



















- [2] Jens Allwood and Loredana Cerrato. 2003. A study of gestural feedback expressions. In *Proceedings of the 1st Nordic Symposium on Multimodal Communication*. Copenhagen, Denmark, 7–22.
- [3] Jens Allwood, Stefan Kopp, Karl Grammer, Elisabeth Ahlsén, Elisabeth Oberzaucher, and Markus Koppensteiner. 2007. The analysis of embodied communicative feedback in multimodal corpora: A prerequisite for behaviour simulation. *Language Resources and Evaluation* 41 (2007), 255–272. <https://doi.org/10.1007/s10579-007-9056-2>
- [4] Jens Allwood, Joakim Nivre, and Elisabeth Ahlsén. 1992. On the semantics and pragmatics of linguistic feedback. *Journal of Semantics* 9 (1992), 1–26. <https://doi.org/10.1093/jos/9.1.1>
- [5] Janet B. Bavelas, Linda Coates, and Trudy Johnson. 2000. Listeners as co-narrators. *Journal of Personality and Social Psychology* 79 (2000), 941–952. <https://doi.org/10.1037/0022-3514.79.6.941>
- [6] Janet B. Bavelas, Linda Coates, and Trudy Johnson. 2002. Listener responses as a collaborative process: The role of gaze. *Journal of Communication* 52 (2002), 566–580. <https://doi.org/10.1111/j.1460-2466.2002.tb02562.x>
- [7] Elisabetta Bevacqua, Sathish Pammi, Sylwia Julia Hyniewska, Marc Schröder, and Catherine Pelachaud. 2010. Multimodal backchannels for embodied conversational agents. In *Proceedings of the 10th International Conference on Intelligent Virtual Agents*. Philadelphia, PA, USA, 194–201. [https://doi.org/10.1007/978-3-642-15892-6\\_21](https://doi.org/10.1007/978-3-642-15892-6_21)
- [8] Paul Boersma and David Weenink. 2016. Praat, a system for doing phonetics by computer. (2016). <http://www.praat.org/>
- [9] Susan E. Brennan. 1990. *Seeking and Providing Evidence for Mutual Understanding*. Ph.D. Dissertation. Stanford University, Stanford, CA, USA.
- [10] Harry Bunt. 2011. Multifunctionality in dialogue. *Computer Speech and Language* 25 (2011), 222–245. <https://doi.org/10.1016/j.csl.2010.04.006>
- [11] Hendrik Buschmeier. 2018. *Attentive Speaking. From Listener Feedback to Interactive Adaptation*. Ph.D. Dissertation. Faculty of Technology, Bielefeld University, Bielefeld, Germany.
- [12] Hendrik Buschmeier, Timo Baumann, Benjamin Dosch, Stefan Kopp, and David Schlangen. 2012. Combining incremental language generation and incremental speech synthesis for adaptive information presentation. In *Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. Seoul, South Korea, 295–303.
- [13] Hendrik Buschmeier and Stefan Kopp. 2011. Towards conversational agents that attend to and adapt to communicative user feedback. In *Proceedings of the 11th International Conference on Intelligent Virtual Agents*. Reykjavik, Iceland, 169–182. [https://doi.org/10.1007/978-3-642-23974-8\\_19](https://doi.org/10.1007/978-3-642-23974-8_19)
- [14] Hendrik Buschmeier and Stefan Kopp. 2012. Using a Bayesian model of the listener to unveil the dialogue information state. In *Proceedings of the 16th Workshop on the Semantics and Pragmatics of Dialogue (SemDial)*. Paris, France, 12–20.
- [15] Hendrik Buschmeier and Stefan Kopp. 2014. A dynamic minimal model of the listener for feedback-based dialogue coordination. In *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (SemDial)*. Edinburgh, UK, 17–25.
