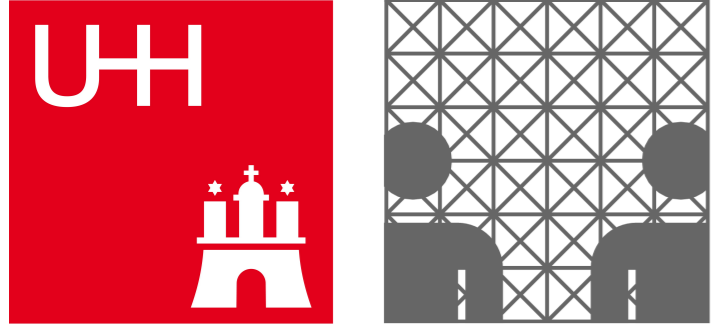


Feedback in Adaptive Interactive Storytelling



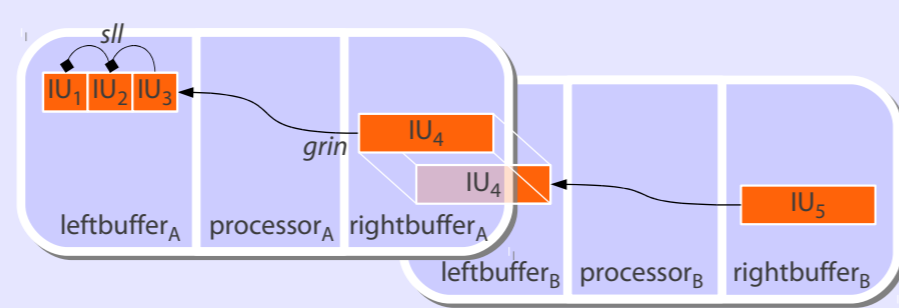
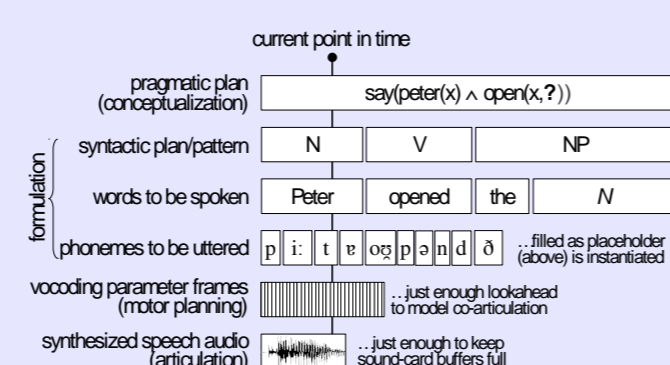
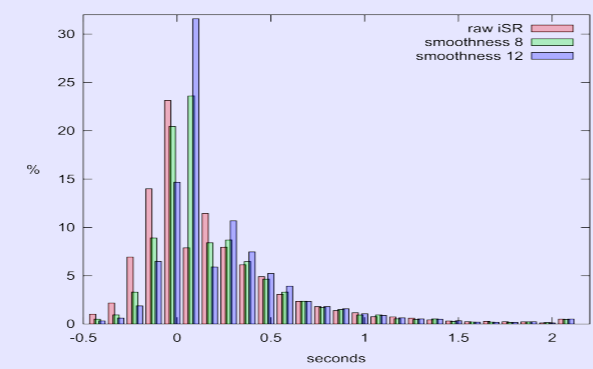
Timo Baumann • Natural Language Systems Division • Department of Informatics
Hamburg University • Germany • baumann@informatik.uni-hamburg.de

