From action to communication: the role of intention and social cues in the perception of gestures as communicative acts

This project investigates how two domains of human cognition, language and action, are linked by studying the cognitive and neural processes underlying gestures used in everyday talk (i.e., co-speech gestures). Past research has shown that speakers use gestures with communicative intent, and that our brain integrates the information it receives from speech and gesture. The proposed project will take this line of research one crucial step further by exploring when and how we perceive gestures as communicative acts - that is, the continuum from action to communication. The focus of the project is on the cognitive and neural architecture that enables us to distinguish everyday instrumental actions from communicative ones; further, the project explores the role of social contextual cues (i.e., indicators of communicative intention) which accompany gestures in situated language use and which may modulate our perception of the communicativeness of different acts. The project is of direct relevance to important recent proposals claiming that human communication ability evolved out of, and has its roots in, actions. However, there is one important distinction between the domains of language and action. This is that the former is designed for communication, while the latter is not (necessarily). This project investigates the design consequences of the communicative versus non-communicative intentions in these two domains by looking at a domain that shows interesting overlapping features with both. This is the domain of co-speech gestures.

More specifically, the project involves the completion of five different experimental studies using behavioural (reaction times) and neuro-cognitive (functional magnetic resonance imaging [fMRI] and event related potentials [ERP]) methods to explore:

1) the continuum from communicative gesture to instrumental action, as well as how our brain processes those two different forms of manual motor movements when they occur together with spoken language

2) how our brain processes gestures that accompany speech in social context, specifically, context in the form of another visual communicative modality, namely eye gaze.

*Comparing communicative gestures and instrumental actions*

Hand gestures combine with speech to form a single integrated system of meaning during language comprehension. However, it is unknown whether gesture is uniquely integrated with speech or is processed like any other manual action. In order to gain insight into this issue, sentences such as ‘he found the answer’ were accompanied by either an instrumental action (actually typing on a calculator) or by gesture (imitating to be typing on a calculator). Further, two incongruent conditions were created in which the same sentence (‘he found the answer’) was accompanied by either an instrumental action mismatching the sentence (e.g., stirring a drink with a spoon) or by a gesture mismatching the sentence (e.g., imitating to be stirring a drink with a spoon). The findings showed that the incongruent verbal and visual information resulted in more difficult phonemic and semantic processing compared to the congruent case for gestures but not for actions. This suggests that, despite both gestures and instrumental actions being manual motor movements of a very similar kind, our brain processes them differently, with speech and gesture being more strongly integrated during comprehension than speech and action. These findings speak against recent claims that gestures are merely stripped down versions of instrumental actions. Rather, the findings suggest that gestures are special forms of manual action specifically designed for communication.

*Gesture processing in the context of social eye gaze*

The aforementioned research allowed us to conclude that co-speech gestures are a ‘special species’ of manual action, one that is readily integrated with the information contained in speech. Building on this conclusion, the second strand of the project aimed to find out how fixed this integrated processing of speech and gesture is, and to what extent it may be modulated by social context variables. In order to address this issue, a range of behavioural and fMRI studies were carried out, all employing the same basic paradigm: a triadic communicative situation was created involving one speaker as well as two recipients whose task was to listen and make certain judgements about the information communicated by the speaker. Both the speaker and one of the recipients were confederates, while the second recipient was an actual participant whose responses in the tasks were measured. The crucial manipulation in this experimental set-up was the speaker’s eye gaze direction, who looked sometimes and one, and sometimes at the other recipient while she communicated her messages. This rendered our actual participant sometimes an addressed recipient (when being directly gazed at), and sometimes an unaddressed recipient (gaze averted to the respective other recipient). This basic paradigm was combined with a variety of additional manipulations, such as the redundancy between speech and gesture (using both complementary and redundant iconic gestures), the nature of the prime and target items used in the different tasks (visual or verbal focus), the stage of processing tapped into by the respective task (on-line or off-line), and the extent to which speech and gesture were required to be semantically integrated (tasks tapping into the speech and gesture components of utterances separately versus tasks tapping into the combined meaning of the utterance).

The findings from these different studies speak towards a coherent interpretation: First and foremost, the findings show that co-speech gestures are indeed processed differently by recipients of different status, at least when recipient status is signaled through eye gaze direction in a triadic communication setting. Secondly, the findings show that speech processing can also be affected by recipient status – in some of our tasks (using a more visual focus) unaddressed recipients showed impaired processing of uni-modal verbal messages. However, interesting is that, when recipients are presented with the same verbal messages accompanied by gestures (i.e., multi-modal messages), the information contained in the gestures appeared to facilitate unaddressed recipients’ processing of the accompanying speech. Together, the findings from these studies suggest that speaker eye gaze direction does influence how we process speech and gesture, and that, while unaddressed recipients may experience a disadvantage in their speech processing, they tend to process gestures *more* than addressed recipients. One interpretation of this finding is that, because unaddressed recipients are not obliged to engage in mutual gaze, they have more cognitive resources available to process information from another visual modality, that is, gesture. We have termed this the ‘Competing Modalities Hypothesis’. Furthermore, in addition to a benefit in processing the information from gesture, and a disadvantage in processing speech, our fMRI study (which tapped more into the online processing of speech and gesture) revealed that unaddressed recipients are impaired regarding the integration of speech and co-speech gestures, at least in the case of complementary iconic gestures (results from some of our behavioural studies suggest that this might be different in the case of the information in speech and gesture being largely redundant).

The research carried out as part of this project is a first attempt to investigate multi-modal language processing by taking into account three modalities, speech, gesture and eye gaze. As such, it has led to the development of a new research paradigm and a new avenue for future research building on the current findings. Of particular interest in this domain are expansions of the current paradigm to capture more interactive communication, communication in different languages and cultures, communication with children (both with intact as well as with impaired pragmatic communication) in interaction with their caregivers, and especially in educational contexts where teaching staff address a multitude of recipients at the same time, but often to different degrees.

In sum, the findings from our project have provided important insights with respect to our objectives: Firstly, they show that co-speech gestures are, at least to some extent, more communicative than actions and thus ‘special acts’ which could have bridged the evolution from instrumental action to verbal language. Secondly, they provide insight into the influence social context has on gesture comprehension, the neural basis for situated gesture comprehension and the influence of pragmatic knowledge on the semantic integration of gesture and speech. And, thirdly, the findings illuminate the debate on the modularity of the human mind versus cross-modal semantic unification, since our findings clearly demonstrate that language processing is influenced by information from the visual modalities, both semantically and pragmatically, with this influence being modulated by our perception of communicative intentions.