AMIDA

Augmented Multi-party Interaction with Distance Access

`http://www.amidaproject.org/`

Integrated Project IST–033812

Funded under 6th FWP (Sixth Framework Programme)

Action Line: IST-2005-2.5.7 Multimodal interfaces

Deliverable D1.3: Qualitative Analysis of Interactions in Face to Face and Remote Meetings

Due date: 30/04/2009 **Submission date:** 30/04/2009
Project start date: 1/10/2006 **Duration:** 39 months
Lead Contractor: UT **Revision:** 2

Project co-funded by the European Commission in the 6th Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

D1.3: Qualitative Analysis of Interactions in Face to Face and Remote Meetings

Abstract:

This deliverable presents Amida WP1 work on qualitative analysis of interactions in face to face and remote meetings. It covers a number of research areas: subjectivity in decision making discourse, speaker listener interactions, the function of deictic expressions, addressing and turn-taking, the role of annotators in the research in computational models of human conversational behavior. The report contains some recommendations for meeting support technology based on our findings.

Contents

1 Introduction

Amida is a research project that aims the automatic identification and classification of human conversational behavior. The specific application area chosen is that of interactive behavior in small group meetings. Amida will show how the research on automatic multi-modal recognition can be used to support humans to have more effective meetings. The main technology tracks should lead to: (a) the automatic production of meeting summaries, the storage, retrieval and presentation of meeting content, and (b) the support of remote meeting participation.

These technologies require our understanding and models of meeting behavior. Thus one of the aims is to find the appropriate structures of meeting embedded discourse, to provide a basis for computational models of regularities in behavior that can lead to automatic classification of these behaviors (WP4 and WP5). Such models should model for example the relation between speaker's gaze behavior during the utterance and addressing and turn-giving.

Another result of this work package's analytical groundwork is the development of annotation schemes that are used for obtaining annotated data to build supervised machine learning classifiers on. Insight obtained in the mechanisms that play in human-human communication in face-to-face and remote meetings, are usefull for understanding the effects of impaired communication channels (lack of good quality audio and video, caused by network delay or a static camera view) on the quality of the interaction and eventually on the quality of the group and the group's achievements. This informs the engineers what technologies they could build to support meeting behavior, to overcome these communication problems.

The data that we analysed consists of:

- Scenario-based face-to-face design meetings. The AMI corpus Carletta [2007], Post et al. [2004].
- Scenario-based remote design meetings, where 3 people are in the same room and one participant is remote. The b and c meetings of the Amida corpus.
- Scenario-based remote design meetings, with two designers discussing a prototype of a device. The X meetings of the Amida corpus (see section 5)
- Three participant remote meetings with two in the same room and where the remote participant is represented by a robot (section 6).

It is nowadays a platitude to say that conversations are embodied, social behavior, joint and situated activities between speakers and listeners, embedded in all kinds of other activities that the partners are involved in. Understanding conversations involves all these aspects. In the following we shortly review these aspects of conversations.

Communication is embodied Whether we want or not, our body, posture, stance, facial expression always presents in some way or another our internal vegetative and cognitive live, our mood, feelings, the way we receive and evaluate impressions from our environment, for others with whom we are physically present, others for whom we are observable. For others our observable body expresses our personal being. This physical co-presence in sharing with others our practical life is the basis for every form of commu-

nication. Language comes into being where we explicitly use our body to present something to someone else, by means of a gesture. Communication then becomes more explicit through the use of gestures, words, sounds that become pointers to objects, feelings, for others with whom we share the practice in which these gestures are made. Gestures, sounds, pictures, become a language if they are recognized and re-used and memorized in a community for which they stand. Language is essentially a social entity, in that it exists and evolves through interaction in social practice.

Communication is a joint activity While speaking speakers monitor the audience and respond to the signals that addressees and other listeners send simultaneously to the speaker. These signals tell the speaker if and how his words are received. Is it clear what he says? Do they agree with what he says? How do they assess his message? The listeners on their side monitor the speaker not only to see and hear what he is saying but also to see and hear if the speaker is receiving their attention and the way the listener perceive the speaker's message.

Grounding is a central proces in every form of communication Clark [1996], Cahn and Brennan [1999], Schegloff [1982], Traum and Heeman [1996], Traum [1994], Paek and Horvitz [1999]. Grounding is the way that interlocutors establish common ground, or shared belief. They do not start from scratch however when they start a conversation or meet each other. They also build on common ground: assumptions about shared knowledge, rules of conduct, they are members of a social community that shares a language.

Grounding operates on all levels of communication. The physical level is the signal processing level. If there is no channel of communication then there can be no audio or visual contact between the partners. A second level is the level of attention. Does the other part listen to the channel? Is he aware that he is spoken to? Another level is that of understanding: does the other side understand what was said or is there mis-understanding? And finally, there is the level of reception: how did the other receive the message? On all these levels grounding operates and mutual feedback is required. Synchrony, or co-temporality of events is a foundational requirement for grounding. We experience two events that frequently happen at the same time as somehow related to each other. Co-temporality is a strong indication for some "internal" relation, or that both participate in one and the same action. Audio delay in mediated and asynchronous communication leaves more room for uncertainty of mutual participation in the same activity, because we miss the co-temporality of the act of speaking and the re-action, so we are not sure anymore that there is a direct relation between the two. The experience of presence is basically the experience that we have if we press our hand against the wall and at the same time we feel the wall pressing our body. Or, if we open our eyes and see the world as it is at the same time we perform this act.

Grounding is basic in the way speakers select the way they refer to things and ideas. To objects in the environment as well as to ideas or events. Does a local meeting participant include the remote participant when he uses the word "we" ? Using this reference means that the speaker assumes that his listeners know who he refers to. But do they? If someone says "the yellow box" pointing at the slide shown on the slide screen in the meeting room does he care about the remote participant? Whether or not he can use this way to refer to

the object depends on 1) whether he sees the remote participants as a participant (if not an addressee than at least a side-participant) or as an overhearer. Or does he even not care at all about the presence of a remote participant? The same way as people are often not aware of the presence of a camera in a public place.

Communication is situated In conversations partners refer to the situation that they believe to share. In section 5 we analyse how people in a remote design meeting use hand gestures, and point at physical objects and drawings in discussing a design prototype. In section 4 we study eliciting acts, how they are addressed and embedded in the course of action and how we can match elicit act and responses to it. Also responses to eliciting acts sometimes only consist of a non-verbal act. That makes the automatic task of matching these paired actions a multi-modal processing task.

Communication is a social activity Interaction involves individuals, and these are members of several communities. Asking, requesting, warning, they all are social activities, in the sense that they express a relation between individuals. Two peoples can have different roles and they often stand in different relations to each other. Not everybody talks with everybody about everything, not about the same things, and not in the same way either. The way people interact, talk to each other, or not talk to each other, expresses their relation. The stance that a speaker takes towards his addressees, can change with the role and the relation they have. Goffman refers to this changing of stance, and roles, by the word "footing", Goffman [1981]. Participants over the course of their speaking constantly change their "footing". He gives a rough summary of what is involved (Goffman [1981] p.128):

1. Participant's alignment or stance, posture, or projected self is at issue.
2. The projection can be across a strip of behavior that is less long than a grammatical sentence, or longer. Prosodic, not syntactic, segments are implied.
3. Changes can range from gross changes in stance to subtle shifts in tone
4. For speakers, code switching is usually involved, and if not then at least the sound markers that linguists study: pitch, volume, rhythm, stress, and tonal quality.
5. The bracketing of a "higher level" phrase or episode of interaction is commonly involved, the new footing having a liminal role, serving as a buffer between two more substantially sustained episodes.

Here is an fragment from one of the Amida remote meetings, that illustrates how a change of footing or stance is expressed[1].

```
3 A > C(ELI): Lawrence , can you hear us ?
4 D > G(CAU): What does he say ?
5 C > A(INF): Yeah I can hear you ,
6 C > A(ELC): can you ?
7 A > C(CAU): Yeah I can hear you .
8 D > C(ELI): Can you hear me ?
```

1. The appendix contains the complete transcripts of meetings ED1005b and ED1005c. The linenumbers in the fragments that we use in this report to illustrate and that come from these transcripts, refer to the numbers in listings in the appendix.

In 4 speaker D refers to Lawrence - the remote participant - by the third person pronoun "he" - addressing the group, excluding the remote participant from the "we" and "use" of the local group. In 8 his stance has changed. Now, that he is confident that there is a line of communication, he is directly addressing Lawrence, using the second personal pronoun "you". Moreover, he does not say "can you hear us", but: "can you hear me".

How do local partners in hybrid meetings refer to remote participants? How do they address the remote participant? How do remote partners address local participants? In section 7 we look at the participation of the remote participant in remote meetings. Section 3 is about the way speakers express their stance towards the group and others by means of deictic expressions. Section 2 is about the way speakers express their own or the group's stance towards some target.

Speech acts differ not only in their illocutionary force or expressive force but also in the way they are directed to others, in the type of addressing, in the way they express that the speaker wants to convince or warn someone in particular, or ask or request something to someone in particular. This selective directness shows in postures, head movement and eye gaze and sometimes in the use of a specific style of addressing. In section **??** we will discuss ways in which partner in conversation address each other, focusing on the relation between turn-giving and addressing of eliciting acts.

1.1 Can we rely on annotated data?

In our analysis we use annotated data. Does this data allow us to infer conclusions that are valid for a larger set of events than the data we analysed? In other words: can we abstract from the specifics of the design meetings and the settings in which they have been recorded? And, can we abstract from the specific annotators that have annotated this data? Or: are annotators replaceable by other annotators, and deliver annotated data that lead us to similar conclusions about the content of the data?

Let us reflect a while on the latter question. It is a serious question in all research that relies on annotated data (see Krippendorff [2004] and Carletta et al. [1997a]). As an example we look at speech act annotation. In principle, the number of different speech acts is unbounded. Moreover a speech act is a multi-faceted act, the speaker can express a number of different intentions at the same time. But, every model of speech acts forces us to distinguish a small number of different categories and confronted with a concrete speech act it is often a tour de force to decide what is the category that best fits the intention of the speaker. However large or multi-dimensional the scheme is, the decision to call it an act of type *X* or *Y* will often be arbitrary. What is true for dialogue acts also holds for addressing (does the speaker address the group or some individual?) and for dialogue act segmentation (to these words express two different dialogue acts in a sequence or one?), and maybe for every form of classification into a fixed number of fabricated abstract classes.[2] What categories we distinguish depends on the aim of the annotation. In the AMI Dialogue Act Scheme for example no distinction is made between opinions and informs, so that for the

2. Note that the AMI Dialogue Act Annotation Procedure (AMI Consortium [2005]) confronts the annotator who has to segment the utterance "As John said this morning, correct me if I'm wrong (directing to John), we shouldn't take this too seriously." (directed to Mary), with a nasty problem, since the scheme doesn't allow discontinuous word sequences as dialogue act segments.

research in the ways people express their opinions a new annotation scheme had to be developed (see section 2).

