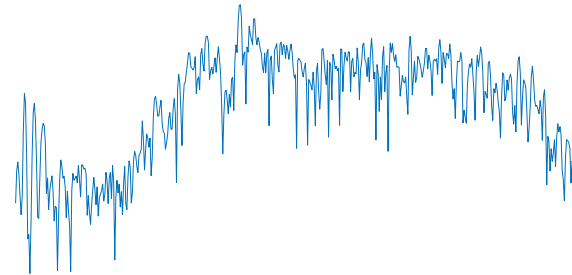
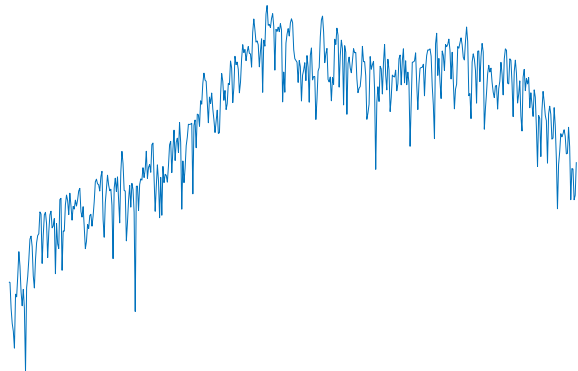
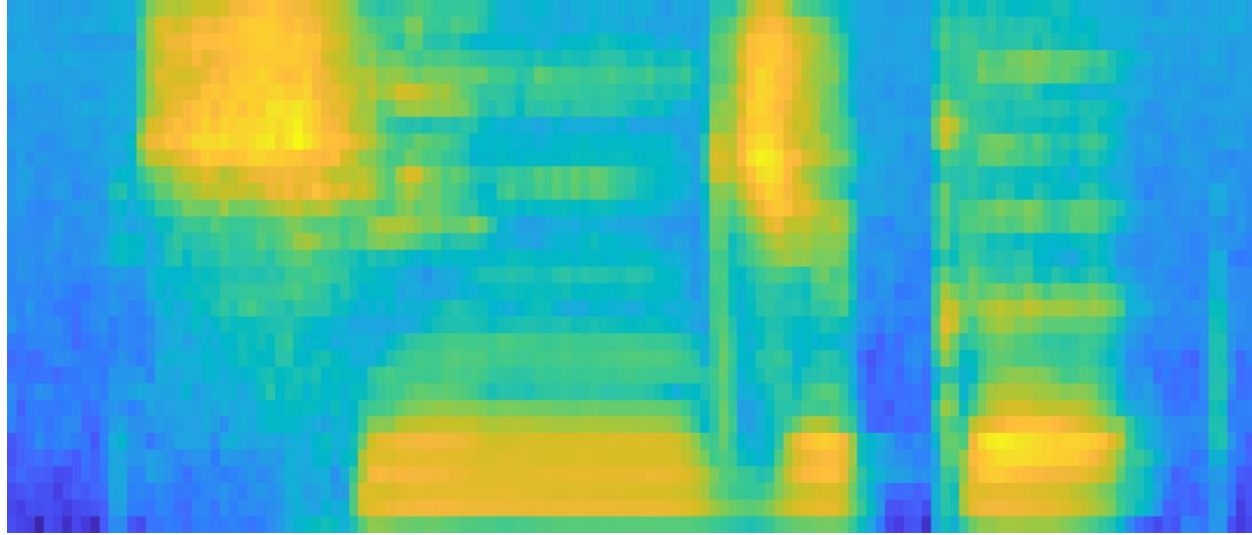


