African Journal of Social Sciences and Humanities Research

Vol.1, No.1, pp.124-154, 2018

www.abjournals.org

WATCH AND LISTEN TO MY FACE: THE POTENTIAL FOR FACIAL COMMUNICATION IN HUMAN-AGENT INTERACTION

Abdulmalik Yusuf Ofemile[1], Innocent Ejimofor Agu[2] and Evangelista Chimebere Agu[2]

[1]FCT College of Education, Zuba, Abuja, Nigeria. [email: abdulmalikofemile@yahoo.co.uk]
[2]Federal University Wukari, Nigeria

ABSTRACT: *Most research around listenership in Human-Agent Interaction has focused on assessing listener feedback using participant utterances during interaction, narratives after interaction or posed facial actions. However, little attention is paid to spontaneous facial actions displayed when interacting with software agents in instruction-giving contexts. This paper reports a study aimed at developing a better understanding of the nature and communicative potentials of spontaneous facial actions displayed during these interactions. Forty-eight participants were tasked with assembling two Lego models using verbal instructions from a computer interface. The interface used three voices of which two were synthesised and one provided by a voice actor. A 24-hour-long multimodal corpus was built and marked instances analysed from these interactions. The results suggest that it is possible for humans to show their perceptions of agent identity through their facial actions as positive, negative or indifferent during interaction. Furthermore, there is a potential for formulating a theoretical basis for researching interaction in similar contexts. Findings suggest that agents enhanced emotive functionality may enhance Human-Agent interaction in emerging contexts, but this requires further research.*

KEYWORDS: Human-Agent Interaction; Facial Actions; Listenership; Software agents

INTRODUCTION

Most research around listener perception of speaker identity has focused on interaction between people. However, there is now a marked increase in everyday interaction with software agents due to the pervasive nature of computing (Jennings et al. 2014). This has expanded interaction to broader discourse contexts, for example, satnav systems giving instructions to drivers, automated phone instructions, and intelligent personal assistants managing itineraries.

Facial actions are facial movements in terms of component actions used in nonverbal communication (Ekman,1997) in Human-Human Interaction(HHI) however, the communicative role of listener facial actions in Human-Agent Interaction (HAI) requires understanding. For the purposes of this paper, an agent refers to an intelligent software run within a simulated body and environment where a program takes in commands and returns appropriate percepts (Poole & Mackworth, 2010). The agent in this paper is simulated and does not possess agency as such but can create a believable agent-like interaction.

With regards to assessing user experience, researchers in HAI use post-task questionnaires (Clark et al., 2014; Nass & Lee 2001; Nass et al., 1999), participants' nested narratives and semi-structured interviews (Luger & Sellen 2016; Clark 2016; Williams et al., 2009) and facial actions as simulated feedback (Jabbi et al., 2007) while spontaneous facial actions get little or no attention. Thus, this study uses spontaneous listener facial actions displayed during task to assess user experience in HAI because research indicates that, people tend to forget exact details of interactions after they have taken place and often re-construct events that may have inaccuracies (Bach & Goncalves 2004).

Although, human emotions have been used in HAI to determine facial recognition in robots and intelligent agents, measure user frustration when interacting with agents and monitor emotion regulation with computers (Miles, Nind, & Macrae, 2009; Klein et al., 2002; Cowie, et al., 2001) however, the agent’s ability to use a human co-interlocutor’s facial actions communicatively at human levels still requires understanding because with time agents may need to respond to human behaviour at that level to make interaction effective.

Furthermore, this study uses spontaneous listener facial actions to assess listener attitudes towards agent-instructors as perceptions of agent identity during interaction. The face has super-colossal ‘sending capacity’. In 1969 Paul Ekman and Wallace Friesen developed the concept of sending capacity as a means of evaluating the information sending capacity of any part of the body using three factors: the average transmission time of any message from that part, the number of discernible stimulus patterns which can be emitted from it, and its visibility.

This ties into research suggesting that nonverbal cues including facial actions are important when communicating feelings and attitudes and measuring bias (Meadors & Murray 2014). Mehrabian, (2007), citing Mehrabian and Ferris (1967) posits that nonverbal cues are used for making inferences of positive, negative or neutral attitudes towards others in interaction and validating behaviour towards specific communication as positive or negative evaluations, or as submissive or dominant.

RELATED LITERATURE

Agent Identity in HAI

One of the interpersonal functions of communication is the assertion of identity because this portrays us as individuals and members of a group (Pearson et al., 1999). For the purposes of this paper, identity is defined as a property of the individual that emerges from social interaction and can be regarded as residing in the mind or in concrete social behaviour that can be anchored to the individual or to group (De Fina, 2006: 365).

De Fina’s (2006) definition agrees with Goffman’s (2002) position that identities are created performatively, because in ordinary everyday situations, people behave in ways that tell others who they are, what they do and what their life expectations are ((Edgley, 2016;Taylor 2009; Hollway 2009) thus, a teacher is distinguishable from a student.

In addition, it suggests that identity is owned by individuals which ties in with the view that, personal identity is a person’s own idea or view of who he/she is (Taylor 2009) but, in HAI, the agent’s personal identity emerges from a process that begins with the designer’s proposed intended or real affordances (Ofemile 2015; Hollway 2009). Intended affordances are the uses or qualities that a designer proposes and introjects through programming into an agent to furnish its self-identity.

The personal identity of the agent used in this study is linked to its voice and projected as its social identity through its role (as instructor) to be received by others i.e. instructees (Marine-Roig, 2015) during interaction. Thus, personal identities are not wholly separate from social identities because, the agent's identity of being an instructor as separately suggested by Taylor (2009), Hollway (2009) and Zimmerman (1998) is both a social identity and an important part of who the agent is. This implies that, while being a simulated instructor makes it different from embodied or tactile agents but its role as an instructor is discursive because it builds interaction just as some embodied agents would.

Agents may also have a relational identity conferred or given to them by co-interactants as a label. Relational identities contain the differences that say "it is them not us" (Taylor 2009: 178) and are ascribed by others through the process of "othering" (Chauhan & Forster 2014). Othering is a process of engaging with others we perceive as mildly or radically different from ourselves. It also involves a manner of space purification that serves to mark and name others as different from us (Tope et al., 2014; Canales 2000).

Research indicates that othering in human society is achieved through representational absence, representation of difference and representation of threat (Chauhan & Forster, 2014). However, pervasive computing has ensured that agents are ubiquitously represented in our every day life for example, in smartphones, watches, and advent of software in standalone speakers (Amazon Echo, Google Home) and their perceived threats are being better understood. However, the representation of difference is still relevant as this concerns recognizing other objects or personalities as distinct from oneself by objectifying them or reducing them to a preconceived image that is often negative (Chauhan & Forster 2014).

It is hoped that agent projection of identity in HAI may elicit facial actions in the same way as in HHI. Furthermore, such facial action data may allow us to gain a real-time insight into the attitudes and identities created by the users of an agent, without having to rely solely on post-interaction data.

Spontaneous Facial Actions in HAI

It is established that facial actions are either spontaneous or simulated. Spontaneous facial actions are generated by biologically given processes that operate automatically, eliciting facial muscle reactions quickly and independent of conscious cognitive processes (Dimberg et al. 2000; Ekman 1992a, 1992b) for example, faces made when a person sees a sudden flash of lightning. In contrast, simulated facial actions are approximations of facial behaviours that occur when people are making little attempt to manage facial appearance (Ekman & O'Sullivan, 2006) for example, faces deliberately made for a photoshoot.

Furthermore, spontaneous facial actions align with other emotional behaviours -such as voice, gesture or posture- and are more reflexive and smoother with fewer phases than deliberate behaviour (Hess & Kleck 1997) showing fluency in communication. Conversely, simulated emotions more assymetric because their timing of muscular contraction is more irregular with missing components in the frequency of physiological movements in different parts of the face thereby causing dysfluencies in these expressions (Ekman, 2013).

