

Closed Doors – Modelling Intention in Behavioural Interfaces

Anders Kofod-Petersen
SINTEF ICT
Trondheim, Norway

Rebekah Wegener
Macquarie University
Sydney, Australia

Jörg Cassens
University of Lübeck
Lübeck, Germany

Abstract

Truly smart systems need to interface with the behaviour of human and non human actors in their surroundings. Systems with such interfaces could prove beneficial in supporting those with non standard communication practices, the elderly living alone, people with disabilities, and many others. While the benefits are clear, the means of achieving true behavioural interfaces are yet unclear. In this paper we outline how semiotics helps us to understand behaviour. We show how such an approach may be put to use in modelling the intention to walk through a door. We begin by outlining the semiotic approach and then discuss the behaviours which need to be described to model intention. We also discuss how this varies according to context and suggest the potential for a more general model of behaviour.

1 Introduction

The emergence of research areas such as *ubiquitous computing* [1], *pervasive computing* [2], *ambient intelligence* [3], and most recently *everywhere computing* [4]¹ has given artificial intelligence methods and techniques a renaissance. While the goal of such areas of study is often to build intelligent artefacts that approximate human behaviour, the underlying processes which characterise human behaviour are often ignored. Instead, many ambient intelligence scenarios assume a lot of common-sense reasoning and elaborate problem-solving where the particularities are either ignored or just assumed as a black-box.

To realise the abilities of an ambient intelligent system, three main areas of responsibility can be identified [5]: first, the initial responsibility of perceiving the world that the system inhabits; second, the responsibility of being aware of the environment and reason about ongoing situations, which traditionally has been labelled as *context-awareness*; and third, exhibit appropriate behaviour in ongoing situations by being *context-sensitive* [5, 6].

The ability of being context-aware, and in a broader sense, *situation-aware* [7], is arguably the most important aspect of these systems. Without an assessment of

This paper was presented at NAIS-2009; see <http://events.idi.ntnu.no/nais2009/>.

¹Although all these terms can be viewed as synonyms, a particular term typically indicates a particular perspective, e.g., a physically distributed system perspective vs. a functional-oriented service perspective. The term ambient intelligence will be used consistently throughout this text.

an ongoing situation the ability to assist in implicit or explicit problem-solving by an ambient intelligent system is hardly possible.

As ambient intelligent systems are characterised by being able to perceive their environments, be aware of the presence of people and other agents, interpret their own role in that context, and respond intelligently to one or more agents' needs [8], the intentions of human partners in ambient systems are of utmost importance. However, very often human behaviour is left out of the equation and reduced to a single correlation that is taken to be a causality. For example, with automatic doors, being within a certain defined proximity to a door is taken to imply the intention of going through the door. While this is often a good approximation and will work well in most instances, it does not explain the relationship between the behaviour and the intention. Understanding human behaviour is important for creating intelligent artefacts that are able to understand behaviour across contexts.

Interpretation of human intention is by no means an easy task. Interpretation is based on the reading and understanding of human behaviour and includes gesture and facial expression. However, these are highly complex and meaning here often emerges from an ensemble of a great many different elements. Not only does assessing human intention include many technical challenges with regard to sensing human behaviour, it also poses interesting challenges when modelling. Modelling behaviour of this complexity relies on establishing some crucial boundary conditions and parameters. The work presented here approaches modelling of human behaviour by employing a semiotics perspective, in particular a Systemic Functional Linguistic approach to semiotics.

2 Motivation: modelling space and behaviour as meaningful by using semiotics

Humans can be considered to be social beings who are inclined to interact. As such, most people, regardless of intellectual ability will acquire some form of symbolic-linguistic communication [9, p. 411]. As Bateson [10, p. 244] asserts, all biological systems are capable of adaptive change, which takes many forms according to the size and complexity of the system under consideration. It may be that the prime means of communication in humans, as in other primates, is behaviour, and in the absence of verbal communication, individuals will adapt and find new ways to communicate in a given context, often relying more heavily on behaviour.

In situations where verbal communication is not possible behaviours will often be read as meaning bearing and made to carry a significantly greater portion of the meaning load². Situations where this is more likely include group homes for people with disabilities, communicating with people with a disability who exhibit challenging behaviour, and independent living for the aged, although many other situations meet these criteria. Behaviours such as our daily routines, movements, gestures, facial expressions and posture may all carry crucial information that can help provide a safer living environment and facilitate communication.