If we evaluate a machine classifier trained on a set of annotated data, we have to be aware of the non-computational character of the classification that human annotators were forced to perform. Participants in a conversation can have different interpretations of the intention of the speaker, (some speech act have different functions for different addressees). This sometimes, but not always, reveals in the course of the conversation. What holds for participants, addressee and overhearers (Clark and Schaefer [1992]), also holds for outside observers, remote participants, and for human annotators as well. Shouldn't we expect that different observers see different things? But if that is the case how then do we assess the outcome of a machine classifier, or a machine that is trained to produce a summary of the decisions made in a meeting? What authority does it have in saying how to interpret a contribution of a speaker in a conversation?

These questions that fit in a reflection on the methodology of research that aims at developing technology based on computational models of human interactive behavior have been part of the work of WP1. The results of this research is reported in a journal papers Reidsma and Carletta [2008], a PhD thesis Reidsma [2008] and in workshop proceedings Reidsma et al. [2008], Reidsma and op den Akker [2008].

This report concludes with a section where we give some implications of our findings for the development of technology that aims at improving the participation and effectiveness of remote meetings. These could be taken as recommendations for the User Engagement and Floor Control demonstrator being developed in WP6. We end with a section in which we present our research plans for the final year of Amida.

2 Subjectivity

Opinions, emotions, agreements, disagreements, and other types internal mental and emotional states (*private states*) are a crucial part of communication. Within the context of meetings, being able to recognise and understand when such private states are being expressed is critical, both for the meeting participants seeking to have an effective meeting and for applications tasked with helping to facitilate the meeting or retrieve information from meeting archives.

When a meeting involves remote participation, communication and interaction naturally must adapt to changes in the communication medium. How the expression and recognition of private states changes with remote participation is an important question. For some types of subjective content, e.g., direct statements of opinions and sentiments, we might expect fairly little change. A direct expression of an opinion or sentiment will likey be the same whether the communication mode is face to face or over a remote channel. However for other private states that may be expressed less overtly, such as agreements agreements and disagreements, differences certainly may be expected as participants seek to overcome the challenges posed by remote communication.

In this section, we investigate how the expression of two particular types of private states, agreements and disagreements, may differ in the remote versus face-to-face setting. Previously, we developed an annotation scheme for marking various types of subjective content in multiparty conversation Wilson [2008]. Although this annotation scheme distinguished positive and negative subjective content, it did not specifically represent agreements or disagreements. Thus, we first modify the AMIDA subjectivity annotation scheme to better represent agreements and disagreements, and then use the revised scheme to mark agreements and disagreements in 12 AMIDA meetings, four of which involve participants. With these new annotations, we then consider what effect remote participation has on the expression of agreements and disagreements.

2.1 Representing Agreement and Disagreement

The AMIDA scheme for marking subjective content in multiparty conversation has three main categories of annotations, *subjective utterances*, *objective polar utterances*, and *subjective questions*. Each of these main categories are further divided into subcategories. Table 1 gives the categories used in the original version of the scheme.

In the original version of the AMIDA scheme, agreements and disagreements were included in the *positive subjective* and *negative subjective* categories. Specifically, the positive subjective category included agreements, positive sentiments (emotions, evaluations, and judgments), positive suggestions, arguing for something, beliefs from which positive sentiments can be inferred, and positive responses to subjective questions. The negative subjective category encompassed private states that are the opposite of those in the positive subjective category, specifically, disagreement, negative sentiments, negative suggestions, arguing against something, beliefs from which negative sentiments can be inferred, and negative responses to subjective questions.

For the revised version of the annotation scheme, we break out agreement and disagreement from the positive and negative subjective categories and create two new categories of

Subjective Utterances
positive subjective
negative subjective
positive and negative subjective
uncertainty
other subjective
subjective fragment
Objective Polar Utterances
positive objective
negative objective
Subjective Questions
positive subjective question
negative subjective question
general subjective question

Table 1: AMIDA Subjectivity Annotation Types

subjective utterances. The agreement category is used for marking utterances expressing agreement with a previous statement, idea, or opinion, including expressions of private states like concession, in which there is agreement even if it is reluctantly or grudgingly given. Disagreements are the opposite of agreements, and include utterances that express disagreement with a previous statement, idea, or opinion.

Verbal agreement and disagreement may be expressed in a number of different ways. For example, expressions of agreement are not limitted to just positive words and phrases. In the following example, participants D and B are agreeing with a negative sentiment expressed by A. Although they are both agreeing, D agrees by uttering "no" and B agrees by uttering a "yeah."

A: And ⟨NEGATIVE SUBJECTIVE; neither of them were very pretty⟩ , you know?
D: ⟨AGREEMENT; No⟩ .
B: ⟨AGREEMENT; Yeah⟩ .

In the next example, disagreement is expressed with a positive subjective utterance that is in contrast to the previous statement.

B: Okay, ⟨POSITIVE SUBJECTIVE; let's have our buttons all be one color⟩ .
D: ⟨DISAGREEMENT; Mm, ⟨POSITIVE SUBJECTIVE; I kind of like the buttons⟩ ⟩.

2.2 Agreement and Disagreement in Remote versus Face-to-Face Participation

3 Using Participant Deixis in Conversational NLP

This section is about personal pronouns and deictics in general (John Niekrasz - UEDIN)

The words I and you are the most frequently used lexical nominals in conversational English, from the Switchboard telephone conversations Godfrey et al. [1992] to the AMI multi-party meetings McCowan et al. [2005], and in the British National Corpus where they are in fact the two most common of any words in the "demographic" (i.e. conversational) subcorpus Burnard [2007]. Even in some textual sources, they are still by far the most frequent. Google's Web 1T 5-gram statistics Brants and Franz [2006] list I and you as more frequent even than the word it. The word we generally falls within the top 10 most frequent words as well. The frequent use of first- and second-person (egophoric) pronouns in conversation is an indication of a dominance of personal beliefs and attitudes in conversational arenas. But how, when, and why are speakers using these words? And why so frequently? What is the role of these words in conversation? And ultimately, within the interests of this workshop community, how can they be modelled computationally and used in conversational technologies?

We think person reference, particularly **participant deixis**, is an important, non-trivial, but neglected area of research. It is our hypothesis that studying it is necessary for continued progress in the computational modelling of human-human conversation. In particular, we expect that progress will have direct impact on three major problems in conversational NLP: (1) determining the roles played by individuals in conversational activities and actions, (2) understanding participants' beliefs, intentions, sentiments, and other attitudes, and (3) segmentation and dialogue structuring at multiple levels. In this paper, we justify our hypotheses by presenting qualitative empirical evidence for each. We finish by highlighting the central challenges and our ongoing work to solve them.

Problem 1: Participant roles in conversational activities and actions

Our idea to study person reference comes from previous research on information extraction in multi-party conversation Gruenstein et al. [2005], Purver et al. [2006a,b, 2007]. As part of a broader project to extract and identify important information in organizational meetings, Purver et al. identified action items – public commitments to perform a given task – as one of the by-products of project meetings which are most important to participants. What these studies revealed was the need for a solution to the problem of resolving references to people. Without such a solution, it would be impossible to summarise who was responsible and to whom. Consider the following dialogue from a project planning meeting:

(1) ISL–m063–u0680

01 SAQ: yeah. also, 'cause **you** said **you** were gonna send **me** an email about how to set up **our** travel

02 HHI: yeah, **I**'m gonna send- yeah, **I**'ll send **you** the email uhm uhm when **I** go back. send **you** the email. uhm and **you**'re gonna have to contact **him**, and **they** have a travel agency.

03 SAQ: okay.

In this short extract there are 11 occurrences of the personal pronouns you, me, I, him, and they and one occurrence of a possessive determiner our. Clearly, the ability to understand this exchange relies heavily on the resolution of these expressions in context.

Problem 2: Participant beliefs, sentiments, and other attitudes

The words I, you, and we are often used as discourse markers (e.g. I mean, you know), as the subjects of mental and communication verbs (e.g. I [think—guess—dunno], you [want—think—see]), and with modals (e.g. you [can—could], I [would—can], we [can—should—could]). These contexts of use reflect the attitudes of speakers toward the subject matter, their own thoughts, and the thoughts of others. Consider the following monologue within a project design meeting:

(2) AMI–ES2002b–434.06s

01 B: well **i think** it's a valid point
02 B: **i mean** like the one on the left looks quite uh quite complicated
03 B: and that p r t p r o t thing is incredibly confusing
04 B: uh so **i see i see** why **you know you might** prefer the simpler design
05 B: but yes **you don't want** to lose out on **you know** what it does
06 B: so maybe **you know**
07 B: **you know you get** a lot of remote controls where **you kind of** flip the thing open
08 B: **i think** that's a good idea

This example highlights the need for distinguishing discourse markers from other uses. It also highlights some other phrases which act as hedges and expressions of opinions.

Problem 3: Segmentation and structuring at multiple levels

Thinking of a conversation as an information-bearing object is very limiting. Neglecting to consider it as action within situational, social, and interactional contexts impairs the choosing of problems and the success of applications. For example, topic-based segmentation of spontaneous speech is notoriously difficult for even humans to do. But annotation schemes which model discourse structure in terms of the joint activity have far better agreement Carletta et al. [1997b]. Participant deixis depends greatly on the joint activity (e.g. sharing opinions, providing information, and narrating stories). As such, considering participant deixis in models of cohesion and segmentation is likely to produce benefits. At the same time, analysis of finer-grained structure like adjacency pairs will also benefit from consideration of the contextual interpretation of participant deixis, exemplified by the following dialogue:

(3) SWBD–2023–429.21s

01 B: oh **you**r up in Memphis
02 A: no **i** used to be **i**'m in Texas
03 B: oh **you**r in Texas oh okay **i** was **i** was going to go goodness they really got uh this out far
04 A: **i** don't know **i** don't know how far it goes
05 A: are **you** in Texas
06 B: yeah yeah **i**'m down in Houston

Main challenges: ongoing and future work

Our work is currently in the annotation phase. Our first achievement has been the development of a three-step annotation method as follows: (1) all participant references are resolved (pronouns are classified for their referentiality, and non-pronominal and elliptical references are identified), (2) the commanding syntactic unit is identified (commanding

clause, phrase, or particle is marked and verbs and modals are marked where appropriate), and (3) pragmatic functions of the units are marked (hedging, subjectivity, sentiment, epistemic attitudes). By the time of the workshop, we expect to have completed our annotations and a quantitative corpus analysis.

We also expect to have preliminary results from machine learning experiments at the time of the workshop. Our first use of these data will be experiments using CRFs to do disambiguation and classification of each of the pronouns. Building on Gupta et al. [2007], our main aim will be the integration of more meaningful features which make use of the syntactic and attitudinal context provided by our annotations. We will also be experimenting with unsupervised lexical chaining using resolved egophoric pronouns rather than those in their original form.

