Year: 2016

Laughter research: a review of the ILHAIRE project

Dupont, Stéphane ; Çakmak, Hüseyin ; Curran, Will ; Dutoit, Thierry ; Hofmann, Jennifer ; McKeown, Gary ; Pietquin, Olivier ; Platt, Tracey ; Ruch, Willibald ; Urbain, Jérôme

Laughter Research: a Review of the ILHAIRE Project

Stéphane Dupont, Hüseyin Çakmak, Will Curran, Thierry Dutoit, Jennifer Hofmann, Gary McKeown, Olivier Pietquin, Tracey Platt, Willibald Ruch, Jérôme Urbain

1 Introduction

Laughter is all around us. So much that we often don't even notice it. First, laughter has a strong connection with humour. Most of us seek out laughter and people who make us laugh, and it is what we do when we gather together as groups relaxing and having a good time. But laughter also plays an important role in making sure we interact with each other smoothly. It provides social bonding signals that allow

––––––––––––––––––––

Stéphane Dupont
University of Mons, Belgium e-mail: stephane.dupont@umons.ac.be

Hüseyin Çakmak
University of Mons, Belgium e-mail: huseyin.cakmak@umons.ac.be

Will Curran
Queen's University, Belfast, UK e-mail: w.curran@qub.ac.uk

Thierry Dutoit
University of Mons, Belgium e-mail: thierry.dutoit@umons.ac.be

Jennifer Hofmann
University of Zurich, Switzerland e-mail: j.hofmann@psychologie.uzh.ch

Gary McKeown
Queen's University, Belfast, UK e-mail: g.mckeown@qub.ac.uk

Olivier Pietquin
University of Lille, France e-mail: olivier.pietquin@univ-lille1.fr

Tracey Platt
University of Zurich, Switzerland e-mail: tracey.platt@psychologie.uzh.ch

Willibald Ruch
University of Zurich, Switzerland e-mail: w.ruch@psychologie.uzh.ch

Jérôme Urbain
University of Mons, Belgium e-mail: jerome.p.urbain@gmail.com

our conversations to flow seamlessly between topics; to help us repair conversations that are breaking down; and to end our conversations on a positive note.

Currently, attempts are being made at creating computer agents that interact with humans in the same way humans do, understanding their social signals (cfr. review by Vinciarelli *et al.* [133]). However, laughter has not received so much attention in the area of Human Computer Interaction (HCI), despite its huge importance in social interactions, and as we will see, being one of the most important non-verbal vocal social signal. Gathering new knowledge on laughter, from both engineering and human sciences perspectives, and transferring this understanding to the design of computer agents will enable them with both emotional and conversational competence that will increase their impact and application potential. This was the goal of the ILHAIRE [1] collaborative research project. It has been supported by the Future and Emerging Technologies branch of the 7th framework program for research of the European Union. It started in September 2011 and has been running until September 2014. It gathered an interdisciplinary team in different areas of human sciences and engineering: psychology of laughter, emotion-oriented computing, automatic recognition and synthesis of communicative behaviors and signals, study of non-verbal social communication cues, and virtual agents.

This chapter provides a summarized account of the context of this research topic and a comprehensive overview of the ILHAIRE research results, as well as references to pioneering works. It is organized as follows. Section 2 covers the roles and functions of laughter, during both hilarious as well as conversational interactions. Such knowledge is fundamental when reaching towards endowing computer agents with the capability to understand and use this fundamental social signal. Section 3 covers another fundamental building block, the more precise understanding of the different characteristics of the signal itself, and how they affect the way laughter is perceived and impact us, together with the personality factors that matter. Such studies are heavily relying on observations drawn from multimodal corpora of human-human and human-computer interactions. The advances in available databases and their annotation are reported in Section 4. Corpora are also key in designing computational approaches enabling to embed laughter in HCI, which imply methods to accurately detect and recognize laughter, to produce natural-looking and natural-sounding laughter, to perceive the communication scene and make use of the communication context in order to laugh the right way at the right time. These are respectively addressed in Sections 5, 6 and 7. Those cover the different facets of laughter, in particular the sound, the gestures and body motion/posture and the facial expressions. Finally, application perspectives are drawn in Section 8 while research perspectives are outlined in Section 9.

---

[1] http://www.ilhaire.eu/

2 Roles and Functions

Laughter is an ubiquitous social signal in human interaction; it occurs very often and only a fraction of the occurrences seem to be related to humour. The ILHAIRE project made an initial distinction between laughter that was social and conversational in nature and laughter that was hilarious, with the latter kind of laughter being more directly related to humour. The research conducted within ILHAIRE bore out this distinction, with laughter intensity being the important distinguishing variable. To understand why these separate forms of laughter may have arisen, and what functions each serves, we can look to their evolutionary origins and to the situations in which social and hilarious laughter are found.

Laughter is an evolutionarily ancient behaviour which arguably precedes spoken language (Davila-Ross *et al.* [18]; Davila-Ross *et al.* [17]). This is borne out by the observation of laughter-like behaviours in many primate species. Preuschoft and van Hooff [100] argue that the smile and laughter distinction has its origin in two different sets of signals, and that both sets of signals have important social functions in regulating and enhancing social affiliation.

Laughter-like behaviour in chimpanzees prolongs play actions (Matsusaka [69]), suggesting that it is an important tool for promoting social affiliation and developing cooperative and competitive behaviours (Davila-Ross *et al.* [17]), and similar behaviour is observed in human infants (Rothbart [105]).

Smile-like behaviours have been observed in primates signalling affiliation, appeasement, reassurance, and submission; again a social bonding element seems to unify these functions. Thus from an evolutionary perspective, laughter and smiling can be viewed as key adaptive behaviours because of their facilitative effect on social cohesion. Preuschoft and van Hooff [100] propose that, within a number of primate species and especially humans, the boundaries between smile- and laughter-like signals have become blurred and intensity-related phenomena enable an ambiguous yet gradual transition of function from social and (in the case of humans) conversational laughter to stronger amusement related laughter.

Social and conversational laughter predominates at low levels of intensity. Here there seems to be a strongly ordered and rule bound nature to laughter production. These rule sets have been most firmly elucidated within the Conversation Analysis tradition (Glenn [36], Holt [**?**], O'Donnell-Trujillo & Adams [87]), but also confirmed in some more experimental work (Bonin *et al.* [7]). Laughter in conversation tends to be initiated by one member of an interaction, but others can join in and it is typically a shared rather than competitive activity. Laughter can be received as an invitation to shared laughter that can lead to reciprocated laughter, or responded to with silence or serious talking. A response to silence can be the further pursuit of shared laughter while serious talking is an active declining of the invitation (Glenn [36]). While the latter response can often extend the exploration of a topic of conversation, accepting an invitation to laughter can lead to a topic change. In this way laughter has a regulatory function in conversations by serving as a turn-taking cue or signalling that the speaker may be approaching a transition point in their topic or theme (Holt [**?**]; O'Donnell-Trujillo & Adams [87]). Other conversational rules

suggest that in dyadic conversations the speaker is more likely than the listener to laugh first, while in group conversations the listeners are more likely to laugh first (Glenn [36]).

An important feature of laughter that facilitates these conversational dynamics is that it has an inherent ambiguity. This is most apparent at low levels of intensity, but the transition to more humour related laughter at higher levels of intensity also seems to be reasonably weakly defined. Within the ILHAIRE project experimentation showed that, when removed from its conversational context, laughter becomes hard to classify in terms of classic emotional or sociolinguistic labels. This suggests that context plays a strong role in the interpretation of laughter, as has been recently highlighted in more general valance and emotion-based stimuli (Kayyal *et al.* [53]). McKeown *et al.* (under review) have argued based on ILHAIRE research that ambiguity plays an important role in servicing and maintaining social interaction. Holt [**?**] suggests that, as laughter as a signal has no propositional content, it can be safely interspersed within conversations; because of its lack of propositional content, laughter serves as a social bonding signal that allows an abstraction from the content of the conversation, it is this that allows safe turn-taking transitions, topic changes and terminations. McKeown *et al.* suggest that this ambiguous property of laughter allows multiple interpretations of the same content to be held by different interlocutors, which facilitates the repair of conversations. Further, they propose that laughter ambiguity also facilitates the holding of two separate interpretations of the same content in the minds of both interlocutors at the same time. Consequently laughter can aid the safe exploration of possible taboo areas, hypothetical scenarios, and impropriety while retaining the possibility of plausible deniability. Laughter requires an ambiguous nature to enable this social exploration, allowing it to be abstracted from the content and to be interpretable in multiple ways. The ambiguous nature of laughter has been underlined by an important study within the ILHAIRE project, in which similar-intensity laughter from different contexts within the same conversation, as well as from different conversations, was interchageable with minimal impact on the perceived genuineness of the conversational interaction.

Laughter also has clear relationships with humour. In ratings of laughs extracted from their conversational context, consistent strong positive correlations (in the range of r=.65 to r=0.7) were found between ratings of intensity and rating of association with humour. Not surprisingly, high intensity laughter appears to be strongly related to humour. McKeown [76], has argued that humour production is a hard-to-fake signal of creativity (following Miller [78] and Greengross and Miller [38]) and of mind-reading ability. Correspondingly, McKeown *et al.* [73, 74] have argued that laughter also serves as a hard-to-fake signal of humour appreciation. The construct of exhilaration indeed describes the effective response to humour (Ruch [106]; [109]). The laugh features that arise with increasing laugh intensity are important to convince an interlocutor that the laughter is related to a genuine felt emotion state. Yet, the boundaries of where this transition to laughter associated with the felt state of amusement and laughter that serves more socio-communicative functions remain ambiguous.

Beside felt amusement following humour, laughter is actually related to other enjoyable emotion states. Ekman [28] proposed that there are at least sixteen enjoyable emotions rather than simply a global emotion of joy. He hypothesized that even though they would all go along with the Duchenne display, namely, the joint and symmetric contraction of the zygomatic major and orbicularis oculi muscles (pulling the lip corners back- and upwards and raising the cheeks and compression of the eyelids causing eye wrinkles, respectively), the main differences in the facets of joy would lay in the parameters, such as timing of onset and offset or intensity. While investigating the responses of those with a fear of being laughed at within ILHAIRE, Platt *et al.* [99] discovered that, of those sixteen emotions, some were consistently associated with the expression of laughter (e.g., amusement, relief, tactile pleasure, schadenfreude).

Finally, although there are individual differences in the susceptibility or willingness to engage in the moment, signals that are displayed when positive emotions are being experienced are contagious, in as much as others who decode those signals will often feel enjoyment (Hatfield *et al.* [41]). Again, laughter has a strong role in maintaining social connection.

Overall, the roles and functions of both conversational and enjoyment laughter, and the laughter linked to the experience of and the contagion of positive emotions will be key for the smooth interaction with virtual agents.

3 Characteristics, Perception and Effect

Laughter hence appears as one of the most important non-verbal vocal social signal. But beyond understanding its roles and functions and knowing the contexts where it is relevant, it will also be crucial to understand more precisely what are its different characteristics (also referred to as features here under), and how these affect the way the signal is perceived, and impact on people. Laughter can indeed be very varied, and although in many cases it has a strong inherent ambiguity, we will see that specific laughter features can have a significant impact on its perceived naturalness, emotional color (valence, arousal, dominance), maliciousness (or alternatively friendliness), and contagiousness. Besides, inter-individual difference in experiencing laughter signals have been identified, including people with a fear of being laughed at, as well people with autism spectrum disorders. It is hence fundamental when creating virtual agent equipped with laughter to understand which of the social signal facets need to be designed with care.

Works on laughter perception can be divided into subtopics, guided by the modality that was the focus of investigation (face, voice, body). Before ILHAIRE, most work has been done on the perception of only auditorily presented laughs. Within ILHAIRE, we have worked on all three modalities and investigated the perception of naturally occurring laughs (spontaneous), acted laughs, manipulated laughs, and virtually portrayed/synthesized laughs. Novel insights have been obtained, and some

of those then guided technological and experimental developments, exposed in Sections 5, 6 and 7.