Communication is here defined broadly as the making of meaning within a social context. Clearly in contexts where natural language is not an option or is ineffectual, individuals will use other resources to convey meaning. This is particularly the

²Meaning load here refers to the amount of work a modality has to do in a context e.g. the written mode in a book carries the entire meaning load compared with a lecture where the written mode carries a much lighter load.

case for individuals with severe intellectual disabilities and poor language ability. All behaviour has a semantic orientation and although behaviours differ in their consequences and intent, they all carry meaning. As such, they can be studied using approaches such as Systemic Functional Linguistics [11, 12, 13] which, by already including behaviour as a potential meaning making option, can combine behaviour with other modes of communicating such as verbal or signed language, space, architecture and written text to name but a few³.

Semiotic foundations: Systemic Functional Linguistics

Systemic functional linguistics is a social semiotic approach to meaning making that sets out from the assumption that all behaviour is potentially meaning bearing and can be treated as communicative [14]. Taking a functional perspective on language, SFL arranges the resources of language around three broad functions or metafunctions: experiential, interpersonal and textual. We focus here on the interpersonal resources of behaviour because we are interested in how we use behaviour to engage others. It also has a stratified approach to language and thus divides description into layers of abstraction including the expression plane (phonetics), phonology, lexicogrammar, semantics and context. Context in this approach can be defined as what is necessary to understand what is going on and has been treated as a triple of *Field*, *Tenor* and *Mode* [13]. *Field* refers to what is going on, *Tenor* to who is involved and how they are related and *Mode* to the organisation of the interaction [13].

If we think of semantics as meaning potential or “what the person can mean” [14, p.72], then it is possible to see behaviour as semantic since there is a set of behaviours that are at the disposal of the individual within a particular context. While this patterning is specific to the individual and the social context, there is a limit to how truly individual it can be in most social contexts if the intention is to share meaning. After all, to share meaning you must share the code. This means that it should be possible to model the meaning potential available in a particular context.

Because communicating is multimodal, the internal state of an individual's intention will not always be signalled entirely by behaviour in any given social context. Signs occur with other signs and behaviours work together in patterns to create meaning in a specific context in much the same way that the rhetorical device of metonymy⁴ works in a text [13, p. 10]. The task in modelling intention in any situational setting is to find the behaviours or patterns of behaviours which carry the most significant meaning in the context.

Firkins [15] sets out some of the challenges for an analyst treating behaviour as meaning bearing. Firstly, behaviour needs to be seen as part of the interaction that is occurring (a dialogue). Secondly, analysis needs to build on the assumption that the interaction will be multimodal with the relative contributions of linguistic and behavioural forms of expression being context dependent. Thirdly, the boundary conditions for analysis need to be defined, suggesting that the interaction must have some form of cohesion which binds it together. Finally, analysis needs to consider the behaviour as being goal directed and therefore aimed at impacting on other

³Although the theory leaves space for the integration of these modalities this does not imply that the modalities themselves have been theorised or that their integration is without problems.

⁴Calling something not by its own name, but by the name of something associated.

participants involved in the interaction.

While it is difficult to find theoretical accounts which have succeeded in addressing all these concerns, several useful models have emerged which provide a strong basis for developing analysis of behaviour. One such model is established by Martinec [16]. Martinec suggests that behaviour has many of the same functional foundations as language and can thus be treated in much the same way. The value of this is that it can then be mapped into a more complete model of semiosis which includes other modalities providing a common metalanguage or ontology between different modes of communication.

In researching behaviour during surgery, Moore [17] adapted Martinec's [16] model to account for some of the variability that occurs in such settings. He suggests that certain measures of body alignment and proximity, together with visual target, can be taken to realise certain interpersonal meanings, specifically, various levels of engagement. These meanings Moore suggests, are context dependent and the values which signal meanings in one context will not necessarily be the ones which signal the same meanings in another context. However, variation does not mean that there is random variation, on the contrary, the contextual variability is highly predictable and it is this predictability that makes it useful for the current study.

Modelling behaviour in Context

How behaviour creates meaning and how we assign meanings to behavioural acts of expression is significantly related to situation and context. Meaning is constituted in the interaction between the behavioural sign and its function within a context. The context is not simply the environment of use; it contains the factors essential to the interpretation of meaning. Essentially each social context potentially presents opportunities for interactants to use both language and behaviour to create meaning. It is therefore important to see expressive action as part of context and not as the product or effect of context. The behaviour sign only has meaning through its interaction with context. We can only assign meaning to behaviour through its interaction with the context in which it is embedded. If we are to find meaning in behaviour we primarily look to the dynamic relationship between the unfolding interaction and the context. Context goes beyond simply the immediate antecedents of the behaviour or the consequence.

