

Linguistic Expression of Emotion in Human-Machine Interaction: The NIMITEK Corpus as a Research Tool

Milan Gnjatović¹, Manuela Kunze¹, Xian Zhang¹, Jörg Frommer², Dietmar Rösner¹

¹ Department of Knowledge Processing
and Language Engineering
Otto-von-Guericke-University Magdeburg

² Psychosomatische Medizin
und Psychotherapie
Otto-von-Guericke-University Magdeburg

Bellagio, 6 October, 2008

- **The NIMITEK project (*Neurobiologically inspired, multimodal intention recognition for technical communication system*):**

- ❖ since the end of 2005, interdisciplinary research on issues in human-machine interaction.

- **The prototype spoken dialogue system:**

- ❖ a spoken dialogue system for supporting users while they solve problems in a graphics system (e.g., the Tower of Hanoi puzzle),

- ❖ implements multimodal recognition of emotions and an adaptive dialogue strategy for supporting users.



- **The NIMITEK corpus:**

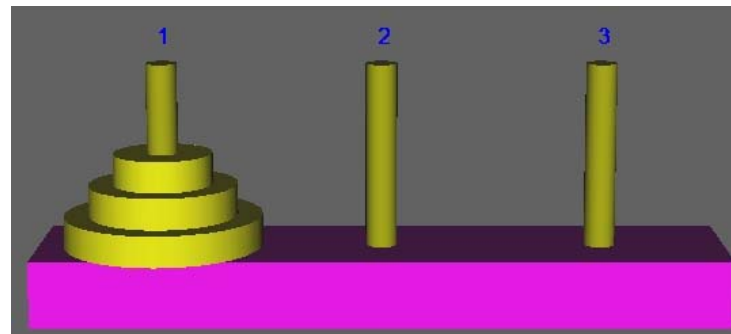
- ❖ 15 hours of audio/video recordings of affected speech and facial gestures in human-machine interaction,
- ❖ 10 subjects (7f, 3m), age from 18 to 27 (mean 22.0),
- ❖ language: German.

- **Proper attention was devoted to the issue of ecological validity:**

- ❖ genuine everyday emotions, extended in time and modality.

- **Wizard-of-Oz experiment:**

- ❖ subjects were given an illusion that they are communicating with a computer system, while a human operator played the role of the system,
- ❖ subjects were instructed that they undergo an intelligence test supported by a spoken natural language dialogue system,
- ❖ **in fact:** subjects were confronting a set of graphically based tasks, that was never intended to meet the criteria of an intelligence test.



- **Evaluation of the NIMITEK corpus with respect to its emotional content:**
 - ❖ performed in two phases.
- **Phase 1 – Data-driven evaluation:**
 - ❖ 6 evaluators (3 German native speakers and 3 non-German speakers),
 - ❖ evaluators were advised to introduce labels according to their own perceptions,
 - ❖ goal: to demonstrate the level of ecological validity of the corpus.
- **Introduced labels:**

nervousness interested
confused disappointed pleased
surprised anger fear
 insecure boredom joy
 accepting
thinking contentment sadness
 neutral stressed impatient

Phase 2 – Evaluation with predefined labels:

- ❖ 5 evaluators (native German speakers),
- ❖ evaluators used labels from the ARISEN model,
- ❖ goal: demonstrating the appropriateness of the ARISEN model.

<i>Introduced labels</i>	<i>Classes</i>
anger, nervousness, stressed, impatient	Annoyed
fear, insecure, confused	Retiring
sadness, disappointed, accepting, boredom	Indisposed
joy, contentment, boredom	Satisfied
thinking, surprised, interested	Engaged
neutral	Neutral

- **Multimodal emotion recognition in the NIMITEK project:**

- ❖ Other projects of the NIMITEK consortium work on detecting emotion from the prosody and from the facial expressions of the user.

- ❖ Here, we report about analyses of emotional content from the linguistic structures in the transcribed conversations.

