

# Words Can Shift:

## Dynamically Adjusting Word Representations Using Nonverbal Behaviors



清华大学  
Tsinghua University



Yansen Wang, Ying Shen, Zhun Liu, Paul Pu Liang, Amir Zadeh, Louis-Philippe Morency  
Department of Computer Science, Tsinghua University  
School of Computer Science, Carnegie Mellon University



MultiComp Lab



Language  
Technologies  
Institute

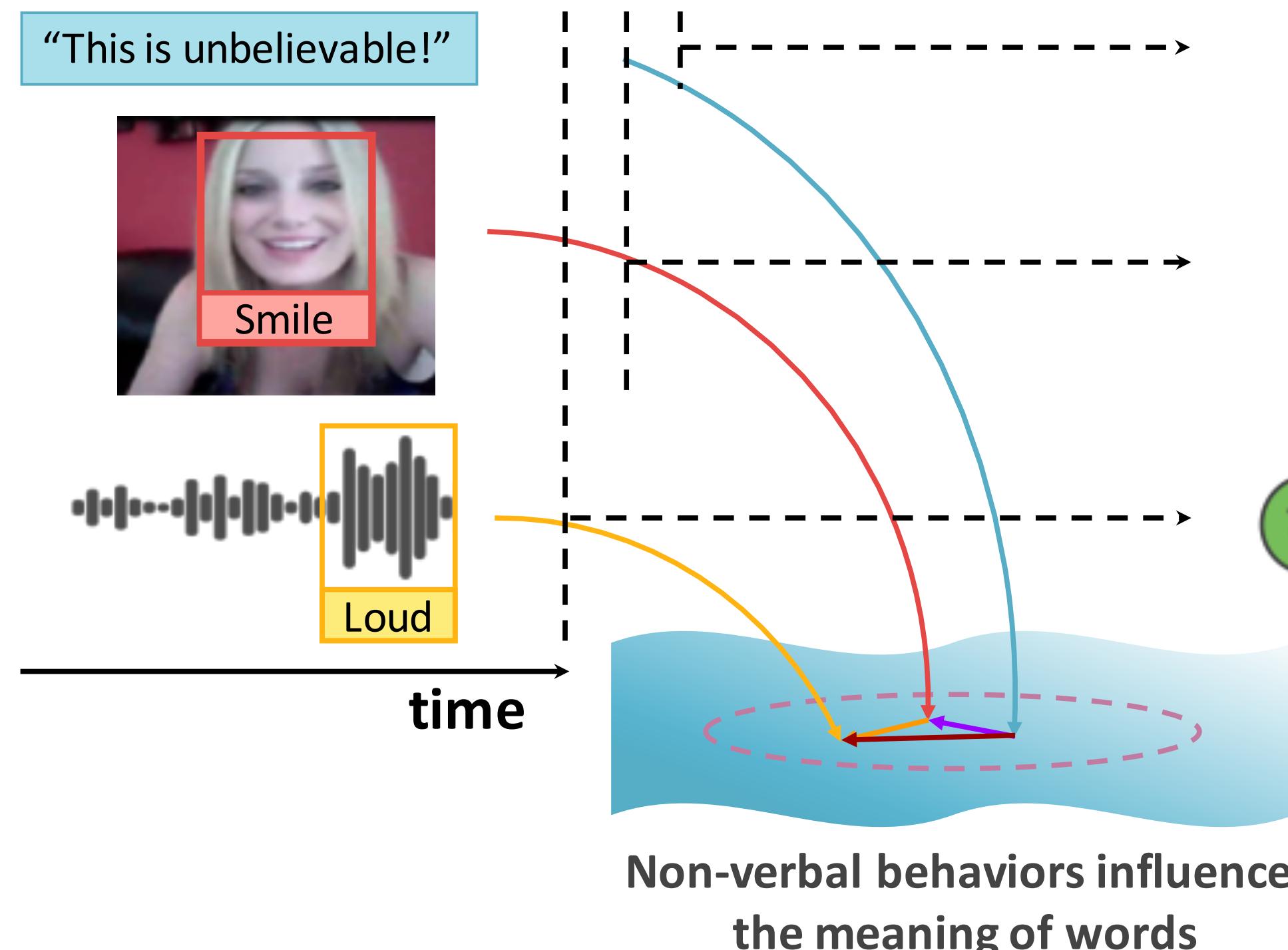


## Introduction

### Motivation

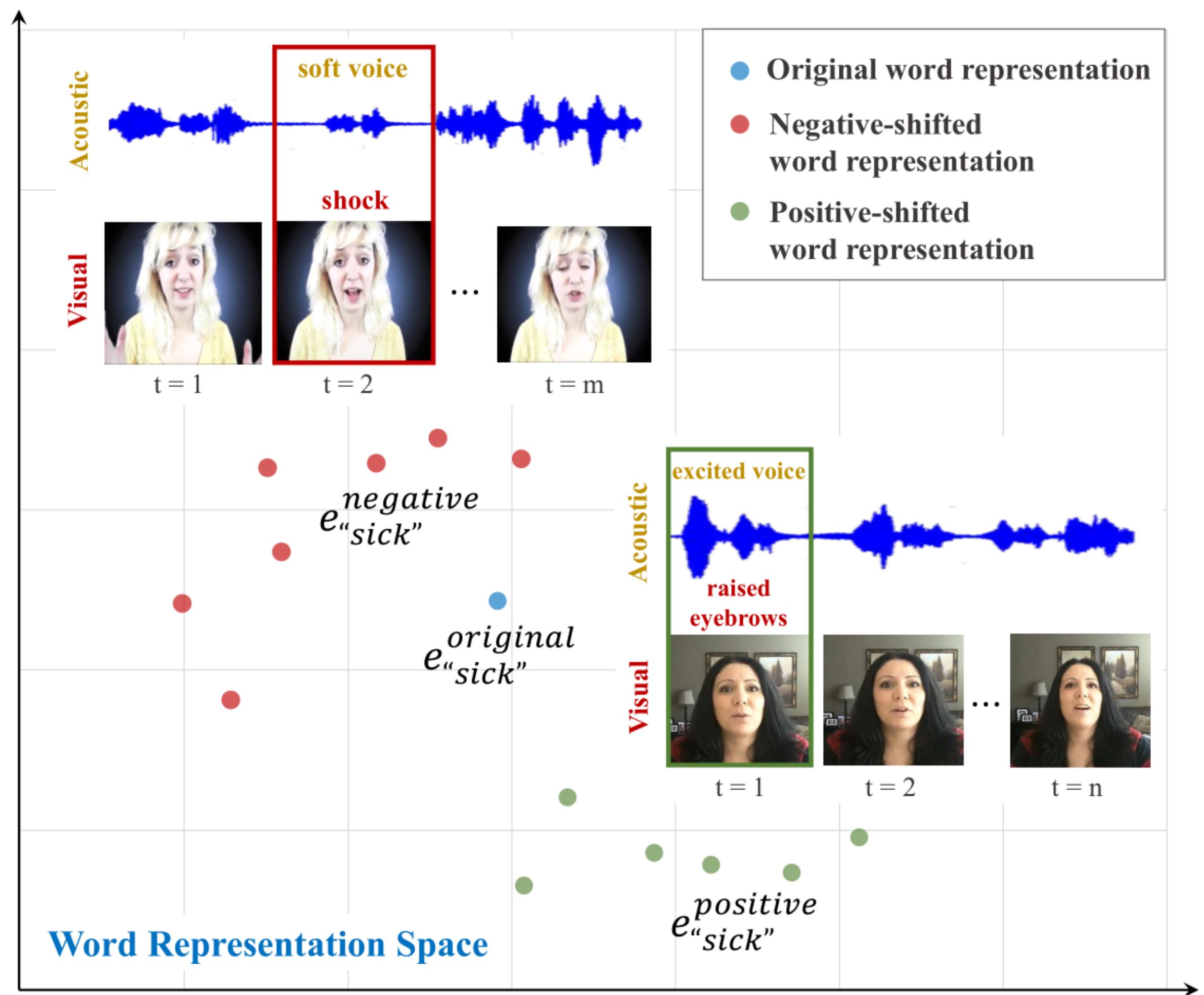
- Multimodal communication happens through two channels: **verbal** and **nonverbal**.
- Words exhibit **dynamic meanings** depending on different nonverbal contexts.