Solving these problems will be difficult, despite some apparent simplicities. Gupta et al. Gupta et al. [2007] show that the word you takes on a generic meaning about half of the time in telephone and meeting conversations. And when referential, words like you, we, us, and they have a multitude of (often vague or ambiguous) interpretations. Additionally, resolution of referential you requires addressee detection and a model of turn-taking.

Despite the frequency, complexity, and importance of egophoric pronouns in conversation, computational linguists have almost completely neglected these issues. This neglect is largely due to two factors. First, research has focused on non-conversational arenas of language use. Second, previous computational treatments of this problem have relied on naive rule-based methods which ignore the complexities outlines above. In other words, these treatments produce poor results in human-human domains because they are designed for human-computer dialogue systems and based on oversimplified assumptions. But this is not to say that the subject has been neglected in non-computational research areas. A wealth of literature exists, primarily ethnographic studies in conversation analysis. The principal challenge going forward will therefore be the transformation of this knowledge into useful computational models.

4 Eliciting Acts in Meeting Conversations

Most of the contributions in a conversation can only be properly understood in the discourse context. There are many aspects of context that can be relevant for this understanding. An important example of discourse dependency is that between questions and answers. In answering a question the speaker often refers to words or phrases, that occur in the question. Short answers, like "Yeah, that's okay", or "Ten" make a strong case. Here we look at elicit acts, encompassing questions and requests. In particular we look at how they are addressed by the speaker, and how knowledge about who is being addressed can help in finding the match between elicit act and the responding act.[3]

The data that we used consists of a part of the AMI corpus of face-to-face scenario based design project meetings, we refer to this part here as the AMI corpus, and a part of the AMIDA corpus. The AMI corpus consists of 14 meetings. They are all manually annotated with dialogue acts, focus of attention, addressing, and adjacency pairs (Schegloff and Sacks [1973]. This corpus was also used by Jovanovic Jovanovic [2007] in her addressing research. The Amida part consists of 3 meetings that were annotated on the same layers except for the focus of attention layer.[4]

Elicit Acts are typical a-parts of adjacency pairs, with responses as b-part. The AMI corpus counts 243 adjacency pairs with an I-addressed elicit act as a-part and a b-part of which the speaker is the addressee of the elicit act (thus, they have speaker-addressee pattern ABBx where the x stands for either A, i.e. the addressee of the b-part is the same as the speaker of the a-part, C, some third partner is the addressee of the b-part or G, the group is addressed by the b-part of the responding act [5].

A rule in a model of turn-taking in face-to-face conversations, says that the speaker selects the next speaker by addressing(see Sacks et al. [1974] and for a critical analyses of the status of these "rules" in the SSJ-model of turn-taking see O'Connell et al. [1990]). Based on this rule, a naive method to find the answer to a question would be: take the first speaker turn by the one who was addressed in the elicit act that "follows" the elicit act (where "follows" means: the start of the response act is after the start of the elicit act). Of the 243 adjacency pairs in the AMI corpus this method gives 196 goods answers, 47 are bad, a performance of 81% correct. See the top row in the left (AMI) part of Table 2. Of the 82 adjacency pairs in the Amida corpus, this method finds the matching b-part of the elicit

3. Questions and request are examples of a type of speech acts that Searle called "directives" Searle [1969]. They have the intention to direct the listener's behavior. Asking someone a question, implies requesting the addressee to respond to that question. Although it is also the case that when a speaker gives information while he is directing himself to a specific listener it has often the effect of eliciting feedback, the "directives" are speech acts that more strongly request for a response than the other types of speech acts, so that it is clearly marked behavior when the addressee does not respond. The fact that elicit acts are more directed also shows in speaker's gaze behavior and in speaker's addressing behavior. We see that elicit acts are more prominently addressed than most other types of speech acts. Human annotators show significantly more agreement when asked about the type of addressing of elicit acts than when asked about the type of addressing of other dialogue acts.

4. Note that for addressing annotation of the remote meeting corpus collected in the Amida project a less elaborate procedure was followed than in the AMI project. The AMIDA procedure asked annotators only to label those segments of word sequences of the speech transcription, that were individually addressed, where in the AMI corpus addressee labels were attributed to (proper) dialogue act segments, where all dialogue act segments together cover the whole word layer. See the Amida addressee annotation manual.

5. The notation is adopted from Gibson [2003].

AMI

pattern	GOOD	WRONG	Total
ABB	196	47(25s)	243
ABC	0	26	26
ABA	0	8	8
Total	196	81	277

AMIDA

pattern	GOOD	WRONG	Total
ABB	57	25(11s)	82
ABC	0	18	18
ABA	0	2	2
Total	57	45	102

Table 2: Tables with performance of the naive method

a-part in 57 cases. In 25 cases the result is wrong. A performance of 70% correct. See the top row in the right (AMIDA) part of Table 2.

It happens that more than one speaker responds to an elicit act in more than one act. In general: acts can participate as an a-part in more than one adjacency pair. In 25 of the 47 wrong predictions the elicit act was a-part of only one relation. Interestingly, of these there are 9 APs in the AMI corpus with a b-part that has a start time before the start time of the a-part. In these cases our simple method -which, for obvious reasons, expects the response after the elicit act- always gives a wrong answer. If we ignore these cases our method still only finds the good answer in 82% of the cases. How can we improve this? What is the position of these answers in the course of the interaction? do they always follow the elicit act? Do they partly overlap with the elicit act? Are their other speakers active in between the a-part and the b-part? Are all elicit acts actually responded by the addressee or do others answer the question? What indices of the context of the act, and what aspects of the elicit act, are relevant or made relevant by the act for finding answers to these questions?

There are some preliminary questions to be answered as well. How do we find elicit acts and how do we know that the act is I-addressed and to whom, so that we know what speaker we have to attend to find the answer to the question. And, how reliable is the annotated material? Can we draw any conclusion that go beyond the personal interpretation of the annotators? We found that I-addressed elicit acts have higher agreement between annotators than other acts. In section 4.2 we present results from inter-annotator agreement in this specific context.

[6] Hence, there is not a fixed context in which utterances should be interpreted, instead the speaker make specific aspects of context relevant. There is an interaction between

6. A quote from Stanton Wortham -

> All approaches to discourse share a commitment to studying language in context. But "context" is notoriously indeterminate, and different approaches to discourse analysis emphasize different aspects of context as potentially relevant to understanding language use. The basic question for contemporary discourse analysts is: how do participants and analysts know which of an indefinite number of potentially relevant aspects of the context are relevant to understanding a given utterance?

(`http://www.gse.upenn.edu/~stantonw/research/discourse-analysis.htm.`)
Drawing particularly on work in linguistic anthropology, Wortham argues that an adequate answer to this question must rely on the central concepts of mediation and emergence. Instead of following rules to make their utterances both grammatical and appropriate to particular cultural contexts, speakers deploy indexical cues that could have multiple meanings and hearers must infer which of several possible aspects of the context are in fact relevant to interpret the utterance. Wortham [1996]

Figure 1: The Dialogue Act Annotation (top-frame) and Relation Pair Annotation Viewer (bottom frame). The Suggest act participates in four adjacency pairs.

contributions and what should count as context.

4.1 Annotation layers relevant for this research

The annotation layers relevant for this research are:

- Dialogue act segmentation and dialogue act labeling.
- Adjacency pair annotation: for example question answer relations.
- Addressing: who the speaker is talking to.
-

These layers are described in the AMI Guidelines for Dialogue Act Annotation, AMI Consortium [2005]. Three classes of acts are distinguished in the AMI dialogue act annotation scheme, with four types of eliciting acts. (For an abstract of the annotation scheme with examples for annotators see Appendix 9.3.)

1. Acts about information exchange:
   ELICIT-INFORM
2. Acts about possible actions:
   ELICIT-OFFER-OR-SUGGESTION
3. Acts that comment on the previous discussion:
   ELICIT-COMMENT-ABOUT-UNDERSTANDING and
   ELICIT-ASSESSMENT.

Table 3: Krippendorff alpha values (and numbers of agreed DA segments) for the three pairs of annotators; for addressing, addressing of elicit acts, dialogue acts (all 15 DA classes), and elicit vs non-elicit acts.

pair	adr	adr-eli	da	da-eli
a-b	0.50(412)	0.67(31)	0.62(756)	0.69
a-c	0.37(344)	0.58(32)	0.58(735)	0.64
b-c	0.33(430)	0.62(53)	0.55(795)	0.80

Figure 1 shows the Dialogue Act and Relation Annotation Viewer that we developed and used for visualizing the annotations. The viewer allows to load multiple annotations from different annotators of the same meeting so we can easily see where annotators differ in their labelings. Notice that in the situation (from Amida meeting ED1002b) depicted in the Figure, all four responses to the Suggest act, which was produced by participant *A* are short feedbacks, all labeled as Assessment, and they all fully overlap in time with the Suggest act.

4.2 Reliability Analyses

Focus of attention annotation was done with high agreement, so we may conclude that the annotated data allows a good starting point for research of multi-modal conversational behaviour involved in addressing of eliciting acts and the responsive behaviour that follows in multi-party face to face conversations in general.

One AMI meeting (IS1003d) was annotated by three annotators. Table 3 shows for each pair out of these annotators Krippendorff alpha values for inter-annotator agreement Krippendorff [2004]. For the group of annotators alpha is 0.35 for addressing. The statistics are based on comparing DA-labels of completely agreed DA-segments. Most confusions in the addressing labeling are between I-addressed and G-addressed, between I and U and between G and U; there is hardly any confusion between annotators about who is addressed when they agree that the DA is I-addressed (see also Jovanovic [2007]). The table shows that annotators agree more on the addressing of elicit acts than on DAs in general. For the subset of elicit acts we see hardly any U labels used, and when annotators agree that an elicit is I-addressed (which happens in 50-80% of the agreed elicit acts), they agree on who is addressed, without exception. Annotators agree more on the addressee of a DA in situations where the speaker clearly gazes at the addressed person.

Are there types of dialogue acts that are more difficult than others to label with an addressee type?

Total number of proper dialogue acts is 454. Krippendorff's α for all eleven annotators is 0.62 (leave one out range form 0.60 to 0.64).

Krippendorff's alpha for multiple annotators on the set of 50 elicit acts is 0.80 (leave one out values range form 0.77 to 0.80).

