Talking Through Your Arse: Sensing Conversation with Seat Covers

Sophie Skach (s.skach@qmul.ac.uk)
School of Electronic Engineering and Computer Science,
Queen Mary University of London, E1 4NS

Patrick G. T. Healey (p.healey@qmul.ac.uk)
School of Electronic Engineering and Computer Science,
Queen Mary University of London, E1 4NS

Rebecca Stewart (rebecca.stewart@qmul.ac.uk)
School of Electronic Engineering and Computer Science,
Queen Mary University of London, E1 4NS

Abstract

People move in characteristic ways during conversation and these movements correlate with their level of particpation. For example, speakers normally gesture significantly more than listeners. These visible, overt movements are normally analysed using full body video or motion capture. Here we explore the potential of a 'minimal' approach to sensing these participatory movements in part of the natural environment of everyday interactions; chair seat covers. Using custom built fabric sensors we test whether we can detect people's involvement in a conversation using only pressure changes on the seats they are sitting in. We show that even from this impoverished data we can distinguish between talking, backchanneling and laughter; each state is associated with distinctive patterns of pressure change across the surface of the chair. We speculate on the possible applications of this new, unintrusive form of social sensing for architecture, performance and augmented human interaction.

Keywords: human interaction; dialogue; non-verbal communication; social sensing; smart textiles; posture analysis; fabric sensors;

Introduction

People make a variety of distinctive body movements during conversation. The most commonly studied of these are the gestures that speakers produce while talking. These include gestures that contribute to the content of what is said, such as iconics, metaphorics and pantomimes (McNeill, 1992; de Ruiter, 2000), as well as gestures that help to orchestrate the interaction such as beat gestures and gestures that can hold or hand over the turn to someone else (Bavelas, Chovil, Lawrie, & Wade, 1992; Healey & Battersby, 2009). Listener's body movements are also organised in characteristic ways. Most obviously through the production of concurrent feedback or 'backchannels' (Yngve, 1970). Although these are often produced as non-interruptive verbal acknowledgements such as a brief "aha" or "mmhm" people also frequently backchannel by nodding in response to an ongoing turn. Listeners are also distinguished from speakers by their relative lack of hand movement although they move their hands much more when a speaker requests clarification or makes repairs to their turn (Healey, Plant, Howes, & Lavelle, 2015).

The significance of this non-verbal choreography is illustrated by how much we can infer about an interaction from the observation of body movements alone. We can often tell just by looking at who is talking to whom, who -if anyone- is listening, who is likely to speak next, whether the interaction is hostile or friendly and so on (Kendon, 1990). These inferences from non-verbal performances can be striking; people appear to be able to make reliable estimates of the quality of someone's teaching over a whole semester from a single 5 second video of body movements alone (Ambady & Rosenthal, 1992).

Research on non-verbal communication has tended to focus on these relatively large scale overt body movements; they are the easiest signals for participants to perceive and respond to and the most tractable for analysis (although see e.g. Ekman & Friesen 1969). Research typically takes advantage of video and, more recently, motion capture equipment to capture and analyse these movements (e.g. Healey & Battersby 2009; Vinciarelli et al. 2009. The rapid development of new sensor technologies and their application to social signal processing has opened an intriguing new space of possibilities for detecting patterns of interaction (Vinciarelli, Pantic, & Bourlard, 2009). For example, it is possible to detect people's levels of interest, stress and intoxication in conversation using the speech signal alone i.e. without knowing anything about the content of what is said (B. W. Schuller & Rigoll, 2009; B. Schuller et al., 2013). In contrast to relatively intrusive technologies such as video or automatic speech recognition, this approach makes it possible to create anonymised 'minimal' forms of social sensing by using textile technology (see also e.g. Rekimoto 2001).

Here we explore the potential of this approach for one of the most commonly used parts of the physical environment for social interaction; chairs. Even the shape and position of unoccupied, uninstrumented chairs can indicate a great deal about interaction; chairs around a small table suggest something very different from chairs in rows (see also Anderson 1996). Moreover, chair covers are often made out of stretch and soft fabric that makes the textile surface itself a potentially promising sensing material. Using metallic yarn gives a fabric conductive properties so that it can be turned into a pressure sensitive surface. Different possibilities of using textile surfaces as sensing materials or interfaces for electronic devices have been explored in recent years, for example by turning a jacket into an interface to a mobile phone (Poupyrev et al., 2016) or by measuring biomechanical data for healthcare applications (e.g. Pacelli et al. 2006). Here we apply a similar fabrication technique to chair covers to address the basic question of whether it is possible to detect patterns of conversational interaction from movements on the chair surface alone.