3.1 Perception of Facial Features

Coeval writers of Charles Darwin, stemming from the historic German field of Ausdruckspsychologie (expression psychology; for example Piderit, 1867 or Bore, 1899, cfr [110]) delivered extensive descriptions of the vocal and facial markers of qualitatively different laughter types. Thus, the perception-related studies within IL-HAIRE started with two investigations of the facial features of laughter basing on historic knowledge [48, 110]. The two studies concentrated on the facial features of four different laughter types (joyful, intense, schadenfreude laughter, and grinning). These four laughter types were chosen because a) at least four historic authors had described them in their laughter classification, and b) the authors had delivered a visual illustration as well as verbal description (cfr. [110] for details). A total of 18 illustrations were first examined for their facial features with a technique allowing for the objective assessment of all visually discernible facial actions (Facial Action Coding System [29]). Then, the decoding of these laughter types by laypersons was investigated in two online studies. The results showed that illustrations of laughter involving a Duchenne Display (DD; the symmetric and simultaneous contraction of the zygomatic major muscle and orbicularis oculi, pars orbitalis muscle) were perceived as joyful laughter, irrespective of their initial classification by the historic writers. Only the DD configuration could be reliably morphologically differentiated and was recognized at high rates. In intense laughter, the intensity of the FACS coded zygomatic major muscle action predicted the perception of intensity by the laypeople. The proposed changes in the upper face highlighted in the literature, i.e., the presence of an additional eyebrow-lowering frowning, did not predict the perception of intensity. Even more, the presence of eyebrow-lowering frowning was antagonistic to the perception of joy. Schadenfreude and grinning did not have high recognition rates, but these displays were also highly heterogeneous in their portrayals. For schadenfreude laughter, two hypotheses were put forward [48]: Schadenfreude may either be a blend of a positive and negative emotion (entailing facial features of both), or expressed by a joy display with regulation or masking attempts (as it is not socially desirable to laugh at the misfortune of others, [28]). Hofmann and colleagues have tested these hypotheses in two encoding studies within the IL-HAIRE project [45, 47], showing that indeed, schadenfreude was often dampened or down-regulated when expressed in social contexts. While many historic writers had claimed facial morphological differences from joyful to intense joyful laughter, our decoding studies did not support the proposed changes, but indicated that the presence of markers beyond the Duchenne markers did not increase the perceived intensity, but led to a change in the perceived valence of the laughter. If eyebrow-lowering frowning (a proposed marker of laughter intensity) was present, the laughter was consequently rated as more malicious. Therefore, we investigated this notion fur-

ther within ILHAIRE by manipulating the presence or absence of eyebrow-lowering wrinkles in synthesized avatar laughter. Basing on synthesized laughter animations with refined facial wrinkles [85, 93], two studies were conducted to investigate the influence of the presentation mode (static, dynamic) and eyebrow-lowering frowning on the perception of laughter animations of different intensity [42]. In a first study, participants (N = 110) were randomly assigned to two presentation modes (static pictures vs. dynamic videos) to watch animations of Duchenne laughter and laughter with added eyebrow-lowering frowning. Ratings on the intensity, valence, and contagiousness of the laughter were completed. In a second study, participants (N = 55) saw both animation types in both presentation modes sequentially. Our results confirmed that the static presentation mode lead to eyebrow-lowering frowning in intense laughter being perceived as more malicious, less intense, less benevolent, and less contagious compared to the dynamic presentation, just as we found in the study of historic illustrations. This was replicated for maliciousness in the second study, although participants could potentially infer the "frown" as a natural element of the laugh, as they had seen the video and the picture. Hofmann (2014) concluded that a dynamic presentation is necessary for detecting graduating intensity markers in the joyfully laughing face. While these studies focused on general differences in laughter perception, we also investigated inter-individual differences within the ILHAIRE project. Hofmann and colleagues (2015) [44] studied the responses to photos of different smiles and laughter and found that gelotophobes assigned the joyfully laughing face not only joy, but also contempt. Thus, for gelotophobes, the "smiling face may hide an evil mind". Ruch and colleagues [113] looked at how the fear of being laughed at (gelotophobia, cfr. [112] for a recent review on this topic) influenced the perception of laughter animations (face and upper body avatar portrayals, synthesized laughter sounds with four different modifications, faceless full body stick figure animations). For the perception of the face, the results showed that gelotophobes found medium intensity laughs that gave the impression of being contrived or regulated as most malicious.The shape and appearance of the lips curling induced feelings that the expression was malicious for non-gelotophobes and that the movement round the eyes, elicited the face to appear as friendly. This was opposite for individuals with a fear of being laughed at: they perceived those features as indicative of maliciousness.

3.2 Perception of Acoustic Features

Most previous studies on laughter acoustics have focused on the decoding of natural, posed, and manipulated laughs. These studies of laughter mainly followed the notion that single laughter elements [2] and changes in acoustic parameters are important for the identification and evaluation of a given laugh [56, 57]. In Table 1, findings on the perception of laughter features are summarized (adapted from [43]).

---

[2] Laughter elements correspond to individual bursts of energy, whose succession is characteristic of laughter.

Table 1 Findings on the Perception of Laughter Acoustics.

Dimension	Acoustic Features	Stimuli
Naturalness/ Realness/ Genuineness	• Serial patterns with varying parameters = ratings close to natural laughs • Faster is perceived as more real/natural • Series with stereotyped patterns are perceived as less natural and genuine • Descending F0 in successive laughter elements is perceived more genuine than laughter series of elements with a constant fundamental frequency	Natural + manipulated laughs [55, 56, 57, 8, 126]
Arousal	• More rapid is perceived as higher aroused • Higher laugh rate is perceived as higher aroused • Lower inter-bout duration is perceived as higher aroused • Higher pitch is perceived as higher aroused • Higher levels of high-frequency energy (F1) is perceived as higher aroused	Posed laughs [120]
Dominance	• Higher intensity parameters more dominant • More precise articulation (lower F0 band width, lower jitter) more dominant • Energy more strongly concentrated in the high frequency range more dominant • Prolonged vocalic segments more dominant • Temporal distance between bouts shorter perceived more dominant • Lower harmonic energy (less voiced elements) perceived more dominant	Posed laughs [120]
	• Dominance is predicted by the interval from vowel to vowel (58%)	Natural + forced laughs [61]
	• Dominance is predicted by F0 (mean, max) of the noise vowel reiteration (31-31%) • Dominance is predicted by Small vs. large amplitude diminishment (31%)	
Valence	• Higher number of segments, higher laugh rate, lower inter-bout duration is perceived more positive	Posed laughs [120]
	• Voiced laughs are perceived more positive than unvoiced	Natural laughs, modified laughs [55, 56, 57, 19, 4, 2, 3]
	• Duration of the initial expiratory noise predicts 42% of the positive valence rating	Natural, forced laughs [61]

F0 = Fundamental frequency.

Table 1 shows that voicing is a potent predictor of the perception of the positive valence in laughter [19, 4, 2, 3]. Basing on this finding, many studies concentrated on voiced laughs and modified acoustic parameters of such laughs to investigate perceptual changes. F0 variations were found to influence the perception of valence, arousal, and dominance (see Table 1). For example, the descending fundamental frequency (F0) in successive laughter elements was evaluated as more friendly or genuine than laughter series of elements with a constant F0 [55]. In subsequent work, [56] modified naturally occurring laughter in different laughter series. The results indicated that experimentally modified series with decreasing pitch and variable duration, as well as the series with sub-phrases, were evaluated as good as spontaneous laughter. In particular, series with decreasing parameter courses such as decreasing pitch or declining durations in successive elements were rated as "friendly" and "laugh-like". Also, series with rhythmic patterns "long-short" and "long-short-long" (accents within the rhythm of a laughter-series) evoked more smiles and laughs in listeners than all other series [55, 56, 57]. Moreover, stereotyped patterns in the course of the F0 received less good evaluations. Kipper and Todt [56] concluded that the evaluation of laughter depends on the dynamic changes of acoustic parameters in successive elements of laughter.

A different methodology was utilized by Tanaka and Campbell [121]. In their study, students first labeled laughter examples from natural conversations in four pre-defined laughter categories. Most laughs were categorized as sounding "polite" or "mirthful". Second, they performed an acoustic analysis of all the laughs that were labeled as "polite" or "mirthful" to identify acoustic features discriminating between these two types. Polite laughter was related to low maximum power and mirthful laughter to high maximum power, long duration, and a high number of bouts. They further reported that the best predictors of the two laughter types were variations in F0 (mean and maximum value), the number of bouts, power, spectral slope, and measures of prosodic activity. With these features, a classification was performed, leading to 79% classification accuracy between polite and mirthful laughter.Thus, listeners distinguished between laughter types in auditory stimuli, which indicated that those types have a distinct signal value. Nevertheless, there were also group-related differences. For autistic individuals, differences in global laughter evaluations were found. Hudenko and Magenheimer [50] found that autistic childrens voiced laughs were perceived as more positive than normally developed childrens voiced laughs. The latter were generally lower pitched, and shorter (but no differences in F0 were found). When comparing individuals with a fear of being laughed at to individuals with no fear, Ruch and colleagues [113] found within ILHAIRE that the fundamental frequency modulations and the variation in intensity were indicative of perceived maliciousness. Fast, non-repetitive voiced vocalizations, variable and of short duration were perceived as most friendly by individuals with a fear of being laughed at.

3.3 Perception of Bodily Portrayals

Enjoyable emotions are aligned with laughter, that involves open mouth smiling, vocalized laughter sounds, but also bodily changes such as dropping or relaxing the shoulders, and shaking of the trunk. Within ILHAIRE, another strong focus did hence lay in the investigation of laughter in the body and the perception of such cues. Within ILHAIRE, Griffin and colleagues [39] analyzed participants perception of laughter from body movements. The participants task was to categories animations of natural laughter from motion capture data replayed using faceless stick figures (characters with trunk, limbs and heads simply represented by edges). In general, animations that were perceived as representing a laugh differed in torso and limb movements compared to stimuli categorized as non-laughter. Also, the distinguishing features differed for laughter stemming from sitting or standing avatar positions. Perceived amused laughter differed from perceived social laughter in the amount of bending of the spine. Similarly, Mancini *et al.* [66] found that laypersons were generally very good in distinguishing full-body animations of laughter from non-laughter (79.70% of the stimuli were categorized correctly), with high levels of confidence in rating either stimulus category. When assessing the perceptions of gelotophobes and non-gelotophobes, Ruch and colleagues [113] found that in the virtual body portrayals (faceless full body stick figures), backwards and forward movements and rocking vs. jerking movements distinguished the most malicious from the least malicious laugh.

3.4 Perception of Multimodal Portrayals

Sestito and colleagues [116] investigated the decoding of audio-visual laughter stimuli. They found that the correct decoding of laughter was high above chance rate and that in audio-visual incongruent stimuli, the visual modality was prioritized in the decoding over the acoustic dimension. Using electromyography measurements of the zygomatic major muscle activity, they also reported that rapid and congruent mimicry toward laughter stimuli occured. Within ILHAIRE, several studies have investigated the perception of multimodal portrayals of human laughter (spontaneous and acted/ fake laughs; presented with visual-auditory stimuli). McKeown and colleagues [73] conducted two experiments to assess perceptions of genuine and acted male and female laughter and amusement facial expressions. The main results showed that participants were good in detecting fakeness in laughs by males. When women faked laughs, males distinguished cues of simulation, but judged fake laughs also to be more genuine. When judging other women, female participants perceived genuine laughs to contain higher levels of simulation.With a focus on inter-individual differences, Ruch, Hofmann, and Platt [111] investigated the verbal and facial responses of 20 gelotophobes and 20 non-gelotophobes towards videos of people recalling memories of laughter-eliciting positive emotions (amusement, relief, schadenfreude, tactile pleasure). The facial expressions of the participants

were clandestinely filmed and evaluated by the FACS [29]. Smiles of enjoyment and "markers of contempt" were coded and verbal ratings of the participants obtained. Gelotophobes responded with less joyful smiles and with more expressions of contempt to laughter-eliciting emotions than did non-gelotophobes. Gelotophobes also perceived the degree of joy expressed by participants in the video clips of tactile pleasure and relief lower than non-gelotophobes. No differences occurred in the perception of joy for schadenfreude and amusement. Thus, spontaneous affective responses and cognitive responses through ratings have to be distinguished.

