

Laughter and Smile Processing for Human-Computer Interactions

Kevin El Haddad, Hüseyin Çakmak, Stéphane Dupont, Thierry Dutoit

TCTS lab - University of Mons
31 Boulevard Dolez, 7000, Mons Belgium
kevin.elhaddad@umons.ac.be

Abstract

This paper provides a short summary of the importance of taking into account laughter and smile expressions in Human-Computer Interaction systems. Based on the literature, we mention some important characteristics of these expressions in our daily social interactions. We describe some of our own contributions and ongoing work to this field.

Keywords: laughter, smiling, Human-Computer Interaction, synthesis, recognition

1. Introduction

Computers are increasingly becoming part of our lives, making a lot of our daily tasks easier. As interactions with them increase, so too does the need to have interfaces that are more natural to use than the traditional keyboard or mouse. These interfaces included speech, which is our most common means of communication. An ideal interface would be one with which we could communicate as if we were talking to person. This would entail access to the same expressiveness and emotions as in a conversation with another person. Laughter and smiling have been the subject of several studies in the social sciences, including psychology, anthropology, and paralinguistics, because of their importance in social interaction. They are indeed multifunctional and extremely common. Therefore in order to create a natural Human-Computer Interaction (HCI) system, these expressions have to be integrated in it. However, the need to consider laughter and smiling in HCI systems may not be immediately obvious to researchers not related to this field of study in particular, or to affective computing in general. This might be because integrating emotions in general might not seem important for some, or, at first glance, laughter and smiling may seem to be no more important than any other paralinguistic expression. The goal of this paper is provide a broad overview of studies in different research fields highlighting the necessity of considering laughs and smiles into an HCI system in order to make the interaction more natural. Below, we will first discuss the relevance of these non-verbal paralinguistic expressions to HCI by providing a short survey of previous work. We will then sketch some applications in which they have been considered, and finally we will describe our own contributions to this field.

2. Laughter and Smiles in Interactions

Laughs and smiles are among the most, if not the most, important non-verbal expressions in our daily interactions and this makes them worthy to be considered in HCI systems. The reason of their importance are cited in the following paragraphs:

Frequent Occurrences in Conversations: Laughs and smiles occur frequently in conversations. Indeed, in the ICSI meeting corpus, Laskowski and Burger reported 9.5%

of the total verbalizing time being laughter (Laskowski and Burger, 2007; Vogel et al., 2013). In other work, Chovil did not consider smiles in his analysis of affect in conversation as smiles were so overwhelmingly frequently present in the data compared to other expressions (Chovil, 1991). This high frequency of occurrence is the first reason to work on including smiling and laughter in HCI systems which aim to replicate human-human interactions.

Expression of Different Emotions: Laughter can express several different affective states. Although intuitively and commonly related to emotions with positive valence (generally amusement, joy and sympathy), laughter can also express other negative emotions such as disappointment, stress and embarrassment (Devillers and Vidrascu, 2007).

Smiling can also express different emotions of different valence, such as joy or embarrassment (Ambadar et al., 2008; Keltner, 1995; Frank et al., 1993).

Emotions are crucial to understanding (recognition) or creating (synthesis) a certain context or mood. Being able to automatically and accurately assimilate this dimension in a dialogue would improve the interaction by increasing its naturalness.

Social Functions: Laughs and smiles are more likely to happen with someone rather than alone (Glenn, 2003; Fridlund, 1991). They have been shown to be somewhat related to the cultural background (Soury and Devillers, 2014). In social interactions, they are used not only to express emotions, but also to apply certain social functionalities that do not really contain emotions. In fact, people do sometimes laugh and smile without really feeling any emotion.

Laughter and smiling can be used in the course of a conversation, with social functions, punctuating the dialogue with social information (Provine, 2010), expressing politeness (Hoque et al., 2011) or changing the topic (Bonin et al., 2014; Vogel et al., 2013).

Both laughter and smiling can be used as backchannels to show interest in the speaker and to encourage him or her to carry on talking (Duncan, 1972; Poggi and Pelachaud, 2000).

Being able to use these expressions with these social functionalities in dialogue systems will increase the naturalness of an agent's reaction during an interaction.