Sensing Chairs

Informal observation suggests that people frequently change the position of the torso, lower body, and feet during seated conversations. These movements necessarily cause pressure changes on the surface of the chair and are therefore potentially detectable by measuring changes in resistance. Previous work has investigated the use of chairs to classify postures through pressure sensors, creating pressure maps of both, static and dynamic postures - posture identification versus continuous tracking (Tan, Slivovsky, & Pentland, 2001; Slivovsky & Tan, 2000). A commercially available pressure measurement system, BPMS (Body Pressure Measurement System) by Tekscan [1] has been used in some of these research projects (e.g. D'Mello et al. 2007 and Arnrich et al. 2010), which consists of a plastic mat with 64 integrated pressure sensors that allow for the creation of detailed pressure maps. The main applications for these sensing systems have been in the analysis of posture to improve seating comfort (e.g. Milivojevich et al. 2000), designs for objects involved in rehabilitation (e.g., wheelchairs) and Human-Computer-Interaction. For example, presenting chairs as novel haptic interfaces for computer games (Tan, Slivovsky, & Pentland, 2001), or as a system to measure people's cognitive states in various situations Arnrich et al. (2010), including measuring a car driver's fatigue (Furugori, Yoshizawa, Iname, & Miura, 2003) or identifying drivers (Riener & Ferscha, 2008), as well as measuring boredom in students (D'Mello, Chipman, & Graesser, 2007). However, this approach has not previously been applied to sensing aspects of social interaction. With this study, we explore what information about social behaviour can be retrieved from pressure sensor data on a chair.

Methods

Drawing on informal observations of people's leg and torso movements in meetings we decided on a configuration of eight sensors that were integrated in the chair cover and distributed in a symmetric arrangement; four in the seat of the chair and four on the back (see Figure 1), dividing the chair into four key areas to be sensed in order to determine postures: shoulders (at the top of the back rest), waist (lower back), buttocks and thighs. These observations also laid the basis upon which initial hypotheses about different states in a conversation were built.

[1] see https://www.tekscan.com/

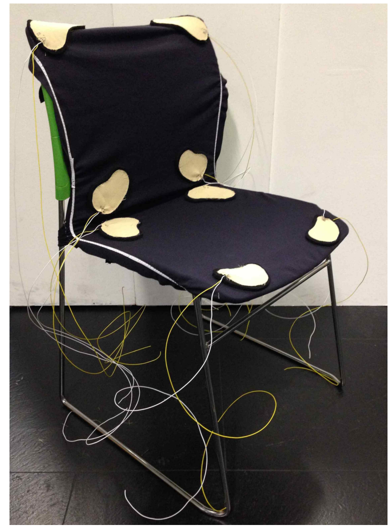

Figure 1: Reverse side of the chair cover showing sensors.

Sensor Development

The textile sensors were made from conductive fabric and resistive foam, hand sewn into soft sensor patches that were manually attached to the backside of a chair cover (which was made of jersey knit fabric). The conductive fabric, SaniSilver, was purchased from LessEMF and woven with a silver yarn showing on one side of the fabric and a cotton yarn visible on the other. The sensors are constructed such that two swatches of conductive fabric are facing the resistive foam on both sides. When pressure is applied to the conductive fabric on either side of the foam, the foam compresses and reduces the resistance between the two fabric swatches. This change in resistance is measured by the microcontroller. The sensors have the advantage of behaving like an ordinary fabric that could also be used in other wearable applications, such as garments (since, through the use of cotton fibre, the fabric retained a soft touch and remained comfortable to wear).

Data Collection

A microcontroller (a Teensy 3.2) collected the pressure data from the sensors and stored it on micro SD cards. The sampling frequency of the sensors was 4 readings per second. Using these piezoresistive sensors, the unit of measure is Ohm (Ω). Since the aim was to investigate postural behaviour in a situation of social interaction, three chair covers were manufactured, each housing one micro controller that were placed underneath the chair. Wires were hooked into the conductive fabric and connected the sensors with the Teensy (to ground and to an assigned analog pin providing 3.3 volts to run the programme, which read analog output values from the sensors).

Participants

Participants were recruited in groups of three friends or colleagues to ensure they all had some initial level of familiarity with each other. We conducted 9 trials in total, collecting data from 27 participants, of which were 11 female and 16 male and between the age of 20 and 40.