Table 4 shows for each dialogue act type how many instances 0,1,2,...11 annotators the instance labeled as addressed to the group. Thus -for example- of the total of 4 instances of dialogue acts of the Elicit-Offer-Or-Suggestion type

Dialogue Act Type	Counts											
	0	1	2	3	4	5	6	7	8	9	10	11
Assess	13	11	13	12	10	6	7	2	3	1	4	11
Comm-About-Under	13	9	3	2	1	1	0	1	0	0	0	0
Elicit-Inform	17	4	2	0	1	0	0	0	2	0	0	2
Elicit-Offer-Sug	0	1	0	0	0	0	1	1	0	0	1	0
Inform	17	15	22	15	19	15	7	12	9	6	17	37
Elicit-Comm-Under	0	2	0	0	0	0	0	0	0	1	0	5
Offer	0	0	0	2	0	0	1	0	0	0	1	7
Elicit-Assess	5	1	2	0	0	0	0	0	1	0	1	0
Be-Positive	0	2	1	1	0	0	0	1	2	1	0	0
Suggest	15	13	8	3	3	2	3	4	3	3	5	9

Table 4: Counts of judgements of eleven annotators in telling if a dialogue act was addressed to the group or not. Rows show the counts for the dialogue acts per type.

- one time 1 annotator labeled is a G-addressed (and all other 10 annotators agreed that is was I-addressed)
- one time 6 annotators labeled it as G-addressed
- one time 7 annotators labeled it as G-addressed
- one time 10 annotators agreed and labeled it as G-addressed (and 1 labeled is as I-addressed)

Thus the higher the numbers in the middle sections of a row the more confusion about the addressee type of the associated type of dialogue acts.

These counts suggest that the Inform and Assess types of dialogue acts are more difficult for annotators to tell if they are G-addressed or I-addressed than the dialogue acts of types Elicits and Suggest.

Can we say that these types are the typical "initiating" types of dialogue acts? Can we say that speakers performing initiating acts often more clearly show whether they address an particular individual or the group?

Table 5 shows results from an agreement analysis for dialogue act and addressing of AMI meeting IS1003d. It shows that annotators agree more on the addressing of elicit acts than on DAs in general. For the subset of elicit acts we see hardly any U labels used (for uncertainty about addressee label), and when annotators agree that an elicit is I-addressed (which happens in 50-80% of the agreed elicit acts), they agree on who is addressed, without exception. Annotators agree more on the addressee of a DA in situations where the speaker clearly gazes at the addressed person. We did not find any indication that annotators systematically confused speaker's gaze with addressing. Addressing is a complex phenomenon and we believe that the low agreement between addressee annotations is due to this complexity.

4.3 Addressing and Eliciting Acts

Speakers have to make clear whom they address and therefore they select a way that they think will be succesfull in communicating this part of there intention. Addressing is a joint

pair	adr	adr-eli	da	da-eli
a-b	0.50(412)	0.67(31)	0.62(756)	0.69
a-c	0.37(344)	0.58(32)	0.58(735)	0.64
b-c	0.33(430)	0.62(53)	0.55(795)	0.80

Table 5: Krippendorff alpha values (and numbers of agreed DA segments) for three pairs of annotators; for addressing, addressing of elicit acts, dialogue acts (all 15 DA classes), and elicit vs non-elicit acts.

activity (in the sense of Clark) and needs to be grounded to an extend that is required to go on with the conversation. This process thus can be described in the grounding framework. op den Akker and Theune [2008].

Lerner Lerner [2003], Lerner [1996] distinguishes explicit methods of addressing, which are speakers' gaze and naming (the use of vocatives, address terms), from "tacit forms of addressing that call on the innumerable context-specific particulars of circumstance, content, and composition to select a next speaker" (Lerner [2003], p.177). Lerner examines the context-sensitivity of addressing practices employed by a current speaker to make evident the selection of a next speaker. His discussions are restricted to those turns-at-talk that implement sequence-initiating actions, the first parts of adjacency pairs.

Addressing by gaze works only if the addressee notices the speaker's gaze and picks up the signal as a sign of addressing; moreover both have to believe that they share this common belief. Mutual gaze between speaker and addressee is basic for grounding in face-to-face conversations. Only mutual gaze between A and B is the most reliable way to establish the belief of A a) that A sees B, b) that A sees that B sees that A sees B, and c) that both share this belief. Accompanied with other messages sent by A (an utterance of a question for example, or a gesture) this may lead B to believe that A's gazing at her means that B is being addressed by A. By looking at B, A checks whether B is ready to receive his message. Others also have to understand that they are not selected as next speaker. Thus, "gaze is an explicit form of addressing, but its success is contingent on the separate gazing practices of co-participants" (Lerner [2003], p.180).

According to Lerner there is one form of address that always has the property of indicating addressing, but that does not itself uniquely specify who is being addressed: the recipient reference term 'you':

> The use of 'you' as a form of person reference separates the action of addressing a recipient from the designation of just who is being addressed. In interactional terms, then, 'you' might be termed a recipient indicator, but not a recipient designator. As such, it might be thought of as an incomplete form of address. Lerner [2003], p.182.

The speaker will try to complete the addressing act by gazing at the selected recipient, a completion that needs the joint gazing of the intended recipient, and of others present as well, so that they know they are not selected. Thus, for addressing to be complete it requires the joint actions of all participants. This is illustrated by the following fragment from the AMI meeting corpus Carletta [2007]. In the first utterance by speaker P3 'you" is not supported by disambiguating gaze; both conversation participants P2 and P0 are gazed at by P3. P2 feels addressed and responds, but P0 also. P2's response overlaps with

the elicitation and he is interrupted by P0. It is as if P2 then recognizes that not he but P0 was selected as next speaker. P2 and P0's "Uh" may also indicate the confusion in the situation.

P3>P0: What do you think, is it fancy?
P2>P3: Uh, it's really
P0>P3: Uh, I think that fancy, we can say it is fancy.

Lerner [2003] discusses an example of use of referring 'you' directed to a specific individual in a multi-party conversation, where the addressing does not need the support of speaker's gaze at the intended addressee. In a situation where four people are having dinner together, and everybody knows who has prepared the dinner, and the speaker assumes that everybody knows that, the speaker asks: "Did you cook this all the way through?" (Lerner [2003], p.192). Here, the content and context are sufficient to determine the identity of the addressee without the need for explicit addressing behaviour. Schegloff mentions usages of "you" by which the speaker refers to himself, or himself and his listeners Schegloff [1996].

The most explicit form of addressing is by use of an address term (which may or may not take the form of a name). This is either used in pre-position, in post-position, or in mid-position, as illustrated by the following examples.

So, mister money, what's your opinion according to this remote control?
What do you think, Ed?
They wake up fast, Jessie, if they have to.

In almost all usages of address terms in talk in face-to-face conversations their function is not purely to call the addressee's attention. If it is, the term is used in pre-position, more often than elsewhere, but most often it seems to be used to put more stress on the addressing, maybe to signal the addressing to co-participants, or to express some affective or social relation with the recipient.

Some activities center around one specific actor; a presenter, or someone drawing on the white board, or someone holding the clay prototype that is being discussed. If someone says "is it heavy?" it is clear who is being addressed. Or, when a person is drawing his favorite animal on the white board and the speaker makes a guess "a horse?", asking the artist to reveal his secret animal. Actors of activities that are in focus are more salient than others for addressing. Moreover, these actors can tacitly be addressed by others when they comment on, or ask about, the action they perform.

Sometimes, the speaker uses the wrong name or a wrong attribute for his addressee. In such cases an unaddressed listener might feel more entitled to answer the question than the addressee. The speaker uses the referent term 'you' and gazes at P2 to make clear whose identity he is after. But the real marketing guy is called by the attributive use of "marketing guy".

P3>P2: You are the marketing guy ? Or
P0>P3: I'm marketing .

In the following fragment it is unclear who is being addressed: a non-addressed attendant tries to answer but is interrupted by the addressee. The speaker indeed gazes at P1 at the end of his question which could easily be taken as if he has selected P1 to speak next.

P2>P0: so how many units should we sell to have a
P1: Well . Uh
P0>Group: Well each unit is is sell uh twenty five Euros .

Another example of unclear addressing: a you-utterance without speaker gaze to select the designated addressee

P0>P3: D D Is is there anything you want to add ?
P2>Group: Is there any fruit that is spongy ?

We have seen a number of cases that make clear that proper addressing uses beliefs that speakers have about saliency of persons because of their role in the activity that the group is busy with. Successful addressing is constrained by the general conditions about sharing beliefs about who are salient and who are gazed at as a signal of addressing.

Our analysis has shown that speaker's gaze at addressee differs for elicit acts and other I-addressed acts: speakers gaze more at addressee during elicit act than with other I-addressed acts. This could explain also outside observers, like annotators, have less problems to identify who the speaker is addressing.

4.4 A Naive Method for Finding Responses to Elicit Acts

Table 6 shows statistics for elicit acts in the corpus of 14 AMI meetings. Table 7 contains statistics of the AMIDA meetings. The columns in these two tables contain the following data:

1 **name**. The name of the meeting (observation) and an identifier for the annotator. The annotator did dialogue act segmentation, labeling, addressee labeling, as well as relation (adjacency pairs) annotation.
2 **acts**. The number of dialogue acts segments including those labeled as Stall, Fragment, Backchannel, and Other (these have not been annotated with an addressee label.
3 **r-acts**. The number of proper dialogue acts.
4 **I-racts(pc)**. The number of I-addressed proper dialogue acts and the percentage of this against the total number of proper dialogue acts (column r-acts).
5 **el**. The number of Elicit acts.
6 **I-el(pc)**. The number of I-addressed Elicit Acts and the percentage of this against the total number of I-addressed Elicit Acts (column el).
7 **aps**. The number of relations (adjacency pairs).
8 **el-aps(pc)**. The number of relations of which the a-part (target) is an elicit acts, and the percentage of this against the total number of adjacency pairs (column aps).
9 **I-el-aps(pc)**. The number of relations of which the a-part (target) is an I-addressed elicit acts, and the percentage of this against the total number of adjacency pairs with an elicit act as a-part (column el-aps).
1. **abbx**. The number of relations of which the a-part (target) is an I-addressed elicit acts and of which speaker of the b-part (source) is the same as the addressee of the a-part, the elicit act. The pattern form *abbx* covers the speaker-addressee patterns *ABBA*, *ABBC*, and *ABBG*.

name	acts	r-acts	I-racts(pc)	el	I-el(pc)	aps	el-aps(pc)	I-el-aps(pc)	abbx
ES2008a-vk	386	273	67(24,54)	28	15(53,57)	31	15(48,39)	11(73,33)	11
IS1000a-vk	658	417	245(58,75)	43	24(55,81)	108	22(20,37)	13(59,09)	12
IS1001a-s9	323	199	112(56,28)	33	24(72,73)	44	15(34,09)	11(73,33)	11
IS1001b-dh	897	568	182(32,04)	68	47(69,12)	196	50(25,51)	38(76)	37
IS1001c-dh	565	370	104(28,11)	47	27(57,45)	127	37(29,13)	25(67,57)	23
IS1003b-vk	693	454	270(59,47)	50	35(70)	147	47(31,97)	35(74,47)	26
IS1003d-vk	1563	838	394(47,02)	67	41(61,19)	247	60(24,29)	30(50)	23
IS1006b-vk	953	643	261(40,59)	55	31(56,36)	152	38(25)	28(73,68)	25
IS1006d-s9	1232	785	393(50,06)	104	52(50)	226	55(24,34)	24(43,64)	20
IS1008a-s9	263	192	117(60,94)	30	25(83,33)	43	25(58,14)	20(80)	20
IS1008b-vk	640	467	182(38,97)	53	30(56,6)	87	20(22,99)	15(75)	12
IS1008c-s9	584	351	129(36,75)	24	17(70,83)	69	18(26,09)	12(66,67)	11
IS1008d-s9	589	401	145(36,16)	23	13(56,52)	67	11(16,42)	10(90,91)	9
TS3005a-vk	641	389	142(36,5)	27	6(22,22)	107	34(31,78)	5(14,71)	3
Totals	9987	6347	2743(43,22)	652	387(59,36)	1651	447(27,07)	277(61,97)	243

