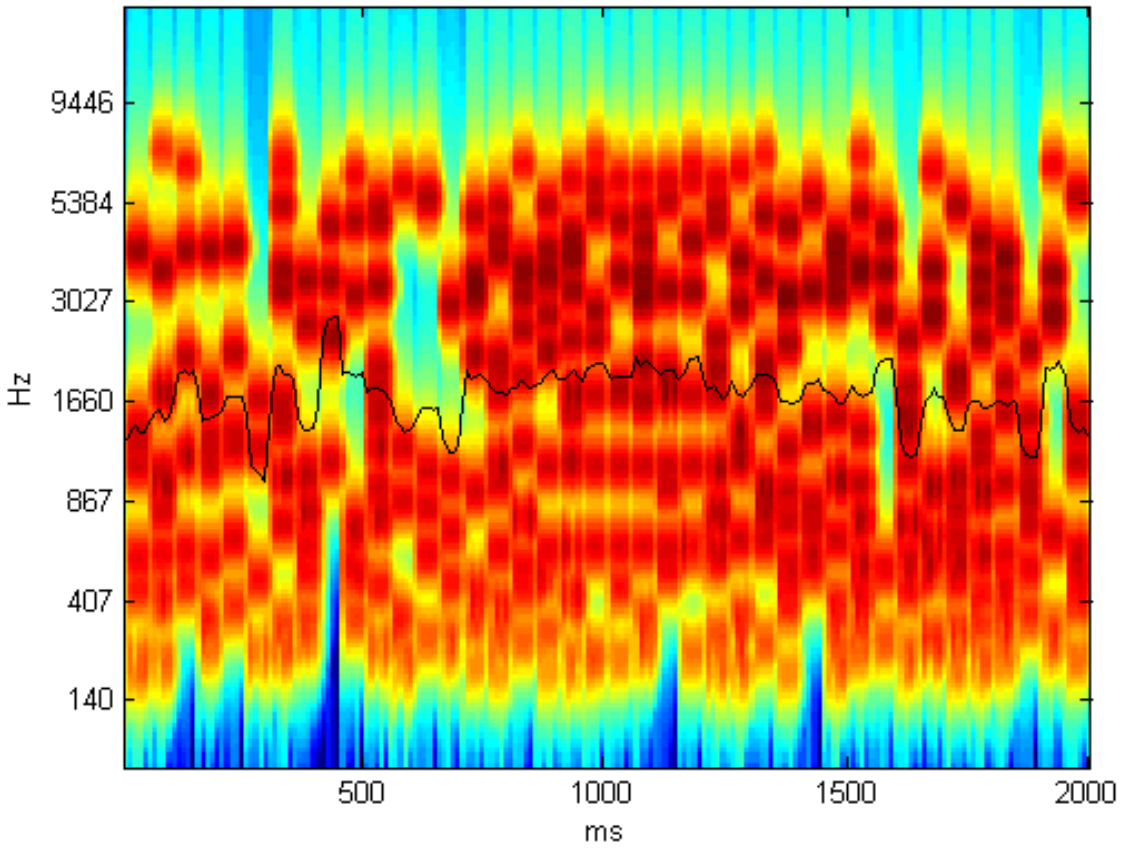
SIG. Effect of duration: $F(3,54) = 142$; $p < 0.001$

No significant effect of ramp rate: $F(2,36) = 0.058$; $p = 0.944$

No significant effect of condition (ramp direction): $F(1,18) = 0.776$; $p = 0.390$