Objective

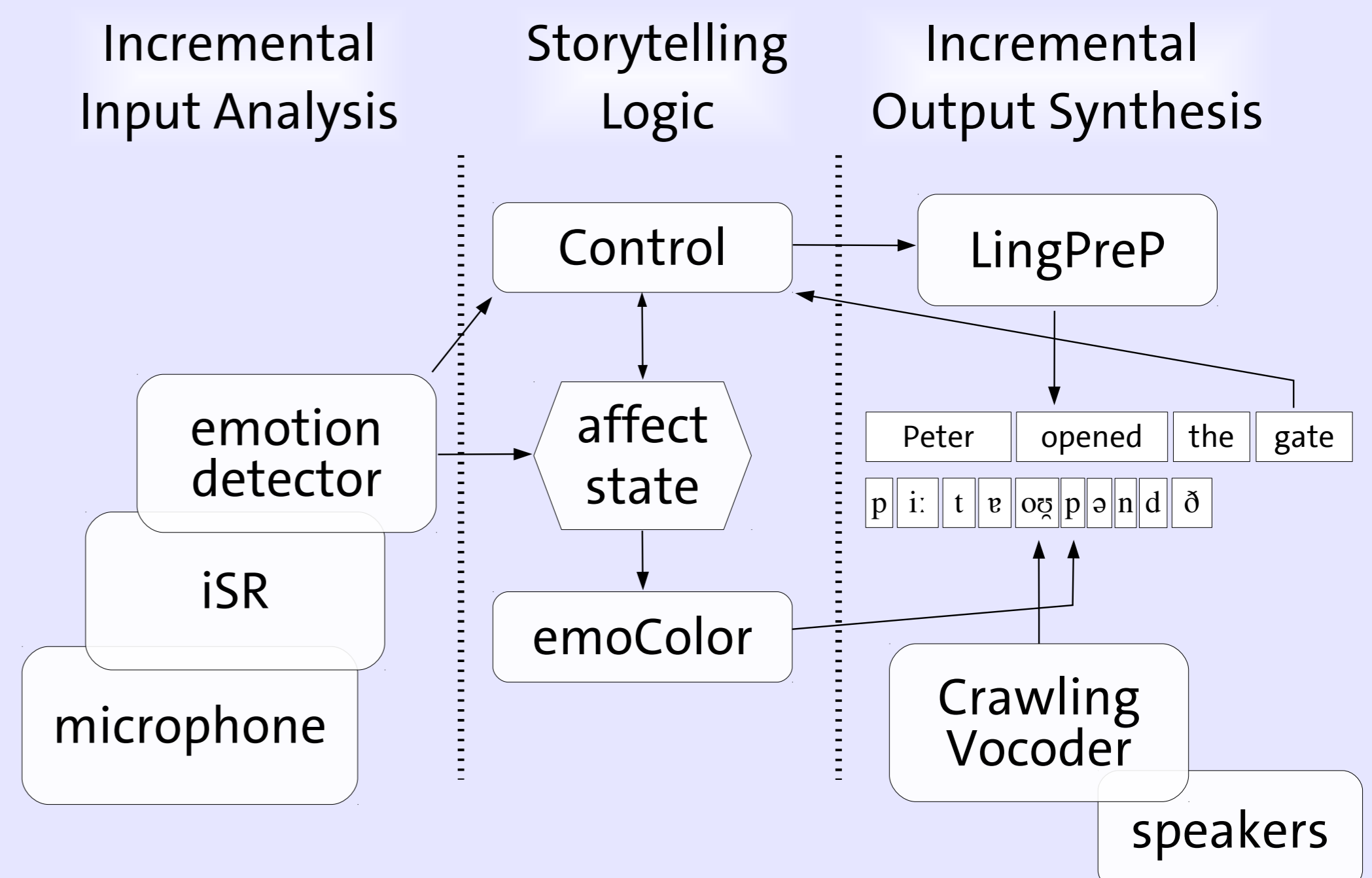
- Spoken Dialogue Systems (SDSs) are rarely confronted with feedback utterances
 - several reasons: SDSs don't produce feedback, turns are short and precise, users don't assume feedback to be interpreted by the system, ...
 - most importantly: the focus is on the content
 - thus, there is little data on feedback behaviour towards SDSs.
 - Storytellers, in contrast, invite their listeners to give feedback and react on it
 - the focus is on delivery
 - reactions mostly prosodically/emotionally
 - but also embellish parts of the story or change it altogether
- *Interactive Storytelling*
- We propose **Interactive Storytelling as a setting for learning more about feedback behaviour in human-computer interaction**. In this domain,
 - the system can actively elicit feedback, to increase the feedback rate
 - it can be used to study how reactions to feedback feed back to the user
 - dialogue management requirements are very low (the story is given)
 - purely system-driven domain

Existing Components for Very-low Latency Dialogue Processing

- Incremental Speech Recognition (iSR) [1] outputs partial hypotheses about user input while the user is speaking
 - stability estimates enable the management component to decide based on partial results
 - fully incremental prosody extraction can be used as a basis for emotion detection
 - Incremental Speech Synthesis (iSS) [2] is able to change, extend, or plausibly interrupt output
 - prosodic adaptations are possible with extremely low latency (~30ms)
 - Dialogue Flow Estimation [3] enables predictions about the short-term future (upcoming ends of words/turns/hesitations)
 - The IU framework provides an architecture to integrate the above components and to allow reasoning over time [4]
 - IUs hold minimal amounts of information (at the respective level: words, phonemes, phrases, ideas, ...)
 - interconnected processors collaboratively build and extend an IU network that represent the system's information state
- all of this is available** in our toolkit for Incremental Spoken Dialogue Processing, **InproTK** [5].