Spontaneous facial actions are also mediated by neurological pathways that are different from those mediating simulated facial actions. Ofemile et.al, (2016) citing Weiss et al. (1987) posits that changes in body temperature, fluid movements and oestrogen injection into the body system observed when spontaneous emotions are expressed, are never realised when emotions are simulated.

Facial actions can be different due to the ontogeny or development of each individual from conception. However, some are universal because of their presence in all human populations as we are hardwired to produce and recognise facial actions. The universal ones are basic facial actions that evolved within man to deal with fundamental life tasks, such as achievement or failure. Examples include anger, disgust, sadness, fear, happiness, surprise, and neutral expressions (Ekman, 2016).

Non-basic emotions are made up of a combination of basic emotions because facial muscles are sufficiently complex to display a blend of emotions (Ekman 1997) such as smug and scorn. The blending process involves one emotion running into another to produce a new one (Ekman & Friesen, 2003) in a manner similar to articulation of diphthongs where one sound runs into another to sound as one.

Emotional attitudes are more sustained than basic emotions and may involve more than one emotion being elicited during interaction (Ekman, 1997). Listeners have attitudinal reactions like frowns that may indicate their willingness to accept or reject a message nonverbally (Allwood 1993) in ways that are different from emotions yet, can be used to indicate personal preferences, affective disposition and intrapersonal stances (Scherer, 2005).

Ekman (1997) and Miller (2014) separately hold that people can control facial actions because social pressures sometimes dictate such control and the very behaviour expected in each context is concealed thus, people tend to deliberately give off emotions that do not reflect their true feelings. However, a close examination of the process reveals the true feelings usually leaked as microexpression or “hot spots” (Miller, 2014). Microexpressions flash on and off the face in less than one-quarter of a second – so quickly that they are usually missed, however they constitute the greatest source of information leakage from the human face (Ekman, 2007).

Ekman (2004) suggests that humans acquire display rules early in life that are enlarged through learning and increased interaction. However, people in different circumstances depend on different display rules usually assessed through norming. Norming is an analytical process that establishes the baseline behaviour of a person in a neutral nonconfrontational environment (Miller, 2014). Baseline behaviour provides a basis to compare and contrast that person’s behaviour patterns during interaction because, display rules govern facial actions on a habitual basis and it becomes noticeable when violated. When the information provided by the face contradicts that provided by other behavioural modalities, researchers tend to look for the regularity of these discrepancies and how the subject mitigates them. Hence there is the need to employ a measurable and repeatable framework to assess facial actions such as used by Ekman and Friesen in 1977 and 2004.

METHODOLOGY

This section outlines the research methodology that attempts to answer the following questions:

RQ1: Which listener facial actions are observed during interaction and do these indicate their attitudes towards agents in interaction?

RQ2: Are there differences in listener perceptions of agents (Positive, Neutral, and Negative)?

Agent Design

An instruction-based task was used in this study, so that the research would somewhat reflect an existing context in which humans and agents interact. The study context requires that users follow verbal instructions provided by the agents to complete two Lego model assembly tasks.

A simulated agent was created on a computer interface instead of a real agent as simulated agents can provide users with experiences similar to those of real agents (Clark et al 2014). The interface (Figure 1 below) issued forty-seven assembly instructions. These instructions are contained within Hypertext Markup Language (HTML) files connected to a database of individual assembly instructions stored as .wav files. Together these make up the agent interfaces that users interact with.

Figure 1: Simulated Agent Interface

The agent interface was housed within a MacBook Pro 10.1 and provided instructions to participants verbally from a pool of three voices. The interface (Figure 1) above displays the model that participants are assembling (Nex or Aquagon), task steps covered, and repeat and next instruction buttons. The agent cannot go to a previous instruction but can repeat a current instruction and record task timings and number of repetitions requested by participants.

The agent uses one of three voices – two synthesised and one human voice. The synthesised voices were Cepstral Lawrence[1] (CL) and CereProc Giles[2] (CP).The human recorded voice (HR) provided by a professional voice actor hired from http://voicebunny.com. Each voice was male, aged between 40-55 years, and used a Received Pronuniciation (RP) accent.

Participants

48 participants that speak English as a first language were recruited through self-selection via email responses to participate in the study and were remunerated with a £10 Amazon voucher.

[1] https://www.cepstral.com
[2] https://www.cereproc.com

21 participants were male (43.8%) and 27 were female (56.2%), with their ages ranging from 18 – 32 years old. As participants engaged in two tasks from the pool of three agent voices, three main groups were created based on voice pairings: CP + CL; CP + HR; CL + HR. Participants were randomly assigned to these groupings using a free online list randomiser[3].

Experiment and Task Procedure

Participants took part in two tasks within their individual sessions. Prior to these tasks, they filled out brief demographics and informed consent was obtained from all participants using the consent form, followed by being briefed on the upcoming session. Participants were informed that each task would cease upon a fifteen-minute time limit expiring or completing the model assembly before this limit. Participants were also informed that they could repeat as many instructions as was necessary. Following the briefing, participants undertook the tasks.

Data Collection and Analysis

Repetition of instructions is used as a lens to understand participant's nonverbal feedback that may indicate attitudes towards instructors and assembly strategies during interaction because repetition is tied to task performance (Kim & Tracy-Ventura, 2012) and listenership (Tsuchiya, 2013). Following Clark et al. (2016) occurrences when participants asked the agent to repeat instructions were systematically selected as the sampling frame for the study because there is no guarantee that participants in assembly tasks will complete all the steps.

Each interaction is recorded from two angles shown (Figure 2) below. A Panasonic HDC-SD900 captured the eye level shot of the participant and a Canon Legria HFR306 recorded interactions shots from the side to capture both the nuances of interaction with the interface and the model assembly. Although the cameras could record in high definition, the smaller .mp4 format was used to save storage space without compromising on quality.

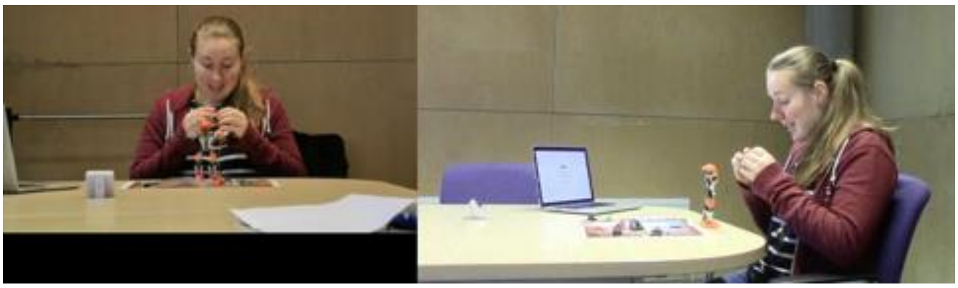


Figure 2: Front and side level camera shots of a participant assembling a Lego model

Interactions recorded were used to create twenty-four hours of multimodal corpora. The corpus was cleaned, segmented and tagged as follows. QuickTime Player [4] was used to edit and align videos recorded from two angles then saved in mp4 and .mov formats. This was uploaded as media unto CLAN[5] (Computerized Language ANalysis – software designed specifically to analyse data transcribed in the CHAT format) which was used to earmark interactions as bullets

[3] https://www.random.org/lists
[4] https://www.apple.com/quicktime/player
[5] http://childes.psy.cmu.edu/clan/

of repeated instructions. The CHAT files were exported to ELAN[6] -EUDICO Linguistic Annotator 4.9.0 (Lausberg & Sloetjes 2009) for annotation.

Data analysis was conducted through systematic codification to ensure thick data analysis (Dörnyei 2007) and reduce distraction. Earmarked interactions were watched severally in mute and sound modes then frame-by-frame to identify the most suitable presentation of participants' facial actions. Research indicates that specific facial muscles are associated with specific emotions thus, facial actions were segmented and annotated using the Facial Action Coding System (FACAID) by Ekman and Friesen in1977 (Ekman & Friesen, 2003) outlined in Table 1 below.

