

## Introduction

This volume presents a selection of the papers presented at MMSYM 2016, the 4th European and 7th Nordic Symposium on Multimodal Communication. The symposium aimed to provide a multidisciplinary forum for researchers from different disciplines who study multimodality in human communication as well as in human-computer interaction. It was organised by the Centre for Language Technology (CST), University of Copenhagen, and took place at the University of Copenhagen on September 29-30, 2016. The invited speaker was Adam Kendon which discussed the relation of language and gesture in communication.

The symposium followed up on a tradition established by the Swedish Symposia on Multimodal Communication held from 1997 till 2000, and continued by the Nordic Symposia on Multimodal Communication held from 2003 to 2012. Since 2013 the symposium has acquired a broader European dimension, with editions held in Malta, Estonia and Ireland. In 2016 the symposium returned to the University of Copenhagen.

As in the preceding editions, we aimed to present a broad view of the field by accepting papers on a wide range of topics and employing different methodologies in order to reflect the largely interdisciplinary nature of this research field. This breadth is also maintained in the choice of papers presented here, which span from modelling and automatic detection of different kinds of communicative movement to the use of non-verbal signals for specific purposes in different communicative situations and different languages. The methods adopted also vary a great deal, encompassing quantitative analyses, experimental and neurocognitive investigations, and semiotic taxonomies. The individual contributions are briefly summarised below, and positioned with respect to this range of topics and methods.

Frid, Anbrazaitis, Svensson-Lundmark, and House present a method for the automatic detection of head movements in audiovisual recordings of read news in Swedish. Head movements are annotated manually, and a machine learning classifier is trained to predict presence or absence of head movement with reference to words based on automatically extracted velocity and acceleration features.

Detection of head movement is also the topic of the paper by Jongejan, Paggio, and Navarretta. In their approach, an SVM classifier learns to classify head movements based on three movement features, i.e. velocity, acceleration, and jerk. The paper also explains that the results of the classifier can be integrated seamlessly with the annotation produced by the ANVIL annotation tool.

Still in the area of automatic recognition, the paper by Stefanov and Beskow describes a method for automatic recognition of isolated Swedish Sign Language signs for the purpose of educational signing-based games. Signs are performed by experienced and inexperienced adult signers and captured with a Kinect sensor. A recognizer based on manual components of sign language is tested on these datasets, with good recognition rates for signer-dependent signing and encouraging ones for signer-independent signing.

A number of papers deal with the use of different non-verbal signals. Among these, the study by Navarretta deals with the relation between audience response and the use of filled pauses and gestural behaviour in humorous speech. The topic is discussed based on statistical analysis of data from two semi-automatically annotated speeches by Barack Obama.

Audience engagement is also investigated in the paper by Curtis, Jones, and Campbell. Their paper looks at correlations between emphasised speech, and increased levels of audience engagement during audio-visual presentations and lectures. In turn, the authors show that emphasised speech corresponds to video sequences characterised by an increase in pitch accompanied by high visual motion.

Jehul, Brône, and Feyaerts study the use of the fillers ‘euh’ and ‘euhm’ in Dutch conversational data, particularly how they co-occur with gaze. Based on this multimodal analysis, they distinguish between a deeper cognitive thinking function, mainly associated with the nasal filler, and a word search function reserved for the pure vocalic filler. Their findings confirm a similar distinction also described in other Germanic languages.

The study by Hoetjes presents an experimental investigation of how informative gestures are when produced in repeated references after negative feedback. The results of the experiment show that, when presented with gestures produced after negative feedback, subjects are more likely to identify correctly the object referred to by the gesture, thus indicating that gestures after negative feedback become more informative.

Hand gestures are also the topic of the paper by Wessel-Tolvig and Paggio. This paper presents the results of two judgment tasks in which the authors investigate how Italian speakers understand complex prepositions in ambiguous motion event constructions. The results show that these constructions are subject to certain semantic constraints, but also that co-speech gestures change the reading of event constructions and thus override default meaning expressed only in speech.

The two papers by Ousterhout both deal with how the semantic processing of emojis can be studied by means of EEG technology. The first of these papers, on BCI effectiveness, establishes the viability of commercial EED equipment to measure the N400 effect, which is known to be an effective way of detecting processing of semantically incongruous signals. The second paper investigates, through survey data and measurements obtained with a commercial EEG, how users understand moving facial emojis occurring in different sentence positions. The results indicate that there is no preference for particular positions of the emojis, and that some of the unusual emojis are ambiguous and are presumably ignored in the processing of the sentence in which they are embedded.

A qualitative methodology is largely followed in the paper by Gerwing, which focuses on the use of gesture, particularly body-oriented gesture, in the area of healthcare. The data for the study come from publically available videos from actual patient-physician encounters. About 150 examples of body-oriented gestures are found and analysed, and it is shown that these gestures have a range of functions related to grounding and cohesion in the interactions.

The paper by Allwood, Ahlsén, Lanzini, and Attaran is also related to the area of healthcare. It compares video health information about overweight and obesity in two different countries – Sweden and Malaysia. Interesting differences both at the level of content and representation styles are described, pointing to possible cultural differences in the rhetorical approach to the topic.

In the study by Vincze and Poggi, videotaped monologues in which doctors and scientists explain their scientific findings are analysed to understand how epistemic stance is expressed through words

and body markers. The authors focus in particular on signals of high certainty and obviousness, which they illustrate and discuss from a semantic and cognitive point of view.

Poggi and Ansani take the reader to the realm of music practice, by analysing gestures of intensity in orchestra and choir conduction. Based on examples from a corpus of concerts and rehearsals, they posit a multimodal lexicon of musical intensity where postures and gestures are categorised along a number of dimensions. These gesture types are shared by different conductors, and form thus a specific lexical area governed by systematic semiotic devices.

A semiotic methodology is also followed by Kreydlin and Khesed, who are interested in how parts of the body are expressed in words and non-verbal expressions. In particular, they discuss how 'hair' is conceptualised and expressed in the Russian language, including the expression of actions involving hair.

In conclusion, this collection of papers provides a snapshot of a field in which research from quite different disciplines can hopefully create synergies and knowledge that can contribute to the further development of multimodal studies, as well as a continuing interest for the MMSYM symposia.

Patrizia Paggio and Costanza Navarretta

Copenhagen, 26 June 2017