4 Naturalistic Databases

At the start of the ILHAIRE project, there were a small number of existing databases that provided examples of laughter for research purposes. The most useful of these were the AudioVisual Laughter Cycle database (AVLC) (Urbain *et al.* [130]) and the MAHNOB laughter Database (Petridis *et al.* [94]), both of which contain laughter from individuals watching funny video clips. While these databases serve their purpose well, they are limited to a certain style of laughter and context and both had similar aims in the style of laughter targeted. The ILHAIRE project sought to collect a much larger amount of laughter and from a much more diverse range of settings and contexts than had been previously gathered. The goals of the ILHAIRE laughter database were to collect laughter from a broad range of contexts. Thus in addition to collecting laughter of people observing amusing scenarios (e.g. watching comedy), we also targeted laughter occurring in social interaction and, importantly, in situations that led to what was termed hilarious laughter and what was deemed to be more social and conversational in style. There were further goals of collecting laughter data from more than one culture and from interactions that took place in more than one language. In addition the range of sensors that gathered the information was to be extensive, including high quality audio and video, but also incorporating depth, motion capture information, respiration and facial expression information where possible. The project also sought to provide as much annotation as resources would allow and devise an annotations scheme for laughter to facilitate this.

As a result of these diverse goals the ILHAIRE laughter database is not a database in the traditional sense but something of a meta-database; it incorporates a number of different databases and sources with the overarching goal of providing a useful laughter-focused set of resources and stimuli for the research community. There were three main phases to this: the collection and annotation of laugh stimuli from existing databases (also summarized in McKeown *et al.* [70]), the collection of hilarious laughter, and finally the collection of conversational and social laughter. The first phase is distinct due to the nature of the task. However, the other two phases are less distinct as it is not straightforward to define what distinguishes a hilarious laugh and a social laugh and these laugh types both commonly occur in any given

social interaction and as argued on the basis of ILHAIRE research the boundaries between these laughs are ambiguous (underdetermined).

4.1 Existing Databases

Belfast Naturalistic Database

The Belfast Naturalistic Database (Douglas-Cowie *et al.* [24]) was an early attempt to gather a broad swath of audio-visual material of people who at least appeared to be experiencing genuine emotion with material sourced mainly from television programmes. 53 of the total of 127 video clips contain laughter, but only five can be made available due to copyright issues.

HUMAINE Database

The HUMAINE database (Douglas-Cowie *et al.* [25]) was created with the purpose of demonstrating the existing breadth of material related to a broad understanding of the word emotion - termed "pervasive emotion". From fifty video clips 46 instances of laughter of variable quality were extracted, and are useful as illustrations of the variety of situations in which laughter occurs.

Green Persuasive Database

The Green Persuasive Database (Douglas-Cowie *et al.* [25]) contains audiovisual clips recorded to capture interactions with strong feelings, but not basic emotions. The scenario involves one participant convincing another to adopt an environmentally friendly lifestyle. There is a strong power imbalance between participants as the persuader is a University Professor and the listeners are students. There were eight interactions in total lasting between 15 and 35 minutes. From these eight participants, 280 instances of laughter were extracted.

Belfast Induced Natural Emotion Database

The Belfast Induced Natural Emotion Database (BINED) (Sneddon *et al.* [118]) represents a deliberate effort to induce specific kinds of emotional behaviour. The goal was to use a series of tasks to generate spontaneous and dynamic emotional material that could replace the posed static photographs often used in studies of emotion. Laughs were extracted from Set 1 of the database including tasks designed to elicit: amusement, frustration, surprise, disgust and fear. 289 instances of laughter were extracted from a total of 565 clips with 113 participants (43 female, 70 male).

SEMAINE Database

The SEMAINE database (McKeown *et al.* [75]) provides high quality audio-visual clips from a Sensitive Artificial Listener (SAL) task. In this task one participant took the role of the user and another played the role of an embodied conversational agent, using one of the four SAL characters in the SAL system. The laughter in these interactions was largely conversational and social, and was incidental to the task of interacting with the avatar or with a person pretending to be an avatar. In total 443 instances of laughter were extracted from 345 video clips.

4.2 Hilarious Laughter Collection

UCL Motion Capture Stick Figure Stimuli

There were a variety of data gathering sessions dedicated to gathering hilarious laughter. One of the goals of these sessions was to gather data related to body movement, so a focus was on motion capture elements. Two data gathering sessions were dedicated to collecting this data. One in Belfast developed the laughter induction techniques (McKeown *et al.* [72]), and was followed by a similar session conducted at UCL in which the motion capture data was made available as part of the database (Griffin *et al.* [40]; McKeown *et al.* [71]). The available data consists of 126 animated "stick figure" video stimuli of laughter that has been categorized as either hilarious, social, fake, awkward or not a laugh.

Multimodal Multiperson Corpus of Laughter in Interaction (MMLI)

The MMLI database (Niewiadomski *et al.* [80]) focused on gathering multimodal full body movement laughter data. This data was collected during recording sessions of French speakers made in Paris. It contains both induced and interactive laughs from human triads. 500 laugh episodes were collected from 16 participants. The data consists of 3D body position information, facial tracking, multiple audio and video channels as well as respiration data.

Belfast StoryTelling Corpus

The Belfast Story-telling corpus was comprised of six sessions of groups of three or four people telling stories to one another in either English or Spanish. The storytelling task was based on the 16 Enjoyable Emotions Induction Task [49]. Participants prepared stories related to each of 16 positive emotions or sensory experiences and were seated around a central table, and each participant wore a head-mounted microphones. HD webcams and depth cameras (Kinect) captured audiovi-

sual streams, facial features, face direction and depth information. Participants took turns at recalling a story with occasional open discussion. Synchronized recording of data streams was achieved using the Social Signal Interpretation (SSI) framework [135].

4.3 Conversational Laughter Collection

Although there was much laughter that could be termed conversational laughter within the Belfast Story-telling Database, the activity that the participants were engaged in was not strictly a conversation. To ensure that the database contained laughter that was taken from conversations between people, we devised a very minimal task to capture conversations that were as natural as possible given the presence of cameras, microphones and depth sensors, and between only two participants. There were two versions of this task, one recorded in Belfast, and the other in Peru. Participants were asked to talk on a topic randomly selected from a pre-determined list, but to continue talking freely until they felt a new topic was needed. Sessions lasted for an hour. Some dyads used up to 10 topics in their session, whereas many used only one.

Belfast Conversational Dyads

While the task differed from the one used in the Belfast storytelling database (the random topic task instead of the 16 enjoyable emotions task) the recording set up was almost identical. HD webcams, head-mounted microphones and Kinect sensors were used to record interactions between participants who sat opposite each other. The various data streams were once again synchronized using the SSI software. 10 pairs were recorded.

Peruvian Conversational Dyads

A mobile version of the data capture system was devised and taken to Peru to capture interactions between Peruvian conversational pairs. Unfortunately, the depth sensors could not be made to function, but the same HD webcam and headmounted-microphones were sued to ensure quality recordings. The data streams were synchronized using the SSI software, and 20 interacting dyads were recorded. These recordings involved the speakers interacting in Spanish.

4.4 Annotation

An annotation scheme, which had seemed a straightforward endeavor at the outset of the ILHAIRE project, proved to be a much more difficult proposition. We have argued that the inherent ambiguity and underdetermined nature of laughter mean that it does not yield easily to simple a classification system (ref underdetermined). However, within the project we did develop a set of guidelines for the segmentation of laugh episodes; these were used to create a significant amount of annotated material that was then rated along a number of dimensional scales. A substantial amount of laughs have been annotated and rated and this information is available at the ILHAIRE laughter database website [3]. In conclusion, the ILHAIRE laughter database has achieved its goals and now represents the most comprehensive set of laughter resources currently available for use by the research community. It remains an active research database and new materials and annotations continue to be added. We encourage researchers to send us any annotations or data gathered using the database materials so we can incorporate them into the database for the use of other researchers.

5 Automatic Detection, Recognition and Characterization

Systems that understand and can also positively impact on social communication require algorithms for the detection and recognition of laughter.

This can rely on knowledge, models and technologies in sensor systems, signal processing, machine learning, and in particular related to automatic detection and classification of audio-visual signals. Microphone and video cameras can provide raw signals to analyse acoustic and visual modalities. With the availability of cheap depth cameras (providing for each pixel a measure of the it is also possible to get improved accuracy in the capture of 3D cues from the face and the body, where video-only cameras could otherwise fail. In some cases, researchers can also have recourse to more specific equipment, such as respiration sensors, or accurate tracking solutions, for instance based on optical motion capture and gaze tracking.

Then, it is necessary to research on the features to be extracted from the raw sensor signals, on the categories and characteristics that are useful to be recognized and measured, and on the models to be used. This can rely on knowledge about the morphological features of laughter (cfr. Section 3) that really matter.

As machine learning remains a state-of-the-art tool for detection and recognition, also in this area of social signal processing, large annotated corpora containing many instances of the phenomena of interest are also necessary (cfr. Section 4).

Early work essentially focused on designing laughter detectors from audio signals and recognized facial expressions. Such work was actually fueled by research projects starting around the year 2000. At the time, data collection was first achieved

---

[3] http://www.qub.ac.uk/ilhairelaughter

through microphones only, sometimes in naturalistic social settings, such as group meetings [52]. Multimodal capture systems including video cameras came a bit later. Since 2010, it is also possible to capture 3D information using cheap hardware, to the benefit of research on leveraging body motion and multimodal approaches. However, this had not been done for laughter analysis before ILHAIRE. Most modalities have actually been further considered, and multimodal approaches for combining them have been proposed. Studies covered detection, characterization of important dimensions such as laughter intensity, and classification of laughter in generic categories. On the side of acoustic analysis, the project came up with novel approaches for making use of automatic recognition of either phonetic or else purely data-driven symbolic units whose local statistics have been used for laughter recognition, yielding beyond state-of-the-art performance. Also, a special focus has been on estimating laughter intensity, one of its most important facet, with first studies on that aspect. Besides, to our knowledge, there was also no previous work on automatic laughter detection and analysis from body movement and respiration signals. The project closed these gaps. Finally, on the side of multimodal integration, a novel approach outperfoming previous proposal of early and late fusion has also been published. The following sections are organized according to the covered modalities.

5.1 Acoustic

Early studies on laughter detection often used acoustic features initially designed for speech analysis as well as automatic recognition, classification, statistical modeling and time series modeling, in particular Hidden Markov Models (HMMs) with Gaussian Mixture models (GMMs). Spectral coefficients and HMMs have been used in [64]. In [11], Mel-Frequency Cepstral Coefficients (MFCCs) and perceptual features were applied together with HMMs. In [54], the authors used a different approach relying on Support Vector Machines for classification, still fed with MFCCs though. In [123], the authors used Artificial Neural Networks (ANNs) fed with Perceptual Linear Prediction (PLP) features. In [60], a similar approach but with longer temporal feature windows allowed to reach better performance. Reported results in terms of detection equal error rate (false positive rate and false negative rate) range from about 30% down to 8% in [60]. A figure of 3% has been achieved but using speaker dependent models and pre-segmented laughs.