Linguistic annotation consisted of four steps. First, the neuro-biological processes or muscles generating the facial actions were described and codified as Action Units (AU). Following this, the neurological processes surrounding the facial action are described. The next stage classifies the facial action into family and sibling buckets based on FACAID and related disciplines. This level of granularity enables similar facial actions to be grouped together and the differences between and within groups highlighted. The last step locates the communicative and task function of facial actions as done with multimodal repertoires in Atkinson and Bradley, (2017). Table 1 provides the annotation guide and outlines facial actions observed, the AU code, muscles responsible and level of intensity.

Table 1: Facial Action Annotation Scheme (Adapted from Ekman, 2007, 2002; Ekman, Friesen & Ellsworth, 2013)

Annotation Group	Facial Action	Action Unit Number/Code	Muscular Basis
	Neutral Face	AU0	
	Inner Brow raiser	AU1; AU1B	frontalis (pars medialis)
	Outer Brow Raiser	AU2	frontalis (pars lateralis)
	Brow Lowerer	AU4; AU4B	depressor glabellae, depressor supercilii, corrugator supercilii
	Upper Lid Raiser	AU5; AU5B	levator palpebrae superioris, superior tarsal muscle
	Cheek Raiser	AU6; AU6B	orbicularis oculi (pars orbitalis)
	Lid Tightener	AU7; AU7B	orbicularis oculi (pars palpebralis)
	Pulls or tighten lips towards each other	AU8; AU8B	the orbicularis oris muscle
	Nose wrinkler	AU9; AU9B	levator labii superioris alaeque nasi
	Lip corner puller	AU12; AU12B; R12A	zygomaticus major
	Dimpler	AU14; R14A	Buccinator

[6] https://tla.mpi.nl/tools/tla-tools/elan/

	Lip corner depressor	AU15; AU15B	depressor anguli oris (also known as triangularis)
	Lower lip depressor	AU16; AU16B	depressor labii inferioris
	Chin raiser	AU17	mentalis muscle
	Lip Stretcher	AU20	risorius w/ platysma
	Neck Tightener	AU21	platysma
	Lip Funneler	AU22	orbicularis oris
	Lip Pressor	AU24	orbicularis oris
	Lips part	AU25	depressor labii inferioris muscles
	Jaw drop	AU26	masseter; relaxed temporalis and internal pterygoid
	Mouth Stretcher	AU27B	pterygoids, digastric muscle
	Dimpler	R14A	Buccinator
	Lip Bite	AU 32	
	Swallowed/compressed Lips	AU8B	orbicularis oris
	Glabella Lowerer	AU41	Separate Strand of AU 4:depressor glabellae (Procerus) muscle
	Inner Eyebrow Lowerer	AU42	Separate Strand of AU 4: depressor supercilii
	Eyebrow gatherer narrows eyes	AU44; AU44B	corrugator Supercilii (Darwin's muscle of difficulty)
	Wink	AU46	orbicularis oculi
	Head downward movement	AU54	
Intensity of Facial Action	Trace Slight Pronounced or maximum	A B C	
Eye/Ocular motor movement	Saccade-Left Saccade-Right	AU61 AU62	Lateral rectus and Medial rectus muscles; AU46 orbicularis oculi muscle;

Facial action units are expressed with some level of intensity and Ekman (1997) suggested five levels, given as: A Trace; B Slight; C Pronounced; D Severe or extreme; E maximum. However, Ekman's (1997) classification does not specify the differences between A and B or D and E. To avoid confusion, this study adopts only three levels of intensity given as: A- Trace in micro-expressions; B- slight; C- pronounced or maximum. C is the default and AUs without any alphabet are assumed to have maximum intensity.

RESULTS

The results suggest that participants expressed emotions and attitudes, they also projected moods and could facially display task coping strategies during interaction. In answering RQ1, Table 2 is used to classify listener facial actions displayed during interaction.

Table 2: Listener Facial Actions Displayed during Assembly Task

S.NO	Facial Action Family	ILLUSTRATION	
		Sibling Bucket	**Within Sibling Buckets**
1	Basic Facial Actions	Neutral Face	Neutral, Neutral Face down; Neutral hard; Neutral concentration on task; Neutral concentration on instruction
		Surprise	Surprise, slight surprise
		Smile	Felt smile; Nervous smile; Tight-lipped smile; GW Bush Grin;
		Fear	Controlled Fear
		Disgust	Slight Disgust
2	Non-Basic and Blends	Smug	Smug
		Angry-disgust	Angry-Disgust
3	Emotional Attitudes	Frown	Frown
4	Moods	Aha! Moment	Aha! Moment
		Workman Face	Work Face
5	Tense Mouth and Lip Action	Compressed or Swallowed lips	Swallowed lips
		Biting Lips	Biting Lips
		Pouty face	Pouty face
6	Oculomotor Movement or Eye Movement	Saccade	Saccade
7	Micro-expression	Basic Hotspots	Micro-anger; Micro-smile; Micro-tight-lipped smile; micro-disgust; micro-sadness

Basic Facial Actions

The basic expressions include neutral face, surprise, smile, fear, and disgust.

Neutral Faces

Neutral faces (AU0) are often expressionless with all the muscles relaxed (Ekman & Friesen 2003). However, the results suggest that listeners' faces are never completely expressionless because of their involvement in a task. There is a distinction between a completely expressionless face or soft expression and a hard one that comes with a trace of expressions. In the case of the soft expression, all the muscles are relaxed (Ekman, Friesen & Ellsworth, 2013) without any emotion as shown (Figure 2) below

Figure 2: Neutral Face_Soft expression

Sometimes expressionless listeners lean forward - indicated by the broken green arrow (Figure 3) below given by AU54 (Ekman,2002) with the eyes focused on some point in space and the hands inactive.

Figure 3: Neutral Face_ Head bending forward

These actions suggest that participants may be paying close attention for comprehension as the instruction is repeated.

Sometimes, the face is predominantly expressionless yet showing traces of other emotions. For example, the participants' lips are held tight (Figure 4) below. Ekman (2007) suggests that tight-lips may indicate a trace of anger however, this may also a mood or way of concentrating on the task.

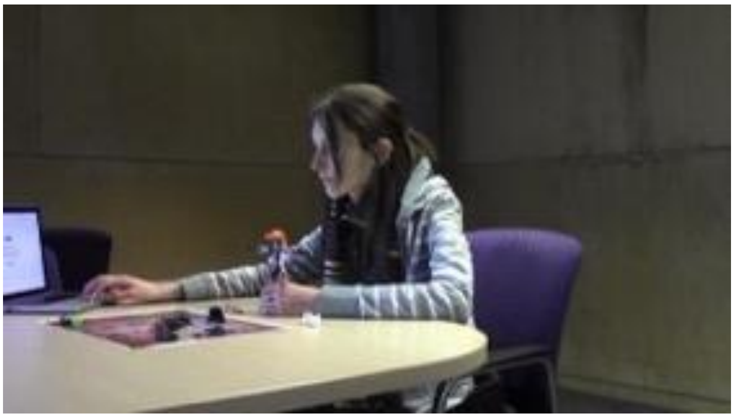


Figure 4: Neutral Face _Hard Expression

Another form of the neutral face is intense neutral face shown (Figure 5) below. The indicator in this facial action is the partially furrowed forehead drawn together by AU44 (Ekman 2007).

