Gesture Signs in Social Interaction: How group size influences gesture communication

Tracy MacLeod

Department of Psychology University of Glasgow

Submitted for the degree of Ph.D to the Highers Degree committee of the Faculty of Information and Mathematical Sciences.

April, 2009

Acknowledgements

Many thanks to Prof Simon Garrod for his supervision, to Dr Jean Carletta for her technical assistance with NXT, to Marc Becirspahic for assistance with programming and to Dr Judith Holler for helpful discussions about gesture coding.

Declaration

I declare that this thesis is my own work and was completed under the usual terms of supervision.

Abstract

This thesis explores the effects of group size on gesture communication. Signs in general change, in the kind of information they convey and the way in which they do so, and changes depend on interactive communication. For instance, speech is like dialogue in smaller groups but like monologue in larger groups. It was predicted that gestures would be influenced in a similar way by group size. In line with predictions, communication in groups of 5 was like dialogue whereas in groups of 8 it was like monologue. This was evident from the types of gesture that occurred with more beat and deictic gestures being produced in groups of 5. Iconic gesture production was comparable across group size but as predicted gestures were more complex in groups of 8. This was also the case for social gestures. Findings fit with dialogue models of communication and in particular the Alignment Model. Also in line with this model, group members aligned on gesture production and form.

Table of Contents

ACKNOWLEDGEMENTS ……………………………………………2

DECLARATION ……………………………………………….....3

ABSTRACT ………………………………………………….....4

EXECUTIVE SUMMARY ……………………………………….8

CHAPTER 1. WHAT, HOW AND WHY GESTURES SIGN?…….13

1.1 INTRODUCTION ……………………………………………….. …..13

1.2 MCNEILL'S CLASSIFICATION SCHEME ……………………………..........13

1.2.1 Iconic gestures and perspective………………………………………........16

1.2.2 Metaphorical gestures and the case of the conduit……………………………19

1.3 WHY GESTURES SIGN……………………………………………...20

1.3.1 Interaction effects in monologue and dialogue………………………………......20

1.3.2 Shared Knowledge…………………………………………………………30

1.4 GESTURE AS AN INDEPENDENT SIGN……………………………………..34

CHAPTER 2: DIALOGUE THEORY ……………………………36

2.1 COMMON GROUND ………………………………………………..36

2.2 INTERACTIVE ALIGNMENT ………………………………………………38

2.2.1 Speech Alignment…………………………………………………………….....40

2.2.2 Graphical Alignment……………………………………………………………40

2.2.3 Body Alignment………………………………………………………………......41

2.2.4 Gesture Alignment……………………………………………………………....43

CHAPTER 3: GROUP RESEARCH……………………………...47

3.1 A GROUP DESIGN FOR GESTURE…………………………………...47

3.2 ALIGNMENT AND SIGN CONVENTION……………………………….52

CHAPTER 4: PILOT…………………………………………...58

4.1 INTRODUCTION……………………………………………………………..58
4.2 METHODS………………………………………………………………….59
4.2.1 Subjects…………………………………………………………………59
4.2.2 Materials……………………………………………………………….59
4.2.3 Procedure………………………………………………………………60
4.3 CODING……………………………………………………………………62
4.3.1 Gesture coding…………………………………………………………62
4.3.2 Speech coding …………………………………………………………66
4.3.3 Coding Tool…………………………………………………………….66
4.4 ANALYSES…………………………………………………………...…68
4.4.1 Gesture and speech patterns…………………………………………68
4.4.2 Alignment and emergence of conventions……………………………69
4.5 RESULTS………………………………………………………………...69
4.6 DISCUSSION……………………………………………………………83

CHAPTER 5: DYAD EXPERIMENT……………………………….90
5.1 INTRODUCTION……………………………………………………………..90
5.2 METHODS………………………………………………………………….92
5.2.1 Subjects…………………………………………………………………*92*
5.2.2 Materials……………………………………………………………….92
5.2.3 Procedure………………………………………………………………92
5.3 CODING……………………………………………………………………94
5.3.1 Gesture coding…………………………………………………………94
5.3.2 Speech coding …………………………………………………………95
5.3.3 Coding Tool…………………………………………………………….95
5.4 ANALYSES……………………………………………………………...96
5.4.1 Gesture and speech patterns…………………………………………*96*
5.4.2 Alignment and emergence of conventions……………………………99
5.5 RESULTS……………………………………………………………….100
5.6 DISCUSSION…………………………………………………………..117

CHAPTER 6: OVERSEER………..……………………………..123
6.1 INTRODUCTION…………………………………………………………123
6.2 METHODS………………………………………………………………124

6.2.1 Subjects……………………………………………………………………………………124
6.2.2 Materials…………………………………………………………………………………124
6.2.3 Procedure……………………………………………………………………………...125
6.3 ANALYSES……………………………………………………………………….126
6.4 RESULTS……………………………………………………………………....…127
6.5 DISCUSSION……………………………………………………………………..127

CHAPTER 7: GROUP EXPERIMENT……………………………..130
7.1 INTRODUCTION…………………………………………………………………130
7.2 METHODS………………………………………………………………………..131
7.2.1 Subjects……………………………………………………………………………….131
7.2.2 Materials……………………………………………………………………………...131
7.2.3 Procedure……………………………………………………………………………...132
7.3 CODING…………………………………………………………………………..133
7.3.1 Gesture coding………………………………………………………………………….133
7.3.2 Speech coding …………………………………………………………………………..134
7.3.3 Coding Tool……………………………………………………………………………...134
7.4 ANALYSES……………………………………………………………………….135
7.4.1 Gesture and speech patterns……………………………………………………………..135
7.4.2 Alignment and emergence of conventions………………………………………………..136
7.5 RESULTS…………………………………………………………………………138
7.6 DISCUSSION……………………………………………………………………..158

CHAPTER 8: GENERAL DISCUSSION…………………………..166

Executive Summary

Spontaneous hand gestures co-express important information with speech. They can enhance the information in the speech or even signify information not in the speech signal (McNeill, 1992). In this way, spontaneous gestures are like other signs such as speech and graphics. Moreover, spontaneous gestures have the other general properties of signs. As they have these properties, we might expect spontaneous gestures to be influenced in a similar way to other signs. Like speech and graphical signs, spontaneous gestures should be shaped by the interactive context and the communicative setting. Research findings show this to be the case as both the content and form of gestures, or what information gestures sign and how they do so, is influenced by being in dialogue and particular aspects of dialogue such as visibility and shared knowledge.

For instance, gestures providing feedback about the communication, by directly referencing addressees, are produced more often in interactive dialogue and moreover, when interactive communication has a visible component. This contrasts with iconic gestures that convey content information as these are produced more often in monologue. In addition, iconic gestures are qualitatively more complex when there is a lack of shared knowledge. Deictic gestures also carry content information but, as they rely on visibility, deictic gesture production drops in monologue. That the interactive setting influences spontaneous gesture production in a similar way to other signs highlights the signifying properties of these gestures and emphasises the need to treat them as independent signs.

Chapter 1 provides a descriptive definition of spontaneous gesture signs from current gesture schemes (McNeill, 1992; Beattie & Shovelton, 2002; Bavelas, Chovil, Lawrie & Wade, 1992; Bavelas, Chovil, Coates and Roe, 1995). In a review of research findings, the chapter goes on to look at the effects of the interactive context and more specifically, whether communication is in monologue or dialogue. The effects of related aspects, such as visibility and shared common ground, are also reported. From these findings, it is apparent that considering whether communication is in monologue or dialogue is important for the descriptive definition of gesture signs. The chapter

concludes by rounding up the reasons for treating spontaneous gestures as independent signs.

Dialogue theories of communication can account for the sign modification reported in chapter 1. Both Clark's (1992) theory of common ground and Pickering & Garrod's (2004) alignment model propose that sign modification occurs as representations are shared or aligned over the course of an interaction. By both accounts, for this mutual representational state to occur, the communication must allow for shared knowledge and a constant on-line checking of the shared knowledge state. It therefore relies on interactive communication, or being in dialogue, and importantly on interactive feedback. As information is held in a representational form (ie through shared knowledge of the prior interaction), it does not need to be carried in the sign and this leads to reduced expressions. This contrasts with less interactive communication, like monologue, where there is no such shared representational state and therefore more information needs to be carried in the sign. This account explains the findings in chapter 1, whereby iconic gestures carrying content information are both produced at a higher rate and are more complex in monologue as opposed to dialogue where there is a reduction in overall production and complexity.

The models can also explain why gestures providing feedback should occur more often in dialogue than in monologue. As interactive communication relies on interactive feedback, these gestures should be produced more often in interactive dialogue. Interactive feedback can occur in dialogue because when in the dialogue situation, there is the ability to provide and monitor feedback, be it verbal or gestural feedback, whereas in monologue this is difficult. If gestures can be monitored better in dialogue, those relying heavily on visibility, which is the mode of monitoring in gesture, should be produced more often in dialogue. Monitoring ability can therefore also account for the reduced production of deictic gestures when in monologue. In addition, deictic gestures and some social gestures function as a reduced expression and this is another reason why they should be produced more often in dialogue.

Chapter 2 reviews Clark's (1992) theory of common ground alongside Pickering & Garrod's (2004) alignment model. As well as providing a theoretical framework for the findings reported in chapter 1, these models of communication propose gestures be treated as independent signs. There are however important distinctions between the grounding and alignment processes. One distinction is that, when alignment occurs, as well as changing in the ways described above to provide more or less information and different kinds of information, signs are said to converge to become more similar between those communicating. The final sections of chapter 2 therefore review findings specific to alignment in speech, graphics, body and finally gesture.

A further distinction is that whereas common ground relies on complex partner-modelling, alignment most often occurs through implicit grounding and so, can occur in situations other than two-person dialogue, like in group communication. The group setting therefore affords this prediction of the alignment model to be tested. As there is a need for alternatives to one-person monologue and two-person dialogue, group communication is also a useful paradigm for the study of communication in general and more specifically, gesture communication.

Chapter 3 discusses the benefits of a group design for gesture and reviews group findings in speech that mirror those in one-person monologue and two-person dialogue. For instance, group research by Fay (2000) shows groups of different size elicit different speech styles, with smaller groups eliciting a dialogue style and larger groups a monologue style. Although these findings sit well with Clark's theory of common ground (1996), and indeed Fay (2000) interprets them within this framework, they sit particularly well with the alignment model and it's proposal that grounding can occur outside of two-person dialogue such as in group communication.

Gesture research has so far empirically tested groups with a maximum of three people (Ozyurek, 2002). Here, although increasing size from dyads to triads did not influence gesture production, a shift in seating, and so change in the shared gesture space did influence what was produced in the gesture. Applying the group design to gesture is then a novel approach to the study of gesture communication.

Chapter 3 goes on to address the assumption that it is through a process of alignment, or convergence, that signs become conventionalised. The alignment model assumes communication, and communicative sign, evolves within the dialogue context. This assumes the cognitive processes behind sign production to be a product of social cognition. It is then through their use that they become identifiable as usable signs. Indeed, like the studies of two-person communication reviewed in chapter 2, group research finds signs converge, or align, over the course of an interaction. However, depending on the group dynamics, signs can also become conventionalised over the course of an interaction (Garrod & Doherty, 1994). Chapter 3 goes on to discuss the relationship between convergence and conventionalisation and how this relates to gesture sign.

On the basis of the findings reported and predictions of the dialogue models of communication, in particular the alignment model, a group experiment was designed to explore the effects of group size on gesture communication. To test the group design and explore the gestures that emerged from it, this was first piloted in chapter 4. Here groups of 5 and 8 collaborated to find the correct order of a cartoon story. Problems in the pilot, with the experimental design and coding, led to a modified group design and suggested coding scheme.

To test the new group design and get a baseline coding scheme for interactive communication, the design was modified for pairs of participants who communicated in dialogue. This dyad experiment is reported in chapter 5. An issue with the new group design was that members in the different sized groups would have different amounts of information. Since varying the amount of information had no effect on the interaction, the design could be applied to group communication. The coding scheme that emerged could also be applied to groups, as it was applicable to interactive communication as well as monologue styles. In addition to testing the design and coding scheme, the dyad experiment found alignment, or convergence, between the communicating pairs on gesture rates and amounts. Alignment also occurred on the gesture form. This finding of alignment in communicating pairs provides support for the alignment model.

Chapter 6 investigated the issue of alignment further by testing whether gestures from the same pairs were more similar in form, or aligned, than those from different pairs. To do this, overseers were presented with three gestures, two from the same pair and one form a different pair. Their task was to choose an odd gesture out and they more often chose gestures from different pairs indicating gestures from the same communication were more aligned on form. This experimental chapter therefore more fully addressed the issue of alignment in interactive communication and discussed the findings in relation to the conventionalisation of gesture.

The group experiment is reported in chapter 7. As in the pilot, groups of 5 (G5) and 8 (G8) collaborated to find the correct order of a cartoon story. Groups of 5 were more collaborative than groups of 8 and this was evident from the interaction times alongside the speech and gesture patterns that emerged. Across the groups, changes occurred in the gesture sign. Groups of 5 had more interactive beat gestures and more deictic gestures whereas the production of iconic gestures and social gestures were comparable across group size. There were however qualitative differences in iconic and social gestures. For instance, the information these gestures carried was more complex when they were produced in larger groups and reduced when in the smaller groups. In addition, social gestures involved in turn taking were more explicit in larger groups indicating a less fluid turn taking process. All of these findings fit with dialogue theories of communication. Alignment on gesture rate, amount and form was also evident in the groups. That alignment occurred in the groups and moreover in G8 as well as G5 can be explained by the Alignment model in terms of how automatic the alignment process is.

Chapter 8 gives an overall discussion of the thesis, the main aim of which was to explore the effect of the interactive setting on gesture signs and relate this to signs more generally.

Chapter 1: What, how and why gestures sign?

1.1 Introduction

Spontaneous hand gestures communicate important information both alongside speech and in the absence of speech. Spontaneous gestures are therefore independent signs that can stand alone in any communication. Yet, they are also part of a wider sign system where they are co-ordinated with other signifying movements. Most notably, gesture is co-ordinated with speech but both speech and gesture are co-ordinated with other meaningful movements. They are co-ordinated with movements of the head as in nodding, and of the eyes, as in gaze for example. Spontaneous gestures are then one of a range of kinetic movements that communicate meaningful information. They are a part of what Clark (1996) has called the composite signal, a signal made up of all, or at least some, of the aforementioned signs. Since gesture is an important sign, in terms of the amount and kind of information it conveys, it is vital to consider gesture when investigating how we communicate. Moreover, because gestures emerge spontaneously in the dialogue context, rather than through formal teaching, gesture research could enlighten the field of communication studies, with regards to the development of signs, more generally. This first chapter explores *what how* and *why* gestures sign in the way they do.

1.2 McNeill's classification scheme

McNeill's (1992) classification scheme of spontaneous gesture is widely used. The scheme classifies spontaneous gestures by considering how gesture relates to both the object of representation and accompanying speech. This provides a descriptive definition of the different ways gestures sign. Spontaneous gestures are first classed as imagistic, if they conjure up an image or picture in mind, or as non-imagistic, if they do not. Imagistic gestures are classed as iconic, when they refer to a concrete entity and '… bear a close formal relationship to the semantic content of speech' (McNeill, 1992; p12). McNeill (1992; p 12) gives as example an iconic gesture, which occurs at the emphasised point in brackets, where the gesture refers to the same act as the speech;

(1.1) and he [bends it way back]

Iconic: hand appears to grip something and pull it from the upper front space back and down near to he shoulder.

When they convey abstract concepts, imagistic gestures are said to be metaphorical. The difference between an iconic and metaphorical gesture concern; (1) how the sign relates to the object of representation and (2) how the sign relates to the speech. Whereas iconic gesture signs map directly onto both the object of representation and the speech, metaphorical gestures convey more distant concepts that are not so directly mapped. In dealing with abstract content, metaphorical gestures can convey information about the thoughts of the gesturer as well as information about the topic of conversation. Gestures reflecting the thinking behind the discourse are treated as distinct from other types of metaphorical gesture and called conduit metaphors. To distinguish between these different types of metaphor, McNeill (1992; p158 & p14) gives an example of a metaphoric (see example 6.15) and a conduit gesture (see example 1.2);

(6.15) … she's flying into the area [wondering] why all the animals are running away

Metaphoric: hands radiate away from the head for beams of "wondering why".

(1.2) it [was a Sylves]ter and Tweety cartoon

Conduit: hands rise up and offer listener an "object".

Non-imagistic gestures are classed as deictic, as in the case of pointing, if they index and locate meaning by referencing the gesture space. They convey information relating to the discourse content and in this way are like iconic gestures. Although deictic gestures can be concrete and refer to an entity present in the gesture space, they are most often abstract and refer to an entity not actually present in the gesture space (McNeill, 1992; p18) gives as example;

(1.8) [where did you] come from before?

Deictic: points to space between self and interlocutor

However beats, described as small and repetitive biphasic movements that manage the grammatical structure, are also classed as non-imagistic. Beats convey information to do with the structure of the communication because they deal with organisational aspects of the discourse. In doing so, beats are the non-imagistic counterpart to the conduit metaphor. For example, beats are said to organise the discourse by conveying information about the gesturer's thinking. They can, for instance, mark new as opposed to given information and in doing so convey assumptions about what is already known. Or, they can more generally convey the thinking behind the discourse models of those communicating. Here, for McNeill (1992), as the object of representation and accompanying speech is indirectly mapped onto the sign, beat gestures convey information abstractly. In this way, beats are again like the conduit metaphor. McNeill's (1992; p16) example of a beat gesture is given below;

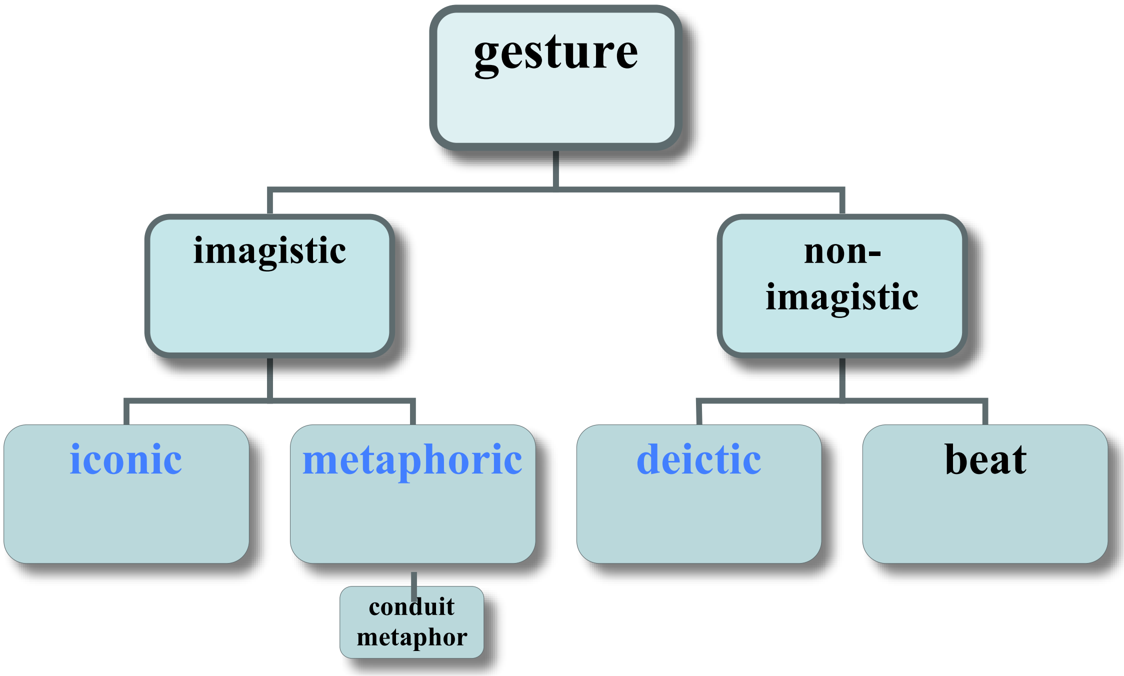
(1.4) when [ever she] looks at him he tries to make monkey noises
Beat: hand rises short way up from lap and drops back down.

To summarise, McNeill's (1992) scheme addresses ***what*** kind of information the gesture sign conveys by considering whether the information is to do with the content or structure. It also addresses ***how*** signs convey this information by considering whether the gesture sign is imagistic, alongside whether it is abstract or concrete. By addressing what and how gestures sign, McNeill (1992) explores the relationship between the object of signification and the sign. This shows the way in which gesture's sign depends on the relationship between the object and sign as well as on what the object (of signification) is. Since McNeill (1992) also uses the accompanying speech to define spontaneous gestures, the relationship between the gesture sign and speech sign is also a consideration.

McNeill's (1985, 1992) paradigm typically has one person relaying a movie cartoon to another who has not seen the cartoon. From the ensuing conversation, spontaneous gestures are observed and described and from this are seen to communicate important information. It is through such descriptive analyses, that the coding scheme described above has emerged. With gestures coded by the categories in the scheme, the scheme provides a description of the gesture sign properties, properties that in turn determine what is produced in the sign. Although gestures are categorised by these existing properties, they are considered to lie on a dimension. Categories are therefore not distinct and the properties of any gesture will be more or less like a given category.

The placement of any gesture on this dimension is determined by a process of best fit. The gesture definitions forming the basis of McNeill's (1992) scheme are represented in a Mind Map (see figure 1.1). Depicted in this way, the connectors of the Mind Map represent the thinking process of the coder and the boxes the coder's gesture category decision. Gesture types coloured in blue convey content information whereas those in black convey structural information. Particular aspects influencing the production of these different types of gesture are considered further in the following sections.

Figure 1.1: Mind Map of McNeill's (1992) gesture categories.



1.2.1 Iconic gestures and perspective

Iconic gesture signs are influenced by perspective. Hence, the same imagistic information, referring to the same concrete entity or concept, can be conveyed differently depending on the perspective taken. The gesturer can adopt either a character or observer viewpoint (Cassell & McNeill, 1991; McNeill, 1992). For example, to convey a running character through an observer viewpoint the gesturer might outline the route of the runner by pointing with the forefinger along a trajectory. When a gesture is from an observer viewpoint, it is as though the gesturer is a spectator of the story looking on from outside. In the case of a character viewpoint gesture, the gesturer might move their own arms as though they themselves were running. Through a character viewpoint perspective, the gesturer and so the gesturer's body becomes one of the characters inside the story. Distinct viewpoints are apparent from the different ways iconic gestures sign, with respect to how the gesturer incorporates their own body when gesturing and the size of gesture. Since the gesturer uses their own body as a frame of reference, either wholly or in part, character viewpoint gestures make reference to specific body parts and they are often bigger in size because of this.

It is important to take into account the perspective of iconic gestures in order to understand the communicative aspects of these gestures. From descriptive analysis, Cassell & McNeill (1991) observe character viewpoint (Cvpt) gestures to occur on events central to the narrative and with simple single clause sentences that create closeness to the action whereas they observe observer viewpoint (Ovpt) gestures to occur with peripheral events and with complex sentences that create distance from the action.

The difference between Cvpt and Ovpt gestures is similar to differences in the frame of reference for speech. Whereas the relative frame of reference in speech is an egocentric perspective, the intrinsic frame of reference is object centred (see Coventry & Garrod, 2004 for a review). Cvpt gestures are therefore like the relative frame and Ovpt gestures like the intrinsic frame of reference.

Differences in the communicative strength of Cvpt and Ovpt gestures have also been found. Beattie & Shovelton (2001; 2002) found character viewpoint gestures communicated more information than observer viewpoint gestures. They modified McNeill's (1985; 1992) cartoon paradigm to elicit spontaneous gesture and then used these gestures in further experiments. In the first stage of the experiment, participants narrated three different cartoon stories to the experimenter. This generated material for the main experimental conditions. Gestures elicited in the narrations were isolated and played to another group of participants. They were played in; an audio/visual presentation, an audio only presentation (Beattie & Shovelton, 2001) and a visual only presentation (Beattie & Shovelton, 2002). Beattie & Shovelton (2002) found Cvpt gestures conveyed more information about size and relative position than Ovpt gestures. They also found gestures conveyed a substantial amount of information by themselves, in the absence of speech. Interestingly, Cvpt gestures again communicated more when there was no speech. In addition to showing an effect of perspective, on both the form and function of the gesture sign, these findings indicate spontaneous gestures communicate essential information and that they do so as independent signs.

Beattie & Shovelton's (2001; 2002) methodology differed in several ways from the descriptive analyses of McNeill (1985; 1992). First of all, Beattie & Shovelton's (2001; 2002) cartoon narration was manipulated in an experimental set up so that variables could be controlled and cause and effect outcomes reported. Their gesture analyses also differed. Rather than linking gestures to the accompanying speech, as in McNeill's (1985; 1992) analyses, Beattie & Shovelton (2001; 2002) devised and used a semantic-features approach. This method, where particular features of interest are analysed like say how size is conveyed, links the gesture directly to a reference in the comic. Directly linking the gesture to a referent in this way makes the task a referential communication task. The definition of the gesture sign is then based wholly on the object the gesture refers to (the object of representation) and relations with this object. Like McNeill (1985; 1992), this again addresses the question of what and how gestures communicate but, in having different object-sign relations, it does so by treating gesture as an independent sign. The accompanying speech and so what the speech refers to, does not define the gesture in any way. The reasons for defining gestures independently from speech will be discussed in the final section of this chapter. For now, it is important to note that the two methodologies differ substantially in terms of the level of manipulation in the experimental design and interpretation of gesture.

1.2.2 Metaphorical gestures and the case of the conduit

As mentioned, metaphorical gestures signify differently depending on the metaphor. As distinctions between different types of metaphorical gestures are complex, they are described more fully here. McNeill (1992) describes the various ways gesture can convey abstract content, to do with space and movement for instance, and considers these all to be metaphorical. An interesting example given by McNeill (1992) is the concept of a mental state (see example 6.15), which he notes is often conveyed as an energy force in the form of an aura or beam that emanates from the body. McNeill (1992) compares metaphorical gestures to metaphorical speech, since both convey abstract content as though it were concrete. Metaphorical gestures do this by conveying information as though it were a physical substance that is localised, bound and contained (McNeill, 1992).

The conduit metaphor (see example 1.2) is a particular kind of metaphorical gesture categorised by McNeill (1992). Like other metaphorical gestures, abstract information is again bound in the gesture to give it concrete substance. It is again passed over only here it is passed on to a recipient over a conduit (McNeill, 1992). Also like other metaphorical gestures, conduit metaphors function to convey information about the topic of conversation. McNeil (1992, p147) gives as example a passing across gesture in the context of getting an idea across. Alternatively, the ‘idea’ could be held up, as though bound in a cup (McNeil, 1992, p147). For McNeill (1992), these gestures depict the imagery of the idea more, when the idea is bound and held up, or less, when the idea is passed over a conduit.

Conduit gestures can also make reference to the ‘… interpersonal context of the narrative situation …’ (McNeill, 1992; p147). Such gestures are said to function pragmatically by conveying information about what the gesturer is thinking about the on-going discourse. In the case of getting an idea across then, the idea would refer to what the gesturer thinks about the on-going discourse rather than conveying an idea from the topic of discussion. These conduit gestures therefore convey information about interpersonal aspects of the communication. An example from McNeill (1992) is;

(6.3) I have [a question]

Metaphoric: hand forms a cup for the image of a question (a container) or the hand out to receive an answer (substance).

The cup shape of the hand is said to be a container that either holds the gesturer's question or is expectantly waiting to receive an answer from the addressee (McNeill, 1992). Here, as in the example above, since the question is held in the hand as if it were a physical object bound by the laws of physicality, the gesture directly depicts the imagery of the question.

The conduit metaphor has additional forms and functions but these become apparent only when the interactive context is taken into account. The next section describes findings from monologue and dialogue and how they are of particular importance for the conduit metaphor.

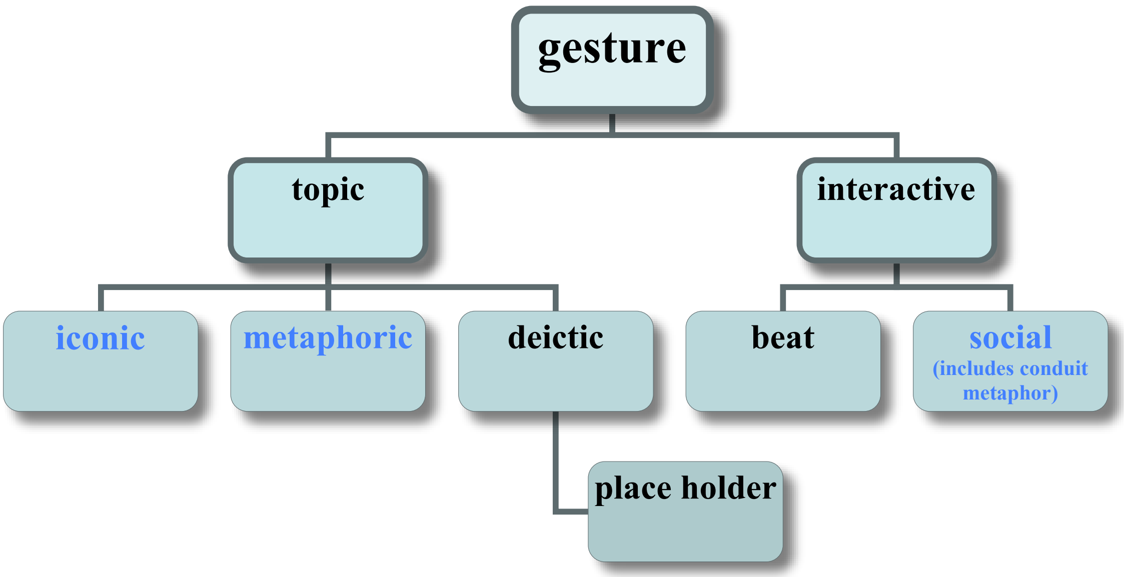
1.3 Why gestures sign?

1.3.1 Interaction effects in monologue and dialogue

Descriptive analyses like McNeill's (1992) tend to focus on monologue styles of communication and, due to their prevalence in monologue, on gestures conveying content information. However, studies looking at dialogue alongside monologue show monologue is not representative of dialogue. It is therefore crucial to consider the interactive setting as a factor in exploring the production of signs. Indeed, being in dialogue influences gesture production in interestingly different ways. In particular, being in dialogue highlights the prevalence and variety of gestures that convey information about aspects of the interaction, which are akin to McNeill's (1992) beats and conduit gestures.

Bavelas, Chovil, Lawrie & Wade (1992) had participants take part in two tasks; one where participants gave a set of instructions and another where participants watched and re-told a cartoon narrative. The tasks were completed in either a dialogue or monologue context so the level of interaction could be treated as a variable. When in dialogue, the conversation was free flowing and either interlocutor could contribute at any time. As both of the participant's gestures were analysed in the dialogue condition and one person's analysed in the monologue condition, analyses and gesture distinctions were based on these two different styles of communication rather than being based solely on monologue. This led to two distinct gesture categories, what Bavelas et al (1992) called 'interactive' gestures and what they called 'topic' gestures. The coding scheme of Bavelas et al (1992) is illustrated a Mind Map (see figure 1.2). These categories crossed McNeill's (1992) descriptive definitions as interactive gestures consisted of beat and conduit gestures whereas topic gestures consisted of iconic and all other metaphorical gestures. By these definitions then, Bavelas et al (1992) more clearly treated beat and social gestures as having an interactive or social function.

Figure 1.2: Mind Map of Bavelas et al's (1992) gesture categories.



Bavelas et al (1992) found the interaction influenced both the rate of interactive gestures, which were higher in the dialogue condition, and rate of topic gestures, which were higher in the monologue condition. This showed how information was conveyed in the sign differed depending on the interaction (Bavelas et al, 1992). More specifically, information was packaged differently depending on whether the interaction was in monologue or dialogue. As well as showing the interaction to have an effect on gesture production, qualitative analyses highlighted the very different functions the gestures performed. Whereas topic gestures carried meaningful information about the topic, interactive gestures were found to be organisational in their function. As interactive gestures were mostly oriented towards interlocutors, this led Bavelas et al (1992) to conclude that instead of relaying previously mentioned content information (in say a topic gesture), interactive gestures cited what had already been said by oneself or, by the other person.

Supporting the idea that interactive gestures made reference to the other person, and moreover for the other person to see, a second experiment found visibility influenced the rate of interactive gestures. Here, Bavelas et al (1992) manipulated the visibility of interlocutors whilst they spoke of a personal incident so as to compare a face-to-face communication with a non-visible condition. They found interactive gesture rates to be lower in their non-visible than in their face-to-face condition, whereas topic rates were comparable. On the basis of the combined findings, Bavelas et al (1992) compare the function of interactive gestures with the function of other 'topic free acts' such as speaker sociocentric sequences and listener back channel responses (1992; p486). They claim interactive gestures, by citing information, mark what those communicating hold to be common knowledge and, in doing so, co-ordinate the dialogue by providing feedback about the communication (Bavelas et al, 1992). In conclusion, interactive gestures are said to be social gestures influenced by social variables and that their presence indicates the conversation is in a 'social system' (Bavelas et al, 1992).

Studies addressing the issue of face-to-face communication v's non-face-to-face communication have however had mixed results. For instance, two separate studies by Cohen & Harrison (1973) and Cohen (1977) found what they called illustrator gestures (which are akin to topic gestures) to be produced at a higher rate in face to face communication than in non-face to face communication. Also, a study by Alibali, Heath & Meyer (2001) found comparable rates of beat gestures across a visible and non-visible condition. Such differences in findings can, as Alibali et al (2001) point out, be attributed to differences in the classification of gesture. For instance, whereas Bavelas et al (1992) classify gesture on the basis of both the gesture function and form, other studies classify gesture typologically. Such differences in findings highlight how important it is to consider the gesture classification scheme and moreover, Bavelas et al's (1992) findings suggest a consideration of function as well as form is necessary.

A later study by Bavelas, Chovil, Coates and Roe (1995) found interactive gestures to have additional forms and functions, all of which helped co-ordinate the communication. In a within design, dyads narrated a cartoon story together so as to be in dialogue. They then narrated the story in alternating monologues, so as to be in sequential monologue. Interactive gesture rates were higher in alternating dialogues from the same pair than in alternating monologues. As interactive gesture rates were dependant on the level of interaction, with more being produced in dialogue than monologue, this effect of dialogue was in line with Bavelas et al (1992). Descriptive analyses were also in line with Bavelas et al (1992) as interactive gestures were found to reference the other person in the communication. Again, interactive gestures did this by citing information. A typical way for these gestures to cite information was to acknowledge an understanding of what was said. For example, in a quick flick of the fingers directed towards an addressee to indicate engagement. As in Bavelas et al (1992), such gestures provided feedback about the communication.

As well as providing feedback by citing information, interactive gestures were involved in the provision of feedback in a number of other ways. They could seek a response such as understanding. When doing this, they were usually in the form of a conduit gesture, with the hand held out as though waiting for information. These gestures were also often presented with the hand moving in a circular motion. This movement is described as being like the motion of a conveyer belt or water wheel (Bavelas et al, 1995). Bavelas et al (1995) suggest this is a gestural equivalent of Bernstein's (1962) verbal sociocentric sequences, such as 'you know'.

Another group of interactive gestures provided feedback about and organised the turn exchange. These gestures indicated when speakers wished to hold on to the floor or give it up. They did so in a pulling motion, as though pulling the turn towards the person speaking/gesturing, or in an analogous motion, as though pushing the turn away.

Interactive gestures also marked speaker assumptions. For instance, new information was commonly marked by McNeill's (1992) conduit gesture where information was metaphorically handed over in the palm of the hand (Baveals et al, 1995). Shared information was often indicated in a swift hand movement where the palm was held upwards and rotated towards the addressee (Bavelas et al, 1995). Interactive gestures also marked digressing and elliptical information through hand flicks that were often made to the side. This fits with McNeill's (1992) observation that metaphorical and beat gestures are often made in lower central or peripheral regions (p90-91).

In line with interactive gestures referencing addressees, Bavelas et al (1995) found interactive gestures elicited appropriate responses from the addressee. For example, where a gesture questioned whether or not an addressee understood and/or followed the speaker/gesturer, a typical response was a spoken 'Mhm' or 'Yeah'. The typical response to gestures seeking help was to provide the help requested. In the case of citing given information, addresses (who had often first mentioned the information) were not required to respond, however a response was usually given. Despite receiving a response, interactive gestures co-ordinated the conversational turns with minimal disruption to the accompanying speech. Bavelas et al (1995) suggest these kinds of gesture allow the communication to be organized with minimal disruption to the accompanying speech.

Overall, Bavelas et al's (1992; 1995) findings show how the level of interaction influences gesture sign production and in doing so, highlight just how important it is to consider the interaction in any exploration of gesture. It is only by considering the structural and functional aspects of gestures within the context of the interaction that the complexity of their form and function becomes apparent. In particular then, findings highlight the importance of the dialogue setting for the production of interactive beat and conduit metaphors, and show how looking solely at monologue underestimates the complex form and function of these gestures.

The observation that interactive gestures are involved in interpersonal aspects of conversation management is in accord with McNeill's (1992) description of beat and conduit gestures. However, from the orientation of interactive gestures (towards an addressee) and the accompanying speech (often 'you'), Bavelas et al (1995) propose interactive gestures directly reference the other person. In directly referencing an addressee, both beat and conduit gestures map directly on to their object of representation. This differs from McNeill's (1992) description of beat and conduit gestures, which are said to relate indirectly and so, abstractly to the management of the communication. The findings show then how being in monologue or dialogue can influence the function assigned to a gesture and how coding schemes with different definitions emerge.

In Bavelas et al (1992; 1995), the effects of dialogue and visibility were mainly for interactive gestures therefore findings were concerned with these gestures. However, in both studies topic gesture rates were lower in dialogue than in monologue and this is also an important finding. Bavelas, Gerwing, Sutton & Prevost (2007) note that these earlier studies did not separate out dialogue from visibility and that it may be important to do so. To explore the effects of monologue and dialogue, whilst treating visibility as a separate factor, Bavelas et al (2007) separated these factors out within the same experimental design. In doing so, this experiment addresses the conflicting findings for visible v's non-visible conditions reviewed earlier for both illustrator (Cohen & Harrison, 1973; Cohen, 1977) and beat gestures (Alibali et al, 2001).

To separate out visibility from dialogue, Bavelas et al (2007) constructed three conditions; a monologue condition, where communication was directed towards a tape recorder, a face-to-face dialogue condition, where interlocutors were both co-present and visible and a telephone dialogue condition, where interlocutors were neither co-present nor visible. The experimental task was to describe a picture of a fancy 18^{th} century dress worn by a lady of the time.

