EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Philosophische Fakultät
Seminar für Sprachwissenschaft

Distinct ERP signatures of word frequency, phrase frequency, and prototypicality in picture naming

Peter Hendrix, Patrick Bolger & Harald Baayen

Introduction

- Primed picture naming
- Primes: preposition + definite article (e.g. “on the”, “into the”)
- Targets: photographs of nouns (e.g.; “strawberry”, “onion”)

Introduction

+

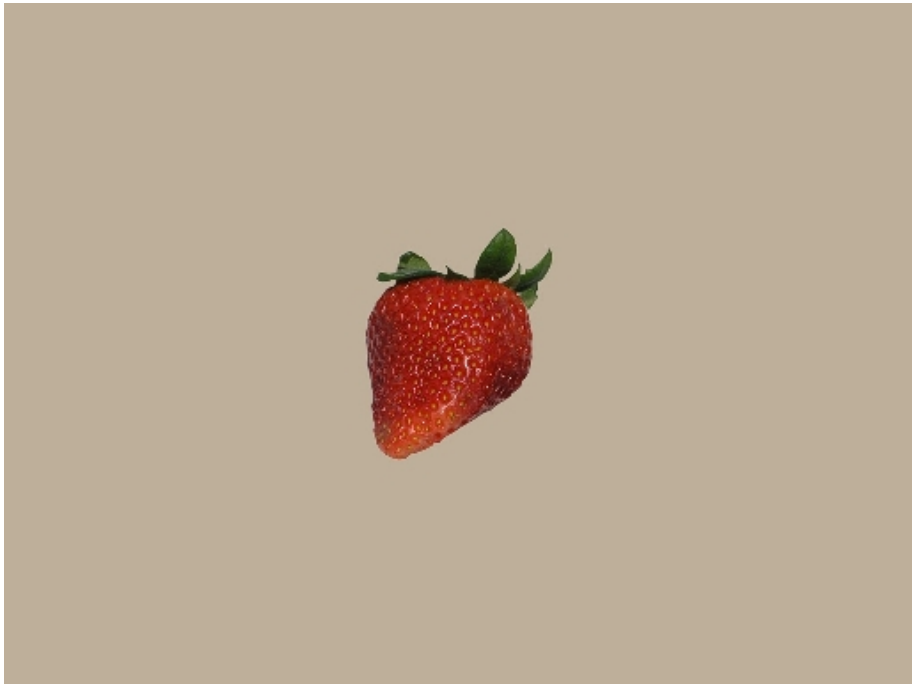
Introduction

on the

Introduction

+

Introduction



Introduction

+

Introduction

into the

Introduction

+

Introduction

Predictors

- Word Length
- Word Frequency
- Phrase Frequency
- Relative Entropy

Relative entropy

- Given two vectors of probabilities p and q, relative entropy measures the similarity of these vectors:

$$\text{Relative Entropy} = \sum_{i=1}^{n} (p_i * \log_2 (p_i / q_i))$$

  where n is the length of the vectors

- Applied to prepositional phrases relative entropy is a measure of prototypicality