Figure 5: Neutral Face_ Intense

The participants (Figure 5) above have visible furrows in the forehead that may also suggest they are keenly concentrating, or possibly confused or determined (Ekman 2002). Another variation of the neutral face co-occurs with the listener's hands engaged in tasks suggesting that the participant is concentrated on the task while listening as shown (Figure 6) below

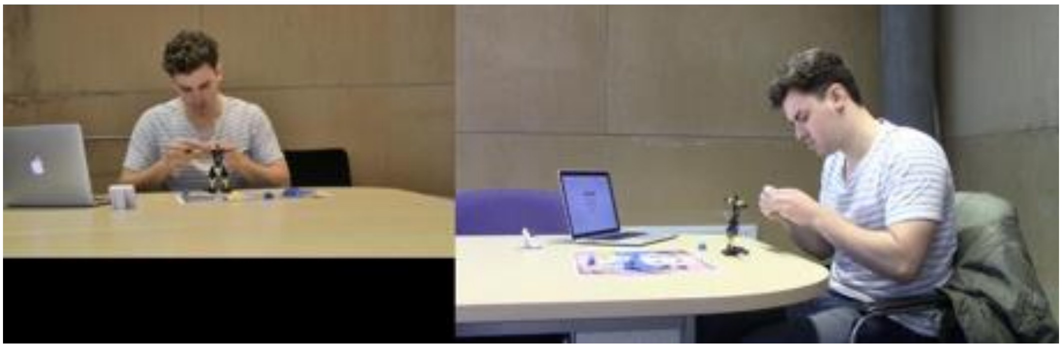


Figure 6: Neutral Face_Concentrating on Task

Surprise

Surprise is often triggered by something unexpected such as a sudden loud noise or sighting unexpected phenomena. Ekman (2002) suggests that it is very brief probably lasting a few microseconds. In Figure 7 below, surprise is triggered when the participant hears the agent's voice for the very first time and as Ekman and O'Sullivan (2006) suggest, the participant may have been unsettled by the instructor's voice. The expression is made possible by the following muscle actions (Table 1); AU1 raises inner brow, AU2 raises outer brow, AU5B slightly raises the upper lip, while AU26 makes the jaw drop down (Ekman 1997).

Figure 7: Surprise

Slight surprise is expressed with a lower intensity in (Figure 8) as the inner and outer brows are raised, but with AU5B, the upper lid is raised slightly when the jaw drops down (Ekman 2002).

Figure 8: Slight Surprise

In figure 8 surprise is triggered by the sudden realisation that previous assembly steps executed were incorrect after the participant repeats instruction 12. This leads to participant's self-correction during the task.

Smiles

The participants displayed felt, miserable and tight-lipped smiles during interaction.

Felt Smiles

The felt, full, or Duchenne smile is a positive emotion (Figure 9 below) elicited by positive stimulation such as amusement, contentment, and enjoyment of another person (Malik 2010; Matsumoto & Ekman, 2008). The facial movements responsible (Table 1) include AU6, that gathers the skin around the eyes inwards and narrows eye apertures, and AU7, which tightens the eye lids, raises the lower eye lid creating crow's feet wrinkles below it. AU12 pulls the lips sideways showing the teeth.

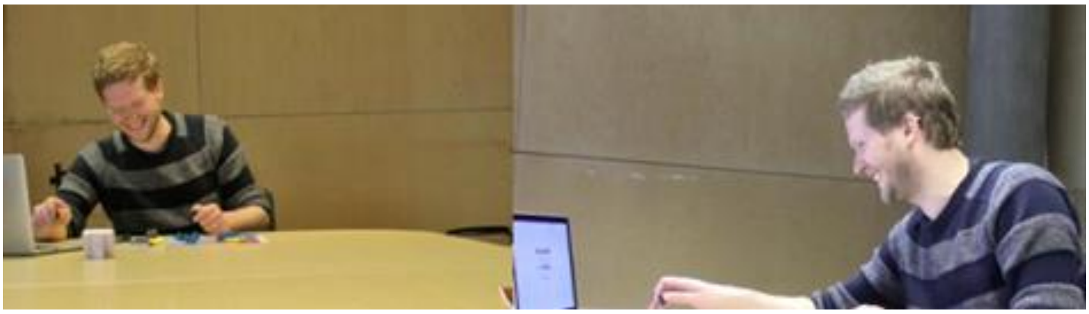


Figure 9: Felt smile

Sometimes a felt smile breaks into laughter as shown (Figure 9) above. Unlike the false smile, which is a deliberate attempt to project a positive emotion that does not exist the felt smile suggests that the participant is enjoying the interaction and probably found the agent amusing (Pease &Pease 2004).

Nervous Smile

The nervous or miserable smile (Ambadar et al. 2009;) is produced deliberately. It may be superimposed on a negative expression or come after it with the negative expression persisting (Ekman & Friesen 2003). The participant (Figure 10) below displays a mixed feeling of smiling and embarrassment elicited by the failure to execute the assembly task. The muscles responsible for this emotion (Table 1) include AU6, the cheeks raiser, AU12, the lip corners puller and AU54 provides the downward head movement.

Figure 10: Nervous Smile

The participant (Figure 10) above displayed embarrassment while smiling, similar to the participants in Ekman et al.'s 1980 study who looked unhappy while smiling (Ekman & Friesen 2003). The nervous smile is not as intense as the felt smile and may have a longer duration, greater amplitude, more head downward movement and open mouthed than the tight-lipped smile (Ambadar et al., 2009).

Tight-lipped Smile

The tight-lipped smile (Figure 11) below is performed by the lips stretching tight across the face in a straight line with the teeth unexposed (Ekman, 2002). Research suggests that this smile is sometimes used by people trying to hide their true feelings ((Malik, 2010; Pease &Pease 2004).

Figure 11: Tight-lipped Smile

The muscles responsible for this emotion (Table 1) include AU6B, which slightly raises the cheeks without visibly reducing the apertures of the eyes or creating very visible crow's feet around the eyes. AU12B slightly pulls the corners of the lips creating some wrinkles.

G. W Bush Grin

This facial action is named after the trademark smirk of President George W. Bush. The Bush grin shown (Figure 12) below is a smirk with a smile-like appearance (Pease & Pease 2004; Ekman, 2002) which becomes clearer when the upper part of the face is covered.

Figure 12: George W. Bush Grin

The Bush grin is indicated by the archetypal oblique eyebrows, lowered lip corners, slightly stretched mouth and raised cheeks. The muscles responsible for this facial action (Table 1) include AU1, raises the inner brow, AU15, depresses the lip corner, AU27B the mouth stretcher and AU6 raises the cheek. Furthermore, the eyelids are lowered by AU5, helping the eyes focus on the piece held up for examination, while an instruction is repeated.

Controlled Fear

Laboratory studies suggest that fear may be inbuilt rather than learned (Ekman & Friesen 2013). The clue for controlled fear (Figure 13 below) is slightly stretched and downwards shaped lips. The muscles responsible for this emotion (Table 1) include AU1, the inner brow raiser, AU2, the outer brow raisers AU4, the brow lowerer, AU5, the upper lid raiser; AU7 lip tightener AU 20 lip stretcher and AU 26 the jaw dropper (Ekman, 1997).

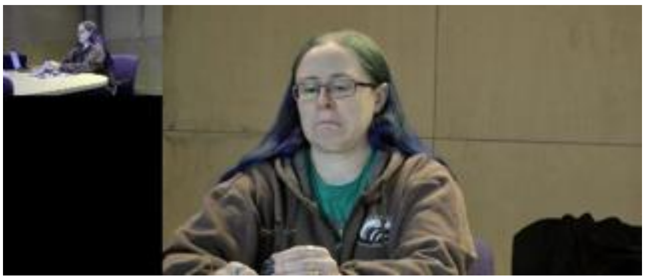


Figure 13 Controlled Fear

Slight Disgust

Disgust is "a feeling of aversion" (Ekman 2003: 190) that may be triggered by the senses encountering something offensive, such as a hated person, location or experience.

