Towards a dynamical pattern recognition model of gamma activity in auditory cortex

Will Penny

Wellcome Trust Centre for Neuroimaging, University College London, UK

2nd International Conference on Neural Field Theory,
Centre for Integrative Neuroscience and Neurodynamics,
University of Reading, 19th April 2012

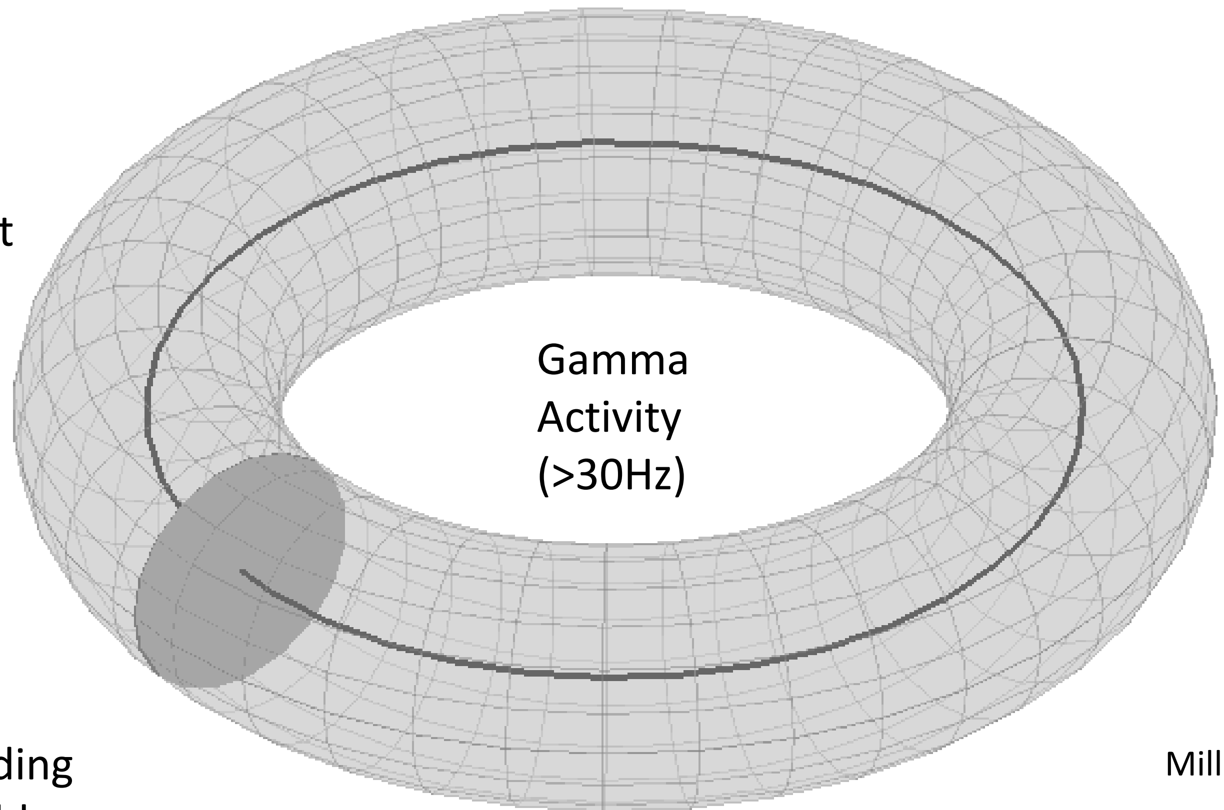
GAMMA ACTIVITY

Imaging Neuroscience

Spike
Time
Dependent
Plasticity

BOLD
signal

Goense et al
Curr Biol, 2008



Localisation

Miller, PLOS-CB,2009

Binding
Problem

Singer,
Neuron,
1999

Computational
Neuroscience

Overview

Neural network model of spatio-temporal pattern recognition in the brain (Hopfield & Brody, PNAS, 2001). Applied to speech recognition.

Algorithmic Level: Recognition based on pattern of Occurrence Times (OTs) of level-crossings of power in different frequency bands.

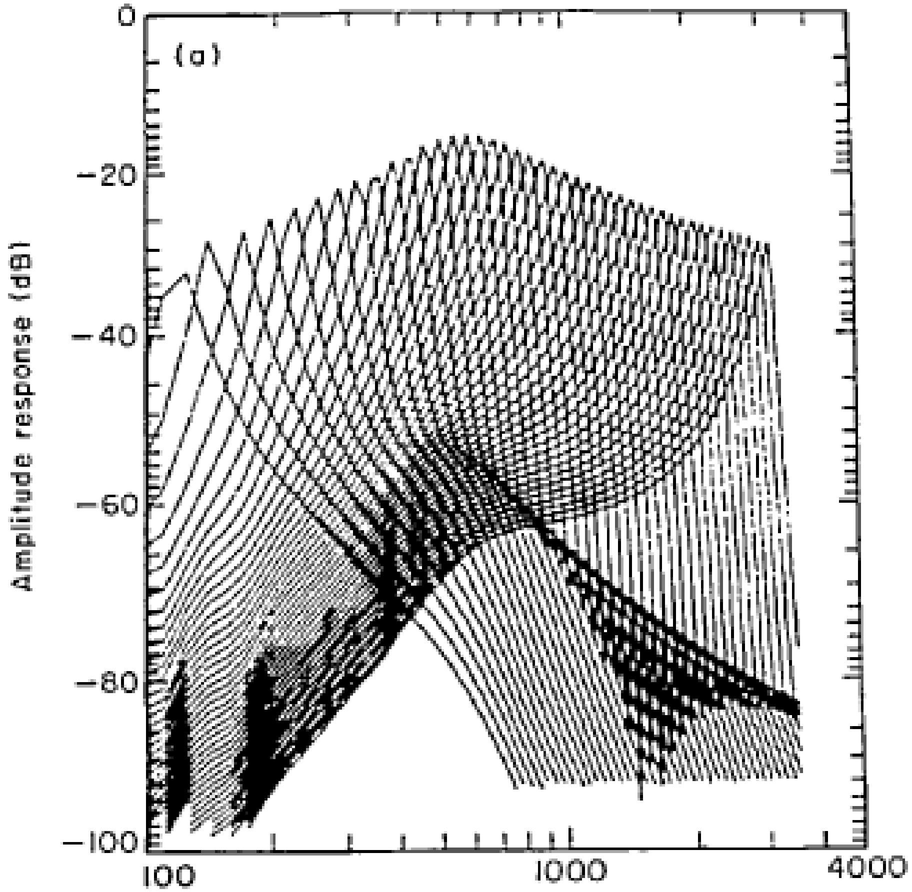
Implementation Level: Transient synchronization mechanism signalling recognition with a gamma burst (more details later !)

Idea: Use dynamical process as both a computational model and forward model of neuroimaging data.

Algorithmic Level

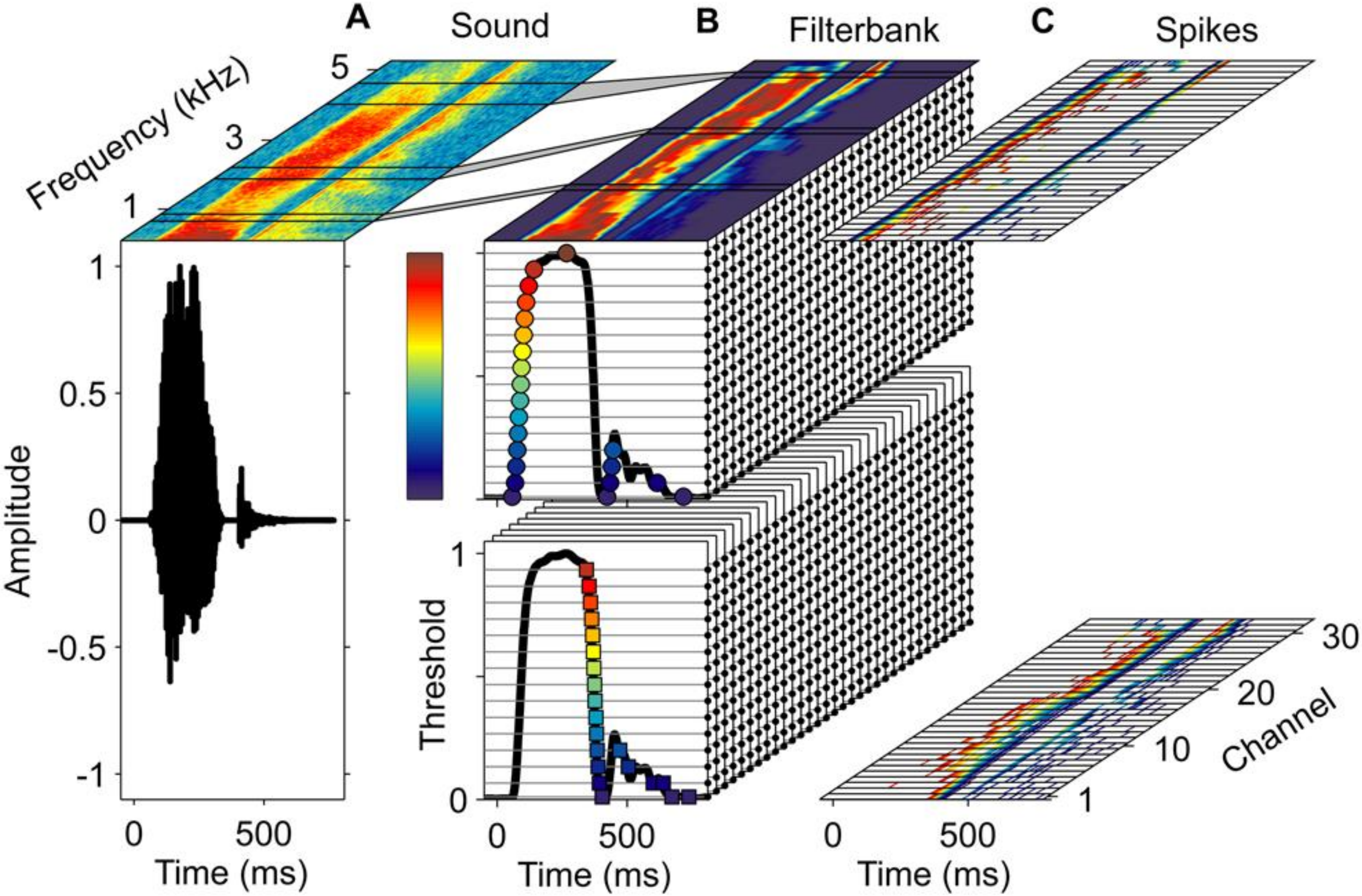
- Bandpass filtering
- Onset/offset or ‘level-crossing’ detectors
- Pattern recognition based on Occurrence Times (OTs)

