



Progress to a VOCA with Prosodic Synthesised Speech

Jan-O. Wülfing M.A. & Prof. Elisabeth André



Human Centered Multimedia
Institute of Computer Science
Augsburg University
Universitätsstr. 6a
86159 Augsburg, Germany



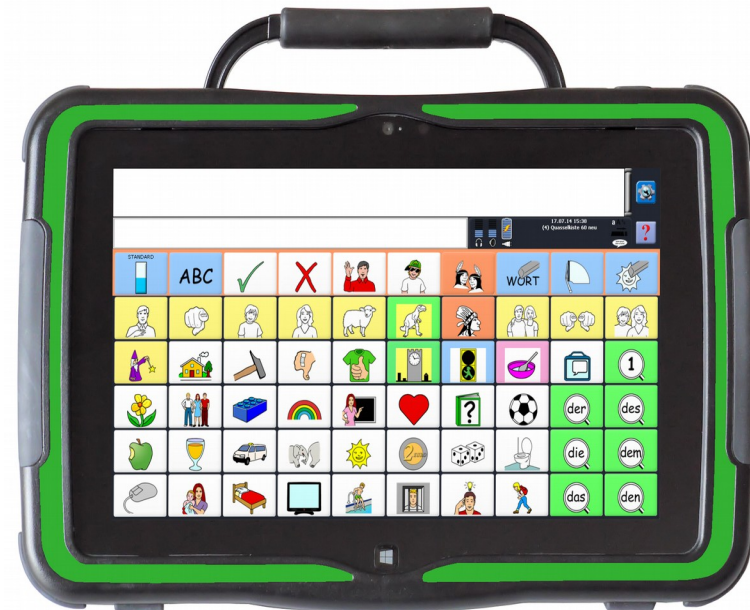
Don't forget to turn your mobile
on after this session



Agenda

- definition
- problem
- why emotional utterances
- hypothesis
- techniques used
- study
- results
- further steps

defined as „all forms of communication (other than oral speech) that are used to express thoughts, needs, wants, and ideas.“ [1]



by courtesy of REHAVISTA GmbH



Problem

though, current VOCAs sound intelligent,
they miss the ability of expressing the
user's emotion



why are they so important?

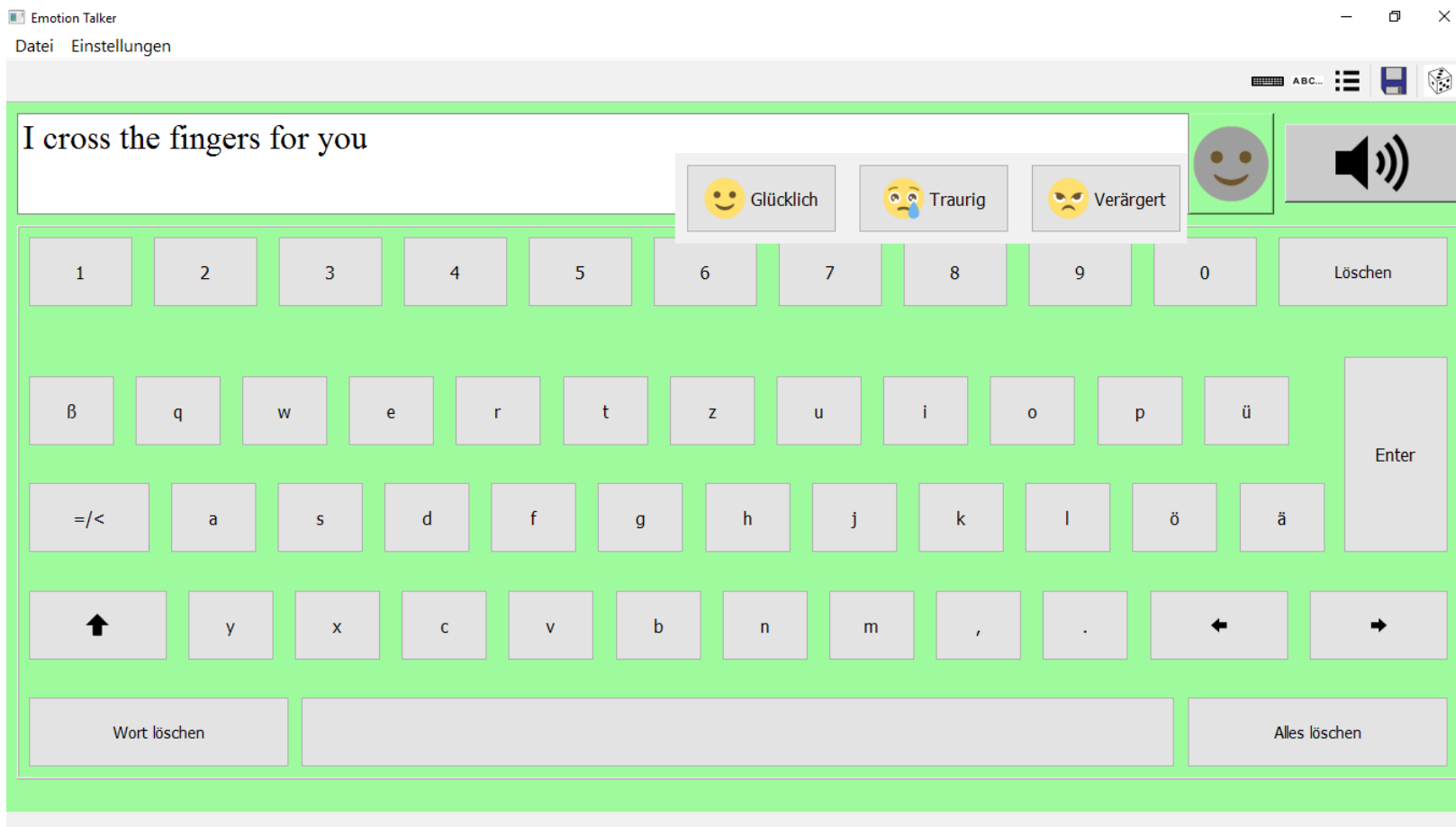
“Children have low development of their emotional competence if they cannot express themselves orally.” [2]