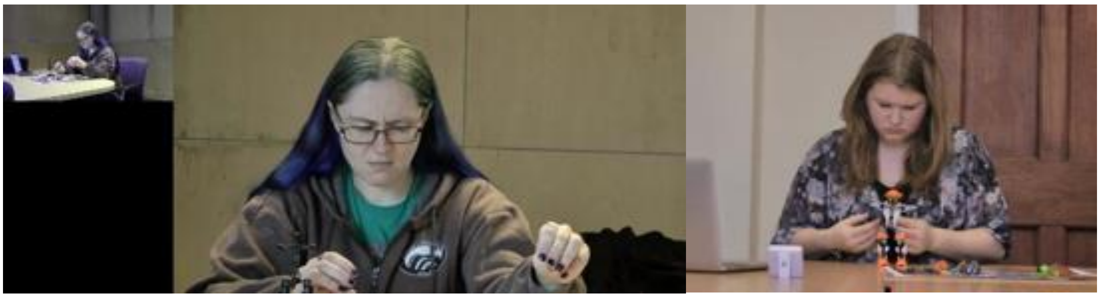


Figure 14: Slight Disgust

Disgust is operated by the slightly raised upper and lower lips, with the lower lip protruding. The nostril wings are raised slightly with wrinkles appearing on the sides, while the bridge of the nose and eyebrows are pulled down. These actions are produced by AU9 – the slight contraction of the muscle that wrinkles the nose – along with AU15B, the lip corner depressor, AU16B the lower lip depressor, and AU44B, a separate strand of AU4 that visibly narrows the yes even when the head is bent.

Non-Basic Emotion/Blends

Smug Expression

Blended emotion is formed as the participant (Figure 15) below displays a combination of enjoyment and contempt creating a smug expression (Ekman, 2003) during interaction.

Figure 15: Blend of emotions or smug expression

There are several muscles responsible for creating this facial action (Table 1). AU4, working in conjunction with AU42 a separate strand of AU4, lower and gather the eyebrows together. Simultaneously, AU8, brings the lips toward each other and compresses them while AU12, in conjunction with AU6B, tighten lip corners and raises the cheeks to form a slight smile.

Angry-Disgust

This blend of anger and disgust could be triggered by irritation or experience of a hated phenomenon or even frustration during interaction (Figure 16).

Figure 16: Angry Disgust

Although, observers may sometimes confuse anger for disgust because of intensity (Ekman, 2009), the emotions are basically characterised by a lowered brow, raised upper lip, flared nostrils, and the open mouth curved downwards. The facial muscles responsible (Table 1) are as follows: AU4, lowers the brow, AU7 lowers the eye lid, AU9 wrinkles the nose, AU15 depresses the lip corner and AU16 the lower lip depressor.

Emotional Attitude: Frown

Ekman (2003) describes a frown as an emotional attitude produced primarily by the furrowing of the eyebrows through the action of 'muscle of difficulty' because frowning occurs with mental or physical difficulties (Mutlu 2011). AU44 and AU46 lower and pull the brows together (Table 1) sometimes with the neck jerking forward.

Figure 17: Frown

The participants (Figure 17) above are looking down while one has fingers on the chin suggesting that they may be trying to concentrate or pay attention while listening to the repeated instruction, as opposed to being wrongly perceived as feeling unpleasant, sad, or angry (Ekman 2001)

Moods

Mood: Aha! Moment

The *Aha! Moment* is "a moment of sudden realisation, inspiration, insight, recognition, or comprehension" (Merriam-Webster Dictionary 2016). The form and function of the Aha! Moment is given in the vignette (Figure 18) below.

Phase 1- Impasse: A period of mental fixation indicated by an expressionless face thinking outside the box
Phase 2: Break in the mental fixation and the insight or solution appears suddenly and elicits positive effect like the full smile that occurs.

Figure 18: Aha! Moment

The Aha! Moment is realised through the experience of processing fluency (Topolinsky& Reber, 2015) in a two-phased process. In the first phase, the participant reaches an impasse or a period of mental fixation (Mai et al.,2004) and concentrates on the task without assembling the required parts. The second phase occurs suddenly when a break in the mental fixation occurs and the solution suddenly appears eliciting the participant's full smile in the second vignette.

Mood: Workman's face

This is disgust showing all over the face, blending with contempt and some anger, to form the workman's face under pressure. This emotion is realised by the following muscles: AU9, wrinkling the nose, AU15 the lip corner depressor raises the upper lip, AU16 the lower lip depressor raises the lower lip making it protrude, while AU7 the lid tightener, and AU12B slightly pulls the lip corner.

Figure 19: workman's face

In figure 19 above, the participant is under pressure and is exerting force to assemble the kit and people often press their lips together when doing anything that requires physical exertion such as pulling things apart or trying to lift heavy things.

Tense Mouth and Lip Action

Compressed or swallowed Lips

Compressed or swallowed lips (figure 20) below is a tense-mouth action showing obvious muscular tension produced by compressing, in-rolling, and narrowing the lips to a thin line position in which they are visibly tightened and pressed together. This is done through AU8 pulling or tightening the lips towards each other (Ekman 1997).

Figure 20: Compressed Mouth

In assembly tasks, lip compression may indicate cognitive processing such as pondering, thinking, or feeling uncertain about the instructions (Givens, 2015). The participant on the top-left of the vignette displays slight compression without swallowing the lips done by AU8B (Table 1)

Biting the Lip

The biting lip is shown (figure 21) below. The muscle bases (Table 1) for this action include AU12 pulling the lower lip corner over the lower jaw teeth and AU32 executing the lip bite (Ekman 2007).

Figure 21: Lip biting

Morris (2015) suggests that the lip bite coincides with anger. This may be tenable in HHI however, during assembly tasks, lip biting may indicate uncertainty while participants are processing instructions in the same manner that Kita's (2013) co-thought gestures indicate cognitive processing of information.

Pouty Face

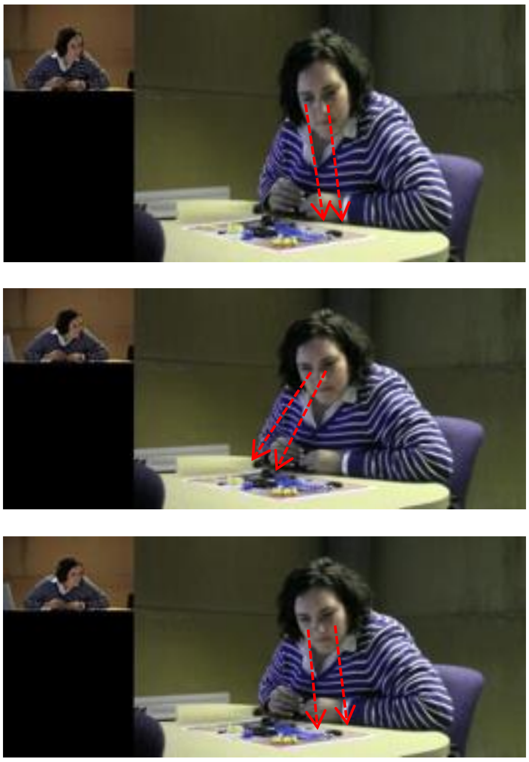
Givens (2010) observes that adults spontaneously pout when disagreeing with comments during interaction. When people pout, AU17 contracts the chin concurrently with AU25 pulling lips apart while AU 21 stretches the neck. Meanwhile, AU20 and AU7 collaboratively stretch the direct labial tractor muscles of the lower lip (Givens 2002) while AU 22 funnels the lips making the pout more prominent.

Figure 22: Pouty face

People use pouting lips in certain cultures to indicate direction when their hands are busy (Morris 2015). However, in assembly contexts the participant (Figure 22) may be pouting for concentration or self-comfort during the interaction.

Saccadic eye movement: Static searching face

This facial action is created by eye movement during interaction. The saccade eye movement is used to track an object without head movement (Dragoi 2015). The participant (Figure 22) below leans forward and listens to instruction repeated while searching for pieces described with her eyes sweeping across the table using three pendulum (right-left- right arrows) eye movements.



Initiation: Saccade eye movement from right to the left.

Return: Saccade eye movement back from left to the right

Final: the locating movement from right to left again.