Bandpass filters



Frequency

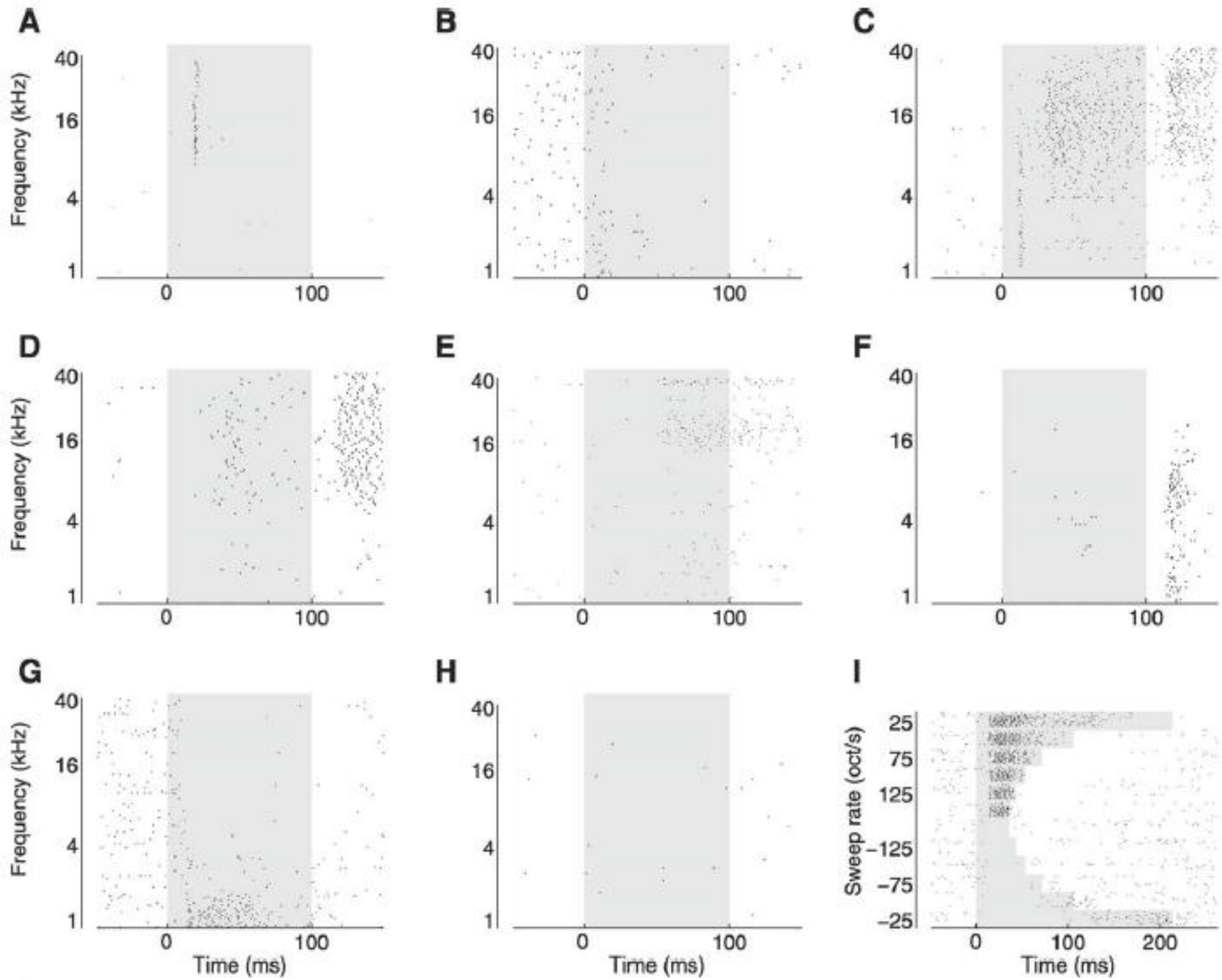
Allen (1985) Cochlear Modelling IEEE ASSP



LEVEL CROSSINGS

From Gutig and Sompolinsky (PLOS-Biology,09)

Onset Cells, Offset Cells, etc.



Cell attached recordings from A1 in rat

Hromadka et al, PLOS B, 2008

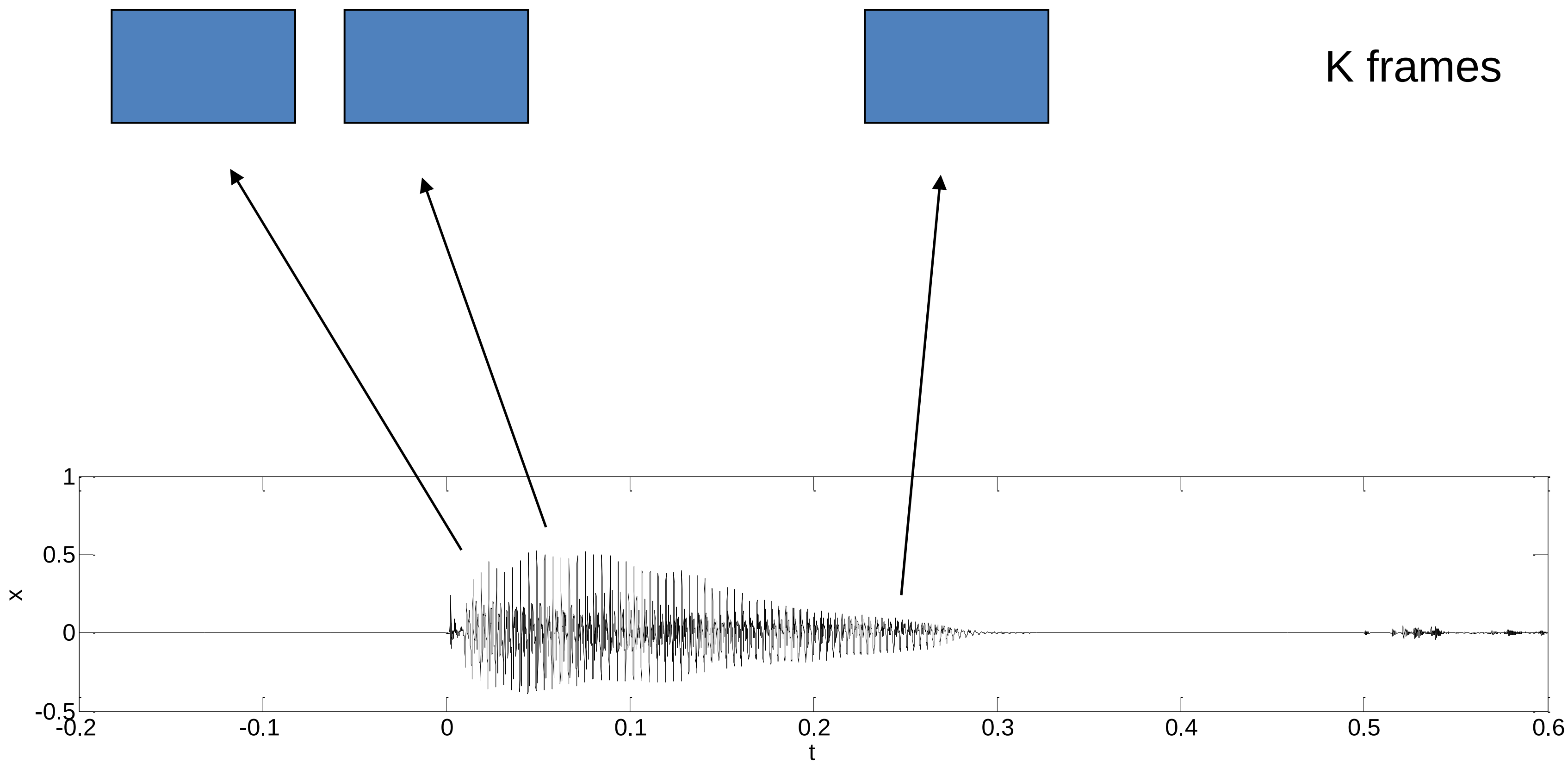
Recognising Patterns

Fukushima, Rolls etc.