We assign meaning to behaviour through observing closely the *Field*, *Tenor* and *Mode* of the behaviour in interaction. Hasan [18] argues that these elements of context are applicable for all interactions and cannot be ignored by the Interactants.

Schegloff [19] suggests for each enquiry into interaction, we need to consider what represents relevant social context. In other words, what serves as context – whose orientation to context is the consequential and warrantable one for an analysis of meaning? Schegloff also asks a series of important questions relevant to the use of situation as a background to the interpretation of meaning, the connection between the sign and its interpretive situation [20]. Principally, how should we formulate context or setting that will allow:

1. to connect to the theme or social structure,
2. that will do so in a way that takes into account not only the demonstratable orientation of the participants

3. but will also allow us to make a direct ‘procedural’ connection between the context so formulated and what actually happens in the interaction and
4. to eliminate aspects that do not actually inform the production and grasp the details of its conduct [20, p. 112].

According to Schegloff we need to show the characterisation of setting or context in which the interaction is occurring and to demonstrate how this is relevant for the parties. By showing how that aspect is demonstrably relevant to the parties we are able to see how it may therefore be procedurally consequential to the interpretation of meaning. What aspects of context and situation adds to an understanding of how the interaction proceeded in the way it did and came to have the trajectory that it ended up having [20].

3 Modelling intention: automatic sliding doors

When you watch Star Trek you will notice that the doors on the Enterprise are automatic and open whenever somebody wants to go through them. The scene which caught our attention showed the first Officer, William Riker, walking towards the door but turn in response to a question from Captain Picard and then stand and talk within a proximity that would otherwise have triggered the doors to open. Even though the first officer was “near to the doors”, they did not open until after the conversation had ended and Riker showed some intention of going through the door. The crucial point to note here is “intention to go through the door”. The doors do not open and close simply on the basis of proximity as is typically the case. There is an important difference between activating simply because “he is near the door” and activating because “he is near the door and wishing to go through it”.

Intention is something which is dynamic and emergent from interaction rather than a static and predetermined feature of interaction, thus, intention can be considered context sensitive. Because of this, we have not attempted to model intention as a general or context free concept, rather, we have modeled it as dependent on the context. As such, we are modelling intention to walk through a door rather than intention in general. We suggest that the model of intention set out here may be generalisable to contextually similar situations such as waiting for a bus or train or engaging in an interaction or sales encounter.

The test case of automatic sliding doors was chosen specifically because of its rather restricted behavioural set and because the link between behaviour, intention and outcome is much clearer and simpler than in other typical, but more complex situations. The doors either open appropriately or they do not. Despite the possible energy saving benefits of having doors respond to intention, we are in no way suggesting that automatic doors should respond to intention. On the contrary, the volume of people using automatic doors based on proximity daily without much of a problem suggests that using proximity is a good approximation of intention to go through a door. In fact, it might very well be argued that standing in a doorway to have a conversation is a dis-preferred choice and one which doors opening might discourage. This response could even be construed as the physical space interacting with humans to shape their behaviour.

Intention to walk through a door has been chosen as a means of testing the viability of coding behaviour in such a way that it means something for machines as

