This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Deposited on: 21 November 2017

Body Language without a Body: Nonverbal Communication in Technology Mediated Settings

Alessandro Vinciarelli

University of Glasgow - School of Computing Science - Glasgow, United Kingdom

Alessandro.Vinciarelli@glasgow.ac.uk

ABSTRACT

Cognitive and psychological processes underlying social interaction are built around face-to-face interactions, the only possible and available communication setting during the long evolutionary process that has resulted into Homo Sapiens. As the fraction of interactions that take place in technology mediated settings keeps increasing, it is important to investigate how the cognitive and psychological processes mentioned above - ultimately grounded into neural structures - act in and react to the new interaction settings. In particular, it is important to investigate whether nonverbal communication - one of the main channels through which people convey socially and psychologically relevant information - still plays a role in settings where natural nonverbal cues (facial expressions, vocalizations, gestures, etc.) are no longer available. Addressing such an issue has important implications not only for what concerns the understanding of cognition and psychology, but also for what concerns the design of interaction technology and the analysis of phenomena like cyberbullyism and viral diffusion of content that play an important role in nowadays society.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**;

KEYWORDS

Social Signal Processing, Nonverbal Communication

1 INTRODUCTION

Human-human and human-machine interactions involve nonverbal communication, i.e., the use of nonverbal cues (facial expressions, vocalisations, gestures, etc.) that add layers of meaning to the words being uttered [5]. In particular, nonverbal cues convey information about social and psychological aspects of an interaction, including attitude, stance, emotional states, conflict, social verticality, roles, etc. Typically, psychologists group nonverbal cues into five major classes, typically called the *codes* [6]:

- Physical Appearance: attractiveness, clothes, ornaments, body shape, etc.;
- Face and Head: facial expressions, head nods, head shakes, gaze, etc.;
- Gestures and Posture: orientation of the body with respect to others, self-touching, arm folding, spontaneous hand movements, etc.;
- Vocal: prosody, intonation, vocalizations (laughter, crying, filers), pauses, voice quality, etc.;
- Space and Environment: mutual distances, spatial formations, patterns of movement in public spaces, etc.

An important aspect of the taxonomy above is that it revolves around face-to-face interactions, namely settings in which the people interacting with one another are physically co-located and, hence, can communicate through cues that can be perceived with their senses, in particular sight and hearing. This is not surprising given that nonverbal communication is a natural phenomenon and it is the result of a long evolutionary process during which the only possible interactions where face-to-face [4]. In other words, nonverbal communication developed to be functional in the only communication scenario available and possible in evolutionary times, namely the interaction with physically co-located others.

Nowadays, face-to-face interactions still play a crucial role in everyday life and they are well known to be the basis for a correct development of an individual, especially when it comes to the ability to establish meaningful interactions with others [2], the acquisition of the skills necessary to "read" the mind of others [1], and the very development of those areas in the brain that are known to regulate the behaviour during interactions [8]. However, an increasingly significant fraction of human-human interactions take now place in technology mediated settings, i.e., with the help of technologies aimed at allowing people that are not physically co-located to interact with one another.

One of the main aspects of communication technologies is that they inhibit, partially or totally, the use of nonverbal cues [12]. For example, the phone allows one to use speech, but no cue that can be perceived through the eyes can be adopted (e.g., facial expressions and gestures). Similarly, online textual chats allow one to roughly reproduce the dynamics of a conversation, but none of the natural nonverbal cues listed above can be used (emoticons are an attempt to reproduce them and their success shows how important and necessary nonverbal communication is). Finally, social media do not involve any nonverbal cue, but still allow the use of a wide spectrum of nonverbal messages such as, e.g., images, likes, connections, etc.

The natural question that arises in the scenario outlined so far is whether nonverbal communication in technology mediated settings

is still possible and, if yes, whether it works in the same way as it does in face-to-face settings. In other words, the question is whether human brain - wired for face-to-face interaction - processes the signals exchanged in technology mediated settings in the same way as it processes the social signals that are exchanged between physically co-located interactants. Besides being interesting from a psychological and cognitive sciences point of view, the question has important implications for technology. The reason is that it can provide important indications for the design of new interaction technologies and, furthermore, it can contribute to explain phenomena observed in technology mediated settings such as, e.g., cyberbullyism, virality, etc.