Within ILHAIRE, such approaches have been applied while novel techniques have also been developed. Although laughter is mostly non-verbal and non-articulated, it nevertheless exhibits some vocal tract configurations close to phonetic sounds. Automatic laughter transcription through phonetic labels has hence been found useful [127]. Following up, in [134], detection has been addressed through a two stages approach. Instead of estimating the probability of laughter using a model based on acoustic features directly as in previous work, it first uses a generic phonetic recognizer to obtain a symbolic sequence. Short audio segments are then characterized by

one or two histograms describing the distribution of symbols in the sequence. These constitute feature vectors on which to base a classification model. When combined with more traditional features, it was possible to detect laughter and filler events with 88% accuracy (unweighted average area under curve (AUC)), hence 4% absolute above a baseline approach on a published benchmark [115]. A related approach is to use n-gram models of such symbolic sequence to model the patterns of laughter, as evaluated in [91]. A detection F-measure of 75% has been reached. An additional specificity of that development is to use non-phonetic symbolic units that are defined automatically using the Automatic Language Independent Speech Processing (AL-ISP) method, hence presenting some potential for improved modeling of non-verbal vocal sounds.

Categories and Intensity

Once laughter is detected, one would want to identify some of its important characteristics. Early results by others on recognizing laughter categories have been obtained in [12], where phonetic features and HMMs were applied to classify among four types (hearthy, amused, satirical and social) with an accuracy close to 75%.

As explained earlier, an important facet to be measured is laughter intensity, which has been shown to correlate with the social vs. amused categories, while having the advantage of being continous instead of categorical. Within ILHAIRE, in [86, 129], a research on the audio and visual cues that can be extracted automatically and correlate with the perception of laughter intensity has been conducted. On the acoustic side, the range of variation of several of the MFCC coefficients, but also of the spectral flux, loudness and pitch were found to be the most important acoustic features, with a correlation coefficient of the best features with laughter intensity ratings reaching 83%.

5.2 Facial expressions

Recent possibilities for automated tracking of specific facial features or action units have started to be applied to laughter detection and characterization. In [51], spatial locations of facial feature points are tracked using markerless video processing, and used for laughter detection. In [102], principal component analysis of tracked spatial location of feature points is used to obtain features from video signals, and GMMs as well as SVMs were used for classification. In [95, 96], head movement and facial expressions, obtained through facial feature points tracking from the video channel too. These different publications actually applied a multimodal framework, combining acoustic features with visual features reaching detection and classifications accuracies above those obtained when using individual modalities. In [95], an accuracy of 75% to distinguish three classes, namely unvoiced laughter, voiced laughter and speech, is reported.

Within ILHAIRE, evaluations of face tracking approaches for estimating Facial Animation Parameters (FAPs) and intensities of Action Units (AUs) have been performed [101]. A smile detector based on estimated action units has also been used as one of the component of laughter detection and laughter intensity estimation within a multimodal approach, evaluated in interactive settings.

Categories and Intensity

The previously mentioned works [86, 129] also covered the visual side of intensity estimation. It was shown that the maximum opening of the mouth/jaw as well as the lip height and lower lip protrusion were identified as the most important, with a correlation coefficient of the best features with laughter intensity ratings reaching 68%. A study of audio and visual features that differ in laughter inhalation and exhalation phases was also proposed. Some features present different patterns finally enabling to distinguish these phases automatically.

5.3 Body Movement and Gestures

Body movement during laughter has less background work and available corpora to base the studies on. Although some previous studies described the morphological attributes of laughter, it was still necessary within ILHAIRE to gather more detailed statistics related to specific motion patterns appearing during laughter [39]. These studies used recordings done within the project and including motion capture using high-end optical hardware. Following a preliminary visual inspaction, statistics were drawn from laughter segments exhibiting cues from all three modalities (audio, facial and body), hence covering essentially laughter of high intensity (amused emotion). Shoulder shaking has been identified as the most frequent. Then, torso rocking, torso throwing, knee bending and torso leaning were second, each two times less frequent than shoulder shaking. Finally, head shaking and shoulder contraction appeared, but only rarely. Given the importance of torso movement, subsequent research on automatic detection focused essentially on that aspect.

In [67, 132], shoulder movement was tracked either based on a body skeleton extracted from a depth sensor signal, of from shoulder tracking based on machine vision. Features characterizing this motion are then computed, including the correlation between the movement of the two shoulders, the overall kinetic energy, and the periodicity of the movement. Based on that, a system for the automatic detection using commodity hardware (video and depth cameras used in conjunction) has then been proposed and evaluated in [66, 68]. Several motion features were extracted, accounting for shoulder motion correlated with torso (trunk) motion, but also directly torso and head. Evaluation relied on Kohonen self-organising maps, showing significantly above chance estimation results.

Categories and Intensity

The previously referenced work from ILHAIRE [68] also proposed an evaluation of the selected body motion features for automatic estimation of intensity, showing significantly above chance estimation results too.

Another proposal was made earlier [40, 39] within the project where body motion was investigated for laughter type recognition among five categories (hilarious, social, awkward, fake, and non-laughter). Features characterizing hand gesture, shoulder movement, neck/spine bending, as well as kinetic energy of several upper body articulations were extracted. Several classification approaches were compared (k-nearest neighbors, multilayer perceptron, linear and kernel ridge and support vector regression, random forests) with the random forest method yielding the best performance. Tests were made using motion data captured using full-body motion capture equipment. Automatic recognition performance reached about 66%, which was also shown to approach human rating levels.

In addition to movement directly induced by laughter, the phenomenon may also trigger complementary movement due to the amused state or the need to replace speech with other cues (using pointing, clapping, illustrator, or other gestures). Studies towards understanding these are ongoing.

5.4 Respiration and Muscular Activity

Physiological responses such as elevated heart rate can accompany laughter. Also, its production is intimately linked to respiratory patterns, as described in [108]. A distinctive pattern can indeed be observed, consisting in a rapid exhalation followed by a period of smaller exhalations at close-to-minimum lung volume. This pattern, accompanied by contractions of the larynx and epiglottis and facial patterns, result in the specific sound, facial and body movement being observed. On may hence try to detect and characterize laughter directly from respiration measurements, and their underlying muscular activity. In previous work by other [35], myoelectric signals from the diaphragmatic muscle were measured. Detection of laughter was shown to be possible through a threshold-based approach on the amplitude of the high-frequency component of the captured signal. No formal evaluation of the detector itself was proposed however.

Another approach studied in ILHAIRE was to rely on measurements of the thoracic (chest) circumference [132]. Features enabling the detection of laughter from these respiration signals are computed including the following sequence of events: a sharp change in current respiration state, a period of rapid exhalation resulting in rapid decrease in lung volume, a period of very low lung volume. Formal evaluations were performed later, using recognition through HMM models, with a classification accuracy of 69%, validating the approach (results not published yet).

5.5 Multimodal Fusion

Previous studies by others on combining acoustic and visual cues have been published. In [95, 96] (aleady introduced here above), late fusion (decision level) has been applied using a sum rule or an artificial neural network. In [51], the output of the acoustic and facial detectors are combined with an AND operator, hence a form of late fusion too. Decision level late fusion was also used in [102].

Within ILHAIRE, late fusion of estimations from audio, body and respiration modalities has been implemented in [132] for laughter detection. In [63], an event driven real-time fusion system was proposed. It rather corresponds to a late fusion mechanism, with some additional time-based accumulation. This approach proved particularly robust for the case of laughter detection, since it does not directly fuse identical time frames throughout modalities, but rather computes probabilities indirectly by accumulating shorter, detection-indicating and possibly time-shifted events. Evaluation was performed on an enjoyment detection task (enjoyment defined as an episode of enjoyable emotion, which may hence also include just smile segments). From 54% and 72% accuracy for audio-based and image-based detection, the fusion approach reached about 79%, outperforming more traditional early or late fusion schemes (only reaching between 65% and 68%).

Experimental studies on tuning of laughter analysis to genre and culture were also carried out with promising results. They rely on a range of techniques actually enabling the parameters of detection/classification models to be adapted to specific demographic subgroups or subjects, similarly to what is done in speech synthesis and recognition technology.

Finally, for more information, a survey on multimodal fusion within human-agent dialogue has been proposed by André *et al.* [1].

6 Automatic Generation and Synthesis

A proper understanding of the nature of multimodal signals during laughter is necessary not only to inform on the proper features and models for detection and characterization purposes but also on the proper models for generating signals that sound and look natural, or enable experimental protocols. Generation and synthesis is covered here.

6.1 Acoustic Synthesis

Despite previous work on synthesizing "emotional speech" (cfr. review by Burkhardt and Campbell [9]), acoustic laughter synthesis is an almost unexplored domain. In 2007, Sundaram and Narayanan [119] synthesized laughter vowels by Linear Pre-

diction. To obtain the repetitive laughter pattern, they modeled the laughter energy envelope with the equations of an oscillating mass-spring system.

The same year, Lasarcyk and Trouvain [62] compared laughs synthesized by a 3D modeling of the vocal tract and diphone concatenation. The articulatory system gave better results, but synthesized laughs could still not compete with human laughter naturalness. Beller [5] proposed an original approach to laughter synthesis, as voiced laughter is synthesized from a neutral speech sentence.

Sathya *et al.* [114] synthesized voiced laughter bouts by controlling several excitation parameters of laughter vowels: pitch period, strength of excitation and amount of frication. After analyzing these features on a range of human laughs, Sathya *et al.* concluded that the pitch contour and the strength of excitation of laughter calls can be approximated by quadratic functions, while the amount of frication tends to decrease within and across calls.

Cagampan *et al.* [10] synthesized laughs by concatenating syllables. Laughs were segmented into syllables with different labels denoting laughter vowels ('ha', 'he', 'hi', 'ho', 'hu'), grunt- and snort-like syllables as well as laughter onset and offset. These units were then combined to form laughs with four syllables in the apex, plus possibly an onset and an offset.

A real-time laughing instrument has been developed by Oh and Wang [89]. Their main objectives were expressivity and control, rather than quality of synthesis or laughter naturalness. They synthesized vowels by formant synthesis (source-filter decomposition).

Recent works from Oh and Wang [88] to modulate speech and make it sound like speech-laugh, as opposed to all previous attempts on pure laughter. The method takes speech as input and segments it into syllables, based on the energy envelope. Then they provide control over several parameters of the syllables that can be affected by laughter: intensity contour, maximum pitch value, tempo regularity (the degree to which segmented speech syllables are fetched to an isochronous tempo), rhythm (the periodicity of syllables between 4 and 6 Hz).

More recent studies within ILHAIRE followed an HMM-based approach. The method used is based on a framework which became popular in the field of speech synthesis in the last decade. In HMM-based parametric speech synthesis, the spectrum, f0 and duration of phonemes are modeled in a unified framework [136]. Based on the resulting HMM, a maximum-likelihood parameter generation algorithm is used to predict the source/filter features [122], which are then sent to a parametric synthesizer to produce the waveform. Urbain *et al.* exploited this technique to perform HMM-based acoustic laughter synthesis [124, 125, 128, 126, 127]. They investigated the synthesis audio laughter from arousal curves by comparing the arousal curve given as input to the arousal curves of laughter syllables available in the database. The transcriptions of the best matching syllable were used to drive the HMM-based acoustic laughter synthesis system. In [6], Bollepalli *et at.* compared the use of different vocoders for the specific purpose of acoustic laughter synthesis. Using the same approach, "speech-laugh" which is the phenomenon of laughter occurring at the same time as speech by intermingling with it or by interrupting it, an HMM-based speech-laugh synthesis system has been developed by El Haddad *et*

al. [32]. This system involves first creating HMM models of laughter and speech-smile. Then, some vowels in the synthesized speech-smile sentences are replaced by laughter bursts.

6.2 Visual Synthesis

As for acoustic laughter synthesis, very few studies exist concerning visual laughter synthesis. Since only few studies focused only on facial synthesis or body synthesis, this section is organized as a chronological survey of studies related to visual laughter synthesis without further categorization.

In 2008, a parametric physical chest model which could be animated from laughter audio signals was proposed by DiLorenzo *et al.* [20]. The model is able to produce realistic upper body animation but facial animation is not addressed.

The next year, Cosker *et al.* [16] studied non-verbal articulations including laughter. They explored the possible mapping between facial expressions and their related audio signals. Hidden Markov Models (HMMs) were used to model the audio-visual correlation. As for DiLorenzo *et al.*, the animation was audio-driven.