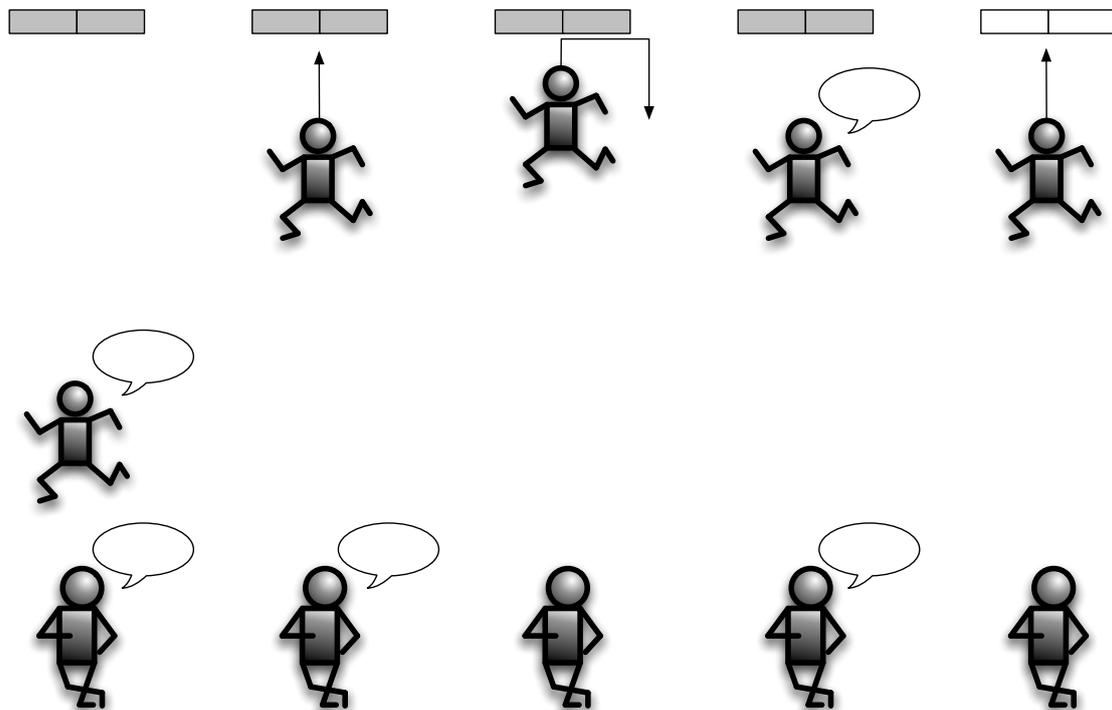


Figure 1: Riker, Picard and a door

well as humans. Compared to other forms of human behaviour the behaviours associated with showing intention to walk through a door are relatively coarse grained and emerge from very few behavioural features which allow us to test the model without having to control too many elements. In the section below we compare the behaviour of people walking through a door with that of people standing near a door and talking or waiting in order to distinguish the behaviours which signal intention to walk through a door.

Drawing on Moore's [17] adaptation of Martinec [16] work, we have focused on the representation of *body alignment*, *proximity* and *visual target*. These three features of human behaviour have been associated with the interpersonal function of communication. That is, they are said to be associated with representing our relation to others. Of the three, *body alignment* and visual target typically vary together, however, particularly when people are moving they will vary independently (as we often see when people look at something else while walking). It is for this reason that they are coded separately so that meanings that might be made through visual target alone are accessible. Because there is an unstable relationship between body angle and visual target, particularly in this context, we have discarded visual target as a useful measure and focused on body alignment and proximity. To this we have added a further measure of dynamism.

Body alignment is made up of a number of different measures and is made relative to the origin of interest – in this case, the doors. The point of origin to which measurements are made relative can of course vary depending on what it is that we are coding for and this is also the case for proximity. When walking through a door a person's body angle is measured relative to the door while those standing talking in front of a door are measured relative to each other and to the door.

Measurements for body alignment used in this study take the angle of shoulder and hips to the point of origin. Shoulders and hips are used because each alone can vary, particularly when people are in motion, however the combined measurement gives a good indication of the average alignment of the person to the origin. When showing intention to walk through a door people have their shoulders and hips aligned approximately square to the door while those engaged in conversation or waiting typically stand with shoulders and hips perpendicular to the door and in the case of conversation, square or obtuse to their conversation partner. One of the options for waiting or talking is to have shoulders and hips square to the door as they would be for walking through the door but to have visual target away from door, so that the back to the door. This is one situation where visual target can be used to discriminate. However, the value of making this discrimination is outweighed by the other situations where it is unhelpful. This also shows why body alignment alone can not display intention to walk through the door since front to door or back to door are not distinguished in such cases⁵.

To distinguish between front to door and back to door we can use the measure of proximity and motion. This measure takes the proximity of the person or object at two points and establishes whether they are static or dynamic and if dynamic in what direction. This allows us to discriminate between people walking away from the door and towards the door and those standing with their back to the door.