Table 6: Statistics of Elicit Acts in 14 AMI meetings

name	acts	r-acts	I-racts(pc)	el	I-el(pc)	aps	el-aps(pc)	I-el-aps(pc)	abbx
ED1002b-vk	1278	806	330(40,94)	81	42(51,85)	262	82(31,3)	36(43,9)	31
ED1005b-vk	698	498	139(27,91)	54	17(31,48)	84	38(45,24)	17(44,74)	17
ED1005c-vk	969	669	318(47,53)	94	51(54,26)	237	106(44,73)	49(46,23)	34
Totals	2945	1973	787(39,89)	229	110(48,03)	583	226(38,77)	102(45,13)	82

Table 7: Statistics of Elicit Acts in the 3 AMIDA meetings

name	I-eli-acts	hasBpart	NextEqNextByAdr		NextByAdrEqBpart	
ES2008a-vk	15	10	12	(80)	9	(90)
IS1000a-vk	24	12	11	(45,83)	8	(66,67)
IS1001a-s9	24	11	18	(75)	11	(100)
IS1001b-dh	47	35	36	(76,6)	33	(94,29)
IS1001c-dh	27	19	17	(62,96)	18	(94,74)
IS1003b-vk	35	26	23	(65,71)	21	(80,77)
IS1003d-vk	41	21	27	(65,85)	14	(66,67)
IS1006b-vk	31	23	20	(64,52)	17	(73,91)
IS1006d-s9	52	22	21	(40,38)	14	(63,64)
IS1008a-s9	25	19	21	(84)	18	(94,74)
IS1008b-vk	30	11	9	(30)	8	(72,73)
IS1008c-s9	17	11	12	(70,59)	10	(90,91)
IS1008d-s9	13	7	9	(69,23)	6	(85,71)
TS3005a-vk	6	4	4	(66,67)	3	(75)
Totals	387	231	240	(62,02)	190	(82,25)

Table 8: Results on the AMI corpus, using the naive method (take the first act performed by the person addressed by the elicitor that follows the elicit act as response act to the elicit act)

Of the 277 aps with I-addressed Elicit acts as a-part in 141 cases the annotated b-part is the next dialogue act.

Based on the rule that the speaker selects the next speaker when he addresses a elicit act to an individual and that the addressee will take up the turn and respond to the elicit act, we implement a naive method for finding the response to a question: we simple take the next proper dialogue act as the response. We analysed the results of two variants: in the first variant we return the first proper dialogue act that follows the elicit act (variant N). In the second variant we return the first proper dialogue act performed by the addressee of the the elicit act and that follows the elicit act (variant NbyA). Tables 8 and 9 show results of both variants of this naive method evaluated on our corpora. The columns in these tables contain the following data:

1 **name**. The name of the meeting (observation) and an identifier for the annotator.
2 **I-eli-acts**. The number of I-addressed Elicit Acts
3 **hasBpart**. The number of I-addressed Elicit Acts that are a-part of at least one adjacency pair.
4 **NextEqNextByAdr**. The number of times that the variant N returns the same act as the method NbyA and the percentage of this against the total number of I-addressed Elicit Acts Elicit Acts (column I-eli-acts).
5 **NextByAdrEqBpart**. The number of times that the variant NbyA returns the b-part of the elicit act and the percentage of this against the total number of I-addressed Elicit Acts that occur as a b-part of an adjacency pair (column hasB-part).

Since not all elicit acts are annotated as a-part of an adjacency pair, we looked at all the elicit acts to see how often the method returns the correct act. There are in total 156 I-

name	I-eli-acts	hasBpart	NextEqNextByAdr	NextByAdrEqBpart
ED1002b-vk	42	27	31 (73,81)	21 (77,78)
ED1005b-vk	17	14	14 (82,35)	11 (78,57)
ED1005c-vk	51	31	37 (72,55)	22 (70,97)
Totals	110	72	82 (74,55)	54 (75)

Table 9: Results on the Amida corpus, using the naive method (take the first act performed by the person addressed by the elicitor that follows the elicit act as response act to the elicit act)

addressed Elicit acts that do not have a b-part defined, so that we could not compare the outcome of our method with the hand-annotated relations. Of these 156 in 46 cases the next speaker is the addressed person.

An example of an I-addressed Elicit acts that have no AP annotated. (In meeting IS1006d there are 8 of them).

```
1 D: (ELI): And when does it turn off?
2 B: (INF): When you don't touch the control
3 B: (INF): but you go out of the
4 D: (CAU): Oh
5 D: (ELI): SO YOU HAVE A
6 B: (INF): For for enough time
7 B: (FRG): like uh you
8 D: (ELI): sensing sensor machine that uh knows
9 B: (INF): It's a question to our technical design,
     our two engineers.
```

In this example the elicit act is interrupted.

In the following example the speaker doesn't pass the floor to the addressee after the elicit act, but continues with an other question (which is indeed answered by the addressee, and which has an annotated b-part):

```
1 D: (ASS): Uh I'm not sure about the screen ,
2 D: (ELI): wha what is the use usefulness of the screen ?
3 D: (ELI): Uh is it a touch screen by the way ?
4 B: (STL): So
5 B: (INF): I think it can be just a menu which can be
     controlled with a left , right , up , down and enter
     .
```

The final example shows an elicit act which is more rhetorical, at least it shouldn't be taken as a serious request for information.

```
1 D: (STL): Well
2 D: (INF): I I heard of devices where you just uh
     whistle them
3 D: (FRG): and and they
4 D: (INF): because of the the frequency they they
     just answer to that .
```

```
 5 D: (CAU): You can't whistle .
 6 D: (SUG): Or a clap .
 7 A: (FRG): And uh
 8 A: (INF): And it's answered .
 9 C: (BCK): Yeah .
10 C: (BCK): Yeah .
11 B: (INF): I can't whistle .
12 B: (INF): No , no , I can't .
13 B: (BCK): Mm .
14 D: (ELI): YOU CAN CLAP. CAN YOU ?
15 D: (ASS): Clap is good .
16 D: (FGR): Tak
17 A: (ASS): Clap clap clap it's a good
18 A: (INF): I I think it's universal .
19 C: (BCK): Yeah .
20 C: (BCK): Yeah .
21 D: (INF): [laugh] Just a [laugh] suggestion .
```

Improving the results of our naive method requires identification of elicit acts that are not really meant as a question, or that are followed by a question that overrules an earlier question, or that elaborate an earlier question before the elaborated question was answered. It would certainly be worth further exploration to see how follow-up questions can be distinguished from initial questions that initiate a new issue. Length of utterances seems like a good cue for distinguishing initial questions from follow-up questions.

5 Analysis of Corpus-based Remote Design Meetings

5.1 Introduction

This report presents an analysis of remote design meetings within the AMIDA corpus-based research. It specifically focuses on the phase-2 of AMIDA project's scenario-based meetings, where remotely located User-Interface Designer assists Industrial Designer - who is located in the meeting room. The report provides an analysis of 13 design sessions - all approximately 40 minutes each. The focus of this report is on understanding different patterns of non-verbal communications of participants. The analysis shows several important non-verbal aspects that should be taken into account when we design technologies that support remote collaboration between different participants.

In the following we will briefly describe the recording set-up used in AMIDA project. Next, an analysis of these remote design meetings is discussed.

5.2 Remote Meetings: The AMIDA Setting

(a) (b)

Figure 2: The AMIDA design meeting recording setting

The AMIDA project has developed a large corpus of video recorded meetings. Figure 2 shows the setup of the remote meeting scenario which was hosted by University of Edinburgh. In this setting we had 4 participants working on a design project. Their aim was to design a prototype of a remote control using provided materials. Through several cameras and sensors participants' interactions are recorded, specifically their behaviors and conversations. Amongst the participants there was a Project Manager (PM), a User-Interface Designer (UID), an Industrial Designer (ID) and a Marketing Expert (ME). The design project was divided into three planned phases.

1. Kickoff and introduction of the project
2. Concept design
3. Detailed design and budgeting

In the first phase, all the 4 participants were in the same room and they had to go through the earlier work done on this project and discuss pros and cons of different way forward. In this meeting they also allocated the tasks and role to each other. In the second phase, remotely located UID collaborates with ID for developing a remote control prototype. In the third meeting, ID reports the work (done with UID) to the PM and ME and discusses the budgeting issues. In the later part of the third meeting, the UID remotely joins in for the final discussion of the project. As it can be seen in Figure 2, participants used several digital and physical objects and tools to complete their tasks. The participants were equipped with microphones and 4 different high-resolution cameras capturing different views for supporting efficient communication.

This report focuses specifically on the phase-2 of the design scenario - where two designers ID and UID work together in order to develop a prototype remote control. Figure 2 (b) shows an example of this part.

5.3 Analysis

Our analysis shows different mechanisms used by both the remote and local participants for establishing communication and common-ground amongst themselves. We will specifically focus on the non-verbal aspects of communication utilized by the participants through common visibility of projected actions, gestures and use of design prototype and other relevant materials.

We will also describe the kind of activities that were supported by these non-verbal communication patterns. We will show how these communication patterns allowed adjustment in the design of the remote control prototype, its functionalities, its interaction mechanisms and future uses and so on. Our analysis also sheds light on how the information related to work-in-progress of design artefact is communicated by these participants.

5.3.1 Projecting Actions

(a) (b)

Figure 3: Projecting artefacts and actions on the camera

We observed that projecting actions towards the camera allowed participants to coordinate their design activity. By projecting actions, gestures and artefacts (e.g. drawings, prototype remote control) on the camera the information is intentionally made commonly visible which in turn supports shared understanding amongst the two remote participants. As it can be seen in Figure 3, several physical actions were projected so that the intended participant can see these actions and their meanings. Common examples were showing objects towards the camera view, pointing, and raising hands. Some of the actions were more frequent than others, depending on the stage of the design process. When UID and

ID where discussing the prototype we saw a higher level of dependence on these projected actions to coordinate work.