Procedure

The experiment was carried out in the Human Interaction Lab at Queen Mary University of London. Groups of three participants were asked to resolve a moral dilemma: the balloon task. This is a fictional scenario describing three people in a hot air balloon that is about to crash, if not one of the passengers jumps to their certain death. The task is then to come to an agreement of who to throw off. The participants were told that the aim of the experiment is to investigate collaborative interaction. They were seated at a round table and asked to discuss options and come to an agreement on how to resolve this dilemma. We aimed to record 15 to 20 minutes of conversation, so if not having come to an agreement after this time, participants were given the option to stop the conversation or carry on (vice versa, if they came to an agreement faster, alternative scenarios were provided to encourage further discussion). Due to the materiality of the sensors, the presence of the sensor patches was not noticed by the participants, so that the experience wasn't different to sitting on a common chair.

Data Analysis

The interactions were captured on two cameras placed in different corners in the room. Lapel microphones were used to facilitate speaker-specific analysis of the audio for transcription. The data from the video recordings was annotated using Elan (Brugman, Russel, & Nijmegen, 2004). Coding focused on three key behaviours with: speaking, laughter and backchannels. When determining speaking modes, periods of overt speech were coded, regardless of postural and gestural changes, or nodding. But focusing on postural movement overall, it was noticed that often, a postural or gestural change was performed immediately prior to speaking. This makes the start of an utterance ambiguous. For the purposes of this study, the beginning of utterances was defined as the onset of speaking. For laughter, responsive as well as speakers concurrent laughter was noted. Therefore, laughter is annotated for both, speakers and listeners. Backchannels were coded for all continuous verbal particles of response, as well as repair initiations. An overview of the coding scheme for these behavioural cues can be seen in Table 1.

Table 1: Coding scheme used in Elan.

Tiers per participant	Social behaviour
speaking	verbal utterance
laughter	responsive and concurrent
backchannel	responsive, repair initiation

Following this coding scheme, all elements that mark listening modes are created from the gaps of the annotations for laughter, speaking and backchannels. This means that within the listening mode, any gross and subtle body movement, as well as nodding or any other conversational action is included. With the aim of distinguishing speakers from listeners, this level of detail in annotations is sufficient, although the sensitivity of the sensors allows for richer and more fine-grained distinctions.

Results

The data from all eight sensors were analysed in a General Linear Model Multivariate Regression using SPSS v.24. Talking, Laughing and Backchanneling were included as binary predictors coded as 1 or 0 for presence / absence of each behaviour. All two and three-way interactions of these three factors were included in the model. Participants were also included as a main effect to ensure individual variation was accounted for.
Since the relative changes for each participants were calculated, changes of weight had no effect on the outcome of the analysis.

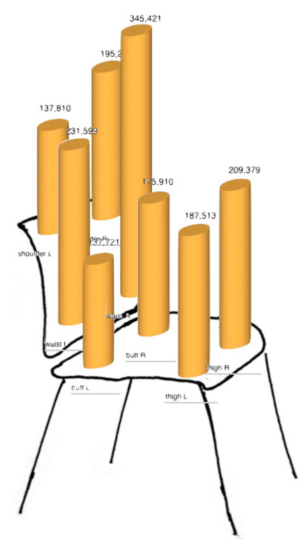

Figure 2: Estimated means of all participants for TALK: thighs: left(187.513), right(209.379); butt: left(137.721), right(175.910); waist: left(231.599), right(345.421); shoulders: left(137.810), right(195.288)

Multivariate Tests (Pillai's Trace) show all three dialogue factors reliably predict the outputs of the pressure sensors (Talk: $F_{(8,82933)} = 9.68, p < 0.00$; Backchannel $F_{(8,82933)} = 10.2, p < 0.00$; Laugh: $F_{(8,82933)} = 6.95, p < 0.00$;). The effects are very small with Partical Eta Squared of 0.001 and observed power for Alpha = 0.05 of 1. The contribution of individual variation is, by contrast, much larger: Participant: $F_{(8,82933)} = 6.95, p < 0.00$, Partial Eta Squared = 0.71).