子音



ことばの特徴

- 語ごとに音韻列が決まっている

- 謝る /a y a m a r u/

× /a p y a m a r u/

× /a y a n a r u/

辞書

- 次に来る単語は文法で制限される

- 「今日」の次によく来る語 「今日/は」

「今日/も」

「今日/今日」 余りない

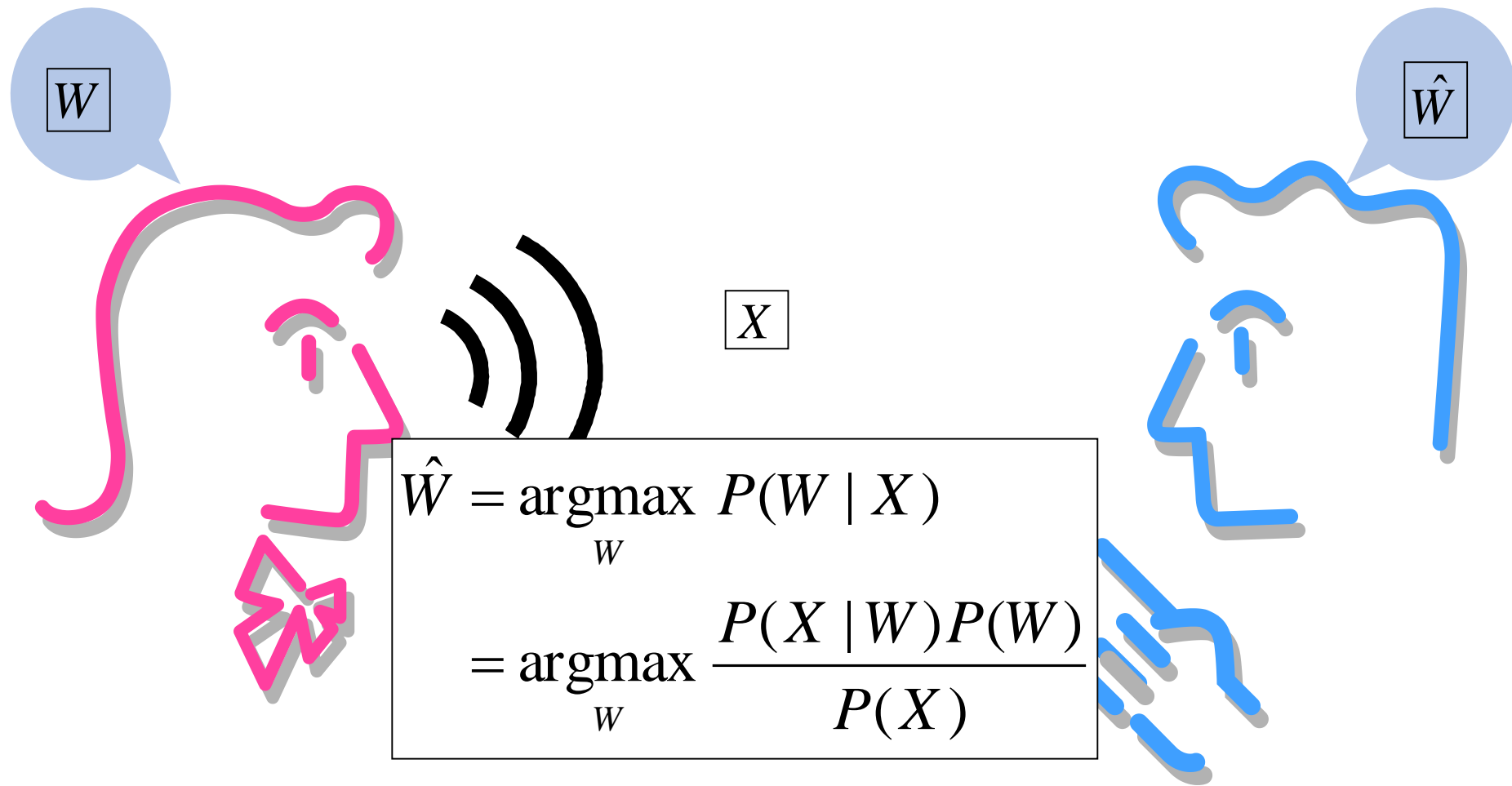
言語モデル

- 使われる語は分野で異なる

- 「代打」 野球に関する文には出る

サッカーに関する文には余り出ない

音声コミュニケーションの確率モデル



結果(X)から原因(W)を推測する

音声コミュニケーションの確率モデル

$$\hat{W} = \operatorname{argmax}_W P(W | X)$$

X 音声

$$= \operatorname{argmax}_W \frac{P(X | W)P(W)}{P(X)}$$

W 単語列

$P(W)$ 言語モデル

$P(X|W)$ W が発声されたときに X が観測される確率

音韻音響モデルから計算できる

ある音韻が発声されたときに X が観測される確率

音声コーパス

- テキストを発声した音声とそのテキストから構成される
- 全ての単語を含むわけではない
- 音声認識用
 - 多数の人がそれぞれ多数の文章を読み上げる
- 音声合成用
 - 一つの声に対し、一人が多数の文章を読み上げる
- ボランティアベースのものもある
 - Mozilla Common Voice

