Expressions Causing Differences in Emotion Recognition in Social Networking Service Documents

Tsubasa Nakagawa, Shunsuke Kitada, Hitoshi Iyatomi | *Graduate School of Science and Engineering, Hosei University, Tokyo, Japan*

法政大学 HOSEI University

■ Summary

- Detection of expressions that cause **differences in the writer's and reader's perception** of emotion
- Identification of *hidden emotional expressions* unconsciously used by writes

■ Background

- With the increase in online communication via SNS, the discrepancy between the perceptions of writers' and readers' has become a problem

Reader 😄
Writer 😡

- Identifying the writer's true emotions and the expressions that cause differences in emotional perception is important
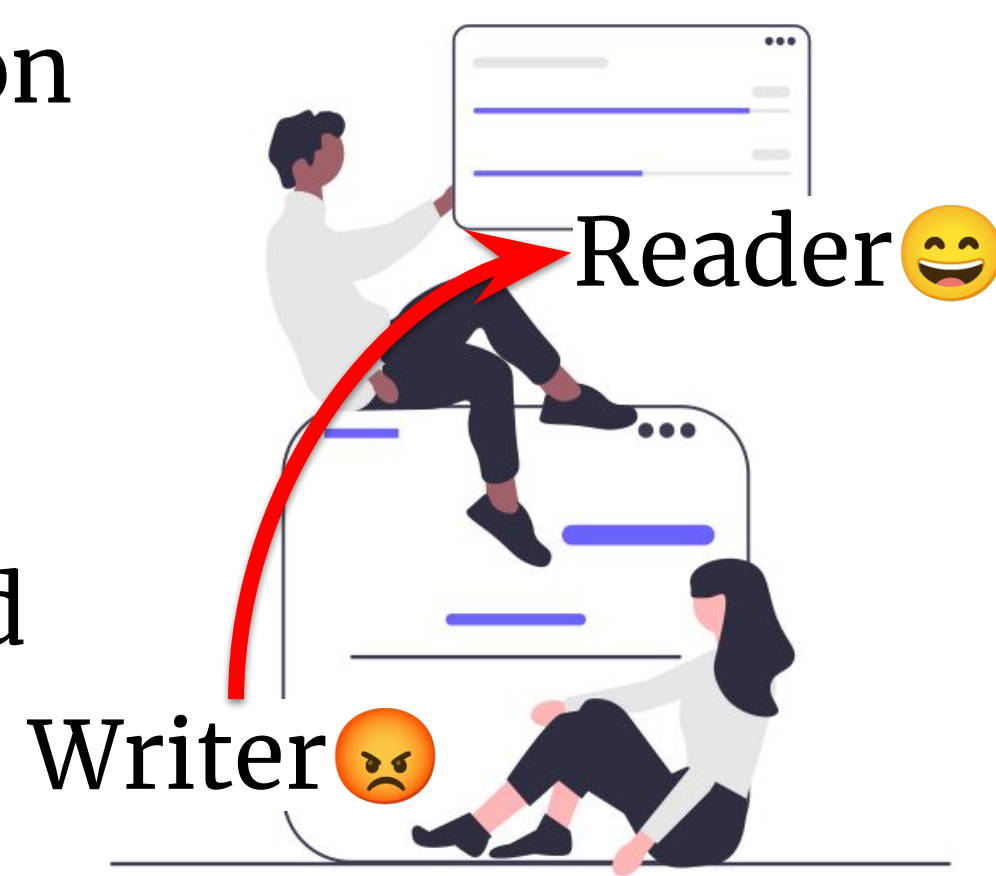
■ WRIME Dataset

WRIME: A New Dataset for Emotional Intensity Estimation with Subjective and Objective Annotation [Kajiwara+ NAACL'21]

- Japanese SNS-document dataset with 4 levels of emotional intensity annotated by the writer and 3 readers
  - 8 emotional labels are annotated (😄, 😭, 🤩, 😮, 😡, 😨, 😖, 🥰)

Examples of "joy 😄" emotional labels

Text 📝	🇯🇵 早寝するつもりが飲み物がなくなりコンビニへ。ん、今日、風が涼しいな。 🇺🇸 *I was going to go to bed early, but ran out of drinks and went to the convenience store. Hmm, the wind is cool today.*
Joy 😄	Writer: **1**　　Reader 1: **0**　　Reader 2: **0**　　Reader 3: **0**

- Total number of emotional labels with intensity of 2 or more in 4 levels (0 to 3)
  - Anger 😡 Writer: 3,040 ➜ Reader: 543
  - Trust 🥰 Writer: 5,167 ➜ Reader: 185