Based on Bavelas et al's (1992) distinctions, 'topic' gestures, conveying information about the discourse content, and 'interactive' gestures, conveying information about the discourse structure, were coded for all participants in all three conditions. In line with and expanding on earlier findings, interactive gestures were influenced by visibility, as rates were higher when they could be seen, in face-to-face dialogue, than when they could not be seen, in telephone dialogue and in monologue. For topic gestures then, there was an effect of dialogue but it was in the opposite direction to prior findings. Here, in face-to-face and telephone dialogues topic gesture rates were comparable but were higher in comparison to being in monologue.

As well as influencing the overall rate of interactive gestures, visibility influenced the more qualitative aspects of topic gestures as when visible, topic gestures were more often bigger, deictic and with a Cvpt perspective. Bavelas et al (2007) suggest visibility is important for Cvpt gestures because addressees must see the relationship between the speaker's body parts and that of the character's body parts they are adopting. The necessity to see Cvpt gestures lends support to Beattie & Shovelton's (2002) finding that Cvpt gestures are more communicative since the more communicative a gesture, the more need there will be to see that gesture. In addition, gestures increased in redundancy (that is they didn't add anything over and above the accompanying speech) across the three conditions showing an effect of both visibility and dialogue.

Overall, Bavelas et al (1992; 1995; 2007) found particular aspects of the interaction, namely visibility and dialogue, to influence how information was conveyed in gesture signs but that these factors influenced gestures differently. Interactive gestures, found to co-ordinate the communication by directly referencing addressees, were produced at a higher rate when in dialogue. Moreover, their rate of production was to do with the visible component of dialogue. The main way interactive gestures co-ordinate the communication is by providing feedback about it. Since feedback in the gesture modality relies heavily on visual processing, visibility should be crucial for feedback to be monitored. But all gestures rely on the visual channel therefore visibility should be crucial for all gestures. This was the case as the more qualitative aspects of topic gestures were also influenced by visibility (Bavelas et al, 1992; 1995; 2007). The drop in the overall rate of interactive gestures however, indicates visibility is of particular importance to these gestures.

Interactive gestures also occurred less frequently with speech than topic gestures and were less redundant when they did occur with speech (Bavelas et al, 1992). In light of their independence from speech, Bavelas et al (1992) suggest interactive gestures may be devised to function in the absence of speech as by being non-disruptive, they are particularly useful for co-ordinating the interaction. Being less dependent on speech, interactive gestures should rely heavily and at times solely on the visual channel. This again distinguishes them from topic gestures that more often rely on combined auditory and visual channels. As was the case then, effects of visibility ought to be more dramatic for interactive gestures, which depend on the visual channel, than for topic gestures, which utilise both modes. In order to receive and monitor feedback, it is vital for the other person to see an interactive gesture as the accompanying speech, if indeed there is any, provides little or no information. Similarly, when there is no other person to see the gesture, as in a monologue, and there is no opportunity to provide or receive feedback, the rate of gesturing would be expected to fall, as it does.

The importance of monitoring feedback for the success of the communication will be explored within a theoretical framework in chapter 2. For now, it is sufficient to note that monitoring feedback is a necessary component of dialogue.

Bavelas et al (1992; 1995; 2007) also found being in dialogue had an effect on the production of topic gestures conveying content information. However, findings here were mixed, with two out of three studies finding topic gestures to be produced at a higher rate in monologue as compared to dialogue (Bavelas et al, 1992; 1995) and another study, that more clearly separates out dialogue effects from visibility, finding comparable rates but more subtle qualitative changes (Bavelas et al, 2007). All of these findings fit with findings in spoken (Krauss and Wheinheimmer, 1964; Clark & Wilkes-Gibbs, 1986; Schober & Clark, 1989) and graphical (Garrod, Fay, Lee, Oberlander & MacLeod, 2007) signs where the amount of information carried in the sign tends to remain the same or become more complex in monologue as compared to dialogue, where it is reduced.

So, what is it about dialogue that makes it different from monologue? Speech and graphical sign findings are interpreted within an information-processing framework. Here, in any communication information is packaged for an addressee with speakers making and directing contributions towards the addressee so that information in the sign will vary depending on what the speaker assumes the addressee to know. Where more is assumed to be known, less needs to be conveyed in the sign and vice versa. This is then to do with the amount of shared knowledge held between the speaker and addressee as it is perceived by the speaker to be held by the addressee.

Findings in graphical sign are of particular relevance to gesture as, like many gestures, graphical signs are imagistic. Using a graphical equivalent of Krauss and Weinheimer's (1964) verbal referential communication task, Garrod et al (2007) considered graphical sign both in relation to the object of representation and the interaction. In a Pictionary style experiment items were drawn for an interlocutor who was to guess what the items were. Items were drawn over the course of six blocks so that the way in which each item was represented, and any change in representation, could be analysed. The experiment consisted of three different conditions; (1) a single Director repeatedly drawing for an imaginary audience so there was no feedback (SD – F); (2) a single Director drawing and another person matching allowing for minimal feedback (SD + F) and (3) alternating Director-Matcher roles so there was dual feedback (DD + F).

As dialogue requires a specific level of interaction and, in addition to this, the ability to monitor feedback, varying these determined whether the communication was in a dialogue or monologue style of communication. This initial experiment tested whether graphical signs were simplified and refined. It also tested whether refinement came about through a process of repeat production or, through an interactive process requiring feedback such as in dialogue. Garrod et al (2007) found items were best identified when there was feedback and when feedback was maximised, in the highly interactive conditions (SD+F and DD + F respectively). The refinement of graphical signs also depended on the provision of feedback and therefore, on whether the interaction was in a dialogue.

During interactive graphical communication (ie the two feedback conditions), over the course of the interaction iconic representations became increasingly less complex and more symbolic whereas in the less interactive conditions, where there was no feedback, iconic representations became more complex. Relating this back to the definition of signs more generally through the gesture research findings, just as gestures can be more or less abstract, iconic graphical signs were more directly representative of the information they conveyed than symbolic graphical signs. Graphical signs could then be said to vary in terms of how they related to their object of representation. Moreover, what determined this relationship was the level of interaction and whether feedback was afforded. The authors concluded that, with the ability to monitor feedback as in dialogue, '… the burden of information-carrying shifts from individual signs into history' (p5) so that less needs to be produced in the sign. When the interaction is highly interactive, as in dialogue, those communicating reach a shared conceptual state that facilitates the communication. This is evident from the finding that less needs to be conveyed in the sign in order to get the message across.

Overall, spoken, graphical and gesture sign findings show what is signified in a communication depends on the level of interaction in that communication. Moreover, they find the interaction level is determined by the amount of feedback afforded. They show that when the interaction is highly interactive, as in dialogue, less information needs to be conveyed in the sign in order to get a message across. Returning to gesture signs, Bavelas et al's (1992; 1995) finding of a reduced rate of topic gestures coincides with an increase in the production of interactive gestures. That is, whilst less topic gestures were produced in dialogue than in monologue, more interactive gestures were produced in dialogue, when there was a visible component at least, as compared to monologue. As interactive gestures cite given information, these gestures can function to make a reduced reference to information conveyed at an earlier point in the communication. This means, rather than re- referencing (through gesture and/or speech) a second mention of something gestured and/or spoken about earlier on in the conversation, interactive gestures can refer to the same reference without having to fully reference it. In this way, the gesture sign is essentially modified so that it carries less information. Overall, these findings suggest there is a reduction in the gesture sign.

Framed in terms of the way information is packaged, the studies reported here suggest what is important for dialogue is the shared knowledge state held between those communicating. In line with this are studies directly testing the effect of shared knowledge as they support the idea that gesture signs are altered, to carry more or less information, depending on the needs of the communication. The findings reported in the next section therefore lend further support to the claim that sign modification, resulting in a reduced expression, is a general property of all signs.

1.3.2 Shared Knowledge

Studies contrasting shared knowledge at the outset of a communication with when it develops over the course of the communication with when there is no such knowledge, find signs are modified depending on the needs of the communicative context. As in dialogue, shared knowledge influences gesture signs in a way that reduces the amount of information carried in the sign. As the shared knowledge between those communicating varies, so do the properties of content gestures with what information as well as how it is conveyed differing so as to reduce the expression. Gerwing & Bavelas (2004) designed an experiment where speakers held and worked with particular objects and then, in the absence of the objects, described them to addressees. Experiences were varied so that addressees had either also worked with the same objects as the speakers or had no direct experience of the objects. Since those communicating had either a shared experience or not, this altered the level of shared knowledge. Gerwing & Bavelas (2004) found gestures were influenced by the original instrumental action. Essentially then, gestures were influenced by what was to be signified and how they related to this object of representation.

In addition, Gerwing & Bavelas (2004) found the level of shared experience to influence gesture production. They found gestures were more informative, complex and precise when interlocutors lacked common ground. By looking at how new and old the information conveyed was, they also investigated the effects of accumulated shared knowledge. As predicted, gestures presenting new information were most salient as they were more precise, life-like and well-formed. Similarly, earlier references conveying original information were more salient than later mentions conveying old information. Since particular features (eg location) were retained, Gerwing & Bavelas (2004) suggest gestures were systematically transformed to provide less information.

Here then, as well as the referent (in this case an instrumental action) influencing gesture sign production, the amount of shared knowledge and, in relation to this, the saliency of information also influenced gesture signs. Moreover, it did so in a systematic way. Overall, the authors conclude that since both shared knowledge and saliency are related aspects of the local context, gestures are influenced locally at the level of the interaction.

Saliency is related to shared knowledge as when a knowledge state is shared, with information known and assumed to be known, information needs to be produced in the sign less often. In being given, and therefore not new, information needs to be less salient. Moreover it is the saliency of information that influences what is relayed in the gesture sign. For example, a first mention of an entity might be represented through a complex imagistic iconic gesture with say, a Cvpt. Think of an elaborate sweeping movement that frames the head of the gesturer and then continues to run down the gesturer's body from shoulders to torso and to finally sweep outwards and upwards to represent 'little red riding hood' (wearing her coat of course). In a later mention however, 'little red riding hood' might be represented as a less complex iconic (with just the head framing bit or just the final swish for example) or be reduced even further to a deictic gesture that merely indexes an abstract point in the gesture space. Later gestures mark the speaker's (or gesturer's) assumption of the addressee's mental model, that they know (from the reduced expression) what is being talked about since it refers back to the initial fuller expression. Conveying information through a deictic gesture, based on the assumptions made about the original iconic gesture, would be on a par with replacing a complex noun phrase with a pronoun in speech.

Indeed, McNeill (1992) notes that because gesture signs lack the constraints of language, they are freer than language and so can convey salient aspects of the discourse more directly. To determine what is salient, McNeill (1992) compares gestures in terms of their complexity. Exemplifying simple and more complex gestures, McNeill (1992, p 125-126) notes that variation in the complexity of the gesture can indicate what is salient to the gesturer. McNeill (1992) employs a qualitative methodology in exploring the complexity of gestures but points to Kita (1990) for a more quantitative approach where gesture complexity was scored on a point system. Here, gestures were given additional points if they had certain features. Additional features, such as whether the gesture was two handed and/or had movement in the fingers (see McNeill, 1992, p126), could then be taken to highlight what was salient and at what point it became salient.

Gestures depicting motion events can then be broken down into components to serve as markers of assumed knowledge similarly to the way syntactic and lexical components can be in speech (Parrill, 2004). To better define the complexity of gestures, Parrill (2004) investigated motion events by breaking them down into ground, manner and path. More complex gestures consisting of multiple components could then be distinguished from less complex gestures consisting of only one or two components. This de-compartmentalisation was used to manipulate the saliency of information (by using components or not) in the experimental conditions and in analysing the semantic complexity of the gestures produced. The experimental conditions either referred to a ground prompt, and made information old/less salient, or did not refer to a ground prompt, so as to make the information new/salient. In addition, by having conditions where both or only one of two interlocutors watched a cartoon clip, the amount of shared information varied. Gestures varied as a function of both factors, as when information was less salient and knowledge shared, information was less often coded in the gesture medium than when either factor worked alone. Again then, this study found both shared knowledge and, in relation to this, the saliency of information, to influence particular qualitative features of the gesture sign.

A further study by Holler & Stevens (2007) shows the effects of shared knowledge on particular semantic features and importantly, how the effect of shared knowledge works across different modalities. Having both speaker and listener (what they called the knowing condition) or speaker only (what they called the unknowing condition) privy to information, Holler & Stevens' (2007) varied the amount of shared knowledge about overall information, and in particular size information. They found speakers in the unknowing condition represented information (such as banana) in gesture alone or in gesture and speech, whereas speakers in the knowing condition represented information verbally and without gesture. In addition, size markers (such as 'big') were represented gesturally when there was no shared size information (the unknowing condition) whereas they were represented verbally when size information was shared (the knowing condition). Also, size information in gesture was less pronounced for knowing than unknowing addressees. Holler & Stevens (2007) conclude that shared knowledge has different effects on speech and on gesture and on how speech and gesture interact.

When both gesture and speech represented size information in the condition with no shared knowledge, adding shared knowledge led to information being represented in the verbal channel alone. This indicates a reduction in the amount of information conveyed. In contrast, when gesture alone represented information in the condition with no shared knowledge, shared knowledge led to a swapping of channels over to the verbal modality. Importantly, whether it was to do with qualitative aspects or more quantitative measures like overall production, in both cases the gesture sign was reduced.

Being in dialogue, shared knowledge and the saliency of information are all found to influence what and how spontaneous gestures sign. Combining the findings, it seems what is important for dialogue is the shared knowledge state and in relation to this, the saliency of information. As is noted by Gerwing & Bavelas (2004), shared knowledge is an aspect of the local context. Therefore sign production is influenced by the needs of the local context. Signs are essentially modified through everyday use. These findings support and are interpreted within dialogue theories of communication. Such theories hold the monitoring of feedback, shared common ground and being in dialogue to be integral parts of the communicative process so much so that their presence or absence influences gesture sign production. More specifically, shared knowledge and being in dialogue are said to be dependant on the ability to monitor feedback.

Chapter 2 of this thesis reviews the dialogue theory of Clark (1996), alongside a mechanistic model proposed by Pickering & Garrod (2004) that has followed on from it. Dialogue theories like these treat speech and gesture as independent signs that are a part of a wider sign system. Before moving on to a theoretical explanation of the findings reported in this chapter, a final section summarises the reasons for treating gesture as an independent sign.

1.4 Gesture as an independent sign

As reported throughout this chapter, spontaneous gestures communicate a significant amount of information. What's more, they often do so completely on their own, in the absence of speech, and this is found for gestures with an organisational function as well as for gestures conveying content information (Bavelas, 1992; 2007; Beattie & Shovelton, 2002). When gestures do accompany speech, they often convey complimentary information (McNeill, 1992). Such instances have been called gesture-speech mismatches (Church and Goldin-Meadow, 1986; Goldin-Meadow, 2003). An example of a complimentary gesture from Church and Goldin-Meadow (1986) is where the gesture [how short and wide the dish is] conveys more information than the speech ["the dish is wide"]. As well as the gesture conveying complimentary information, gesture-speech mismatches can convey different information from the accompanying speech. Church and Goldin-Meadow (1986) give as example, where speech conveys information about an action carried out on an object ['you poured water from the glass into the dish"] with the accompanying gesture conveying information about the object's dimensions [how tall the glass is and how small the dish is]. All gestures then, albeit some more than others, appear to have a degree of independence from speech.

Findings also show particular properties of signs are shared across different signs. For instance, when communication meets the requirements of dialogue, spoken, graphical and gesture signs are modified in a similar way so as to reduce the expression. That this property of reduction is shared across signs again suggests signs can stand alone in terms of the information they represent.

A theoretical reason for treating signs independently is to avoid imposing the properties of other signs on the sign under study. Acknowledging the independence of sign systems, like gesture, in which signs relate to their own object of representation, in various ways and more or less in these ways, affords the properties of the sign to be explored. Treating signs as independent at the outset is preferable as it is questionable whether properties of language, as it is written and read, can be ascribed to speech, as it is spoken and heard (Pickering & Garrod, 2004), let alone to gesture as it is performed and watched. Only once a sign is explored in it's own right can it then be compared to other signs, to explore any shared properties and how signs work together as part of an integrated system. As results reported in this chapter show, the way in which information is distributed across signs will depend upon how it is carried across the different modalities. Sometimes the signal will rely more on speech, making the gesture redundant, and at other times it will rely more on gesture, making the speech redundant.

This first chapter highlights the need to consider why gestures sign as well as what and how they do so. Whilst signs should be considered independently at the outset, findings reported in this chapter show they must also be considered in relation to the context within which they emerge. There is a great deal of variability in how signs relate to their object of representation. Most notably, imagistic content can be represented more or less directly and where it is, how closely related the gesture sign is to the object of representation varies. This, alongside finding other local level factors to influence gesture sign production, indicates signs also have a degree of independence from their object of representation. This being so, signs should not be defined purely in terms of their object of representation.

Clark's (1996) dialogue theory of communication treats signs as independent from the outset whilst acknowledging that all signs are a part of a larger composite signal. By this account, signs are defined in relation to an object of representation whilst taking into account aspects of the interaction found to influence them.

Chapter 2: Dialogue Theory

Dialogue theories of communication observe that, whilst we formally learn the linguistic elements of speech, through reading and writing, by the time we come learn in this way our language has already begun to develop on a much more informal basis. Language develops through everyday talk and this is chiefly in dialogue. Since there is no formal teaching of comprehension and production in spontaneous gesture, it is especially interesting to consider these gestures within a dialogical framework. Finding spontaneous gestures to be influenced by local level factors highlights the importance of the dialogue setting for the development of gesture signs. Moreover, that other signs share this property highlights the importance of the dialogue context for language development as a whole. In doing so, findings support a further assumption of dialogical theory; that language has developed in and through dialogue.

2.1 Common Ground

One of the most influential accounts of dialogue is from Clark (1996). According to Clark (1996), the communicative message is a composite signal made up of various signs that signify in different ways and it is how well these different signs are co-ordinated, alongside how well the methods of signalling within signs are co-ordinated, that's important for the communication. Such co-ordination is perceived as a joint activity between those communicating. When this joint activity is well co-ordinated, a representational history of the interaction builds up. This representational history is the shared knowledge state, or what Clark (1996) calls the shared common ground, in which interlocutors have 'common experience, expertise, dialect and culture' (Schober & Clark, 1989; p211). This shared history is then consulted and taken into account in the making of contributions. On the basis of their shared knowledge, interlocutors need only contribute what is necessary to get a message across. When meaningful information is held in a shared conceptual state, the communication is said to be grounded and when it is grounded, less information needs to be produced in the sign. In comparison, when the interaction is poorly co-ordinated with little or no information held in a shared representational state, information must be carried explicitly in the sign. Shared common ground is said then to be beneficial to the communication. The empirical findings reported in chapter 1 support this as they find information is packaged differently, with signs being relied upon more or less, depending on how well co-ordinated or grounded the interaction is.

Grounding is a cumulative process, where common ground is met and with the introduction of new information, is lost. It requires interactive communication that affords the ability to monitor both contributions made and feedback about contributions. This allows a previously shared state, which has been lost with the introduction of new information, to be repaired and regained. The ability to provide and monitor feedback is therefore crucial in determining whether those communicating reach a shared conceptual state. That the monitoring of feedback has a central place in the grounding process is also in line with findings reported in chapter 1. Since co-ordination relies on monitoring feedback, dialogical theory predicts that the conditions of monitoring, which can be crudely considered as hearing for speech and seeing for gesture, should influence how well co-ordinated the interaction is and whether or not a shared conceptual state is reached. It is crude because there is some cross-over of modalities in monitoring. For example, visually monitoring, movements of the eyes for instance, facilitates the speech channel by providing turn taking cues. Whether relying on visual or auditory signs, monitoring is optimal when co-present and when other conditions of dialogue are met. Dialogue is therefore considered to be the optimum setting for the co-ordination of action and for reaching a state of shared representation. Again, this is in line with findings reported in chapter 1 that emphasis the importance of monitoring feedback.

By Clark's account, when those communicating consult their shared history or common ground, they model each other's mental state. To do this, they must consider their own mental and knowledge states alongside another's. Contributions, in the form of what and how information is carried or is packaged in a sign, are said to be addressee specific as they are formed on the basis of what's assumed known by the addressee. As the gesture and speech findings reported in chapter 1 are framed within an information processing account, they again provide support for Clark's (1996) theory of common ground.

To summaries, what and how signs signify depends on the representational history of the interaction. More specifically, it depends on whether those communicating are grounded or not, both at the start of the communication and as it accumulates, since contributions are made on the basis of this shared mental state. Reaching such a state requires the interaction to be well co-ordinated. This is dependant on the monitoring of feedback, and on factors that influence it such as visibility. In broader terms, reaching a shared knowledge state depends on whether or not the communicative setting is optimal, like when in dialogue. For Clark (1996) then, these are the particular aspects of the joint social interaction that influence the cognitive production processes.

2.2 Interactive Alignment

In their model of 'interactive alignment', Pickering & Garrod (2004) provide a mechanistic account of the processes involved in reaching a shared representational state. The model assumes that when in dialogue there is an interactive alignment process whereby, in and through the interaction, representations are aligned at various levels. Alignment is based on an automatic priming mechanism where the output representations of production match the input representations of comprehension. Furthermore, it is suggested that alignment occurs as part of a within and between person process. Such priming of output by prior input, where simply hearing an utterance or seeing a gesture primes output representations, produces highly repetitive and fixed expressions. By narrowing down the choice of contribution, the model makes for better prediction, leading to a well co-ordinated or aligned interaction that enhances overall communication. The alignment process makes it more likely that the same utterance, or gesture, will be re-used later on. Therefore, as well as changing with regard to how much information they carry, or feedback they provide, signs also become more similar between those communicating. In support of this, Pickering & Garrod (2004) point out that dialogue is indeed highly repetitive and routinized. They go on to suggest it is through such repetition that contributions are interactively repaired.

As in grounding, when new information is introduced, aligned interlocutors can become misaligned. The process of representational alignment therefore also relies on the ability to monitor feedback. It is in and through dialogue that the model proposes comprehension and production processes become tightly coupled, to give parity of representation. Again then, the model considers the monitoring of feedback to be an integral part of interactive dialogue.

Alignment at one level is assumed to enhance alignment at others and it is through a 'percolating' process that linguistic alignment eventually leads to alignment at the critical level, that of the situational model. Alignment at the situational model is akin to the shared knowledge state held in common ground but there is a crucial difference between the two. Where those communicating are aligned, as in dialogue, speakers need only monitor and modify contributions on the basis of their own situational model since this model is aligned with the listeners. This leads to a state of implicit common ground. Implicit common ground is held to be sufficient for communication, with the fuller common ground that necessitates some inference of the other's mental state, only occurring under unusual circumstances like in deception or when there is difficulty (Pickering & Garrod, 2004). This contrasts with Clark's (1996) account where signs are modified on the basis of a shared common ground that involves a consideration of both speaker and listener models. In alignment, although utterances are self-specific, being made on the basis of ones own situational model in the first instance, since speaker and listener models are implicitly aligned, they are indirectly specified for the addressee's model.

Both accounts of the dialogue process therefore hold co-ordinated action and representation to be dependant on particular aspects of the social interaction. They highlight the monitoring of feedback, which can only occur through a process of interactive turn taking, as facilitating such co-ordination or alignment. They also show the level of co-ordination in any communication influences what, how and why information is produced in a sign. Since monitoring and coordination processes function at a local level, the local context is said to influence sign production. Both of these theoretical standpoints predict signs produced in interactive settings, where there are different levels of monitoring and co-ordination, like monologue and dialogue, will carry more or less information respectively. They predict then that sign modification depends on the needs of the communication and therefore on localised aspects. Although the research findings reported in chapter 1 are framed within Clark's (1996) theory of common ground, they also support the alignment model as this predicts the same local effects. However, the alignment model offers a different and more mechanistic account for them. Several studies specifically looking at alignment, in speech, graphical, body and gesture communication, are therefore reported in the following sections.

2.2.1 Speech Alignment

Research findings from speech show alignment occurs at various levels of the speech signal. It occurs on rhythm, sounds, structure, words and meaning (see Pickering & Garrod, 2004 for a review). All of the findings suggest interlocutors come to be aligned and that they do so through aligned representation (Pickering & Garrod, 2004). The findings also show that less information needs to be explicitly carried in the sign as the communication becomes aligned.

Most relevant to this thesis, in terms of the experimental design and findings from it, is an experiment by Garrod and Anderson (1987). Here, pairs of participants co-operated over a maze game task in order to elicit spontaneous descriptions about the maze map. As descriptions reflected the mental models of those communicating, they afforded insight into the state of the mental models of those communicating (or at least the speaker's mental model and assumptions about the addressee's model). Garrod and Anderson (1987) found the lexical and semantic choices of those communicating to be influenced by what speakers had just heard in the dialogue. Those communicating tended to converge on their speech over the course of the interaction and this resulted in more abstract descriptions. In other words, speaker outputs matched their inputs. This was an effect of precedence but Garrod and Anderson (1987) also found that the saliency of information influenced choices. They found communicators initially focussed on what was salient but later in the interaction focussed on to what was precedent. What communicators converged on shifted then over the course of the interaction as those communicating aligned. This entrainment was accounted for in terms of what the authors call an output/input co-ordination principle, whereby production of an utterance (or output) is formulated by the same principles as comprehension of an utterance (or input). These findings fit with the alignment model. As does the overall conclusion that the co-ordination of mental models is achieved through a process of collaboration at a local level.

2.2.2 Graphical Alignment

Similarly, convergence occurs in graphical signs. In a second experiment by Garrod et al (2007) the highly interactive condition from experiment 1 (see chapter 1) was modified. In the original condition, two participants had taken turns at drawing and providing graphical feedback throughout the process of drawing. In the newly modified condition, they did the same but in one of two feedback conditions. In a concurrent feedback condition, partners stood side-by-side to draw items whereas in a non-concurrent feedback condition, partners were separated by a visual barrier during drawing and could only interact graphically once drawing was complete. Graphical signs from both conditions became less complex and more symbolic over the course of the interaction.

Graphical signs evolve to converge over the course of the interaction and they do so irrespective of the type of feedback. Irrespective of the kind of feedback, graphical signs became less complex, more symbolic and converged. Moreover, they communicated just as well when they had been modified. As in speech then, when in dialogue where feedback is provided, graphical signs converged over the course of an interaction. Convergence occurred in such a way that the iconic, indexical and symbolic properties of graphical signs were combined so that any one sign had a mix of all or some of these properties. The distinctions between iconic and symbolic and iconic and indexical were said be graded and best placed on a dimension. This is just like McNeill's (1992) categories within spontaneous gesture sign which are treated as dimensional (see chapter 1).

2.2.3 Body Alignment

Interpersonal alignment also occurs at the level of gross kinetic movements such as in the way we hold and position ourselves (LaFrance, 1982; Shockley, Santana & Fowler, 2003). LaFrance (1985) compared intra and intergroup postural mirroring, a measure taken to indicate interpersonal rapport, whilst participants read pairs of letters with the aim of selecting a genuine letter from the two. Groups were made up of four participants consisting of two dyads in each. Instruction was first given to the two dyads and then to the group as a whole. There were four conditions; a cooperative condition where participants were informed that task performance was dependent on both dyads; a competitive condition where participants were informed that they were in competition with the other dyad; a co-acting condition where participants were informed each individual score would be compared to a mean average score so that performance depended on every individual's sole performance but where they were encouraged to collaborate and a control condition where they were informed of the same as in the co-acting condition but without collaboration as here there was no other person to interact with.

La France adopted Scheflen's (1964) definition of postural mirroring described as '… the degree to which two or more people simultaneously adopt identical or mirror-imaged bodily positions vis a vis one another' (cited La France, 1985, p207). Since previous studies had shown a relationship between congruent arm movements and rapport, postural mirroring was measured by coding arm positioning. Coders coded the arm positioning of both arms for each person with each arm being coded separately. Coders then categorised arm movements into twelve pre-defined categories. Next, notations of all possible combinations of movements in the pairs were coded for the presence or absence of mirroring which was indexed by an intra-intergroup index across the dyads.

It was predicted that the cooperative and co-acting conditions would show more intergroup (across dyad) postural mirroring as compared to intragroup (within dyad) postural mirroring than the competition and control conditions and that this would be strongest for the more extreme conditions (cooperative/competitive) with the other two (co-acting/control) falling in-between, in the predicted direction. Findings followed predictions as participants in the cooperative condition mirrored each other more often (indicated by an increased between-group mirroring relative to within-group) than those in the competitive condition (more within-group mirroring relative to between-group). This finding indicated dyads came together more when they were cooperative than when they were competitive. An unexpected finding was that co-acting produced the most inter-group mirroring. LaFrance (1985) proposed the explicit instruction to 'collaborate' when co-acting to account for this. In addition, all postural mirroring varied as a function of the expected (anticipated) interpersonal involvement confirming it to be an indicator of rapport.

Shockley, Santana & Fowler (2003) describe everyday conversation as '… a suprapostural task that involves gesturing, listening and visual inspection. These activities likewise require adjustments to the postural state for success' (p327). Interested in the interpersonal co-ordination of communication, they investigated the synchronisation of postural sway. They created four puzzle solving conditions; (1) participants worked together whilst facing each other, (2) participants worked together whilst facing away from each other, (3) confederates worked together whilst facing each other, (4) confederates worked together whilst facing away from each other. These conditions treated body orientation (towards/away) and partner (participant/confederate) as variables.

Shockley et al (2003) found an effect of partner (participant/confederate). Here, postural trajectories converged more often when participants communicated over the task than when participants and confederates communicated. In line with this, they found postural trajectories diverged less over time in the participant condition than in the confederate condition. Moreover, it was when the interaction was cooperative that postural sway converged or aligned. There was therefore a clear effect of level of interaction on movement entrainment. Shockley et al (2003) conclude that when those communicating act to jointly cooperate, not only is the solution to the problem a joint activity but the actions behind the outcome are also a joint activity.

2.2.4 Gesture Alignment

From a single group of communicators, Kendon's (1970) observed that the gesture movements between speakers and listeners are synchronised during conversation so that those communicating are in rhythm or what Condon and Ogston (1966) called 'interactional synchrony' (cited Kendon, 1970). This observation led Kendon (1970) to conclude that the level of involvement in a group can be judged by the degree to which postures and movements coincide with other group members. When in synch, listener movements were often a mirror image of speaker movements and Kendon (1970) illustrates this with the following example from the group when; 'As T leans back in his chair, B leans back and lifts his head; then B moves his right arm to the right, just as T moves his left arm to the left, and he follows this with a head cock to the right, juts as T cocks his head to the left. We might say that here B dances T's dance.' (p110). In addition, throughout listening, B's eye blinks, shifts and mouth moves were co-ordinated with T's speech. Interestingly, whereas B's movements were timed to the speech of the speaker when listening, when B initiated a response (head nod) this was no longer in synch with T's speech but was re-arranged in B's time frame. Although B initiated a head nodding response here, B was still in a listening phase. Later in the communication, when B took the floor and was speaking, B initiated a movement (a left and forward tilt of the head) and T mirrored it.

For Kendon (1990), the two persons directly involved in the communication make up what he calls the 'axis' of interaction. Although the others present are not a part of this axis, they can become a part at any point and where they do, the same patterns of behaviour occur between the various parties. For instance, on-lookers/listeners sitting to the side often display aligned movements that are in accordance with either the listener or the speaker. For example, Kendon (1990) describes the case of a side participant's movement [trunk to right + forward], which was analogous to the speaker's movement [head to left] and in line with the listener's move [head to right] (p116). Kendon (1990) frequently observed this pattern, where a side participant adopts a similar movement to the listener indicating that the side participant has adopted a listener role, albeit a more distant listener role than the listener in the axis.

We even align on the space we use in which to perform gestures, or what Kendon (1990) has called the actor's transactional segment. Kendon (1990) gives as example the watching space in front of a television to illustrate the possible depth of such a space. The transactional segment is the space directly in front of the person. In the case of more than one person, the space becomes a joint transactional space that is managed by the actors sharing it. Where individual transactional segments come together so as to overlap, or align, this is called an F-formation (Kendon's, 1990).

How spaces are managed to become aligned depends on the particular number of people in the interaction and how they are grouped. For instance whether the group forms '… clusters, lines or circles …' (p 209). The amount of overlap in the joint space is influenced by the group arrangement, which in turn dictates rights of passage in the communication. In group arrangements of more than two people, a circular arrangement is an equal setting for all group members. Other logistical aspects also impact upon this as, when seated for example, seating arrangements maintain and manage the arrangement. However, at other times and when logistical aspects change, interactive manoeuvres serve to acquire overlap in the spaces of those communicating (Kendon, 1990).

Parrill & Kimbara (2006) explored alignment in gesture by testing how sensitive observers were to alignment, or what they called mimicry. They did this by having observers view different video conditions of two people discussing the route they would take through a model town. The two people communicating either had; (1) mimicked gesture and speech, (2) non-mimicked gesture and speech, (3) mimicked gesture but no mimicked speech, (4) mimicked speech but no mimicked gesture. After watching one of the conditions, observers relayed the route discussed in the video. The observer's gesture and speech, and whether they mimicked it or not, was taken as a measure of the effect observed mimicry had on them. In line with their predictions, Parrill & Kimbara (2006) found observers produced more mimicked gesture when they had watched conditions with mimicked gesture and more mimicked speech when they had watched conditions with mimicked speech but no effect across modalities (ie mimicked gesture did not elicit mimicked speech and vice versa). Although these findings are of interest in that they suggest observers are sensitive to mimicry and moreover that seeing mimicked behaviour leads to it's production, they test the effect of seeing mimicked behaviour rather than whether mimicry, or alignment, occurs, and to what extent it does so, in particular contexts.

Like sign modification more generally then, alignment is most likely a feature of all signs. However, there is a need for more empirically based studies of gesture alignment. As well as showing gesture alignment, Kendon's (1970; 1990) studies highlight the dynamics of group communication. Group communication is an additional setting in which gesture occurs and is one that provides further scope for investigating gesture signs. Moreover, group studies manipulating group size in an experimental setting can elicit different speech styles like monologue and dialogue. This finding sits particularly well with Pickering & Garrod's (2004) alignment model where the model predicts that implicit grounding, necessary for alignment, can occur outside of two-person dialogues, such as in the group situation. As noted earlier, this differs from Clark's (1996) notion of common ground which is based on partner-modelling in two person dialogue. As well as being an alternative to the dialogue and monologue context, group experiments therefore test the prediction of alignment theory, that alignment is an implicit process. The following chapter (chapter 3) explores a group design for gesture by reviewing relevant group experiments.

Chapter 3: Group Research

3.1 A group design for gesture

As the group setting elicits different speech styles, it is a useful paradigm for studying communication. It also affords a prediction of the alignment model (Pickering & Garrod, 2004), that grounding and alignment can occur within group communication, to be tested.

Alternatives to one-person monologue and two-person dialogue are also attractive because of difficulty in re-constructing these styles of communication in an experimental setting. To treat the level of interaction as a factor, studies have paid particular attention to what constitutes dialogue and monologue. Bavelas et al (2007) in particular strived to meet as many of the pre-requisites of dialogue defined by Clark (1996) as possible. However, in their telephone condition, where they treated visibility as a separate factor, they did not adhere to the visibility pre-requisite nor did they adhere to the co-presence pre-requisite. Bavelas et al (2007) report constructing a monologue to be more difficult than constructing a dialogue, mainly due to the problem of people having an implicit audience in mind when in monologue, and suggest monologue might only ever be approximated in an experimental setting. However, as is illustrated by their telephone condition, the issue of dialogue is equally as complex. Although it is useful to consider visibility and dialogue as independent factors, especially in the study of gesture where the mode of communication relies on the visual channel, it is important to bear in mind that these are both necessary components of dialogue. It may be more conducive to consider the three conditions of this experiment as different approximations of monologue and dialogue. Whereas the tape recorder condition was like monologue, the face-to-face condition was like dialogue, and the telephone condition somewhere in-between.

Such a continuum of dialogue has been suggested by Pickering & Garrod (2004) who give examples of approximations in everyday dialogue and monologue. For example, in everyday dialogue-like communication, we chat on the phone or, via tele-computer link or, face-to-face. When in monologue-like communication, we lecture, present or make a speech, but whatever the occasion we usually have an implicit audience in mind. The group setting then is another approximation of these styles of communication as, depending on group size, communication is either more like monologue or dialogue. To say a group communication is in dialogue or monologue would be incorrect but it can be dialogue-like or in serial monologue.

A group design was employed by Fay (2000) to explore the speech styles that emerged in different sized groups, in relation to monologue and dialogue. In their experiment, Fay (2000) had groups of five and ten members discuss an issue of plagiarism. The task involved a degree of role play as participants were to think of themselves as being in a decision making group, like that of a University board or jury, where they were to reach a decision on what was the best course of action to take. Speech from the discussion was then analysed. In groups of ten members, speech was explicit and there tended to be a dominant speaker. In contrast, speech in the groups of five members carried information less explicitly and was more evenly distributed amongst members. In addition, larger groups had fewer verbal back channel responses, considered to provide feedback about the communication and related to this, speech turns were longer. Fay (2000) found the group speech patterns fitted with earlier findings in monologue and dialogue (Krauss and Wheinheimmer, 1964; Clark & Wilkes-Gibbs, 1986; Schober & Clark, 1989), as speech in larger groups was like serial monologue whereas speech in smaller groups was like dialogue. This led to the conclusion that, varying the number of people in the communication influenced monitoring ability and the processes dependent on it, such as whether the communication was in a dialogue style or serial monologue and the suggestion that difficulty in monitoring in large groups leads to a poorly co-ordinated interaction and a compensatory monologue style of speech (Fay, 2000).

Overall, these group findings show that monitoring cues (for feedback and repair to the on-going communication), had a bearing on how well co-ordinated the interaction was. Moreover, they show co-ordinational aspects of the interaction, said to determine the shared knowledge of those communicating, influenced what was produced in speech, in terms of the amount of information and way in which it was carried. The findings therefore fit with the dialogue models of communication discussed in chapter 2. In addition, as group findings mirrored speech findings in one-person monologue and two-person dialogue, group communication could be said to be representative of both dialogue and monologue styles of communication. They therefore fit particularly well with the alignment model (Pickering & Garrod, 2004) as this predicts implicit grounding and alignment can occur within group communication.