Perception of Laughs & Smiles: Laughs and smiles are contagious as shown by Provine (Provine, 2013) and Wild (Wild et al., 2003) respectively. Indeed it is likely that a subject will smile or laugh under when they are exposed to another's laughter or smiling. They can also affect the perception of a subject: viewing a smiling photograph versus a photograph of the same person with a neutral expression has been reported to result in an increased perception of characteristics such as attractiveness, trustworthiness and sociability (Reis et al., 1990).

Gelotophobia: Gelotophobia is the fear of being laughed or smiled at (Ruch et al., 2014). This disease is another example of the importance of these two expressions in our social communications and show the influence they can have on individuals.

3. Laughter and Smiling Embedded in HCI Applications

Several HCI system have already been developed which include laughter and smiling detection systems.

Melder et. al. (Melder et al., 2007) presented a multi-modal real-time HCI system with the goal to detect and elicit laughter. In this application, a user's behavior is monitored, interpreted and regulated by the system in an interactive loop. An audio laughter (Truong and van Leeuwen, 2007) detection system and visual smile recognition system were developed and contributed to assess the user's emotions state.

Some HCI experiments were also conducted in the framework of the European project Ilhaire (Dupont et al., 2016), which was dedicated to study laughter. For example, in (Pecune et al., 2015a), a laughing avatar is used to study the contribution of a virtual agent to enhancing a user's experience. A user is presented with stimuli in the presence or not of the avatar. When the avatar is present, it either copies the user's behavior of laughs at predefined times and intensities. The multimodal synthesis system developed in (Ding et al., 2014) is used here to generate laughter animation. The detection is made using a platform based on a Eye-Web XMI platform (Mancini et al., 2014).

A facial smile detection system was integrated in a Perceptual User Interface (PUI) in (Deniz et al., 2008). This PUI was used in an application to control the status and insert smile/big smile emoticons in an Instant Messaging client conversation window. The system can assess the level of the facial smile and map it to the emoticon to be inserted.

One of the goals of the European project JOKER (Devillers et al., 2015) is to study the impact an emotional social agent showing empathy and compassion might have on a user's mood during a conversation. Interfaces related to laughter and smiling are crucial for obtaining such a virtual social companion.

4. Interfaces for HCI applications

In order to take into account laughter and smiling in HCI systems, interfaces must be developed for this task. These interfaces take care of the generation (synthesis) and detection (recognition) of laughs/smiles. This section will present our work and main contribution in this field which

concern interfaces related to laughs and smiles. It will also mention interesting work of others in this field. Please also note that all the synthesis and recognition/detection modules mentioned in Section 3., even though relevant here, will not be repeated.

4.1. Synthesis Applications

Being able to synthesize laughter and smiles would, in general, increase the naturalness of an HCI and therefore make the interaction more comfortable to the user(s) as shown in (Theonas et al., 2008). Application examples of laughter and smiling synthesis systems in HCI can be, first, the control of a conversation flow by using the social functions that smiles and laughs have. It could provide the user(s) with feedbacks while he/she is speaking and thus encouraging him/her to carry on. It could, for instance, change the subject of the conversation, or express agreement or even disagreement (with a mockery laugh for example). A second example would be to influence the user's mood or emotional state. Indeed this could be used to express empathy in order to make the avatar more likable (Devillers et al., 2015), or trigger amusement by uttering amused laughs or smiles (Niewiadomski et al., 2013; Pecune et al., 2015b). Such synthesis systems could also be used for medical purposes, helping to study the phenomenon of gelotophobia and even treating it. This was one of the purposes of the Ilhaire European project (Ruch et al., 2015; Ruch et al., 2014). It can also help reducing stress since it has been found that laughter helps reduce stress (Bennett et al., 2003).

Urbain et. al. (Urbain et al., 2014) presented a Hidden Markov Model (HMM)-based audio laughter synthesis system in which the level of the arousal intensity or of the laughter is controllable. Other work on audiovisual laughter synthesis can be found in (Çakmak, 2016). In this thesis, the author presents synthesis and evaluation of audio and motion capture cues of laughter. He also presents synchronization rules between the audio and visual cues for synthesizing laughter from a virtual agent. In (de Kok and Heylen, 2011), the authors present an attempt on predicting the types of smiles that should be generated, based on the context. But no actual synthesis is presented. In (Ochs et al., 2010), a decision tree is used to predict the type of smile to be generated. The generation system was also evaluated using a subjective perceptual test.