Relative Entropy

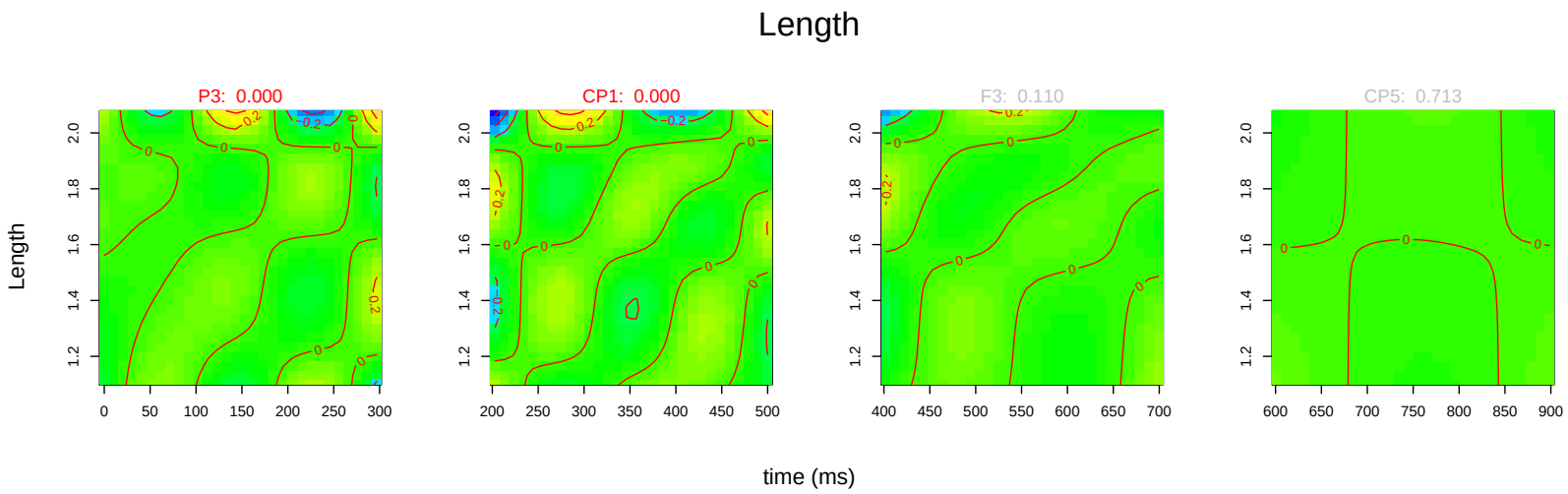
The two vectors of probabilities determining the relative entropy of "onion":

Phrase	**Freq.**	**Prob. *p***	**Prep.**	**Freq.**	**Prob. *q***
"with the onion"	8,867	0.305	"with"	2,171,020	0.074
"in the onion"	7,058	0.243	"in"	10,212,008	0.347
"to the onion"	5,734	0.197	"to"	4,148,449	0.141
"from the onion"	2,213	0.076	"from"	2,150,946	0.073
"on the onion"	1,922	0.066	"on"	4,010,429	0.136
"into the onion"	1,337	0.046	"into"	1,296,889	0.044
"up the onion"	1,091	0.038	"up"	403,114	0.014
"over the onion"	826	0.028	"over"	269,847	0.009
...	...	...	...	...	...
Total	29,048	1.000		29,401,403	1.000

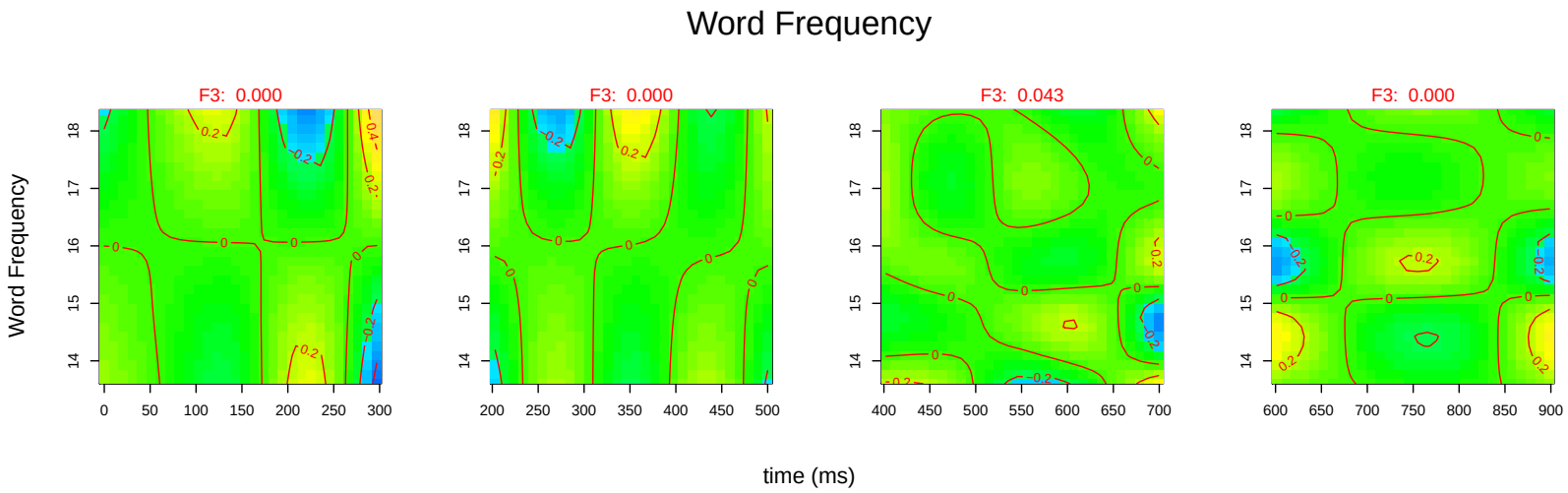
Analysis

- Dependent variable: ERP signal from picture onset at 32 electrodes
- Generalized additive models (GAMs)
- Models control for main trend over time, as well as effects of item, participant, trial and picture complexity