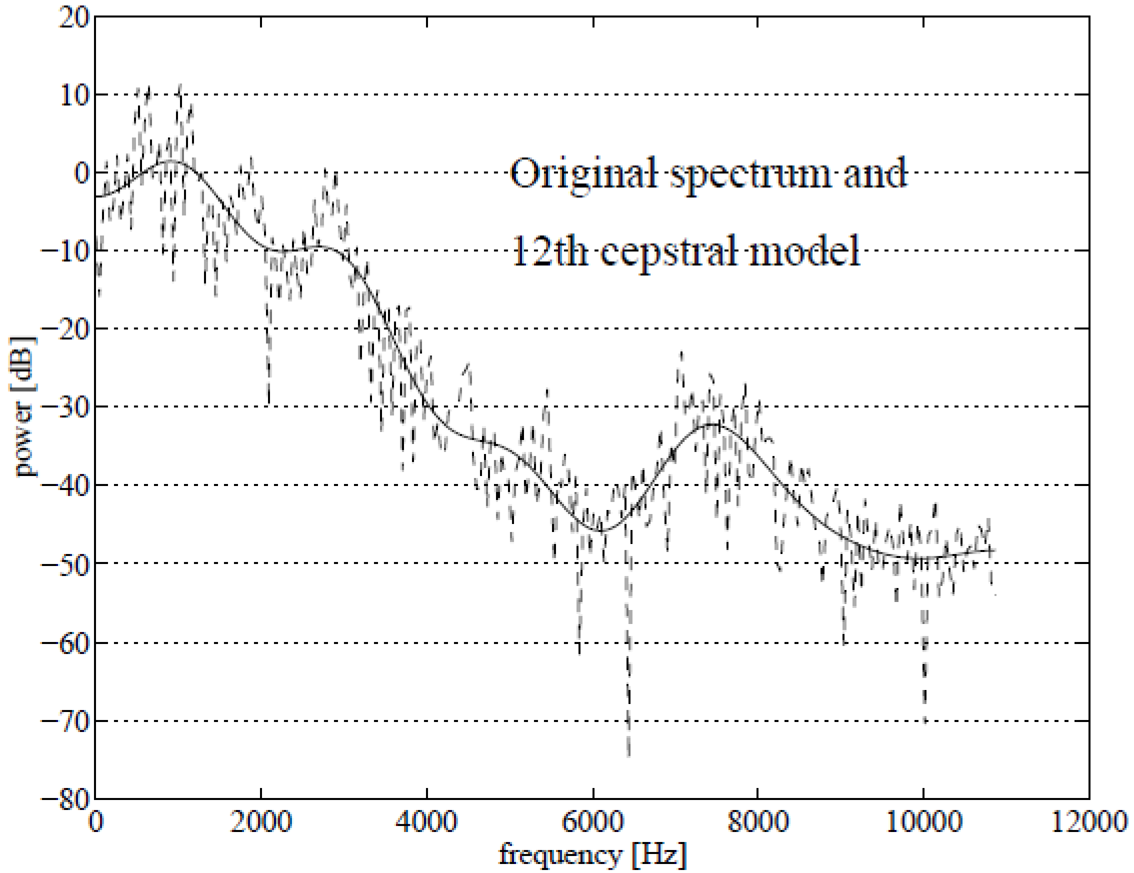
Bandpass filtering at multiple spatial scales	Level detection	Spatial configuration of Features
Bandpass filtering at multiple temporal scales	Level detection	Temporal configuration of Features

Standard Automatic Speech Recognition (ASR)

- Bandpass filters (Mel Frequency scale)
- DCT of log spectrogram keeping subset of coefficients
- Cepstral Coefficients (MFCC)



Get Cepstral Coefficients for each frame



S.Roweis, Speech Processing Tutorial, 1998

Classification

Use MFCC features in a nearest-neighbour classifier.

Use OT features in a nearest-neighbour classifier

With same classifier (‘back-end’) we can see what are the best features.

Speech Database

Subset of TI46 database (spoken digits 0 to 9)

5 female speakers

10 repetitions of each digit

Algorithms trained on 5 reps from all speakers (a total of 250 utterances)

Algorithms tested on remaining 5 reps in different additive noise environments (250 different utterances)

Speaker 1

Dig	1	2	3	4	5	6	7	8	9	10
0										
1										
2										
3										
4										
5										
6										
7										
8										
9										

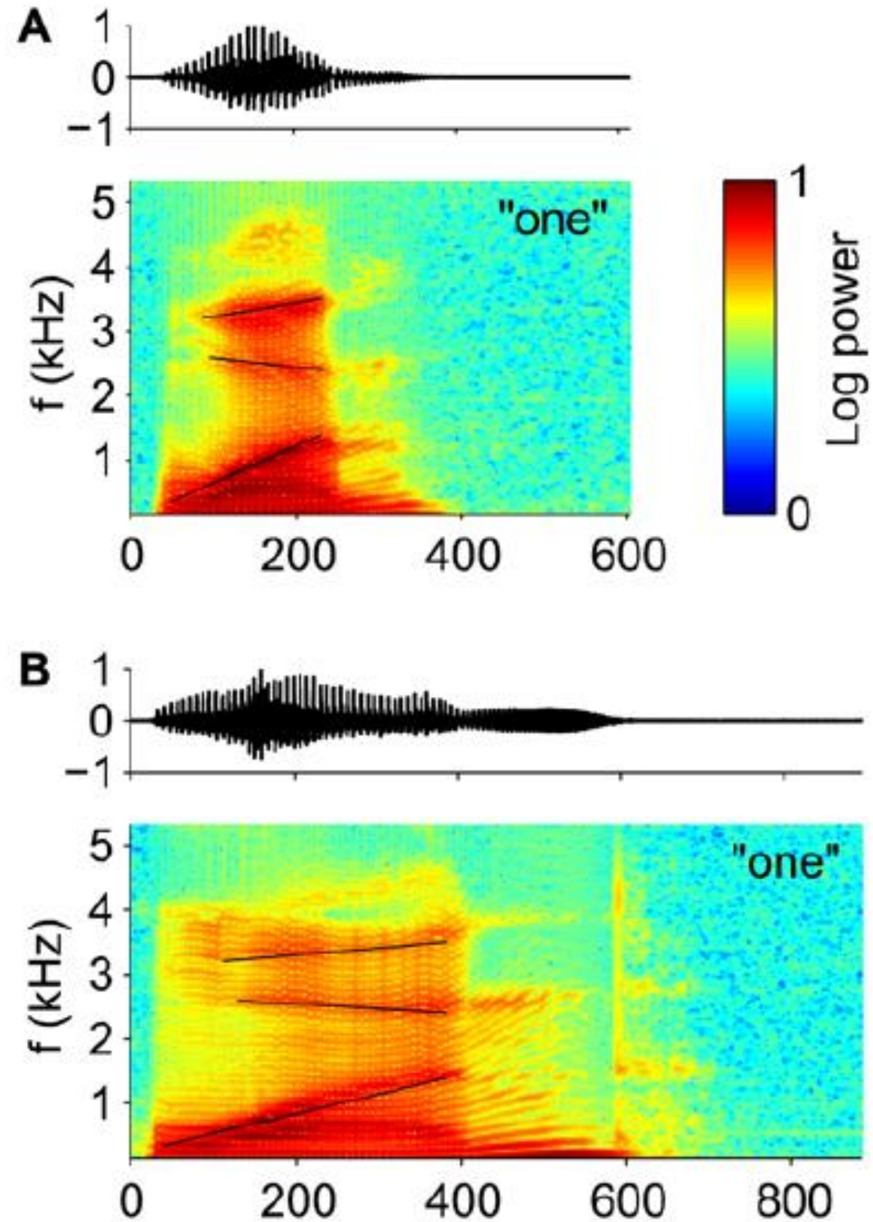
Speaker 2

Dig	1	2	3	4	5	6	7	8	9	10
0										
1										
2										
3										
4										
5										
6										
7										
8										
9										

Natural speech contains a four-fold variation in the speed at which words are spoken (Miller et al, Phonetica,84)

Readily handled by linear scaling of OTs.