Our contribution in this field focused on adding smiles and laughs to synthesized speech, thus creating speech-smiled and speech-laugh. Hidden Markov Models (HMM)-based systems were used to synthesize speech-smiles (El Haddad et al., 2015e; El Haddad et al., 2015b) and control the arousal level of smiling in an utterance. A speech-laugh synthesis system was also created based also on HMM and proved to increase the naturalness perceived compared to neutrally synthesized sentences (El Haddad et al., 2015f; El Haddad et al., 2015a). In order to do this, databases were collected containing laughter and smiled speech. The next step is first to be able to synthesize in real time sentences will controlling the level of amusement in speech. This includes varying the level of smiling and adding laughter bursts. We will also work on reproducing this system

in different languages. In order to do that, a multilingual database similar to the one in (El Haddad et al., 2015f) has been collected. We also intend to create the same speech-laugh/smiling synthesis systems audiovisually. This means synthesizing also motion capture speech-laugh and controllable smiling data synchronized with the synthesized audio cues.

4.2. Recognition Applications

Since smiles and laughs can express different types of emotions and can also have several social functions, their detection and recognition would help understanding the emotional state of the user(s) and therefore also the context. A context understanding will help an agent react more adequately. This would improve the quality of the interaction (Yang et al., 2015). This can also be used for user mood monitoring, for instance, to detect the level of amusement and estimate the level of stress since they are related (Bennett et al., 2003). In addition, being able to recognize/detect smiles and laughs in speech, would increase the robustness of an automatic speech recognition system by differentiating between speech and non-speech.

Knox et. al. (Knox and Mirghafori, 2007) presents an automatic audio laughter detection using a neural network.

In (Yang et al., 2015), presents a multimodal laughter and smiling recognition system to be used in a human-robot interaction with elderly people. In (Ito et al., 2005), Ito et. al. also present a laughter and smiling audiovisual detection system. This system was developed for application in natural conversation videos.

The main contribution we have in this field is work related the arousal level assessment of amusement. In (El Haddad et al., 2015c; El Haddad et al., 2015d) we defined so called Amused Speech Components (ASC), collected data and presented analyses and classification systems for them. This work is in the larger framework of assessing the amusement arousal level in a given sentence. Indeed, we aim at building an ASC detection system and then accurately assess a level of amusement arousal in the given sentence based on the detected ASC. A multimodal system will be used based on a database containing motion capture data as well as audio data.

5. Acknowledgements

This work was partly supported by the Chist-Era project JOKER with contribution from the Belgian Fonds de la Recherche Scientifique (FNRS), contract no. R.50.01.14.F. H. Çakmak receives a Ph.D. grant from the Fonds de la Recherche pour l'Industrie et l'Agriculture (F.R.I.A.), Belgium.

6. Bibliographical References

Ambadar, Z., Cohn, J. F., and Reed, L. I. (2008). All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *Journal of Nonverbal Behavior*, 33(1):17–34.

Bennett, M. P., Zeller, J. M., Rosenberg, L., and McCann, J. (2003). The effect of mirthful laughter on stress and natural killer cell activity. *Alternative therapies in health and medicine*, 9(2):38.

Bonin, F., Campbell, N., and Vogel, C. (2014). Time for laughter. *Knowledge-Based Systems*, 71:15 – 24.

Çakmak, H. (2016). *Audiovisual Laughter Synthesis - A Statistical Parametric Approach*. Ph.D. thesis, University of Mons, February.

Chovil, N. (1991). Discourse oriented facial displays in conversation. *Research on Language and Social Interaction*, 25(1-4):163–194.

de Kok, I. and Heylen, D. (2011). When do we smile? analysis and modeling of the nonverbal context of listener smiles in conversation. In *Affective Computing and Intelligent Interaction*, volume 6974 of *Lecture Notes in Computer Science*, pages 477–486, Berlin, Germany, October. Springer Verlag.

Deniz, O., Castrillon, M., Lorenzo, J., Anton, L., and Bueno, G., (2008). *Advances in Visual Computing: 4th International Symposium, ISVC 2008, Las Vegas, NV, USA, December 1-3, 2008. Proceedings, Part II*, chapter Smile Detection for User Interfaces, pages 602–611. Springer Berlin Heidelberg, Berlin, Heidelberg.

Devillers, L. and Vidrascu, L. (2007). Positive and negative emotional states behind the laughs in spontaneous spoken dialogs. In *Interdisciplinary Workshop on The Phonetics of Laughter*, page 37.

