

# Speech Perception and Deception Linked to Neural Representations of Prediction Error

**Helen Blank (hblank@uke.de)**

Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf  
Martinistr. 52, 20248 Hamburg, Germany

**Marlene Spangenberg (marlene.spangenberg@psy.ox.ac.uk)**

Department of Experimental Psychology, University of Oxford,  
15 Parks Road, Oxford OX1 3PH, UK

**Matthew H. Davis (matt.davis@mrc-cbu.cam.ac.uk)**

MRC Cognition and Brain Sciences Unit, 15 Chaucer Rd,  
Cambridge, CB2 7E, UK

## Abstract:

Perception depends on combining sensory input with prior knowledge. When sensory signals are degraded, informative priors can improve perception but may also lead to misidentification (deception). Here we contrast two functionally-distinct computational mechanisms by which prior expectations can influence sensory representation of degraded speech. Expected features of the speech input can be enhanced or sharpened (Sharpened Signals), or in Predictive Coding accounts, expected features are suppressed and unexpected signals (Prediction Errors) are represented and processed further. The present work aimed at distinguishing between these two accounts. In two fMRI experiments we explore how prior knowledge (written text) influences perception of degraded speech. By combining behavioural, univariate and multivariate fMRI measures of how prior expectations lead to speech perception or speech deception, we provide evidence uniquely consistent with Prediction Error computations.

**Keywords:** speech perception; predictive coding; representational similarity analysis; fMRI

Accurate perceptual inference requires that sensory input and prior knowledge are combined to generate a representation of our environment. However, despite extensive experimental work in many sensory modalities the neural and computational mechanisms by which prior knowledge guides perception are unclear. One proposal is that neural representations of expected sensory signals are enhanced or tuned (Kok, Jehee, & de Lange, 2012). According to this account, perceptual representations are sharpened by relevant prior expectations as if the quality of the sensory input was increased. Alternatively, Predictive Coding schemes suggest that expected sensory input is explained away and unexpected information is represented in the form of Prediction Errors (Rao & Ballard, 1999). One intuitively attractive aspect of

Predictive Coding both for engineering and neuroscience is its assumption that minimal effort should be invested in representing or processing sensory information that is already known or expected.

## Representations of Prediction Error Not Sharpened Signals for Degraded Speech

In recent work, we were able to distinguish these two fundamental coding schemes for how matching prior expectations improve perception of degraded speech (Fig 1A/B, Blank & Davis, 2016). By combining behavioural, univariate and multivariate fMRI measures of sensory detail and priors during speech perception with computational modelling, we obtained evidence for Prediction Error computations. Prior expectations were manipulated by presenting written words before degraded spoken words. Sensory detail was manipulated by noise-vocoding spoken words (4- or 12-channel). Increasing the amount of sensory detail in vocoded speech and providing informative expectations had additive effects on behavioural and univariate neural measures. Both manipulations improved word report accuracy, and reduced BOLD signal in the left superior temporal sulcus (STS, Fig 1C). Conversely, they had interacting effects on multivariate speech representations in the STS. When prior knowledge was absent, increased sensory detail enhanced the amount of information in speech representations. However, with matching prior knowledge, increased sensory detail reduced the amount of information (Fig 1D). Computational simulations of Sharpened Signals and Prediction Errors could both explain these behavioural and univariate fMRI observations, only multivariate fMRI observations were simulated by a Prediction Error and not a Sharpened Signal model (Fig 1A/B).

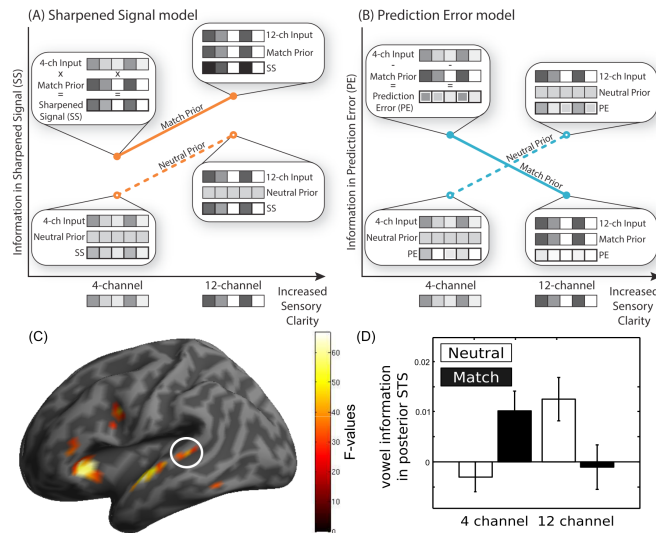


Figure 1: Two computational simulations of how matching and neutral prior expectations influence processing of degraded speech: (A) Sharpened Signal model, (B) Prediction Error model. (C) Reduced fMRI responses for degraded speech following matching text compared to neutral text, and (D) Multivariate fMRI analysis shows an interaction consistent with the Prediction Error model (from Blank & Davis, 2016).

## Speech Perception and Deception Explained by Prediction Error

In a second fMRI study, we assessed whether representations of Prediction Error can explain misidentification of degraded spoken words caused by misleading prior knowledge (speech deception). Prior expectations were manipulated by presenting written words before 6-channel vocoded spoken words in one of three conditions: (1) Match: written and spoken words were identical (e.g., kit – kit). (2) Total Mismatch: spoken and written word were phonologically unrelated (e.g., kit – ball). (3) Partial Mismatch: the spoken and written words had different final segments (e.g., kit - kick) or initial segments (e.g., kit - pit). Partially-mismatching trials were often incorrectly perceived (i.e., participants report that written and spoken words are identical) allowing us to compare neural responses when human listeners were deceived into falsely accepting or correctly rejecting prior expectations.

With data from partial mismatch trials we can assess whether perceptual outcomes and neural representations are better explained by: (1) Overlapping segments in written and spoken words (e.g., kit - kitsch, kip - kick both share the onset /k/; kit - pit and writ - wit both share the offset /t/) or (2) Mismatching segments (e.g., kit - pit, kitsch - pitch have

the same mismatching onset /k/-/p/; kit - kitsch, pit - pitch have the same mismatching offset /t/-/f/).

**Behavioural Results** Perception of partial mismatch pairs was inconsistent with listeners incorrectly responding “same” to 38.5% of these trials. The rate of correct perception and deception was more similar for word pairs that shared the same mismatching segments (i.e., kit - pit and kitsch - pitch) than for pairs that shared the same overlapping segments (i.e., kit - pit and writ - wit).

**Univariate fMRI Results** We again observed reduced fMRI responses for match compared to total mismatch trials (cf. Blank & Davis, 2016). Partial mismatch trials in which listeners responded “same” led to reduced fMRI responses in left posterior STS compared to trials in which listeners detected the mismatch.

**Multivariate fMRI Results** Representational similarity analysis of STS responses showed that mismatch representations but not overlap representations were related to rates of speech deception for single participants and single word pairs.

In combination this work demonstrates that perceptual outcomes (speech deception) and neural responses to degraded speech signals are best explained by computations that represent mismatch between prior expectations and speech input (i.e., prediction error)

## Acknowledgments

Funded by the UK Medical Research Council (MC-A060-5PQ80 to MHD) and the EU Horizon 2020 programme (703635, Marie Curie fellowship to HB).

## References

- Blank, H., & Davis, M. H. (2016). Prediction Errors but Not Sharpened Signals Simulate Multivoxel fMRI Patterns during Speech Perception. *PLOS Biology*, 14(11), e1002577.
- Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*, 75(2), 265-270.
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*, 2(1), 79-87.