Further research has been pursued within ILHAIRE. In 2012, Niewiadomski *et al.* compared three possible approaches regarding visual laughter animation [83]. The same year, Niewiadomski and Pelachaud [84] consider how laughter intensity modulates facial motion. A specific threshold is defined for each key point. Each key point moves linearly according to the intensity if it is higher than the corresponding threshold. So, if the intensity is high, the facial key points concerning laughter move more. In this model, facial motion position depends only on laughter intensity.

More recently, in 2013, two studies [132, 79] included the use of laughter capable avatars for human-machine interactions. Two different avatars animated from recorded data are proposed. One of them is the Greta Realizer [82] which takes as controls either high level commands using the Facial Action Coding System (FACS) or low level commands using Facial Animation Parameters (FAPs) of the mpeg-4 standard for facial animation. Greta generates an animation corresponding to the initially recorded laughter (copy-synthesis). The other avatar is the Living Actor [4] which plays a set of manually drawn animations.

A recent study published in 2014 aiming at synthesizing facial laughter was proposed by Çakmak *et al.* [15]. The approach followed was to model facial deformations by means of landmark trajectories. The basic steps followed throughout the work are : recording of the 3D data using a motion capture system, post-processing to shape the data for training, training of Hidden Markov Models (HMMs) on this data, synthesizing trajectories based on this data and retargeting the synthesized trajectories on a 3D face model to finally render a video output. To be able to build accurate models, this study needed the building of a specific audio-visual laughter

---

[4] http://www.cantoche.com/

database containing synchronous audio and 3D motion capture data in sufficient amount for a single subject [13].

The same year, Ding *et al* [21, 23] developed a generator for face and body motions that takes as input the sequence of pseudo-phonemes and their duration [21, 22]. Lip and jaw movements are further driven by laughter prosodic features based on a contextual Gaussian Models approach. The relationship between input data (pseudo-phonemes and acoustic features) and motion features is first modeled and then the model is used to produce laughter in real-time. Head an eyebrow generation is based on the selection and concatenation of motion segments from the database. Torso and shoulder motions are derived from head motion.

Another study in 2014 by Niewiadomski *et al.* [81] propose a procedural method to synthesize rhythmic body movements of laughter based on spectral analysis of laughter episodes. For this purpose, they analyze laughter body motions from motion capture data and reconstruct them with appropriate harmonics.

Niewiadomski and Pelachaud [85] studied the identification and perception of facial action units displayed alone as well as the meaning decoding and perception of full-face synthesized expressions of laughter.They focused on three factors that may influence the identification and perception of single actions and full-face expressions: their presentation mode (static vs. dynamic), their intensity, and the presence of wrinkles. They used a hybrid approach for animation synthesis that combines data-driven and procedural animations with synthesized wrinkles generated using a bump mapping method.

7 Interaction Modeling

Enabling conversational agents with laughter capabilities is not only about being able to recognize and synthesize audio-visual laughter signals. It is also concerned by an appropriate management of laughter during the interaction. There is thus a need for a laughter-enabled interaction manager (IM), able to decide when and how to laugh so that it is appropriate in the conversation. Despite the body of work in so-called dialog modeling, there was no previous work specifically adressing laughter as indeed, previous research was essentially focused on verbal social communication.

It remains uneasy to define what an appropriate moment to laugh is. This can be seen as a decision making process. These decisions have to be taken according to the interaction context which can be inferred from laughter, speech and smile detection modules (detecting social signals) but also by the task context (for example, if the human is playing a game with the agent, what is the status of the game). Formally, the IM is thus a module implementing a mapping between interactional contexts and decisions. Lets call this mapping a policy.

Describing the optimal policy of the agent is a very tricky task. It would require the perfect knowledge of rules prevailing to the generation of laughter by humans. Interpreting sources of laughter or predicting laughter from a cognitive or psychol-

ogy perspective is non-trivial. Therefore, a data-driven method has been preferred in the ILHAIRE project. Especially, we adopted a Learning from Demonstrations (LfD) framework to learn the IM policy. Indeed, humans are implementing such a policy and they can provide examples of natural behaviors.

LfD is a paradigm in which an artificial agent learns by observing another agent (artificial or human) performing optimally the task at sight. Several generic methods can be used to implement this paradigm among which two have been explored within the ILHAIRE project: 1) Imitation Learning (IL); 2) Inverse Reinforcement Learning (IRL).

Imitation learning reduces the problem of learning the optimal policy to a classification problem. Indeed, one can see the policy as the result of a process that assigns a decision to an interactional contexts which is similar to the standard classification problem consisting in assigning labels to inputs. For instance, a simple K-nearest neighbors algorithm has been used in [79]. Taking as input the results of audio-visual processing systems able to detect speech and laughter of other participants, this method generated a laughter/silence decision every 200 ms. Yet, classification algorithms usually underperform when trained on unbalanced data. This is the case with laughter which is way less frequent than speech and silence in human-human interactions. For this reason, we introduced a structured-classification method in [97], enabling to emphasize more on laughter and improve classification performances.

Inverse Reinforcement Learning considers laughter as a sequential decision making process. In this framework, the decisions taken by the conversational agent at a given time are supposed to have an impact on the reactions of the user(s) right after or even later and, so, on the course of the interaction. Therefore, the optimization of this module should take the whole interaction into account, including the impact of local decisions onto the future. Sequential decision making processes are generally addressed under the Reinforcement Learning (RL) paradigm in the machine learning literature. In RL, the agent is assumed to optimize a cumulative function of immediate rewards, provided after every decision. Because it learns to maximize the sum of rewards, it learns a sequence of decisions instead of local decisions as in imitation learning. Yet, the problem of defining the reward remains. IRL is a method by which, observing an expert agent, another agents learns the unknown reward the expert is optimizing. This is unfortunately an ill-posed problems. Indeed, the null reward makes any decision policy optimal. In addition, IRL often requires observing expert and non-expert data which is obviously impossible in laughter studies (what is a non-expert laughter?). During the ILHAIRE project, we thus developed innovative IRL algorithms that could learn non-trivial rewards from expert-only data [58, 59] and applied them successfully to laughter/silence decision in the final project demonstrator.

8 Application Perspectives

Many applications can be imagined, some of which already having been prototyped. On the side of non-interactive ones, the automatic annotation of laughter events and their characteristics could be useful for searching content and gathering interaction statistics from conversational databases, such as meeting recordings. Laughter presence has shown relation with dialog turns and topic changes and, because of this role as regulator of conversation, its detection could facilitate the interpretation and enriched transcriptions of such content and detecting jokes, laughter being often related to amusement. Beyond that, laughter use imply different perceptions of emotional valence, arousal and dominance, which would provide an additional layer of enrichment, that should help in inferring roles and moods within such multimodal conversational data. The interpretation of responses to laughter would also offer indications about the emotional state of the subjects, as well as their personality. There are indeed demonstrated inter-individual differences in the way laughter is perceived and which feedback expressions it triggers in return.

Further ahead, the proper integration of "laughing skills" into interactive systems would offer several benefits, that we structure here under into four aspects: content, emotion, social bonding, and personality-related.

First, laughter would not only make them more natural, but would also offer an additional tool for such systems to influence on the course of the interaction and hence its content. Using computers to regulate multi-party conversations/meetings has already been suggested in [104]. This remains to be researched whether such system will face negative perceptions (akin to the uncanny valley), how much work will be necessary to bridge that gap if any, and what are the ethical concerns that need to be sorted out.

Being a very frequent natural indicator and determinant of positive moods and emotions (amusement but also other enjoyable emotions), such laughing machines would then be better equipped to align their expressive behaviour to the emotional context of the interaction. This would lead to increased naturalness, as well as impacts on people. Positive mood changes induced by conversational systems equipped with laughter have been demonstrated within ILHAIRE and by others in media consumption setups [131, 79, 46, 35], "affective" installations [77, 117, 92], and experiments are also covering games with a virtual conversational agent [65]. More specifically, in Niewiadomski *et al.* [79], Hofmann *et al.* [46], When a virtual agent was interactively responding to laughter, the level of contagion associated with spontaneous laughter as well as exhilaration was experienced by the users (Niewiadomski *et al.* [79], Hofmann *et al.* [46]).

Some of the positive effect of laughter on learning have also been studied by others [34]. As a consequence, laughter could contribute to improved motivation and self-image in various applications, such as learning, rehabilitation or fitness exercises. In ILHAIRE, a virtual tutor equipped with laughter increased the positive perception of the learning experience, if not the learning outcome itself (results not published yet). The fact that some properties of laughter trigger different perceptions valence and arousal would only enhance the active repertoire of such systems. Here

again, ethical issues related to the flip side of laughter when it is associated with negative feelings (such as ridicule) will need to be considered.

Laughter is also an important marker of social bonding, reducing the sense of threat and facilitating cooperation in a group [90, 26]. It could then be used within computer mediated inter-personal communication setups to favour collaborative and positive outcomes, and to complement inter-personal skills of interacting humans. Laughter, together with its accompanying body postures and movements, can also convey sexual bonding messages that range from solicitation to aversion, depending on which and how many different signals are present [37]. Novelists and filmmakers may be better equipped to speculate on such functions being used by computer systems.

Empathy is a significant and particularly interesting component of social bonding. Recent experiments provide support for the hypothesis that anthropomorphism positively affects empathy towards robots. This was tested in [103], where Riek *et al.* used robots (along a chosen anthropomorphic range) shown to be experiencing mistreatment by humans, hence targeting the pro-social behavioral component of empathy (i.e. the component of empathy that leads to helping behavior toward others in need) in the subjects. Although those experiments were covering the physical appearance of anthropomorphism only, the authors speculate that these results are in fact compatible with Simulation Theory which states that people mentally simulate the situation of other agents in order to understand their mental and emotive state. Laughter being so pervasive, it is also likely to help in getting humans to empathize with virtual agents and robots.

In all these, the underlying social context is a fundamental facet. But personality traits constitute another crucial element to be considered, especially given the inter-personal differences in perception and feedback. Analysing such reactions hence offer information that could enable systems to better understand and adapt to individual personality. Gelotophobes constitute a particular group which has been strongly considered within ILHAIRE. They exhibit a fear of being laughed at and display negative expressions in reaction. This constitutes a significant social handicap given the pervasiveness of laughter in society. Another opportunity to be pursued in the future would hence be to build interactive systems enabling interventions designed to help such population to either participate more frequently to social activities through virtual agents designed with laughter expressions that are not perceived negatively, or better combat their fear through progressive accustomization to natural laughter expressions.

We believe these applications will be facilitated by technologies for the recognition, understanding, interaction modeling, generation and synthesis of a wide range of laughter expressions, adaptable to a wide range of social contexts, contents, moods and personalities; and grounded in strong knowledge from both psychology and machine learning.

9 Research Perspectives

Bringing together the fields of engineering and psychology can only strengthen both. Understanding the psychological impact of the interface between computer and human allows for the evaluation of the AIs success. However, safeguards must be taken in order to benefit from extrapolating the information gathered. As discussed in Platt *et al.* [98], this is especially so when evaluating humourous laughter, as many aspects of humour revolve around the individual differences in personal preferences in the type of humour the individual finds funny and also personality traits. Researchers should hence continue to taking take care of formalizing the experiments and experimental environment as much as possible. Also, further research on the sense of humour are of particular interest [107].

Future research should also concentrate on eliciting laughter in natural situations and study the morphology, subjective experience, and social consequences. The studies should also aim towards the inclusion of multiple modalities in laughter coding. Here, the respiration, body movements, as well as the interplay of the features may be considered. Future works should also attempt to arrive at a classification of laughter that spans over more than one modality.

The performance achieved by HMM-based laughter synthesis is significantly higher than other laughter synthesis attempts, but remains far from actual human laughs. Future work need to be directed towards improving the naturalness of the methods, and a better synchronization between the audio and the visual channels, as shown in a first exploratory study by Çakmak *et al.* [14]. Other aspects include the desired level of control and diversity (of laughter categories and characteristics), and the desired reactivity in interactive systems. Reactive synthesis is characterized by the ability to change the input of the synthesis system on the fly and to see the effect on the output immediately, and a pilot study [27] focusing on laughter gave promising results. Further research on augmenting the diversity [31, 30] as well as reactivity [33] of such systems are ongoing.