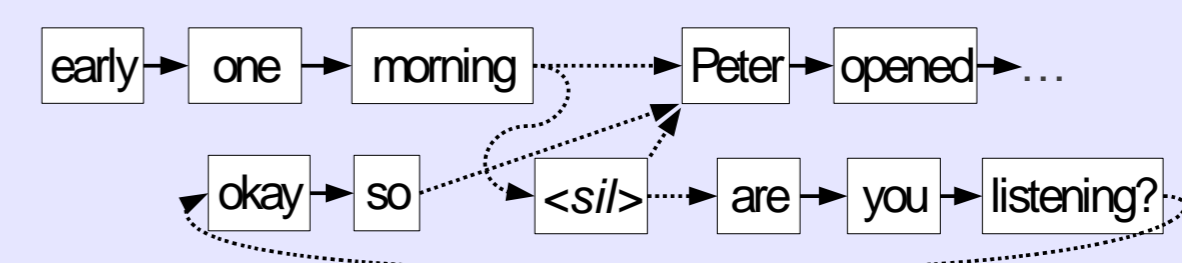


System Architecture



An interactive storyteller that adapts to user feedback would consist of:

- iSR to detect and understand feedback incrementally
- emotion detection to measure emotional aspects of the feedback utterance
 - based on words, prosody, and possibly temporal relation to system utterance)
- Storytelling Logic that generates and follows utterance plans [6]
 - plan selection and adaptation based on affective state
 - some additional rules to deal with the content of user feedback
- iSS to generate output just-in-time with little overhead
 - some support for adaptation of high-level prosodic aspects (stress, tones) [7]
- emotional coloring that changes prosody of generated output, based on emoSpeak[8]



Further Requirements and Future Work

- allow the user to influence the story content
 - *Interactive, Collaborative Storytelling*
- examine influence on different feedback elicitation strategies
- how should a speaker (i.e. system) react to feedback?
 - is it important to react in precisely the 'right' way?
- more insight into temporal alignment of speaker/listener,
 - including possibilities for the systematic manipulation of system behaviour

Open Source!

Our software for incremental dialogue processing is available as open source:

- inprotk.sourceforge.net for the source code and documentation
- www.inpro.tk for more information on the Inpro project

We value your feedback to inprotk-devel@lists.sourceforge.net !

References

- T. Baumann, Michaela Atterer, and D. Schlagen: „Assessing and Improving the Performance of Speech Recognition for Incremental Systems,” in *Proceedings of NAACL-HLT*, Boulder, USA, 2009.
- T. Baumann and D. Schlagen: „Inpro_iSS: A component for just-in-time incremental speech synthesis,” in *Proceedings of ACL System Demonstrations*, Jeju, Korea, 2012.
- T. Baumann and D. Schlagen: „Predicting the Micro-Timing of User Input for an Incremental Spoken Dialogue System that Completes a User's Ongoing Turn,” in *Proceedings of SigDial*, Portland, USA, 2011.
- D. Schlagen and G. Skantze: „A general, abstract model of incremental dialogue processing,” in *Proceedings of EACL*, Athens, Greece, 2009.
- T. Baumann and D. Schlagen: „The InproTK 2012 release,” in *Proceedings of SDCTD*, Montréal, Canada, 2012.
- G. Skantze and A. Hjalmarsson: „Towards incremental speech generation in dialogue systems,” in *Proceedings of SigDial*, Tokyo, Japan, 2010.
- M. Schröder and J. Trouvain: „The German Text-to-Speech synthesis system MARY: A tool for research, development and teaching,” *Int. J. of Speech Tech*, 6(3), 2003.
- M. Schröder, R. Cowie, E. Gouglas-Cowie, M. Westerdijk, and S. Gielen: „Acoustic correlates of emotion dimensions in view of speech synthesis,” *IEEE Trans. on Pattern Analysis and Machine Intelligence* 31(1), 2009.