### Speaker's behaviors

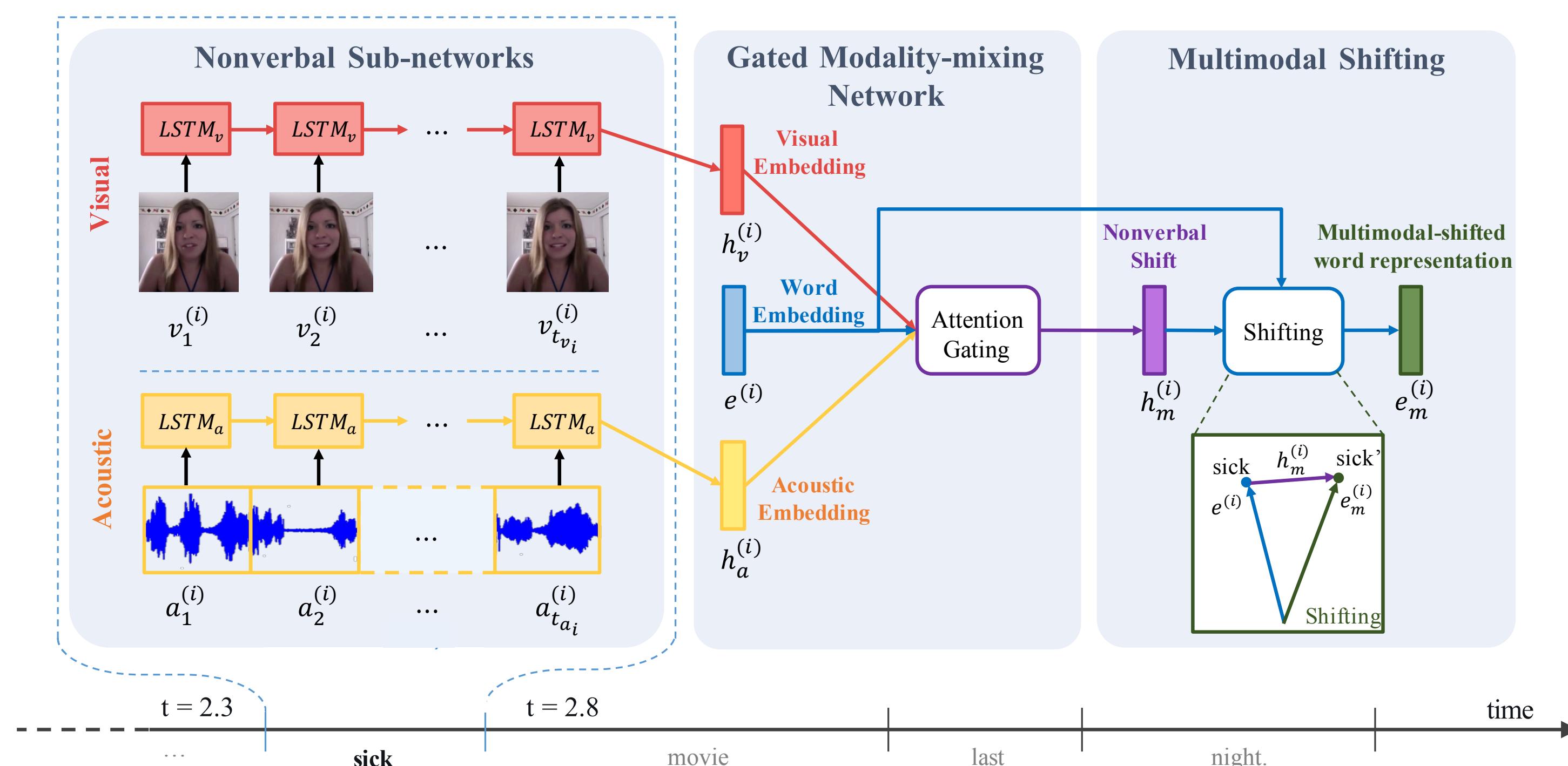


### Sentiment Intensity

## Motivating Example



## Recurrent Attended Variation Embedding Network (RAVEN)



## Results

### RAVEN: Recurrent Attended Variation Embedding Network

#### Variation Proposal Sub-networks

$L$ : the sequence of uttered words.

$L^{(i)}$ : the  $i$ -th word in the language modality.

$V^{(i)}, A^{(i)}$ : the visual and acoustic sub-word units.

Calculate the **Visual and Acoustic Variation Proposal**:

$$\begin{aligned} h_v^{(i)} &= LSTM_v(V^{(i)}) \\ h_a^{(i)} &= LSTM_a(A^{(i)}) \end{aligned}$$

#### Variation Attention Gates

Variation Proposals are weighted before assembled.