- [16] Hendrik Buschmeier and Stefan Kopp. 2014. When to elicit feedback in dialogue: Towards a model based on the information needs of speakers. In *Proceedings of the 14th International Conference on Intelligent Virtual Agents (IVA)*. Boston, MA, USA, 71–80. [https://doi.org/10.1007/978-3-319-09767-1\\_10](https://doi.org/10.1007/978-3-319-09767-1_10)
- [17] Herbert H. Clark. 1996. *Using Language*. Cambridge University Press, Cambridge, UK. <https://doi.org/10.1017/CBO9780511620539>
- [18] Herbert H. Clark and Meredyth A. Krych. 2004. Speaking while monitoring addressees for understanding. *Journal of Memory and Language* 50 (2004), 62–81. <https://doi.org/10.1016/j.jml.2003.08.004>
- [19] Herbert H. Clark and Edward F. Schaefer. 1989. Contributing to discourse. *Cognitive Science* 13 (1989), 259–294. [https://doi.org/10.1207/s15516709cog1302\\_7](https://doi.org/10.1207/s15516709cog1302_7)
- [20] Kohji Dohsaka and Akira Shimazu. 1997. A system architecture for spoken utterance production in collaborative dialogue. In *Working Notes of the IJCAI-97 Workshop on Collaboration, Cooperation and Conflict in Dialogue Systems*. Nagoya, Japan.
- [21] Konrad Ehlich. 1986. *Interjektionen*. Max Niemeyer Verlag, Tübingen, Germany. <https://doi.org/10.1515/9783111357133>
- [22] Chris D. Frith and Uta Frith. 2006. The neural basis of mentalizing. *Neuron* 50 (2006), 531–534. <https://doi.org/10.1016/j.neuron.2006.05.001>
- [23] Mattias Heldner, Jens Eklund, and Julia Hirschberg. 2010. Pitch similarity in the vicinity of backchannels. In *Proceedings of INTERSPEECH 2010*. Makuhari, Japan, 3054–3057.
- [24] Harold Jeffreys. 1961. *Theory of Probability* (3rd ed.). Clarendon Press, Oxford, UK.
- [25] Daphne Koller and Nir Friedman. 2009. *Probabilistic Graphical Models. Principles and Techniques*. The MIT Press, Cambridge, MA, USA.
- [26] Stefan Kopp. 2010. Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors. *Speech Communication* 52 (2010), 587–597. <https://doi.org/10.1016/j.specom.2010.02.007>
- [27] Stefan Kopp, Jens Allwood, Karl Grammer, Elisabeth Ahlsén, and Thorsten Stockmeier. 2008. Modeling embodied feedback with virtual humans. In *Modeling Communication with Robots and Virtual Humans*, Ipke Wachsmuth and Günther Knoblich (Eds.). Springer-Verlag, Berlin, Germany, 18–37. [https://doi.org/10.1007/978-3-540-79037-2\\_2](https://doi.org/10.1007/978-3-540-79037-2_2)
- [28] Robert M. Krauss and Sidney Weinheimer. 1966. Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology* 4 (1966), 343–346. <https://doi.org/10.1037/h0023705>
- [29] François Mairesse and Marilyn A. Walker. 2010. Towards personality-based user adaptation: Psychologically informed stylistic language generation. *User Modeling and User-Adapted Interaction* 20 (2010), 227–278. <https://doi.org/10.1007/s11257-010-9076-2>
- [30] Zofia Malisz, Marcin Włodarczyk, Hendrik Buschmeier, Joanna Skubisz, Stefan Kopp, and Petra Wagner. 2016. The ALICO corpus: Analysing the active listener. *Language Resources and Evaluation* 50 (2016), 411–442. <https://doi.org/10.1007/s10579-016-9355-6>
- [31] Teruhisa Misu, Etsuo Mizukami, Yoshinori Shiga, Shimichi Kawamoto, Hisashi Kawai, and Satoshi Nakamura. 2011. Toward construction of spoken dialogue system that evokes users' spontaneous backchannels. In *Proceedings of the 12th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. Portland, OR, USA, 259–265.