„I want to be able to sound sensitive or arrogant, assertive or humble, angry or happy, sarcastic or sincere, matter of fact or suggestive and sexy.“ [3]



Hypothesis

- interlocutor is more motivated to listen
- individual with CCN participate in lasting longer conversation



front-end: the EmotionTalker



Techniques used (2)

- Python 2.7
- Qt 5.6
- Cereproc's LTD.
Cerevoice synthesiser SDK 3.0
- Speech Synthesis Markup Language 1.1



Techniques used (3)

Speech Synthesis Markup Language [4]

```
<?xml version='1.0'?>
```

```
<parent>
```

```
<prosody pitch="high" rate="fast"  
volume="+60">
```

```
"I cross the fingers for you"
```

```
</prosody>
```

```
</parent>
```

happy

sad



Study (1)

- interview
asking participants, for example,
in which situations they would like to
utter emotional
- implementation
- testing/ interview



Study (2)

following an inspiration of the User-Sensitive Inclusive Design [5]

- due to the variety of impairments, finding a suitable way to present ET
- visiting individuals with CCN in their convenient environment, instead in our lab



Study (3)

	Sex	Age	Disability	Communication Method
P1	f	8	CP	Tobii C15
P2	f	10	CP	Accent 1000
P3	m	15	CP	Tobii I12
P4	f	45	CP	EcoTalker
P5	f	56	ASD	Faciltated Communication



child with CCN is playing Ludo
using the ET



Results

- it was not always easy to recognise the subtle emotions produced by the synthetic speech output
- overall, the participants agreed: a VOCA like ET would be a great benefit
- P2 said, she enjoyed telling her mom how lovely she is – in a happy sound
- P3 said, he loved to talk about football in an excited and louder manner



Further steps

- investigation of speeding up the input
- employing other modalities, e.g. facial expression recognition, to enhance user's emotional utterance by prosodic synthetic speech



- [1] American Speech and Hearing Association
asha.org/public/speech/disorders/AAC
(accessed 05/17/18)

- [2] Blackstone, S.W. & Wilkins, D.P., (2009).
Exploring the Importance of Emotional Competence
in Children With Complex Communication Needs.
In: Perspectives on Augmentative and Alternative
Communication (78-87).

- [3] Portnuff, C. (2006). AAC - A User's Perspective.
aac-lerc.psu.edu/index.php/webcasts/show/id/3
(accessed 05/17/18)



- [4] Speech Synthesis Markup Language – SSML
www.w3.org/TR/speech-synthesis11 (accessed 05/24/18)
- [5] Newell, A.F., Gregor, P., Morgan, M. et al.
User-Sensitive Inclusive Design.
Univ Access Inf Soc (2011) 10: 235-243.
<https://doi.org/10.1007/s10209-010-0203-y>



thank you for your attention

questions are very welcomed