Analysis of the contributions of each sensor show that different patterns of pressure changes across the chair are associated with the different dialogue states. The sensors most sensitive to talking were in the seat of the chair and correspond

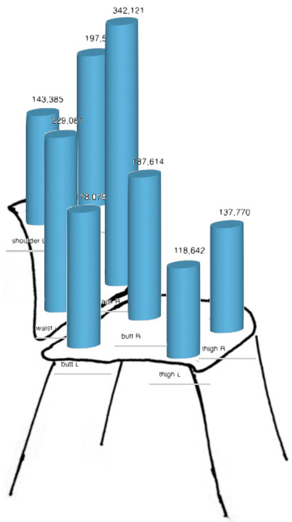

Figure 3: Estimated means of all participants for LAUGHTER: thighs: left(118.642), right(209.379); butt: left(178.079), right(187.614); waist: left(229.081), right(342.121); shoulders: left(143.385), right(179.532)

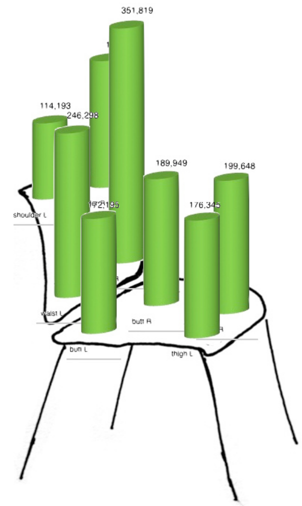

Figure 4: Estimated means of all participants for BACKCHANNELS: thighs: left(176.345), right(199.648); butt: left(172.195), right(189.949); waist: left(246.298), right(351.819); shoulders: left(114.193), right(189.709)

to increased pressure from the thighs and reduced pressure from the buttocks. In contrast to this laughter corresponded to reduced pressure in the thighs and increased pressure in the buttocks with no significant changes detected in the seat back. The pattern of pressure changes for the relatively brief backchannels were distributed across both the seat and back of the chair and corresponded to increased pressure across thighs, buttocks and waist but a reduction across the shoulders. The estimated means for changes at each sensor are illustrated in Figures 2, 3 and 4 (numbers based on modified population marginal mean).

Discussion

The results show that it is possible, in principle, to detect significant aspects of social interaction from quite limited, indirect and noisy data. The small movements detected by pressure sensors embedded in chair seats are small-scale and almost completely invisible correlates of the gross body movements that typically distinguish speakers from hearers and laughter from silence. Interestingly, even the relatively small nodding movements of the head associated with backchannels appear to create a distinguishable pressure signature on a chair.

This is the first attempt to detect significant conversational states from simple 'homemade' pressure sensors and the signal to noise ratio is low. Individual variations in movement in particular account for far more of the variance than differences in dialogue state. Further work to optimise the size and position of the sensors would doubtless improve the quality of the sensing. It is also likely that other approaches, such as training person-specific classifiers and machine learning mechanisms, would improve the accuracy and robustness of the approach although this would also undermine the advantages of anonymity. The demonstration that even relatively crude sensors can detect minimal changes in posture, suggests that future work should explore the possibility of capturing more complex social behaviour, especially relational questions such as whether interactions are, for example: convivial or combative; autocratic or egalitarian, or whether it is possible to characterise regularities in multiparty interaction (see e.g. (Abney, Paxton, Dale, & Kello, 2014)).

What could this form of sensing be used to do? The principle opportunities for application are in any situations where there is value in the ability to unintrusively gather information about general patterns of social interaction including levels of interest and engagement. One example is architecture where the ability to sense a building's energy performance and patterns of air flow is highly valued but currently has no social counterpart. We speculate that the ability to make simple, systematic assessments of a building's 'social performance' by instrumenting the chairs in a building could also have a significant positive impact on domestic and workplace design. A second example is in the evaluation of audience responses (e.g. continuous audience response measure, CARM, which is used by broadcast hosts to evaluate their programs). The deployment of such a sensor network in an auditorium, meeting room or a classroom could help to assess levels of engagement of students and other audiences. In addition, there are possibly applications to augmented human interaction where, for example, live feedback about how much people are dominating (or not) a conversation can have significant effects on the conduct of the interaction (Donath, 2002). If nothing else these results shed some light on Stephen Fry's (1984) advice that when delivering Shakespeare one should "always gather from the buttocks".

Summary

This paper presents a new sensing system using textile pressure sensors that are designed to be integrated in a chair cover and that are able to reliably distinguish speakers from listeners and detect laughter and backchannels. These fabric sensors provide a non-intrusive way to measure conversational engagement. Data about pressure changes on the seat and back rest alone make it possible to differentiate various behavioural states in a seated conversation. The ability to extract such patterns of social interaction from sensing pressure changes could replace other, more complex motion detection systems and mitigate privacy concerns, since the data collection is anonymous involves no audio or video data and does not capture any of the content of the conversation.

Acknowledgments

We thank the EPSRC and AHRC Centre for Doctoral Training in Media and Arts Technology, who funded this research. In particular, we also thank Adan Benito Temprano for his support, as well as the Digital Catapult Centre.