Time-Warp Invariance.



From Gutig and Sompolinsky (PLOS-B,09)

ASR in White Noise

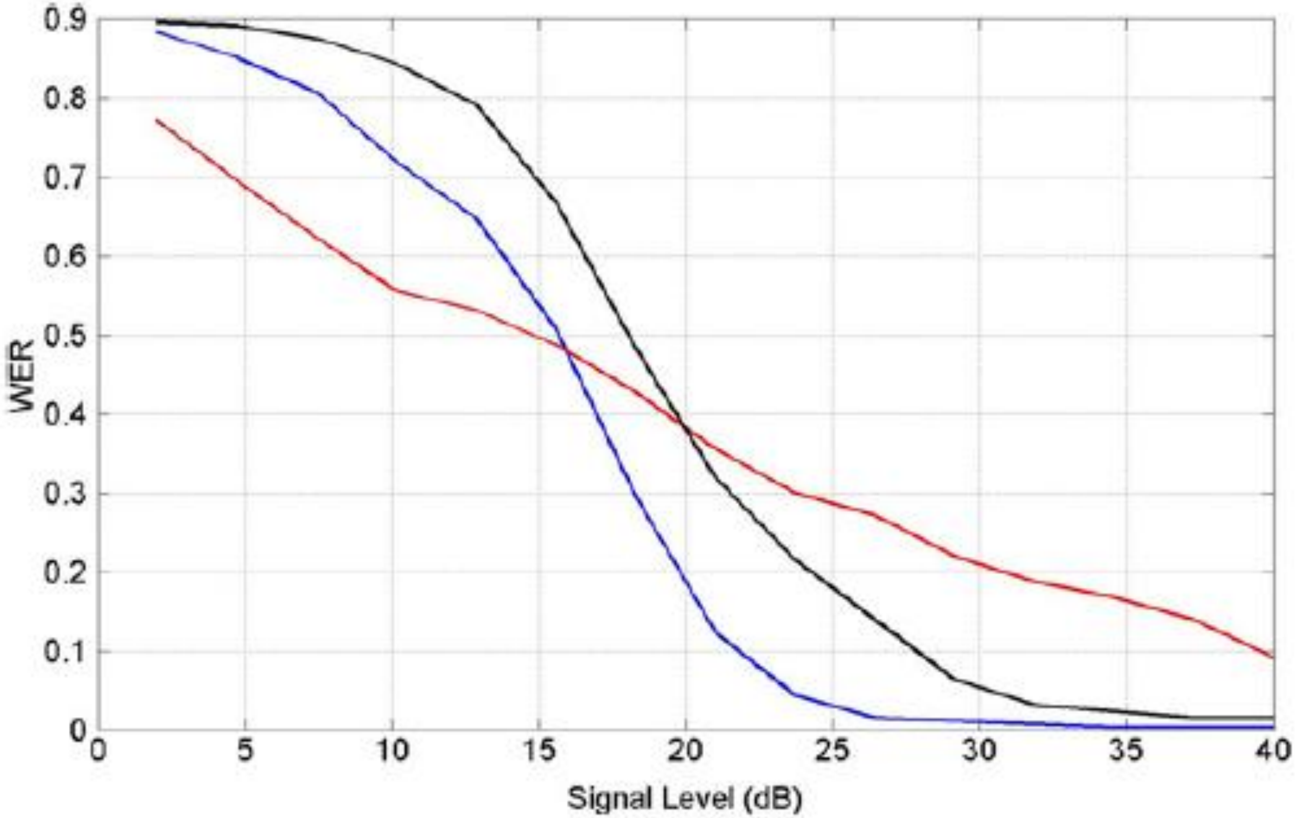


Fig. 6. Speech recognition in additive white noise. We plot Word Error Rate (WER) against signal level for optimized speech recognition systems using Occurrence Time (OT) features (red curve), Mel Frequency Cepstral Coefficients (MFCC) (blue curve) and MFCC coefficients but with the number of features matched to that of the OT system (black curve). (For interpretation of the references to colour in this

ASR in Speech Babble

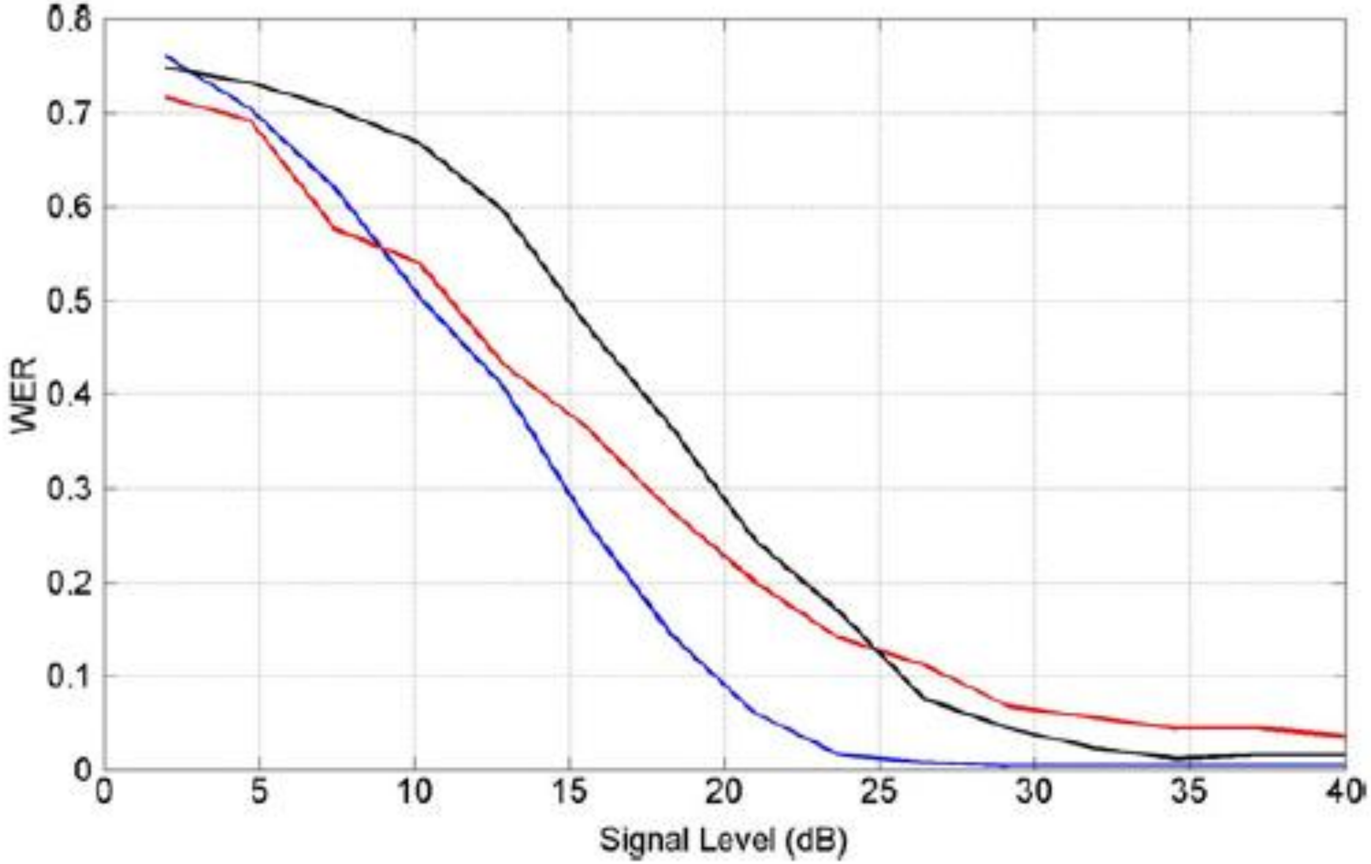


Fig. 7. Speech recognition in additive speech babble. We plot Word Error Rate (WER) against signal level for optimized speech recognition systems using Occurrence Time (OT) features (red curve), Mel Frequency Cepstral Coefficients (MFCC) (blue curve) and MFCC coefficients but with the number of features matched to that of the OT system (black curve). (For interpretation of the references to colour