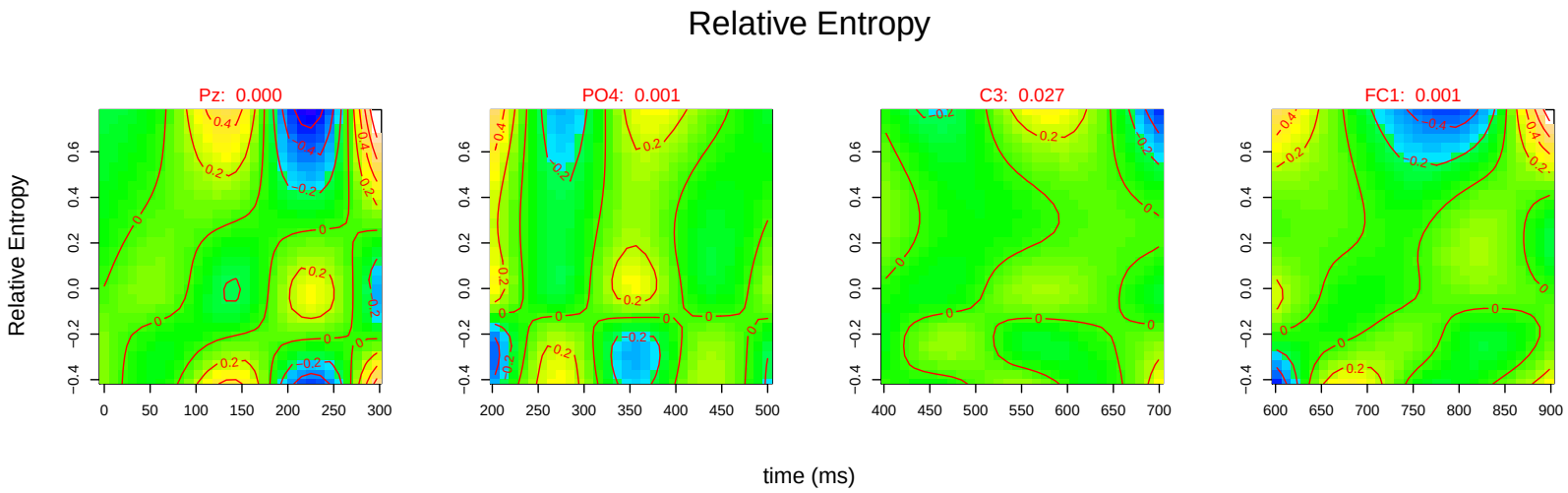
Results



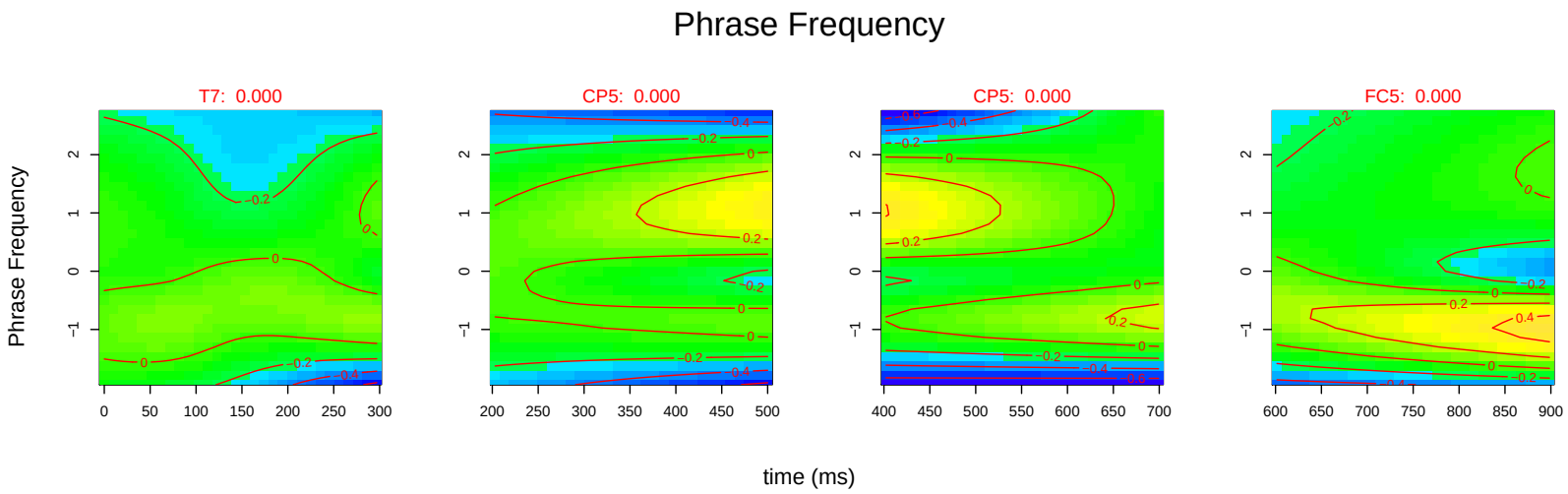
Results



Results



Results



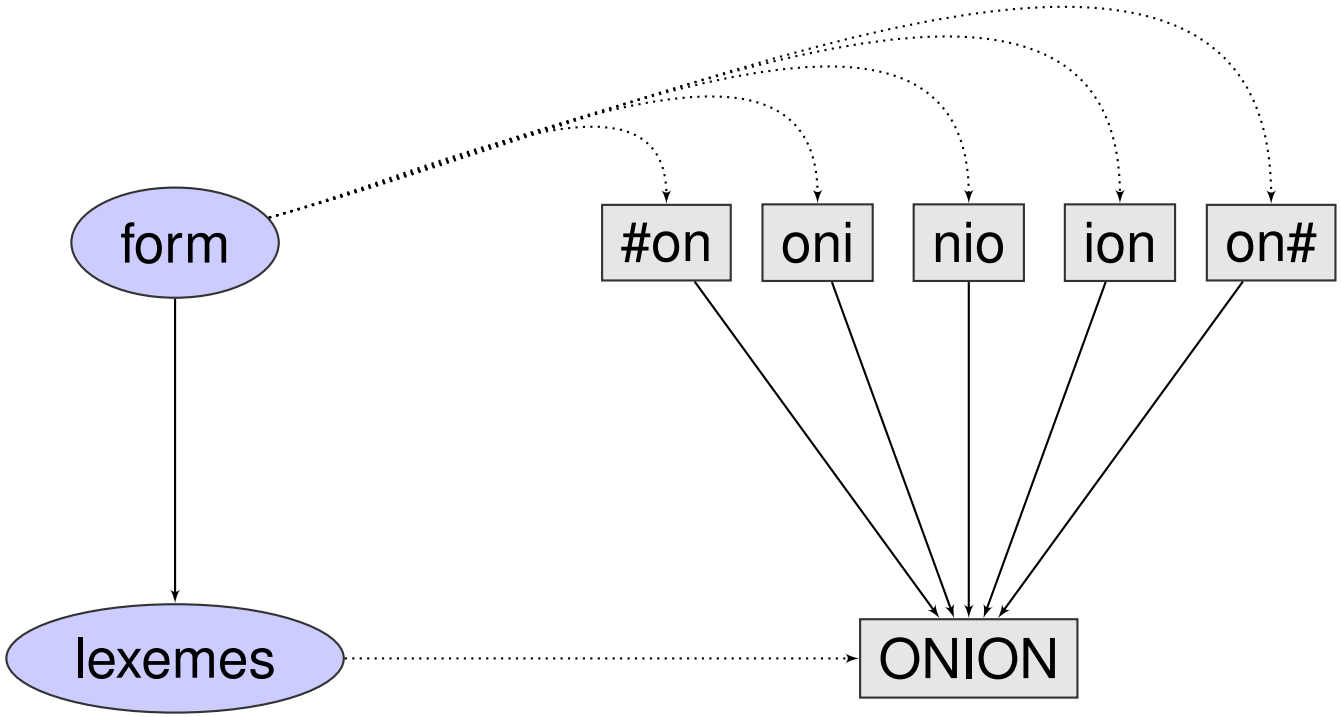
Discussion

- GAMs reveal non-linear predictor effects over time in the ERP signal
- Distinct effects of Word Frequency, Relative Entropy and Phrase Frequency
- What do these effects mean?