References

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2014). Complexity matching in dyadic conversation. *Journal of Experimental Psychology: General*, *143*(6), 2304.

Ambady, N., & Rosenthal, R. (1992). *Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis.* American Psychological Association.

Anderson, R. J. (1996). A security policy model for clinical information systems. In *Security and privacy, 1996. proceedings., 1996 ieee symposium on* (pp. 30–43).

Arnrich, B., Setz, C., La Marca, R., Tröster, G., & Ehlert, U. (2010). What does your chair know about your stress level? *IEEE Transactions on Information Technology in Biomedicine*, *14*(2), 207–214.

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse processes*, *15*(4), 469–489.

Brugman, H., Russel, A., & Nijmegen, X. (2004). Annotating multi-media/multi-modal resources with elan. In *Lrec.*

de Ruiter, J. P. (2000). 14 the production of gesture and speech. *Language and gesture*, *2*, 284.

D'Mello, S. S., Chipman, P., & Graesser, A. (2007). Posture as a predictor of learner's affective engagement. In *Proceedings of the cognitive science society* (Vol. 29).

Donath, J. (2002). A semantic approach to visualizing online conversations. *Communications of the ACM*, *45*(4), 45–49.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, *32*(1), 88-106.

Furugori, S., Yoshizawa, N., Iname, C., & Miura, Y. (2003). Measurement of driver's fatigue based on driver's postural change. In *Sice 2003 annual conference* (Vol. 1, pp. 264–269).

Healey, P. G., & Battersby, S. A. (2009). The interactional geometry of a three-way conversation. In *Proceedings of the 31st annual conference of the cognitive science society* (pp. 785–790).

Healey, P. G., Plant, N., Howes, C., & Lavelle, M. (2015). When words fail: collaborative gestures during clarification dialogues. In *2015 aaai spring symposium series: Turn-taking and coordination in human-machine interaction.*

Kendon, A. (1990). Spatial organization in social encounters: The f-formation system. *Conducting interaction: Patterns of behavior in focused encounters*, 209–238.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.

Milivojevich, A., Stanciu, R., Russ, A., Blair, G., & Van Heumen, J. (2000). *Investigating psychometric and body pressure distribution responses to automotive seating comfort* (Tech. Rep.). SAE Technical Paper.

Pacelli, M., Loriga, G., Taccini, N., & Paradiso, R. (2006). Sensing fabrics for monitoring physiological and biomechanical variables: E-textile solutions. *Proceedings of the 3rd IEEE-EMBS International Summer School and Symposium on Medical Devices and Biosensors, ISSS-MDBS 2006*, 1–4. doi: 10.1109/ISSMDBS.2006.360082

Poupyrev, I., Gong, N.-W., Fukuhara, S., Karagozler, M. E., Schwesig, C., & Robinson, K. E. (2016). Project jacquard: Interactive digital textiles at scale. In *Proceedings of the 2016 chi conference on human factors in computing systems* (pp. 4216–4227).

Rekimoto, J. (2001). Gesturewrist and Gesturepad: Unobtrusive Wearable Interaction Devices. *ISWC '01 Proceedings of the 5th IEEE International Symposium on Wearable Computers*, 21–27.

Riener, A., & Ferscha, A. (2008). Supporting implicit human-to-vehicle interaction: Driver identification from sitting postures. In *The first annual international symposium on vehicular computing systems (isvcs 2008)* (p. 10).

Schuller, B., Steidl, S., Batliner, A., Burkhardt, F., Devillers, L., MüLler, C., & Narayanan, S. (2013). Paralinguistics in speech and languagestate-of-the-art and the challenge. *Computer Speech & Language*, *27*(1), 4–39.

Schuller, B. W., & Rigoll, G. (2009). Recognising interest in conversational speech-comparing bag of frames and supra-segmental features. In *Interspeech* (pp. 1999–2002).

Slivovsky, L. A., & Tan, H. Z. (2000). A real-time sitting posture tracking system.

Tan, H. Z., Slivovsky, L. A., & Pentland, A. (2001). A sensing chair using pressure distribution sensors. *IEEE/ASME Transactions On Mechatronics*, *6*(3), 261–268.

Vinciarelli, A., Pantic, M., & Bourlard, H. (2009). Social signal processing: Survey of an emerging domain. *Image and vision computing*, *27*(12), 1743–1759.

Yngve, V. H. (1970). On getting a word in edgewise. In *Chicago linguistics society, 6th meeting* (pp. 567–578).