Figure 23: Static Searching Face Saccade eye movements

The following AUs (Table 1) are responsible for saccadic movement: AU61 for turning left and AU62 for turning right. Ekman (1997) does not state the muscles responsible but research in human physiology indicates that the lateral and medial rectus muscles[7] may be responsible for saccadic eye movements (Dragoi, 2015).

Microexpressions

The participants displayed micro-anger, micro-disgust, micro-sadness, and micro-smile during interactions.

Micro-Anger

Ekman (2002) suggests that anger and fear often occur together, with fear preceding anger, while anger galvanises people to action thus, they have similar features. Anger may also be mistaken for confusion.

[7] Details on http://droualb.faculty.mjc.edu/Lecture%20Notes/Unit%203/muscles%20with%20figures.htm

Figure 24: Micro-Anger

The pointer to anger is that lips become thinner as participants (Figure 24) display the brief slight open-mouth anger (Ekman 2002). Even though, participants are trying to remain calm, emotion leaks probably due to confusion during the task (Ekman and Friesen 1997). Micro-anger is realised (Table 1) by the AU4B brow lowerer, AU5 upper lid raiser, AU7B lid tightener, and AU20B lip stretcher making the lips thinner. AU25B opens the mouth and keeps lips wide apart as if in speech, but the narrowing lips forming a square mouth is an early indicator of anger in the top-left picture.

Micro-Disgust

Micro-disgust observed shows on one side of the face with the mouth bending to the right. This emotion is realised (Table 1) by the AU9B nose wrinkle, AU15B lip corner depressor raising the upper lip, and the AU16B lower lip depressor raising the lower lip making it protrude.

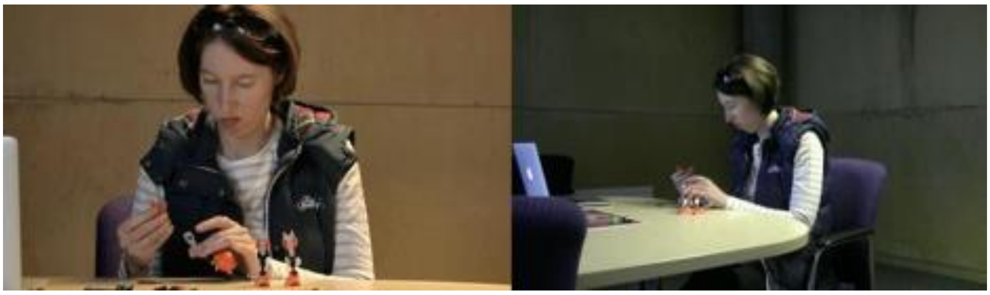


Figure 25: Micro-disgust

The participant (Figure 25) below displays a disgust that is different from the disgust displayed when adults encounter something unpleasant (Ekman& Friesen 2003). The participant directs this emotion involuntarily towards the effort she is putting into executing the instruction.

Micro-Sadness

Micro-sadness (Figure 26) is performed by the lips being turned down slightly. The muscles responsible for micro-sadness include AU15B, the lip corner depressor. This becomes more pronounced below as participants display slightly oblique eyebrows, slightly lowered lip corners, stretched mouth and slightly raised cheeks (Ekman 2007). The additional muscles responsible for the facial actions include the AU1B inner brow raiser, AU27B mouth stretcher and the AU6 cheeks raiser. The lips are slightly parted by AU25B (Ekman 2007).

Figure 26: Micro-Sadness

Sadness may have been triggered by failure to achieve a goal (Ekman 2002) such as unsuccessfully executing the agent's instruction, but the emerging pout may suggest determination.

Micro-Smile

Participants taking instructions carry out an appraisal of each instruction using the listener's automatic processing strategy. This process may trigger micro-smiles shown (Fig 27) below. Morphologic features responsible include the slight contraction of AU6B, raising the cheeks slightly and AU12B, the lip corner puller.

Figure 27: Micro-smile

A quick drawing up of eyebrows with lips spreading suggest that participants are experiencing enjoyment or feeling good with the task. Conversely, it may be that the participants are having a tough time and smiling to bear the pressure of having to execute the agent's instruction.

Micro Tight-lipped smile

The tight-lipped smile (Figure 28) is indicated by lips stretched tight across the face in a straight line with the teeth unexposed (Ekman & Friesen, 2003). The muscles responsible include the AU6B slight cheek raiser, while the eyelids are lowered by AU5B thus visibly reducing the eye apertures. AU12B slightly pulls the corners of the lips thus creating some wrinkles and the slight action of AU27B gives the smile a Bush-like grin.

Figure 28: Micro Tight-lipped smile

The sad element becomes clear when the upper part of the face is covered suggesting that the participant may be managing negative emotions with a smile (Ekman, 2002).

DISCUSSION

The discussion answers RQ2 and focuses simultaneously on how listener facial actions indicate their attitudes towards agents and the differences in listener perceptions of agent identity as positive, neutral, and negative during interaction**.** The results suggest that listener attitudes might be reliably detected and measured nonverbally through facial actions just as Meadors and Murray (2014) measured and classified bias through body language. Listener feelings and attitudes towards the interaction may also be distinguishable as positive, neutral or negative (Clark 2016; Mehrabian 2007).

Categories of facial actions displayed when participants repeat instructions may reflect positive, neutral, and negative attitudes towards the agent (Figure 29). Positive facial actions include felt smile, slight smile, George W Bush grin, tight-lipped smile, and controlled laughter and these represent 21.21% of the distribution. Neutral includes neutral, neutral concentration, workman effort face, and static searching head representing 59.09%. Negative facial actions represent 19.70% of the distribution and includes puzzled face, sadness, compressed or swallowed lips, slight anger, disgust, slight disgust, slightly compressed lips, frown and nervous smile.

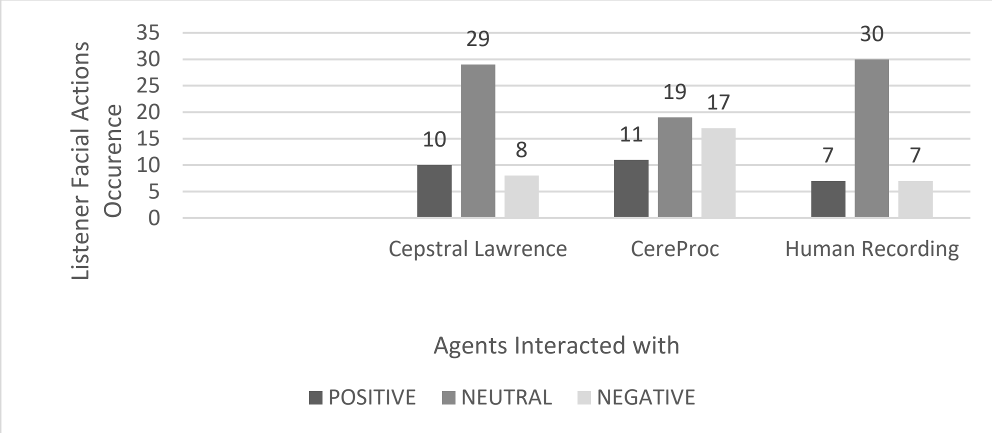


Figure 29: Inference of Listener perceptions of Agent Identity

Table 3: Degree of Listener Attitudes towards Simulated Agents

	Inferred attitude scores corresponding to Agent desired					
AGENT	POSITIVE		NEUTRAL		NEGATIVE	
	M	SD	M	SD	M	SD
CP	2.00	1.73	5.80	7.05	1.00	0.38
CL	2.75	1.50	3.80	3.83	1.70	0.86
HR	2.33	1.41	7.50	0.89	1.40	0.89

Facial actions were used to communicate three degrees of listener attitudes (table 3) classified as positive or likeability, negative or dislike and neutrality towards the agent's identities (Mehrabian, 2007). The facial communication of three degrees of attitude (Table 3) was assessed by positioning positive, neutral and negative on a scale of standard deviation from the mean of individual listener spontaneous facial actions displayed during interaction with the agent.