Naive Discriminative Learning

- Naive Discriminative Learning (see Baayen et al., 2011)
- NDL models learn associations between orthographic cues and lexemes

Naive Discriminative Learning



Naive Discriminative Learning

- Simulation of the experiment using Naive Discriminative Learning (NDL)
- Cues: #in, int, nto, to#, o#t, #th, the, he#, e#o, #on, oni, nio, ion, on#
- Outcomes: ONION

Naive Discriminative Learning

- The activation of a word's lexeme (a_{word}) given its set of input cues (C_k) is defined as:

$$a_{word} = \sum_{j \in \{C_k\}} V_{jword}$$

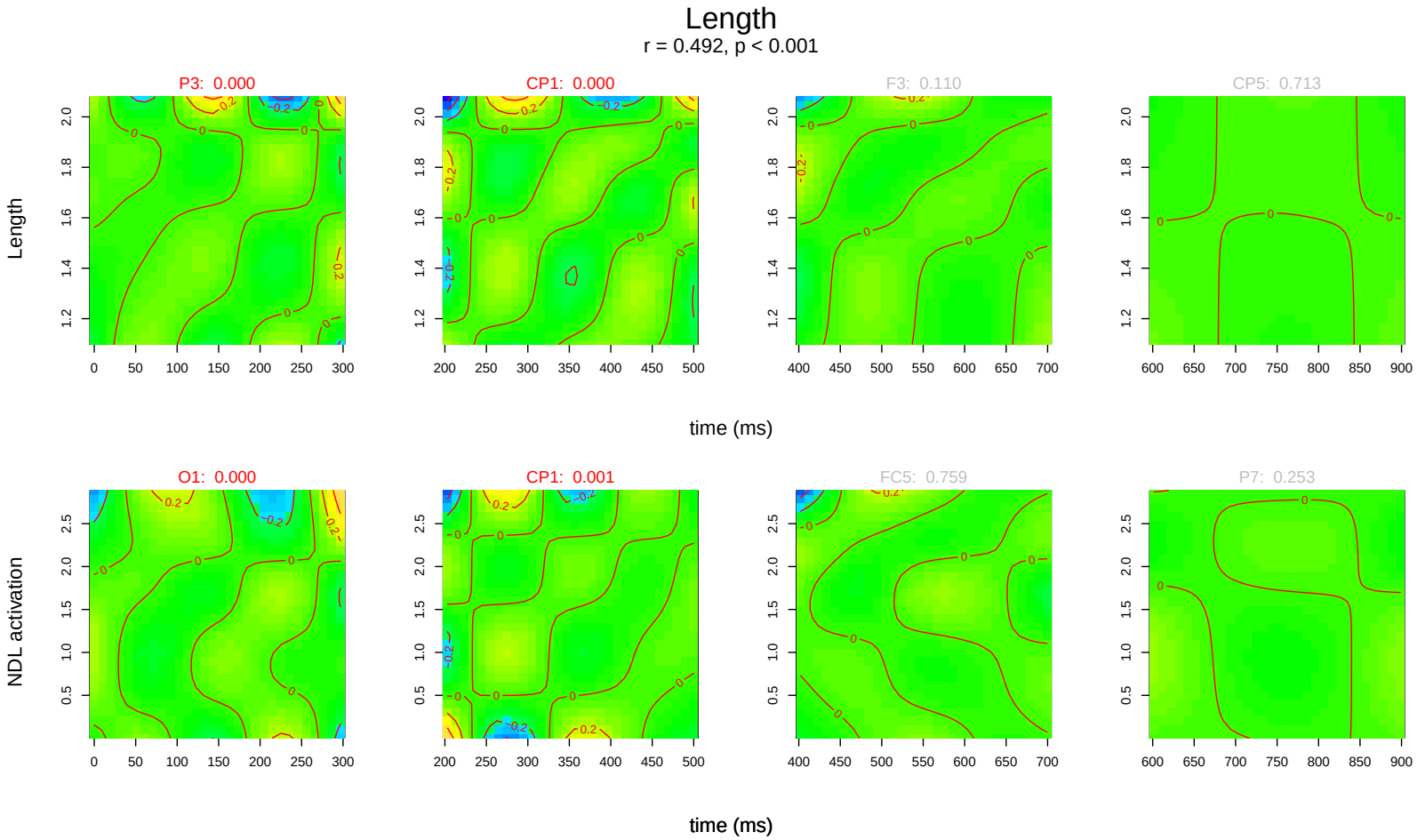
- Simulated reaction times are inversely proportionally to a_i:

$$RT \propto \log(\frac{1}{a_{word}})$$

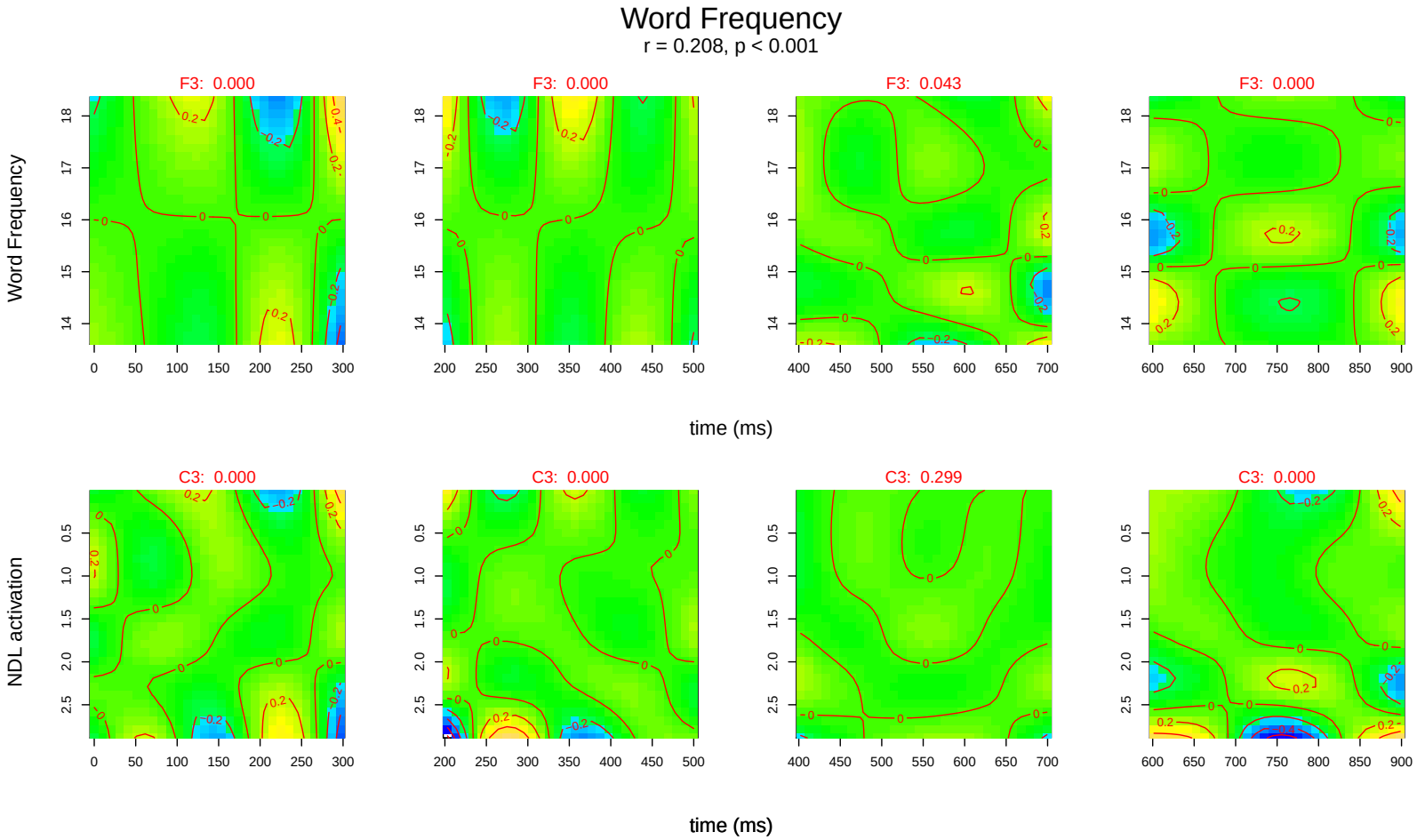
Naive Discriminative Learning

- Can NDL activations predict the ERP signal in a picture naming task?
- Use NDL activations as a predictor
- Compare results to word length, word frequency and relative entropy effects