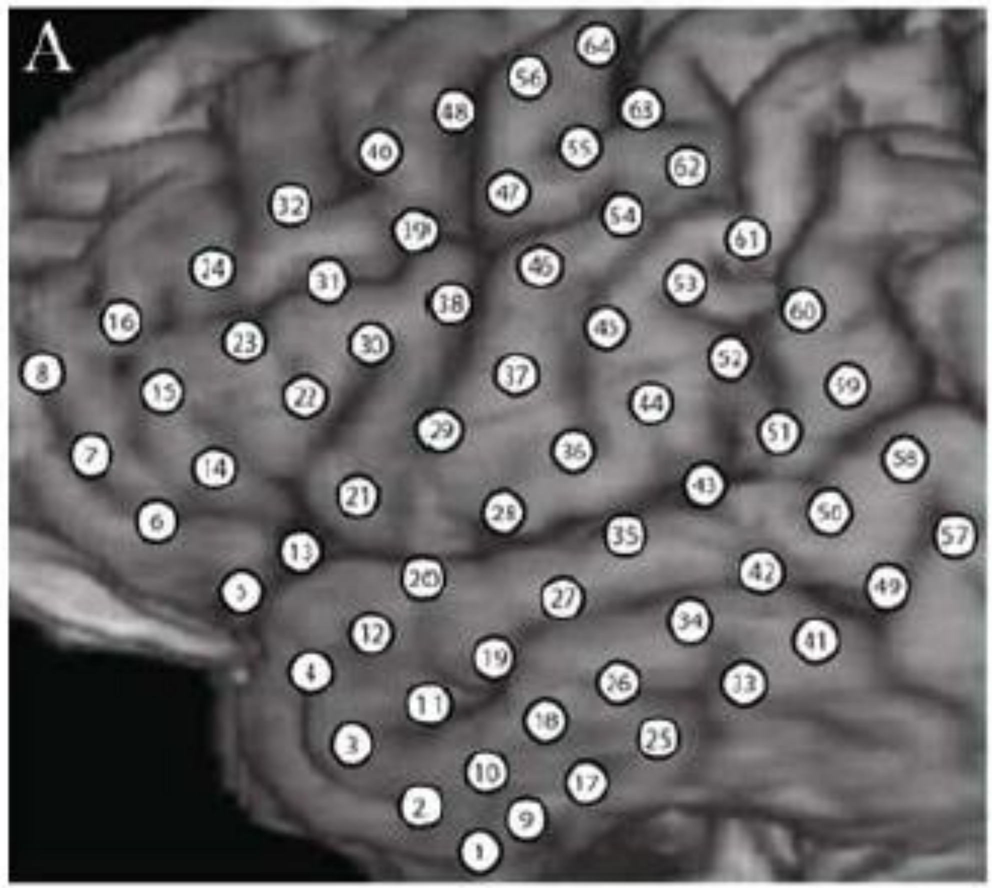
Interim Summary

- OTs are as good as MFCCs in high noise environments.
- But, MFCC-kNNs can be much improved. The standard is an MFCC-HMM.
- However, OT-kNNs can also be much improved. Gutig and Sompolinsky (PLOS-Biology, 2009) show an OT-Tempotron has performance equal to MFCC-HMM on full TI46 database.

ECOG recordings over fronto-temporal cortex

Electrodes are
4mm diameter.
Centres are
10mm apart.

Later work:
Smaller electrodes,
4mm apart.



Canolty et al. Front. Neuro, 2007

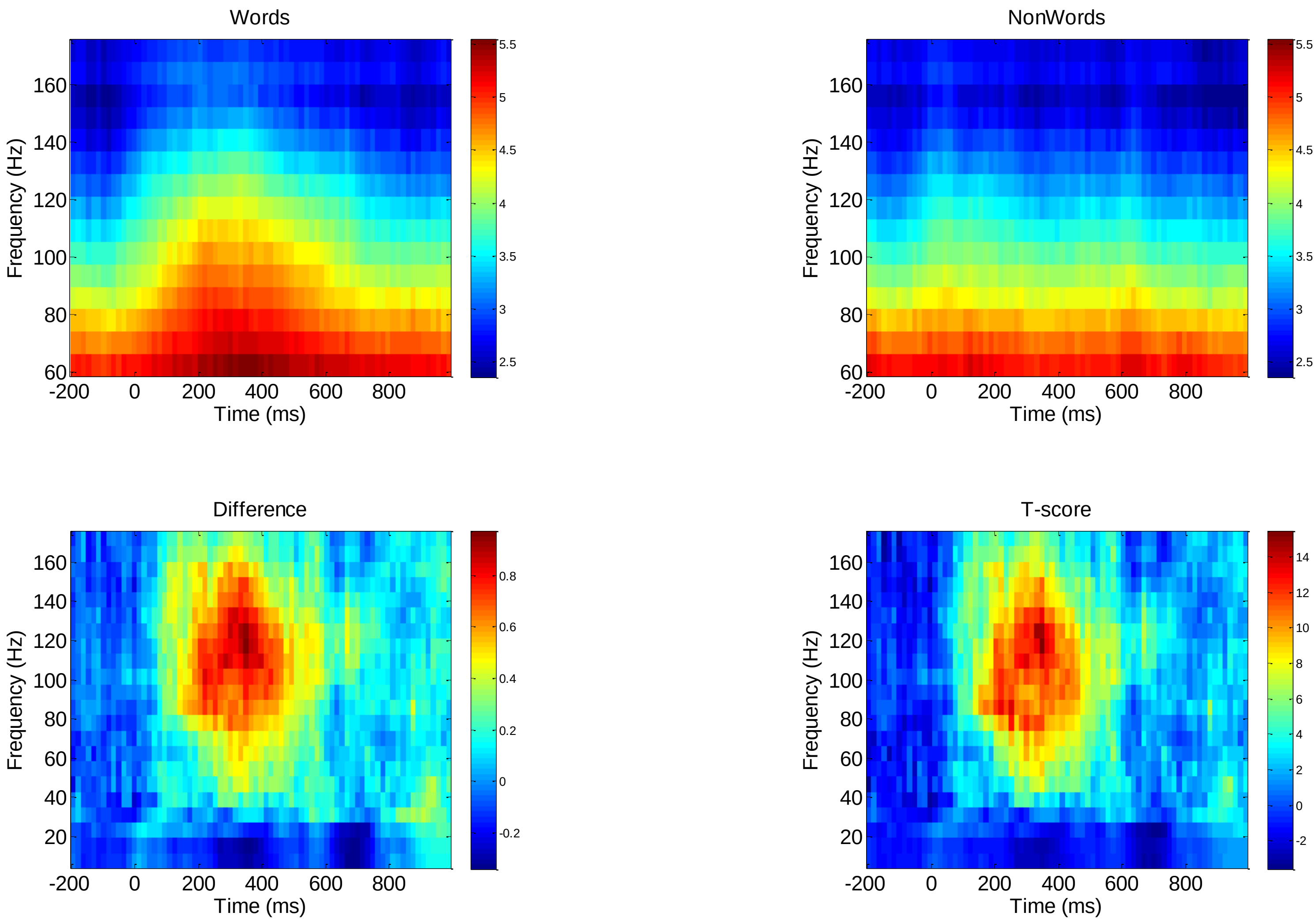
Subject listened to words and “nonwords”

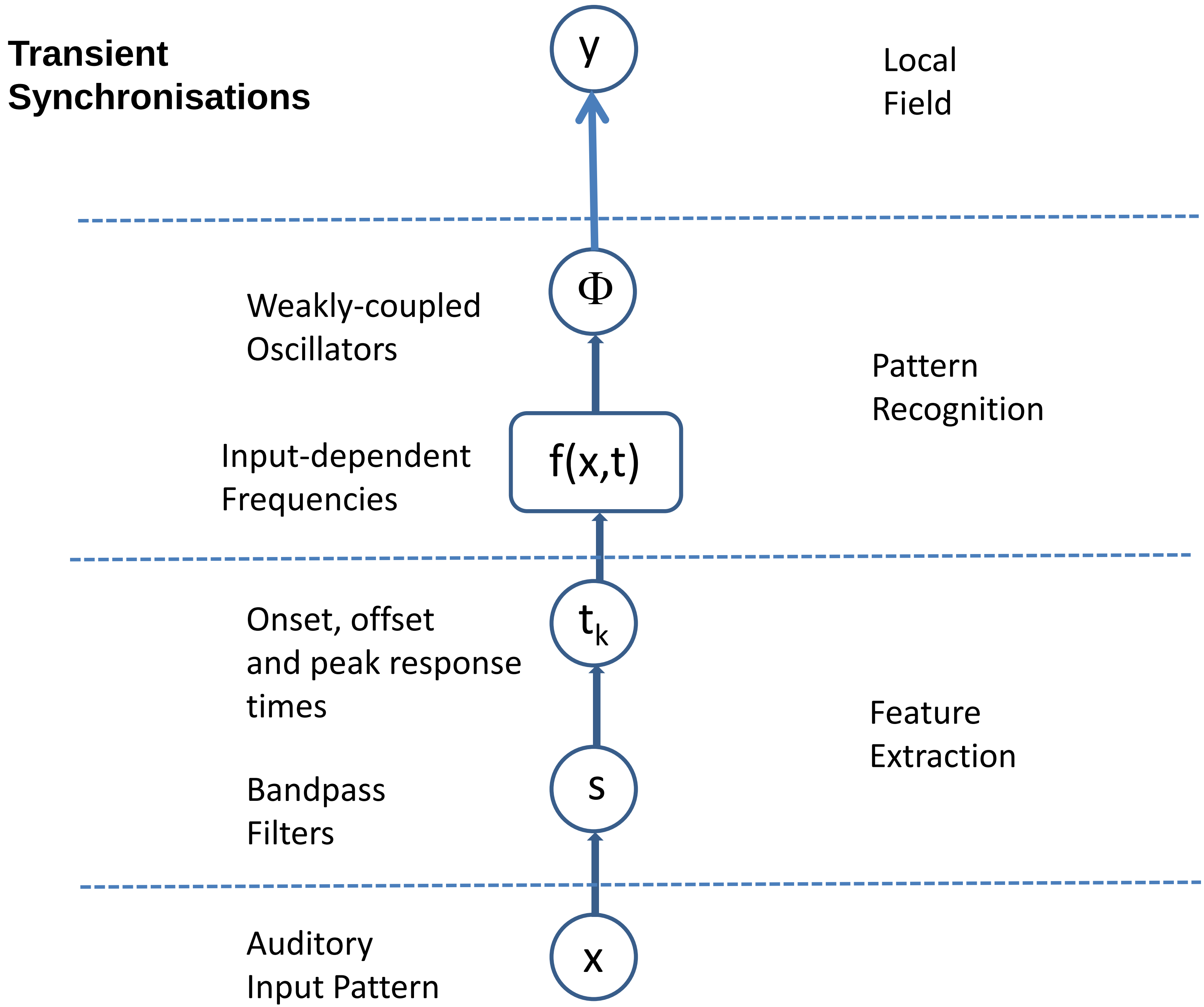
- Task: press button if word is a persons name (this data not analysed)