➜ **Readers tend to underestimate writers' emotions**

■ Proposed Framework

- This study focuses on anger, which has a large difference in emotion recognition between writers and readers in the dataset
- We define sentences in which the writer's anger intensity exceeded the reader's by 2 or more as *hidden-anger sentences*

Our proposed framework consists of two stages:

- *Prediction of hidden-anger sentences*
  - *Building a BERT-based detector:* fine-tuned a pre-trained the BERT on Japanese Wikipedia to predict hidden-anger sentences
- *Detection of hidden-anger expressions*
  - *Analyzing sentences obtained by the detector:* extracted the top 10 words with the largest differences in frequency of occurrence between hidden-anger sentences and other sentences

Examples of sentences that are extremely difficult to estimate emotionally

Text 📝	🇯🇵 雨の日1日1回は見るよねこれ 🇺🇸 *I see this at least once a day on rainy days.*
Anger 😡	Writer: **3**　　Reader 1: **0**　　Reader 2: **0**　　Reader 3: **1**

Text 📝	🇯🇵 アッヒョヒョ！！！ファ????????！！！！！！！！！！！！！！みたいな鳴き声しとる 🇺🇸 *It's making a chirping sound like "AHHHHHHHHH!!!! Fa????????!!!!!!!!!!!!".*
Anger 😡	Writer: **3**　　Reader 1: **0**　　Reader 2: **0**　　Reader 3: **0**

■ Results & Discussion

Performance of the detector

The BERT-based model showed effectiveness as a detector of hidden-anger sentences

AUC = 0.772
ROC curve for detecting hidden-anger sentences
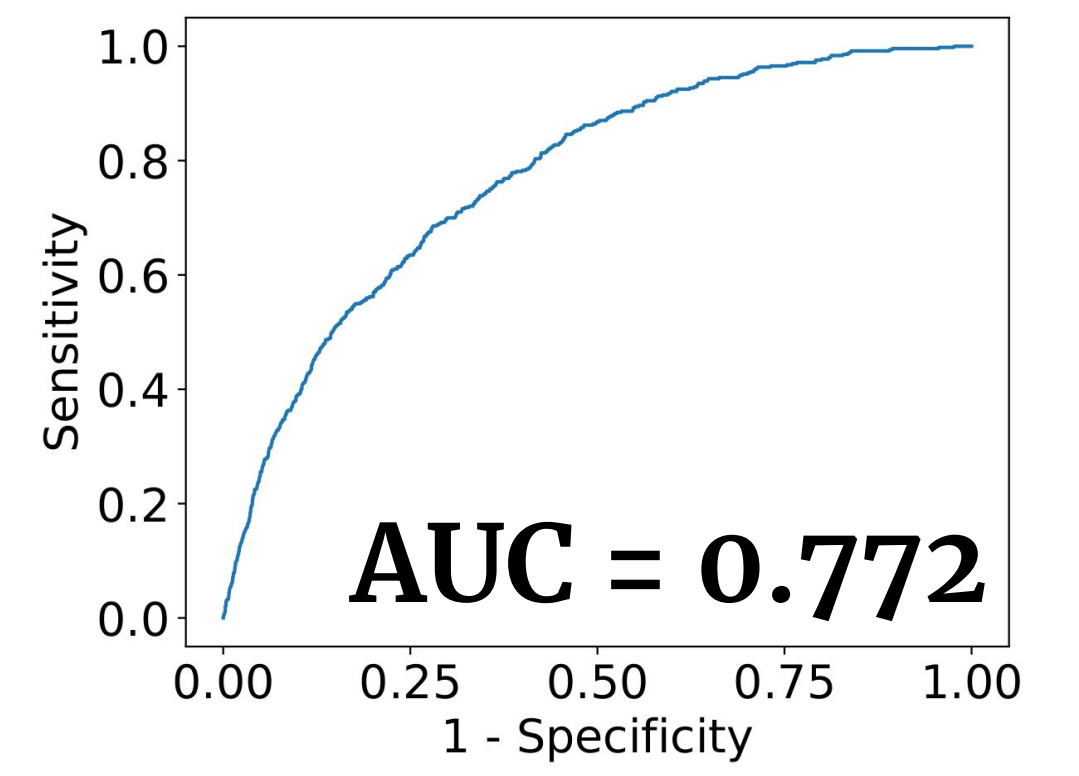
Comparison of (a) only label information and (b) our proposal

In (b), **our proposed framework**:

- The differences in their frequency are more clearly shown than in (a), indicating their level as *hidden-anger expressions*
- Pre-removal of inappropriate sentences enabled us to detect expressions that could not be detected in (a)