Fay (2000) investigated the speech sign, and so the spoken and heard aspects of monitoring. However, they suggest visual aspects of monitoring, like eye gaze, would also have been disrupted. For instance, the increase in group size would have limited the ability to monitor all parties. In addition, spatial arrangements differ with changes in group size and group members can be more or less occluded depending on this arrangement. Even when communicators are not occluded, Ozyurek (2002) has shown how a shift in seating position and so the shared gesture space, influences the communication in terms of gesture production and this study is reported next.

In gesture research so far, the maximum number of interlocutors experimentally investigated is three. This was in an experiment by Ozyurek (2002), where the number of people communicating varied with the re-telling of a story so that where a story was originally told to a dyad, it was re-told to different recipients in either a dyad or triad. At the same time as varying the recipient number, the recipient location also varied. Dyads sat side-by-side, having shared space to the side of the narrator, whereas triads sat face-to-face in a triangle, having shared space to the front. The orientation of gestures differed depending on the condition, with more lateral gestures occurring in dyads and more saggital gestures in triads. To test whether this was an effect of location or of number of recipients, a second experiment held the addressee number constant whilst varying their location. The initial findings were replicated leading Ozyurek (2002) to conclude that addressee location, rather than recipient number, had an effect on how speakers utilised the gesture space they shared. Interestingly, effects were strongest for spatial prepositions of motion descriptions bound by a beginning and end point, such as 'into' and 'out'. That effects were strongest for particular types of information again shows how the object of representation, as well as the relationship the sign has with the object, influences gesture sign production and that, depending on this relationship, certain types of information may be better represented in gesture than others. As changes in the gesture signal were similar to those in the speech signal, they support dialogue theory in the idea that gesture is a composite signal to speech (Ozyurek, 2002).

Ozyurek's (2002) finding of an effect of location, rather than of recipient number, is interesting in terms of applying the group design to the study of gesture. It suggests a group design might not replicate findings in speech. However, by returning to findings in speech this can be explained further. Groups have an optimum size; those with less than five members have too few resources, such as information and channels of communication, and those with more members have too many resources (see Fay, 2000 for a review). Moreover, findings suggest there is a switch in the communication around the number seven, since real world working groups consisting of eight members (Fay, 2000) have a different communication style to those consisting of five members (Carletta, Garrod, Fraser-Krauss, 1998). Taking these group findings into account then, there should be no difference between the communication style of dyads and triads, as Ozyurek (2002) found.

The group findings sit well with the gesture and graphical sign findings reported in chapter 1, as these also suggest a reliance on the visual channel to monitor feedback. As the group design affords an additional context within which to explore monologue and dialogue, and the effect of different styles of communication on the production of signs, it is of interest to apply such a design to the study of gesture. Considering the findings to date, in particular the importance of visibility for gesture and especially for those involved in monitoring feedback, any experimental design used to explore gesture should address the issue of monitoring in the visual channel. A group design, modified to suit the study of gesture, does just this as it manipulates monitoring load without explicitly violating the co-presence and visibility pre-requisites of dialogue. This thesis therefore employed a group design (see chapter 7) to explore any effect of group size (taken as an indicator of level of interaction) on gesture production. In addition, by specifically employing groups sized five and eight the group experiment of this thesis further explored where the switch in the communication style is. As the group design is a novel way to study gesture communication, before running the group experiment, a pilot (chapter 4) first explored both the design and emerging communication. On the basis of this pilot, a dyad experiment (chapter 5) tested the design further and got a baseline for the groups.

A number of predictions were made for the group experiment. These were based on the findings reported throughout chapters 1, 2 and 3 alongside the dialogue models of communication but more specifically, since the design was based on group communication, the alignment model. Firstly, it was predicted that members of the groups of five would be able to monitor feedback well and this would make for a better co-ordinated communication resulting in a dialogue style. In groups sized ten, it was predicted that the increase in group membership would make for less efficient visibility and difficulty in monitoring feedback, resulting in a poorly co-ordinated communication with a monologue style. It was expected that these different communication styles would be evident from the emerging gesture rates and patterns.

Since non-content social and beat gestures are involved in the provision of feedback and so, rely heavily on the visual modality to be monitored, it was predicted that more of these gestures would be produced in groups of five. As well as the presence of these gestures indicating the occurrence of monitoring, or their absence a problem in monitoring, interactive gestures can also function as a reduced expression. As reduced expressions should be predominant in more interactive settings like dialogue, this led to a prediction in the same direction.

Predictions for content gestures varied. It was predicted that more deictic and place holder gestures, which are reliant on both interaction and visibility (Bavelas et al, 2007), would be produced in groups of five. Predictions for Cvpt and Ovpt iconic gestures were less clear. In prior research, visibility is found to influence the more qualitative aspects of these gestures whereas being in monologue or dialogue is found to influence the overall rate (Bavelas et al, 1992; 1995; 2007). Iconic gestures are more often produced at higher rates in monologue, as compared to dialogue, and this is in line with findings in other signs that show there to be a reduction in the sign when in dialogue. However, results are conflicting with one study finding rates of iconic gestures to be higher when in dialogue (Bavelas et al, 2007).

Based on the majority of findings, where speech, graphical and gesture signs are reduced in dialogue, it was tentatively predicted that more Cvpt and Ovpt gestures would be produced in groups of eight, with their monologue style of communication *than in the groups of five, with their dialogue style. Such a prediction was also in line with dialogue theories of communication. However, since iconic gestures rely on visual monitoring, larger groups could encounter problems due to the increased monitoring load. This could counteract any effect of interaction such that the production of these gestures would drop in the larger groups. In addition, Cvpt gestures rely heavily on visual aspects and so occur more often in dialogue (Bavelas, 2007). More Cvpt gestures could therefore be produced in the smaller groups. Lastly, rather than there being a change in the overall production rate, more qualitative aspects of iconic gestures could change. Predictions were tentative then because effects of visibility, saliency and the possibility of more qualitative changes could counteract any effect of interaction level.*

Regarding the gesture patterns, it was predicted that groups of eight would have a more dominant gesturer and groups of five a more evenly distributed gesture pattern.

3.2 Alignment and Sign Convention

The group design has also been employed to study the emergence of conventions in speech. On a maze game task, Garrod & Doherty (1994) compared isolated pairs of speakers with pairs matched in a community of speakers. Here isolated pairs played several games (9 in total) to give them a lengthy exchange with the same partner. The linguistic community played the same amount of games but with different partners so as to form a community of speakers. In both the pairs and the community group, as in the pairs of Garrod and Anderson (1987), there was evidence of an output/input co-ordination process that led to converging speech.

Convergence in the isolated pairs was not however the same as convergence in the linguistic community. In isolated pairs, inter-speaker convergence occurred quickly and reached a plateau (at game 3). In comparison, the community was poorly co-ordinated to start with but came to be co-ordinated at a steady pace as players were introduced (over all games), until eventually the community was better co-ordinated than the isolated pairs. Observations revealed that convergence in the isolated pairs relied heavily on local aspects, to do with salience and precedence. Although speech converged in the pairs, whenever a member failed to support the converging bit of speech it was no longer valid. In comparison, convergence within the community was less reliant on precedence and salience. Here, convergence relied on these local aspects at the level of the interacting pair but it also involved a global process whereby successful schemes from previous pairs were re-used in subsequent pairs. This resulted in successful schemes being adopted and used by the whole community. In this way, convergence in the community was accepted by and had the support of a community of speakers and could better withstand any violation. Convergence in the community therefore occurred through a process of conventionalisation (Garrod & Doherty, 1994).

Differences in the way speech converged put different constraints on the available choices. Whereas isolated pairs were sensitive to what was salient and precedent, the community were sensitive to what was held to be an accepted choice, or convention, within their community. Although converging speech was sensitive to different aspects and occurred at different local and global levels depending on condition, the reasoning behind the language refinement was considered to be the same (Garrod & Doherty, 1994). It was said to have the same underlying function, to facilitate semantic co-ordination of the interaction. The authors therefore conclude they were related processes. Moreover, they suggest local level convergence, observed in pairs, was the mechanism underlying global convergence, observed in the community as convergence in the community was essentially a two-step process, whereas in the isolated pairs it involved only the first step.

The findings from this group study support a further facet of alignment theory; that through alignment signs evolve to become routinized and more conventional (Pickering & Garrod, 2004). Considering local effects alongside global effects, the alignment model addresses how the modification of signs at a local level might relate to the evolution of signs at a global level. As stated earlier by Garrod & Doherty (1994), the model holds local alignment, where signs are modified through a process of convergence, to be related to global alignment, where signs are modified through a process of conventionalisation.

The question is, do gesture signs evolve in a similar way? In order to address this question, it is necessary to consider spontaneous gesture alongside the issue of conventionalisation. Kendon (1995) does this by comparing conventionalised gestures, or what have been called emblems (Eckman & Friessen, 1969), with spontaneous gestures in Southern Italians. Comparing the conventional Mano a borsa, or purse hand, to what he calls the finger bunch, which is spontaneous, and the Mani giunte, or praying hands, to what he calls the spontaneous ring, Kendon (1995) finds these gestures are of similar form and perform similar functions. The variation in form is evident from their names, which are based on the gesture form, and whereas the emblems function as illocutionary markers, the spontaneous gestures function as discourse markers. These gestures are therefore like some of the social gestures described by Bavelas et al (1992; 1995). The reason the first two are considered conventional and the other two not, is because conventional gestures can be detached from speech. By this definition of conventionalisation then, whether a gesture is considered conventional or not is to do with how 'detachable' the gestures are from speech (Kendon, 1995, p267). Kendon (1995) suggests that although the spontaneous gestures are not conventional in the same way, they are related to conventional gestures. Moreover, it is suggested that conventional gestures originate form spontaneous gestures (Kendon, 1995). Indeed, Kendon (see 2004 for a review) and Muller (2004) find social gestures taking the form of McNeill's (1992) conduit gesture, described as the palm up open hand gesture, also have a degree of conventionalised. This leads to the suggestion that this gesture belongs to a family of gestures that are all related (Kendon, 2004).

Kendon (1997) suggests all gestures should be treated as 'a range of forms that vary in their degree of conventionalisation' (pp 119) and notes that whilst gestures can be more or less conventional, most are 'intermediate' (pp119) in their degree of conventionalisation. Kendon's (see 2004 for a review) dimensional approach is depicted by McNeill (1992) in what he calls Kendon's continuum. The continuum places spontaneous gesture at one end and conventional gesture at the other as can be seen below;

gesticulation (or spontaneous gesture), language-like-gestures, pantomimes, emblems (or conventional gestures) and sign languages (see McNeill, 1992, p37).

Just like McNeill's (1992) categories within spontaneous gesture then, as gestures are on a dimension, any gesture placed on the continuum is considered to have more or less of the given properties along the continuum. Despite this dimensional approach, in practice there has been a tendency to study the gestures placed at either end, like spontaneous and conventional gestures, in isolation. This has led researchers to consider the issue of conventionalisation separately from spontaneous gesture. Such a division does not allow the issue of local convergence and how it relates to global conventionalisation to be addressed. In addition, emblems are only one type of conventionalised gesture suggesting that conventionalisation may occur elsewhere on the continuum.

In a review of cross-cultural studies, Kita (2009) describes how emblematic gestures arise as a result of form-meaning conventions within different cultures. Yet Kita (2009) reviews other types of conventionalisation in gesture that is due to different factors such as cognitive diversity, linguistic diversity and pragmatic diversity. Like the field of gesture communication more generally, most studies providing support for conventionalisation in gesture are observational. Or, where there is empirical evidence, studies do not address the question of conventionalisation in gesture directly. An exception is studies looking at linguistic diversity across cultures where experiments with gesture have been devised. These studies show differences in language influence gesture production in a way that could lead to gesture sign convention.

Kita & Ozyurek (2003) show gestures are shaped by the linguistic properties of a language. Testing languages that package speech differently, they found gestures were influenced by the way in which the speech was formulated. They contrasted Turkish and Japanese speakers with English speakers in a story telling experiment. In the experiment, participants re-told a cartoon they had watched to a naïve listener and their re-telling of the story was video recorded so that it could be analysed. The cartoon contained two analysable scenes. The first was called the swing scene because in it a cat swung across an imaginary rope to catch a bird. Whereas the arc trajectory is easy to verbalise in the English language, as the intransitive verb 'to swing', there is no equivalent to this intransitive verb in Turkish and Japanese and so, the trajectory is not easily verbalised in these languages. In Turkish and Japanese, where verbalising the trajectory is difficult, it was predicted that the trajectory would be produced less often in gesture as compared to in English, where verbalising the information is easier and so would be represented in gesture more often. In line with predictions, in addition to the arc trajectory being missed out in Turkish and Japanese speech, it was also missed out in gestures as speakers more often made an arc trajectory with a straight line or just a straight line. In comparison, English speakers referred to the arc with the verb swing and more often made an arc trajectory to represent the swing gesture.

The second scene of the story was called the rolling scene in which a cat rolls down the street. To convey this scene, Turkish and Japanese speakers more often produced two separate gestures for manner and trajectory. Although single gestures producing both manner and trajectory were comparable across languages, these gestures were more often accompanied by separate gestures for manner and path in Turkish and Japanese. These differences in gesture again reflect the organisation of linguistic properties. They therefore show how language influences gesture sign production and how cross-linguistic variation can occur in gesture signs. Findings from this cross-cultural study then, alongside others reviewed by Kita (2009), provide scope for the conventionalisation of gesture.

Further evidence of the conventionalisation of gesture comes from homesign languages. Morford (1996) describes homesign as the gesture communication of non-hearing individuals who have not learnt a spoken language and have not been exposed to any formal sign language. Typically, homesign develops over a single generation of users and is used by a limited community. This prevents it from becoming as linguistically complex as formal signed languages such as American or British sign languages (ASL/BSL). However, despite being less complex than more formal sign languages, homesign meets the needs of those communicating. Moreover, the signs used to communicate evolve from within a specific community as those in it use the signs to get a message across.

There is also a Bedouin sign language described by Arnoff, Meir, Padden & Sandler (2008) which is both distinct from homesign and other more mature sign languages like ASL or BSL. Bedouin sign is distinct from homesign because it emerged 70 years ago and so, over several generations. It is also used widely by both non-hearing and hearing individuals in the community. Bedouin sign is also distinct from ASL and BSL as, although it is more mature than homesign languages and is a complex sign language with similar features to language, it's structure is less complex than more formal sign languages. As a sign system then, Bedouin sign seems to fall somewhere in-between homesign and more formal sign languages. Altogether, the development of homesign and Bedouin sign language, alongside more formal sign languages, suggests that the conventionalisation of gesture may indeed lie on a continuum.

Observations and empirical findings alongside the evolution of informal sign languages within given communities suggest gesture can be conventionalised in a similar way to other signs, like speech and graphics. Although this thesis does not address the issue of global conventionalisation in gesture, it does investigate whether local convergence, the first step in the conventionalisation process (Garrod & Doherty, 1994; Pickering & Garrod, 2004), occurs in gesture sign. Convergence was measured quantatively by correlating gesture amounts in both the dyad (chapter 5) and group data (chapter 7). Convergence was also measured qualitatively in an overseer experiment on gestures from the dyads to investigate alignment on gesture form (chapter 6).

Based on findings of alignment (chapters 2) and findings supporting the idea that the conventionalisation process begins locally (chapter 3) it was predicted that local convergence or alignment, in being a particular quality of all signs, would occur in spontaneous gestures. This led to the prediction that gesture rates would be highly correlated within pairs and that gesture rates would be more highly correlated in groups of five than in groups of eight. For the overseer, it led to the prediction that gestures from the same isolated pairs would be rated as more similar by over seers than gestures from different pairs. As Cvpt gestures are more egocentric than Ovpt gestures, it was predicted that Ovpt gestures would align more than Cvpt gestures.

Chapter 4: Pilot

4.1 Introduction

This pilot study modified McNeill's (1992) cartoon paradigm in which one participant typically describes a moving comic strip to another who has not seen the clip. Such a paradigm is useful for exploring hand gestures since the description of comic action and motion events encourages their use. There were several differences between McNeill's (1992) paradigm and the one used in this pilot experiment. The stimuli differed in that static comic stills from a comic strip were used rather than moving clips from a film cartoon. Here, instead of one person having all the information and re-laying it to another a number of participants took part. These participants each had an equal amount of information but it was different information. As the overall task was for participants to construct a story based on all of the information in the stills, the task encouraged group members to share information in a collaborative way.

To create different levels of interaction, a group design similar to that used by Fay (2000) to investigate speech patterns was used. The group design was chosen because it elicits different styles of speech depending on the number of participants, with larger groups having a serial monologue style of communication and smaller groups, a dialogue style (Fay, 2000). These different styles of communication emerge, because monitoring feedback, which is a necessary component of dialogue, is easier in smaller groups. The pilot experiment therefore was a first step in investigating the effect of the group context, alongside the level of interaction it afforded, on the communication style of the groups. Since group communication can evoke both monologue and dialogue styles of communication, it is an alternative context for considering the role of the interaction. The main aim of the pilot was to test the feasibility of the group design for gesture study and the feasibility of gesture coding in groups.

Participants were grouped under two group size conditions; a group of 5 members (G5) and a group of 8 members (G8). As every group member received one comic still each, the number of stills in the group as a whole varied across group conditions with group membership. Whereas G5 received five stills, G8 received eight stills. Although based on Fay (2000), different group sizes were chosen to investigate more precisely when communication shifts from dialogue to monologue. As research suggests this is around a size of 7 (Carletta, Garrod, Fraser-Krauss, 1998; Fay 2000), this was the chosen cut off point.

The pilot also served to check whether a cartoon with a more imagistic and abstract story content (The Beano) elicited more gestures than one with a concrete story content (Tom and Jerry). It also checked any effect of the absence/presence of stills on both gesture rates and memory for story construction. The preferred design was one with stills absent, to eliminate the possibility of referring to comic stills directly (with say a point to the referent) and so that hands were unoccupied, both of which could confound gesture production rates.

As this was a pilot, no predictions were made but based on findings in speech, it was expected that smaller groups would have a dialogue style of communication whereas larger groups a monologue style (Fay, 2000). It was expected then that more content gestures and speech would be produced in the monologue style of the larger groups, whereas more non-content gestures would be produced in the dialogue style of the smaller groups.

On the basis of alignment findings and the alignment model, gesture alignment was expected to be more evident in G5 than in G8 and, more often on Ovpt gestures since these are less egocentric than Cvpt gestures. It was also expected that gestures would also align on form. To investigate alignment on gesture form, iconic gestures with the same referent were identified for qualitative analysis

4.2 Methods

4.2.1 Subjects

Thirteen students from the University of Glasgow took part in this pilot. Students participated in a mixed gender group of 5 (consisting of 4 females and one male) and a mixed group of 8 (consisting of 7 females and one male). All were paid cash for their participation.

4.2.2 Materials

Sixteen comic stills from two static cartoon strips were selected. Eight stills were taken from a Tom and Jerry cartoon (from now on called T&J). These depicted Tom the cat chasing Jerry the mouse with various accidents happening along the way, such as tins of paint being knocked over. These scenes showed the cause and effect of an action and were therefore grounded in concrete events and ideas. Eight stills were taken from a Beano cartoon and depicted Denis dreaming of a journey through the Beano Book. The story was a dream sequence, in which the character Denis was already in the book he was dreaming about. The comic strip used Alice in Wonderland as a metaphor and so strange events could not always be explained in terms of the usual conditions of cause and effect. The Beano story content was therefore more abstract than that of the T&J story content. Stills were selected depending on how well they elicited gestures, for instance those that encouraged spatial reference were preferable, and on the strength of their story theme, so as to enhance story construction. With this basis for selection, stills were not necessarily in chronological order. Stills were blown up to A4 size and verbal expressions blanked out. Onomatopoeic motion and noise words (eg whiz; bang) that exaggerated action events, and in doing so encouraged gesture production, were left in provided they were not part of a longer expression.

4.2.3 Procedure

5 participants were seated on chairs arranged in a circle to make the group5. Participants were placed in a circle as this was the preferred arrangement for seating groups of two or more since it gives each group member an equal share of the gesture space (see Kendon, 1990 in section 2.2.4).

In the experimental room, two Sony digital video cameras were set to record all participants, one from a side angel and the other from a bird's eye view. The camera mounted in the ceiling, to capture the bird's eye view, was fitted with a fish eye lens. Both cameras were set to capture the interaction at 25.00 frames per second (fps) with a frame size of 320 by 240 at 48, 000 Hz and 32 bit float in stereo.

A random 5 of the 8 stills from the T&J comic were selected. Participants were given the comic stills and asked to place them face down on their knees whilst awaiting instruction. They were informed that each group member had a different still and that stills had been handed out in a totally random order. Therefore, still order did not necessarily equate with where each person sat. Next they were informed that the task was to discus and order the stills into a cohesive story by using all the information they had between them. At this point all participants were given some time to look at the detail of their own stills. They attended to the detail in isolation so that they could only ever see their own still. They were then asked to place the stills face down on their knees as before. Although the comic stills were tangible, in that participants had access to them, it was requested that they be kept face down on their knees throughout the discussion. Participants were informed that if they needed to check information on their still, they should do so by breaking from the conversation for as brief a time as possible (ie. not converse at this point). As well as reducing the likelihood of hands being occupied whilst talking, these measures ensured other group members could not see the stills, both factors that could have confounded gesture production. It was explained that, information could be missing and so, they should think of and use possible events outside of the information they had especially if constructing a story was proving to be difficult. Once the group members felt they had found the best order for a cohesive story, the first interactive task was complete.

In a second task, the same 5 participants received 5 randomly chosen stills from 8 of The Beano comic stills. In this second task, the procedure was the same as for that in the T&J cartoon but for a memory component. The same procedure was explained, as in the first part of the experiment, but with additional instruction to memorise the still content. Participants memorised the still content in isolation so that they could see only their own stills. They were then asked to place the stills face down under their own chair. They were instructed to always attempt to retrieve the detail of their stills from memory. Once all participants were confident they had memorised the details of their own stills, the interactive discussion began. Again the task was complete when they were happy with the story constructed.

The same procedure was followed for a group of 8 different participants. The group8 were first given 8 of The Beano comic stills in a no memory (stills present) condition. They were then given 8 T&J stills in a memory (no stills) condition.

As the conditions of memory and cartoon type were across two different sized groups, this gave a mixed group design for these two conditions and a fully between group design for exploring group size.

4.3 Coding

4.3.1 Gesture coding

After viewing the video recordings, only the bird's eye view footage was used as this captured all participants and their gesture space best. This footage was converted into the .mov file format using Pixela image maker. As the aim was to have gesture coding categories emerge from the data set, the video footage was first viewed so that gestures could be observed and annotations of these gestures made.

Observations at this stage were of both gesture form (eg. hand configuration, movement and trajectory) as well as on the typical function (eg. whether the gesture conveyed content information and was iconic, metaphoric, or deictic, or non-content information and was a beat or social gesture) associated with the gesture. To treat gesture as an independent sign, gestures were defined in terms of what they referred to rather than in terms of what the speech referred to. Like the methodologies of Beattie & Shovelton (2002) and Bavelas et al (1992; 1995; 2007), gestures were linked to a referent in the story content so as to treat the task as a referential communication task specific to gesture. Like Bavelas et al (1992; 1995; 2007), non-content gestures were treated in the same way and were linked to a social referent, namely the addressee. Treating gesture as an independent sign was important for the reasons discussed in chapter 2.

The pioneers of gesture research propose different types of methodology for analysing gesture with some using form analyses (see Kendon, 2004 for a review of his methodology) and others using functional analyses (McNeill, 1992). Since both categorise gesture on the basis of a subjective judgement about what constitutes a movement, hand shape and/or indeed function, there are issues with both of these methodologies. The main concern with form analyses is how form descriptions can be explained in a functional framework. With functional analyses it is that speech biases the gesture interpretation. Treating gesture as an independent sign addresses the issue of whether analysis should be based on form or function. For any sign, both the form and function must be considered in order to define it. That is, when coding a sign, the functional definition is dependent on the formal signification. Here, although gesture categories were functionally distinct, they were based on the gesture form. In treating gesture as a sign then, this pilot addresses the methodological shortfalls that separate form/functional types of analyses can encounter.

A related methodological issue is whether gestures should be analysed with the speech signal on or off. The current trend in form analyses is to analyse gestures in the absence of speech (Kendon, 2004). In functional analyses it is to analyse gestures by linking them to their associated speech on a precise timeline (McNeill, 1992). The issue of speech being on or off is especially important for functional analyses where it is of concern that speech biases the interpretation of the gesture. Here, the possibility of speech biasing the gesture interpretation was eliminated by following the methodology of Beattie & Shovelton (2002) and linking gestures to references in the stills, rather than to references in the accompanying speech. What's more, gesture coding was always carried out before speech coding. Gesture coding was therefore independent from speech coding, which as will be discussed further in the speech coding section, was only ever carried out to a superficial level. Having addressed these methodological short falls, it was considered preferable to code gestures with the speech signal on. This coded gestures within their overall context, which may or may not have included speech, whilst treating them as independent from other signals in the communication. Considering the overall context improved the clarity of gesture coding. As a result, all gestures were coded and very few were omitted due to problems with classification.

A second stage of coding more formally identified gestures. Coding at this stage was again based on both function and form with analysis following along the lines of McNeill (1992) in looking for the semantics of the gesture. Like McNeill (1992), the first step in coding a gesture movement was to identify a meaningful kinetic movement. That is, gestures carrying meaningful information (eg to convey the idea of head scratching) were distinguished from other movements that performed some other function such as scratching the head to relieve an itch or to self adapt to the social environment. In this first pass of coding, gestures were coded in the data set by identifying the relevant parts of what has been called the 'gesture phrase' or G-phrase (Kendon, 2004**).** Several movements are involved in the G-phrase. The preparatory phase is where the limb moves from a resting position into the gesture space in preparation for the stroke. The stroke is the peak of the gesture movement and is where meaning is expressed. There can then be retraction, where the hand returns to a resting position which may be the same or different to the position held before the G-phrase began. In between the preparation and stroke, there can be a pre-stroke hold or a longer hold and in-between the stroke and retraction, a post-stroke hold. The stroke is the only necessary movement for the G-phrase. In this data set, most gestures consisted of a preparation and stroke phase but none had a retraction phase. This was most likely due to gestures being produced in quick succession with no break between them. Where a gesture involved a two-way movement, such as up and down, this was coded as one gesture. Movements such as these are treated as one gesture because the two movements belong to the same phrase (McNeill, 1992).

During this initial coding pass, any gesture considered to be conventionalised, or an emblem as they have been called (Eckman & Friessen, 1969), was filtered out (see chapter 3 for a discussion about emblems in relation to spontaneous gestures). An example from the English language would be the ok sign. Emblems were treated as distinct from spontaneous gestures as the aim of this study was to explore spontaneous gesture. The identification of emblems was based on their familiarity to the coder in terms of their use as conventionalised signs in English language communication. Filtering out conventional gestures left a data set that consisted of only spontaneous gestures. No other gesture type occurred in this data set probably because of the task focus.

Next, spontaneous gestures were classified according to the kind of information they conveyed. Gestures were classed as content gestures when they explicitly communicated information from the stills and as non-content gestures when they did not communicate such information. As in the earlier stage of coding, the methodologies of Beattie & Shovelton's (2002) and Bavelas et al (1992; 1995; 2007) were followed and gestures were linked to a referent in the story content or to an interactive aspect of the communication. With gestures coded as either (1) content or (2) non-content gestures, categories were like the topic and interactive gesture categories in Bavelas et al (1992; 1995; 2007). Although this was a rather crude scheme, these two categories were considered sufficient for this initial study of group communication and test of the current experimental design. Moreover, these two categories fit with findings in dialogue as well as in monologue. Having fewer categories at the outset allowed a coding scheme, which was expected to consist of additional sub-categories, to emerge from the data set. It should be noted that the reason for naming the categories content/non-content gestures rather than topical/interactive gestures like Bavelas et al (1992; 1995; 2007) was to make fewer assumptions at the outset.

Notes taken at the first stage of gesture coding would illustrate the gesture sub-categories that emerged within the two main categories so that the scheme could be extended to include these additional categories in the actual experiment. The experimental coding scheme would then be constructed from the gesture sub-categories.

Coding was for all members within G5 and all members within G8 who were distinguished by the labels p1-p5 and p1-8 respectively.

In this pilot, definitions and means of identifying gestures were based largely on McNeill (1992) and Kendon (2004) but were extended by turning to methodologies that contextualise gesture. These were the methodologies of Beattie & Shovelton (2002), who ground gesture within the story content and Bavelas et al (1992; 1995; 2007), who do so in relation to the social context.

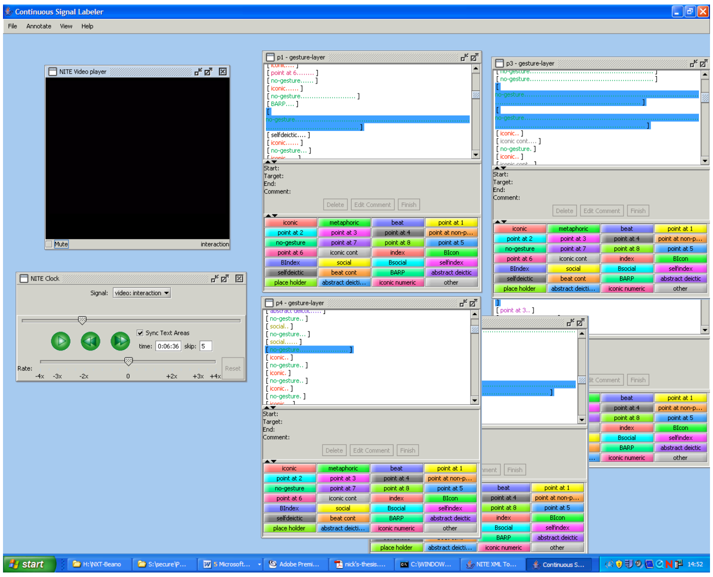
4.3.2 Speech coding

Once gesture coding was complete, the speech signal was coded as either on or off. Speech was coded as ‘on’ when it was a speech turn. That is, when the speaker held the floor. This definition of a speaking turn is based on Sacks, Schegloff and Jefferson (1974) and the modification of the original definition by Fay (2000) to suit group communication. Fay’s (2000) modification treated back channel responses, simultaneous speech and irrelevant chatter differently and excluded it from the turn analyses. Here then, speech outside of a turn that did not gain or hold on to the floor, was coded as the default speech off. Speech coding was for all members of G5 and G8 who were distinguished by the labels p1-p8 as described in the gesture coding section.

4.3.3 Coding Tool

The coding scheme and actual coding were managed within Nite (Carletta, Evert, Heid and Kilgour, 2005). This is a flexible programme that affords alterations to be made to coding schemes as categories emerge and is therefore ideal for the data driven analyses reported here. Gesture/speech categories were first defined within the programme. Once defined, the beginning and end of every gesture/speech event was coded. This was done in real time by mouse clicking on a pre-defined gesture/speech event which was housed in a box on the interface (see screenshot in Figure 4.1). As the screenshot shows, events were coded along a vertical timeline rather than along the horizontal. Coding is on a vertical time line because the programme is tailored towards the frequency of event coding at a macro level, an additional feature that suited the research questions of this pilot.

Figure 4.1: Screenshot of NXT for pilot coding



Parts of the analyses were also conducted in NITE, which has an inbuilt query language for counting gesture frequencies within the programme. This query language consists of a number of variable declarations that can be constrained in various ways. Examples of queries are;

How many events were 'iconic' and by 'person 4'?

($g gest) ($t gesture-target): ($g > $t) && ($t@name = "iconic") && ($g@who = "p4")

Likewise all events with the attribute speaking by a particular person (in this case person 1) can be extracted.

($s spks) ($a sp-att): ($s > $a) && ($a@name = "speaking") && ($s@who = "p1")

With queries like these, the frequency of gesture events were counted within the program to get gesture frequency counts. They were then extracted to an excel spreadsheet for further analyses. Similarly, start and end times of speech events were extracted and exported to an excel spreadsheet where the duration of speech was then calculated.

4.4 Analyses

4.4.1 Gesture and speech patterns

As many of the original comic stills were missing from this task, there was no optimum task solution. The basic aim of the experiment was to generate a natural and free flowing conversation with gesture that would reference the comic stills. Analyses focussed on the type of gesture that emerged alongside the amount and distribution of both gesture and speech. All of this was considered in relation to the interactional context and, more specifically, to the level of interaction in the groups.

Content and non-content gestures were counted to give nominal frequencies for every participant in the group and speech turns were added up. The total amount of gesture and speech alongside the interaction time was then reported for every group.

To explore the relationship between gesture and speech, overall amounts of speech and gesture were correlated for every group member. This involved correlating the overall amount of speech with overall amount of content gestures and overall amount of speech with overall amount of non-content gestures.

Gesture frequencies were transformed into the ordinal measure of rate per minute of speech. This was done by dividing each participant's gesture frequency by their total amount of speech. Participant's average rates per minute were then averaged across conditions to give the mean average rate for the group of interest. Rates were averaged within the groups to explore any effect of group size on gesture rates. They were averaged between groups for both the memory and cartoon conditions.

A central aim of the pilot was to establish the best measures for analyses therefore additional measures were considered. Gesture rate per speech turn is a measure that can account for any differences in the rate of speech. However, since speech turn rate is not a direct measure of the amount of speech, the speech turns must be of a similar length across the groups of interest for this to be a viable measure. This measure was rejected as G5 was expected to have shorter speech turns, since these are indicative of a dialogue style, compared to G8, where a monologue style with longer turns was expected.

A measure that addresses the issue of variable speech rates is gesture rate per amount of words, say 100. This measure was not calculated since the aim was to focus on the gesture channel and to plough analytical resources into that. Moreover, the gesture rate by total amount of speech is reported to be as good a measure as rate per number of words (Bavelas et al, 2007).

The additional measure used to explore gesture rates was the gesture rate per minute of interaction time. To get this, gesture frequencies were divided by the total time taken to complete the task (total interaction time) for every participant in the group. These rates per minute were calculated and then averaged across conditions. Additional arguments for using this measure are that gestures could, and often did, occur outside of the speech turn. In addition, speaker back-channel responses, where gestures could also occur, were not coded as speech here.

4.4.2 Alignment and emergence of conventions

Occurrences of gesture alignment were noted and then described.

4.5 Results

Gesture Categories

Although coding consisted of two broad categories (content/non-content), sub-categories of gestures within these categories were observed so that they could be used to build an extended coding scheme for coding experimental data at a later stage. Like the content/non-content categories, gesture sub-categories were identified based on the descriptions outlined in chapter 1. Gesture sub-categories are described and examples given within the document.

For gestures conveying content information several sub-categories emerged. These were iconic, metaphoric, deictic, and place holding gestures. Following McNeill (1992), iconic gestures were imagistic gestures that referred to concrete entities and following Beattie & Shovelton (2002), that communicated semantic information about the still content (rather than communicating information about the speech content as in McNeill, 1992). An example of a stream of iconic gestures by one group member is presented below;

4.1 He's like [burst into] his room …
Iconic: hand in flat upright position sweeps forwards with some force.

4.2 … looking really ticked off [holding a chicken] … and …
Iconic: hand goes to right hip of gesturer as though holding a chicken under arm.

4.3 … and … he's got .. he's got his [belt round his waist] …
Iconic: hand outlines round gesturer's waist.

4.4 … [with the key hanging off it] and uhm …
Iconic: hand moves up and down right side of hip to indicate keys hanging.

4.5 … jerry's like [hiding in this sort of pink vase] in the middle of the room …
Iconic: flat hand moves outwards and makes circular outline of vase some distance in front of gesturer.

As can be seen from the descriptions above, the iconic gesture category could be broken down further as iconic gestures were either of a character (Cvpt) or observer (Ovpt) viewpopint. In the above stream of gesture and speech, examples 4.1 and 4.5 are Ovpt whereas all other examples are Cvp gestures.

Metaphorical gestures conveyed the same kind of information as iconic gestures but the underlying concept was abstract. There were very few metaphorical gestures in this data set. It should be noted that conduit metaphors, referencing the on-going discourse (McNeill, 1992), were not included here since they do not reference the story content. As conduit gestures refer to non-content information, they are discussed in the non-content section below. An example of a metaphoric gesture from the data set is presented below;

4.6 The [posh guy] with the …
Metaphoric: circular motion made with forefinger in front of the gesturer's nose.

Iconic and metaphoric gestures convey content information and they do so imagistically but information about story content was also conveyed in a non-imagistic way. For example, pointing gestures indexed and/or located meaning by referencing the gesture space. Although deictic gestures can be concrete, when referring to an actual entity present in the gesture space, they are most often abstract, when referring to an entity in the gesture space not actually present (McNeill, 1992). In line with this, most deictic gestures in this data set were abstract. An example of a deictic gesture is given below;

4.7 [He] was sitting on the grass …
Deictic: point to abstract space on Rh side of gesturer.

Note that the deictic gesture occurs on an anaphoric speech reference and so, like the speech, functions as a reduced expression. This deictic gesture was followed by two iconic gestures with different viewpoints so these have been included as further examples of iconic gestures with specific viewpoints;

4.8 … as if he was [pulling on boots]
Cvpt: both hands clasped as though holding the tops of boots and oriented down towards gesturer's legs then pulled upwards and back as though putting boots on.
note: there is also an accompanying leg movement with this Cvpt gesture.

4.9 … and there's [bushes and stuff behind him]
Ovpt: open and flat hand and waves outline of bushes.

A sub-set of deictic gestures, known as place holding gestures, also occurred. Place holders performed the same function as other deictic gestures but the hand configuration differed in that it indexed a larger portion of the gesture space (eg some or all of the fingers as opposed to a singular index finger). A place holding gesture from the data set is illustrated below;

4.10 It was [denis the menace] wasn't it?
Place Holder: open hand with fingers in claw shape held at a downwards angle as though holding the place of denis in the hand.

As well as making reference to the still content, deictic gestures referenced other group members by pointing at them. In this case, the deictic gestures were concrete. However, the function of these gestures was not clear because when points referred to a group member they seemed to perform two very different functions. At times, they directly referred to group members. At others, they treated group member as a bit of information and pointing was used as a means to order this information as though each person was a still. In the latter case, where group members were treated as bits of information to be ordered, there seemed to be two different functions. Pointing could have an intrapersonal function, where it facilitated the gesturer's thinking processes, or have a more interpersonal function, where it was for the thinking process of the group as a whole. In addition, these gestures usually only occurred in bursts at the end of the communication and, when they did, tended to occur simultaneously being made by several group members at once. Two examples of these pointing gestures are given below. In the first example (see 4.11), the speaker refers to specific content in the stills when pointing at the other group member whereas in the second example (4.12), speakers refer to group members as though they were bits of information to be ordered.

