

# 情報理論は言語学に何を語るのか？

梶川康平

Ph.D. student

Department of Linguistics, Georgetown University  
kk1571@georgetown.edu



GEORGETOWN UNIVERSITY



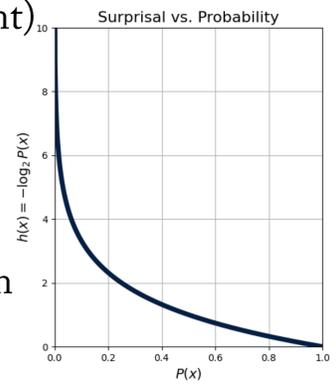
[Website]

## 情報理論とは

- 事象  $x$  の情報の「量」や「複雑さ」を data-driven で測る
  - 事象  $x$  の確率分布さえ用意できればいい
  - 複数事象  $x, y$  同士の関係・繋がり強さも測れる
- なぜいま？
  - (i) 大規模言語データ
  - (ii) 高性能計算機たち
  - (iii) 大規模言語モデル ...

➡ 従来よりはるかに尤もらしい確率推定ができる状況！！

- ▶ Surprisal (information content)
 
$$h(x) := -\log_2 P(x)$$
- ▶ Entropy
 
$$H(X) := \mathbb{E}_x [h(x)]$$
- ▶ Pointwise mutual information
 
$$\text{pmi}(x; y) := \log_2 \frac{P(x, y)}{P(x)P(y)}$$
- ▶ Mutual Information
 
$$I(X; Y) := \mathbb{E}_x \mathbb{E}_y [\text{pmi}(x; y)]$$

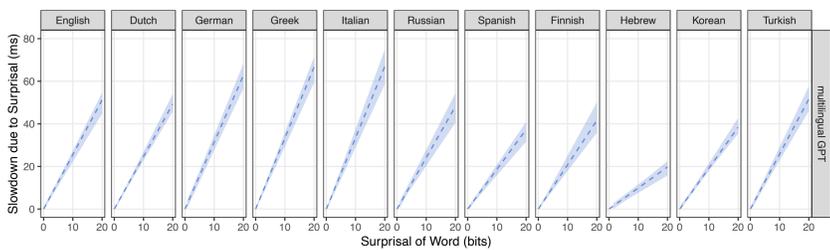


## 情報理論の強み弱み

### 理論中立的な指標になる 😊

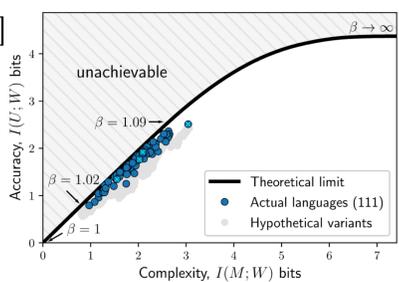
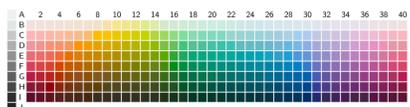
- 特定の表象やアルゴリズム、実装から独立
  - 「事象  $x$  が 3 bits」 = 「0と1を3つで書ける」
- ある単語  $w_i$  の処理コストは  $h(w_i | w_{<i})$  に比例する

[Hale 2001; Levy 2008; Wilcox et al. 2023]



- 色彩語彙は複雑さと正確さのトレードオフで形作られている [Zaslavsky et al. 2018]

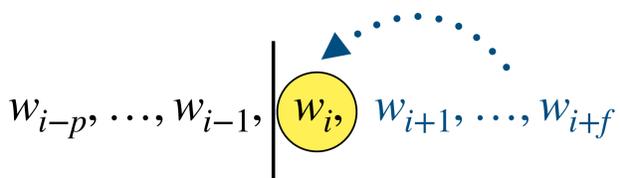
$$\mathcal{F} = I(M; W) - \beta I(W; U)$$



- 人間の言語獲得は1.5 MB。そしてそのほとんどが語彙の意味情報 [Mollica and Piantadosi 2019]

### 形式的に厳密なモデルを作るのに役立つ 😊

- 言語処理が合理的ならば、未来の情報  $w_i$  の処理に有益そうなとき、 $w_i$  の処理を遅らせる？



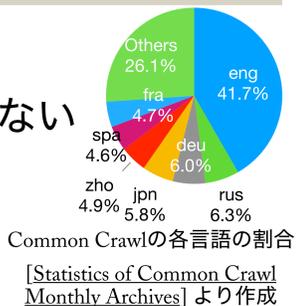
- 現在  $w_i$  の処理にどれだけ有益そうかは、Mutual Information で測れる [Kajikawa & Wilcox 2026]

$$\begin{aligned} \Delta \bar{h} &:= h(w_i | w_{<i}) - h(w_i | w_{<i}, w_{>i}) \\ &= \text{pmi}(w_i; w_{>i} | w_{<i}) \end{aligned}$$

$$\mathbb{E}_{w_i | w_{<i}} \mathbb{E}_{w_{>i} | w_i, w_{<i}} [\Delta \bar{h}] = I(W_i; W_{>i} | w_{<i})$$

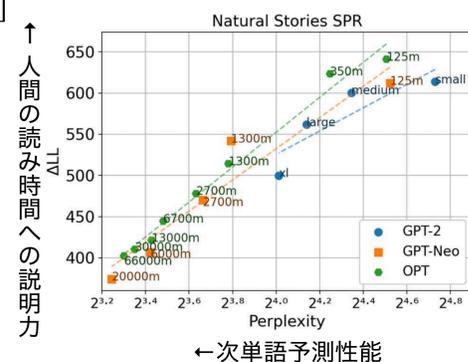
### 「技術 🤖」 周りの問題はたくさん 😞

- 情報量の推定は使う道具依存
  - 数式通りのことができるとは限らない
  - 巨大モデルの学習・推論は高価
  - 英語一強の言語資源
  - 「単語」の定義？ ; tokenization
  - 測れるのはテキストだけ？ [c.f. Wilcox et al. 2025]
  - 目的が手段になりがち
  - 当たり前のことを 🤖 を通して再確認してるだけ？



### 人間の言語って統計だけで語れるんですか 😞

- Super-human? 機械と人間の言語処理のズレ [Oh & Schuler 2023]



- 統語知識による予測は、読み時間のうち、 $h(w_i | w_{<i})$  とは違った部分を説明可能 [Kajikawa & Isono 2026]



- 人が判断する argumentらしさと、統計的依存関係は、そこそこは相関するが... [Kajikawa et al. in prep]

- *Put a book on the table*
- *Met her at the park*
- *Cut an apple with a knife*