The difference in frequency of occurrence between hidden-anger sentences and other sentences
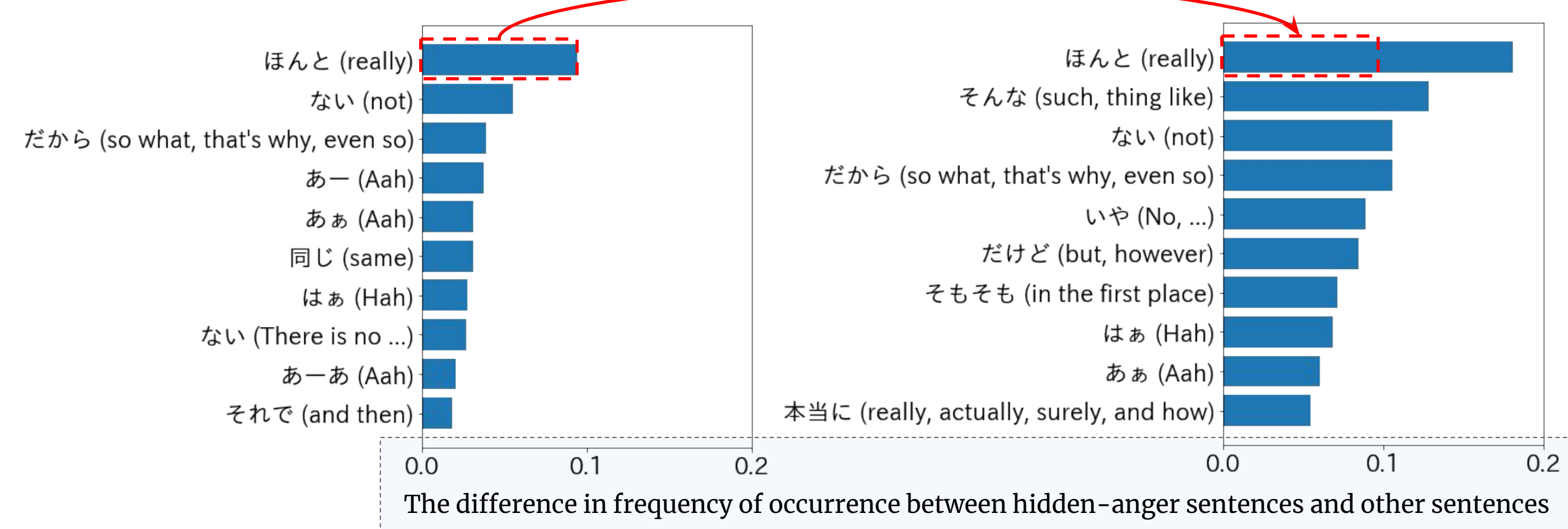
(a) Detection using only label information　　　(b) Detection using **our proposed framework**

Detected word-level hidden-anger expressions

Examples of hidden-anger sentences containing the detected words

Text 📝	🇯🇵 草取りと朝マラソンと持久走大会だけは**ほんと**解せなかった。 🇺🇸 *I **really** didn't understand the weeding, the morning marathon, and the endurance running competition.*
Anger 😡	Writer: **3**　　Reader 1: **0**　　Reader 2: **0**　　Reader 3: **0**

Text 📝	🇯🇵 マックでハンバーガーを注文したら店員に「ハンバーガーは無いです」と言われたので、**そんな**わけ無いでしょと思いながらもチーズバーガーを注文した 🇺🇸 *When I ordered a hamburger at McDonald's, the staff told me that they didn't have hamburgers, so I ordered a cheeseburger instead, even though I thought there was **no way** that was possible.*
Anger 😡	Writer: **2**　　Reader 1: **0**　　Reader 2: **0**　　Reader 3: **0**

Text 📝	🇯🇵 私のバッシュが売り切れになったんだっててさ…。**だから**、勝手に私の分がキャンセルになりやがった 🇺🇸 *Apparently, the basketball shoes I ordered were out of stock… And **so** my order was canceled without my notice.*
Anger 😡	Writer: **3**　　Reader 1: **1**　　Reader 2: **2**　　Reader 3: **0**

- These sentences do not involve explicit expressions of anger
  - ➜ It is difficult for readers to estimate the writer's anger
- However, a careful reading of each sentence, with an awareness of the possibility of hidden-anger due to our framework, reveals why the writer may have been angry

■ Conclusion & Future Work

- We proposed a framework for detecting expressions that may cause differences in emotion recognition between writers and readers
- Pre-removal of inappropriate sentences enabled us to detect expressions that could not be detected using only label information
- Sharing the findings of this study may ultimately reduce the frequency of mutual misunderstandings

Future Work

- Analysis using the writer's recent tweets as well

CIKM