We denote the embedding of  $L^{(i)}$  as  $e^{(i)}$ .

$$w_v^{(i)} = \sigma(W_h h_v^{(i)} + W_e e^{(i)} + b_v^{(i)})$$

$$w_a^{(i)} = \sigma(W_h h_a^{(i)} + W_e e^{(i)} + b_a^{(i)})$$

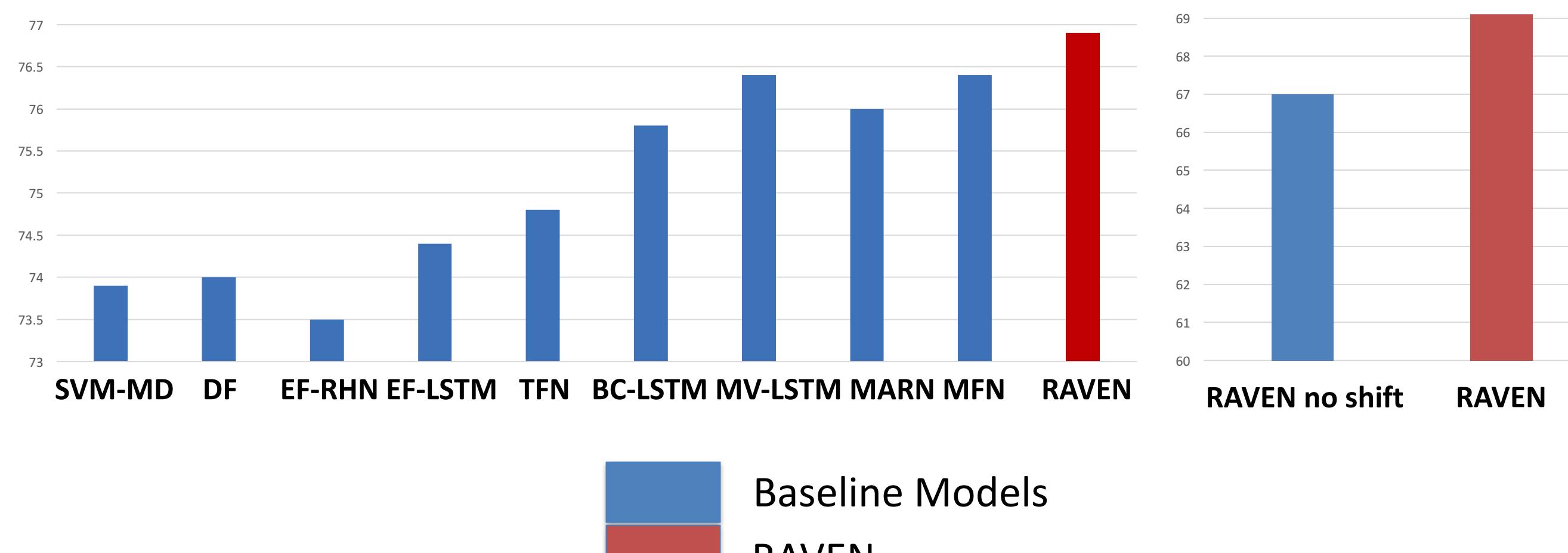
$$h_m^{(i)} = w_v^{(i)} \circ (W_v h_v^{(i)}) + w_a^{(i)} \circ (W_a h_a^{(i)}) + b_h^{(i)}$$

Here,  $h_m^{(i)}$  is the multimodal variation vector. Then we can get the final multimodal-augmented language representation:

$$e_m^{(i)} = e^{(i)} + \alpha h_m^{(i)}$$

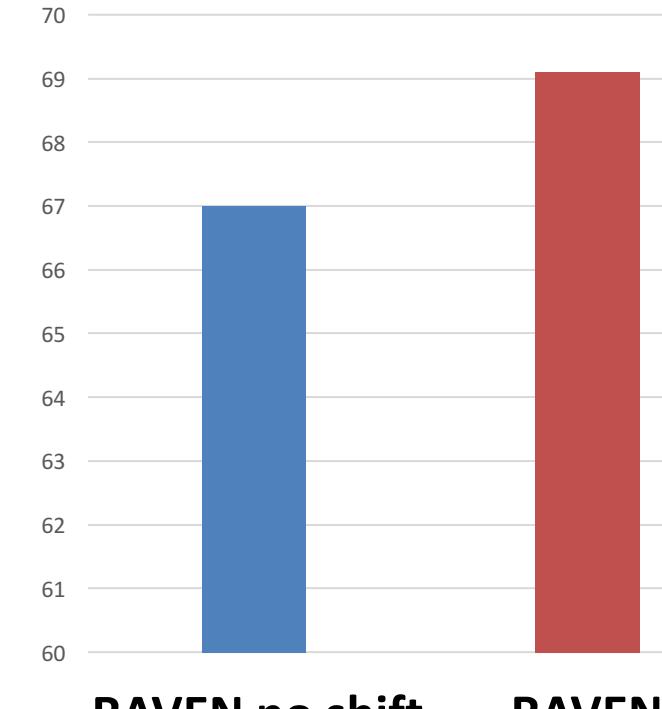
$\alpha$  is a hyper-parameter to control the effect of the variation vector.

IEMOCAP Emotion Recognition (Binary Accuracy)

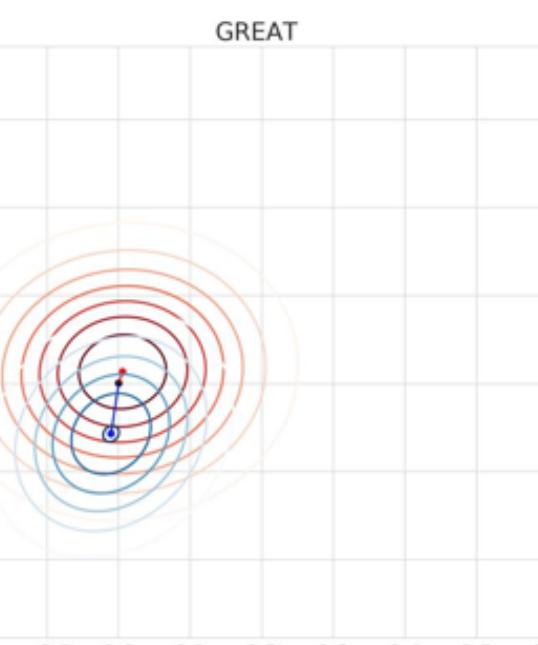


RAVEN allows us to visualize the nonverbal shifts on words

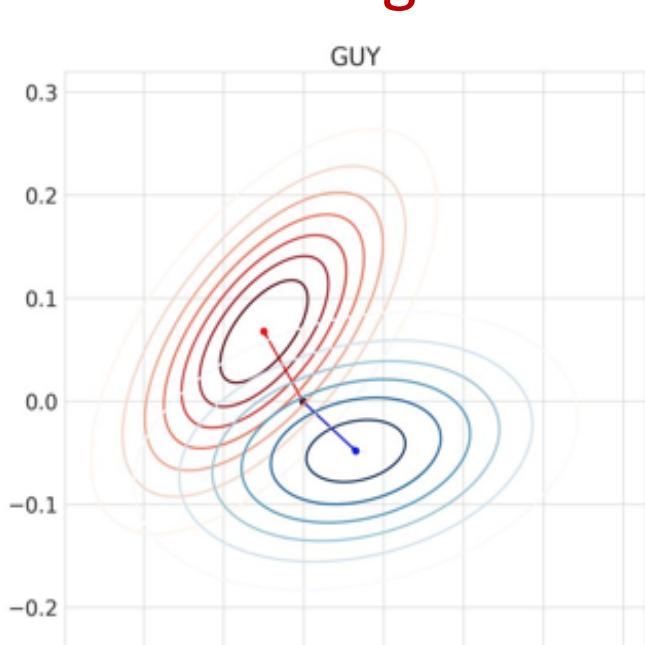
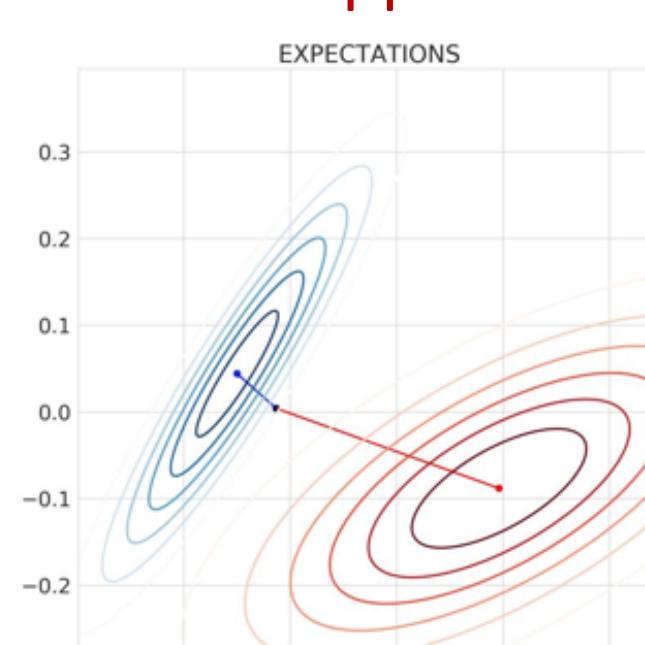
IEMOCAP Emotion Recognition (Binary Accuracy)