Laughter-enabled interaction management remains a tricky task even though Learning from Demonstrations (LfD) proved to be a promising way to handle it. Yet it requires a lot of annotation work, new data collection campaigns for each task, integration of non-trivial contextual information. To address these issues, machine learning offers several perspectives that can be envisioned among which semi-supervised learning to train the models from little amounts of labeled data, transfer learning to capitalize knowledge acquired from task to task and automatic feature selection combined with non-parametric methods to scale with the dimensions of the input space. The Imitation Learning (IL) and Inverse Reinforcement Learning (IRL) algorithms developed within the ILHAIRE project could be easily extended to implement most of these methods. On another hand, even though IRL considers the laughter/silence decision process as a sequential decision process, it only models the long term influence of the artificial agent on the interaction. Another direction of research would therefore be to model the interaction as a multi-agent problem where the human users would also try to influence the agents behavior (like trying to generate contagion to the agent).

10 Conclusions

Laughter will be a key component of future socially believable behaving systems. This chapter reviewed recent work towards endowing computer systems with the ability to master this crucial social cue in humourous as well as more general conversational contexts. It covered a range of fundamental and technological topics including fundamental studies regarding the role, function and perception of laughter, as well as technological developments on recognizing/synthesizing social signals, and on modeling such interactions.

Given the currently increasing popularity of research on laughter in human communication and its use by computers, an exhaustive review within the scope and limits of a book chapter is already becoming quite difficult to achieve. Here, we focused essentially on the work done within the ILHAIRE project, together with pioneering work by others. We invite the readers to continue their exploration through recent papers by us, as well as other researchers active in those topics. Within the areas of affective conversational human-computer interaction, you are invited to look as a starting point in particular for Pelachaud, André, Camurri, Berthouze (participants in ILHAIRE), as well as Devillers, Campbell, Trouvain, Truong, and numerous other researchers we can not cover here. We hope this chapter will be the starting point for a fascinating journey, back in time through the evolutionary roots of laughter, amusement and social behaviour; and into the future of people engaging with the digital media.

11 Acknowledgements

References

1. André, E., Martin, J.C., Lingenfelser, F., Wagner, J.: Multimodal fusion in human-agent dialogue. In: M. Rojc, N. Campbell (eds.) Coverbal Synchrony in Human-Machine Interaction. CRC Press, Inc., Boca Raton, FL, USA (2013)
2. Bachorowski, J.A., Owren, M.J.: Not all laughs are alike: Voiced but not unvoiced laughter readily elicits positive affect. Psychological Science **12**(3), 252–257 (2001)
3. Bachorowski, J.A., Owren, M.J.: Sounds of emotion. Annals of the New York Academy of Sciences **1000**, 244–265 (2003)
4. Bachorowski, J.A., Smoski, M.J., Owen, M.J.: The acoustic features of human laughter. Journal of the Acoustical Society of America **110**(3,Pt1), 1581–1597 (2001)
5. Beller, G.: Analysis and generative model for expressivity. applied to speech and musical performance. Ph.D. thesis, Université Paris VI Pierre et Marie Curie (2009)
6. Bollepalli, B., Urbain, J., Raitio, T., Gustafson, J., Cakmak, H.: A comparative evaluation of vocoding techniques for hmm-based laughter synthesis. In: Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on, pp. 255–259 (2014). DOI 10.1109/ICASSP.2014.6853597
7. Bonin, F., Campbell, N., Vogel, C.: Laughter and Topic Changes: Temporal Distribution and Information Flow. In: CogInfoCom 2012 – 3rd IEEE International Conference on Cognitive Infocommunications, pp. 53–58. Kosice, Slovakia (2012)
8. Bryant, G.A., Aktipis, C.A.: The animal nature of spontaneous human laughter. Evolution and Human Behavior **35**(4), 327–335 (2014)
9. Burkhardt, F., Campbell, N.: Emotional speech synthesis. In: R. Calvo, S. D'Mello, J. Gratch, A. Kappas (eds.) The Oxford Handbook of Affective Computing. Oxford University Press, Oxford ; New York (2015)
10. Cagampan, B., Ng, H., Panuelos, K., Uy, K., Cu, J., Suarez, M.: An exploratory study on naturalistic laughter synthesis. In: Proceedings of the 4th International Workshop on Empathic Computing (IWEC'13). Beijing, China (2013)
11. Cai, R., Lu, L., Zhang, H.J., Cai, L.H.: Highlight sound effects detection in audio stream. In: Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on, vol. 3, pp. III–37–40 vol.3 (2003). DOI 10.1109/ICME.2003.1221242
12. Campbell, N., Kashioka, H., Ohara, R.: No laughing matter. In: Proceeding of INTERESPEECH, pp. 465–468. Lisbon, Portugal (2005)
13. Çakmak, H., Urbain, J., Dutoit, T.: The AV-LASYN database : A synchronous corpus of audio and 3d facial marker data for audio-visual laughter synthesis. In: Proc. of the 9th Int. Conf. on Language Resources and Evaluation (LREC'14) (2014)
14. Çakmak, H., Urbain, J., Dutoit, T.: Synchronization rules for hmm-based audio-visual laughter synthesis. In: Acoustics Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on (2015)
15. Çakmak, H., Urbain, J., Tilmanne, J., Dutoit, T.: Evaluation of HMM-based visual laughter synthesis. In: Acoustics Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on (2014)
16. Cosker, D., Edge, J.: Laughing, crying, sneezing and yawning: Automatic voice driven animation of non-speech articulations . In: Computer Animation and Social Agents (CASA) (2009)
17. Davila Ross, M., Allcock, B., Thomas, C., Bard, K.A.: Aping expressions? Chimpanzees produce distinct laugh types when responding to laughter of others. Emotion **11**(5), 1013–1020 (2011)
18. Davila Ross, M., Owen, M.J., Zimmermann, E.: Reconstructing the evolution of laughter in great apes and humans. Current Biology **19**(13), 1106–1111 (2009)
19. Devillers, L., Vidrascu, L.: Positive and negative emotional states behind the laughs in spontaneous spoken dialogs. In: Interdisciplinary Workshop on The Phonetics of Laughter, p. 37 (2007)

20. DiLorenzo, P., Zordan, V., Sanders, B.: Laughing out loud: control for modeling anatomically inspired laughter using audio. ACM Trans. Graph (2008)
21. Ding, Y.: Data-driven expressive animation model of speech and laughter for an embodied conversational agent. Ph.D. thesis, Télécom ParisTech (2014)
22. Ding, Y., Huang, J., Fourati, N., Artières, T., Pelachaud, C.: Upper body animation synthesis for a laughing character. In: Intelligent Virtual Agents, pp. 164–173. Springer (2014)
23. Ding, Y., Prepin, K., Huang, J., Pelachaud, C., Artières, T.: Laughter animation synthesis. In: Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems, pp. 773–780. International Foundation for Autonomous Agents and Multiagent Systems (2014)
24. Douglas-Cowie, E., Campbell, N., Cowie, R., Roach, P.: Emotional speech: Towards a new generation of databases. Speech Communication **40**(12), 33 – 60 (2003). DOI http://dx.doi.org/10.1016/S0167-6393(02)00070-5. URL http://www.sciencedirect.com/science/article/pii/S0167639302000705
25. Douglas-Cowie, E., Cowie, R., Sneddon, I., Cox, C., Lowry, O., McRorie, M., Martin, J.C., Devillers, L., Abrilian, S., Batliner, A., Amir, N., Karpouzis, K.: The humaine database: Addressing the collection and annotation of naturalistic and induced emotional data. In: A. Paiva, R. Prada, R. Picard (eds.) Affective Computing and Intelligent Interaction, *Lecture Notes in Computer Science*, vol. 4738, pp. 488–500. Springer Berlin Heidelberg (2007)
26. Dunbar, R.: Mind the gap: Or why humans are not just great apes. In: Proceedings of the British Academy. Volume 154. Joint British Academy/British Psychological Society Annual Lecture (2008)
27. dAlessandro, N., Tilmanne, J., Astrinaki, M., Hueber, T., Dall, R., Ravet, T., Moinet, A., Cakmak, H., Babacan, O., Barbulescu, A., Parfait, V., Huguenin, V., Kalayc, E.S., Hu, Q.: Reactive statistical mapping: Towards the sketching of performative control with data. In: Y. Rybarczyk, T. Cardoso, J. Rosas, L. Camarinha-Matos (eds.) Innovative and Creative Developments in Multimodal Interaction Systems, *IFIP Advances in Information and Communication Technology*, vol. 425, pp. 20–49. Springer Berlin Heidelberg (2014)
28. Ekman, P.: Sixteen enjoyable emotions. Emotion Researcher **18**(2), 6–7 (2003)
29. Ekman, P., Friesen, W.V., Hager, J.C.: Facial Action Coding System: A technique for the measurement of facial movement (2002)
30. El Haddad, K., Çakmak, H., Dupont, S., Dutoit, T.: Towards a speech synthesis system with controllable amusement levels. In: Proc. of 4th Interdisciplinary Workshop on Laughter and Other Non-Verbal Vocalisations in Speech. Enschede, The Netherlands (2015)
31. El Haddad, K., Dupont, S., d'Alessandro, N., Dutoit, T.: An hmm-based speech-smile synthesis system: An approach for amusement synthesis. In: Proc. of 3rd Intl Workshop on Emotion Representation, Analysis and Synthesis in Continuous Time and Space (EmoSPACE15). Ljubljana, Slovenia (2015)
32. El Haddad, K., Dupont, S., Urbain, J., Dutoit, T.: Speech-laughs: An HMM-based Approach for Amused Speech Synthesis. In: Internation Conference on Acoustics, Speech and Signal Processing (ICASSP 2015) (2015)
33. El Haddad, K., Moinet, A., Çakmak, H., Dupont, S., Dutoit, T.: Using mage for real time speech-laugh synthesis. In: Proc. of 4th Interdisciplinary Workshop on Laughter and Other Non-Verbal Vocalisations in Speech. Enschede, The Netherlands (2015)
34. Fredrickson, B.: The broaden-and-build theory of positive emotions. Philosophical Transactions of the Royal Society B: Biological Sciences **359**, 13671378 (2004)
35. Fukushima, S., Hashimoto, Y., Nozawa, T., Kajimoto, H.: Laugh enhancer using laugh track synchronized with the user's laugh motion. In: CHI '10 Extended Abstracts on Human Factors in Computing Systems, CHI EA '10, pp. 3613–3618. ACM, New York, NY, USA (2010). DOI 10.1145/1753846.1754027. URL http://doi.acm.org/10.1145/1753846.1754027
36. Glenn, P.J.: Laughter in Interaction. The Discourse Function of Laughter in Writing Tutorials. Cambridge University Press, Cambridge (2003)
37. Grammer, K.: Strangers meet: Laughter and nonverbal signs of interest in opposite-sex encounters. Journal of Nonverbal Behavior **14**(4), 209–236 (1990). DOI 10.1007/BF00989317. URL http://dx.doi.org/10.1007/BF00989317

38. Greengross, G., Miller, G.F.: Humor ability reveals intelligence, predicts mating success, and is higher in males. Intelligence **39**(4), 188–192 (2011)

39. Griffin, H., Aung, M., Romera-Paredes, B., McLoughlin, C., McKeown, G., Curran, W., Berthouze, N.: Perception and automatic recognition of laughter from whole-body motion: continuous and categorical perspectives. Affective Computing, IEEE Transactions on **PP**(99), 1–1 (2015). DOI 10.1109/TAFFC.2015.2390627

40. Griffin, H., Aung, M., Romera-Paredes, B., McLoughlin, C., McKeown, G., Curran, W., Bianchi-Berthouze, N.: Laughter type recognition from whole body motion. In: Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on, pp. 349–355 (2013). DOI 10.1109/ACII.2013.64