4.11 Is it [you that has onions]? Oh, its [you].
Concrete point: points at one person and then another.

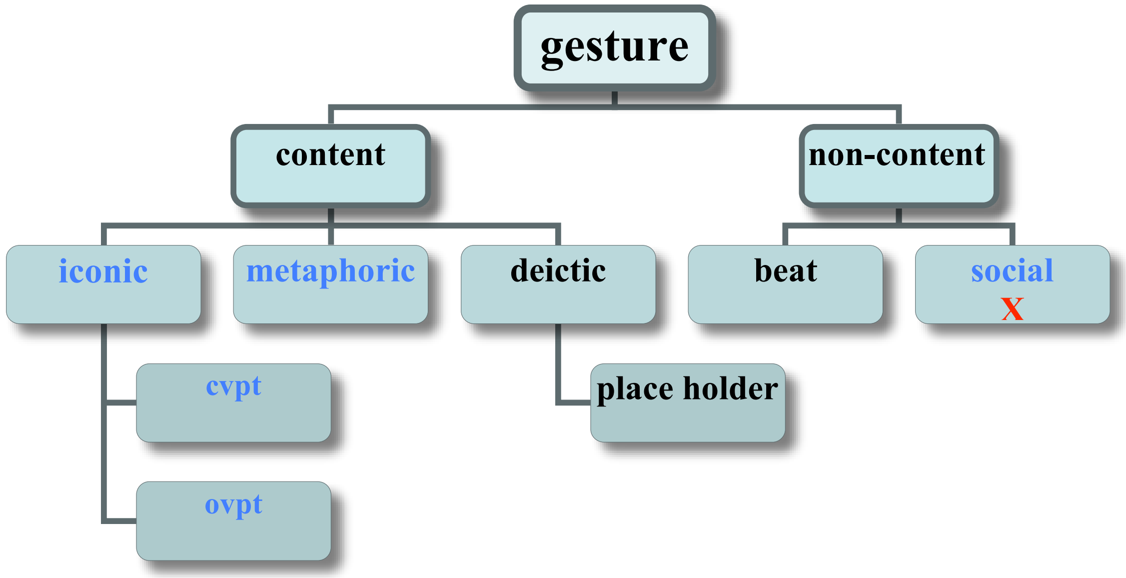
4.12 It's your first then …
Concrete points: numerous points for ordering the stills

The pointing gestures described above are non-content gestures but are different to the interactive gestures described by Bavelas et al (1992; 1995; 2007). Interactive gestures include beat and social gestures (which include McNeill's (1992) conduit metaphor) and are said to reference addressees in particular ways. Although social gestures did not occur in this data set, beat gestures did. As described in chapter 1, beat gestures are also considered to have a social function in dealing with organisational aspects of the discourse (McNeill, 1992). They can for instance act as markers of the communicators thinking, say by marking given information. Like social gestures then, beats perform a social function by managing the social interaction (Bavelas, 1992; 1995; 2007). Interestingly, despite not exclusively filtering for beats, as only one non-content category emerged this category was in essence a beat filter. McNeill (1992) uses such a filtering technique for coding beat gestures. An example of a beat gesture is given below;

4.13 … looks kind of [shocked] to see them
Beat: hand flips out then retracts.

All of the gesture categories that occurred in the data set are depicted in a Mind Map (see figure 4.1). This is with the exception of points towards other group members which were not included due to their uncertain function, their timing overlap and uneven distribution. In a tree structure, the branches of the map show the decision making process of the coder when deciding where any given gesture should be placed. As well as the content/non-content gesture classification and sub-categories within each, Mind Map1 illustrates whether gestures were imagistic or not with imagistic gestures depicted in blue and non-imagistic gestures in black. In addition, Mind Map1 shows the perspective of iconic gestures. Interestingly, gesture sub-categories fit with McNeill's (1992) scheme, which is based on monologue styles of communication. For example, with the exception of beats no other interactive gestures emerged as there was little or no evidence of the social gestures observed in dialogue by Bavelas et al (1992; 1995; 2007). Although absent in this particular data set, social gestures were included in Mind Map 1 but are identified by the symbol X to indicate their absence in this data set. As sub-categories fit with monologue styles of communication more than dialogue styles, social gestures were included, with a question mark so to speak, to ensure coding in the experimental data would fit interactive communication.

Figure 4.1: Mind Map of gesture sub-categories emerging from the pilot.



The scheme that emerged here differs from previous schemes. For instance, the metaphorical gesture category adopted by McNeill (1992) and carried in to the dialogue context by Bavelas et al (1992; 1995; 2007) is broken down more precisely. Rather than having one metaphorical category with sub-categories of metaphorical types like the conduit (McNeill, 1992) or, classing social gestures as metaphorical whilst claiming they directly reference addressees (Bavelas et al, 1992; 1995; 2007), these gestures were treated purely as non-content/social gestures. They were treated as social and not as metaphorical because, when they did occur, they directly referenced others in the communication. Classifying them as metaphorical would therefore have been misleading.

In addition, social and beat gestures were treated separately. This contrasts with the earlier studies of Bavelas et al (1992; 1995; 2007) where social and beat gestures formed one 'interactive' gesture category. They were separated out here because social gestures more evidently reference an addressee than beat gestures. Both beats and social gestures highlight or reference bits of the discourse (like what you said) but social gestures also reference where the information has come from (an addressee) or is to go to as they are oriented towards addressees whereas beats are not oriented in this way. Another difference is that, social gestures are imagistic whereas beats are not.

Other important features of gestures in the data set.

Gestures involving two-way movement, such as up and up, were coded as one gesture because the two movements belong to the same gesture phrase (McNeill, 1992). An example of a gesture with a two way move is given below;

4.14 It's as if he's trying to go [up the stairs] …
Ovpt: flat hand moves up one step then up another.

Conventional gestures or emblems were removed;
4.15 Emblem: hand slaps head twice making circular motions in-between.

Total amount of gesture

The total amounts of gesture (see table 4.1) were very different in G5 but were comparable in G8. Content gestures in G5 T&J (no-mem) = 13 whereas in G5 Beano (mem) = 74. Non-content gestures in the G5 T&J (no-mem) = 5 whereas in G5 Beano (mem) = 16. Both G8 T&J (mem) and G8 Beano (no-mem) had exactly the same amount of content gestures = 127. Non-content gestures in G8 T&J (mem) = 36 whereas in G8 Beano (no-mem) = 32.

Table 4.1: total amount of gesture (frequency)

condition	content gestures	non-content gestures
G5 T&J (no-mem)	13	5
G5 Beano (mem)	74	16
G8 T&J (mem)	127	36
G8 Beano (no-mem)	127	32

Total amount of speech (duration)

The total amount of time spent speaking (see table 4.2) differed in all groups across and between group size with speaking duration in G5 T&J (no-mem) = 4.4 mins; G5 Beano (mem) = 7.6 mins; G8 T&J (mem) = 14.1 mins and G8 Beano (no-mem) = 21.6 mins. Speaking time was longest for the Beano comic in both groups irrespective of the memory condition.

Table 4.2: total amount of speech (duration)

condition	speech (mins)
G5 T&J (no-mem)	4.4
G5 Beano (mem)	7.6
G8 T&J (mem)	14.1
G8 Beano (no-mem)	21.6

Interaction time (duration)

The total duration of the interaction was shorter in G5 (7.15 mins/5.2 mins) than in G8 (17.34 mins/28 mins) and in the T&J cartoon than in the Beano. In G5, the no-memory condition (with still) was shorter than the memory condition (without still), whereas in G8, the memory condition (without still) was shorter than the non-memory (with still) condition. Therefore, the shortest time taken to complete the task was in G5 T&J (no-mem), the same group that had the lowest amount of gesture.

Amount of speech (as duration) correlated with amount of gesture (as frequency)

Descriptive Statistics

The Shapiro-Wilks (W) test of normality results for gesture rate by total amount of speech and speech durations are presented in table 4.3. As can be seen from the table, the data were mixed in terms of being normally and non-normally distributed. Speech in G5, Beano (mem) and in G8, Beano (no-mem) had non-normal distributions as did content gestures in G5, Beano (mem) and G5, TJ (no-mem). All other distributions were normally distributed.

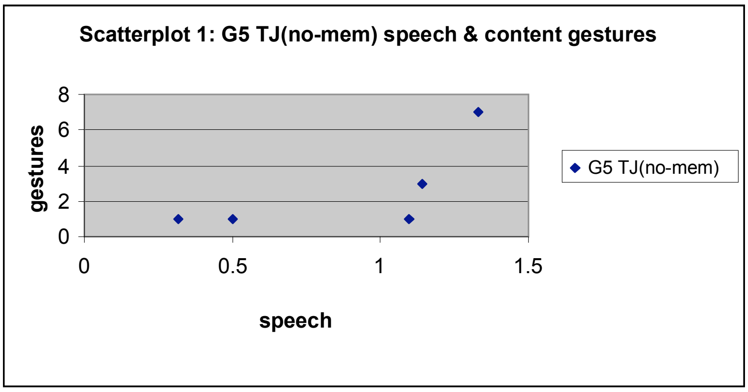
Table 4.3: Shapiro-Wilks (W) Tests of Normality

	Df	W	Sig <.05
spch G5Beano(mem)	5	.682	.006
spch G5TJ(nomem)	5	.877	.294
spch G8TJ(mem)	8	.860	.230
spch G8Beano(nomem)	8	.767	.042
content, G5Beano(mem)	5	.762	.039
contentG5TJ(nomem)	5	.735	.021
contentG8TJ(mem)	8	.870	.268
contentG8Beano(nomem)	8	.963	.827
non-contG5Beano(mem)	5	.962	.823
noncontG5TJ(nomem)	5	.883	.325
non-contG8TJ(mem)	8	.950	.740
non-contG8Beano(nomem)	8	.914	.492

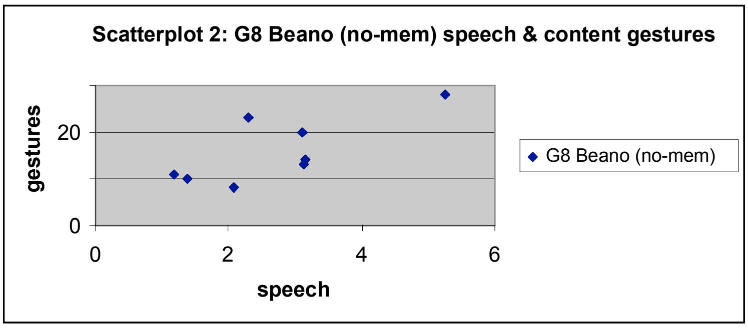
Correlations

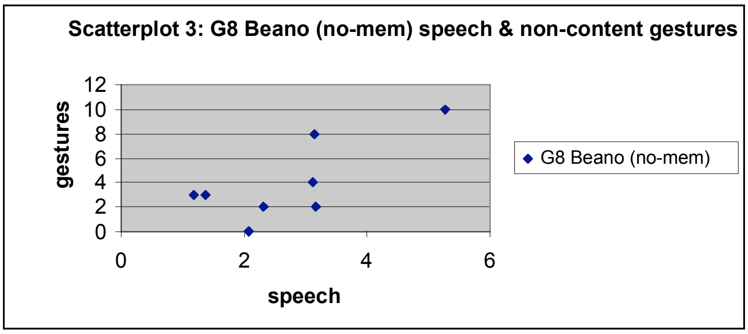
In G5, Beano (mem) both Pearson's and Spearman's correlation coefficients found no significant relationship between speech and content gestures or speech and non-content gestures.

In G5, TJ(no-mem), Spearman's correlation coefficient found a significant relationship between speech and content gestures ($r = .894$, $p<.05$) only. Scatterplot 1 shows this relationship to be a positive one.

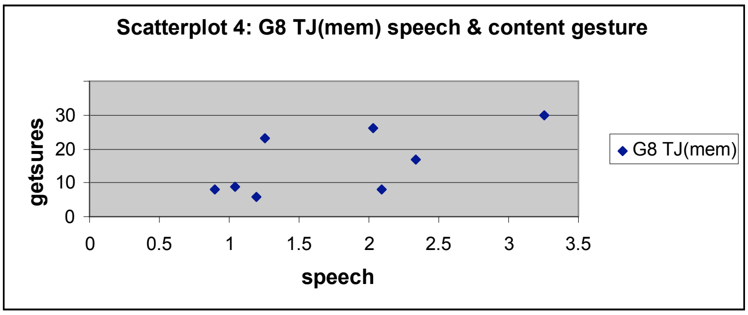


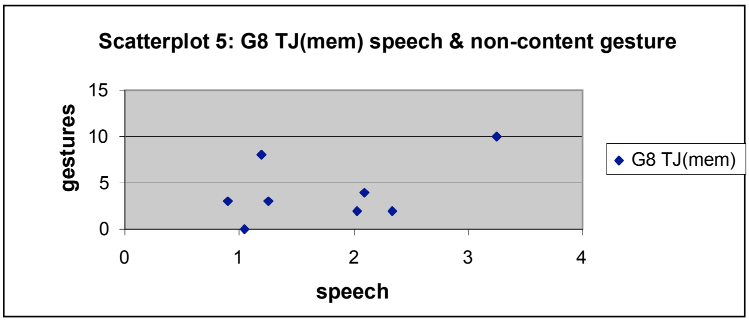
In G8 Beano (no-mem), Pearson's correlation coefficient found a significant positive relationship (see scatterplot 2) between the amount of speech and amount of both content ($r = .751$, $p<.05$) and non-content gestures ($r = .736$, $p<.05$). Speech and content gestures ($r = .690$, $p = .058$) were also positively correlated on Spearman's correlation (see scatterplot 3)





In G8, TJ(mem), Pearson's correlation coefficient found the relationship between speech and content ($r = .668$, $p =.07$) and speech and non-content gestures ($r = .669$, $p =.07$) to be marginal and just missing significance. Scatterplots 4 and 5 show this relationship to be a positive one.





Gesture Proportions

Looking at the proportion of gestures across the content and non-content gesture categories (see table 4.4), content gestures accounted for most gestures with non-content gestures accounting for a much smaller proportion. Both G8 had comparable proportions of content/non-content gestures with G8, TJ, mem = 77%/23% and G8, Beano, no-mem = 82%/18% whereas proportions in G5, Beano (mem) = 83%/17% and in G5 T&J (no-mem) = 64%/36%. G5, Beano (mem) was then comparable to G8 but the low interactive G5 T&J (no-mem) showed a different pattern of proportions to all the other groups. That this group had different content/non-content proportions within the G5 and across G8 fits with the gesture rate findings in suggesting there was a problem with the communication of this group.

Table 4.4: Proportion of Gestures by Type

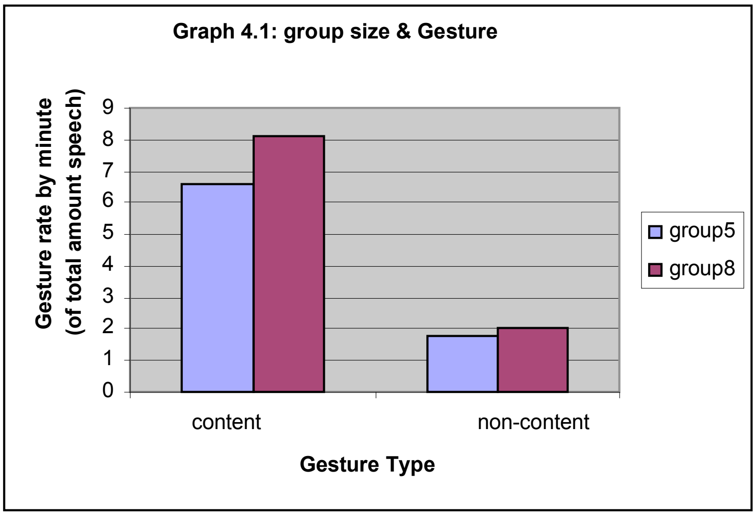
Group	content gestures	non-content gestures
group5, TJ (no-mem)	64	36
group5, Beano(mem)	83	17
group8, TJ, (mem)	77	23
group8, Beano, (no-mem)	82	18

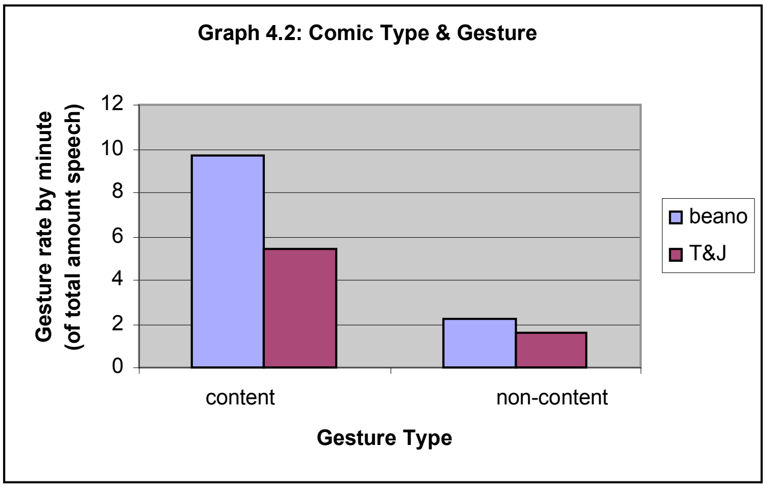
Gesture rate per minute of speech

As can be seen from graph 4.1, both content and non-content gesture rates per minute of total amount of speech were higher for G8 (Av = 8.2/2.02) than for G5 (Av = 6.6/1.78).

As can be seen from graph 4.2, content gesture rates per minute of total amount of speech were higher for The Beano cartoon than for the T&J cartoon (Av = 9.69/5.39). Similarly, non-content gesture rates per minute of total amount of speech were higher for the Beano cartoon than for the T&J cartoon (Av = 2.26/1.60).

Rates for the memory condition were mixed across the two gesture categories. Average content gesture rates were higher for the memory condition, when no still was present, than for the non-memory condition, when stills were present (Av = 8.3/6.75). Non-content gesture rates were higher for the non-memory condition than for the memory condition (Av = 2.07/1.78).



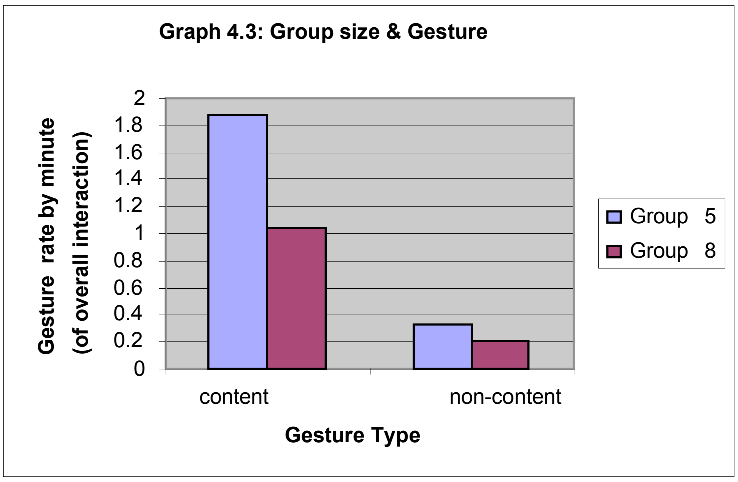


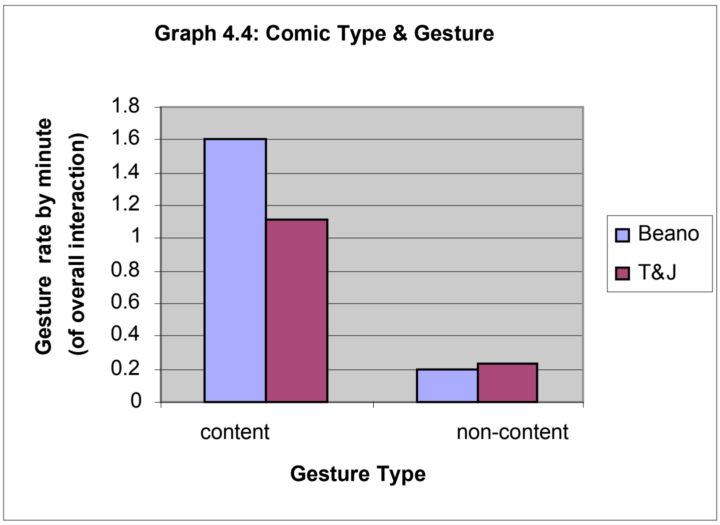
Gesture rate per minute of interaction time

As can be seen from graph 4.3, both content and non-content gesture rates per minute of the interaction were higher in G5 (Av = 1.87/.32) than in G8 (Av = 1.04/.20).

As can be seen from graph 4.4, gesture rates per minute of the interaction were higher on the Beano cartoon than on the T&J cartoon (Av = 1.60/ 1.11), whereas non-content gesture rates were comparable (Av = .20/.23).

Content gesture rates per minute of the interaction were higher in the memory condition, with stills absent, than in the non-memory condition, with stills present (Av = 2.08/.63). The pattern was similar for non-content gestures and more were produced in the memory than in the non-memory condition (.33/.16).





Alignment and emergence of conventions

There were instances in the data set when group members referred to the same entity. Reference to the same entity could occur when a group member referred back to what another member had spoken about or, as some of the entities were the same across stills, when a group member referred to an entity in their own still that another had mentioned from their own. In this data set, it most often occurred in the later case. Gestures referring to the same entity were often similar in form. This is illustrated in the following example (4.16), where p1 first mentions a pink vase and then p2 mentions it but p2 uses a different gesture form. Since both p1 and p2 are referring to entities in their own stills (Jerry is in a pink vase in both), p1 clarifies this is the same pink vase and when he does, uses the same gesture form as p2;

Example 4.16
p1: … jerry's like [hiding in this sort of pink vase] in the middle of the room …
Ovpt: flat hand moves outwards and makes circular outline of vase some distance in front of gesturer.

p2: Jerry is in the pink vase …
Ovpt: hand in point shape makes circular outline of vase in front of gesturer.

p1: The vase with pink flowers on it?
Ovpt: hand in point shape makes circular outline of vase in front of gesturer.

Whereas example 4.16 is an Ovpt gesture, the example below (4.17) shows alignment on the gesture form of a Cvpt gesture;

Example 4.17
p1: [with the key hanging off it] and uhm …
Cvpt: hand moves up and down right side of hip to indicate keys hanging.

p2: … he's got the key stuck on his belt.
Cvpt: hand moves up and down right side to centre of hip to indicate keys hanging.

4.6 Discussion

This pilot investigated a novel way of exploring gesture. With empirical research being on one-person monologues and two-person dialogues or at most on triads, the aim of the pilot was to adjust methodologies found to work for these interactional contexts so as to apply them to the group setting. As discussed in chapter 3, experimental groups can be manipulated to elicit styles of communication similar to monologue and dialogue (Fay, 2000). Groups therefore provide an additional setting for investigating effects of the interaction on communication and, as findings reported in earlier chapters show, considering the interaction is vital in studying gesture or indeed any sign.

The interactive group setting is also an additional context within which to consider current gesture coding schemes. The pilot therefore tested the application of established gesture coding schemes to groups, whilst treating the interaction level as an influential factor.

The coding methodology in the pilot treated gesture as an independent sign by linking gestures directly to a referent. Importantly, it was recognised that referents could be in the story content (as in the case of content gestures) or be in the social context (as in the case of non-content gestures). Defining gestures in terms of a referent was considered to be of the utmost importance since defining gesture in this way highlights the variety of forms and functions gestures have. This is of particular significance for non-content social gestures as when linked to a referent they are observed to reference an addressee (Bavelas et al, 1992; 1995; 2007) but is also important for content gestures as, when attention is paid to the semantic features these gestures convey, gestures with a Cvpt are found to convey more information than those with an Ovpt (Beattie & Shovelton, 2002). As well as being defined in terms of a referent, the way in which gestures conveyed information was also considered (ie whether the gesture was imagistic or not).

As this was a pilot rather than an experiment, no predictions were made. However, on the basis of prior research and dialogue theories of communication, in particular the alignment model, it was expected that smaller groups would have a dialogue style of communication whereas larger groups a monologue style (Fay, 2000). It was expected then that more content gestures and speech would be produced in the monologue style of the larger groups, whereas more non-content gestures would be produced when in the dialogue style of the smaller groups. Expectations for specific gesture types within these categories were also framed within prior research findings.

Although the coding scheme was somewhat crude, the intention was to extend the scheme based on what emerged from the data set. Additional sub-categories did emerge. For content gestures these were iconic, metaphoric and deictic (including place holding) gestures however, metaphorical gestures occurred infrequently. In addition, iconic gestures could be broken down by perspective and so, Cvpt and Ovpt gesture sub-categories also emerged.

Non-content gestures however consisted solely of beats as no social gestures occurred, at least not frequently enough to identify them as a separate category. That social gestures did not emerge as a category was surprising, since these gestures are predominant in dialogue and moreover when dialogue has a visible component (Bavelas et al, 1992; 1995; 2007). They were therefore expected in G5, which should have been like dialogue. That a social gesture category did not emerge in any group suggests the group communication was more like monologue than dialogue. This is in line with subsequent analyses which also showed groups to have a monologue style of communication. Such a scheme would not then be appropriate for a dialogue style of communication. This highlights a problem of generating coding schemes from the data set alone. It suggest that if the interactive context is not considered and analysed alongside gesture production, it may lead to an under inclusive scheme. For this reason, social gestures were included in Mind map 1 to be considered as a category for coding in the group experiment. As Bavelas et al (1992) suggest, social gestures may be present in monologue, and so monologue styles, but they are likely to be of poor form and so difficult to identify.

As this was a pilot it was not appropriate to draw strong conclusions from the analysis. This is even more so the case if the interactions were all like monologue as is suggested by the coding scheme that emerged. Nonetheless, the data was explored further to investigate the kind of communication that emerged. Overall amounts of gesture and speech, alongside the length of interaction, are a first measure of how interactive a communication is. Expectations rather than predictions were then as stated earlier in this discussion. Although, as expected, more content gestures were produced in G8, non-content gesture amounts were also higher in G8. Comparing amounts within groups of the same size, G8 had comparable amounts of content and non-content gestures regardless of cartoon type or stills being present or absent whereas G5 had different amounts. Amounts of gesture in G5 Beano (mem) were exactly half that of G8, whereas amounts of gesture in G5, TJ (no-mem) were extremely low and the lowest of all the groups. The extremely low amounts of gesture in G5, TJ (no-mem) suggest something other than the manipulated condition influenced the amount of gesture. In addition, finding G5 to have lower rates of non-content gestures than G8 suggests communication was less interactive and therefore less dialogue like in the smaller groups.

In line with the amount of gesture, interaction times for G5 were much shorter and particularly for G5, T&J (no-mem). As overall time taken to complete the task is an indicator of how collaborative a task is, this again shows communication was less interactive in G5 as compared to G8.

Amounts of speech were in line with content gestures and were lower in G5 than in G8. However, amounts of speech were variable across all of the groups, were particularly low in G5 and especially low in G5, T&J (no-mem). Although less speech was expected in G5, such low levels of speech again suggest a low level of interaction in these groups and especially in G5, T&J (no-mem).

To explore gesture-speech relations and whether gesture could be taken as a measure of the communication style, amounts of speech were correlated with amounts of gesture. If speech and gesture were correlated gesture, like the speech in Fay (2000), could be taken to measure the communicative style. Speech was correlated with content gestures in G8, Beano (no-mem) and in G5, TJ(no-mem) and with non-content gestures in G8, Beano (no-mem). That any of the gesture-speech correlations reached significance is a good result considering the low power in this study. Considering the small sample size, the results for content gestures and speech and non-content gestures and speech in G8, TJ(mem) should also be considered as they were highly correlated. On the basis of these correlations, gesture rates were reported to further explore communication in the different sized groups alongside effects of cartoon type and memory.

Gesture rates were compared within (group size) and between conditions (cartoon type/stills absent or present) on two different measures. On the gesture rate per minute of speech measure both content and non-content gestures were produced at a higher rate in G8. As before, although rates of content gestures were in line with expectations, the lower rate of non-content gestures in G5 indicated these groups were less interactive. However, the interaction time measure corrected for the low level of interactivity in G5, as both content and non-content rates were higher in G5 than in G8. Nonetheless, although higher non-content rates were expected in G5, higher content rates were not.

In addition to the problem with the G5 interaction, conflicting results for the rates per minute measures could have been because of imprecision in the content/non-content gesture categories. The content category consisted of; iconic (broken down by Cvpt and Ovpt), metaphoric, deictic and place holding gestures. As deixis is produced more often when visible (Bavelas et al, 2007), deictic gestures should be produced at a higher rate in dialogue where there is better visibility. In line with this, deictic gestures can function as a reduced expression and so, if there is reduction in the gesture sign, should again be more evident in dialogue. Yet, if there is sign reduction in dialogue (see chapters 1 and 2), iconic gestures should be produced at a higher rate in monologue. As deictic and iconic gestures are influenced differently then they should be in two separate categories. Similarly, since Cvpt gestures are produced at higher rates when visible and because perspective is found to determine how communicative a gesture is (Beattie & Shovelton, 2002), Cvpt and Ovpt should also be considered as separate categories for the actual experiment. This shows that the original categories were not discrete enough and that, although defining a gesture in terms of a referent is important, how gestures sign is just as important as what they sign.

The anomaly in G5, T&J (no-mem) was most likely an effect of the experimental design. It is possible that G5 did not have a sufficient amount of information to make for an interactive task. As reported in chapter 3, groups have optimal settings and a sufficient amount of information is one of them (Fay, 2000). As G5 only had X5 stills, collectively these groups had less information than G8. In addition, gesture production was lower on the concrete T&J cartoon as compared to the abstract Beano. This was especially so for content gestures which, in relating to the story content, were influenced by the change in it. Combined with too little information, the different story content can account for the low interaction level of G5 T&J (no-mem). Observations of video footage confirmed conversation in G5, T&J (no-mem) was stilted and constructing a coherent story proved difficult. They show the abstract Beano cartoon better engaged participants who had more interactive communication. It is important to consider observations alongside rate findings as the lower rates of gesture for the T&J cartoon could have been due to the anomalous interaction in G5, T&J (no-mem). Low interaction levels in two of the four and in this group especially, had a substantial effect since the pilot only consisted of four groups. However, observations suggest the cartoon type did have an effect.

This highlights the importance of providing an optimum amount of information, considering the kind of information offered and attempting to keep the information constant across groups of varying size. It suggests that in further experimentation, G5 should have more information and that an abstract comic, like the Beano, be used. A solution for the group experiment would be to increase the number of stills so that when in G5, every group member has X2 stills each whilst keeping the number constant for G8, so that they have X1 each. Although every group member would have an extra still when in G5, as a group they would have 10 stills whereas G8 would have 8 stills. Collectively then, overall amounts of information would consist of 2 stills more for G8 and any effect of this difference could be tested beforehand. In addition, a way in which to account for the variation in the amount of information would be to measure the gesture frequency as a function of number of stills.

Another design aspect was the memory component and whether participants had access to stills. Findings for this between group condition were mixed across the two gesture categories. Whereas content gesture rates were higher in the memory condition than for the non-memory condition, non-content gesture rates were higher for the memory condition on the rates per minute of interaction measure, but were higher for the non-memory condition on the rates per minute of speech measure. As reference to stills would be made when speakers were providing a lot of contextual information, this would disrupt the flow of content gestures and explains why content gestures were consistently affected by the memory component. Non-content gestures more often occur outside of presentation phases (like at the beginning and end of the phase) and so would be disrupted less. These gestures can also be more easily performed whilst holding stills. Interestingly, finding more gestures to be produced in the memory condition, when hands were free, differs from Bavelas et al's (2007) finding that whether hands were free had no effect. It may then have to do with the memory component rather than the hands being occupied. This fits with findings that show gesture facilitate lexical retrieval (Butterworth & Beattie, 1978; Krauss, 1998; Morrel –Samuels & Krauss, 1992), as this would be necessary in a more demanding task such as one with a memory component. Although this suggests an intrapersonal function for gesture, as many gesture researchers note, it need not be at odds with the interpersonal function of gesture in communication. However, as memory and cartoon were confounded factors, any conclusions drawn here are tentative. In addition, the anomalous group made for lower gesture rates in the non-memory condition. That the memory component worked in terms of the design of the experiment combined with the possibility of increased gesture production when there is a memory component suggests a group design with a memory component is preferable.

As there were issues with the design of the experiment that created problems in the G5 communication, a proper investigation of alignment was not feasible. However, there were instances of alignment and these were exemplified by way of short descriptions of the gesture and speech. Such instances of alignment provide tentative support for the alignment model's (Pickering & Garrod, 2004) prediction that, as well as changing in other ways, signs become more similar between those communicating.

The pilot was to test the design of the experiment and check whether a group experiment applied to the study of gesture was feasible. Findings from the pilot suggest the group design works for exploring gesture communication but that certain changes should be made to the original design. First, the amount of information needs to be increased in G5 whilst it should be kept as it was in the pilot for G8. A comic with highly imagistic and even abstract story content should be used and these should be memorised so that the communication proceeds from memory.

As the intention was to develop a coding scheme from the pilot and apply this to the group experiment, the pilot also checked the emerging gestures against current gesture coding schemes. Since communication in G5 was less interactive than expected, the gestures sub-categories that emerged were fitting with schemes based on monologue. For this reason, social gestures were included in the group experiment coding scheme.

Coding multi-party groups, especially when the coding scheme was uncertain, proved a difficult task. For this reason and because the scheme here was largely based on monologue, it is suggested that the experiment be modified for pairs of participants. A dyad experiment would acquire a baseline-coding scheme for interactive communication. It would also test the new group design for any effect of varying the amount of information, since different sized groups will have different amounts. A dyad experiment was therefore designed and is reported in the next chapter.

Chapter 5: Dyad Experiment

5.1 Introduction

To resolve problems encountered in the pilot, the group experiment was modified to give a new group design. In the pilot, when groups of 5 (G5) members received only one comic still each, they had too little information to establish a cohesive story. Having only one still each did not evoke a collaborative discussion in G5. In comparison, when groups of 8 (G8) members received only one still each, they had a sufficient amount of information to establish a cohesive story. This was evident from the collaborative discussion G8 had about the made up story. As discussed in the pilot chapter, a solution to this design problem is to give every group member two stills each when in G5, whilst keeping the still number constant, by giving every group member one still each when in G8. By this design, the amount of information each group has, as a collective unit, is almost equal as G5 receive ten stills in total and G8 eight, eight stills in total. However, as G5 will always have a bit more information, it is important to test whether this additional information influences the interaction before applying the design to different sized groups. The dyad experiment reported here, in a modified version of the planned group experiment, checked the design of the group experiment by testing whether varying the amount of information (from 8 to 10 stills) influenced group interactivity.

The dyad experiment also served as a baseline for gesture coding in interactive dialogue that could be applied to the group experiment. Such a baseline would avoid an additional problem encountered in the pilot, the risk of an under inclusive coding scheme emerging from the data set. For instance, as coding in the group pilot was based largely on monologue styles of communication, no social gesture category emerged. This is in line with research (Bavelas et al, 1992; 1995; 2007) that shows coding schemes based on monologue overlook the complexity, importance, and even existence of interactive gestures.

The new group design was then modified for pairs of participants in order to; (1) check the group experimental design and (2) acquire a baseline coding scheme for interactive dialogue (on this particular task). Depending on the still number and comic conditions, participants in a pair received either 4 or 5 static comic stills from two different beano cartoons (flower/book). This meant they received 8 or 10 stills collectively, the same number proposed for G8 and G5 respectively. The task was to memorise their own stills and then discuss them with their partner, with the aim of constructing a story from the shared information.

Memorising stills was the preferred method and so stills were memorised here. Although memory load was substantially greater than in the group pilot, any effect of memory would come out in the by conditions analyses. This therefore checked the memory component of the experimental design.

Gesture rates were calculated per minute of speech to account for the speech. Gesture frequency as a function of number of stills was also calculated to account for the different amounts of information. In addition to the two stages of coding in the pilot, a third and fourth level of coding was introduced here where the viewpoint of iconic gestures and gesture referent was tagged. As the gesture analyses were exploratory, to investigate any effect of information amount, no predictions were made for gesture rates/frequencies.

Gestural alignment was evident in the group pilot, albeit from informal observations. To explore gestural alignment in the pairs, gestures of the same type were correlated by their rates and frequency (corrected for the amount of information). This gave a crude measure of alignment on the rate and amount of gesture production. It was hypothesised that both gesture rates and gesture amounts would be highly correlated within pairs. This prediction was based on findings of alignment (see chapter 2), informal observations of gesture alignment in both the pilot of this thesis and earlier research (see chapters 2 and 3 for a review) and predictions of the alignment model (Pickering & Garrod, 2004).

It was also expected that gestures would align on form. To investigate alignment on gesture form, iconic gestures with the same referent were identified for qualitative analysis.

5.2 Methods

5.2.1 Subjects

24 participants took part in this experiment by signing up to an advert posted within the department of psychology at the University of Glasgow. Individual participants were then grouped into pairs to do the experiment. Two pairs were removed from the data set due to problems with the task. In both cases, one participant thought the other they collaborated with was a confederate and this influenced the interaction in a way that reduced contributions. After removing these pairs, additional pairs were run to get the required number of participants. From 13 female and 11 male participants, this gave; 3 male, 4 female and 5 mixed gender pairs. Participants were paid in course credits or cash depending on their preference.

5.2.2 Materials

20 static cartoon stills were used. These were chosen from two different Denis the Menace comic strips. Ten stills came from a comic strip featuring events based around the discovery of a smelly flower (from now on called ‘flower’ cartoon). The other ten came from a comic strip featuring Denis dreaming of a journey through the beano book with events based around the relevant characters (from now on called ‘book’ cartoon). Both story lines were fairly abstract. Comic stills were chosen on the basis of their content. Of particular importance was the relevance to the story theme and likelihood of eliciting gestures. They were not then necessarily in chronological order. Indeed some of the still sequences were missing. Stills were blown up to A4 size with verbal expressions blanked out. This was with the exception of lone onomatopoeic words (eg whiz; bang) that exaggerated events and in doing so encouraged the use of gesture.

5.2.3 Procedure

Participants were seated opposite one another. A camera fitted with fish eye lens was mounted to capture the bird’s eye view perspective of both participants. The camera was linked to a television to allow the experimenter to check both the participants and their gesture space were captured. The camera was set to capture the interaction at 25 frames per second (fps) with a frame size of 320 by 240 at 48, 000 Hz and 32 bit float in stereo.

