

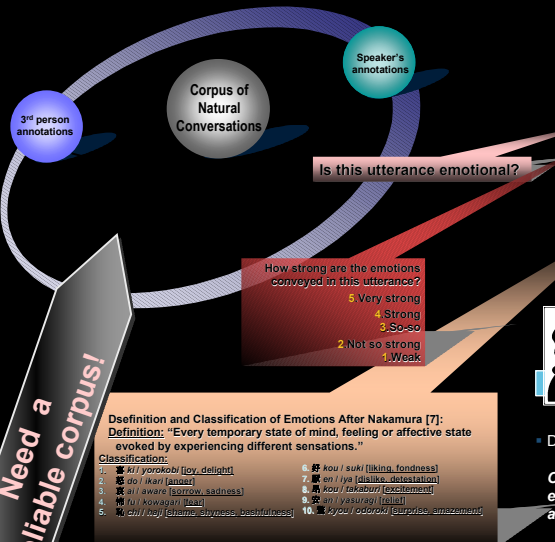
# 日本語の感情推定用コーパスの開発 Development of Corpus for Affect Analysis in Japanese

Michal PTASZYNSKI Pawel DYBALA Rafal RZEPKA Kenji ARAKI  
Graduate School of Information Science and Technology, Hokkaido University

## Abstract

This poster presents a design of a dialogue corpus required for the research on affect analysis in Japanese. The research on analyzing and estimating emotional states of users during their interaction with robots and generally perceived cyber-environment has been focused on developing sophisticated methods rather than gathering a relevant material for evaluation. In this paper we explain an urgent need for the change of this tendency and propose a design of a corpus meant to support research in this field. We also propose the first application of the corpus in the evaluation method proposed by us in the former research.

## Corpus Design for DSEM

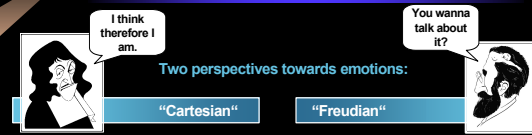


## Annotating the Corpus

### Three levels of annotations

1. Emotive / Non-emotive
2. Emotive Value
3. Emotion Types

### Two perspectives



Descartes about emotions: Only the one who expresses emotions knows exactly what they are.

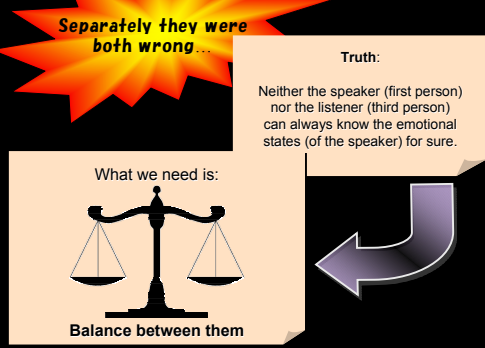
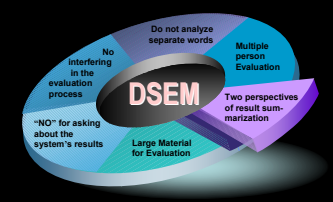
Freud about emotions: The one who expresses emotions know nothing about them – better ask the third person.

## Problems with Evaluation

1. Analyzing only the common words.
2. Performing important parts of the evaluation "manually" by the researchers.
3. Confusing commonsense with recognition.
4. "Are the system's results commonsensical?"
5. Small number of evaluators.

Need a reliable corpus!

## Double Standpoint Evaluation Method



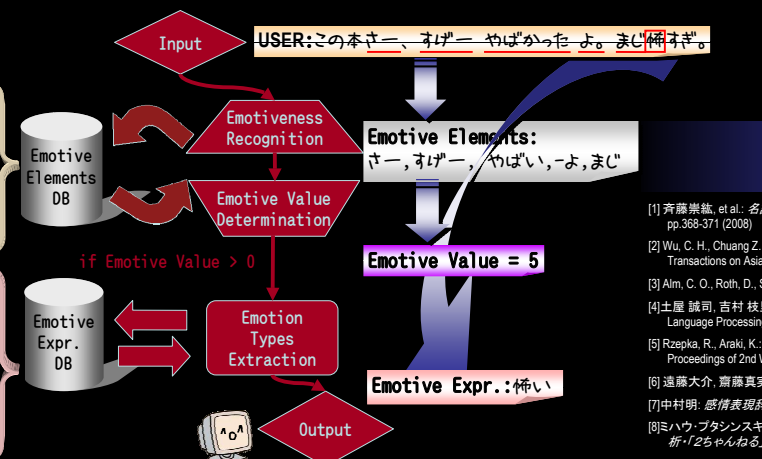
Urgent need for an objective method of evaluation for affect analysis systems!