言語コーパス

- 大量の文章データを集めたもの
 - 言語モデルを作る

- 伝統的な言語モデル

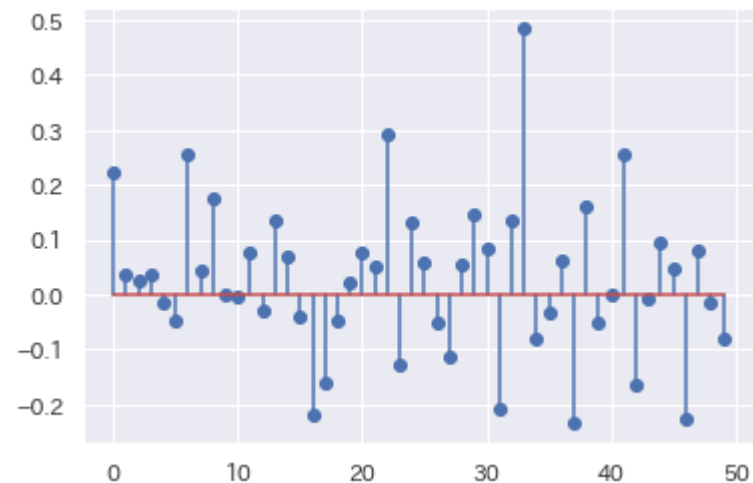
- N-gram

- Trigram

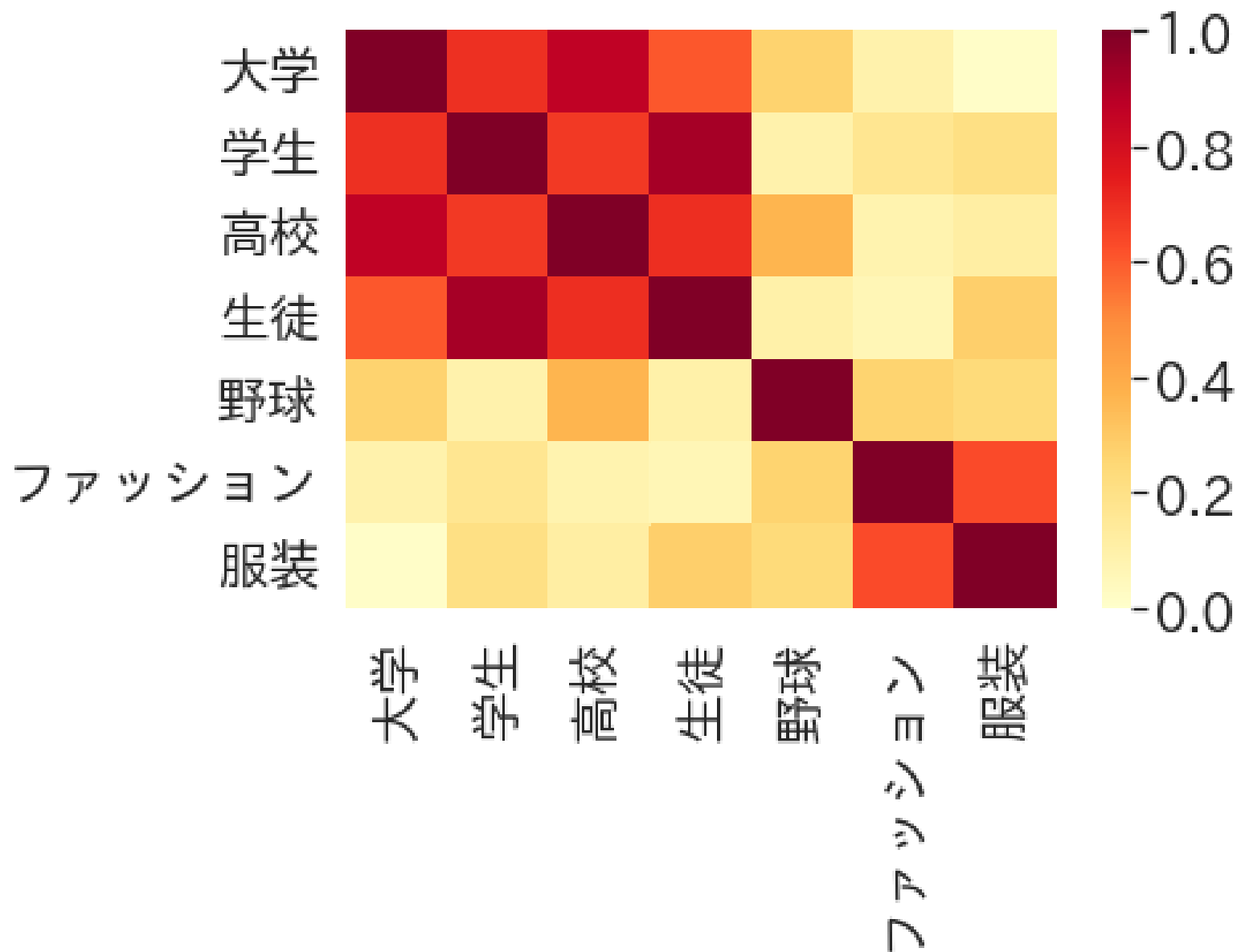
$P(\text{ますか}|\text{エアコン, ついて})$

- Word Embeddings

- 大量のデータで単語をベクトル化



ニュースのコーパスから学習した Word Embedding の例



まとめ

- 音声コミュニケーションの確率モデル
- 言語モデル
- Word Embedding