Devillers, L., Rosset, S., Dubuisson Duplessis, G., Sehili, M. A., Bechade, L., Delaborde, A., Gossart, C., Letard, V., Yang, F., Yemez, Y., Turker, B. B., Sezgin, M., El Haddad, K., Dupont, S., Luzzati, D., Esteve, Y., Gilmartin, E., and Nick, C. (2015). Multimodal data collection of human-robot humorous interactions in the joker project. In *ACII, Xi'an, China*, September.

Ding, Y., Prepin, K., Huang, J., Pelachaud, C., and Artières, T. (2014). Laughter animation synthesis. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems, AAMAS '14*, pages 773–780, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23:283–292.

Dupont, S., Çakmak, H., Curran, W., Dutoit, T., Hofmann, J., McKeown, G., Pietquin, O., Platt, T., Ruch, W., and Urbain, J., (2016). *Toward Robotic Socially Believable Behaving Systems - Volume I : Modeling Emotions*, chapter Laughter Research: A Review of the ILHAIRE Project, pages 147–181. Springer International Publishing, Cham.

El Haddad, K., Cakmak, H., Dupont, S., , and Dutoit, T. (2015a). Breath and repeat: An attempt at enhancing speech-laugh synthesis quality. In *European Signal Processing Conference (EUSIPCO 2015)*, Nice, France, 31 August–4 September.

El Haddad, K., Cakmak, H., Dupont, S., and Dutoit, T. (2015b). An HMM Approach for Synthesizing Amused Speech with a Controllable Intensity of Smile. In *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Abu Dhabi, UAE, 7-10 December.