Some early works addressing the questions above show, e.g., that roughly 25% of the time during phone calls is spent for nonverbal communication, thus confirming that people are ready to spend one quarter of their time in nonverbal cues when it is not possible - like in the case of face-to-face interactions - to speak and display nonverbal cues at the same time [9]. Similarly, other works show that the pictures that people like on internet convey a personality impression, not differently from those that people display in face-to-face interactions [7, 13]. Finally, the typing patterns that people manifest in online textual chats are so specific of an individual, that they can be used for person recognition, not differently from speech patterns, facial appearance or fingerprints [3].

The works above - and the others that the literature presents - provide initial indications, but several questions remain to be addressed (the list is not exhaustive):

- Is there a nonverbal component in communications that take place in technology mediated settings that do not allow the use of any natural nonverbal cue (e.g., social media)?
- Is it possible to identify codes like those that have been identified for nonverbal communication in face-to-face settings?
- Is it possible to link nonverbal cues observed in technology mediated settings (if any) to interactional functions and goals (e.g., to convey an attitude or to send a relational message)?
- Is it possible to improve technology mediated communication through the use of appropriate nonverbal cues (if any available in a particular setting)?

Overall, the question is whether Social Signal Processing (SSP) [10, 11] - the computing domain aimed at modelling, analysis and synthesis of nonverbal communication in human-human and human-machine interactions - can be transferred to domains where natural nonverbal cues cannot be used. If this will be possible, there will be two main advantages: The first is that it will be possible to benefit from the entire body of methodology that has been developed in SSP for face-to-face settings. The second, is that the possibility of transferring SSP to new interaction settings can possibly confirm that these latter - from a cognitive and psychological point of view - work like the face-to-face ones.

REFERENCES

[1] S. Baron-Cohen. *Zero degrees of empathy: A new theory of human cruelty.* Penguin UK, 2011.
[2] J. Bowlby. *The making and breaking of affectional bonds.* Routledge, 1979.
[3] M. Cristani, G. Roffo, C. Segalin, L. Bazzani, A. Vinciarelli, and V. Murino. Conversationally-inspired stylometric features for authorship attribution in instant messaging. In *Proceedings of the ACM International Conference on Multimedia*, pages 1121–1124, 2012.
[4] R. Dunbar, L. Barrett, and J. Lycett. *Evolutionary psychology: a beginner's guide.* Oneworld Publications, 2005.
[5] M.L. Knapp and J.A. Hall. *Nonverbal Communication in Human Interaction.* Harcourt Brace College Publishers, 1972.
[6] V.P. Richmond and J.C. McCroskey. *Nonverbal Behaviors in interpersonal relations.* Allyn and Bacon, 1995.
[7] C. Segalin, A. Perina, M. Cristani, and A. Vinciarelli. The pictures we like are our image: continuous mapping of favorite pictures into self-assessed and attributed personality traits. *IEEE Transactions on Affective Computing*, 8(2):268–285, 2017.
[8] S. Turkle. *Reclaiming conversation: The power of talk in a digital age.* Penguin, 2016.
[9] A. Vinciarelli, P. Chatziioannou, and A. Esposito. When the words are not everything: the use of laughter, fillers, back-channel, silence, and overlapping speech in phone calls. *Frontiers in ICT*, 2:4, 2015.
[10] A. Vinciarelli, M. Pantic, and H. Bourlard. Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, 27(12):1743–1759, 2009.
[11] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schroeder. Bridging the gap between social animal and unsocial machine: A survey of social signal processing. *IEEE Transactions on Affective Computing*, 3(1):69–87, 2012.
[12] A. Vinciarelli and A. Pentland. New social signals in a new interaction world: the next frontier for social signal processing. *IEEE Systems, Man, and Cybernetics Magazine*, 1(2):10–17, 2015.
[13] X. Xiong, M. Filippone, and A. Vinciarelli. Looking good with Flickr faves: Gaussian processes for finding difference makers in personality impressions. In *Proceedings of the ACM International Conference on Multimedia*, pages 412–415, 2016.