Participants were given 4 (condition S4) or 5 (condition S5) stills from either the 'flower' or 'book' cartoons. For every participant, stills were randomly selected and jumbled so that still order did not equate with the way in which they were handed out. Participants were asked to place the stills face down on their knees whilst awaiting instruction. Participants were informed that their initial task was to memorise the still content. They attended to and memorised the still details in isolation so as to ensure they could only visualise their own stills. When the details of their own stills were memorised, they were to place the stills under their own seat. The rest of the experiment was then outlined to them. Participants were instructed to discus and order the stills into a cohesive story. To do this, they were to use all of the stills they had between them, placing them in any order, bearing in mind all stills were different and totally random in order. They were instructed to retrieve details from memory as far as possible but if this proved difficult, they could consult the stills. If there was a need to consult the stills, they were not to discuss or refer to the stills whilst looking at them. At this point they were to break from the conversation, only resuming the conversation when stills were once again placed under their seat. They were informed that because information could be missing, they should think of and use possible events outside of the information presented in the stills to link up the story. The task was complete once they found the best order for a cohesive story.

In this way, the whole procedure was repeated such that each pair did the task twice. This was to counterbalance the conditions of still number (4 or 5) and cartoon type ('flower' and 'book') across the pairs. In the second interaction then, two stills were always either taken away or added depending on the counterbalancing. The same two stills were always removed or added to ensure all pairs received the same story content. This gave a within design for both condition of still number and cartoon type.

5.3 Coding

5.3.1 Gesture coding

Video footage was captured in the .mov file format using Pixela image maker. The first two stages of coding followed the procedure detailed in the pilot study. This was; (1) an initial observational stage where the gesture form/function was observed and (2) coding with a formal scheme constructed from the pilot and initial observations in this data set. The formal coding scheme consisted of iconic, metaphorical, deictic, place holding (Ph), beat and social gestures. This was then a functional analysis that considered the gesture form in defining gestures (see pilot chapter for further details and discussion relating to treating gestures as independent signs). Importantly, gesture definitions were based on what the gesture referred to and how gestures conveyed this information. If difficulty arose in coding a gesture, the location of the gesture in the gesture space was considered (McNeill, 1992).

Expanding on the pilot, an additional third and fourth level of coding was introduced. At the third level, gesture view point was coded for iconic gestures. This was done by modifying iconic gesture events coded at stage 2 to either a character (Cvpt) or observer (Ovpt) viewpoint gesture. At a fourth level, iconic gesture events were linked to a referent in the stills. To do this, iconic gestures were given a descriptive tag of what the gesture referred to and still number (relating to the proper still sequence) the referent was in (eg S8 ‘denis head stuck in ground’).

As detailed in chapter 4, coding was based on the methodologies of Beattie & Shovelton (2002) and Bavelas et al (1992; 1995; 2007) as these methodologies directly link gesture to a referent in the story content or the social context. All gestures in all passes were coded with the speech signal on (see chapter 4 for a discussion as to why). Coding was carried out across both members of a pair who were distinguished by the labels p1 and p2 where p1 was the first to make a contribution in the first condition. These person labels were matched across experimental conditions so that p1 in S4 was also p1 in S5. This enabled the same participants to be distinguished for coding and analyses purposes.

Descriptive tags, alongside the still numbers they referred to, were filed within a word document at the time of coding. This word file would be returned to in subsequent analyses where gestures would be searched for by tags (see 4.3.3) and, once identified as having the same referent, used in a planned overseer experiment (see chapter 6). Examples of all of the gesture types in this scheme are given in the results section.

5.3.2 Speech coding

Once all four stages of gesture coding were complete, speech was coded as either on or off, in accordance with the procedure described in the pilot (see chapter 4). Speech coding was for both members of a pair who were distinguished by labels p1-p2, as described in the gesture coding section.

5.3.3 Coding Tool

Once constructed, the coding scheme, actual coding and gesture frequency counts were managed within Nite (Carletta, Evert, Heid and Kilgour, 2005). Once the gesture/speech codes were defined within the programme, the beginning and end of every gesture/speech event was coded. Coding was in real time by mouse clicking pre-defined events housed in boxes on the interface (see figure 4.1 in chapter 4 for screenshot). Being able to modify events at coding stage 2 (Cvpt/Ovpt) and add tags at coding stage 3 (still number/referent) highlights the way in which Nite can accommodate analyses as it progresses.

Once coding for all pairs was complete, queries within NXT were run to pull out the start and end times of the gesture/speech events. The speech queries were similar to that of the pilot (see chapter 4) but modified for pairs of participants. As well as being modified for pairs of participants, gesture queries were modified to declare additional gesture types (Cvpt, Ovpt, metaphoric, deictic, place holder, beat and social) and accommodate tags (still number/referent) that went with the iconic gesture events. Speech and gesture events were then counted. Gesture counts were for all of the gesture types. Queries were run on still number/referent tags to identify gestures referring to the same entity. More or less specific queries declared events for counting.

As in the pilot, the start and end times of speech events were also extracted and exported to a spreadsheet so that speech durations could be calculated.

5.4 Analyses

5.4.1 Gesture and speech patterns

The coding scheme generated in the dyads was a baseline scheme for the planned group experiment. This scheme would ensure group coding used an inclusive scheme. The coding scheme that emerged was then the first point of interest in the analyses.

Next, interaction times for S4 and S5 were compared.

As the aim was to extend the dyad design to G5 and G8 members, the dyad experiment also checked if varying the amount of information influenced the interaction level. Exploring the production of gesture and speech in relation to the condition of still number did just this. This involved counting gesture events for the different gesture types (with the exception of metaphorical gestures which were excluded from the analyses due to their infrequency) for every participant in all of the pairs. To get the total amount of speech (or speech duration), every speech event for every participant was extracted, and the end point subtracted from the start point, to give the duration of the speech event. These durations were then totalled to give the total amount of speech for every participant.

To explore the relationship between gesture and speech, overall amounts of speech and gesture were correlated for every participant within a pair.

Following on from the correlations, gesture rates were calculated to explore any effect of still number on gesture production. Here, gesture frequencies for every participant, within all twelve pairs, were transformed into rates per minute of total amount of speech by dividing the gesture counts by the total amount of speech. This measure was calculated to give the gesture rate per minute of speech for every participant in a pair. It was calculated for all of the gesture types coded within the iconic type. The gesture rate per minute of speech measure was preferred to rate per minute of interaction time, as the later does not control for amount of speech. It was also preferred to gesture rate per speaker turn as, although reasonable to use as a measure in the dyads, applying this measure to groups is problematic due to differences in turn length between different sized groups. As it was best to keep measures constant across experiments, gesture rate per speaker turn was also not the preferred choice here.

The additional measure chosen was gesture frequency as a function of number of stills. Here, gesture counts were divided by the number of stills to give the frequency as a function of number of stills. Again**,** this measure was calculated to give the frequency for every participant in a pair. This measure takes into account the amount of story content, in the form of number of stills. It also controls for gestures occurring outside of speech. It accounts then for gestures occurring on utterances that were not coded as speech in this data set, such as back channel responses, and for any gestures occurring in the absence of speech.

Whilst it was thought that the gesture rate by amount of speech measure would tap into gestures relating to the speech, that is those pertaining to the story content as well as those with a more interactive function (ie beats and socials), the gesture frequency measure should have captured gestures pertaining to contextual information (Cvpt, Ovpt, place holders and deictics). These measures could also be extended to the planned group experiment to afford consistency in measurement.

Statistical analysis was then applied to the data set. As with any data set, the gesture rate and speech duration distributions were checked in order to apply the correct type of statistics. Although the experiment was designed to reduce any variability, by encouraging all participants to gesture, it was predicted that the distribution of the data would deviate from a normal distribution. This was based on prior research that finds gestures to be variable across subjects. For this reason, particular attention was paid to the data distributions and to what type of analyses should follow, as this had consequences for the type of statistics that could be applied and conclusions that could be drawn.

Tests of normality checked whether the sample distribution was the same as that of a normally distributed population by comparing the sample distribution to a theoretical distribution. The Shapiro-Wilk test of normality was used since this is the recommended test for data sets of less than fifty.

Finding a mix of normal and non-normal distributions, it was decided best to conduct and report both parametric and non-parametric statistical tests. The reasons for this are as follows. Non-parametric tests are considered appropriate for non normal distributions since they make fewer assumptions about the distribution. By ranking the data and using the median rather than the mean, they are less sensitive to outliers. On the other hand, parametric tests assume a normal distribution and homogeneity of variance, and in using the mean as a measure of central tendency, are more susceptible to outliers. However, some would argue that parametric tests can withstand some assumptions being violated so long as the experimental design is well manipulated and there are equal data cells. Therefore, because of the robustness of parametric tests and moreover, that many of the distributions in the planned comparisons were in fact normal, both test types were utilised. All tests were reported and interpreted alongside distributions so that the relationship between the distribution and test results was evident.

The Wilcoxon's signed rank test was conducted on the gesture rate and average amounts of speech. This test works by taking the difference between the matched pairs in the data set and ranking these differences with sign (+ or -). The number of the non tied ranks (N) and sum of the +ively signed ranks (T+) is then determined. It is a powerful test, as it allows for both the direction and magnitude of the results to be considered (Siegel & Castellan, 1989). The exact test in the Wilcoxon signed rank test was used to give more accurate test results (Field, 2005). The equivalent parametric t-test was also carried out. By utilising both test types, issues surrounding the application of statistical analyses to gesture research, where non- normal distributions are most common, could be considered. This was considered important in light of the fact that much of the gesture research to date follows a descriptive type of analyses.

5.4.2 Alignment and emergence of conventions

Alignment theory (Pickering & Garrod, 2004) predicts gestures from the same communication should be produced at similar rates and amounts. To explore these issues of alignment, rates and amounts of the same gesture type were inter-correlated within the pairs. Correlations were on gesture rates per minute of speech and gesture frequency as a function of number of stills. The second measure was an additional one that checked for alignment once frequencies were corrected for number of stills. These measures were inter-correlated for every gesture type within the pairs so that p1 was correlated with p2 for all X12 pairs. It should be noted that for these purposes, the label p1 and p2 was arbitrary and so whether p1 or p2 was held constant was not an issue. In this case it was simply p1 for ease of computation since p1 was the 1^{st} column in the data sheet. For the gesture rate per minute of speech measure, gestures were also correlated between random pairs so that each p1 was correlated with a random other p2 from a different pair, and never with the participant they had interacted with, for all the pairs. Correlations between random pairs gave a baseline for alignment.

The strength of the relationship between gestures was measured using the non-parametric Spearman's and the parametric Pearson's correlations. Again due to a mix of normal and skewed data, the use of both was most appropriate.

Qualitative analyses were also carried out to explore alignment on gesture form. Here, instances of alignment on gestures referring to the same entity were identified and a descriptive analysis for each gesture given in the thesis. These examples make up a sample from the data set to illustrate alignment on gesture form.

5.5 Results

Gesture coding scheme

The coding scheme here differed in several ways form the scheme of the pilot. Firstly, it was more elaborate. Unlike in the pilot, where all content and all non-content gestures were classed together to make two categories, categories were broken down further. The gesture sub-categories observed in the pilot were coded for here. Content gestures were coded as iconic (broken down by perspective), metaphoric, deictic and place holding gestures. Non-content gestures were coded as beats and social gestures. As full description of iconic, Ph and deictic gestures are given in the pilot (see chapter 4). Here, examples from the dyad data set are given below;

Iconic gestures by viewpoint are illustrated in the following examples;

5.1 and [it goes into a piggy's mouth] … it's going into a piggy's mouth
Ovpt: finger points in circular trajectory to mark direction of turnip.

5.2 ahh, the piggy's [like oh yeah like gonna eat it kind of thing]
Cvpt: cupped left hand comes to mouth and then cupped right hand comes to mouth and both move back and forth to mouth as feeding the mouth. .

Metaphorical gestures occurred (as example 5.3 shows) but were infrequent;

5.3 … like [sort of like he's dreaming] or something
Metaphoric: flat hand moves up to the head and away from the head in a wave motion to indicate the mental state of dreaming.

A social gesture category emerged in this data set whereas this category was not identified in the group pilot. The following example describes how the gesture directly references the other person in the communication by being oriented towards them.

5.4 … that could be like [the smell of the flowers … I don't know where that fits in]
Social: hand held in cup shape with palm facing upwards and oriented towards the addressee.

Interestingly, in this data set, deictic gestures only ever referred to still content and never referred to the other person in the communication as they did in the groups of the pilot. This was most likely due to the reduced number of people communicating. The example given below shows how abstract deictic gestures can be as the point is simply made where the hand is resting;

5.5 Gnasher's in it this time …
Deictic: hand resting on leg points where it rests.

Deictic place holding gestures were also present and an example is given below;

5.6 … [in this book]
Place holder: *open hand with fingers in claw shape held horizontally in downwards position to rest on knee as though holding the place of the book.*

Beat gestures occurred and the following example illustrates a beat gestures by p2 when responding to p1;

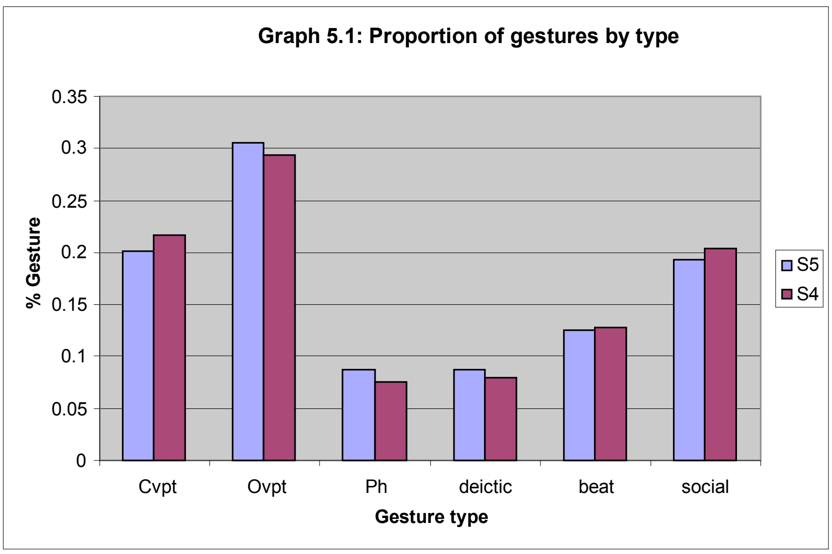
Example 5.7
p1: right at the start of mine [denis is sitting and he's tying his shoe laces …]
Cvpt: hands clasped as though holding shoe laces and move towards one another in tying shoe lace motion.

p2: [right]
Beat: hand in resting position makes simple back and forth movement.

Interaction time (duration)

The two conditions labelled S4 and S5 were described in reverse order so that they were comparable to the groups in terms of testing the experimental design. The average length of the interaction in the S5 (10 stills) condition (M= 13.29 mins, SD = 1.94) was comparable with that of the average length of the interaction in the S4 (8 stills) condition (M = 12.53 mins, SD = 1.43). The Shapiro-Wilks test of normality found the durations of the interaction to be normally distributed in S5 (D(12) = .922 mins, p>.05) and normally distributed in S4 (D(12) = .913 mins, p>.05). A t-test was carried out and found there to be no significant difference in the average length of the interaction between conditions S5 and S4 (t(11) = .364 mins, P >.05). Conditions S5/S4 could not then be said to influence interaction time.

As can be seen from graph 5.1, irrespective of condition, Ovpt gestures accounted for the highest proportion of gestures followed by Cvpt and then social gestures. As well as showing iconic gestures to account for different proportions depending on their perspective, the proportions highlight the importance of social gestures in the data set. Across both conditions, beat gestures accounted for the next highest proportion followed by comparable proportions of Ph and deictic gestures. Comparing within conditions, Ovpt, Ph and deictic proportions were marginally higher in S5 whereas Cvpt, beat and social proportions were marginally higher in S4. Proportionally then, there was little difference across conditions.



Amount of speech (as duration) correlated with amount of gesture (as frequency) for all gesture types.

Descriptive Statistics

The Shapiro-Wilks (W) test of normality results for gesture rate by total amount of speech are presented in table 5.1. As can be seen from the table, the data were mixed in terms of being normally and non-normally distributed. Ovpt gestures had normal distributions in both conditions. Cvpt and Ph gestures had a normal distribution in S4 but not in S5. Deictic, beat and social gestures had non-normal distributions in both conditions as did the speech distributions. These distributions determined whether parametric or non-parametric test were used and so a mix of both are reported in the next section. As speech had non-normal distributions in both conditions, only Spearman's correlations were reported.

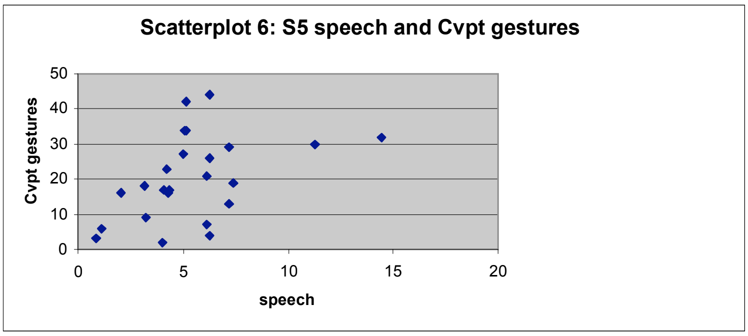
Table 5.1: Shapiro-Wilks (W) Tests of Normality

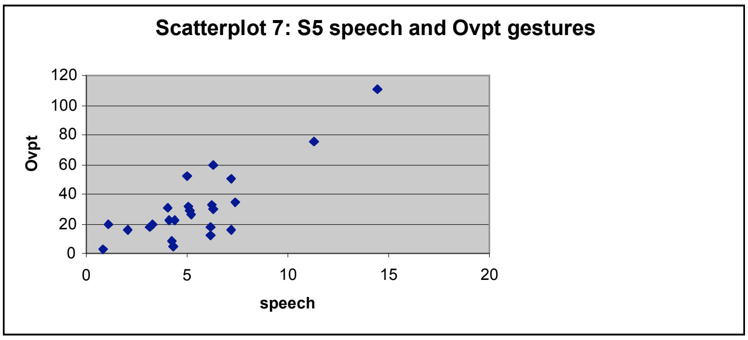
	df	W	Sig <.05
CvptS5	24	.889	.012
OvptS5	24	.965	.540
PhS5	24	.729	.000
deicticS5	24	.591	.000
beatS5	24	.826	.001
socialS5	24	.829	.001
speechS5	24	.591	.000
CvptS4	24	.956	.370
OvptS4	24	.951	.279
PhS4	24	.920	.058
deicticS4	24	.890	.013
beatS4	24	.884	.010
socialS4	24	.917	.050
speechS4	24	.897	.019

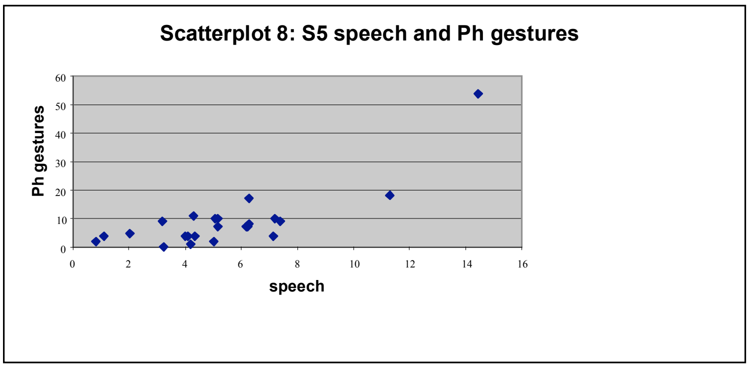
Red = sig different from normal distribution.

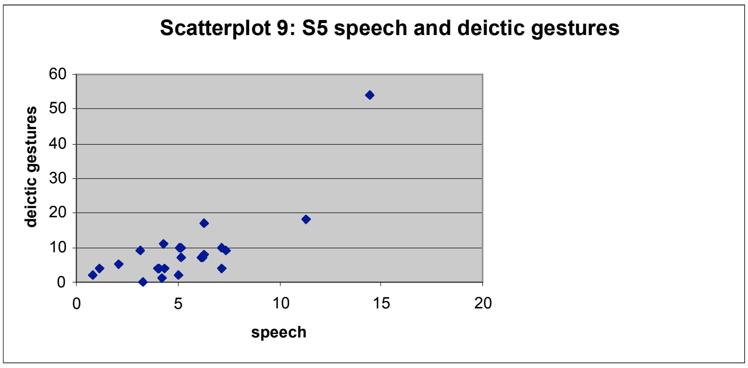
Correlations

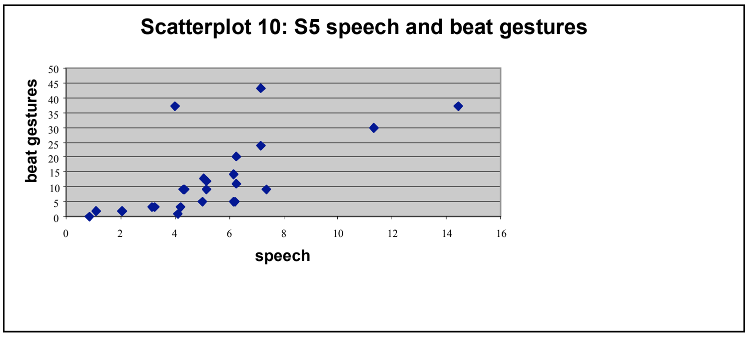
In S5, the Spearman's correlation co-efficient found a significant relationship between participant's amount of speech and amount of Cvpt gestures (r = .511, p=.01), Ovpt gestures (r = .610, p<.01), Ph gestures (r = .629, p<.01), deictic gestures (r = .629, p<.01), beat gestures (r = .740, p<.01) and social gestures (r = .515, p=.01). Scatterplots 6-11 show these relationships to be positive.

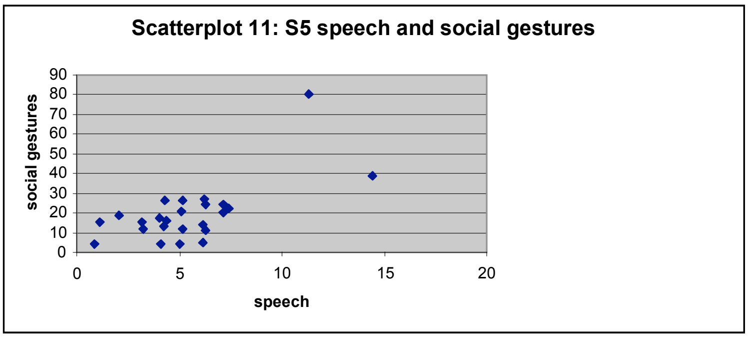




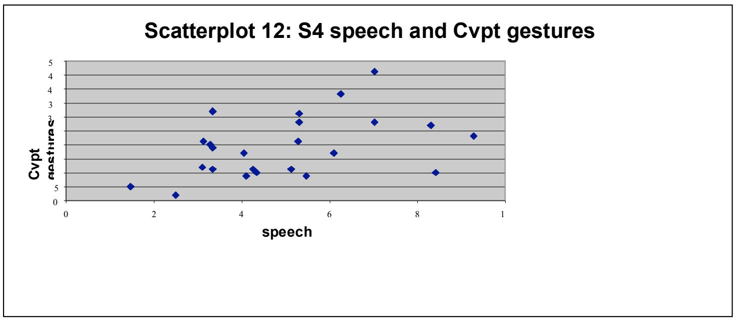


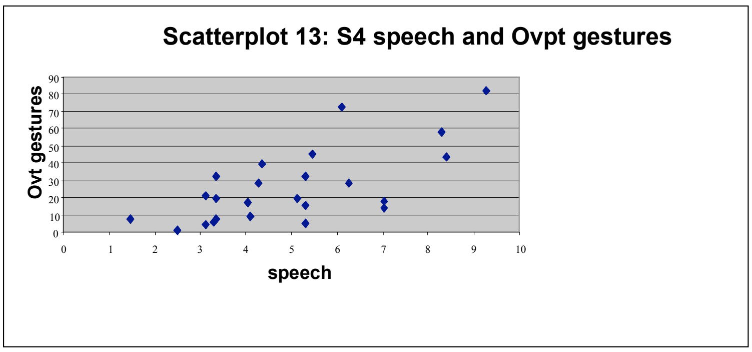


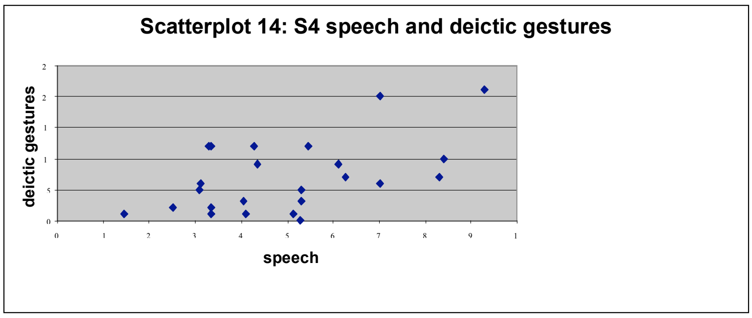


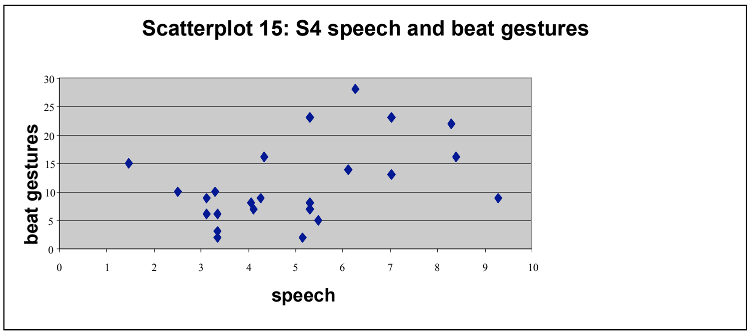


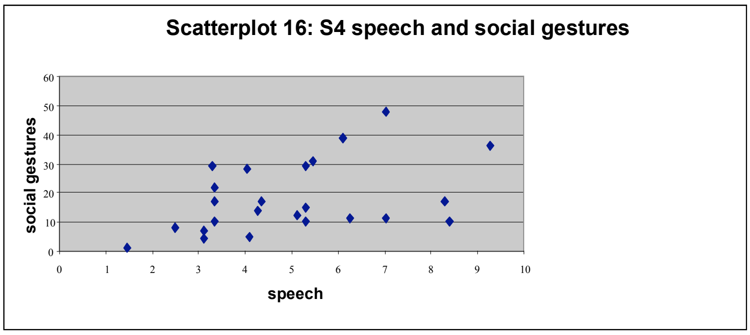
In S4, the Spearman's correlation co-efficient found a significant relationship between participant's amount of speech and amount of Cvpt gestures (r = .422, p<.05), Ovpt gestures (r = .640, p<.01), deictic gestures (r = .448, p<.05), beat gestures (r = .407, p<.05) and social gestures (r = .491, p<.05). Scatterplots 12 - 16 show the relationships to be positive.











All gesture types were therefore correlated with speech and in both conditions indicating speech should be taken into account in further analysis.

Average amount of speech

As the Shapiro-Wilks test of normality found average amounts of speech to be skewed in S5 (D(24) = .889 mins, p<.05) and normally distributed in S4 (D(24) = .956 mins, p>.05), both a Wilcoxon and t-test were carried out. The Wilcoxon test found no significant difference between conditions S5 (Mdn = 5.12 mins) and S4 (Mdn = 4.74 mins), T = 10, z = -.029, p>.05. Similarly, the t-test found no significant difference between the mean of S5 (M = 5.43 mins, SE = .600) and S4 (M = 4.97

mins, SE = .408), t(23) = .791, p>.05. There was then no difference in speech production across conditions S5 and S4.

Gesture rate per minute of speech

Descriptive statistics

The Shapiro-Wilk (W) test of normality results for gesture rate by total amount of speech are presented in table 5.2. As can be seen from table 5.2, the data were mixed in terms of being normally and non-normally distributed. Cvpt gestures had normal distributions in both S5 and S4 whereas Ovpt gestures had a non-normal distribution in S5 but a normal distribution in S4. Ph gestures had normal distributions in both S5 and S4, whereas deictic gestures had a normal distribution in S5 but not in S4. Beat gestures had non-normal distributions in both S5 and S4 whereas social gestures had a normal distribution in S4 but not in S5. As distributions determined whether parametric or non-parametric tests were used a mix of tests are reported in the next section.

Table 5.2: Shapiro-Wilks test of normality (gesture rate per minute of speech)

	df	W	Sig <.05
CvptS5	24	.971	.691
OvptS5	24	.860	.003
PhS5	24	.951	.278
DeicticS5	24	.951	.278
BeatS5	24	.753	.000
SocialS5	24	.836	.001
CvptS4	24	.947	.239
OvptS4	24	.957	.389
PhS4	24	.919	.055
DeicticS4	24	.916	.047
BeatS4	24	.756	.000
SocialS4	24	.921	.062

Red = sig different from normal distribution.

Comparative tests

The Wilcoxon test (see table 5.3) found a significant difference in gesture rate per minute of speech between S5 and S4 for beat gestures with beat S5 (Mdn = 1.76) and beat S4 (Mdn = 1.94), T = 8, z = -2.171, p = <.05, r = -0.44, this is a medium to large effect. There was no significant difference in gesture rate per minute of speech between S5 and S4 on Ovpt, deictic or social gestures.

Table 5.3: Wilcoxon test on gesture rate per minute of speech

Gesture Type	Mdn(S5/S4)	T	Z	p
Ovpt	5.47/4.61	12	.686	>.05
Deictic	1.31/1.03	11	.800	>.05
Beat	1.76/1.94	8	-2.171	<.05, (r = .44)
Social	3.7/3.08	10	.800	>.05

The t-test (see table 5.4) found no significant difference in gesture rate per minute of speech between S5 and S4 on any gesture type. Again, as there was no significant difference in gesture rate per minute of speech on any gesture type, conditions S5/S4 could not be said to influence gesture production.

Table 5.4: t-test on gesture rate per minute of speech

Gesture Type	Mean(S5/S4)	SE	T	p
Cvpt	4.13/3.98	.44/.43	.289	>.05
Ovpt	5.92/4.90	.708/.63	.954	>.05
Ph	1.60/1.36	.204/.22	1.11	>.05
Deictic	1.60/1.39	.204/.215	1.01	>.05
Social	4.08//3.66	.569/.462	.903	>.05

As more beat gestures were produced in S4, when extending the design to groups beat gestures should be interpreted with some caution. Since there was no significant difference in gesture rate per minute of speech on any other gesture type, conditions S5/S4 could not be said to influence their gesture production.

Gesture frequency as a function of number of stills

Descriptive statistics

The Shapiro-Wilk (W) test of normality results for gesture frequency as a function of number of stills are presented in table 5.5. As can be seen from table 5.5, the data were mixed in terms of being normally and non-normally distributed. Whereas Cvpt gestures had normal distributions in both S5 and S4, Ovpt gestures had non-normal distributions in S5 and S4. Ph and deictic gestures had non normal distributions in both S5 and S4. Beats and socials had normal distributions in S4 but not in S5. As distributions determined the type of comparative test used, a mix of tests are reported in the next section.

Table 5.5: Shapiro-Wilk test of normality (gesture rate per still number)

	Df	w	Sig<.05
CvptS5	24	.965	.540
OvptS5	24	.826	.001
PlaceHS5	24	.591	.000
DeicticS5	24	.591	.000
BeatS5	24	.829	.001
SocialS5	24	.729	.000
CvptS4	24	.951	.279
OvptS4	24	.884	.010
PlaceHS4	24	.890	.013
DeicticS4	24	.897	.019
BeatS4	24	.917	.050
SocialS4	24	.920	.058

Red = sig different from normal distribution.

Comparative tests

The Wilcoxon test (see table 5.6) found no significant difference in gesture frequency as a function of number of stills between S5 and S4 on Ovpt, Ph, deictic, beat or social gestures.

Table 5.6: Wilcoxon test on gesture rate per still

Gesture Type	Mdn(S5/S4)	T	Z	P
Ovpt	4.8/4.75	11	.071	>.05
Ph	1.4/1.37	12	.300	>.05
deictic	1.4/1.5	12	.472	>.05
beat	1.8/2.25	8	1.39	>.05
social	3.3/3.62	10	.700	>.05

Similarly, the t-test (see table 5.7) found no significant difference in gesture frequency as a function of number of stills between S5 and S4 on Cvpt, beat and social gestures.

Table 5.7: t-test on gesture rate per still

Gesture Type	Mean(S5/S4)	SE	T	P
Cvpt	4.07/4.77	.49/.557	-.1.115	>.05
beat	2.5/2.82	.512/.360	-.859	>.05
social	3.92/4.49	.630/.622	-.713	>.05

As there was no significant difference in gesture frequency a function of number of stills on any gesture type, conditions S5/S4 could not be said to influence gesture production.

Gesture rates inter-correlated within pairs by type (rate per minute)

Descriptive statistics

The Shapiro-Wilk (W) test of normality results for inter-correlated gesture rates by total amount of speech are presented in table 5.8. As can be seen from table 5.8, the data were mixed in terms of being normally and non-normally distributed. In S5, Cvpt, Ph and deictic gesture rates to be correlated had normal distributions whereas beat rates to be correlated were not normally distributed. Ovpt and social gesture rates to be correlated were a mix of normal and non-normal distributions. In S4, Cvpt, Ovpt, social, Ph and deictic gesture rates to be correlated had normal distributions whereas the beat rates to be correlated were a mix of normal and non-normal distributions. Distributions determined whether Pearson's parametric or Spearman's non-parametric tests were used and so, a mix of tests were carried out with significant findings reported in the next section.

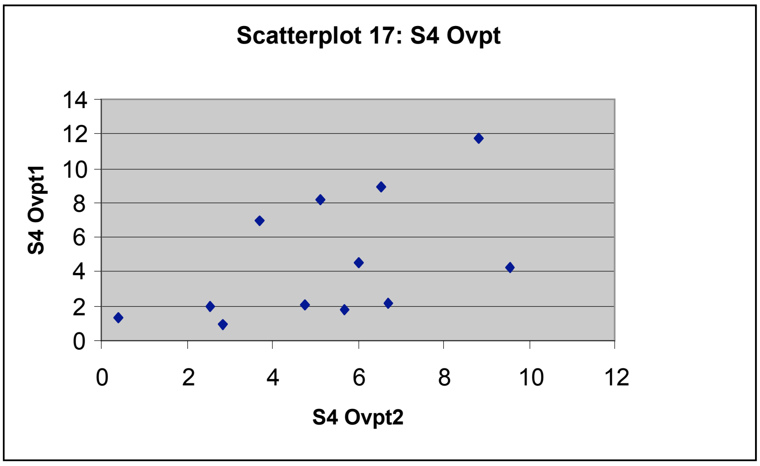
Table 5.8: Shapiro-Wilk test of normality for inter-correlated gesture rate per still number

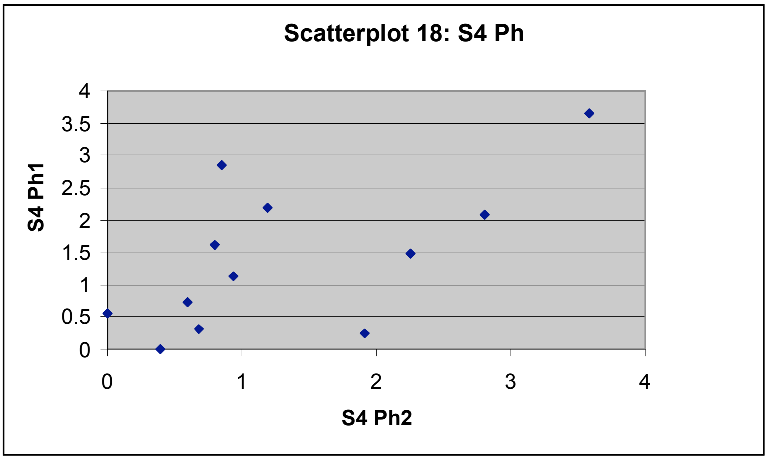
	df	w	Sig<.05
S5Cvpt1	12	.964	.839
S5Cvpt2	12	.974	.948
S5Ovpt1	12	.984	.995
S5Ovpt2	12	.819	.016
S5beat1	12	.715	.001
S5beat2	12	.850	.037
S5social1	12	.944	.547
S5social2	12	.744	.002
S5ph1	12	.947	.599
S5ph2	12	.948	.601
S5deictic1	12	.947	.599
S5deictic2	12	.948	.601
S4Cvpt1	12	.949	.630
S4Cvpt2	12	.915	.247
S4Ovpt1	12	.982	.990
S4Ovpt2	12	.870	.066
S4beat1	12	.713	.001
S4beat2	12	.984	.995
S4social1	12	.960	.782
S4social2	12	.900	.161
S4ph1	12	.901	.163
S4ph2	12	.945	.572
S4deictic1	12	.907	.195
S4deictic2	12	.933	.418

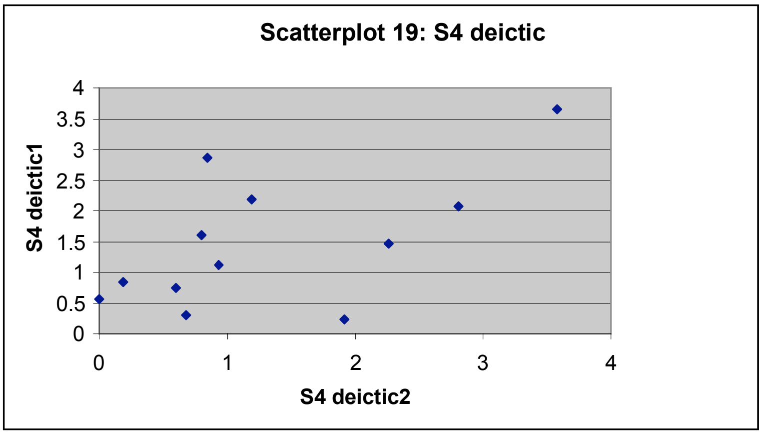
Correlations

In condition S5, there was no significant relationship between p1 and p2's gesture rates on any gesture type.

In condition S4, Pearson's correlation co-efficient found a significant relationship between p1 and p2's Ovpt rates ($r = .524$, $p<.05$), Ph rates ($r = .618$, $p<.05$) and deictic rates ($r = .596$, $p<.05$). Scatterplots 17 – 19 show the relationships to be positive.







As gesture rates per minute were inter-correlated between pairs on Ovpt, Ph and deictic gestures in S4, when one of the communicators in a pair produced these gestures the other was more likely to do so. Communicating pairs therefore aligned on the rate, or how fast, they produced these particular gestures. As no gesture types were correlated in S5, there was no evidence of alignment on gesture rates in this condition.