Although the results indicated that there were more neutral facial actions than any other, listeners were more consistent in the display of negative attitudes than neutral or positive ones. Table 3 above shows that negative attitudes were lower STD(X) = (0.38; 0.86; 0.89) for CP,CL, and HR respectively against neutral's STD(X) = (7.05; 3.83; 0.89) and positive's STD(X) = (1.73; 1.50; 1.41) for the same distribution. Within group results for the positive attitudes suggest that listeners may be happier with HR than CP or CL.

In addition, neutral attitude inferred from the study indicate that the STD(X) = (7.05; 3.83; 2.00) for CL and CP respectively suggests they were far off from the mean of each distribution thus reducing the impact on the interaction. This may be due to the need to regulate emotion (Gross & Levenson, 1997) when paying attention to either the instruction or the process for effective listening. This process suggests that people may develop restraint from systematic, learned, and relevant application of emotional triggers during interaction. Participants may also control their emotions probably because, they were no real danger rather they are trying not to destroy the fragile assembly piece.

The study suggests that facial actions are not restricted to emotions but can be used by participants as strategies employed organise task and interaction. Facial actions in this category include the static searching face used to locate assembly kits, compressed lips and biting lips used to indicate physical effort deployed by participants to handle tasks. Another is the *Aha! Moment*- a self-discovery process used to initiate self-correction during the assembly task.

From the foregoing, the results of the study partly support earlier research indicating that listener facial actions provide information about affective state, cognitive activity, temperament and personality (Donato, et al., 1999).

Emotions are crucial to the development and regulation of interpersonal relationship in human-human interaction context (Clark et al. 2016; Ekman & Rosenberg 1997). However, the findings indicate that participants displayed emotions even when they knew that the agent could not attend to them. These findings reinforce the view that, people may respond to computers in much the same way as they would to people (Nass et al. 1999) because, people process interactions with computers and computerised spaces in the same way they would do in real-life interaction contexts. This implies that, people may also develop some expectations of reciprocal treatment from agents during interaction based on emotional leakages that occurred as micro-expressions.

Listeners thus provided information to agents about the task, perceptions of agent identity, probable next behaviour while projecting their own identities during interaction. These suggest that decoding and understanding human behaviour may make agents more compatible with humans but, there may be issues of trust, effective communication, better understanding and respect that need to be considered.

Findings suggest that implicit attitudes externalised as spontaneous facial actions during interaction may be useful in assessing interaction (Sherman et al. 2003). For example, a positive face like the felt smile may indicate a listener’s positive disposition towards the instructor just as anger may mean the opposite. In this way, listener attitudes may subjectively organise interaction situations and orient them towards agents that are able to understand and use projected attitudes communicatively.

IMPLICATIONS FOR RESEARCH AND PRACTICE

The study has implications for sociolinguistic theories, research methodology and the development of agents with emotional intelligence.

Implications for sociolinguistic theories of nonverbal listenership relates to identity projection and perception in HAI. The study suggests that facial actions are not restricted to emotions but can be used by participants as assembling strategies during task execution for example, compressed lips and biting lips used to indicate physical effort deployed by participants when handling assembly bits.

Regarding the display of emotions, the results confirm earlier studies suggesting that implicit attitudes externalised as spontaneous facial actions during interaction may be useful in assessing interaction (Sherman et al. 2003). For example, a positive face like the felt smile (figure 9) may indicate a listener's positive disposition towards the instructor just as micro anger (figure 24) may suggest a negative disposition. In this way, listener attitudes may subjectively organise interaction situations and orient them towards agents that are able to understand and use projected attitudes communicatively.

Furthermore, these findings reinforce the view that, emotions are crucial to the development and regulation of interpersonal relationship in any interaction context (Clark et al. 2016; Ekman & Rosenberg 1997). Although, facial actions are easily accounted for in HHI as indicated by studies in human only interaction, there is a potential for linguistics to provide an informed basis that may account for the communicative use of listener facial actions in HAI.

Although the results support earlier research indicating that listeners' facial actions may provide information about their affective state, cognitive activity, temperament, personality and truthfulness (Donato, et al., 1999) during interaction, interpreting facial actions is to some extent personal, culturally determined and context-specific.

Following these findings, current theories of discourse identities need to carry out a careful analysis of talk-in-interaction (Schegloff, 1997) that may lead to the development of a better understanding of active role-play in the construction of social identities in HAI.

The results imply that there is the need for applied linguistic research is to devise a multimodal corpus linguistics coding matrix for annotating various co-occuring nonverbal listenership behaviours during interaction with a view to reconciling these fragments into a coherent discourse at the level of analysis. This may clarify how multimodal fusion is achieved in perception and understanding of communication. For example, reconciliation at the level of analysis may enable researchers to understand how interlocutors integrate co-interlocutor's facial actions, voice, and posture to arrive at a multimodal interpretation of information exchanged during interaction. Agent designers have to consider the impact of facial actions alongside language in computer ability to use listener feedback communicatively.

CONCLUSION

This paper has been able to highlight the need to understand how agents can use nonverbal feedback communicatively at near human levels in HAI. The paper used assembly tasks to elicit spontaneous facial actions as against posed ones because, spontaneous nonverbal behaviour will reliably project people's feelings than posed ones. Findings suggest that in the interaction between agents and humans, while agents use voice and speech to create and project their identities, humans can reliably use facial actions to project their perceptions of agent identities and the interaction. As we move into a world where there is increasing pervasive computing and interaction with agents becomes prevalent, it is the view of this paper that we need to understand how these agents may alter how interaction takes place in diverse contexts.

Further Research

Findings suggest that facial actions are meaningful, however, we still need to understand how agents may perceive and utilise their form and functions in HAI. Research could also focus on the use of corpus approaches to perform metacognitive analyses of HAI in non-instructive contexts. As this is a laboratory study, this type of data may be re-evaluated in real-life contexts to enhance our understanding of HAI.

Further research may extend the voice cline used in this study to include a real human instructor as this may provide a basis for comparing HHI and HAI within same experiment contexts. In addition, low-risk studies in this area may provide some guidelines for higher-risk ventures for example, Lego building studies may provide guidelines for designing advice-giving and task-handling agents.

REFERENCES

Allwood, J. (1993). Feedback in second language acquisition. Adult Language Acquisition. Cross Linguistic Perspectives 2, 196-236.

Atkinson, L., & Bradley, J. (2017). Meaning making and collaborative ethnography in transdisciplinary arts. Working Papers in Translanguaging and Translation, WP23. Birmingham, United Kingdom: University of Birmingham.

Bach, J. R., & Goncalves, M. R. (2004). Noninvasive ventilation or paradigm paralysis? European Respiratory Journal, 23(4), 651-651.

Canales, M. K. (2000). Othering: Toward an understanding of difference. Advances in Nursing Science 22.4, 16-31.

Clark, L. M. (2016). Exploring Vague Language Use in Human-Agent Interaction. University of Nottingham, School of Computer Science. Unpublished.

Clark, L. M., Bachour, K., Ofemile, A. Y., Adolphs, S., & Rodden, T. (2014). Potential of imprecision: exploring vague language in agent instructors. Proceedings of the second international conference on Human-agent interaction (HAI '14) (pp. 339-344). New York NY, USA,: ACM. Doi: http://dx.doi.org/10.1145/2658861.2658895

Clark, L. M., Ofemile, A. Y., Adolphs, S., & Tom, R. (2016). A Multimodal Approach to Assessing User Experiences with Agent Helpers. ACM Transactions on Interactive Intelligent Systems (TIIS), 6(4), 1-31.

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. G. (2001). Emotion recognition in human– computer interaction. IEEE Signal Processing Magazine, 18(1), 32–80.