41. Hatfield, E., Cacioppo, J.T., Rapson, R.L.: Emotional Contagion. Cambridge University Press, New York (1994)

42. Hofmann, J.: Intense or malicious? the decoding of eyebrow-lowering frowning in laughter animations depends on the presentation mode. Frontiers in Psychology **5** (2014a)

43. Hofmann, J.: Smiling and laughter in positive emotions: Personality influences and expressive features. Ph.D. thesis, University of Zurich (2014b)

44. Hofmann, J., Platt, T., , Ruch, W., Proyer, R.T.: Individual differences in gelotophobia predict responses to joy and contempt. Sage Open **5**(2), 1–12 (2015)

45. Hofmann, J., Platt, T., Ruch, W.: More than amusement: Laughter and smiling in positive emotions (under review)

46. Hofmann, J., Platt, T., Ruch, W., Niewiadomski, R., Urbain, J.: The influence of a virtual companion on amusement when watching funny films. Motivation and Emotion pp. 1–14 (2015)

47. Hofmann, J., Ruch, W.: Schadenfreude laughter. Semiotika (Special Issue on Laughter) (2016)

48. Hofmann, J., Ruch, W., Platt, T.: The en-and decoding of schadenfreude laughter. sheer joy expressed by a duchenne laugh or emotional blend with a distinct morphological expression? In: Interdisciplinary Workshop on Laughter and other Non-Verbal Vocalisations in Speech Proceedings, pp. 26–27 (2012)

49. Hofmann, J., Stoffel, F., Weber, A., Platt, T.: The 16 enjoyable emotions induction task (16-eeit) - unpublished research instrument. Tech. rep., University of Zurich, Switzerland (2011)

50. Hudenko, W.J., Magenheimer, M.A.: Listeners prefer the laughs of children with autism to those of typically developing children. Autism p. 1362361311402856 (2011)

51. Ito, A., Wang, X., Suzuki, M., Makino, S.: Smile and laughter recognition using speech processing and face recognition from conversation video. In: Proceedings of the 2005 International Conference on Cyberworlds, CW '05, pp. 437–444. IEEE Computer Society, Washington, DC, USA (2005). DOI 10.1109/CW.2005.82. URL http://dx.doi.org/10.1109/CW.2005.82

52. Janin, A., Baron, D., Edwards, J., Ellis, D., Gelbart, D., Morgan, N., Peskin, B., Pfau, T., Shriberg, E., Stolcke, A., Wooters, C.: The icsi meeting corpus. In: Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on, vol. 1, pp. I–364–I–367 vol.1 (2003). DOI 10.1109/ICASSP.2003.1198793

53. Kayyal, M., Widen, S., Russell, J.: Context Is More Powerful Than We Think: Contextual Cues Override Facial Cues Even for Valence. Emotion (2015)

54. Kennedy, L., Ellis, D.: Laughter detection in meetings. In: NIST ICASSP 2004 Meeting Recognition Workshop, pp. 118–121. Montreal, Canada (2004)

55. Kipper, S., Todt, D.: Variation of sound parameters affects the evaluation of human laughter. Behaviour **138**(9), 1161–1178 (2001)

56. Kipper, S., Todt, D.: Dynamic-acoustic variation causes differences in evaluations of laughter. Perceptual and motor skills **96**(3), 799–809 (2003)

57. Kipper, S., Todt, D.: The role of rhythm and pitch in the evaluation of human laughter. Journal of Nonverbal Behavior **27**(4), 255–272 (2003)

58. Klein, E., Geist, M., PIOT, B., Pietquin, O.: Inverse Reinforcement Learning through Structured Classification. In: P. Bartlett, F.C.N. Pereira, C.J.C. Burges, L. Bottou, K.Q. Weinberger

(eds.) Advances in Neural Information Processing Systems 25, pp. 1016–1024 (2012). URL http://books.nips.cc/papers/files/nips25/NIPS2012_0491.pdf

59. Klein, E., PIOT, B., Geist, M., Pietquin, O.: A cascaded supervised learning approach to inverse reinforcement learning. In: H. Blockeel, K. Kersting, S. Nijssen, F. Zelezny (eds.) Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2013), *Lecture Notes in Computer Science*, vol. 8188, pp. 1–16. Springer, Prague (Czech Republic) (2013). URL http://www.ecmlpkdd2013.org/wp-content/uploads/2013/07/327.pdf

60. Knox, M.T., Mirghafori, N.: Automatic laughter detection using neural networks. In: IN-TERSPEECH 2007, 8th Annual Conference of the International Speech Communication Association, Antwerp, Belgium, August 27-31, 2007, pp. 2973–2976. ISCA (2007)

61. Kori, S.: Perceptual dimensions of laughter and their acoustic correlates. Proc. Intern. Confer. Phonetic Sciences Tallinn (4) pp. 255–258 (1989)

62. Lasarcyk, E., Trouvain, J.: Imitating conversational laughter with an articulatory speech synthesis. In: Proceedings of the Interdisciplinary Workshop on the Phonetics of Laughter, pp. 43–48. Saarbrücken, Germany (2007)

63. Lingenfelser, F., Wagner, J., André, E., McKeown, G., Curran, W.: An event driven fusion approach for enjoyment recognition in real-time. In: Proceedings of the ACM International Conference on Multimedia, MM '14, pp. 377–386. ACM, New York, NY, USA (2014). DOI 10.1145/2647868.2654924. URL http://doi.acm.org/10.1145/2647868.2654924

64. Lockerd, A., Mueller, F.M.: Lafcam: Leveraging affective feedback camcorder. In: CHI '02 Extended Abstracts on Human Factors in Computing Systems, CHI EA '02, pp. 574–575. ACM, New York, NY, USA (2002). DOI 10.1145/506443.506490. URL http://doi.acm.org/10.1145/506443.506490

65. Mancini, M., Ach, L., Bantegnie, E., Baur, T., Berthouze, N., Datta, D., Ding, Y., Dupont, S., Griffin, H., Lingenfelser, F., Niewiadomski, R., Pelachaud, C., Pietquin, O., Piot, B., Urbain, J., Volpe, G., Wagner, J.: Laugh when you're winning. In: Y. Rybarczyk, T. Cardoso, J. Rosas, L. Camarinha-Matos (eds.) Innovative and Creative Developments in Multimodal Interaction Systems, *IFIP Advances in Information and Communication Technology*, vol. 425, pp. 50–79. Springer Berlin Heidelberg (2014)

66. Mancini, M., Hofmann, J., Platt, T., Volpe, G., Varni, G., Glowinski, D., Ruch, W., Camurri, A.: Towards automated full body detection of laughter driven by human expert annotation. In: Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on, pp. 757–762. IEEE (2013)

67. Mancini, M., Varni, G., Glowinski, D., Volpe, G.: Computing and evaluating the body laughter index. In: A. Salah, J. Ruiz-del Solar, . Merili, P.Y. Oudeyer (eds.) Human Behavior Understanding, *Lecture Notes in Computer Science*, vol. 7559, pp. 90–98. Springer Berlin Heidelberg (2012)

68. Mancini, M., Varni, G., Niewiadomski, R., Volpe, G., Camurri, A.: 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, pp. 1855–1860. ACM, New York, NY, USA (2014). DOI 10.1145/2559206.2581205. URL http://doi.acm.org/10.1145/2559206.2581205

69. Matsusaka, T.: When does play panting occur during social play in wild chimpanzees? Primates; journal of primatology **45**(4), 221–229 (2004)

70. McKeown, G., Cowie, R., Curran, W., Ruch, W., Douglas-Cowie, E.: Ilhaire laughter database. In: Proceedings of the LREC Workshop on Corpora for Research on Emotion Sentiment and Social Signals (ES 2012). European Language Resources Association (ELRA), Istanbul (2012)

71. McKeown, G., Curran, W., Kane, D., Mccahon, R., Griffin, H.J., McLoughlin, C., Bianchi-Berthouze, N.: Human perception of laughter from context-free whole body motion dynamic stimuli. 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction **0**, 306–311 (2013). DOI http://doi.ieeecomputersociety.org/10.1109/ACII.2013.57

72. McKeown, G., Curran, W., McLoughlin, C., Griffin, H., Bianchi-Berthouze, N.: Laughter induction techniques suitable for generating motion capture data of laughter associated body

movements. In: Proceedings of the 2nd International Workshop on Emotion Representation, Analysis and Synthesis in Continuous Time and Space (EmoSPACE) In conjunction with the IEEE FG. Shanghai, China (2013)

73. McKeown, G., Sneddon, I., Curran, W.: Gender Differences in the Perceptions of Genuine and Simulated Laughter and Amused Facial Expressions. Emotion Review **7**(1), 30–38 (2015)

74. McKeown, G., Sneddon, I., Curran, W.: The underdetermined nature of laughter. In preparation (2015)

75. McKeown, G., Valstar, M., Cowie, R., Pantic, M., Schroder, M.: The semaine database: Annotated multimodal records of emotionally colored conversations between a person and a limited agent. Affective Computing, IEEE Transactions on **3**(1), 5–17 (2012). DOI 10.1109/T-AFFC.2011.20

76. McKeown, G.J.: The Analogical Peacock Hypothesis: The sexual selection of mind-reading and relational cognition in human communication. Review of General Psychology **17**(3), 267–287 (2013)

77. Melder, W.A., Truong, K.P., Uyl, M.D., Van Leeuwen, D.A., Neerincx, M.A., Loos, L.R., Plum, B.S.: Affective multimodal mirror: Sensing and eliciting laughter. In: Proceedings of the International Workshop on Human-centered Multimedia, HCM '07, pp. 31–40. ACM, New York, NY, USA (2007). DOI 10.1145/1290128.1290134. URL http://doi.acm.org/10.1145/1290128.1290134

78. Miller, G.F.: The Mating Mind. Vintage, London (2001)

79. Niewiadomski, R., Hofmann, J., Urbain, J., Platt, T., Wagner, J., Piot, B., Çakmak, H., Pammi, S., Baur, T., Dupont, S., Geist, M., Lingenfelser, F., McKeown, G., Pietquin, O., Ruch, W.: Laugh-aware virtual agent and its impact on user amusement. In: Proc. int. conf. on Autonomous agents and multi-agent systems, AAMAS (2013)

80. Niewiadomski, R., Mancini, M., Baur, T., Varni, G., Griffin, H., Aung, M.S.H.: Mmli: Multimodal multiperson corpus of laughter in interaction. In: A.A. Salah, H. Hung, O. Aran, H. Gunes (eds.) HBU, *Lecture Notes in Computer Science*, vol. 8212, pp. 184–195. Springer (2013)

81. Niewiadomski, R., Mancini, M., Ding, Y., Pelachaud, C., Volpe, G.: Rhythmic body movements of laughter. In: Proceedings of the 16th International Conference on Multimodal Interaction, pp. 299–306. ACM (2014)

82. Niewiadomski, R., Obaid, M., Bevacqua, E., Looser, J., Anh, L.Q., Pelachaud, C.: Cross-media agent platform. In: Proceedings of the 16th International Conference on 3D Web Technology, pp. 11–19. ACM (2011)

83. Niewiadomski, R., Pammi, S., Sharma, A., Hofmann, J., Platt, T., Cruz, R., Qu, B.: Visual laughter synthesis: Initial approaches. In: Interdisciplinary Workshop on Laughter and other Non-Verbal Vocalisations in Speech, Dublin, Ireland (2012)

84. Niewiadomski, R., Pelachaud, C.: Towards multimodal expression of laughter. In: Intelligent Virtual Agents, pp. 231–244. Springer (2012)

85. Niewiadomski, R., Pelachaud, C.: The effect of wrinkles, presentation mode, and intensity on the perception of facial actions and full-face expressions of laughter. ACM Transactions on Applied Perception (TAP) **12**(1), 2 (2015)

86. Niewiadomski, R., Urbain, J., Pelachaud, C., Dutoit, T.: Finding out the audio and visual features that influence the perception of laughter intensity and differ in inhalation and exhalation phases. In: Proceedings of the 4th International Workshop on Corpora for Research on Emotion, Sentiment and Social Signals, Satellite of LREC 2012. Istanbul, Turkey (2012)