Public visibility as a coordinative aspect has been echoed by many others (Heath and Luff [1992], Robertson [1997]. Especially, Robertson suggests that the public availability of different artifacts and embodied actions to the perceptions of distributed participants in a cooperative process could enable their communicative functions. It has also been argued that more flexible or mobile access to the publicly visible information could improve coordination.

Recognizing When and Where to project

Figure 4: ID (Left) adjusts camera to zoom in on the prototype.

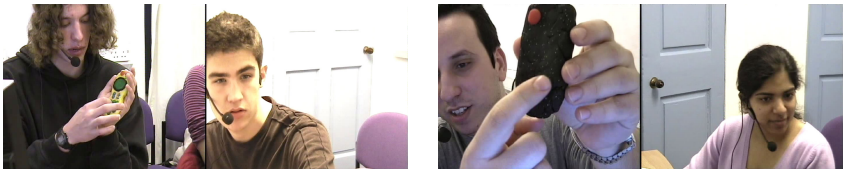
Adjusting Camera Both ID and UID were able to adjust the focus of their own cameras as they were able to see their own view in addition to each other's views. As shown in Figure 4, the ID zooms onto the prototype to see the details of the design. These activities occur when either it was requested by UID or when they both finish an aspect of their specific phase of design activity. It was also seen that sometimes ID forgot to adjust the focus of his camera, which did not provide sufficient information to UID.

Figure 5: ID (left) adjusts the position of his prototype to make it visible for UID.

Adjusting the position of artefacts It was seen very frequently in the meeting recordings that ID, involved in the design activity, comes close to the camera to show the design prototype. This activity is not that intuitive as ID has to look at a screen to be able to adjust the position of his remote control prototype.

5.3.2 Use of Gestures

Head movements and expressions Both participants used head movements and facial expressions to convey agreement and confirmation. This was also a quick way to say 'yes' or 'no' to the other participant.



(a) (b)

Figure 6: ID (left) points to a part of the prototype to communicate with UID.

Pointing to a specific part ID used pointing gestures to be able to describe position, shape and size of the buttons used in the prototype remote control and other related issues. As it can be seen in Figure 6, in order to discuss some detailed information about the prototype, it was necessary for ID to point to a specific portion of the prototype to discuss several relevant design decisions.

(a) (b)

Figure 7: UID (right) describes specific shapes using gestures to communicate with ID (right).

Describing a specific shape As the UID did not have a direct access to the prototype remote control, it was frequently observed that UID used gestures to communicate different shapes and describe some interactional mechanisms of the remote control.

Figure 8: ID (left) uses animated gestures to explain a design mechanism.

Animated gestures Some of the aspects related to the prototype remote control were not easily describable in words or through the prototype only. Participants needed to use animated gesture to be able to clearly explain their ideas. In one example, Figure 8 shows a still image when ID was describing a 'flip' mechanism that was not really possible to incorporate in the prototype. In a different example, Figure 9, ID uses a different animated gesture to describe a scrolling mechanism of a banana shaped remote control prototype. In both these examples ID uses animated gestures in order to describe different interaction mechanisms.

Figure 9: ID (left) explains scrolling mechanism using animated gestures.

Instinctive gestures We also observed that sometimes participants used instinctive gestures to describe their ideas. These gestures are quick reflections of participants while they are discussing the design activities. These gestures are sometimes not visible by the other participants as they are instinctive reactions and participants may not be able to intentionally show these on camera.

5.3.3 Use of Artefacts

We observed that different artefacts were useful in supporting peripheral awareness, continuous coordination, planning, mutual learning, focusing participants' attention and providing each other's status overview. It is important to understand that an artefact - like the prototype remote control is both socially and materially constituted. In the meeting recordings we observed that while creating a prototype remote control the two remotely located designers (UID and ID) discussed, reasoned and changed several aspects of the prototype.

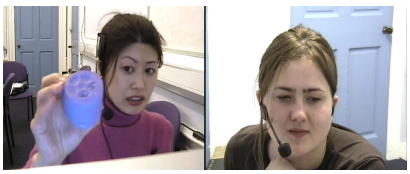
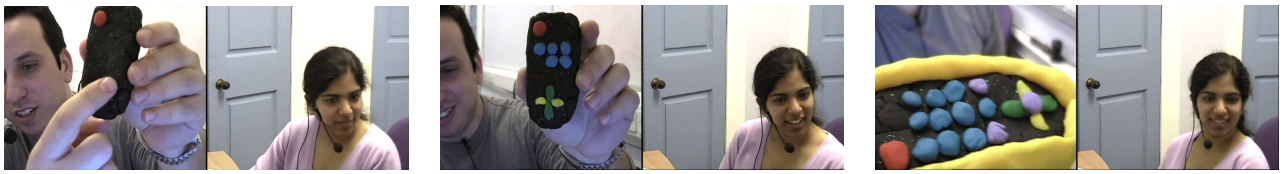


Figure 10: ID (left) shows all the available objects to UID (right)

Available materials We observed that both the participants wanted to have a common-ground about the types of objects they were using for designing. ID has all the design clays and crayons that he uses for designing the prototype. UID needs to know what material ID has in order to be able to better assist the design process. Figure 10 shows ID showing all the design objects she has to the UID in the very beginning of the meeting.

The design object As the remote control prototype is the main purpose of the discussion, ID has to continuously update UID by positioning it close to the camera. Here temporality of design object becomes very important. This temporality could help establishing an understanding of the process that is used in the cooperative design work. Figure 11 shows three different stages of the design of the remote control. Because of the iterative nature of the design process, temporality becomes especially relevant since there will be a need to understand, explain and mediate the design processes involved in it. The temporal dimension of the materiality of artefacts points to different time frames as well.



(a) (b) (c)

Figure 11: Different stages of the remote control projected by ID (left) to UID (right)

Related materials It is also observed in the meeting recordings that participants also use other materials like paper based sketches and drawing diagrams in order to communicate ideas to each other. An example is illustrated in Figure 12.

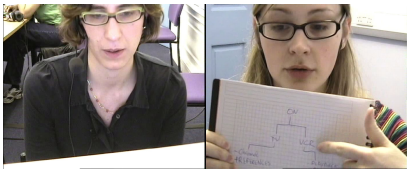
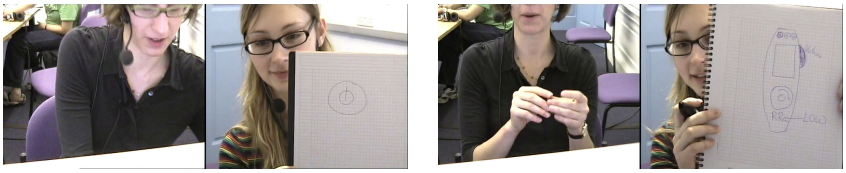


Figure 12: Use of sketches to explain design



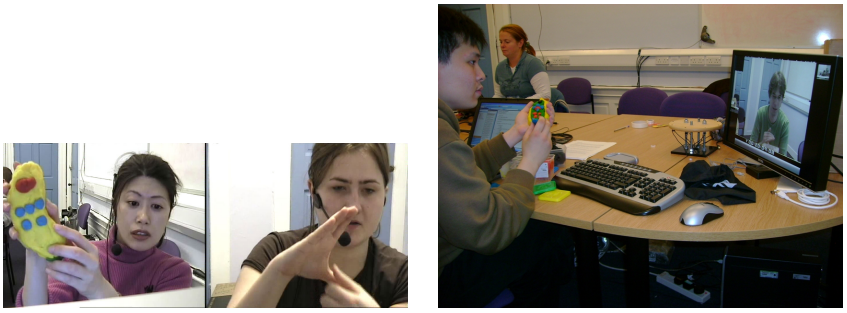
(a) (b)

Figure 13: UID (left) continuously assists ID (right) using drawing sketches.

Continuous consultation through artefacts Figure 12 shows an example of continuous consultation using design sketches. Here UID uses a drawing sketch to constantly assist ID. Importantly, in this case, the development of the physical prototype of the remote control (in the hands of ID) and drawing sketches (in the hand of UID) goes hand in hand. As it can be seen in Figure 13, UID step-by-step works on her drawing while simultaneously explaining her drawings to ID during the meeting.

Artefacts as knowledge landmarks A knowledge landmark is a carrier for information that is left by one participant and can subsequently be used by others to support or foresee any future activities. In the meeting recordings it was seen that artefacts such as drawing in UID's hands or prototype in ID's hands acted as a mean for coordination. During the design meeting, a drawing made by a designer serves as a baseline to carry out further discussions and changes in the drawing. Here the indications left or modifications made by an individual on the artifact provide feedback on themselves and to others. In other sense, the design activities are recorded in the artifact, and this record is used to coordinate work. These knowledge landmarks offered an initiation on collaborative

creativity as participants were able to enhance or change the work done by the previous participants.



(a) (b)

Figure 14: Two examples of material common-ground established by participants.

Material Common-ground There were specific patterns where participants utilized available materials in order to establish common-ground amongst each other. As can be seen in Figure 14 (a) UID (right) mimics the shape of the banana shaped remote control and explains a specific portion that needs to be re-adjusted. Both ID and UID play a role here for establishing a common-ground. In the second example, Figure 14 (b), UID uses a real remote control to suggest changes while the two designers are discussing and collaboratively designing a prototype of a remote control.

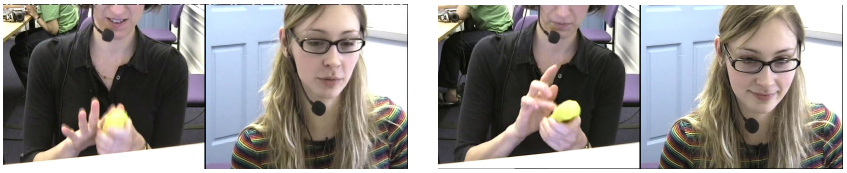
5.3.4 Comments about the video recordings

Because of the limited view of camera on the faces and upper body, the UID does not get frequent updates when ID is working on prototypes (Figure 15). UID has to wait till ID finishes a particular activity of designing. This was observed in almost all the meeting videos. Several times UID asks for updates by asking, ”Show me, how does it look now?” and so on. It was also observed that the UID looses interest in project when he/she does not get updates. This leads to UID focusing away from the camera and doing other activities on the computer (Figure 15). Constant update is needed for the remote participant (UID) to be actively involved in design.

Figure 15: UID (right) distracted while ID (left) works on the prototype.

It was observed in one example where ID changes the projection of the camera on the design object itself. This in fact led to an active participation by the UID (Figure 16).

To avoid this inconsistency, it might be a better idea to have an extra camera focusing on the desk where the ID is building the remote control prototype. This helps UID to choose where she wants to look, especially when she wants to observe the design object.



(a) (b)

Figure 16: UID (right) actively involved in the process as she can see ID's work (left) on the prototype.