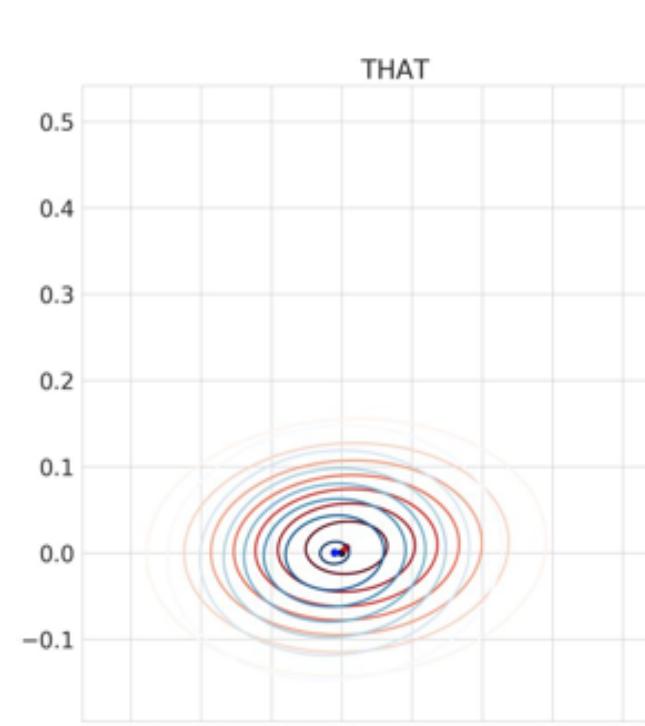
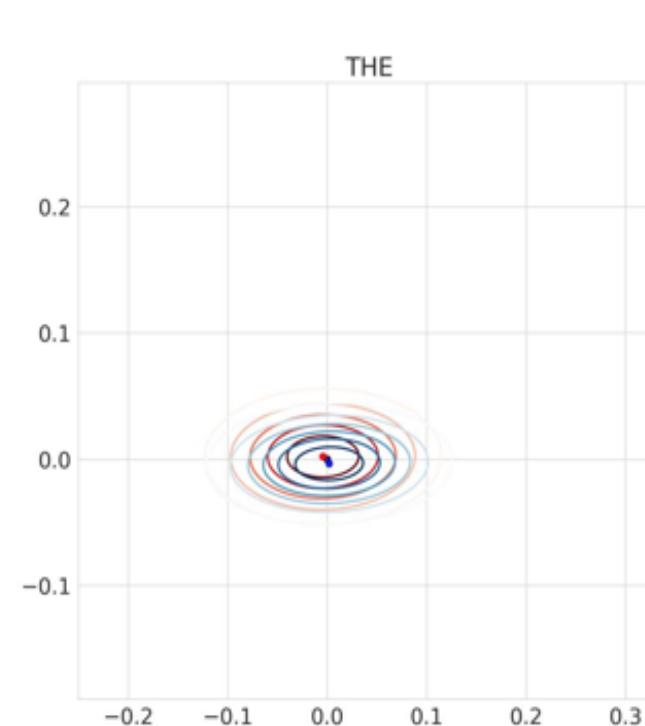
### Words with inherent polarity



Nouns that appear in both positive and negative contexts



Words not critical for expressing sentiment



Code and Models: <http://github.com/victorywys/RAVEN>