Ptaszynski et al [9] Proposed a solution!

- Minimal analysis unit – one utterance
- Evaluators perform the same procedures as the system
- Compare the results of the system to the evaluators
- The more evaluators the better
- Many people create a robust Material for Evaluation
- Change the perspective according to what is evaluated

## ML-Ask System

- In language there are:
1. **Emotive expressions.** Parts of speech, that in emotive sentences describe emotional states.[7]
  2. **Emotive elements.** Indicating that emotions have been conveyed, but not detailing what specific emotions there are. The same emotive element can express different emotions depending on context.[8]
- Examples of emotive elements:
- nominals: 喜 (joy), 憂 (sorrow), 怒 (anger), 悲 (sorrow), 驚 (surprise), 怖 (fear), 恥 (shame)
  - verbs: 喜ぶ (be glad), 怒る (be angry), 悲しむ (feel sad), 驚かす (get angry)
  - phrases / idioms: 喜ぶが喜ぶ (be glad), 怒るが怒る (be angry), 悲しむが悲しむ (feel sad), 驚かす (get angry)
  - adjectives: 喜しい (happy), 怒い (angry), 悲しい (sad), 驚い (surprised), 怖い (scary)
  - exclamatives: すごい (great!), すごい (whoa!)
  - mimetics (gitaigo): ワクワク (heart pounding), ドキドキ (pit-a-pit)
  - vulgarity: やがる (fu'ing do stit), 馬鹿 (stupid)
  - hypocryis: ちゃん (nama stuff)
- textual representations of voice modulation and body language (emoticon):
- ↑ (怒), ↓ (喜), ○ (喜), □ (怒)



## Conclusions

Realization of this project we will:

- Acquire a reliable corpus for Affective Computing Research
- Provide a robust and objective evaluation material for Affect Analysis Systems in Japanese

Annotations not only on textual level, but with audio-video information annotated will provide wide applicability in world-wide research on emotions

## References

- [1] 齊藤崇, et al.: 名詞の感情属性の抽出とそれに基づく名詞間類似度の計算. 自然言語処理学会 Proceedings of NLP pp.368-371 (2008)
- [2] Wu, C. H., Chuang Z. J., Lin Y. C.: Emotion Recognition from Text Using Semantic Labels and Separable Mixture Models, ACM Transactions on Asian Language Information Processing, 2006.
- [3] Alm, C. O., Roth, D., Sproat, R.: Emotions from text: machine learning for text based emotion prediction, HLT/EMNLP, Vancouver, 2005.
- [4] 土屋 誠司, 吉村 枝里子, 渡部 広一, 河岡 司: 連想メカニズムを用いた話者の感情判断手法の提案. Journal of Natural Language Processing, Vol.14, No.3, 2007.
- [5] Rzepka, R., Araki, K.: What About Tests In Smart Environments? On Possible Problems With Commonsense In Ambient Intelligence, Proceedings of 2nd Workshop on Artificial Intelligence Techniques for Ambient Intelligence, ICAI'07, 2007.
- [6] 遠藤大介, 齋藤真実, 山本和英: 係り受け関係を利用した感情生起表現の抽出. 自然言語処理学会 Proceedings of NLP 2006.
- [7] 中村明: 感情表現辞典. 東京堂出版, 2004
- [8] マハウ・プタシンスキ: 前巻る言語, インターネット掲示板の上の日本語会話における感情表現の構造と記号論的機能の分析「いちやんねる」電子掲示板を例として, アダム・ミツキエウヰツキ大学, 2006
- [9] M. Ptaszynski, P. Dybala, R. Rzepka and K. Araki. "Double Standpoint Evaluation Method for Affect Analysis System." Proceedings of JSAI, 2008.