6 Meeting Participant Behavioral Differences in Local and Remote Settings

Summary

This report looks at the differences in communication behaviour between local- and remote multi-party settings. A local setting here is where all participants meet face-to-face, whereas in a remote setting, one of the participants is participating through some tele-conferencing system. Three different settings from two different corpora will be analyzed in a semi-systematic way. For both the iCat Corpus, developed at the University of Twente, and the AMIDA corpus, not enough data is available to do a thorough statistical analysis so we are forced to move into the realm of detailed, but subjective, descriptive analysis of the recorded meetings. This analysis is partly backed up by numerical models of the corpora and point out systematic differences in participant's behaviour in the different settings.

6.1 Introduction and Goal

There is a difference between face-to-face conversations and so-called "mediated" conversations, which take place across a certain medium (like a telephone line or the internet). An obvious example of a difference is that you can not see each other in a telephone conversation. This leads to a change in conversational behaviour, for example having to introduce yourself by saying your name when you call someone, even though you know each other very well. This is an obvious example, but other aspects of conversation are likely to be different as well. How does our addressing behaviour change in the absence of available eye-contact? How do participants who are present via tele-conferencing try getting the floor? Answering such questions may help improve the usefullness of software that assist in mediated conversations like business meetings.

A big research project in the area of mediated meeting analysis is AMIDA (Augmented Multi-party Interaction with Distance Acces) for which an extensive corpus is at the time of writing being created [7]. The project includes a wide variety of research topics ranging from human analysis to audio-video processing specifically aimed at meetings where not all of the participants are present locally. The project is a follow-up of the AMI project which has been running for several years and in which much research in these fields has already been done.

7. http://www.amiproject.org/

The goal of this project is to examine the behavior of participants in a meeting (group conversation), where one of the participants is not fysically present in the same room as the other participants. The findings will be compared to the "normal" behavior of people in face-to-face group conversations.

This project will look into three different situations from two different Corpora: the iCat Corpus, developed at the Twente University; and the AMIDA Corpus, developed for the AMI/AMIDA European Research Project. The iCat Corpus contains meeting sessions where the remote participant is either represented in the meeting room by means of web-cam video on a computer screen, or by a small robot cat called the **iCat** (see Section 6.2 for an exact description).

The iCat Corpus will be hand annotated for speech and gaze and will be the primary source for generating and testing behavioral hypotheses. It also makes it possible to see the effect of different kind of representations of the remote participant on behavior of the meeting participants. The AMIDA corpus is used to check the findings of the small iCat corpus against a larger and more mature dataset.

This report is set up as follows. First, the two corpora that are used in the project will be described in Section 6.2. Following that, important background information from the literature will be outlined in Section 6.3. In Section 6.7 a number of hypotheses concerning behavioural differences in remote meetings will be proposed based on the iCat video corpus and a study of the literature. In Section 6.8 the iCat corpus's annotations will be systematically analyzed to find proof for the hypotheses.

Then, if any of the hypotheses can be proven correct, we will look at the AMIDA corpus to see if the same phenomena apply there (section 6.9). The final section contains a discussion on the results of this project.

6.2 Corpora Definitions

The two corpora that are used in this project are the iCat Corpus and the AMIDA corpus. These two corpora are described here.

6.2.1 The iCat Corpus

The corpus that we use mainly for this project is called the iCat Corpus. It is developed by F. van der Veeken and F.L. de Vries van der Veeken [2008] de Vries [2008], to study the effects of the Philips iCat conversational robot, in particular the effects on conversational functions van der Veeken [2008] and the perception of copresence de Vries [2008].

The iCat robot cat (Figure 17) is "a research product of Philips which is used for developing and evaluating human-robot interaction, aimed to apply the results of this research in smart home environments" van der Veeken [2008].

Figure 17: The Philips iCat

The iCat can be remotely controlled via a computer interface. The user is able to make the iCat look to the right, center or left, make facial expressions (surprised, happy, upset, etc...) and nod. The iCat is equipped with a camera (in the nose), so that the user sees what the iCat is looking at. For more information see `http://www.research.philips.com/technologies/syst_softw/robotics/index.html`.

The corpus consists of 4 meeting sessions where two participants are present locally, and the third is represented by the iCat. Each group of three also did a control session where the remote participant is represented by a webcam screen (see Section 6.2.1). In both sessions the participants received two similar design tasks: the first task was to design buttons for a universal remote control, the second task to discuss the shape and color of the device. All participants of the tests where students of a technical study, between the age of 19 and 25, and most of them had experience with remote conferencing. There was no particular distribution of roles over the participants, and noone was given the role of '*project manager*' to lead the discussion.

Since both studies make use of questionaires as their main source of results, the only available data are the mixed audio/video signals and the logs of the iCat behaviour. To get a more detailed look at the behaviour of the meeting participants, the corpus has been annotated in different layers (see Section 6.2.1).

Setting Description We distinguish between the **icat**-setting and the **remote**-setting. The icat-setting is where the remote participant communicates with the two other participants through the Philips iCat. Figure 18 shows the schematic representation of the meeting table as taken from van der Veeken [2008]. Figure 19 shows a screenshot from the actual video recording of the corpus.

The remote-setting is where the remote participant is represented via a webcam video link on a computer screen. Figures 20 and 21 show the diagram and video still situation for the remote setting.

In the remainder of this report the three participants are identified as Speaker A, B and C, which corresponds to the left, center and right positions in the video stills (Figures 19 and 21).

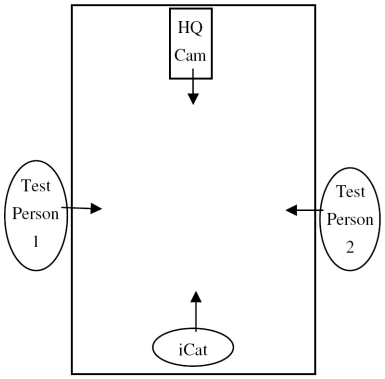


Figure 18: Schematic of the iCat Setting

Figure 19: Video still of the iCat Setting

Annotational layers For the iCat Corpus the following two layers of information have been hand annotated by the author (see Figure 22):

Speaking Layer: for every speaker it is annotated when he or she is talking. Units in this layer are at the 'turn' level, if some one is speaking for a long time without long pauses, this is annotated as a single *speaker turn*. When a participant stops talking for a longer time (usually a few seconds), his *turn* is split into two annotation units.

Gaze: for every speaker (except the remote participant in the **remote**-setting) it is annotated what he or she is looking at. This is either one of 4 options: speaker A, speaker B, speaker C or Other (although speaker A can obviously never look at speaker A). **Other** is a container for whenever someone is not looking at any of the other two participants: this is usually a sheet of paper or the table.

For the **Speaking** layer, the quality of the annotations has not been checked by a second annotator. This is because the annotation task is quite unambiguous. The only hard thing is getting the start- and end- times of a speaker turn exactly right; and some might split certain speaker turns in two where others might not. It is not expected that this has a high impact on the kind of analysis that is done on the data in Section 6.8.

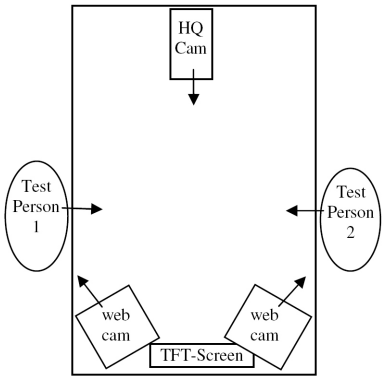


Figure 20: Schematic of the remote Setting

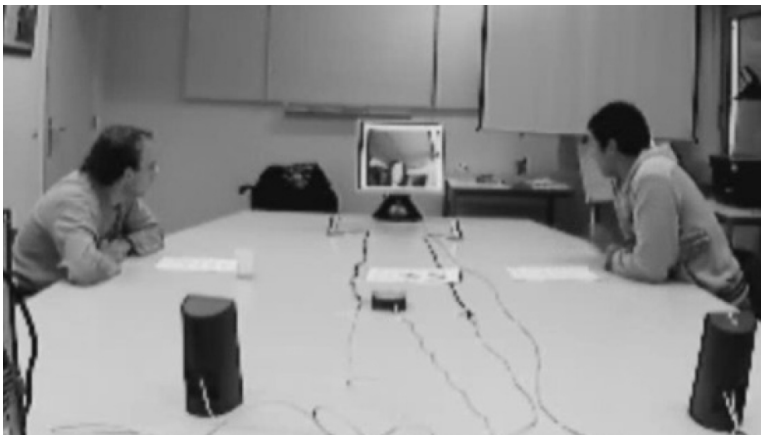


Figure 21: Video still of the remote Setting

The **Gaze** layer is a bit harder to annotate because it can be hard to see where the focus of the participants lie. Therefore, a second annotator[8] annotated two minutes of Gaze layer to see how this compares to the author's annotations. For meeting 1a-remote, the A-LookingAt layer was annotated twice, resulting in 73% exact overlap between the two annotators. Figure 23 shows how the two layers align.

Looking at the similarities in the layers and given the fact that 73% of the layers exactly match each other, the annotations can be considered reasonably accurate. Because the annotation is a continuous labelling, a disagreement of 27% is not that much, considering that a large amount of the disagreement will be at the edges of labellings. An exact analysis of the usefulness of this data is in any case a difficult taks and out of the scope of this research. For more on this topic, see Reidsma [2008].

Currently, the first 4 meetings have been fully annotated: **1a-remote**, **1b-icat**, **2a-icat** and **2b-remote**. See Section 6.8 for the annotation statistics.

8. Thanks goes out to Wilko Wieringa.

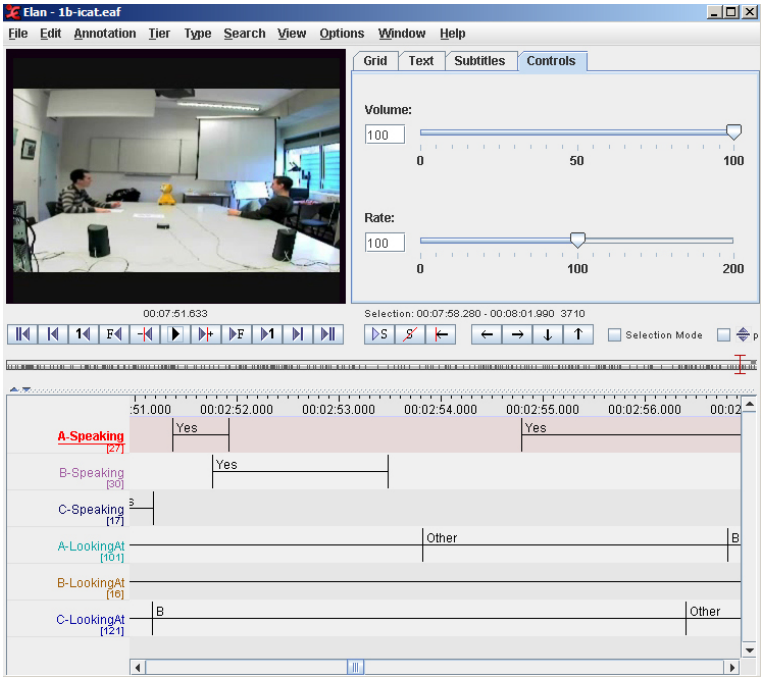


Figure 22: Annotation layers for one of the iCat video's in ELAN

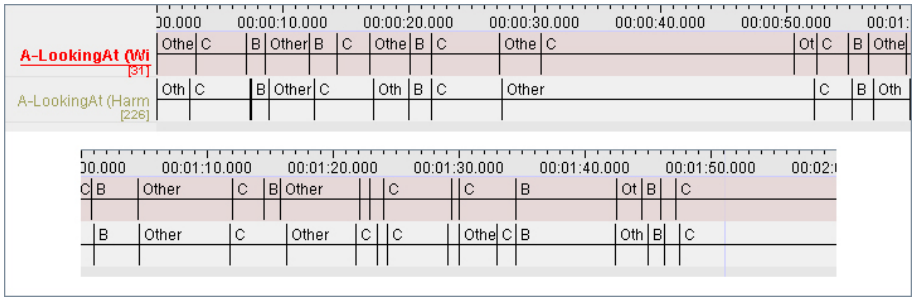


Figure 23: Alignment of the same annotation layer by two different annotators.