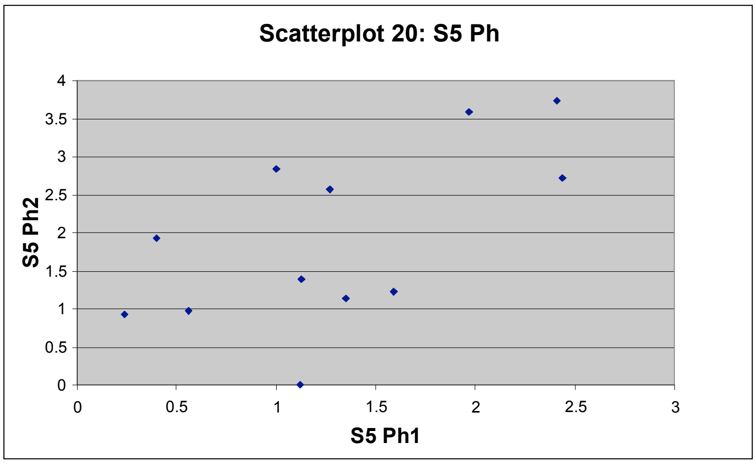
Gestures correlated between random pairs by type (rate per minute)

Descriptive statistics

As distributions were the same as for gesture rates inter-correlated within pairs by type (rate per minute) they were not reported again here. As in the inter-correlations, a mix of a mix of tests were carried out with significant findings reported in the next section.

Correlations

In condition S5, Pearson's correlation co-efficient found a significant positive relationship (see scatterplot 20) between p1 and p2's Ph rates ($r = .615$, $p<.05$) only.



In condition S4, there was no significant relationship between p1 and p2's gesture rates on any gesture type.

As Ph gesture rates per minute were correlated between random pairs in S5, when one of the communicators in a pair produced these gestures another in a random pair was more likely to do so. Rather than suggesting that non-communicating pairs aligned on the rate, or how fast, they produced these particular gestures, this provides a baseline measure of alignment. Such random alignment must be taken into account in interpreting any alignment between pairs. As no gesture types were correlated in S5, there was no evidence of random alignment on gesture rates in this condition.

Gestures inter-correlated within pairs by frequency as a function of number of stills

Descriptive statistics

The Shapiro-Wilk (W) test of normality results for inter-correlated gesture frequencies as a function of number of stills are presented in table 5.9. As can be seen from table 5.9, the data were mixed in terms of being normally and non-normally distributed.

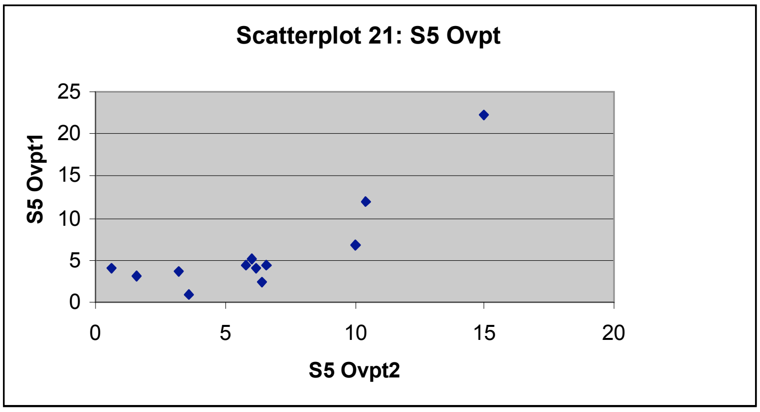
In S5, Cvpt gesture frequencies to be correlated had normal distributions whereas beat frequencies to be correlated were not normally distributed. Ovpt, social, Ph and deictic gesture frequencies to be correlated were a mix of normal and non-normal distributions. In S4, Cvpt, beat, Ph and deictic gesture frequencies to be correlated had normal distributions whereas Ovpt frequencies to be correlated were not normal and social gesture frequencies a mix of distributions. Distributions determined whether Pearson's parametric or Spearman's non-parametric tests were used and so, a mix of tests were carried out with significant findings reported in the next section.

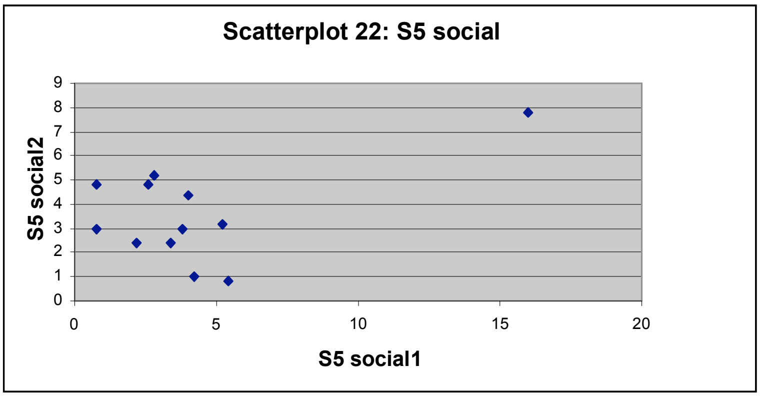
Table 5.9: Shapiro-Wilk test of normality for inter-correlated frequency as a function of number of stills

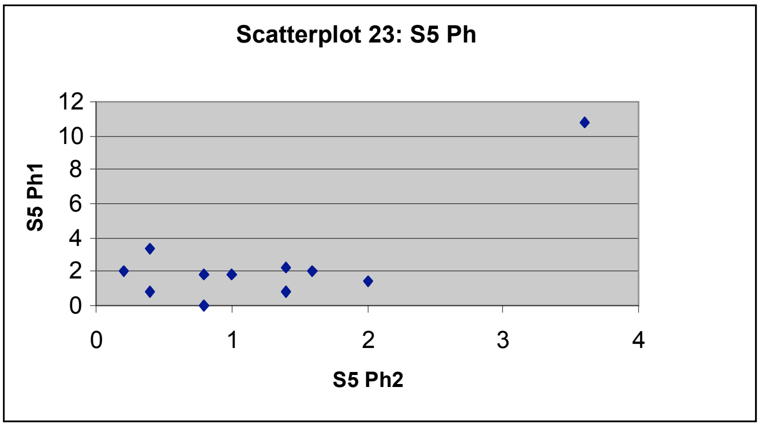
	df	w	Sig<.05
S5Cvpt1	12	.963	.822
S5Cvpt2	12	.917	.264
S5Ovpt1	12	.940	.504
S5Ovpt2	12	.698	.001
S5beat1	12	.885	.101
S5beat2	12	.777	.005
S5social1	12	.686	.001
S5social2	12	.946	.580
S5ph1	12	.867	.060
S5ph2	12	.621	.000
S5deictic1	12	.867	.060
S5deictic2	12	.621	.000
S4Cvpt1	12	.951	.657
S4Cvpt2	12	.895	.136
S4Ovpt1	12	.839	.027
S4Ovpt2	12	.857	.045
S4beat1	12	.940	.502
S4beat2	12	.869	.063
S4social1	12	.840	.027
S4social2	12	.963	.821
S4ph1	12	.877	.081
S4ph2	12	.917	.261
S4deictic1	12	.878	.082
S4deictic2	12	.919	.279

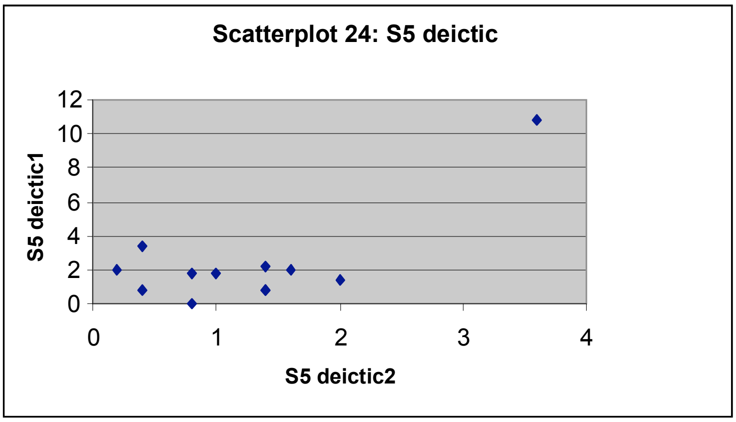
Correlations

In condition S5, Pearson's correlation co-efficient found a significant relationship between the frequency of p1 and p2's Ovpt ($r = .842$, $p = 0$), social ($r = .503$, $p<.05$), Ph ($r = .731$, $p <.01$) and deictic ($r = .731$, $p<.01$) gestures. Spearman's correlation co-efficient found a significant relationship between the frequency of p1 and p2's Ovpt rates ($r = .695$, $p<.01$). Scatterplots 21 – 24 show the relationships to be positive.

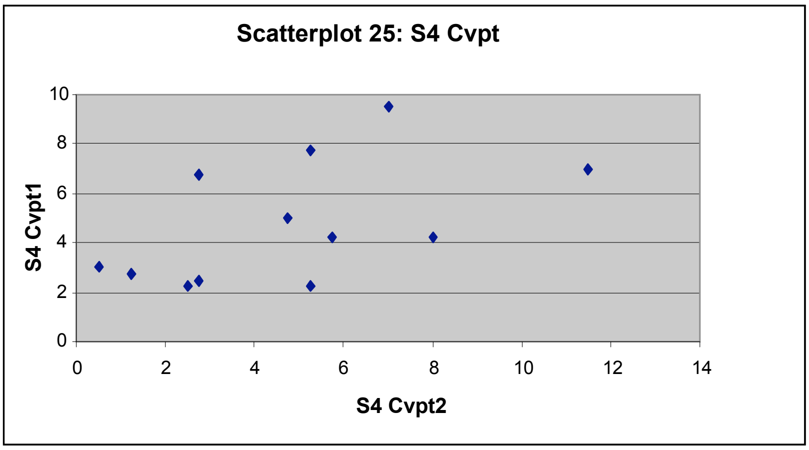

Scatterplot 21: S5 Ovpt
S5 Ovpt1
25
20
15
10
5
0
0
5
10
15
20
S5 Ovpt2

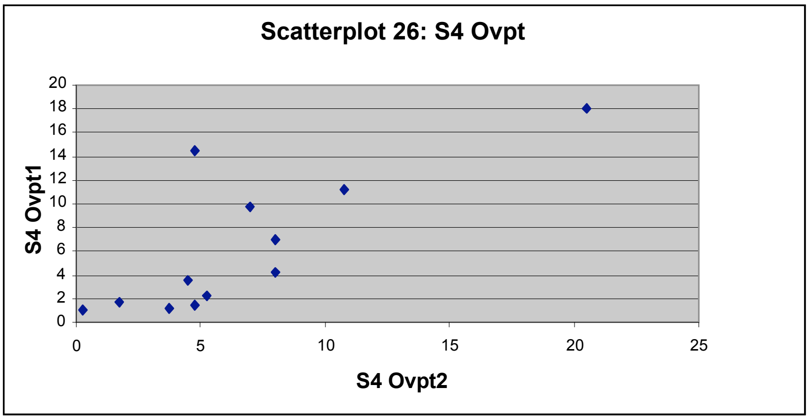

Scatterplot 22: S5 social
S5 social2
9
8
7
6
5
4
3
2
1
0
0
5
10
15
20
S5 social1


Scatterplot 23: S5 Ph
S5 Ph1
12
10
8
6
4
2
0
0
1
2
3
4
S5 Ph2


Scatterplot 24: S5 deictic
S5 deictic1
12
10
8
6
4
2
0
0
1
2
3
4
S5 deictic2

In condition S4, Pearson's correlation co-efficient found a significant relationship between p1 and p2's Cvpt rates ($r = .533$, $p<.05$). Spearman's correlation co-efficient found a significant relationship between p1 and p2's Ovpt rates ($r = .789$, $p<.01$).





As gesture frequencies were correlated on Ovpt, social, Ph and deictic gestures in S5, and on Cvpt and Ovpt gestures in S4, communicating pairs aligned on the production of these gestures.

Descriptive analyses of Alignment.

Alignment on gesture form also occurred in instances where there was reference to the same entity. This indicates that when one of the communicators produced a particular gesture form, the other was more likely to do so. As in the pilot, reference to the same entity could occur when a group member referred to what another member had said or, as some of the entities were the same across stills, when a group member referred to an entity in their own still that another had mentioned from their own stills. In example 5.8, p1 is checking she understands who p2 refers to by describing what the character wears and they both align on the gesture describing the character's top;

Example 5.8
p1: Denis the menace with the stripy shirt?
Cvpt: hand points towards torso and makes back and forth motion to indicate stripes.

p2: yeah, yeah …
Cvpt: hand points towards torso and makes back and forth motion to indicate stripes.

In example 5.9, p1 gestures that denis is getting knocked up into the air and when p2 clarifies that this is what is happening, she aligns on the throwing up in air gesture;

Example 5.9
p1: [knocked up in the air] …
Ovpt: finger point moves in upwards direction to show the trajectory of denis.

p2: … of denis getting flung in the air …
Ovpt: finger point moves in upwards direction to show the trajectory of denis.

5.6 Discussion

The coding scheme consisted of iconic gestures, broken down by Cvpt and Ovpt, metaphoric, Ph, deictic, beat and social gestures. Of these gestures, Cvpt, Ovpt, metaphoric, Ph and deictic gestures were content gestures whereas beat and social gestures were non-content gestures. Although gestures were defined in terms of a referent, and so the kind of information they conveyed (content or non-content), the way in which gestures conveyed this information (imagistic or not/abstract or concrete) was also considered. Defining gesture signs by the way they relate to objects alongside the way in which they convey information fits with the coding scheme of McNeill (1992). However, it differs in one important aspect from McNeill (1992) in that speech is not used to define the gesture in any way. By focusing on the referent of the gesture, coding here more closely followed the methodologies of Beattie and Shovelton (2002) and Bavelas et al (1992; 1995; 2007). Linking gestures to referents in this way treats the gesture as an independent sign (the reasons for doing this are outlined in chapter 1).

As the coding scheme emerged from the data set, it shows the existence and therefore importance of a social gesture category. A reason for running pairs in this experiment was to ensure the coding scheme was an inclusive one. As an additional category emerged in the dyad coding scheme, the coding scheme form the pilot would have been under inclusive for interactive communication. If the pilot scheme were applied to a dialogue situation, social gesture would have been overlooked.

To measure the level of interaction across the two conditions, the interaction time alongside the production of speech and gesture was compared. Since there was no significant difference in length of interaction, having an additional two stills (S5) could not be said to influence the level of interaction. If varying the amount of information influenced the interaction, interaction times would be expected to differ across conditions, with a longer time suggesting a more collaborative communication (Fay, 2000; Bavelas et al, 1992; 1995; 2007).

Gesture proportions showed the prevalence of social gestures, as social gestures accounted for a high proportion of overall gestures and in both conditions. This again highlights the importance of the social gesture category. That Ovpt gestures accounted for the highest proportion, and a higher proportion than Cvpt gestures which followed, is interesting as it suggests Ovpt occurred more often relative to other gesture types. In line with gesture findings is that beat gestures followed by Ph and deictic gestures accounted for the lowest proportions as this indicates these gestures were produced much less often relative other gesture types. (Bavelas, 1992; 1995; 2007). Importantly, as it again suggests varying the amount of information did not influence the interaction, there was no difference in proportions across conditions.

As predicted, speech was positively correlated with all of the six gesture types (Cvpt, Ovpt, Ph, deictic, beat and social) included in the analysis (note that the metaphor category was not included) indicating gesture, like speech, can be taken as a measure of the communicative style. In the same way Fay (2000) used the speech sign as a measure of the communicative style in experimental groups, and whether communication was in monologue or dialogue, gesture production here can be taken to measure the level of communication. That beat and social gestures were correlated with speech is especially interesting as this highlights the importance of treating non-content gestures as communicating signs even though they do not communicate information about the content. As in Bavelas et al (1995), the qualitative examples given in the results section show social gestures directly referenced addressees, and in doing so performed various functions to do with providing feedback about understanding, engagement and so on. Non-content gestures therefore communicate important information about the state of the on-going discourse (Bavelas et al, 1992; 1995).

Another reason the correlation of non-content gestures with speech is of interest is that research shows non-content gestures, in performing an interactive function, have a more distant relationship with speech than content gestures (Bavelas, 1992). From the correlations here, there is no evidence of a weaker relationship and so, more distant relations, between non-content gestures and speech.

Whilst finding that speech and gesture were related allows gesture to be treated as a measure of the communication style because they are related, the speech must be taken into account in further analyses. To do this, gesture rates were calculated per minute of speech. As it was prudent to take the amount of information into account, the gesture frequency as a function of number of stills was also calculated. With the exception of beat gestures, no significant differences were found for any of the gesture types on either measure. As more beat gestures were produced in S4 than in S5 on the rate per minute of speech, any difference for beat gestures on this particular measure between the group conditions should be interpreted with caution. It is however reasonable to conclude that for all other gestures, varying the amount of information from 10 (S5) to 8 (S4) stills did not influence gesture production. In the main then, when there is an optimum amount of information, having two stills more or less doesn't influence gesture production and thus the level of interaction. These findings suggest the experimental design could be extended to a group situation without the amount of information having to be treated as a factor except for on beat gestures.

As this was an exploratory experiment to get a base line coding scheme and check the experimental design for any effect of different amounts of information on gesture communication, no predictions were made. However, different rates of production for the two conditions would have indicated a difference in the level of interaction. Similarly, different amounts of speech across the conditions would indicate a difference in the level of communication. Since average amounts of speech were calculated for the gesture rates per minute of speech measure, these were also reported. As average amounts of speech were comparable across conditions, they were in line with gesture rates and frequencies in suggesting the interactions were comparable.

The same gesture types were correlated within participating pairs in order to explore alignment on both the rate and amount of gesture production. Correlations differed by condition (S5/S4) and depending on the measure used. On the gesture frequency (as a function of number of stills) measure, Ovpt, social, Ph and deictic gestures were correlated in S5 whereas only Cvpt and Ovpt gestures were correlated in S4. Interestingly then, Ovpt gestures were correlated in both conditions. Ovpt gestures were also correlated in condition S4, alongside Ph and deictic gestures, on the rate per minute measure. Ovpt gestures were therefore highly correlated and more so than any other gesture type. As predicted then, gesture rates within pairs were highly correlated but they were more so on the gesture frequency measure and in S4. This shows that even when frequencies were corrected for amount of information gesture amounts were still correlated.

However, as a baseline measure of alignment is necessary, this was calculated for the rate per minute measure (since this was the main measure with the frequency measure being a secondary check corrected for still number). To do this, gesture rates per minute were correlated between random pairs who had not interacted together. The only significant positive relationship found between random pairs was for Ph rates in condition S5. Had there been a significant correlation within pairs on this gesture type, this would need to have been interpreted with caution. As is, correlations of

Ovpt and deictic gestures on the rate per minute measure were above a baseline level of alignment.

That no relationship was found for beat gestures across any of the conditions is interesting considering these gestures are observed to be ‘interactive’ and to ‘seek responses’ from interlocutors (Bavelas et al, 1992; 1995). Alignment on these particular gestures might therefore be expected. However, since Bavelas et al’s (1992; 1995) category of interactive gestures consisted of social as well as beat gestures, and that gestures serve several functions within these categories some of which provide feedback at the outset rather than seeking it, a possible explanation for the lack of correlation on beat gestures is that a response is often not required. That Cvpt and Ovpt gestures often aligned at points in the communication where information needed to be clarified supports this as these gestures did require a response in terms of the information needing to be relayed.

The observation by Kendon (1970; 1990) that back channel types of gestures, such as beats, are often arranged in the listener’s (one’s own) time frame rather than the gesturer’s (another’s) time frame is also important here. Kendon (1970; 1990) suggests back channel type gestures are not aligned like others because they occur outside of a floor hold. Alongside them providing clarification at the outset, this can explain why they do not receive a response and align.

Another explanation for the lack of correlation on beat gestures is that quantitative analysis was not sensitive enough to tap into alignment on these gestures since they are simpler in form than iconic gestures, or even social gestures. Since beat gestures rely heavily on temporal aspects (McNeill, 1992), the best kind of analysis for beats would be a precise temporal analysis. Given the research questions and extent of the coding carried out to answer them, a temporal analysis was not feasible here but this is something that should be considered for further research. Analysing beat gestures is tricky and this is notable by the lack of research on them. Even when considered alongside social gestures, under the umbrella of interactive gestures (Bavelas et al, 1992; 1995; 2007), descriptions are mostly of social gestures. Again, this is most likely due to qualitative analysis not being appropriate due to the simplicity of the beat form.

A reason why Ovpt gestures would align more than Cvpt gestures is to do with their frame of reference. Whereas Cvpt gestures are egocentric to the gesturer, Ovpt gestures are object centred (see chapter 1). To expand on this, Cvpt gestures are egocentric in that they take the perspective of a character the gesturer has experience of but the addressee may or may not have experience of. Even if the addressee has experience of the character, they may not have experience of the character in the particular domain being expressed, whereas the gesturer does. They are egocentric because Cvpt gestures use specific parts of the gesturer's body whilst gesturing. This also makes them more complex in form, and so more difficult to align on, than Ovpt gestures.

Alignment on gesture form was also evident and examples of where this occurred were given in the results section. Although the examples given are a sample, and do not represent all instances of alignment in the data set, they were chosen as good examples and as such illustrate the kind of alignment that occurred.

The findings of alignment on gesture rate, amount and form provide support for the alignment model as they suggest that when one member of a pair produced a particular gesture the other was also more likely to, and moreover at the same rate and in the same form. To explore the issue of alignment further, an overseer experiment was designed to empirically test for alignment on gesture form. This is reported in the following chapter.

Chapter 6: Overseer Experiment

6.1 Introduction

The dyad experiment (chapter 5) explored alignment on the rate and amount of gesture by correlating rates and frequencies of the same gesture type (Cvpt, Ovpt, Ph, deictic, beats and social). As iconic Cvpt/Ovpt gestures are substantial in their form, alignment can also be tested on this basis. The overseer experiment reported here tests whether iconic Cvpt/Ovpt gestures from the same pairs were more similar in form than those from different pairs. By considering the gesture form the overseer investigates alignment more fully. It also follows up the finding that Ovpt gesture rates and amounts were more strongly correlated than Cvpt gesture rates and amounts in the dyads.

Iconic gestures, from same and different pairs, referring to the same entity and with the same viewpoint (eg character or observer viewpoint), were viewed by naive overseers who had not taken part in the original experiment. By choosing an odd man out, overseers decided which two gestures were more similar.

To ensure choices were based on gesture form, overseers watched gestures with the speech signal off and made judgement based on size, movement (eg. trajectory and speed), complexity and hand shape.

It was predicted that if, as the alignment model (Pickering & Garrod, 2004) holds, gesture signs are modified through use so that they evolve to become more similar in form, then gestures from the same pairs would be considered more similar in form than those from different pairs. Based on finding Cvpt and Ovpt gestures, but especially Ovpt gestures, to be correlated, and therefore aligned, on both their rate and amount in the original participating pairs (chapter 5), it was further predicted that Ovpt gestures would be more similar in form than Cvpt gestures. Overseers would then find Ovpt gestures from the same pairs to be more similar than Cvpt gestures from the same pairs.

6.2 Methods

6.2.1 Subjects

Nineteen participants signed up on an advertisement sheet within the department of psychology at the University of Glasgow to take part in the experiment. In total, 9 females and 10 males took part.

6.2.2 Materials

Materials for the experiment were generated from the dyad data set. Iconic gestures had been coded in the dyad data set and given a descriptive tag to define viewpoint (Cvpt/Ovpt). They were also given tags to identify the referent and still number the referent was in. This meant iconic gestures with the same referent could be identified. This was done using the query language in Nite (see chapter 4), by declaring and searching for the still number/referent tags (Carletta, Evert, Heid and Kilgour, 2005). As queries for the still number/referent tags were highly specific, running them involved toggling between a word document, which was generated in the dyad experiment, and the Nite interface. The word document consisted of still number/referent tags for every person in a pair. These tags were copied and pasted into the queries so that the exact still number/referent held in the word document was defined and searched for.

Running the queries identified 74 pairs of iconic gestures that could be matched for same reference with another pair. The 74 pairs of gestures were made up of 40 Cvpt and 34 Ovpt gestures. This gave a total of 148 gesture trials consisting of 80 Cvpt and 68 Ovpt gestures.

Video footage was imported into Adobe Premiere where clips of the relevant gestures were edited by their start and end times so that only the gesture of interest was captured, and each one in a new movie file. Start times were taken from the beginning of the gesture stroke to account for some gestures having a preparation phase and some not having this phase. In order to view only the gesturing participant, and not the other in the communication since this would have given away partners, the clips were masked and edited. To avoid any cue from seating position, the gesturer was always placed in the upper portion of the screen. This involved flipping clips when necessary.

As the experiment was run in Matlab, the edited video clips were organised within the programme. Gesture clips from the same pairs were arranged by participant 1 (p1) and participant 2 (p2). These pairs of clips were then grouped with two gesture clips (p1 and p2) of a different pair that made the same reference. With the gesture clips grouped into quadruplets, Matlab was programmed to present 2 clips from within the same pair and one clip from the other pair in the quad. This meant 3 gesture clips were presented for each of the 148 gesture trials. With a total of 37 quadruplet sets of clips, the quadruplet presentation was randomised across subjects. Presentation was also semi-randomised within every quadruplet set of clips. This randomised presentation ensured those from the same pairs never occurred together more often than with a different pair.

Example of the types of Ovpt and Cvpt gestures used in the sets of quadruplet video clips are given in chapter 5.

6.2.3 Procedure

Participants were seated in front of a computer screen when being given the experimental instructions. Participants were informed that they would see 148 trials of X3 video clips showing different people gesturing and that they were to choose, as the odd one out, the least similar gesture out of the three presented. It was explained that they could view the clips as many times as they wished. All clips could be played as often as was necessary but were always re-played as a set of three rather than as individual clips. This was to eliminate exposure effects. Participants were asked to focus on particular aspects of the gesture form when choosing the odd gesture out and these aspects were placed in order of importance. First, they were to consider; the size, movement (such as the trajectory and speed) and complexity of the gesture. These three aspects of the gesture form were considered most important and equally so. After these, they were to consider hand shape. They were asked not to distinguish on the basis of handedness ie whether the gesture was with the left or right hand.

It was explained that for each trial the computer screen would ask them if they wished to play the videos, at which point they should press a button to indicate ‘yes’. The computer screen would then ask if they wanted to play the clips again or make a decision as to what was the odd one out. They were instructed to follow the instructions on the screen and press ‘R’ to re-play the clips or make a choice by pressing keys numbered 1, 2 or 3 as these related to clip presentation order. Once they had made a decision, by pressing the relevant key on the keyboard, they were informed that the computer would ask them to rate how confident they felt about their decision. Ratings were on a scale of 1-7 and again by using the numbered keys on the keyboard. This gave ratings on a 7 point licker scale. Once the instructions were clear, the experiment began.

6.3 Analyses

The distribution of the overall % correct was checked for normality. A one sampled t-test was then carried out to check whether the proportion of correct responses was significantly above the level of chance. The one sampled t test was chosen as it is recommended over the z-test when $n < 30$.

The distribution of the % correct for Cvpt and Ovpt gestures was checked for normality. A two-sampled related t-test checked whether the mean % correct for Cvpt gestures differed significantly from the mean % correct for Ovpt gestures.

The ratings were not analysed as they had a ceiling effect with all gesture choices from all participants scoring 6 or 7 on the licker scale.

6.4 Results

% correct responses

The Shapiro-Wilks test of normality found the overall % correct data to be normally distributed ($D(19) = .967, p>.05$).

The number of correct responses was greater than the chance score of 33.3% (or 49) on a one sample t-test ($M = 37\%$; $SE = .72$; $t(18) = 51.89, p< .01, r = .99$) and this was a large effect.

% correct responses by perspective

The Shapiro-Wilks test of normality found the % correct by perspective to be normally distributed for both Ovpt ($D(19) = .977, p>.05$) and Cvpt gestures ($D(19) = .941, p>.05$).

The proportion of correct responses was higher for the Ovpt perspective ($M = 41\%$, $SE = 1.299$) than for the Cvpt perspective ($M = 34\%$, $SE = 1.220$). A t-test found this difference to be significant ($t(18) = 3.4, p<.05, r = .62$) and this was a large effect.

6.5 Discussion

Naive participants overseeing gestures more often chose a gesture from a different pair as the odd one out than a gesture from the same pair and they did so above the level of chance. This indicates gestures from the same pairs, or the same communication, were more aligned on form than those form different pairs, or a different communication. This follows up and supports findings from the dyad experiment (chapter 5), where gestures of the same type were correlated on gesture rates and amounts. In line with predictions, this effect was strongest for Ovpt gestures where the odd gesture out (the one from a different pair) was more often chosen correctly. This again supports findings from the dyad experiment where Ovpt gestures were highly correlated by both rate and amount and more so than Cvpt gestures or indeed any other gesture type. Differences in alignment depending on perspective can be explained in terms of frames of reference for speech (see chapters 1 and 5). Cvpt gestures have an egocentric relative frame of reference, which should be difficult to align on, whereas Ovpt gestures have an object centred intrinsic frame of reference, which should be easier to align on.

The overseer shows that there may be several gesture forms for any particular gesture but that communicating pairs were more likely to use the same form. Of course, some gestures will be more restricted in form than others but gestures did substantially vary in form as can be seen from the examples in the results section.

Finding gestures to be more similar in form when from the same 'interactive' communication fits with findings reported throughout this thesis, where signs are modified depending on the needs of the communication. Such evolution of sign also provides support for Pickering & Garrod's (2004) alignment model where it is claimed signs evolve, to become more similar, through interactive communication. Alignment theory also holds that the local convergence observed here is the first step in conventionalisation. The conventionalisation of gesture has not been addressed here but the finding of convergence, alongside other findings of conventionalisation in gesture, suggests gestures can become conventionalised and that they do so through use.

To investigate the conventionalisation process of gesture signs, further research should look at the development of gesture signs over the course of communications in different communities as Garrod & Doherty (1994) did in speech. Further analyses could also involve a qualitative analysis of specific features (such as size) to compare across the gestures rated as more or less similar.

In light of findings of alignment on social gestures (Kendon, 1995; Kendon, 2004; Muller & Posner, 2004) it would also be interesting to test for alignment on the gesture form of these gestures. As social gestures are like iconic gestures, in that they are imagistic, an overseer could also work for these gestures. However, this would be a more difficult task for overseers because social gestures are simpler in form and overseers in this study reported finding the task difficult with more complex iconic gestures. Although participants reported finding the task difficult and this was evident from scoring on the licker scale, they rated gestures successfully. This was remarkable considering gestures were taken out of the wider communicative context and coded with the speech signal off. In addition, overseers received no formal coding training other than the instructions given by the experimenter for distinguishing gestures. Like Beattie & Shovelton's (2002) overseer experiment then, where gestures were also presented in the absence of speech, this highlights the communicative strength of these particular gestures.

The following chapter moves on to the group experiment.

Chapter 7: Group Experiment

7.1 Introduction

The group design, based on the pilot study and tested in the dyad experiment, was with groups of 5 (G5) and groups of 8 (G8). G5 and G8 members were given 10 and 8 stills respectively from one of two beano cartoons (flower/book). This meant in G5, every group member received two stills, whereas in G8, every group member received only one still each. The task was for group members to memorise their own stills and discuss them with other group members with the overall aim of constructing a story from the information they shared. The same five group members took part in both group size conditions to give a within design. Three group members were always either added or subtracted to or from the group for the second group size condition. Group size, alongside still number and comic type (flower/book), were all counterbalanced across 24 groups.

Predictions for the group experiment were based on prior research findings, which are fitting with dialogue models of communication and in particular the alignment model (Pickering & Garrod, 2004). They were further refined by findings in the pilot and dyad experiment in this thesis. For both G5 and G8, it was predicted that speech would be positively correlated with all of the gesture types coded. If this was the case, gesture production (like speech) could be taken as a measure of the communicative style, be it dialogue or monologue. It was further predicted that, G5 members would be more collaborative than G8 members, and this would be reflected in gesture rates and distributions. Whereas the communicative style in G5 was expected to be like dialogue, communication in G8 was expected to be like serial monologue. This led to the prediction that beat and social gestures, involved in providing feedback about the communication, alongside Ph and deictic gestures, which rely on visibility and can serve as reduced expressions, would be produced more often in G5. It was also tentatively predicted that iconic (Cvpt/Ovpt) gestures would be produced more often in G8. Such findings would fit with dialogue models of communication but in particular with the alignment model (Pickering & Garrod, 2004) which predicts dialogue-like and monologue-like styles of communication within groups.

Gestural alignment on the rate and amount of gesture was measured by correlating the same gesture types. On the basis of the alignment model, it was predicted that gesture rates within group members would be more highly correlated in G5 than in G8. On the basis of findings from the dyad (see chapter 5) and overseer experiment (see chapter 6) and the egocentricity of Cvpt gestures (see chapter 1), it was further predicted that Ovpt gestures would be more aligned than Cvpt gestures. Predictions were the same for alignment on gesture form, where gestures referred to the same entity, and these were identified for qualitative analysis.

Although the design was similar for the pairs and groups, it was not feasible to make direct comparisons since these were two different experiments and would be treated as such. The design and analyses were however kept constant across the two experiments for ease of interpretation.

7.2 Methods

7.2.1 Subjects

Ninety-six participants signed up on the experimenter's advertisement sheet within the department of psychology at the University of Glasgow to take part in this experiment. Participants were organised to make 24 groups; 12 groups of 5 and 12 groups of 8 so that 5 of the same group members participated in the two different sized groups. This involved adding or subtracting three participants. All groups were of mixed gender and consisted of 65 females and 31 males in total. Participants were paid in course credits or cash depending on their preference.

7.2.2 Materials

Twenty cartoon stills from two different Beano comic strips were used. Ten came from a story featuring events based around the discovery of a smelly flower (from now on called 'flower') and 10 from a story featuring Denis dreaming of a journey through the beano book with events based around the relevant characters (from now on called 'book'). Stills were chosen on the basis of their content. They were chosen on the basis of how well a story could be constructed from them and if it were likely that reference to the content in them would elicit gestures (eg stills depicting movement). They were not then necessarily in chronological order. Indeed, stills in sequence could be missing. Stills were blown up and verbal expressions blanked out but for onomatopoeic sounds that exaggerated events.

7.2.3 Procedure

Participants were seated in a circle in the experimental room in one of two group size conditions (G5 or G8). A camera fitted with fish eye lens was mounted to capture the bird's eye view perspective of all group members. The camera was linked to a television so the experimenter could check participant's seating and ensure participants, and their gesture space, were captured. Again the camera was set to capture the interaction at 25.00 frames per second (fps) with a frame size of 320 by 240 at 48, 000 Hz and 32 bit float in stereo.

Once seated, the experimenter handed out stills from either one of the comic strips (flower/book). When in G5, participants were given two stills each (collectively making for 10 stills) and when in G8, they were given one still each (collectively making for 8 stills). Participants were asked to place the stills face down on their knees whilst awaiting instruction. Every person in the group had different stills which were in no particular order (ie still order did not equate with seating or way in which they were handed out) and participants were informed of this.

Participants were informed that the task was to use the information they had between them to discus and order the stills into a cohesive story. To do this, participants were to memorise the information in their own stills. They were given sufficient time to do this. The comic stills were tangible, as participants had access to them, however it was requested that stills be kept under the participant's own chairs throughout the discussion with any reference to them being kept brief. This ensured participants saw only their own stills and hands were unoccupied whilst talking. Participants were told information could be missing so they should think up and use possible events in-between the stills in order to link up the story. Once all group members had memorised the detail of their stills, the group discussion began. The first interactive task was complete once group members felt they had found the best still order for a cohesive story.

Five participants from the first group discussion took part in a second group discussion where the size of the group was altered by adding or subtracting three members and the procedure repeated. In the second group discussion, group members received a respective number of stills from the other comic strip. To keep the still reference constant, the same two stills were always added or subtracted. In this way, group size and cartoon type were counterbalanced across groups. This gave a mixed design, with five group members being fully within across the condition of group size and the additional three group members being between the group size condition.

7.3 Coding

7.3.1 Gesture coding

Video footage was captured in the .mov file format using Pixela image maker. Gestures were then coded using the baseline coding scheme generated in the dyad experiment. This scheme consisted of iconic, metaphoric, deictic, place holding, beat and social gestures. In a second level of coding, the perspective of iconic gestures was coded by adding character (Cvpt) or observer (Ovpt) tags. At a third level of coding, iconic gesture events were given a descriptive tag, of what the gesture referred to and still number the referent was in, so as to link to referents to stills (eg S8 ‘denis head stuck in ground’). Gesture coding was always with the speech signal on (see chapter 4 for a discussion as to why) and was for all members (5 or 8) within a group who were distinguished by the labels p1-p8. Group member labels were matched across group conditions so that p1 in G5 was also p1 in G8 and so on. This enabled the same participants to be distinguished for coding and analyses purposes across the condition of group size.

7.3.1 Speech coding

Once all four stages of gesture coding were complete, speech was coded as either on or off in accordance with the procedure described in the pilot chapter. Speech coding was for all members (5 or 8) within a group who were distinguished by the labels p1-p8 as described in the gesture coding section.

7.3.3 Coding Tool

The coding scheme, actual coding and gesture frequency counts were managed within Nite (Carletta, Evert, Heid and Kilgour, 2005). Gesture/speech codes were first defined within the programme. The beginning and end of every gesture/speech event was then coded. As in both the pilot and dyad experiment, coding was in real time by mouse clicking pre-defined events housed in boxes on the interface (see chapter 4 for screenshot).

Once all group members of every group were coded, queries within NXT were run to pull out the start and end times of the gesture/speech events. Gesture/speech queries were like those in the pilot (see chapter 4) but modified for the extra gesture categories and for additional group members.

Speech and gesture events were then counted. The start and end times of speech events were also extracted and exported to a spreadsheet so that speech durations could be calculated. Queries for gestures counted the frequency of the six different gesture types, splitting the iconic type by viewpoint. Queries were also run on still number/referent to identify gestures made to the same reference.

7.4 Analyses

7.4.1 Gesture and speech patterns

The coding scheme applied to the groups came form the dyad experiment. Although this scheme was already in place, as coding analysis was to be data driven, the gestures that emerged were the first point of interest.

The production of gesture and speech was explored in relation to the condition of group size. First, the interaction times for G5 and G8 were compared. Gesture events for each type of gesture (with the exception of metaphorical gestures which because of their infrequency were excluded from the analyses) and for every participant in all of the groups were then counted to give the total amount of gesture for each of the groups. In the first instance, these counts were used to calculate and report gesture proportions by gesture type. For further analyses, the total amount of speech (or speech duration) for every participant was calculated. This was done by extracting every speech event and subtracting the end point from the start point to give the duration of the speech event. Durations were totalled to give the total amount of speech for every participant.