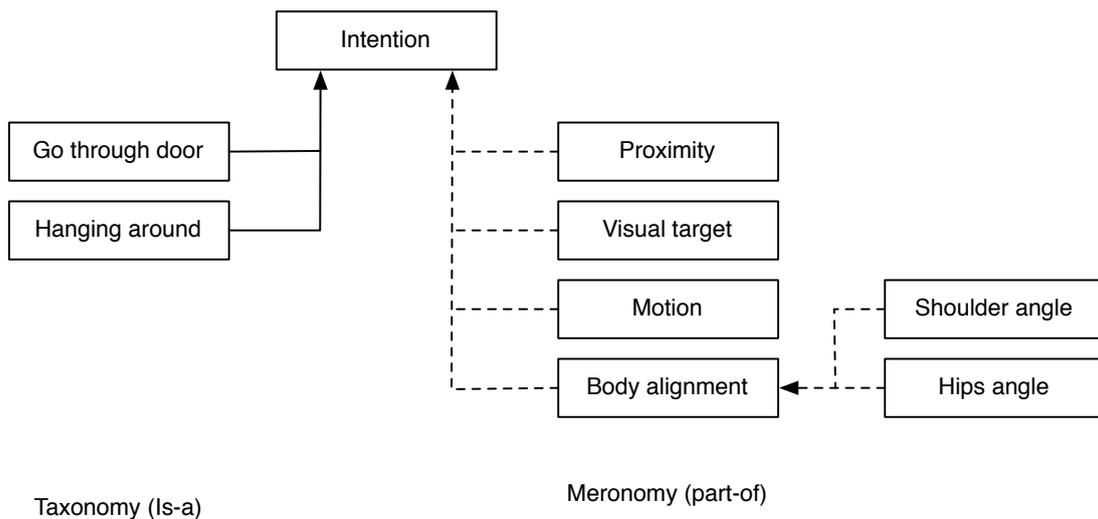


Figure 2: Intentions and behavioural features

While the meaning of “intending to walk through a door” is very simple it is clear that already the features necessary to distinguish between possibilities in this context are quite large although at this point still manageable. It demonstrates that the intention is displayed as emerging from the features “motion towards” and “body alignment square to door (or other object of interest)” but with visual target able to vary and thus not likely to be a reliable measure of intention on its own.

⁵There is actually a distinction between the two in terms of the relative angles of the shoulders and hips given that people walk with slightly angled shoulders and hips and these angles are reversed between front and back, however the difficulty involved in this calculation is greater than using another measure to differentiate.

These signs are all signs that humans are able to read. We are all familiar with watching people's hips and shoulders to avoid collision when walking in the street and in fact this very principle has been the subject of a number of prominent studies such as Goffman's studies of behaviour [21]. Consider also the recent collection by Ingold and Lee Vergunst [22]. These features are also those used by people waiting for a bus to indicate intention to board as opposed to waiting, conversing or walking by. We have used shoulders and hips here since there is some cultural variability in the salience of these in signalling intention.

4 Conclusion and future directions

In this paper we have outlined our underlying communicative approach to modelling behaviour for machine readability. The theoretical approach to semiotics that we have taken allows us to combine all forms of communication within the one approach and to assess how they interact in different contexts. This should prove useful for development of an algorithmic form at a later stage.

Martinec [16] has produced a model of interpersonal engagement that has precise and robust measurements for variability in body alignment and proximity for contexts such as waiting or casually chatting. Moore [17] has shown that the particular values and even the behavioural measures themselves can vary from one social context to another but that this variability is predictable and even useful since it helps us to move beyond context specific descriptions to see what contexts share. In future work we aim to describe the variability in measures for intention to walk through a door and to make these measurements machine readable such that they are useful for opening a door.

The engineering challenges of making these features machine readable are subject of future work, however some of the implications can already be seen. Because what we are attempting to model here is a situation where someone changes his mind and subsequently holds a conversation at a distance that would trigger a normal door to open, lead time required to read intention needs to be longer. This means that the sensor for proximity needs to be set to a greater distance. This also allows more time for the calibration of the different measurements that are required.

When coding of these measures are done by humans very rough approximations are used, particularly if we make these readings while walking on the street or moving in a space. By machine these measurements can be very precise and response triggers would need to be set within ranges. When analysing video material of human behaviour, coding is typically done post production and measurement points are added to the video by hand by the coder. These points are then compared. This same process needs to be done automatically and this requires the ability to identify shoulders and hips on an image. The challenge then is to set response types for different combinations of these features. Beyond the local challenges of coding behaviour is the wider challenge of taking understanding of these behavioural patterns to other contexts. If we are able to use our understanding of body alignments used to signal intention to walk through a door to predict behaviour in other situations such as waiting for a bus or waiting to be served in a queue, then we may have some evidence for the cross contextual generalisability of the approach to understanding behaviour.