87. O'Donnell Trujillo, N., Adams, K.: Heheh in conversation: Some coordinating accomplishments of laughter. Western Journal of Communication (includes Communication Reports) **47**(2), 175–191 (1983)

88. Oh, J., Wang, G.: Laughter modulation: from speech to speech-laugh. In: Proceedings of the 14th Annual Conference of the International Speech Communication Association (INTERSPEECH), pp. 754–755. Lyon, France (2013)

89. Oh, J., Wang, G.: Lolol: Laugh out loud on laptop. In: Proceedings of the 2013 International Conference on New Musical Instruments (NIME'13). Daejon, Korea (2013)

90. Owren, M., Bachorowski, J.A.: Reconsidering the evolution of nonlinguistic communication: The case of laughter. Journal of Nonverbal Behavior **27**(3), 183–200 (2003)
91. Pammi, S., Khemiri, H., Chollet, G.: Laughter detection using alisp-based n-gram models. In: Proceeding of the Interdisciplinary Workshop on Laughter and other Non-Verbal Vocalisations, pp. 16–17. Dublin, Ireland (2012)
92. Pecune, F., Biancardi, B., Ding, Y., Pelachaud, C., Mancini, M., Varni, G., Camurri, A., Volpe, G.: Lol - laugh out loud. In: Proc. of AAAI 2015 (2015)
93. Pelachaud, C.: Interacting with socio-emotional agents. Procedia Computer Science **39**, 4–7 (2014)
94. Petridis, S., Martinez, B., Pantic, M.: The mahnob laughter database. Image Vision Comput. **31**(2), 186–202 (2013). DOI 10.1016/j.imavis.2012.08.014. URL http://dx.doi.org/10.1016/j.imavis.2012.08.014
95. Petridis, S., Pantic, M.: Fusion of audio and visual cues for laughter detection. In: International Conference on Content-Based Image and Video Retrieval, CIVR 2008, pp. 329–337. ACM, New York (2008). URL http://doc.utwente.nl/62669/
96. Petridis, S., Pantic, M.: Audiovisual discrimination between speech and laughter: Why and when visual information might help. Multimedia, IEEE Transactions on **13**(2), 216–234 (2011). DOI 10.1109/TMM.2010.2101586
97. Piot, B., Pietquin, O., Geist, M.: Predicting when to laugh with structured classification. In: Annual Conference of the International Speech Communication Association (InterSpeech) (2014)
98. Platt, T., Hofmann, J., Ruch, W., Niewiadomski, R., Urbain, J.: Experimental standards in research on ai and humor when considering psychology. In: Proc. of Fall Symposium on Artificial Intelligence of Humor (2012)
99. Platt, T., Hofmann, J., Ruch, W., Proyer, R.T.: Duchenne display responses towards sixteen enjoyable emotions: Individual differences between no and fear of being laughed at. Motivation and Emotion **37**(4), 776–786 (2013)
100. Preuschoft, S., van Hooff, J.A.R.A.M.: The Social Function of "Smile" and "Laughter": Variations Across Primate Species and Societies. pp. 171–189. Lawrence Erlbaum Associates, Mahweh, New Jersey (1997)
101. Qu, B., Pammi, S., Niewiadomski, R., Chollet, G.: Estimation of faps and intensities of aus based on real-time face tracking. In: Proceedings of the 3rd Symposium on Facial Analysis and Animation, FAA '12, pp. 13:1–13:1. ACM, New York, NY, USA (2012). DOI 10.1145/2491599.2491612. URL http://doi.acm.org/10.1145/2491599.2491612
102. Reuderink, B.: Fusion for audio-visual laughter detection (2007). URL http://essay.utwente.nl/714/
103. Riek, L., Rabinowitch, T., Chakrabarti, B., Robinson, P.: Empathizing with robots: Fellow feeling along the anthropomorphic spectrum. In: Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on, pp. 1–6 (2009). DOI 10.1109/ACII.2009.5349423
104. Rienks, R.: Meetings in smart environments. implications of progressing technology. Phd thesis, University of Twente (2007). ISBN: 978-90-365-2533-6, Numberofpages: 201
105. Rothbart, M.K.: Laughter in young children. Psychological Bulletin **80**(3), 247–256 (1973)
106. Ruch, W.: The handbook of emotions, chap. Exhilaration and humor, pp. 605–616. Guilford Press, New York (1993)
107. Ruch, W.: Towards a new structural model of the sense of humor: Preliminary findings. In: Proc. of Fall Symposium on Artificial Intelligence of Humor (2012)
108. Ruch, W., Ekman, P.: Emotion, qualia and consciousness, chap. The expressive pattern of laughter, pp. 426–443. World Scientic Publishers, Tokyo (2001)
109. Ruch, W., Hofmann, J.: A temperament approach to humor. Humor and health promotion pp. 79–113 (2012)
110. Ruch, W., Hofmann, J., Platt, T.: Investigating facial features of four types of laughter in historic illustrations. The European Journal of Humour Research **1**(1), 99–118 (2013)
111. Ruch, W., Hofmann, J., Platt, T.: Individual differences in gelotophobia and responses to laughter-eliciting emotions. Personality and Individual Differences **72**, 117–121 (2015)

112. Ruch, W., Hofmann, J., Platt, T., Proyer, R.: The state-of-the art in gelotophobia research: A review and some theoretical extensions. Humor: International Journal of Humor Research **27**(1), 23–45 (2013)

113. Ruch, W.F., Platt, T., Hofmann, J., Niewiadomski, R., Urbain, J., Mancini, M., Dupont, S.: Gelotophobia and the challenges of implementing laughter into virtual agents interactions. Frontiers in human neuroscience **8** (2014)

114. Sathya, A.T., Sudheer, K., Yegnanarayana, B.: Synthesis of laughter by modifying excitation characteristics. The Journal of the Acoustical Society of America **133**, 3072–3082 (2013)

115. Schuller, B., Steidl, S., Batliner, A., Vinciarelli, A., Scherer, K.R., Ringeval, F., Chetouani, M., Weninger, F., Eyben, F., Marchi, E., Mortillaro, M., Salamin, H., Polychroniou, A., Valente, F., Kim, S.: The interspeech 2013 computational paralinguistics challenge: social signals, conflict, emotion, autism. In: INTERSPEECH, pp. 148–152. ISCA (2013)

116. Sestito, M., Umiltà, M.A., De Paola, G., Fortunati, R., Raballo, A., Leuci, E., Maffei, S., Tonna, M., Amore, M., Maggini, C., et al.: Facial reactions in response to dynamic emotional stimuli in different modalities in patients suffering from schizophrenia: a behavioral and emg study. Frontiers in human neuroscience **7** (2013)

117. Shahid, S., Krahmer, E., Swerts, M., Melder, W., Neerincx, M.: You make me happy: Using an adaptive affective interface to investigate the effect of social presence on positive emotion induction. In: Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on, pp. 1–6 (2009). DOI 10.1109/ACII.2009.5349355

118. Sneddon, I., McRorie, M., McKeown, G., Hanratty, J.: The belfast induced natural emotion database. Affective Computing, IEEE Transactions on **3**(1), 32–41 (2012). DOI 10.1109/T-AFFC.2011.26

119. Sundaram, S., Narayanan, S.: Automatic acoustic synthesis of human-like laughter. Journal of the Acoustical Society of America **121**(1), 527–535 (2007)

120. Szameitat, D.P., Darwin, C.J., Wildgruber, D., Alter, K., Szameitat, A.J.: Acoustic correlates of emotional dimensions in laughter: Arousal, dominance, and valence. Cognition and Emotion **25**(4), 599–611 (2011)

121. Tanaka, H., Campbell, N.: Classification of social laughter in natural conversational speech. Computer Speech & Language **28**(1), 314–325 (2014)

122. Tokuda, K., Yoshimura, T., Masuko, T., Kobayashi, T., Kitamura, T.: Speech parameter generation algorithms for hmm-based speech synthesis. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 3, pp. 1315–1318. IEEE (2000)

123. Truong, K.P., van Leeuwen, D.A.: Automatic discrimination between laughter and speech. Speech Communication **49**(2), 144 – 158 (2007). DOI http://dx.doi.org/10.1016/j.specom.2007.01.001. URL http://www.sciencedirect.com/science/article/pii/S0167639307000027

124. Urbain, J.: Acoustic laughter processing. Ph.D. thesis, University of Mons (2014)

125. Urbain, J., Cakmak, H., Dutoit, T.: Development of hmm-based acoustic laughter synthesis. In: Interdisciplinary Workshop on Laughter and other Non-Verbal Vocalisations in Speech, Dublin, Ireland, pp. 26–27 (2012)

126. Urbain, J., Çakmak, H., Charlier, A., Denti, M., Dutoit, T., Dupont, S.: Arousal-driven synthesis of laughter. IEEE Journal of Selected Topics in Signal Processing **8**, 273–284 (2014). DOI 10.1109/JSTSP.2014.2309435

127. Urbain, J., Çakmak, H., Dutoit, T.: Automatic phonetic transcription of laughter and its application to laughter synthesis. In: Proceedings of the 5th biannual Humaine Association Conference on Affective Computing and Intellignet Interaction (ACII), pp. 153–158. Geneva, Switzerland (2013)

128. Urbain, J., Çakmak, H., Dutoit, T.: Evaluation of hmm-based laughter synthesis. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 7835–7839. Vancouver, Canada (2013)

129. Urbain, J., Dutoit, T.: Measuring instantaneous laughter intensity from acoustic features. In: Proceeding of the Interdisciplinary Workshop on Laughter and other Non-Verbal Vocalisations, pp. 18–19. Dublin, Ireland (2012)

130. Urbain, J., Niewiadomski, R., Bevacqua, E., Dutoit, T., Moinet, A., Pelachaud, C., Picart, B., Tilmanne, J., Wagner, J.: Avlaughtercycle. Journal on Multimodal User Interfaces **4**(1), 47–58 (2010). DOI 10.1007/s12193-010-0053-1. URL http://dx.doi.org/10.1007/s12193-010-0053-1

131. Urbain, J., Niewiadomski, R., Hofmann, J., Bantegnie, E., Baur, T., Berthouze, N., Cakmak, H., Cruz, R., Dupont, S., Geist, M., Griffin, H., Lingenfelser, F., Mancini, M., Miranda, M., McKeown, G., Pammi, S., Pietquin, O., Piot, B., Platt, T., Ruch, W., adn G. Volpe, A.S., Wagner, J.: Laugh machine. In: Proc. of eNTERFACE12. The 8th International Summer Workshop on Multimodal Interfaces (2012)

132. Urbain, J., Niewiadomski, R., Mancini, M., Griffin, H., Çakmak, H., Ach, L., Volpe, G.: Multimodal analysis of laughter for an interactive system. In: Proceedings of the INTETAIN 2013 (2013)

133. Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D'Errico, F., Schroeder, M.: Bridging the gap between social animal and unsocial machine: A survey of social signal processing. Affective Computing, IEEE Transactions on **3**(1), 69–87 (2012). DOI 10.1109/T-AFFC.2011.27

134. Wagner, J., Lingenfelser, F., André, E.: Using phonetic patterns for detecting social cues in natural conversations. In: F. Bimbot, C. Cerisara, C. Fougeron, G. Gravier, L. Lamel, F. Pellegrino, P. Perrier (eds.) INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association, Lyon, France, August 25-29, 2013, pp. 168–172. ISCA (2013)

135. Wagner, J., Lingenfelser, F., Baur, T., Damian, I., Kistler, F., André, E.: The social signal interpretation (ssi) framework: Multimodal signal processing and recognition in real-time. In: Proceedings of the 21st ACM International Conference on Multimedia, MM '13, pp. 831–834. ACM, New York, NY, USA (2013)

136. Yoshimura, T., Tokuda, K., Masuko, T., Kobayashi, T., Kitamura, T.: Simultaneous modeling of spectrum, pitch and duration in hmm-based speech synthesis. In: Proceedings of Eurospeech. Budapest, Hungary (1999)