- Each nonword was created by taking a word, computing the spectrogram and removing ripple sound components corresponding to formants. The spectrogram was then inverse transformed (Singh and Theunissen, 2003)

- Each nonword matched one of the words in duration, intensity, power spectrum, and temporal modulation.

ECOG recordings over STS




Transient
Synchronisations
y
Local
Field
Φ
Weakly-coupled
Oscillators
Pattern
Recognition
f(x,t)
Input-dependent
Frequencies
t_k
Onset, offset
and peak response
times
Feature
Extraction
s
Bandpass
Filters
x
Auditory
Input Pattern

Weakly Coupled Oscillator Model

Input-pattern dependent frequencies

Phase Interaction Function

$$\dot{\phi}_i(t) = f_i(x,t) - \sum_j w_{ij} h[\phi_j(t) - \phi_i(t)] + z_i(t)$$

Lateral Connectivity: *Uniform* (A)

LFP $y(t) = \sum_j \cos\phi_j(t)$

Hopfield and Brody (PNAS, 2001) used Integrate and Fire Cells

Frequency Adaptation and Plasticity

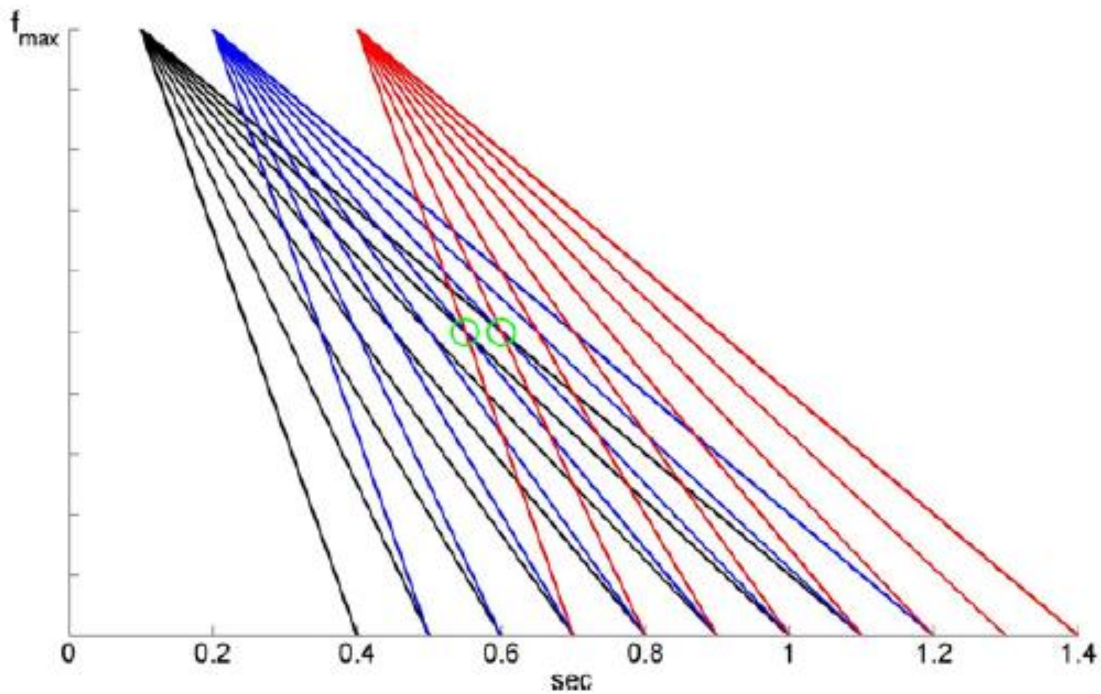
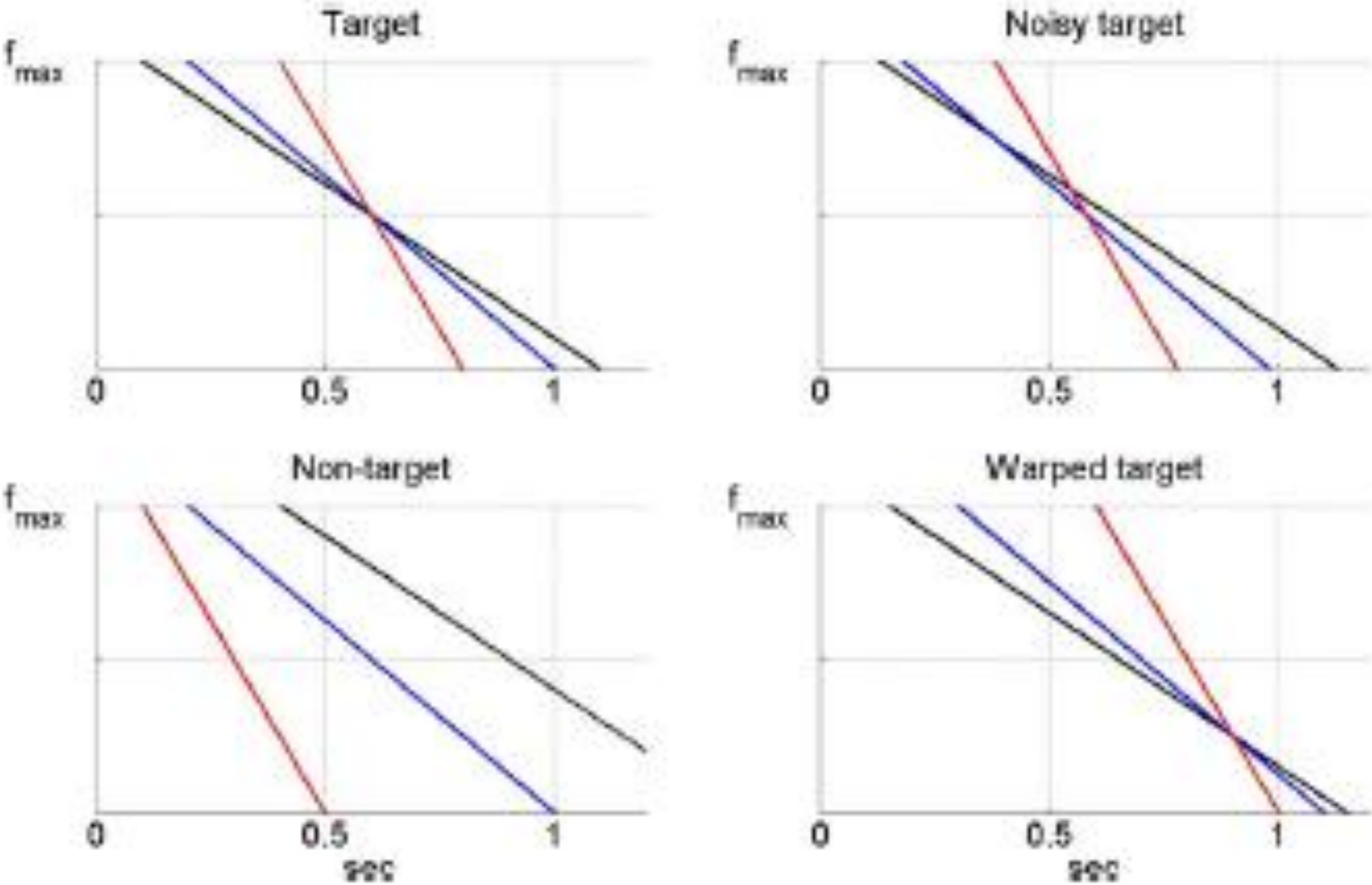