El Haddad, K., Cakmak, H., Dupont, S., and Dutoit,

- T. (2015c). Towards a Level Assessment System of Amusement in Speech Signals: Amused Speech Components Classification. In *IEEE International Symposium on Signal Processing and Information Technology (IS-SPIT)*, Abu Dhabi, UAE, 7-10 December.
- El Haddad, K., Dupont, S., Cakmak, H., and Dutoit, T. (2015d). Shaking and Speech-smile Vowels Classification: An Attempt at Amusement Arousal Estimation from Speech Signals. In *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Orlando, Florida, US, 14-16 December.
- El Haddad, K., Dupont, S., d'Alessandro, N., and Dutoit, T. (2015e). An HMM-based speech-smile synthesis system: An approach for amusement synthesis. In *International Workshop on Emotion Representation, Analysis and Synthesis in Continuous Time and Space (EMO-SPACE)*, Ljubljana, Slovenia, 4-8 May.
- El Haddad, K., Dupont, S., Urbain, J., and Dutoit, T. (2015f). Speech-laugh: An HMM-based Approach for Amused Speech Synthesis. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP 2015)*, pages 4939–4943, Brisbane, Australia, 19-24 April.
- Frank, M. G., Ekman, P., and Friesen, W. V. (1993). Behavioral markers and recognizability of the smile of enjoyment. *Journal of personality and social psychology*, 64(1):83.
- Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology*, 60(2):229.
- Glenn, P. (2003). *Laughter in interaction*, volume 18. Cambridge University Press.
- Hoque, M., Morency, L.-P., and Picard, R. W. (2011). Are you friendly or just polite?—analysis of smiles in spontaneous face-to-face interactions. In *Affective Computing and Intelligent Interaction*, pages 135–144. Springer.
- Ito, A., Wang, X., Suzuki, M., and Makino, S. (2005). Smile and laughter recognition using speech processing and face recognition from conversation video. In *Cyberworlds, 2005. International Conference on*, pages 8 pp.–444, Nov.
- Keltner, D. (1995). The signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology*, pages 441–454.
- Knox, M. T. and Mirghafori, N. (2007). Automatic laughter detection using neural networks. In *INTERSPEECH*, pages 2973–2976.
- Laskowski, K. and Burger, S. (2007). Analysis of the occurrence of laughter in meetings. In *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007)*, pages 1258–1261, Antwerp, Belgium, 27-31 August.
- Mancini, M., Varni, G., Niewiadomski, R., Volpe, G., and Camurri, A. (2014). How is your laugh today? In *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI EA '14, pages 1855–1860, New York, NY, USA. ACM.
- Melder, W. A., Truong, K. P., Uyl, M. D., Van Leeuwen, D. A., Neerinx, M. A., Loos, L. R., and Plum, B. S. (2007). Affective multimodal mirror: Sensing and eliciting laughter. In *Proceedings of the International Workshop on Human-centered Multimedia*, HCM '07, pages 31–40, New York, NY, USA. ACM.
- Niewiadomski, R., Hofmann, J., Urbain, J., Platt, T., Wagner, J., Piot, B., Cakmak, H., Pammi, S., Baur, T., Dupont, S., Geist, M., Lingensfelder, F., McKeown, G., Pietquin, O., and Ruch, W. (2013). Laugh-aware virtual agent and its impact on user amusement. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '13, pages 619–626, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.
- Ochs, M., Niewiadomski, R., and Pelachaud, C. (2010). How a virtual agent should smile? - morphological and dynamic characteristics of virtual agent's smiles. In *10th International Conference on Intelligent Virtual Agents (IVA)*, Philadelphia, Pennsylvania, US.
- Pecune, F., Mancini, M., Biancardi, B., Varni, G., Ding, Y., Pelachaud, C., Volpe, G., and Camurri, A. (2015a). Laughing with a virtual agent. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 1817–1818. International Foundation for Autonomous Agents and Multiagent Systems.
- Pecune, F., Mancini, M., Biancardi, B., Varni, G., Ding, Y., Pelachaud, C., Volpe, G., and Camurri, A. (2015b). Laughing with a virtual agent. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, pages 1817–1818, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.
- Poggi, I. and Pelachaud, C. (2000). Embodied conversational agents. chapter Performative Facial Expressions in Animated Faces, pages 155–188. MIT Press, Cambridge, MA, USA.
- Provine, R. R. (2010). Laughter Punctuates Speech: Linguistic, Social and Gender Contexts of Laughter. *Ethology*, 95:291–298.
- Provine, R. R. (2013). Contagious laughter: Laughter is a sufficient stimulus for laughs and smiles. *Bulletin of the Psychonomic Society*, 30(1):1–4.
- Reis, H. T., Wilson, I. M., Monestere, C., Bernstein, S., Clark, K., Seidl, E., Franco, M., Gioioso, E., Freeman, L., and Radoane, K. (1990). What is smiling is beautiful and good. *European Journal of Social Psychology*, 20(3):259–267.
- Ruch, W. F., Platt, T., Hofmann, J., Niewiadomski, R., Urbain, J., Mancini, M., and Dupont, S. (2014). Gelotophobia and the challenges of implementing laughter into virtual agents interactions. *Frontiers in Human Neuroscience*, 8(928).
- Ruch, W., Hofmann, J., and Platt, T. (2015). Individual differences in gelotophobia and responses to laughter-eliciting emotions. *Personality and Individual Differences*, 72:117 – 121.
- Soury, M. and Devillers, L. (2014). Smile and laughter

- in human-machine interaction: a study of engagement. In Nicoletta Calzolari (Conference Chair), et al., editors, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland, may. European Language Resources Association (ELRA).
- Theonas, G., Hobbs, D., and Rigas, D. (2008). Employing virtual lecturers' facial expressions in virtual educational environments. *IJVR*, 7(1):31–44.
- Truong, K. P. and van Leeuwen, D. A. (2007). Automatic discrimination between laughter and speech. *Speech Communication*, 49(2):144 – 158.
- Urbain, J., Cakmak, H., Charlier, A., Denti, M., Dutoit, T., and Dupont, S. (2014). Arousal-driven synthesis of laughter. *IEEE Journal of Selected Topics in Signal Processing*, 8(2):273–284, April.
- Vogel, C., Campbell, N., Bonin, F., and Gilmartin, E. (2013). Exploring the role of laughter in multiparty conversation.
- Wild, B., Erb, M., Eyb, M., Bartels, M., and Grodd, W. (2003). Why are smiles contagious? an fmri study of the interaction between perception of facial affect and facial movements. *Psychiatry Research: Neuroimaging*, 123(1):17 – 36.
- Yang, F., Sehili, M. A., Barras, C., and Devillers, L., (2015). *Social Robotics: 7th International Conference, ICSR 2015, Paris, France, October 26-30, 2015, Proceedings*, chapter Smile and Laughter Detection for Elderly People-Robot Interaction, pages 694–703. Springer International Publishing, Cham.