References

- [1] Mark Weiser. The computer for the 21st century. *Scientific American*, pages 94–104, September 1991.
- [2] Uwe Hansmann, Lothar Merk, Martin S. Nicklous, and Thomas Stober. *Pervasive Computing: The Mobile World*. Springer Professional Computing, 2003.
- [3] K. Ducatel, M. Bogdanowicz, F. Scapolo, J. Leijten, and J-C. Burgelman. ISTAG scenarios for ambient intelligence in 2010. Technical report, IST Advisory Group, 2001.
- [4] Adam Greenfield. *Everyware: The Dawning Age of Ubiquitous Computing (Voices That Matter)*. New Riders Publishing, 2006.
- [5] Anders Kofod-Petersen and Agnar Aamodt. Contextualised ambient intelligence through case-based reasoning. In T. R. Roth-Berghofer, Mehmet H. Göker, and H. Altay Güvenir, editors, *Proceedings of the Eighth European Conference on Case-Based Reasoning (ECCBR 2006)*, volume 4106 of *Lecture Notes in Computer Science*, pages 211–225, Ölüdeniz, Turkey, September 2006. Springer Verlag.
- [6] Stephen S. Yau, Yu Wang, and Fariaz Karim. Development of situation-aware application software for ubiquitous computing environments. In *29th Annual International Computer Software and Applications Conference (COMPSAC'02)*, volume 1, pages 107–112, Edinburgh, Scotland, July 2002. IEEE Computer Society.
- [7] Mica R. Endsley, Betty Bolté, and Debra G. Jones. *Designing for Situation Awareness: An Approach to User-Centered Design*. Taylor & Francis, 2003.
- [8] E. H. L. Aarts and J. L. Encarnação, editors. *True Visions: The Emergence of Ambient Intelligence*. Springer, 2006.
- [9] T. Deacon. *The Symbolic Species. The Co-evolution of Language and the Human Brain*. Penguin, 1997.
- [10] G. Bateson. A theory of play and fantasy. In *Steps to an Ecology of Mind, Collected Essays in Anthropology, Psychiatry and Epistemology*. Paladin, 1972.
- [11] M. A. K. Halliday. *Explorations in the Function of Language*. Edward Arnold, 1973.
- [12] M. A. K. Halliday and R. Hasan. *Language, Context and Text: Aspects of Language in a socio-semiotic perspective*. Deakin University Press, 1985.
- [13] Michael A.K. Halliday and Ruqaiya Hasan. *Language, Context, and Text: aspects of language in a social-semiotic perspective*. Deakin University Press, Geelong, Australia, 1985.
- [14] M. A. K. Halliday. Text as semantic choice in social contexts. In T. A. van Dijk and J. S. Petofi, editors, *Grammars and Descriptions*. Walter de Gruyter, 1979.

- [15] A. Firkins. The semantics of challenging behavior. In *Australian Group on Severe Communication Impairment Conference*, Adelaide, 2001.
- [16] R. Martinec. Interpersonal resources in action. *Semiotica*, (135):117–145, 2001.
- [17] A. Moore. Surgical teams in action: a contextually sensitive approach to modelling body alignment and interpersonal engagement. In A. Baldry and E. Montagna, editors, *Interdisciplinary Perspectives on Multimodality: Theory and Practice*, Campobasso, Italy, 2008. Palladino.
- [18] R. Hasan. Dynamic view of context in language. In C. Cloran, D. Butt, and G. Williams, editors, *Ways of Saying Ways of Speaking*, Open Linguistic Series. Casell, 1996.
- [19] Emanuel A. Schegloff. Whose text? whose context? *Discourse & Society*, 8(2):165–187, 1997.
- [20] Emanuel A. Schegloff. On talk and its institutional occasions. In P. Drew and J. Heritage, editors, *Talk at Work. Interaction in Institutional Settings*, pages 101–134. Cambridge University Press, 1992.
- [21] Erving Goffman. *The Presentation of Self in Everyday Life*. University of Edinburgh Social Sciences Research Centre. Anchor Books edition 1959, 1956.
- [22] Tim Ingold and Jo Lee Vergunst, editors. *Ways of walking: ethnography and practice on foot*. Ashgate Publishing, 2008.