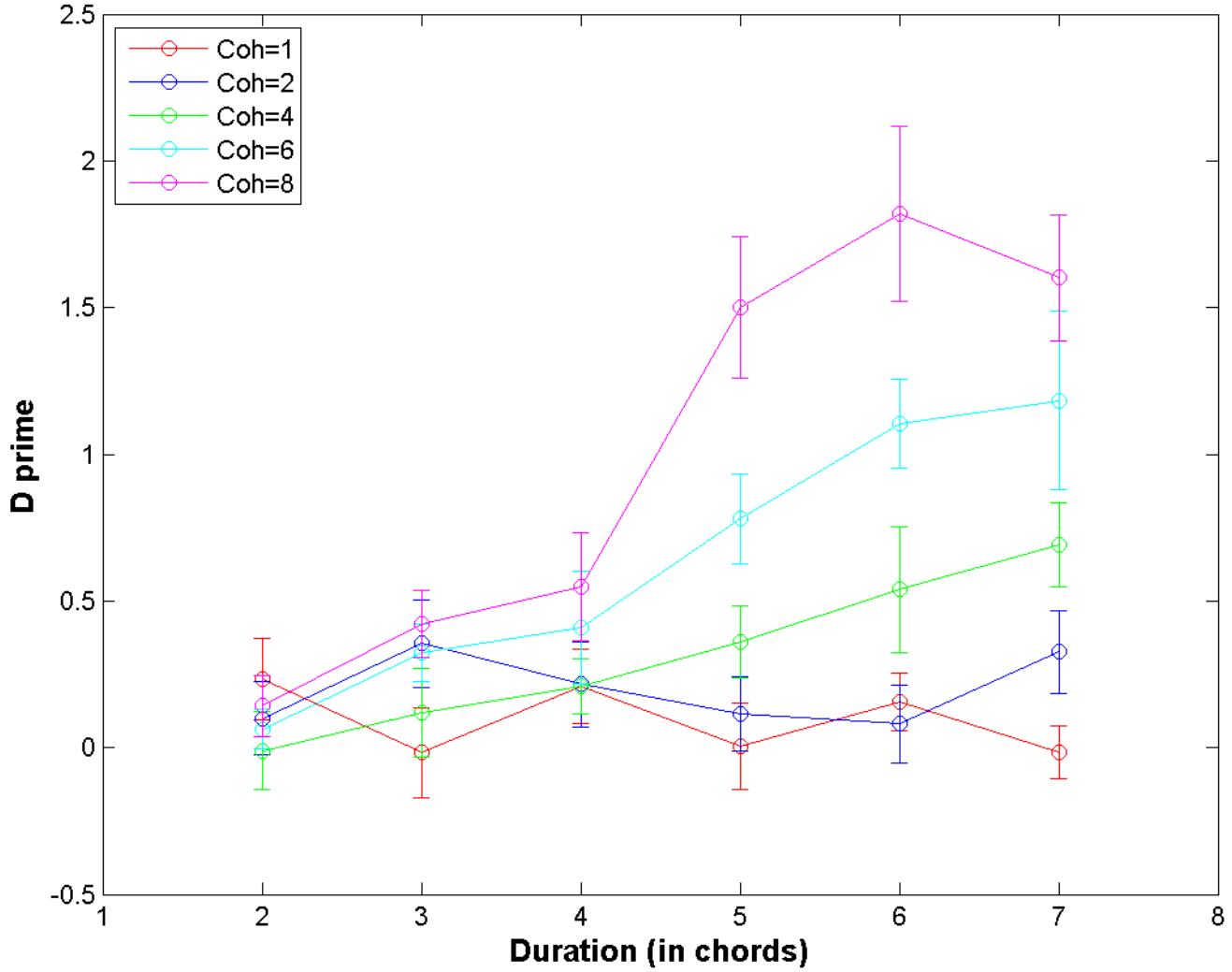
Experiment 7

Stimulus: SFG with 50ms chords with figure; avg. 20 components/chord



Coherence: [1 2 4 6 8] Duration: [2:7]

Experiment 7 (n=10)



Expt. 1 vs. 7

ANOVA

- Coherence and duration as within-subject factors

- Condition (10 vs. 20 components/chord) as between-subject factor.