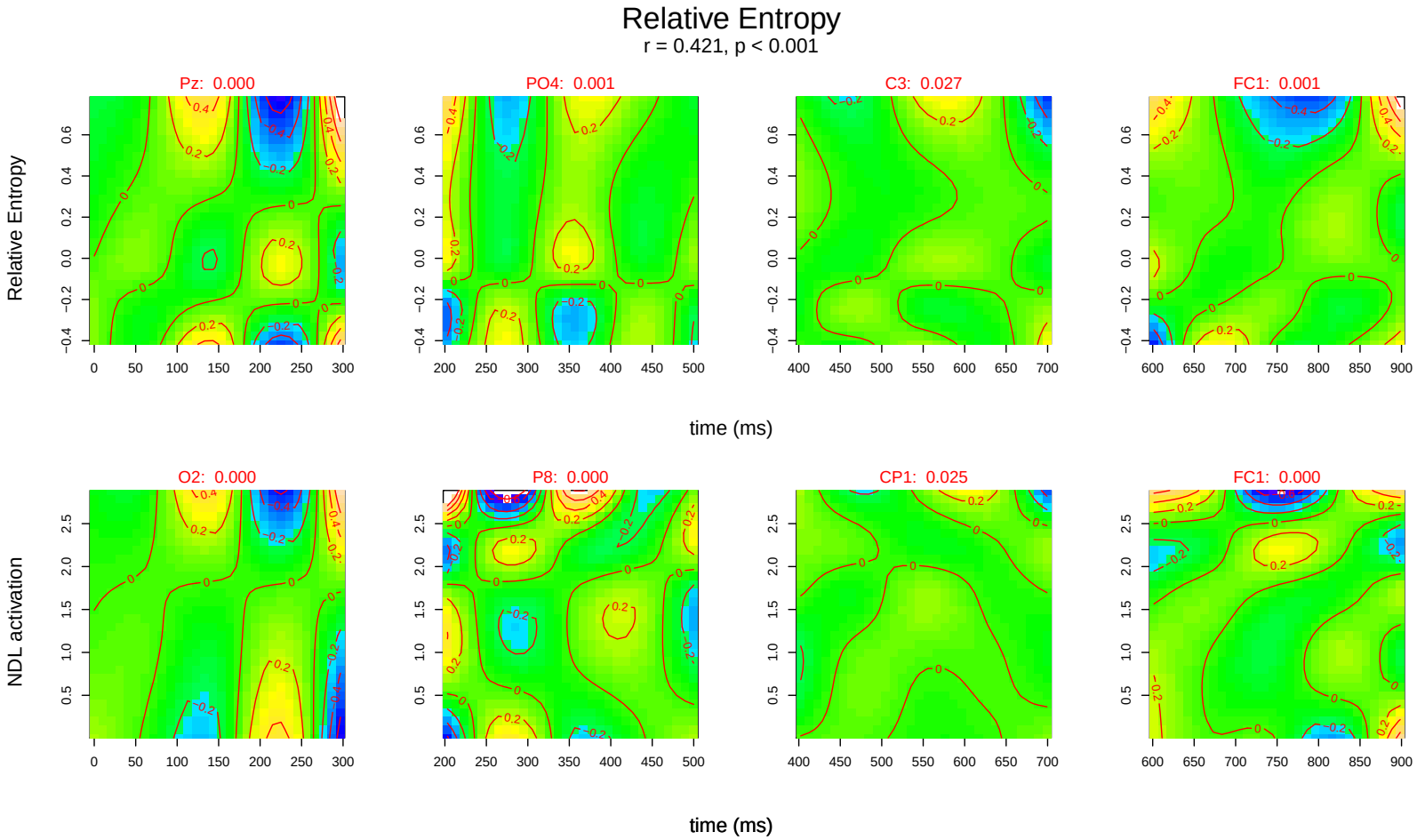
Naive Discriminative Learning



Naive Discriminative Learning



Naive Discriminative Learning

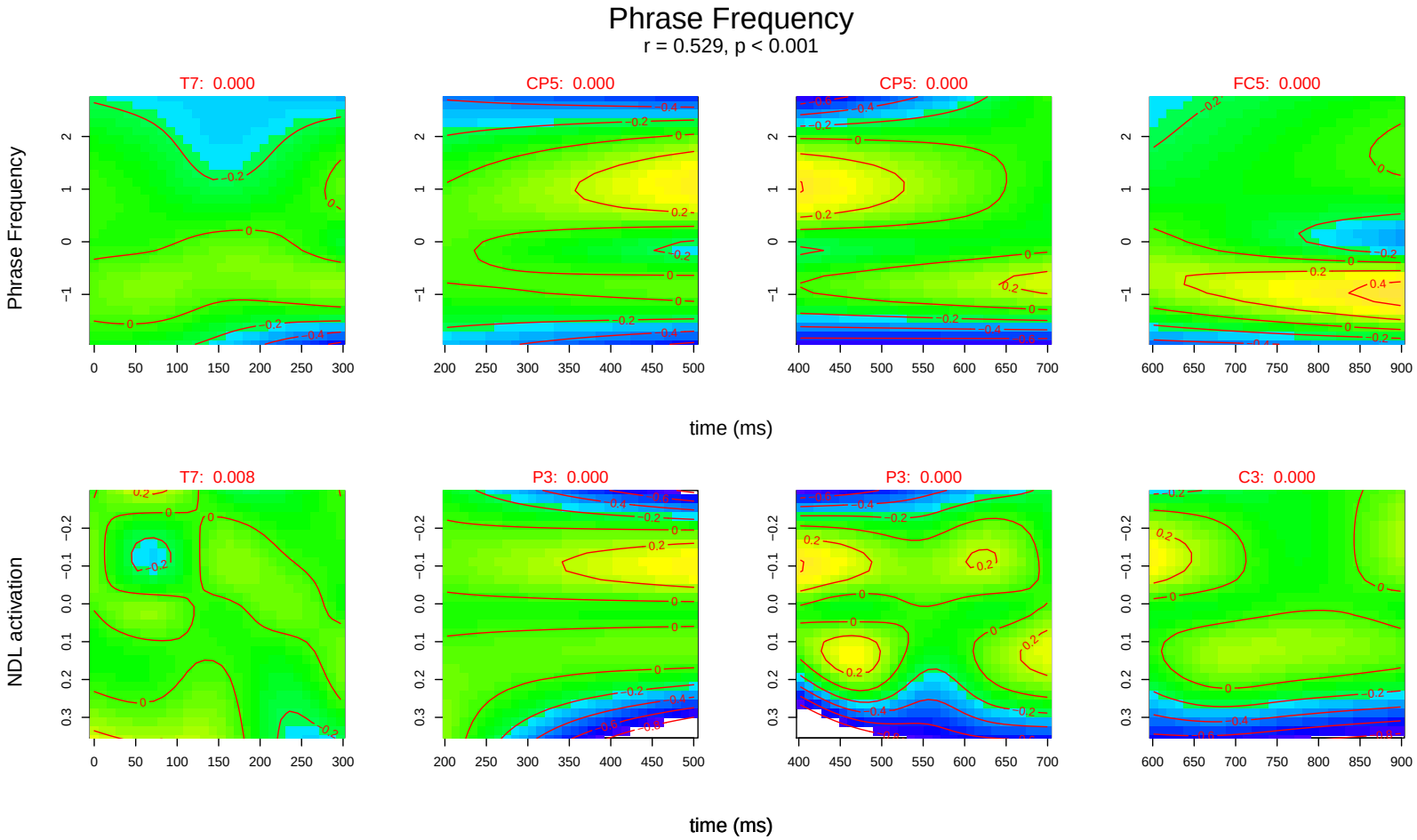


Naive Discriminative Learning

- Does this work at the phrase level as well?
- Cues: #in, int, nto, to#, o#t, #th, the, he#, e#o, #on, oni, nio, ion, on#
- Outcomes: INTO, THE, ONION
- Simulated reaction times:

$$RT \propto \log(\frac{1}{a_{preposition}+a_{determiner}+a_{word}})$$

Naive Discriminative Learning



Naive Discriminative Learning

- NDL activations capture the ERP signatures of the effects of Word Length, Word Frequency, Relative Entropy and Phrase Frequency
- AIC scores are lower for NDL models at all electrodes in all epochs

Summary

- NDL activations are a better predictor of the ERP signal in a picture naming task than traditional lexical predictors
- NDL activations provide a systemic estimate of the information in the language processing system

Thank you!