- [32] Louis-Philippe Morency, Iwan de Kok, and Jonathan Gratch. 2010. A probabilistic multimodal approach for predicting listener backchannels. *Autonomous Agents and Multiagent Systems* 20 (2010), 70–84. <https://doi.org/10.1007/s10458-009-9092-y>
- [33] Richard D. Morey and Jeffrey N. Rouder. 2015. BayesFactor: Computation of Bayes factors for common designs. (2015). <https://CRAN.R-project.org/package=BayesFactor>
- [34] Yukiko I. Nakano, Gabe Reinstein, Tom Stocky, and Justine Cassell. 2003. Towards a model of face-to-face grounding. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*. Sapporo, Japan, 553–561. <https://doi.org/10.3115/1075096.1075166>
- [35] Catharine Oertel, José Lopes, Yu Yu, Kenneth A. Funes Mora, Joakim Gustafson, Alan W. Black, and Jean-Marc Odobez. 2016. Towards building an attentive artificial listener: On the perception of attentiveness in audio-visual feedback tokens. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. Tokyo, Japan, 21–28. <https://doi.org/10.1145/2993148.2993188>
- [36] Dennis Reidsma, Iwan de Kok, Daniel Neiberg, Sathish Pammi, Bart van Straalen, Khiet Truong, and Herwin van Welbergen. 2011. Continuous interaction with a virtual human. *Journal on Multimodal User Interfaces* 4 (2011), 97–118. <https://doi.org/10.1007/s12193-011-0060-x>
- [37] Zsófia Ruttkay, Claire Dormann, and Han Noot. 2004. Embodied conversational agents on a common ground. A framework for design and evaluation. In *From Brows to Trust. Evaluating Embodied Conversational Agents*, Zsófia M. Ruttkay and Catherine Pelachaud (Eds.). Kluwer Academic Publishers, Dordrecht, The Netherlands, 27–66. [https://doi.org/10.1007/1-4020-2730-3\\_2](https://doi.org/10.1007/1-4020-2730-3_2)
- [38] David Schlangen and Gabriel Skantze. 2011. A general, abstract model of incremental dialogue processing. *Dialogue and Discourse* 2 (2011), 83–111. <https://doi.org/10.5087/dad.2011.105>
- [39] Gabriel Skantze, Anna Hjalmarsson, and Catharine Oertel. 2014. Turn-taking, feedback and joint attention in situated human–robot interaction. *Speech Communication* 65 (2014), 50–66. <https://doi.org/10.1016/j.specom.2014.05.005>
- [40] Matthew Stone. 2002. Lexicalized grammar 101. In *Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*. Philadelphia, PA, USA, 77–84.
- [41] Matthew Stone, Christine Doran, Bonnie Webber, Tonia Bleam, and Martha Palmer. 2003. Microplanning with communicative intentions: The SPUD system. *Computational Intelligence* 19 (2003), 311–381. <https://doi.org/10.1046/j.0824-7935.2003.00221.x>
- [42] Marilyn A. Walker, Diane J. Litman, Candace A. Kamm, and Alicia Abella. 1998. Evaluating spoken dialogue agents with PARADISE: Two case studies. *Computer Speech and Language* 12 (1998), 317–347. <https://doi.org/10.1006/csla.1998.0110>
- [43] Nigel Ward. 2004. Pragmatic functions of prosodic features in non-lexical utterances. In *Proceedings of the International Conference on Speech Prosody*. Nara, Japan, 325–328.
- [44] Nigel Ward. 2006. Non-lexical conversational sounds in American English. *Pragmatics & Cognition* 14 (2006), 129–182. <https://doi.org/10.1075/pc.14.1.08war>
- [45] Nigel Ward and Wataru Tsukahara. 2000. Prosodic features which cue back-channel responses in English and Japanese. *Journal of Pragmatics* 38 (2000), 1177–1207. [https://doi.org/10.1016/S0378-2166\(99\)00109-5](https://doi.org/10.1016/S0378-2166(99)00109-5)
- [46] Peter Wittenburg, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 2006. ELAN: A professional framework for multimodality research. In *Proceedings of the 5th International Conference on Language Resources and Evaluation*. Genoa, Italy, 1556–1559.
- [47] Victor H. Yngve. 1970. On getting a word in edgewise. In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, Mary Ann Campbell et al. (Eds.). Chicago Linguistic Society, Chicago, IL, USA, 567–577.