De Fina, A. (2006). Group identity, narrative and self-representations. Studies in Interactional Sociolinguistics. In A. De Fina, D. Schiffrin, & M. Bamberg, (Eds). Discourse and Identity (Vol. 23, pp. 351-375). Cambridge: Cambridge University Press.

Donato, G., Bartlett, M. S., Hager, J. C., Ekman, P., & Sejnowski, T. J. (1999). Classifying facial actions. IEEE Trans. Patt. Anal. Machine Intell. DOI: http://dx.doi.org/10.1109/34.799905 , 21(10), 974–989.

Dörnyei, Z. (2007). Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies. Oxford, UK: Oxford University Press.

Dragoi, V. (2015). Ocular Motor Control Section3: Motor Systems. In J. H. Bryne, Neuroscience Electronic Textbook (p. Chapter 8). Retrieved 12 14, 2015, from http://neuroscience.uth.tmc.edu/s3/chapter08.html

Edgley, C. (2016). The Dramaturgical Legacy of Erving Goffman. In C. Edgley (Ed) The Drama of Social Life (pp. 74-89). London: Routledge.

Ekman, P. (1992a). An argument for basic emotions. Cognition & emotion, 6(3-4), 169-200.

Ekman, P. (1992b). Are there basic emotions? Psychological Review, 99(3), 550-553.

Ekman, P. (1997). What We Have Learned by Measuring Facial Behaviour. In P. Ekman, & E. Rosenberg (eds), What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS) (pp. 469-486). Oxford: OUP.

Ekman, P. (2002). Facial action coding system (FACS). A human face.

Ekman, P. (2004). Emotional and conversational nonverbal signals. In J. M. Larrazabal, & L. A. Pérez Miranda (Eds.), Language, knowledge, and representation: Proceedings of the Sixth International Colloquium on Cognitive Science (ICCS-99) (pp. 39-50). Netherlands: Springer.

Ekman, P. (2007). Emotions revealed: Recognizing faces and feelings to improve communication and emotional life. New York, USA: Macmillan.

Ekman, P. (2009). Darwin's contributions to our understanding of emotional expressions. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364(1535), 3449-3451.

Ekman, P. (2016). What Scientists who study Emotion Agree About. Perspectives on Psychological Science, 11(1), 31-34.

Ekman, P., & Friesen, W. V. (2003). Unmasking the face. Cambridge: Major Books.

Ekman, P., & O'Sullivan, M. (2006). From flawed self-assessment to blatant whoppers: The utility of voluntary and involuntary behavior in detecting deception. Behavioral sciences & the law, 24(5), 673-686.

Ekman, P., Friesen, W. V., & Ellsworth, P. (2013). Emotion in The Face: Guidelines for research and an Integration of findings. New York: Pergamon Press Ltd.

Goffman, E. (2002). The presentation of self in everyday life. Garden City, NY.

Gross, J. J., & Levenson, R. W. (1997). Hiding Feelings: The acute Effects of Inhibiting Negative and Positive Emotion. Journal of Abnormal Psychology, 106(1), 95-103.

Hess, U., & Kleck, R. E. (1997). Differentiating Emotion Elicited and Deliberate Emotional Facial Expression. In P. Ekman, E. Rosenberg, & (Eds), What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS) (pp. 271-286). Oxford: OUP.

Jennings, N. R., Moreau, L., Nicholson, D., Ramchurn, S., Roberts, S., Rodden, T., & Rogers, A. (2014). Human-agent collectives. Communications of the ACM, 57(12), 80-88.

Kim, Y., & Tracy-Ventura, N. (2012). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? System, 41, 829-840.

Kita, S. (2013). Architectural issue in the model of Speech-gesture production: Gesture, Action language.

Klein, J., Moon, Y., & Picard, R. (2002). This Computer Responds to User Frustration: Theory, Design, Results, and Implications. Interacting With Computers, 14(2), 119-140.

Luger, E., & Sellen, A. (2016). "Like Having a Really bad PA": The Gulf The Gulf The Gulf between User Expectations and Experience of Conversational Agents. CHI 2016 (pp. 5286-5297). San Jose, CA USA: ACM.

Mai, X. Q., Luo, J., Wu, J. H., & Luo, Y. J. (2004). Aha!" effects in a guessing riddle task: An event-related potential study. Human brain mapping, 22(4), 261-270.

Malik, A. A. (2010). The secrets of body language. Retrieved 05 2015, from SlideShare: http://www.slideshare.net/arifanees/secrets-of-body-language

Marine-Roig, E. (2015). Identity and authenticity in destination image construction. Anatolia, 26(4), 574-587.

Matsumoto, D., & Ekman, P. (2008). Facial expression analysis. Scholarpedia, 3(5), 4237.

Meadors, J. D., & Murray, C. B. (2014). Measuring nonverbal bias through body language responses to stereotypes. Journal of Nonverbal Behavior, 38(2), 209-229.

Mehrabian, A. (2007). Nonverbal Communication. New Brunswick (USA) and London (UK): Aldine Transaction.

Merriam-Webster Dictionary. (2012). Definition of AHA MOMENT. Retrieved May 05, 2015, from Merriam-Webster Dictionary: https://www.merriam-webster.com/dictionary/aha%20moment

Miles, L. K., Nind, L. K., & Macrae, C. N. (2009). The rhythm of rapport: Interpersonal synchrony and social perception. Journal of Experimental Social Psychology, 45, 585-589.

Miller, J. (Director). (2014). Top Documentary Films - The Secrets of Body Language [Motion Picture]. Retrieved from https://www.youtube.com/watch?v=HTb7XQGKp-g

Morris, D. (2015). Bodytalk: A world guide to gestures. Random House.

Mutlu, B. (2011, fall). Human-Computer Interaction: Step-By-Step Experimental Design. Retrieved January 21, 2014, from University of Wisconsin: http://hci.cs.wisc.edu/co.../HCI-Week06-Lecture07.pdf

Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. Journal of Experimental Psychology: Applied, 7(3), 171-181.

Nass, C., Moon, Y., & Carney, P. (1999). Are people polite to computers? Responses to computer-based interviewing systems. Journal of Applied Social Psychology, 29(5), 1093-1110.

Ofemile, A. Y. (2015). Assessing affordances of selected cloud computing tools for Language Teacher Education in Nigeria. The Journal of Education and Practice, 6(3), 1-10.

Ofemile, A. Y., Clark, L., Adolphs, S., & Rodden, T. (2016). A Multimodal Assessment of Facial Actions as Listener perceptions in Human-Computer Interaction. The Inter-varietal Corpus Studies (IVACS) Conference. Bath Spa University, UK,.

Poole , D., & Mackworth , A. (2010). Artificial Intelligence: Foundations of Computational Agents. Cambridge University Press.

Scherer, K. R. (2005). What are emotions? And how can they be measured? Social science information, 44(4), 695-729.

Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. Journal of Social and clinical Psychology, 13-39.

Taylor, S. (2009). Who do we think we are? Identities in everyday life. In S. Taylor, S. Hinchliffe, J. Clarke, & S. Bromley(Eds), Introducing the Social Sciences: Making Social Lives (pp. 165-202). Milton Keynes, UK: The Open University.

Topolinski, S., & Reber, R. (2010). Gaining insight into the "Aha" experience. Current Directions in Psychological Science, 19(6), 402-405.

Tsuchiya, K. (2013). Listenership Behaviours in Intercultural Encounters. Amsterdam, The Netherlands; Philadelphia, USA: The Benjamins Publishing Co.

Williams, R., Karousou, R., Gumtau, S., & Howell-Richardson, C. (2009). Nested Narratives: Enabling Students to make sense of their Learning. Brookes e-Journal of Learning and Teaching. Retrieved /03/20/2010, from http://bejlt.brookes.ac.uk/article/nested_narratives_enabling_students_to_make_sense_of_their_learning/

Zimmerman, D. H. (1998). Identity, Context and Interaction. In C. Antaki, & S. Widdicombe (Eds), Identities in talk (pp. 87-106). London: Sage.