Results

SIG. Effect of coherence: $F(4, 72) = 104$, $p < 0.001$

SIG. Effect of duration: $F(5, 90) = 63$, $p < 0.001$

Significant effect of condition: $F(1,18) = 36$, $p < 0.001$

Experiment 8

Stimulus: SFG with 50ms chords with figure: pos/neg frequency ramps (ramps within critical band)

Coherence: [2 4 6]

Duration: [3 5 7]

Ramp steps: [± 2 ±5]

Summary

Figure detection performance:

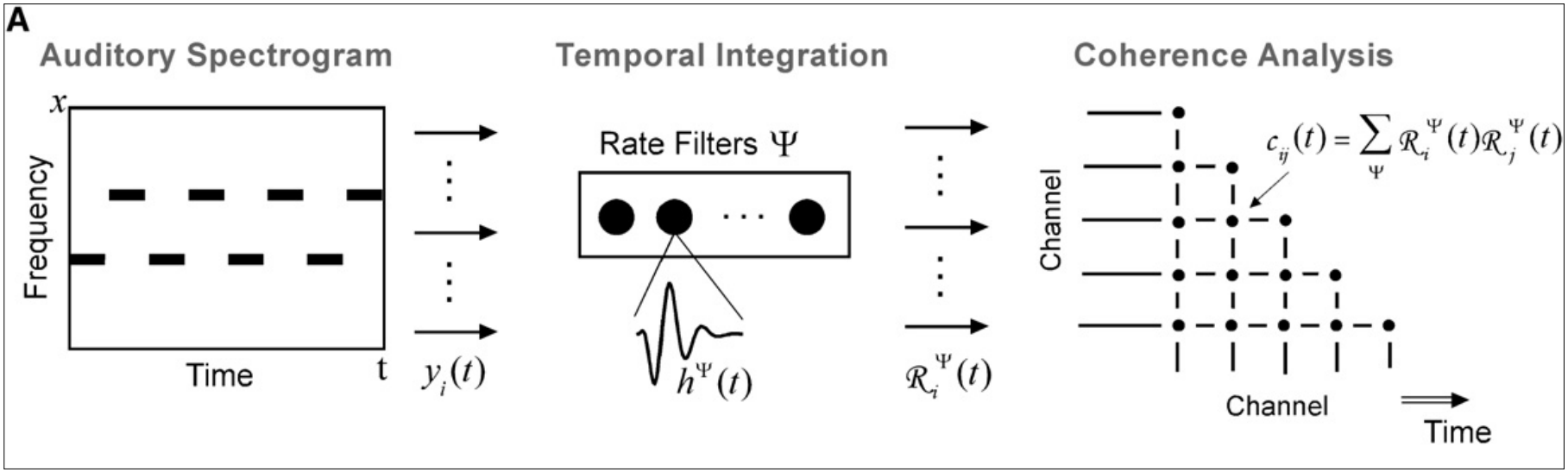
- Invariant to duration of figure, rather depends on no. of repeating components (Expt. 1 & 2)
- Invariant to disruption of signal components with silence or noise (Expt. 2, 3 & 4)
- Invariant to figure pattern – repeating or ramped (Expt. 1, 5 & 6)
- Invariant to ramp direction and ramp size (Expt. 5 & 6)
- Sensitive to background statistics (Expt. 1 & 7)

Discussion

What are the mechanisms underlying figure-ground segregation in SFG stimulus?

- Low-level mechanism, e.g. adaptation?
- A higher order mechanism? Where? IPS?
  - Temporal coherence model *(Shamma, 2009; 2010)* ?

Coherence analysis model



The model takes as input a time-frequency spectrographic representation of sound.

The signal in each channel yi(t) is then processed through a temporal integration stage, implemented via a bank of filters (J) operating at different time constants.

Finally, the output of each rate analysis is correlated across channels, yielding a coherence matrix that evolves over time.

Modelling…

The model takes as input a time-frequency spectrographic representation of sound.

The signal in each channel yi(t) is then processed through a temporal integration stage, implemented via a bank of filters (J) operating at different time constants.

Finally, the output of each rate analysis is correlated across channels, yielding a coherence matrix that evolves over time.