To explore the relationship between gesture and speech**,** the overall amount of speech and gesture, for every group member, was correlated. To further explore gesture in relation to speech the proportion of speech for each group member was ranked, from highest to lowest, within every group. Ranks were then averaged across all groups of the same size and distributions plotted. Gesture proportions, for all six gesture types, were also ranked by % amount so that gesture patterns could be explored independently of speech. Ranking the contributions and plotting them in this way, also afforded a look at the way in which gesture and speech were distributed across group members.

The explore the production of gesture and speech in relation to the condition of group size, gesture rates were calculated per minute of total amount of speech and gesture frequencies as a function of number of stills. As the design of the experiment was a mixed within and between design, to make the rates and frequencies analyses more powerful only members participating in both groups were used. Group member labels (p1-p5) were given in accordance with the onset of speech in the first group communication. They were then kept constant across the condition of group size so that, for example p1 in G5 was also p1 in G8. This allowed participants to be directly compared across conditions. This fully within analysis, allowed rates from the same members, across the two different sized groups, to be directly compared. This meant data from thirty six of the ninety six participants was not included in the rates analyses. To get the gesture rate per minute of speech, gesture counts for the same 5 participants from both group conditions, were divided by the total amount of speech. This was done for every group. To get the gesture frequency as a function of number of stills, gesture counts for the same 5 participants from both group conditions, were divided by the number of stills each participant received. Again this was done for every group and for all of the gestures included in the analysis.

Alongside gestures rates, average amounts of speech for every participant were calculated.

As the Kolmogrov-Smirnov test of normality is recommended for larger data sets, this was applied to the rates of all six gesture types and to the average amounts of speech. As there was a mix of distributions, both the non parametric Wilcoxon's signed rank test and parametric t-test were applied to test for any difference in gesture rates and speech durations on the two group size conditions. The exact test in the Wilcoxon signed rank test was used to give more accurate test results (Field, 2005).

7.4.2 Alignment and emergence of conventions

To investigate the issue of alignment, gesture rates per minute of total amount of speech were inter-correlated within the groups for all of the gesture types. Correlations were carried out over the same five members who participated in both group conditions to give a within subjects analysis. Rates and frequencies were inter-correlated within the groups for all of the gesture types. In order to inter-correlate all possible combinations of pairs within a group, p1 was correlated with p2-p5; then p2 with p3-p5; then p3 with p4-p5 and p4 with p5. This gave a total of X10 paired correlations for each of the five within group members for every G5 and every G8. Correlations were then conducted across all of the groups, since correlations by groups would have had too few data points having only the ten possible combinations. This gave one large correlation across all of the groups for each of the gesture types. The same group member labels (p1-p5) were used as in the original rates and frequencies analysis where they were given in accordance with the onset of speech in the first group communication, and were matched across the condition of group size.

To get a baseline for alignment, gesture were also correlated between random pairs on the main measure of rate per minute of speech so that each p* was correlated with a random other p* from a different group, and never with a participant from their own group.

The gesture frequency as a function of number of stills was also inter-correlated, in the same way as gesture rates per minute of total amount of speech was correlated, as an additional check for alignment once frequencies were corrected for number of stills.

The strength of the relationship between gestures was measured using the non-parametric Spearmans and the parametric Pearsons correlation. Again, due to a mix of normal and skewed data the use of both was most appropriate.

Instances where gestures referred to the same entity and were aligned on form were identified and qualitatively analysed. For this analysis, gestures were described alongside the speech that accompanied them to provide a sample data set illustrating alignment on gesture form.

7.5 Results

Gesture coding scheme

The six gesture types (Cvpt, Ovp, metaphoric, deictic, place holder and beat) are given by example below (see chapter 4 for more extensive description of these gestures);

7.1 … and [he must have hit the ground] and [his heads buried like an ostrich] …
Ovpt: hand held flat with palm facing downwards moves downwards and bounces as though of the ground.
Cvpt: both hands are flat with palms facing inwards and are lifted up to the gesturer's head where they pull the head down as though to bury the head.

7.2 … [yeah he's dreaming] that the characters are out of the year book
Metaphoric: hand in point to head makes circular motion.

7.3 well mines was [denis] …
Deictic: point to abstract space on Lh side of gesturer.

7.4 … and [a mushroom].
Place holder: open hand with fingers in claw shape held vertically as though holding the place of the mushroom.

7.5 What's [happening] here …?
Beat: hands rises up then back down.

Social gestures directly referenced addressees, as the example below shows. Here the social gesture was directed towards the girl to the side of the gesturer as she was the group member who spoke about the witch;

7.6 Could have been [the witch?]
Social: hand held in cup shape with palm facing upwards moves out to the left towards the addressee. .

Interestingly, social gestures were more formal and explicit in G8 as compared to G5, like for instance in the following example where the counduit gesture described by McNeill (1992) and Bavelas et al (1992, 1995) occurred with a pointed finger;

7.7 What (requesting clarification about what the other group member said earlier)?
Social: towards group member

It should be noted that the above gesture is different in both form and function from other types of pointing described in the pilot which were not coded (see chapter 4). The gesture above is a social gesture because it has the conduit form, albeit with an additional point, and it accompanies speech which specifically request for clarification about information given earlier in the communication.

Further aspects of interest

Iconic gestures referring to the same entity were often more complex in G8 as compared to G5. To illustrate this, speech and gesture referring to the same action in the two different sized groups is described below and can also be viewed at the following addresses. Both examples 7.8 and 7.9 are first mentions of the pouring activity but the gesture is more complex in G8 as there is additional information about where the substance is poured from (ie the jar);

7.8 … [pouring some sort of white liquid] on to …
Cvpt: clenched hand held with palm vertical rotates round in pouring motion.

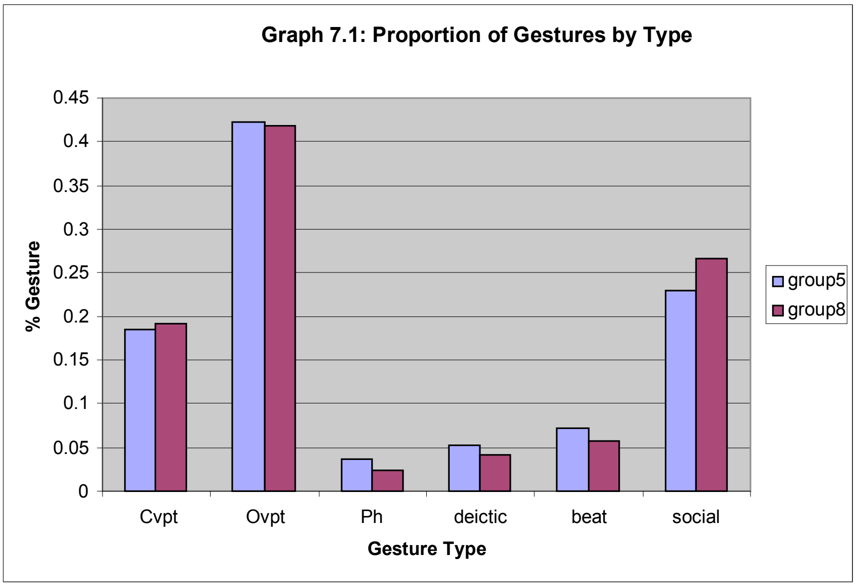
7.9 Seems as though he's like … [pouring a jar of something over a flower …]
Cvpt: both hands semi clenched with palms horizontal as though holding jar and rotate downwards in pouring motion.

Interaction time (duration)

The average length of the interaction in the G5 (M = 21.74 min, SE = 2.57) was significantly longer than that of the average length of the G8 interaction (M = 15.46 min, SE = 1.82). As the Shapiro-Wilks test of normality found the distributions of the durations of the interaction to be normal in both G5 (D(12) = .949 min, p>.05) and G8 ((D(12) = .960 min, p>.05), a t-test was carried out and found this difference to be of significance (t(11) = 2.18 min, p=.021, r = .55), with a large effect size. Being in G5 therefore made for a longer interaction compared to being in G8. Therefore, the group size condition could not be said to influence the length of the interaction.

Gesture Proportions by type

As can be seen from graph 7.1, Ovpt gestures accounted for the highest proportion of gestures followed by social and Cvpt gestures and this was the case for both group sizes. This shows iconic gestures account for different proportions depending on their perspective and the importance of social gestures in the data set. Again irrespective of group size, beat gestures accounted for the next highest proportion followed by deictic and place holding gestures. Comparing within conditions, social and Cvpt proportions were higher in G8 whereas all other gesture proportions were higher in G5.



Amount of speech (duration) correlated with amount of gesture (frequency) for all gesture types.

Descriptive Statistics

The Kolmogorov-Smirnov (K) test of normality results for amount of speech and gesture are presented in table 7.1. As can be seen from the table, all gesture types had non-normal distributions. As these distributions determined whether parametric or non-parametric correlations were used, only the Spearman's correlation coefficient is reported in the next section.

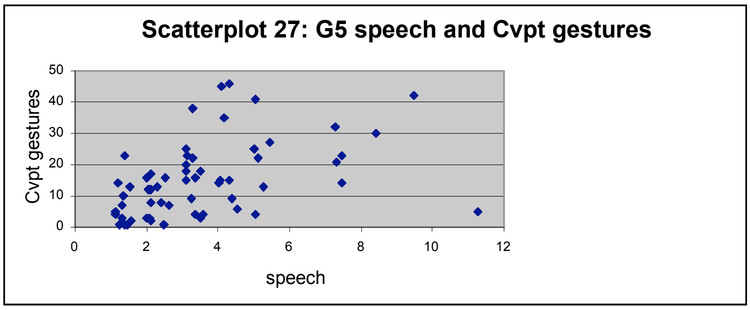
Table 7.1: Kolmogorov-Smirnov test of normality for total amount of gesture and speech

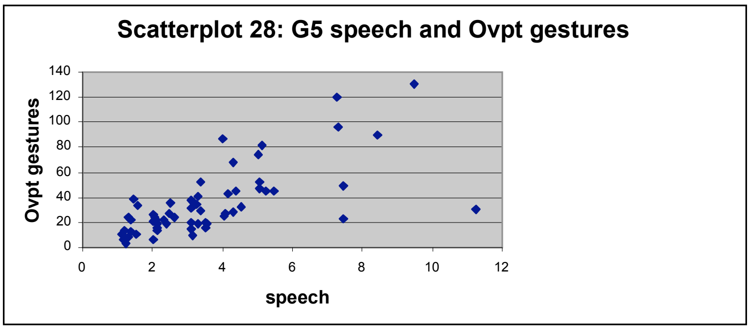
Gesure Type	df	K	Sig<.05
CvptG5	60	.124	.023
OvptG5	60	.177	.000
placeHG5	60	.249	.000
deicticG5	60	.165	.000
beatG5	60	.260	.000
socialG5	60	.133	.010
spchG5	60	.148	.002
CvptG8	60	.182	.000
OvptG8	60	.122	.027
placeHG8	60	.306	.000
deicticG8	60	.261	.000
spchG8	60	.138	.006
beatG8	60	.292	.000
socialG8	60	.211	.000

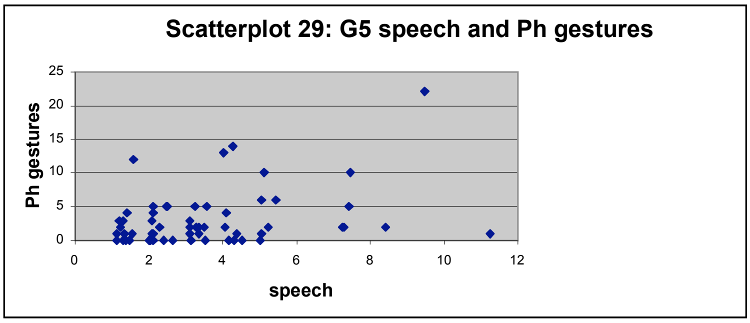
Red = sig different from normal distribution.

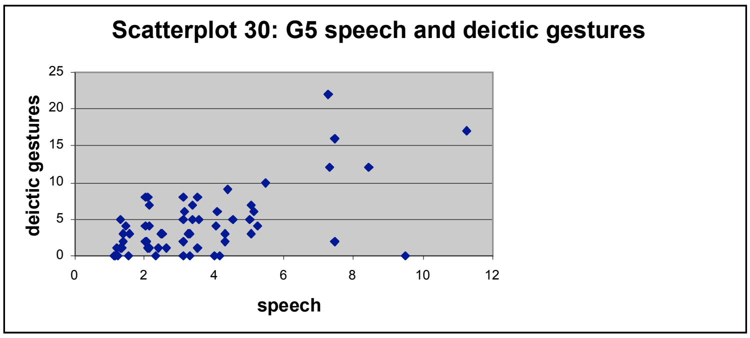
Correlations

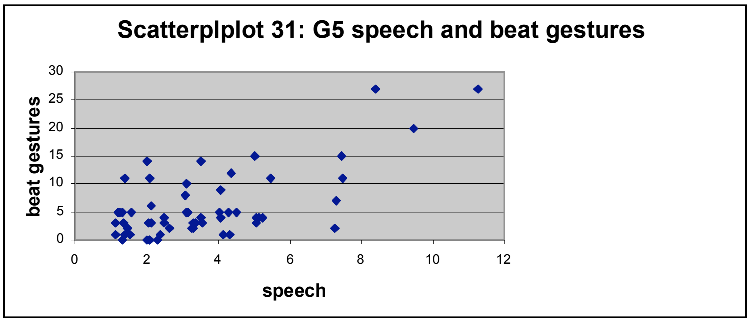
In condition G5, the Spearman's correlation co-efficient found a significant relationship between participant's amount of speech and amount of Cvpt ($r = .526$, $p<.01$), Ovpt ($r = .714$, $p<.01$), Ph ($r = .289$, $p<.05$), deictic ($r = .459$, $p<.01$), beat ($r = .416$, $p<.01$) and social gestures ($r = .713$, $p<.01$). Scatterplots 27 – 32 show the relationships to be positive.

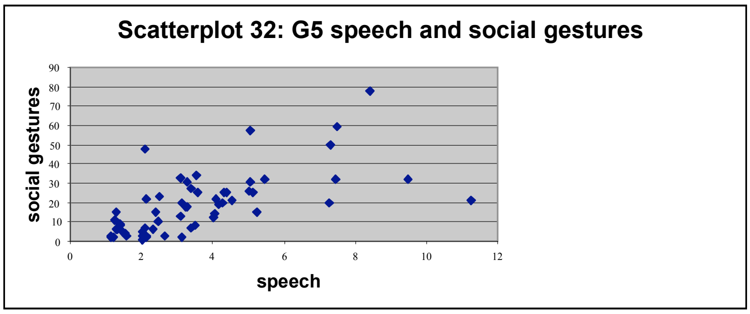




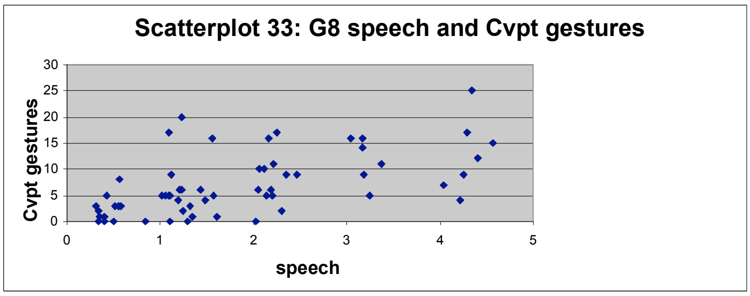


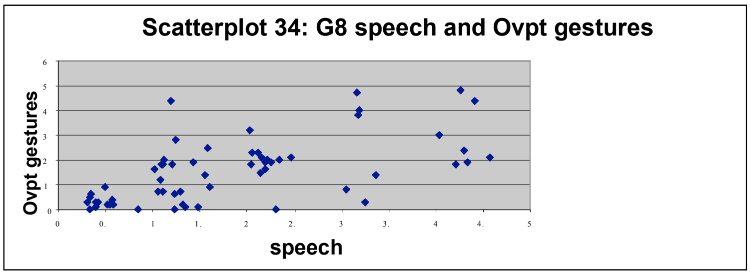


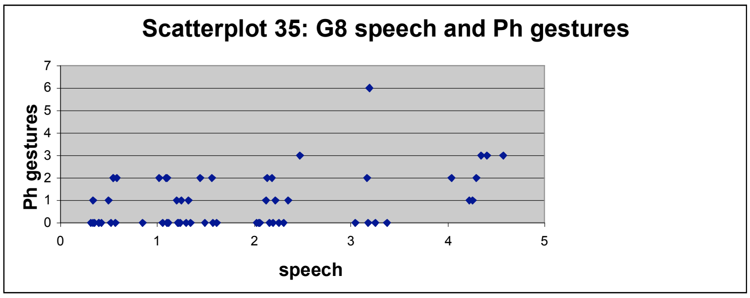


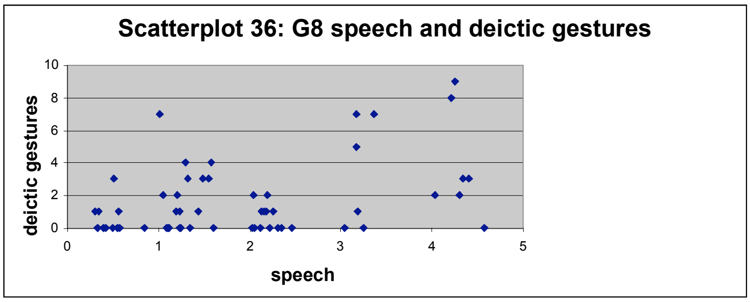


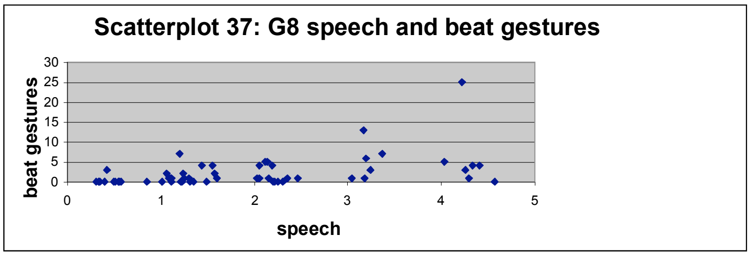
In condition G8, the Spearman's correlation co-efficient found a significant relationship between participant's amount of speech and amount of Cvpt ($r = .615$, $p<.01$), Ovpt ($r = .619$, $p<.01$), Ph ($r = .313$, $p<.05$), deictic ($r = .353$, $p<.01$), beat ($r = .544$, $p<.01$) and social gestures ($r = .622$, $p<.01$). Scatterplots 33 – 38 show the relationships to be positive.

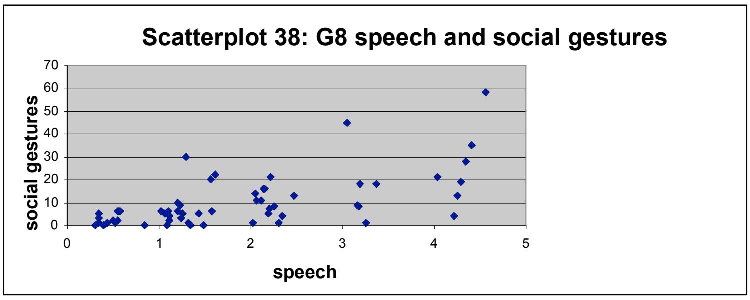

Scatterplot 33: G8 speech and Cvpt gestures
Cvpt gestures
0 5 10 15 20 25 30
0 1 2 3 4 5
speech


Scatterplot 34: G8 speech and Ovpt gestures
Ovpt gestures
speech


Scatterplot 35: G8 speech and Ph gestures
Ph gestures
0 1 2 3 4 5 6 7
0 1 2 3 4 5
speech


Scatterplot 36: G8 speech and deictic gestures
deictic gestures
0 2 4 6 8 10
0 1 2 3 4 5
speech


Scatterplot 37: G8 speech and beat gestures
beat gestures
0 5 10 15 20 25 30
0 1 2 3 4 5
speech

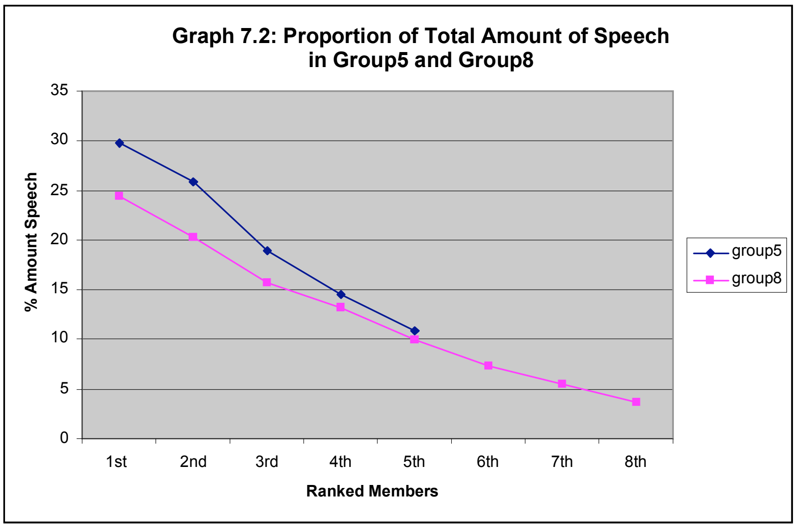


All gesture types were therefore correlated with speech and in both conditions indicating speech should be taken into account in further analysis.

Proportions Ranked by Contributions

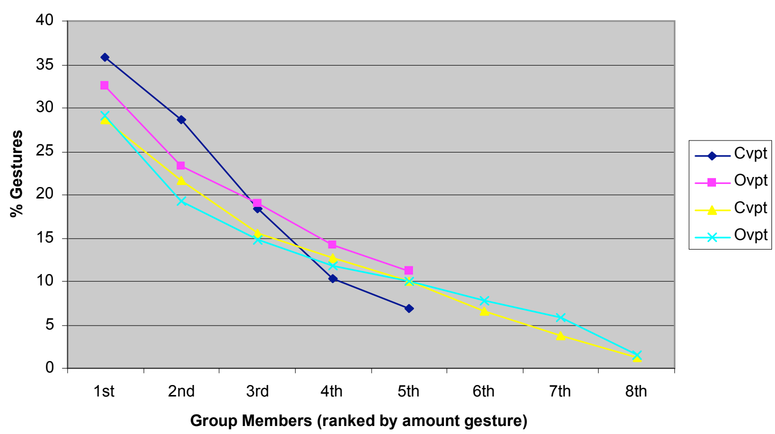
Speech proportions ranked by highest and lowest amount of speech

Plotting the speech proportions (see graph 7.2), by ranking the highest to lowest amounts of contributions across the groups gave a linear curve. In G8, the five highest ranked members accounted for the greatest proportion of contributions (83%) and the three lowest ranked members the remainder whereas in G5 contributions over the five members were more even. G5 members therefore also accounted for a higher proportion than the five highest ranked members in G8. These curves were similar to Fay, Garrod & Carletta (2000), which expands on Fay (2000) with regard to the interpretation of speech curves, though not as exponential, which is the normal curve for free speech. Also, here all group members contributed whereas in free speech and in Fay et al (2000) usually the lowest ranked members did not contribute. Unlike in Fay et al (2000) where large groups had a dominant speaker, a dominant speaker was not evident here.



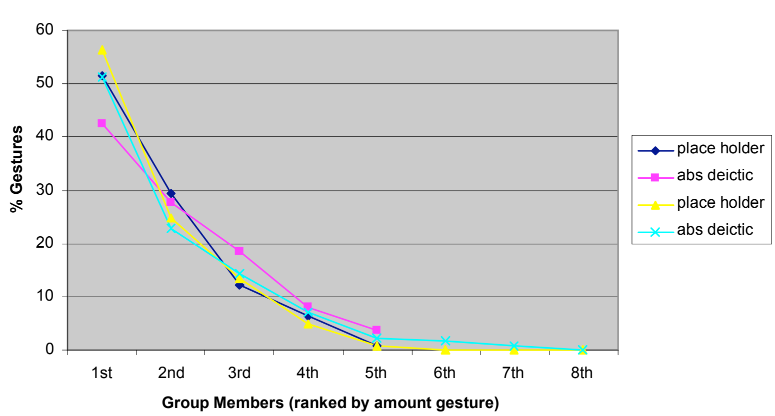
Gesture proportions ranked by highest and lowest amount of gesture

Plotted by the highest and lowest amount of gesture, proportions of Cvpt and Ovpt gestures again produced linear curves and so, followed a similar pattern to that of the speech (see graph 7.3). As in the speech proportions, the five highest ranked members in G8 accounted for most contributions and the three lowest ranked members the remainder. Also like the speech proportions, contributions across G5 members were much more even and accounted for a higher proportion than the five highest ranked members in G8. This was with the exception of the two lowest ranked members in G5 on Cvpt gestures as these dropped below the 3rd and 4th ranked from G8. Both Cvpt and |Ovpt gestures were more variable than speech, as the highest ranked members accounted for a greater proportion of gesture and the lowest a smaller amount. Both gesture types were more variable in G5 than in G8. This was especially so for the Cvpt gesture type which was more variable than Ovpt gestures in G5. Again, a dominant speaker was not evident here.



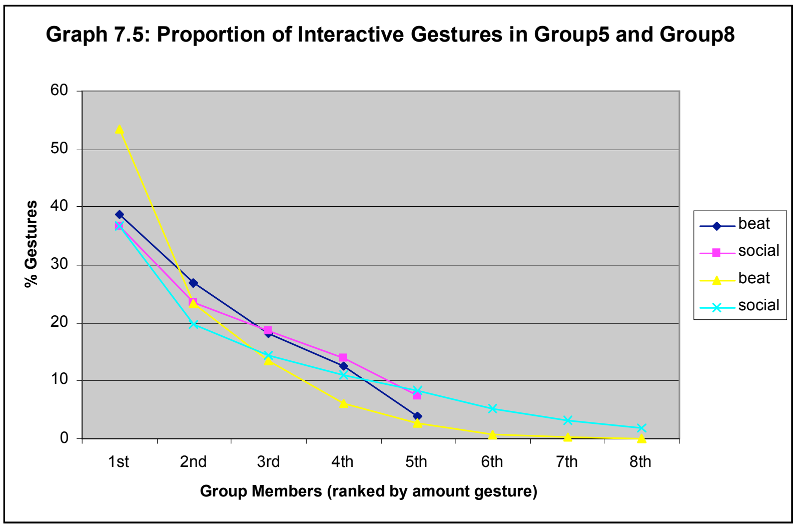
Graph 7.3: Proportion of Iconic Gestures in Group5 and Group8

The proportions of Ph and deictic gestures, ranked by contributions and plotted, produced exponential curves for all of the groups and proportions were comparable within groups of the same size (see graph 7.4). Ph and deictic gestures were not as evenly distributed across the groups as speech and iconic (Cvpt/Ovpt) gestures, as they had a wider range of proportions, indicating they were more variable. These gestures were also more variable in G8 than in G5. There was a dominant gesturer in G8 who accounted for a large proportion of both Ph and deictic gestures but proportion were just above zero for the three lowest ranked members showing some group members did not produce any of these gestures.



Graph 7.4: Proportion of Place Holding and Deictic Gestures in Group5 and Group8

The pattern was mixed for interactive beats and social gestures as proportions in G5 had linear curves whereas proportions in G8 had exponential curves (see graph 7.5). Proportions of both beat and social gestures were more evenly distributed across G5, than across the five highest ranked members in G8. This was especially so for beats where there was a dominant gesturer and some of the lowest ranked members did not elicit any beat gestures.



Average amount of speech

Average amounts of speech were positively skewed in both G5 (D(60) = .148, p<.05) and in G8 (D(60) = .138, p<.05). On the Wilcoxon test, there was a significant difference between G5 (Mdn= 3.13) and G8 (Mdn = 1.46), T = 4, z = -6.088, p< .01. Speech production was therefore influenced by group size with G5 having more than G8.

Gesture rate per minute of speech

Descriptive statistics

The Kolmogorov-Smirnov (K) test of normality results for gesture rate per minute of speech are presented in table 7.2. As can be seen from the table, Cvpt and deictic rates had normal distributions in condition G5 and Ovpt rates were normal in G8 but all other gesture distributions were skewed. The distributions determined whether parametric or non-parametric tests were used and so a mix of tests are reported in the next section.

Table 7.2: Kolmogorov-Smirnov test of normality (gesture rate per minute of speech)

Gesure Type	Df	K	Sig<.05
CvptG5	60	.101	.200(*)
OvptG5	60	.120	.031
PhG5	60	.225	.000
DeicticG5	60	.110	.066
BeatG5	60	.175	.000
SocialG5	60	.126	.019
CvptG8	60	.183	.000
OvptG8	60	.103	.185
PhG8	60	.272	.000
DeicticG8	60	.257	.000
BeatG8	60	.259	.000
SocialG8	60	.138	.006

Red = sig different from normal distribution.

Comparative tests

The Wilcoxon test (see table 7.3) found a significant difference in gesture rate per minute of speech between G5 and G8 for Ph gestures with Ph G5 (Mdn = .579) and Ph G8 (Mdn = 0), T = 14, z = -2.318, p = <.05, r = -.26), with a small to medium effect size. There was a significant difference between deictic G5 (Mdn = 1.04) and deictic G8 (Mdn = .378), T = 20, z = -2.368, p=.01, r = -.29), with a small to medium effect size. There was a significant difference between Beat G5 (Mdn = 1.29) and Beat G8 (Mdn = .445), T = 16, z = -3.155, p=.01, r = -.40), with a medium to large effect size. Ph, deictic and beat gestures were therefore influenced by the group size condition being produced at a higher rate in G5 than in G8.

Table 7.3: Wilcoxon test on gesture rate per minute of speech

Gesture Type	Mdn(G5/G8)	T	z	p
Cvpt	4.34/3.74	25	.773	>.05
Ovpt	9.49/7.94	22	1.480	=.07
Ph	.579/0	14	2.318	<.05, r = -.26
deictic	1.04/.378	20	2.368	<.05, r = -.29
beat	1.29/.445	16	3.155	=.01, r = -.4
social	4.6/4	29	.574	>.05

The t-Test (see table 7.4) found a significant difference in gesture rate per minute of between deictic G5 (M = 1.24, SE = .131) and deictic G8 (M = .909, SE = .180, t(59) = 1.70, p<.05,r =.21), with a small to medium effect. Again then, deictic gestures were influenced by the group size condition being produced at a higher rate in G5 than in G8.

Table 7.4: t-test on gesture rate per minute of speech

Gesture Type	Mean(G5/G8)	SE	t	p
Cvpt	4.64/4.27	.429/.466	.642	>.05
Ovpt	10.05/8.77	.635/.846	1.21	>.05
Deictic	1.24/.909	.131/.180	1.70	<.05, r = .21

Gesture frequency as a function of number of stills

Descriptive statistics

The Kolmogorov-Smirnov (K) test of normality on gesture frequency per still number (see table 7.5), found all gesture distributions to be non-normal and positively skewed. As all distributions were non-normal, only non-parametric tests were used and significant findings are reported in the next section.

Table 7.5: Kolmogorov-Smirnov test of normality (gesture rate per still number)

Gesure Type	df	K	Sig<.05
CvptG5	.124	60	.023
OvptG5	.177	60	.000
placehG5	.249	60	.000
deicticG5	.165	60	.000
beatG5	.260	60	.000
socialG5	.133	60	.010
CvptG8	.182	60	.000
OvptG8	.122	60	.027
placehG8	.306	60	.000
deicticG8	.261	60	.000
beatG8	.292	60	.000
socialG8	.211	60	.000

Red = sig different from normal distribution.

Comparative tests

The Wilcoxon test (see table 7.6) found a significant difference in gesture rate per still between Ph G5 (Mdn = 1) and Ph G8 (Mdn = 0), T = 13, z = -2.311, p=.01, r = -.30), with a medium effect size. There was a significant difference between deictic G5

(Mdn = 1.5) and deictic G8 (Mdn = 1), T = 17, z = -1.964, p<.05, r = -.25), with a small to medium effect size. Group size therefore influenced the production of Ph and deictic gestures as these occurred at a higher frequency in G5 than in G8.

Table 7.6: Wilcoxon test on gesture rate per still

Gesture Type	Mdn(G5/G8)	T	z	p
Cvpt	6.75/5	26	.687	>.05
Ovpt	12.75/15.5	24	1.134	>.05
Ph	1/0	13	2.311	=.01, r = -.30
deictic	1.5/1	17	1.964	<.05, r = -.25
beat	2/1	18	2.553	<.01, r = -.33
social	3.3/6	24	.504	>.05

Gesture rates inter-correlated within pairs by type (rate per minute)

Descriptive statistics

The Kolmogorov-Smirnov (K) test of normality results for inter-correlated gesture rates by total amount of speech are presented in table 7.7. As can be seen from table 7.7, all distributions in G5 and G8 were non-normal therefore only non-parametric Spearman's correlations are reported in the next section.

Table 7.7: The Kolmogorov-Smirnov (K) for inter-correlated gesture rate per still number

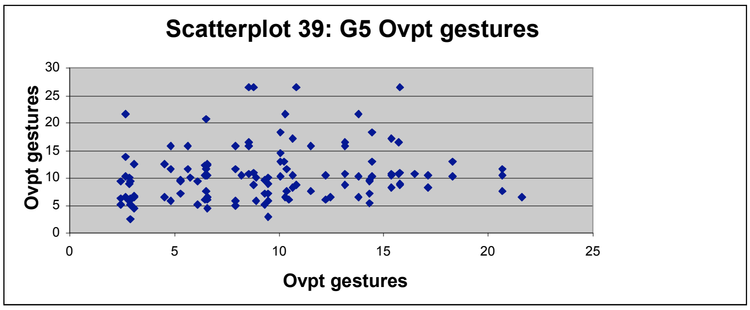
Gesure Type	df	K	Sig<.05
G5 cvpt_a	120	.091	.015
G5 cvpt_b	120	.126	.000
G5 ovpt_a	120	.105	.003
G5 ovpt_b	120	.163	.000
G5 social_a	120	.149	.000
G5 social_b	120	.198	.000
G5 ph_a	120	.092	.015
G5 ph_b	120	.174	.000
G5 deictic_a	120	.229	.000
G5 deictic_b	120	.212	.000
G5 beat_a	120	.096	.009
G5 beat_b	120	.132	.000
G8 cvpt_a	120	.164	.000
G8 cvpt_b	120	.181	.000
G8 ovpt_a	120	.093	.013
G8 ovpt_b	120	.128	.000
G8 social_a	120	.220	.000
G8 social_b	120	.297	.000
G8 ph_a	120	.153	.000
G8 ph_b	120	.146	.000
G8 deictic_a	120	.255	.000
G8 deictic_b	120	.329	.000
G8 beat_a	120	.234	.000

G8 beat_b	120	.288	.000

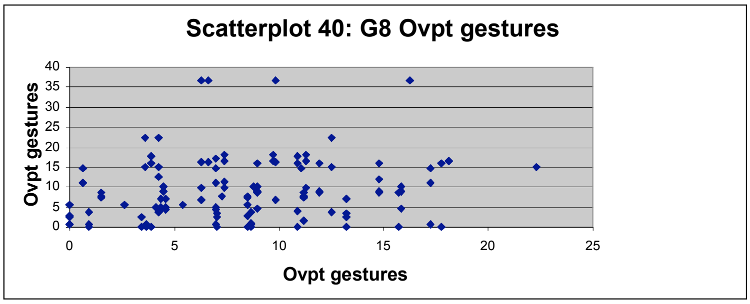
Red = sig different from normal distribution.

Correlations

In condition G5, Spearman's correlation co-efficient found a significant positive (see scatterplot 39) relationship between p1-p5's Ovpt rates ($r = .206$, $p < .05$).



In condition G8, Spearman's correlation co-efficient found a significant positive relationship (see scatterplot 40) between p1-p5's Ovpt rates ($r = .192$, $p < .05$).



As rates per minute were correlated on Ovpt gestures in both G5 and G8, when one of the communicators in a pair produced these gestures the other was more likely to do so. Those communicating therefore aligned on the rate, or how fast, they produced Ovpt gestures.

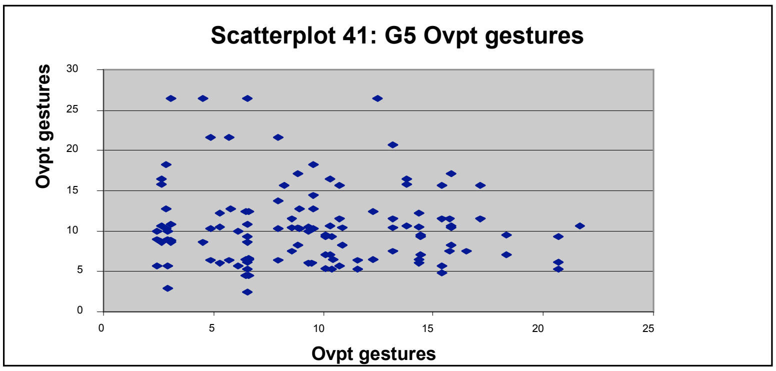
Gestures correlated between random pairs by type (rate per minute)

Descriptive statistics

As distributions were the same as for gesture rates inter-correlated within pairs by type (rate per minute) they were not reported again here. As with inter-correlations, all distributions were non-normal and so only Spearman's correlations are reported in the next section.

Correlations

In condition G5, Spearman's correlation co-efficient found a significant positive (see scatterplot 40) relationship between p1-p5's Ovpt rates ($r = .206$, $p < .05$).



In condition G8, Spearman's correlation co-efficient found no significant relationships between p1-p5's gesture rates.

As Ovpt gesture rates per minute were correlated between random pairs in G5, when one of the communicators in a group produced these gestures another in a random group was more likely to do so. Rather than suggesting that non-communicating groups aligned on the rate, or how fast, they produced these particular gestures, this provides a baseline measure of alignment. Such random alignment must be taken into account in interpreting any alignment between pairs. As no gesture types were correlated in G8, there was no evidence of random alignment on gesture rates in this condition.

Gesture correlated by frequency as a function of number of stills

Descriptive statistics

The Kolmogorov-Smirnov (K) test of normality results for inter-correlated gesture frequencies as a function of number of stills are presented in table 7.8. As can be seen from table 7.8, all distributions in G5 and G8 were non-normal therefore only non parametric Spearman's correlations are reported in the next section.