Fig. 4. The figure shows the detection of three different input features, $k = 1, 2, 3$ coloured black, red and blue. The kth feature is detected at time t_k—these are the Occurrence Times (OTs). For each feature there are $j = 1..J$ cells or circuits that initially respond at frequency $f_{\max}$, and then with linearly reducing frequency. The slopes of the frequency reduction are specified by the constants τ_j. The role of synaptic plasticity is to choose the optimal τ_j for each feature such that there will be a poststimulus timepoint, t_b, at which the frequencies become equal (f_b). These points are indicated by the green circles. Generally, plasticity acts to select long decay constants for early features and short ones for later features. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Specificity and Robustness



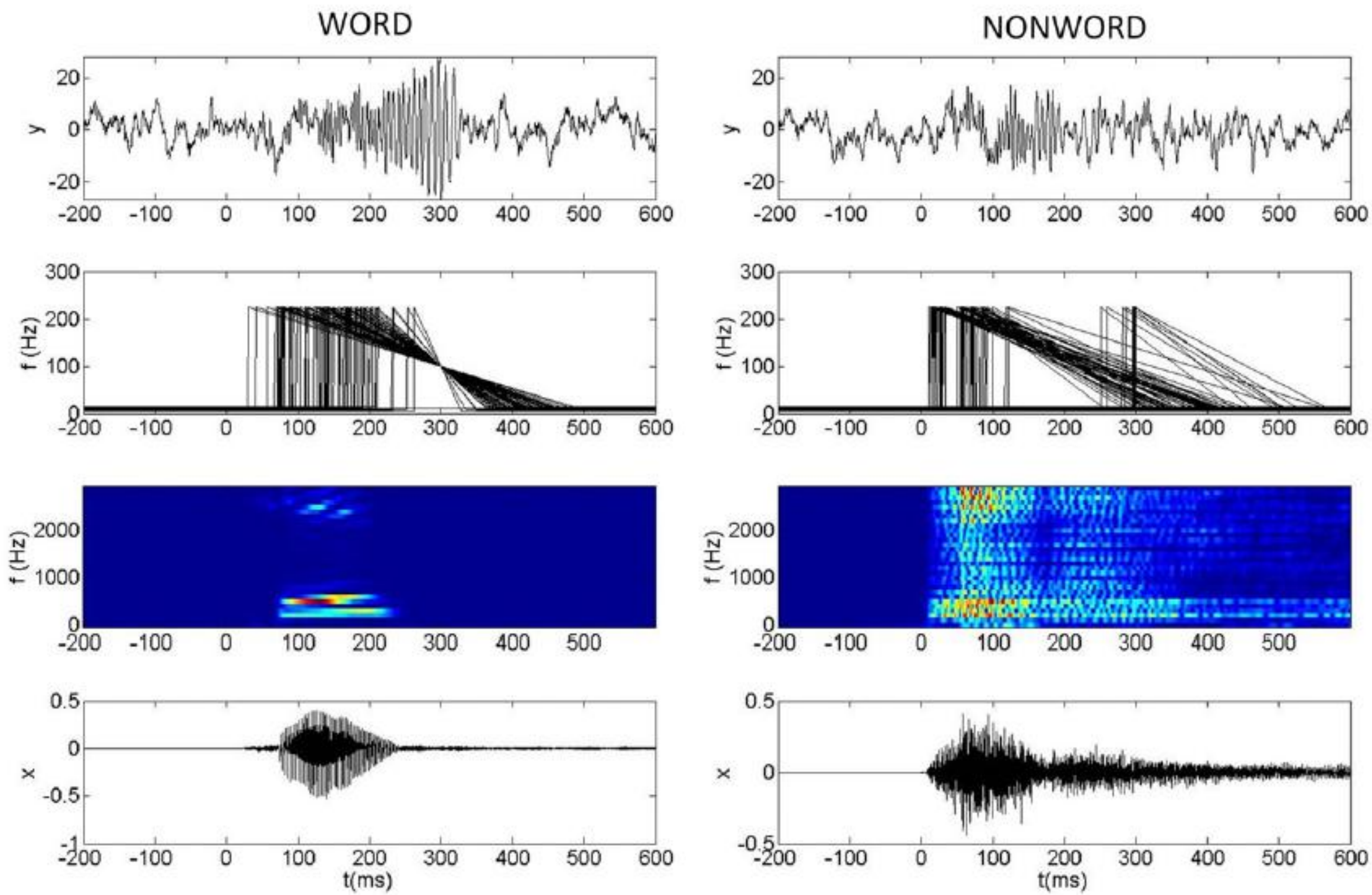
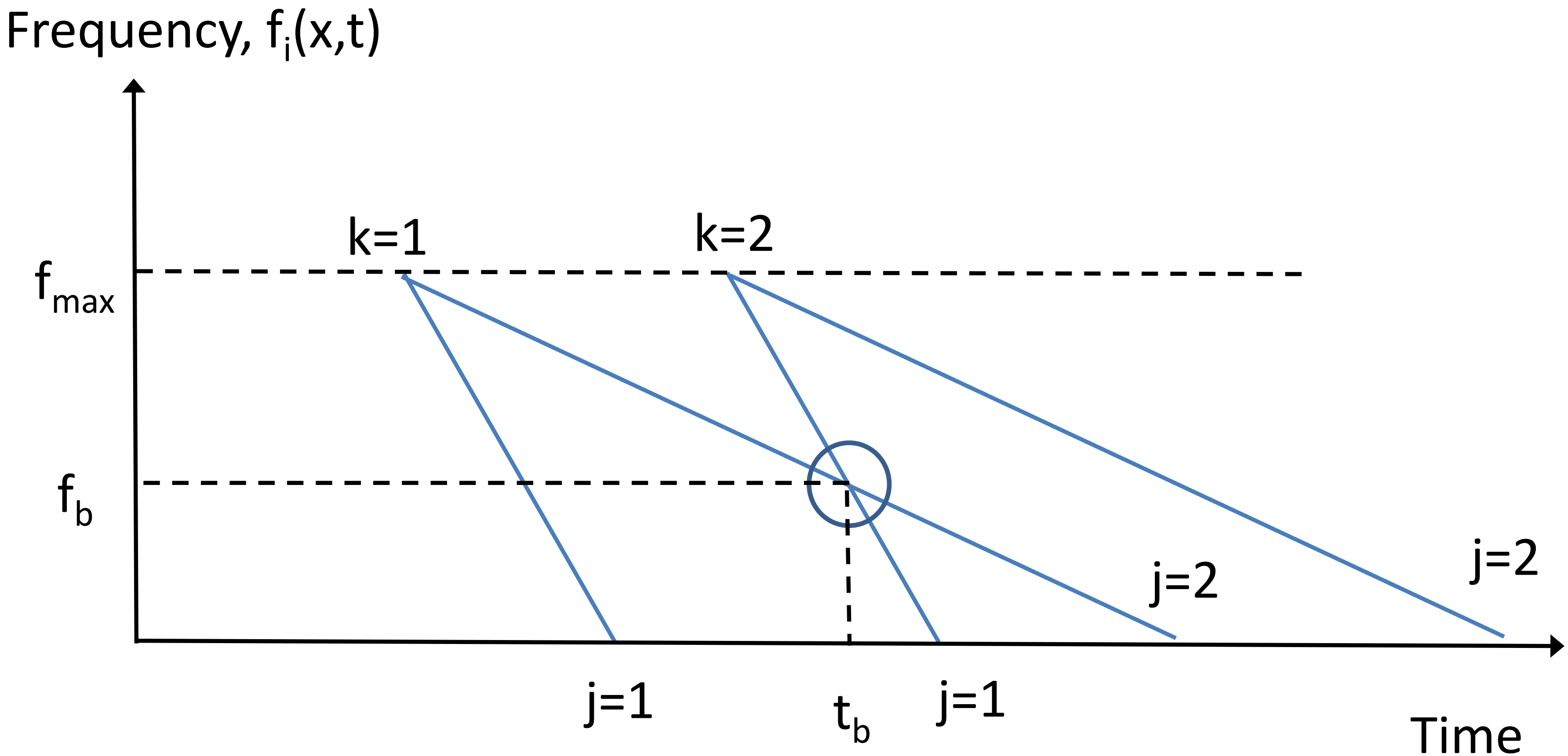