■ Specific key words and phrases:

- ❖ Annoyed: *Sh*t (Sche*ße), stupid (blöd), I've had enough of it (Es reicht mir), ...*
- ❖ Retiring: *I don't understand it (Ich versteh' das nicht),
It's not working at all (Das geht doch gar nicht), ...*
- ❖ Indisposed: *I am going now (Ich geh' gleich),
I don't feel like doing any more (Ich hab' kein' Bock mehr), ...*
- ❖ ...

■ Ellipsis-substitutions:

- ❖ a form of anaphoric cohesion in a discourse, where we presuppose something by means of what is left out, e.g.:

*Why are you **moving** it on peg 2? Why? Why are you **doing** this step?
(Warum **fährst** du auf Säule 2? Warum? Warum **machst** du diesen Schritt?)*

- **The choice of lexical items to create cohesion in the discourse:**

- ❖ Simple repetition:

What is the problem? What is the problem?
(Was ist das Problem? Was ist das Problem?)

- ❖ Repetition with remark:

*Left up. Left up. Left up. **I said** left up.*
*(Links oben. Links oben. Links oben. **Ich habe gesagt** links oben.)*

- ❖ Reformulation:

***Not true** at all. That's **definitely wrong**.*
*(Gar **nicht wahr**. Das **stimmt gar nicht**.)*

- **Questions:**

- ❖ Containing ellipsis-substitution:

What are you doing? (Was tust du?)

Why are you doing this? (Warum machst du das?)

- ❖ Relating to a concrete action that was or should be performed by the system:

Why don't you move the 8 to left?

(Warum schiebst du die 8 nicht nach links?)

- ❖ Containing a reference to a previous utterance :

But there is one more, isn't there?

(Aber es gibt noch eins, oder?)

■ Negation:

❖ Simple negation:

*Right up. **No**, right up.*

*(Rechts oben. **Nein**, rechts oben.)*

❖ Negation with enhancement:

***No**, it is not right, It is **simply** not right.*

*(**Nein**, das stimmt nicht. Das stimmt **einfach** nicht.)*

❖ Negation with confirmation:

*Do I mean it? **No**, I don't mean it, **do I**?*

*(Meinte ich das? **Nee**, das meinte ich nicht, **oder**)*

❖ ...

- **We used the UIMA framework**

- ❖ to implement a first prototype of an automatic annotator for recognition of the user emotional state from linguistic information.

- **We used regular expressions:**

- ❖ to describe specific key words and phrases and additional features from the transcriptions of the dialogues from the corpus.

- **All these patterns are assigned to a concrete class of the ARISEN model.**

- ❖ The selection and the assignment of words and phrases were performed independently from the voting of the human evaluators.

- ❖ Only the transcriptions of observed experimental sessions were used, not video recordings.

- **Evaluation material:**

- ❖ A part of a WOZ experimental session from the NIMITEK corpus.
- ❖ Duration 45 min.; divided into 55 evaluation units.

- **We compare the results of the automatic annotator with the results of the majority voting of the German group of evaluators.**

- **Performance of the annotator for the given 6-classes emotional model ARISEN (Annoyed, Retiring, Indisposed, Satisfied, Engaged, Neutral):**

<i>Hit</i>	<i>Miss</i>	<i>False hit</i>
31.70%	34.35%	33.92%

- **Evaluation using a 3-class model (positive, neutral, negative):**

<i>Hit</i>	<i>Miss</i>	<i>False hit</i>
51.20%	33.67%	17.26%

- **The NIMITEK corpus:**

- ❖ Genuine (i.e., non acted) emotions, extended in time and modality.
- ❖ Available upon request.

- **Analyzing only key words and phrases does not suffice for emotion recognition:**

- ❖ No prosodic information.
- ❖ Ambiguity in defining emotional keywords and phrases.
- ❖ Ambiguity in syntactic and semantic information.

- **Lines of research to improve the performance of the automatic annotator:**

- ❖ Integration with the prosodic classifiers and the facial expression classifier.
- ❖ Considering additional linguistic features (e.g., information about structure of dialogue acts, context, lexical information, etc.)
- ❖ (Corpus specific:) Considering linguistic information from an additional source – the wizard.

Thank you for your attention!

<http://wdok.cs.uni-magdeburg.de/>