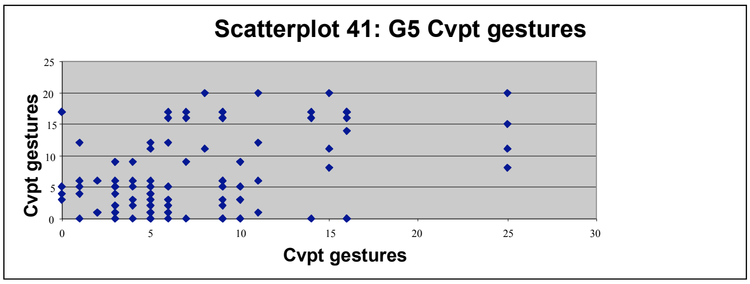
Table 7.8: The Kolmogorov-Smirnov (K) test of normality for inter-correlated frequency as a function of number of stills

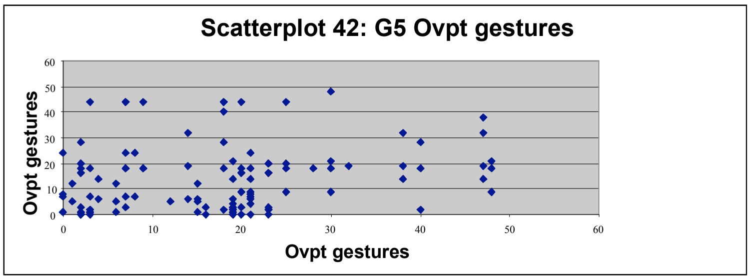
Gesure Type	df	K	Sig<.05
G5 cvpt_a	120	.170	.000
G5 cvpt_b	120	.198	.000
G5 ovpt_a	120	.121	.000
G5 ovpt_b	120	.141	.000
G5 social_a	120	.266	.000
G5 social_b	120	.339	.000
G5 ph_a	120	.177	.000
G5 ph_b	120	.245	.000
G5 deictic_a	120	.259	.000
G5 deictic_b	120	.357	.000
G5 beat_a	120	.272	.000
G5 beat_b	120	.275	.000
G8 cvpt_a	120	.134	.000
G8 cvpt_b	120	.112	.001
G8 ovpt_a	120	.171	.000
G8 ovpt_b	120	.196	.000
G8 social_a	120	.246	.000
G8 social_b	120	.250	.000
G8 ph_a	120	.134	.000
G8 ph_b	120	.150	.000
G8 deictic_a	120	.241	.000
G8 deictic_b	120	.244	.000
G8 beat_a	120	.191	.000
G8 beat_b	120	.162	.000

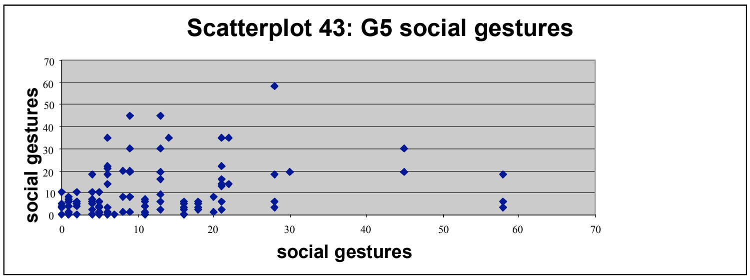
Red = sig different from normal distribution.

Correlations

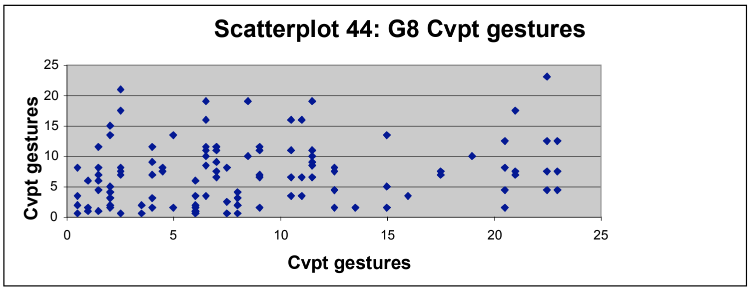
In condition G5, Spearman's correlation co-efficient found a significant relationship between p1-p5's Cvpt rates (r = .196, p<.05), Ovpt rates (r =.255, p<.01) and social rates (r = .356, p <.01). Scatterplots 41- 43 show the relationships to be positive.

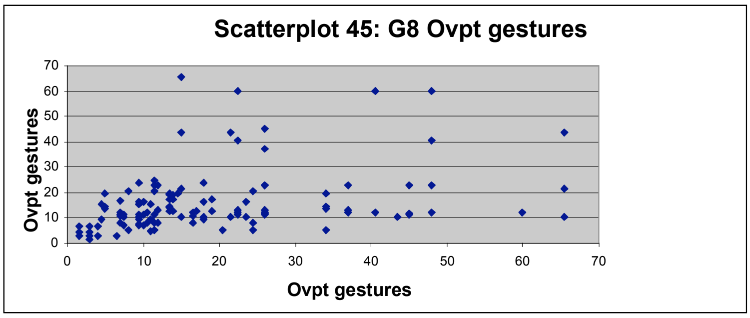


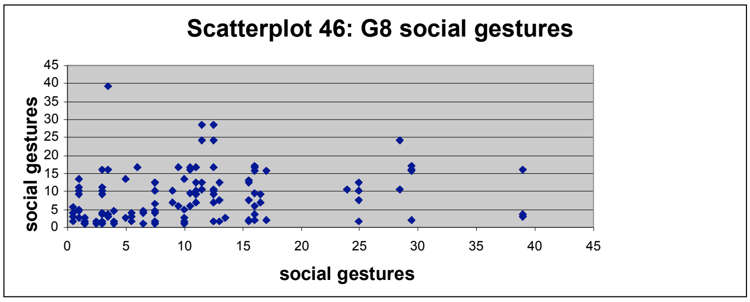




In condition G8, Spearman's correlation co-efficient found a significant relationship between p1-p5's Cvpt rates (r = .274, p<.01), Ovpt rates (r =.437, p<.01) and social rates (r = .340, p <.01). Scatterplots 44- 46 show the relationships to be positive.







In both G5 and G8, gesture frequencies were correlated on Cvpt, Ovpt and social gestures indicating those communicating aligned on amounts of these gestures, even when the amount of information they had to talk about was controlled for.

Descriptive analyses of Alignment.

Alignment on gesture form also occurred in instances when group members referred to the same entity. In example 7.10, p1 is referring to a scene p2 spoke about earlier when p2 interjects what is being said with important information about the potion and p1 agrees on this information at which point they align;

Example 7.10

p1: he's dreaming right … yeah … he's sleeping and … and he [dreams that this dinner lady gives him this] potion that makes …

Cvpt: hand grasped with palm vertical and reaching forwards to pick up then moves in direction of mouth as though to drink.

p2: the [dinner lady gives him that]
Cvpt: hand grasped with palm vertical and reaching forwards to pick up then moves in direction of mouth as though to drink.

In example 7.11, p1 and p2 refer to a cheering action that occurred in both their stills and when they do they align on the gesture form. When p3 refers back to the cheering action that occurred in both p1 and p2's stills she also aligns on the gesture form and then p1 refers again later with the same form;

Example 7.11
p1: … [going like this] …
Cvpt: both hands held flat with palms vertical moved in an upwards direction towards head.

p2: … with [everyone going] …
Cvpt: both hands held flat with palms vertical moved in an upwards direction towards head.

p3: … be like, [oh the glory of the flower it smells so nice I'm gonna collect] it …
Cvpt: both hands held flat with palms vertical moved in an upwards direction towards head.

p1: … [and be like oh] that stinky flower
Cvpt: both hands held flat with palms vertical moved in an upwards direction towards head.

The following example of Ovpt alignment comes from an interaction in G8. Here p1 first outlines the speech bubble in her still and gives it's location at the top of the still. p2 then uses a different gesture form to indicate the blank spaces in his still where speech was cut out. p3 then clarifies where the bubble is and as he does so aligns on p1's original gesture form for outlining the bubble. In p3's next gesture, p3 combines the gesture forms of p1 and p2 and towards the end of the interaction, p3 returns to p1's form when asking p1 to clarify information leading p1 to produce the same gesture.

Example 7.12

p1: why didn't they show the bubble at the top? … It was cut out, .. well I just presumed ..

Ovpt: finger point raised in upwards direction outlines bubble.

p2: well like, we had one earlier and they just [cut out all the text scene]

Ovpt: flat hand moves from left to right to indicate blank space.

p3: ah, … if the [bubbles at the top] …

Ovpt: finger point raised in upwards direction outlines bubble.

p3: it could [cover everything that came above it] so …

Ovpt: flat hand raised in upwards direction moves from left to right then right to left to indicate blank space.

----

p3: [did the thought bubble cover the whole top of the panel?]

Ovpt: finger point raised in upwards direction outlines bubble.

p1: [uhm, no … like that] .. yeah abot half of it.

Ovpt: finger point raised in upwards direction outlines bubble.

7.6 Discussion

The coding scheme from the dyad experiment was applied to the groups. This scheme consisted of iconic gestures (broken down by Cvpt/Ovpt viewpoint), metaphoric, deictic, Ph, beat and social gestures. Whereas Cvpt, Ovpt, metaphoric, Ph and deictic gestures were content gestures, beat and social gestures were non-content gestures. As in the dyad data set, social gestures were present in the group data set highlighting their importance in interactive communication and the under inclusiveness of the pilot coding scheme which did not have this category.

To measure the level of interaction across the two groups, the interaction time alongside the production of speech and gesture was compared. It was predicted that discussion in G5 would be more collaborative than discussion in G8 and this would be reflected in interaction times and production rates. Interaction times were expected to be longer in G5. More beat and social gestures were expected in G5, as these gestures provide feedback about the communication, alongside more Ph and deictic gestures, which rely on visibility and can serve as reduced expressions. On the other hand, more Cvpt and Ovpt gestures were expected in G5 as compared to G8 and the same was predicted for speech. These predictions were based on gesture, speech and graphical sign findings in one and two person communications (see chapter 1 for a review) combined with Fay's (2000) group findings. Fay (2000) found that in smaller groups, when there was ample opportunity to monitor feedback, speech was like interactive dialogue whereas, in larger groups, when increased membership of the group reduced opportunities for monitoring feedback, speech was like monologue.

Interaction times were longer for G5 than G8 suggesting G5 had a more collaborative communication. Irrespective of group size, Ovpt gestures accounted for the largest proportion of gestures followed by social and Cvpt gestures. Beat, deictic and Ph gestures followed. Here the importance of social gestures is emphasised, as these accounted for the second largest proportion of gestures. It is also of interest that Ovpt gestures accounted for a higher proportion than Cvpt gestures. Proportions of social and iconic gesture, broken down by perspective, are in line with the dyads. Across the groups, G8 accounted for a larger proportion of Cvpt and social gestures than G5 whereas G5 accounted for a larger proportion of Ovpt, Ph, deictic and beat gestures. This difference across the groups is interesting as it suggests social and Cvpt gestures occurred more often, in comparison to other gesture types, in G8 and so goes against earlier findings (Bavelas et al, 2007). This will be looked at further in the more precise rates and frequencies analyses which take into account the speech and amount of information.

As predicted, speech was positively correlated with all of the six gesture types (Cvpt, Ovpt, Ph, deictic, beat and social) indicating gestures, like speech (see Fay, 2000), are signs that can be taken as a measure of the communicative style. As in the dyad experiment, speech was correlated with non-content beat and social gestures as well as with content gestures. This again shows non-content gestures, as well as content gestures, were related to speech. As in the dyads, there was no evidence that the relationship between non-content gestures and speech was weaker than the relationship between content gestures and speech. Combined with more qualitative examples from the data set, where non-content gestures were found to directly reference addressees, this highlights the importance of treating non-content gestures as communicating signs even though they do not communicate information about content.

Looking at the way speech and gesture were distributed across group members also highlights the relationship between these two signs but importantly, distributions show how this relationship differs for different types of gesture. Overall, distributions were in line with the prediction that communication would be more collaborative in G5. Both speech and iconic gestures (broken down by Cvpt/Ovpt) had similar patterns (in the form of linear distributions), and in both group conditions, therefore the production of gesture and speech was related. As members in G5 had a more linear curve than the top five members in G8, the distributions show a difference in the level of interaction across conditions, with speech communication being more collaborative in G5 than in G8.

It is interesting to note that speech curves here were not exponential, like they are in free speech (see Fay et al, 2000). Unlike in free speech, where communicators can choose to make a contribution or not, all group members had to contribute in this task as all had still information to share in order to solve the communication problem. This also explains why all members made contributions here, whereas the lowest ranked members usually do not in both free speech and Fay et al (2000), and why there was no dominant speaker/gesturer.

Ph and deictic gestures showed a different pattern to speech as they were less evenly distributed across group members. This was more so the case for G8, where there was a dominant gesturer and some members who did not elicit any gestures. Again then, the distributions show gesture communication was more collaborative in G5 since members made more equal contributions than in G8. They also show Ph and deictic gestures were less tied to speech than Cvpt/Ovpt gestures.

For interactive beat and social gestures, distributions were mixed across the group size condition as G5 proportions produced more even curves than G8. Proportions of both beat and social gestures were then more evenly distributed across G5 members than across the five highest ranked members in G8. This was especially so for beats where there was a dominant gesturer and some members who did not elicit any beat gestures. Like the distributions for speech all other gesture types then beat and social distributions indicate communication was more collaborative in G5. As G5 distributions were as evenly distributed as speech, Cvpt and Ovpt distributions, gestures in this condition were also more ties to the speech. This suggest communication may be better co-ordinated in G5 as in a well co-ordinated interaction, the gestures involved in co-ordinating it, like beat and social gestures, should be organised well with other signs in the communication.

To test for any effect of group size on the emerging communication style, as in the dyad analyses, gesture rates were calculated per minute of speech alongside the gesture frequency as a function of number of stills. Since average amounts of speech were calculated for the rates per minute of speech measure, these were also reported. In line with predictions, both the rate and amount of Ph, deictic and beat gestures were higher in G5 than in G8. These findings are in line with Bavelas et al (1992; 1995; 2007) who found rates of beat and deictic gestures to be higher in dialogue when it had a visible component. Higher rates of these gestures again show communication in G5 was more collaborative than communication in G8. With more beat gestures produced in G5 than in G8, this effect is in the opposite direction to that found in the dyad experiment where more beat gestures were produced in S4 (equivalent to G8) than in S5 (equivalent to G5) on the rate per minute measure. The finding for beat gestures in the group experiment can therefore be taken as was predicted.

An unexpected result was finding rates of social gestures to be comparable across conditions. This is odd considering beat and social gestures, when treated as one gesture category, are found to be produced at a higher rate in dialogue when there is a visible component (Bavelas et al, 1992; 1995; 2007) and especially since this fits alongside Fay et al's (2000) findings for speech.

Social gestures are imagistic like iconic Cvpt/Ovpt gestures and interestingly, findings for these gestures are also at odds with earlier findings. Here, instead of more being produced in the monologue condition (G8), as was predicted if the sign reduced, Cvpt/Ovpt content gestures were produced at a comparable rate across G5 and G8. However, although studies have found rates of production to differ depending on whether communication is in dialogue or monologue, studies have had mixed results. Some studies show iconic gestures are produced at a higher rate in monologue (Bavelas et al, 1992; 1995) whereas another finds they are produced at a higher rate in dialogue (Bavelas et al, 2007). These mixed findings may be a result of how the sign is reduced. Rather than being reduced in terms of the production rate, Cvpt/Ovpt gestures may undergo more qualitative changes. This is in line with Bavelas et al's (2007) finding where visibility influenced the size, deixis and perspective of iconic gestures and graphical findings of Garrod et al (2007) where signs became more symbolic over the course of an interaction. In addition, qualitative changes can occur across different channels of communication (Holler & Stevens, 2007). In line with this**,** Cvpt/Ovpt gestures in this data set, were often more complex in G8 than in G5.

Descriptive analyses of social gestures also show these gestures undergo more qualitative changes. Rather than reducing in terms of their overall production, these gestures changed to be more or less explicit. This is in line with Fay's (2000) findings in speech where in larger groups speech to do with managing turns was more explicit in larger groups, by say using the person's name, rather than using you or even an anaphoric he/she. This more explicit turn taking in speech disrupted the flow of turn taking and is what Fay (2000) suggest made for the monologue style of communication. Similarly, social gestures in G8 were often accompanied by a point to explicitly reference addressees. It seems then that, as social gestures are involved in turn taking (Bavelas et al, 1992; 1995), they too became more explicit in larger groups. This therefore suggests the turn taking process, and therefore the overall interaction, was not as smooth in G8.

With gesture signs being more explicit in larger groups, and therefore carrying more information, this finding is in fact the opposite of what was predicted for social gesture production as, based on findings in gesture research, the production of social gestures was predicted to fall in larger groups. It does however fit with dialogue theories of communication that suggest the provision and monitoring of feedback will be more difficult in monologue style communications.

As was the case for beat gestures in the dyads then, quantitative analyses were not sensitive enough to tap into changes in the imagistic gestures. This again shows how particular gestures may be better suited to certain types of analyses. Being complex in form, imagistic gestures can undergo substantial change so that they reduce, in terms of the amount of information they carry, without the number of gestures being reduced. However, as non-imagistic gestures are less complex in form, to change and reduce, in terms of the amount of information they carry, the amount of gesture drops. This finding also highlights the need to distinguish gestures by how (imagistic or not) the gesture conveys information as well as what (content/non-content) information the gesture conveys. Like deictic gestures, which were influenced differently from other content gestures but in the same way as non-content gestures (see chapter 4), social gestures were influenced differently from other non-content gesture but similarly to iconic content gestures.

To investigate the qualitative reduction of both iconic and social gestures further, it would be interesting to carry out an overseer experiment on gestures elicited across the two groups at different points in the interaction.

It is likely that speech also reduced in a qualitative way as it too was at odds with predictions and earlier findings. Although the focus was on gesture, it was predicted that the average amount of speech would be greater in G8 than in G5 however, the average amount of speech was greater in G5. Like imagistic gestures, the speech signal may have reduced in terms of contributions being shortened, but as speaker turns were the measure of speech, such qualitative changes would not have picked up**.** Another possible explanation for finding more speech in G5 than in G8 is that the increased amount of information in G5 (plus 2 stills) evoked more speech. However, this is unlikely for two reasons. The first is that, if the increase in average amount of speech was an effect of more information, then gestures closely linked to speech (Cvpt/Ovpt) should also have increased. In addition, the group design and specifically effects of varying the amount of information was checked in the dyad design before applying it to the groups. It should be noted that no direct comparison can be made between the group and dyad experiment since they were different experiments. For instance, each participant in a dyad had substantially more information to communicate than each participant in a group. However, by testing for any effect of amount of information, the dyad experiment served as a check on the group design.

Overall, gesture patterns, rates and frequencies here suggest smaller groups elicited a dialogue style of communication whereas larger groups elicited a monologue style. They therefore fit with the prior research findings of Fay (2000) on the speech signal and with dialogue theories of communication. In particular, because these styles of communication are elicited out with two-person communication and in groups, the findings fit with the alignment model (Pickering & Garrod, 2004) outlined in chapter 2. This model states that, representational alignment is necessary for collaborative communication and that because alignment occurs through an implicit process, it can occur more or less depending on the interactional context. It can therefore occur to some degree in groups and more or less depending on group size. However, as dialogue is the optimum interactive context alignment should be most evident in dialogue styles of communication, such as in G5.

To explore whether alignment did occur, rates and frequencies of the same gesture types were inter-correlated within members of the same group. To get a baseline for alignment, they were also correlated between members of different groups on the main measure (gesture rates per minute). As predicted, gesture rates and amounts within members of the same groups were highly correlated in a positive direction indicating that group members aligned on their gestures use. By rate, Ovpt gestures were inter-correlated in both G5 and G8 whereas when correlated between members of different groups Ovpt gestures were correlated in G5 but not G8. Gestures were again correlated when still number was taken into account with Cvpt, Ovpt and social gestures being correlated by frequency in both G5 and G8. As predicted then, Ovpt gestures were more highly correlated than Cvpt gestures.

Rather than gestures aligning more in G5 than in G8, correlations were comparable across group size and so, alignment on rate and frequency was comparable. However, the finding of alignment in G5 on Ovpt gestures must be interpreted with caution since there was also random alignment in Ovpt gestures between non-communicating groups. The alignment model can however accommodate such a finding. By proposing that alignment is implicit and therefore, an automatic process based on a simple priming mechanism, depending on how automatic a process alignment is, it could occur in a group setting where the communication is in serial monologue as well as in a group setting where the communication is in a dialogue style. As well as being of interest for the model in terms of the automaticity of alignment, this is also a reminder that group communication is an approximation of dialogue and monologue situations.

Qualitative analyses on gesture form also indicated this kind of alignment occurred in both G5 and G8. By identifying and describing instances of gesture alignment on the same reference the types of alignment that occurred were explored.

Chapter 8: General Discussion

This thesis explored the effects of group size on gesture communication. As empirical research has so far only looked at the effects of group size on spoken communication (Fay, 2000; Fay et al, 2000), this was a novel approach to studying gesture signs. Predictions were therefore based on findings for one and two person communication in gesture and other signs as well as the group findings for speech. As these findings fit with and are interpreted within Clark's (1996) grounding theory and Pickering & Garrod's (2004) alignment model, predictions were also based on these dialogue models of communication.

A pilot study first tested the feasibility of applying the group design to gesture communication and the feasibility of coding gestures in this context (chapter 4). In the pilot, groups of 5 (G5) and groups of 8 (G8) collaborated to solve the correct order of stills in a comic strip. All group members had the same amount of information (one comic still each) but the information they had was different. They therefore all had to make contributions in order to complete the task. In order to manipulate the comic type and a memory component of the task, G5 and G8 members took part in two separate collaborative tasks.

Although no predictions were made for the pilot, there were some expectations based on earlier findings and these expectations are the eventual predictions for the group experiment. In Fay's (2000) group study of speech communication, speech in small groups was like dialogue whereas it was like monologue in larger groups. This was evident from the type of speech and patterns that occurred, as these mirrored speech findings from monologue and dialogue in one and two-person communication. For instance, in monologue more information was carried in the speech than in dialogue. These findings fit with dialogue models of communication that say because dialogue allows for a shared representational state, less information needs to be carried in the sign. In particular, the group findings sit well with the alignment model's (Pickering & Garrod, 2004) proposal that an implicit shared representational state can occur outside of two-way dialogue. By proposing a continuum of dialogue the alignment model suggests communication can be more or less like dialogue and monologue and that group communication fits along this continuum.

Like speech, and indeed graphical signs (see chapter 2), gesture signs carry more or less information depending on the interaction and whether the communication is in monologue or dialogue. Signs in general then change depending on the interaction. For instance, content gestures are produced at a higher rate in one-person monologue than in two-person dialogue (Bavelas et al, 1992; 1995). As dialogue models hold interactive communication to be necessary for a shared state to occur, finding more information needs to be produced in gesture signs when in monologue fits with the models. However, gestures providing feedback about the communication are more often produced in two-person dialogue and moreover, when the dialogue has a visible component (Bavelas et al, 1992; 1995; 2007). As the models also hold that interactive communication relies on interactive feedback, finding more interactive gestures, to do with feedback, in dialogue is also in line with them.

However, in the pilot G5 had too little information to make for interactive communication and, as communication was largely in a monologue style, gesture types and patterns were not as expected. The pilot design was also confounded which made interpreting the results difficult. The original group design was therefore modified. The best design for the groups was for members in G5 to have two stills each whilst keeping the information constant (as in the pilot) in G8 and giving group members one still each. This meant collectively, G5 had only a little extra information. In addition The Beano comic was chosen as stimuli since this comic stimulated conversation better than the Tom and Jerry cartoon and the conditions were from memory.

As communication was less interactive than expected in the pilot, the coding scheme generated in the pilot was a monologue style scheme (McNeill, 1992) and would not fit interactive communication. For this reason, a coding scheme based on interactive communication (Bavelas et al, 1992; 1995; 2007) was constructed for the actual group experiment.

To check the new group design, with respect to the effect of varying the amount of information and to get a baseline-coding scheme for interactive communication, the design was modified for pairs (chapter 5). Here, the task was the same as in the group pilot with members of pairs discussing and ordering stills from a comic strip. Again members of a pair had different information but the same amount of information. Pairs took part in two separate communications where each member of a pair had four (matched to G8) and five (matched to G5) stills each depending on the condition. The amount of information therefore varied over the two conditions that were matched to the group design. The dyad experiment found a difference for beat gestures, on the rate per minute of speech measure, with more being produced in S4 than in S5 so these gestures were interpreted with caution in the group experiment. As no other effect was found for any other gesture type across the conditions, it was concluded that, in the main, varying the amount of information did not influence the interaction. The design could therefore be extended to groups without treating the amount of information as a confounding factor, except for on beat gestures. An inclusive coding scheme was generated from the dyad experiment that could also be applied to the groups.

The dyad experiment was interesting in it's own right in terms of the results. The alignment model (Pickering & Garrod, 2004) predicts that, as well as providing more or less information and indeed different kinds of information, signs used between those communicating come to be aligned. This led to the prediction that gesture signs from the same communication would be aligned and this was found to be the case. With the exception of beat gestures, all gestures were correlated (Cvpt, Ovpt, social, Ph and deictic) by either rate or frequency and in either condition. Pairs of communicators therefore aligned on the production of these particular gesture types so that, when one member of a pair made one of these gestures the other was more likely to do so. The correlations can be taken to indicate alignment in the pairs since no significant correlations were found between random (non-communicating) pairs, which provided a baseline for gesture alignment.

That beat gestures were not correlated, and thus did not align, is interesting considering Bavelas et al (1992; 1995) claim beats, alongside social gestures, directly reference addressees. If this were the case then, beat gestures would be expected to align and possibly more so than some other gesture types. However, although both beat and social gestures are classed as 'interactive', Bavelas et al (1992; 1995) describe social gestures much more fully in their studies than beat gestures. Moreover, social gestures are said to perform a variety of functions. One such function is to clarify information by providing feedback. As beat gestures add emphasis, this too may be a central function of beats. With gestures providing feedback at the outset, there would be less of a need for addressees to respond to the gesture that has served to clarify information at the outset. An alternative explanation is however that a quantitative type of analyses was not sensitive enough to pick up alignment in beat gestures. Further experimentation should consider a temporal analysis to investigate alignment in beat gestures.

Also of interest was that Ovpt gesture, were more highly correlated, and so aligned more often than any other gesture type. This difference observed between Ovpt and Cvpt gestures was explained in terms of the perspective of the gesture. As Ovpt gestures are observer oriented and Cvpt gestures egocentric, this would have made the latter more difficult to align on.

Alignment was also evident on the gesture form of gestures referring to the same entity and qualitative examples of these were given in chapter 5. These examples show that when one member of a pair used a particular gesture for, the other was more likely to use the same form. To empirically test alignment on gesture form, alignment findings in the dyads were followed up in an overseer experiment (chapter 6). Here naïve overseers viewed Cvpt and Ovpt gestures from the dyad experiment referring to the same entity. Two of these gestures were from the same pair and one form a different pair. The overseer's task was to choose the odd gesture out based on the gesture form. Overseers more often chose gestures from different pairs suggesting gestures from the same pairs were more alike, or more aligned, on form than those from different pairs. This was in line with findings from the dyad experiment where gesture rates and amounts were aligned for both Cvpt and Ovpt gestures.

In addition, overseers more often chose Ovpt gestures from a different pair as the odd gesture out than Cvpt gestures from a different pair suggesting Ovpt gestures were more aligned on form than Cvpt gestures. This was also in line with findings from the dyads where Ovpt gestures were more aligned than Cvpt gestures by both rate and amount and fits with the egocentricity explanation for Cvpt gestures. Again, finding fit with predictions of the alignment model, that in interactive communication such as dialogue, signs from the same communication should align.

To further investigate the effects of perspective on gesture alignment, an overseer study looking at how gesturer's use the gesture space would be useful. In such a study, naïve overseers could rate the use of the gesture space as being more or less egocentric. For instance, it would be predicted that Cvpt gestures would be performed within the gesturer's own gesture space (since they often reference the gesturer's own body parts for example) but that Ovpt gestures would be performed in the shared communication (gesture) space. This different use of the communication space, in particular whether the space overlaps and is shared or not by those communicating, may well be what influences gesture alignment. In relation to how the gesture space is used, a qualitative analysis of specific features (such as size) could be used to compare gestures rated as more or less similar in relation to perspective.

After trying and testing the pilot design and then a new design, the group experiment was conducted (chapter 7). Here groups were again to discuss the best order for stills in a comic strip. The experiment consisted of two group size conditions. In groups of 5 (G5), group members received two stills each whereas in groups of 8 (G8), they received only one still each. The same group members took part in the two conditions so that plus or minus three members were either added or subtracted to the group for the second communication. It was predicted that gesture signs would be influenced by group size in a similar way to speech, with small group communication being like dialogue and large group communication like monologue (Fay, 2000). These predictions are also in line with dialogue models of communication and in particular the alignment model (Pickering & Garrod, 2004).

As predicted, interaction times, alongside speech and gesture patterns, showed communication in G5 was like dialogue but that communication in G8 was like serial monologue. Based on the refined coding scheme, which in being more precise included additional gesture categories, predictions for gesture production were refined. It was again predicted that gestures conveying information about the on-going interaction would be produced more often in G5. These were beat and social gestures. However, as content gestures were broken down into four separate categories, predictions differed for gesture types within these categories. It was predicted that deictic and place holding (Ph) gestures would be produced more often in G5 since these gestures rely on the visible component of dialogue and so would encounter problems in G8. As these gestures also function as a reduced expression and so indicate reduction in the sign, they would also be expected more often in G5. It was tentatively predicted that Cvpt and Ovpt gestures would be produced more often in G5 than in G8. This was a tentative prediction because of mixed findings for these gestures (Bavelas et al, 1992; 1995; 2007), the possibility of visibility influencing viewpoint differently (Bavelas et al, 2007) and the possibility of gestures changing in a more qualitative way.

Findings were in line with predictions as more beat (note that the difference between beat gestures could be taken as predicted since it was in the opposite direction to the effect found in the dyads), deictic and Ph gestures were produced in G5. Although Cvpt/Ovpt gestures were produced at comparable rates across group size, as predicted these gestures changed in a qualitative way to provide more information in G8 as compared to G5. Similarly, social gestures were comparable across group size but changed in a qualitative way being more explicit in G8 than in G5. The finding for social gestures is interesting as it fits with Fay's (2000) finding that turn taking cues in speech were more explicit in larger groups. Interestingly, in Fay (2000) these more explicit cues disrupted the turn taking in larger groups and, in line with dialogue theories, was said to be the underlying problem for large group communication. This was therefore what led to the monologue style of communication. Finding the same qualitative change in social gestures then supports the idea that these gestures are involved in turn taking and that they are influenced similarly to speech by group size.

Gesture alignment also occurred on the rate of Ovpt gestures and amount of Cvpt, Ovpt and social gestures. As was predicted from findings in the dyads, alignment occurred more often on Ovpt gestures, as these are less egocentric than Cvpt gestures. Qualitative analysis also found group members aligned on gesture form. Rather than occurring more often in G5, alignment in the groups was comparable. However, as the baseline measure of alignment, measured by randomly correlating groups that had not communicated together, found Ovpt gestures to be correlated in G5, the correlations taken to indicate alignment in G5 must be interpreted with caution.

That G8 members were correlated and so aligned freely of any baseline measure of alignment, whilst alignment in G5 needed to take a baseline measure into account, can be explained within the alignment model in terms of the automaticity of the alignment process. The state of implicit common ground proposed by the alignment model is an automatic process that can happen in the group situation and this contrasts with the grounding process described by Clark (1996) which is a more complex reasoning process based on partner modelling.

This thesis addressed the issue of alignment or convergence, which according to alignment theory is a first step in the conventionalisation of gesture. However, to investigate the conventionalisation process of gesture signs, research should look at the development of gesture signs over the course of communications in different communities as Garrod & Doherty (1994) did in speech.

To conclude, group size was found to influence gesture communication in fundamental ways. The findings from this thesis show gesture signs are like other signs in the way they communicate. They show that what, how and why information is communicated in gesture signs depends upon the interactive setting. Gesture signs and indeed all signs, must therefore be studied in relation to the interactional context. By showing gesture signs to be modified through interactive communication, the findings provide support for dialogue theories of communication. In particular, the findings reported provide support for the alignment model (Pickering & Garrod, 2004) since this model holds that the alignment necessary for collaborative communication can occur, more or less, in interactive situations other than two-person dialogue. They also support the models assertion that signs are modified in and through everyday use to eventually become a routinized part of dialogue. The findings therefore suggest gesture signs, alongside other signs, evolve within the dialogue context. Rather than being arbitrary then, communicating signs are grounded in everyday communication and change, to become fixed routines and possibly even conventionalised, through a process of social cognition.

References

Alibali, M.W., Heath, D.C., & Myers, H.J. (2001) Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, 44, 169-188.

Arnoff, M., Meir, I., Padden. C.A., & Sandler, W. (2008). The roots of linguistic organization in a new language. Interaction Studies, 9(1), 133-153.

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

Bavelas, J. B., Chovil, N., Coates, L., and Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin, 21,* 394-405.

Bavelas, J., Gerwing, J., Sutton, C.,Prevost, D. (2007) Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language,* 58, 495-520.

Beattie, G. & Shovelton, H. (2001). An experimental investigation of the role of different types of iconic gesture in communication: a semantic feature approach. *Gesture*, 1, 129-149.

Beattie, G. & Shovelton, H. (2002). An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *British Journal of Psychology,* 93, 179-192.

Butterworth, B., & Beattie, G.W. (1978). Gesture and silence as indicators of planning in speech. In Campbell, R., Smith. P., (Eds). Recent Advances in the Psychology of Language: Formal and Experimental Approaches (pp347-360). New York: Plenum.

Cassell, J. & McNeill, D. (1991). Gesture and the Poetics of Prose. *Poetics Today*, 12(3), 375-404.

Carletta, J., Garrod, S., & Fraser-Krauss, H. (1998) Communication and Placement of Authority in Workplace Groups – The Consequences for Innovation. *Small Group Research*, 29(5), 531-559.

Carletta, J., Evert, S., Heid, U., Kilgour, J. (2005). The NITE XML Toolkit: data model and query. *Language Resources and Evaluation Journal*, 39(4), 313-334.

Church, R. B. & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, 23, 43-71.

Cohen. A.A., & R.P., Harrison. (1973). Intentionality in the use of hand illustrators in face-to-face communication studies. *Journal of Personality and Social Psychology*, 28, 276-279.

Cohen, A.A. (1977). The communicative function of hand illustrators. *Journal of Communication,* 27, 54-63.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think,* Cambridge, M.A: Harvard University Press.

Clark, H.H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process, *Cognition*, 22(1), 1-39.

Coventy, K.R. & Garrod, S. (2004). *Saying, Seeing and Acting: The Psychological Semantics of Spatial Prepositions.* Hove: Psychology Press.

Ekman, P. & Friesen, W.V., (1969). The repertoire of nonverbal behaviour: Categories, origins, usage and coding. Semiotica, 1, 49-98.

Fay, N (2000). *The effects of group size upon influence and process in group communication*. Unpublished PhD, University of Glasgow.

Fay, N., Garrod, S., & Carletta, J. (2000). Group discussion as interactive dialogue or serial monologue: The influence of group size. *Psychological Science*, 11(6), 487-492.

Field, A.P. (2005). Discovering statistics using SPSS : and sex and drugs and rock ‘n’ roll (3rd edition). London : Sage.

Garrod, S. & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, 27, 181-218.

Garrod, S. & Doherty, G. (1994). Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition*, 53, 181-215.

Garrod, S., Fay, N., Lee, J., Oberlander, J. & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Psychology,* 31, pp 961-988.

Gerwing, J. & Bavelas, J.B. (2004). Linguistic influences on gesture’s form. *Gesture*, 4, 157-195.

Holler, J. & Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, 26, 4-27.

Kendon, A. (1970). Movement coordination in social interaction. *Acta Psychologica*, 32, 1-25.

Kendon, A. (1990). *Conducting Interaction: Patterns of behaviour in focused encounters.* Cambridge: Cambridge University Press.

Kendon, A. (1995). Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics*, 23, 247-279.

Kendon, A. (1997). *Gesture*. Annual Review of Anthropology, 26, 109-128.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kita, S. (1990). *The temporal relationship between gesture and speech: A study of Japanese-English bilinguals.* MS, Department of Psychology, University of Chicago.

Kita, S., & Ozyurek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, 16-32.

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, 24(2), 145-167.

Krauss, R. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science,* 7(2), 54-59.

Morrel –Samuels, P., & Krauss, R.M., (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology; Learning, memory and Cognition*, 18, 615-622.

Krauss, R. M., & Wheinheimmer, S. (1964). Changes in the length of reference phrases as afunction of social interaction: A preliminary study. *Psychometric Science*, 1, 113-114.

La France, M. (1985). Postural mirroring and intergroup relations. *Personality and Social Psychology Bulletin,* 11, 207-217.

McNeill, D. (1992). *Hand and mind, what gestures reveal about thought*. Chicago: The University of Chicago Press.

McNeill, D. (1985). So you think gestures are non-verbal? *Psychological Review*, 92(3), p350-371.

Morford, J.P. (1996). Insights to language from the study of gesture: A review of research on the gestural communication of non-signing deaf people. Language and Communication, 16 (2), 165-178.

Muller, C. (2004). The palm up open-source: A case of a gesture family? In Muller, C. & Posner, R., (Eds). *The semantics and pragmatics of everyday gestures*. (pp 233-256). The Berlin Conference, Berlin: Weidler.

Ozyurek, A. (2002) Do Speakers Design their cospeech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, 46, 688-704.

Parrill, F. (2004). The hands are part of the package: Gesture, common ground, and information Packaging. 7th Conference on conceptual structure, Discourse and Language. Edmonton, Alberta, Canada.

Parrill, F., & Kimbara, I. (2006). Seeing and hearing double: The influence of mimicry in speech and gesture on event representation. *Journal of Nonverbal Behaviour,* 30(4), 157-166.

Pickering, M.J., & Garrod, S. (2004) Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2), 169-190.

Sacks, H., Schegloff, E.A., & Jefferson, G.A. (1974). A simplest systematics for the organisation of turn taking in conversation. *Language*, 50, 696,-735.

Schober , M. F., & Clark, H.H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21(2), 211-232.

Shockley, K., Santana, M., & Fowler, C.A. (2003). Mutual Interpersonal Postural Constraints Are Involved in Cooperative Conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 326-332.