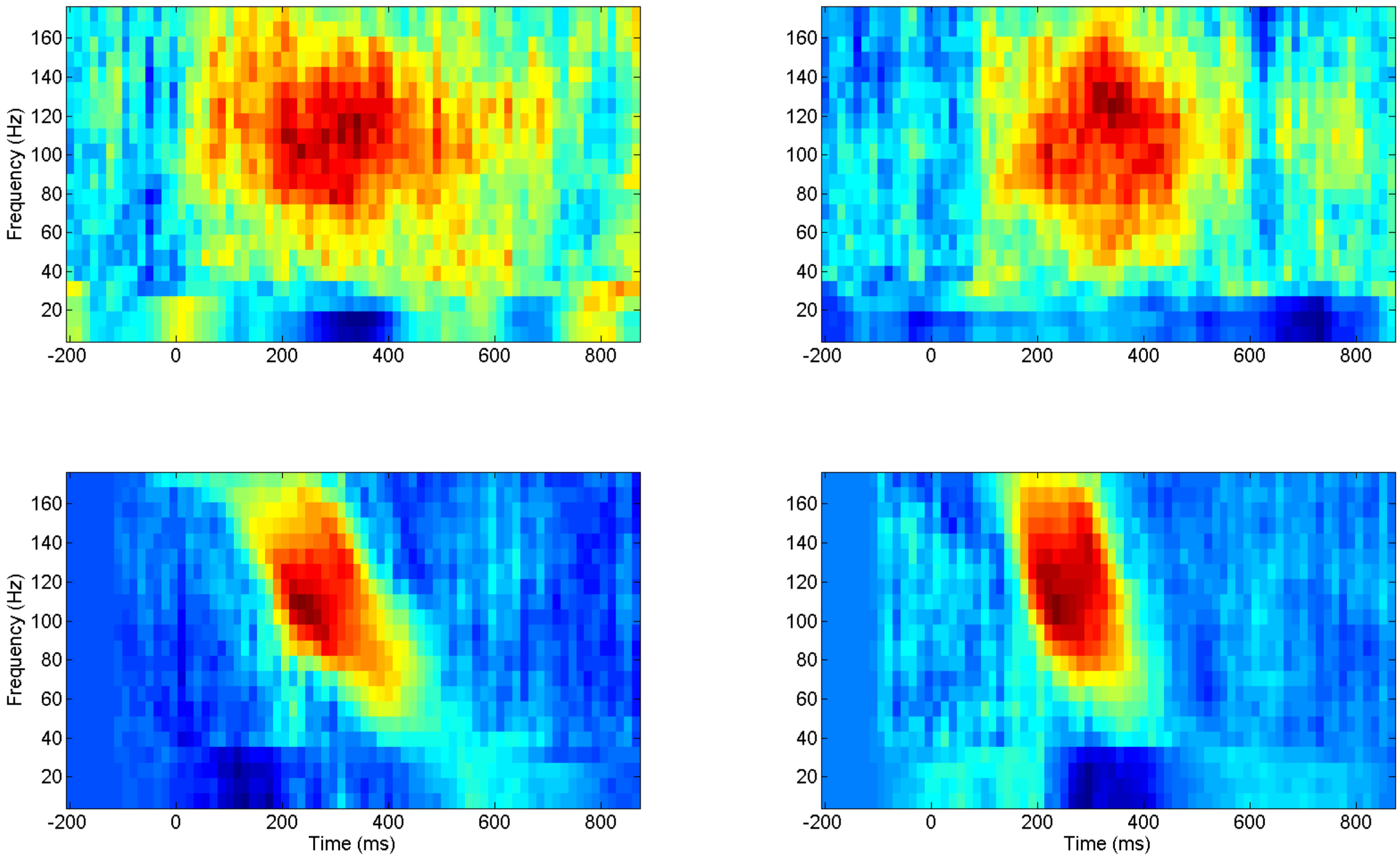
Fig. 3. The figure shows auditory inputs $x(t)$ (bottom row), auditory spectrograms $s(t)$ (third row), input frequencies to the WCO–TS network $f_k(x, t)$ (second row), and LFP signals from the WCO–TS net (top row) for a word (left column) and nonword (right column). The times at which the frequencies ramp up to their maximal value (in the second row) are the Occurrence Times (OTs). The word is 'Hiss' and the nonword was produced using MTF filtering (see text). The adaptation constants τ_j have been optimized such that the input frequencies become equal at $t_b = 300$ ms for the word input. The same τ_j's are used to generate the input frequencies for the nonword. For the nonword input there is no time point at which all the frequencies are equal and consequently no large LFP gamma burst.

Minimal model with four parameters: $\theta=\{A, f_{max}, f_b, t_b\}$



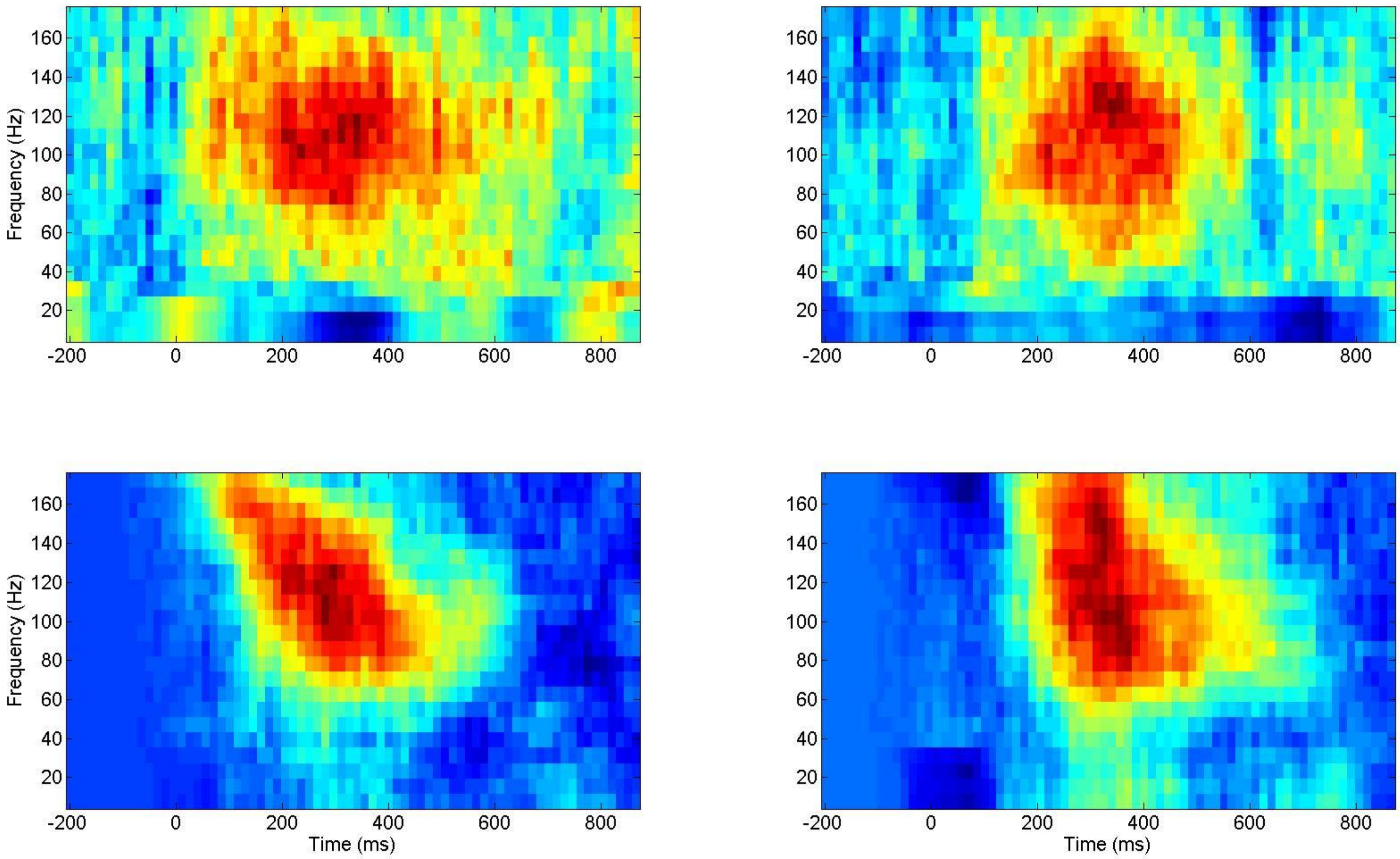
Training

Testing



Minimal Model

Training Testing



Augmented Model (coupling A is frequency dependent)

Summary

- Occurrence Times for Speech Recognition

Gutig & Sompolinsky (2009) PLoS Biology, 7(7):e1000141

- Forward model of neuroimaging data

Zavaglia et al. (2012) Neural Networks 28:1-14.

Transient Synchronisation of spikes (HB) or LFPs (Zavaglia) ?

Would need spike-field coherence to tell (eg Van Rullen, TiNS, 05)

- Have modelled ECOG activity at single electrode only. But word identity can be decoded from multielectrode high-res ECOG data (Pasley et al., PLoS B, 2012)

Acknowledgements

Melissa Zavaglia & Mauro Ursino
DEIS, Cesena, Italy

LIF network simulations

Tom Schofield & Alex Leff
UCL, London.

Cognitive neuroscience of language

Rich Canolty & Bob Knight
UCSF, San Francisco.

ECOG data and analysis

Oded Ghitza
Biomedical Engineering, Boston.

Reviewer