6.2.2 The AMIDA Corpus

For the AMIDA project (Augmented Multi-party Interaction with Distance Acces), a follow-up on AMI[9] a corpus has been created to analyze the behaviour of meetings with remote participants. The corpus is planned to contain 13 meeting series (ED1001-ED1013), each consisting of 4 meetings:

A 4 persons in the meeting, all in an instrumented meeting room.
B 4 persons in the meeting, 3 in meeting room, 1 remote.
C 4 persins in the meeting, 3 in meeting room, 1 remote.
X 2 persons in the meeting, (designers making prototype) via Video Conference.

The duration of these meetings vary in length between 15 and 45 minutes each. Because the recording, editing and annotating of this corpus is at the time of writing still very much a work in progress, there is only a limited amount of these meetings available for analyzing.

9. http://www.amiproject.org/

The **B** and **C** meetings of every series, where there are 3 participants in the meeting room and 1 remote participant behind a desktop are the most suitable to compare to the iCat corpus settings. There are currently 4 of these meeting series for which the recordings are finished, as well as some of the annotational layers; these are: ED1002, ED1003, ED1005 and ED1007. The last two series will be the subset of the AMIDA corpus that will be used in this project.

6.3 Communication with a Remote Participant

We first highlight some important background theory from the literature of mediated conversation and conversational analysis in general. We use this as a starting point to generate the hypotheses on behaviour in the remote meeting setting in Section 6.7.

Clark defines 10 features of face-to-face conversation Clark [1996]. These 10 features will be used to typify the situations that we are looking at with the iCat and the on-screen remote participant. These features are repeated below.

1. Copresence The participants share the same physical environment.
2. Visibility The participants can see each other.
3. Audibility The participants can hear each other.
4. Instantaneity The participants perceive each other's actions at no perceptible delay.
5. Evanescence The medium is evanescent - it fades quickly.
6. Recordlessness The participants' actions leave no record or artifact.
7. Simultaneity The participants can produce and receive at once and simultaneously.
8. Extemporainety The participants formulate and execute their actions extemporaneously, in real time.
9. Self-determination The participants determine for themselves what actions to take when.
10. Self-expression The participants take actions as themselves.

Table 10 describes for the icat-, remote-, and face-to-face settings in how far the characteristics apply.

Table 10: Ten features of face-to-face communication (with applicability to the iCat setting).

		Face-to-Face	iCat Setting	Remote Setting
1	Copresence	++	–	–
2	Visibility	++	-	+/-
3	Audibility	++	+/-	+/-
4	Instantaneity	++	–	–
5	Evanescence	++	++	++
6	Recordlessness	++	+/-	+/-
7	Simultaneity	++	+/-	+/-
8	Extemporainety	++	++	++
9	Self-determination	++	+/-	++
10	Self-expression	++	+/-	++

In the remote meetings, not all of these features (fully) apply. We try to explain how each characteristic differs from the face-to-face setting below:

- Copresence: in both iCat and Remote settings, the participants do not share the same physical environment.
- Visibility: in the iCat setting the local participants can see a "representation" of the remote participant, while the remote participant can only see either one of the other two participants. In the remote setting, everyone can see each other, although the remote participant is represented via a screen for the local participants and vice versa. The remote participant can see the upper body of the local participants, but not the other way around. The remote setting does more closely resemble the face-to-face setting in this way.
- Audibility: although it's probably not exactly what Clark meant, there is a definite problem with audibility due to the bad microphone signal from the remote participant to the meeting room and vica versa.
- Instantaneity: although the remote participant is in the same building as the "local" participants; there is a substantial lag[10] in the communication channel compared to a face-to-face conversation. In real life applications of remote meetings this lag becomes even more apparent, because remote participants tend to be further away from the workfloor.
- Evanescence: this characteristic fully applies to all settings, because there is, for example, no communication in writing.
- Recordlessness: although the meetings are recorded, this is purely for evaluative purposes and should have no effect on the behaviour of the participants. However, in a real remote meeting situation, it would be very easy for the remote participant to record the audio and video layers; this may be in the back of the heads of the meeting participants.
- Simultaneity: although in theory every meeting participant can speak simultaneously, both the iCat and Remote setting differ from a face-to-face situation in that when the remote participant speaks, the local participants must focus to hear what he is saying due to the audio quality.
- Extemporainety: this applies to all settings.
- Self-determination: In the iCat setting, when you see the iCat as the meeting participant, it does not determine for itself what actions it takes.
- Self-expression: Again, in the iCat setting, the remote participant may feel that he takes actions as himself, but in the meeting room, it looks like he takes his actions as an iCat.

It's important to keep in mind that for one of the three participant-pairs (the two locally present speakers) the conversation **is** face-to-face and all 10 features do apply.

Every feature that is missing or limited poses problems in communication; although you could say that the problems caused by absence of *copresence* could be alleviated by

10. *Substantial being: somewhere in the hundreds of milliseconds, being a noticable delay for human-to-human communication.*

perfect *visibility*, *audibility* and *instantaneity*. Remote meetings aim at providing just that, but it is not perfect yet. Therefore, the problems caused by the lack of these three features are explored here.

6.4 Lack of Visibility

Absence of full visibility can cause problems for *addressing* (or *referring* in general op den Akker and Theune [2008]), *turn-taking* and *grounding*. For addressing and turn-taking, you can normally just look at someone to ask something. This way you are addressing him or her and also pass on the next turn. Jovanovic [2007] shows that Visual Focus of Attention is good cue for predicting who is being addressed. This probably has to happen in a more explicit way in a remote meeting setting. The same thing applies to grounding, were normally regular head-nods can signal that you understand what is being said, and the speaker can continue to speak. This also must be done explicitly, or else confusion may arise. We thus expect differences in behaviour in these area's.

6.5 Lack of Audibility

Although there is an audio link between the remote participant and the local participants, the quality is far less than in a face-to-face conversation. The reduced audio quality also cause for long pauses in the conversation ("...did I hear him correctly?").

6.6 Lack of Instantaneity

There is a delay in the audio signal that, although it is small, causes the conversation to loose its pace. This is expected to have a negative influency on turn-taking for the remote participant. Because of the slightly longer delay, he is expected to have more difficulties in grabbing the floor compared to the local participants.

6.7 Hypotheses

Based on the theoretical analysis, we derive the following hypotheses for the remote meeting scenario's:

1. The local participants will be more dominant in the conversation, because the remote participant will have problems grabbing the floor.
   (a) Because the iCat and Remote user lack the same abilities for grabbing the floor, there will be no difference between the dominance of the iCat-user and the Remote-user.
2. The conversation between the local participants will be faster, with more overlaps in speech.
   (a) Pauses in the conversation will be bigger after the remote participant has spoken.
   (b) The remote participant only starts speaking after a relatively long pause.
3. The local participants will feel more comfortable communicating with the remote participant in the **Remote** setting, because the remote participant adheres more to the *Self-determination* and *Self-expression* characteristics of conversation.
   (a) For this reason, the local participants look more often to the Video screen than to the iCat.

6.8 Analysis of the iCat Corpus

6.8.1 General Statistics

Table 11 contains some basic numbers on the iCat Corpus. The first column is the name of the recording, e.g. *1a-remote* is series 1, session a, "remote setting" (remote participant represented by computer screen). The last two columns, *Start-time* and *End-time* depict the 'official' start and end times of the meeting in the raw recordings by de Vries [2008] and van der Veeken [2008].

Table 11: General iCat Corpus Statistics

Name	**Duration**	**Duration (sec)**	**Start-time**	**End-time**
1a-remote	11:37	697	01:12	12:49
1b-icat	08:02	482	14:38	22:40
2a-icat	08:33	513	04:32	13:05
2b-remote	09:45	585	01:30	11:15
3a-icat	16:00	960	08:10	24:10
3b-remote	09:40	580	04:50	14:30
4a-remote	11:29	689	05:59	17:28
4b-icat	09:26	566	07:12	16:38
Time total:	**1:24:32**			

Two remarks should be made about the *1b-icat* recording here. In the first 4 minutes, the sound of another recording is mixed through the session's own sound. Second, at the end of the recording the video signal goes black, so this has been left out of the session data. This means that there is no real end to the discussion.

The next three subsections highlight the statistical information on *speaking*, *gaze* and *turn-taking*. **Speaker A** is always the one on the left (in the video), **Speaker B** is the remote speaker and **Speaker C** the one on the right. The final subsection summarizes the results of the questionnaires done by de Vries [2008] and van der Veeken [2008].

6.8.2 Speaker Statistics

Table 12 shows how often each participant in every meeting is talking. The table shows the total time someone is speaking (in seconds, column 3), the percentage of time this is from the total meeting duration (column 4), the number of annotation units (or roughly: speaker turns) and the average duration per unit (again, in seconds). Note that for example *4.7 seconds* means 4s700ms.

Table 12: iCat Corpus Speaker Statistics

Meeting	Speaker	Time Speaking	Speaking %	Units	s/unit
1a-remote	A	202	29%	43	4.7
	B	145	21%	47	3.1
	C	266	38%	32	8.3
1b-icat	A	123	26%	27	4.6
	B	124	26%	30	4.1
	C	140	29%	17	8.3
2a-icat	A	161	31%	26	6.2
	B	77	15%	27	2.9
	C	161	31%	58	2.8
2b-remote	A	161	28%	62	2.6
	B	80	14%	33	2.4
	C	179	31%	72	2.5

Hypothesis 1: The local participants have on average 169 speaker turns, lasting 697 seconds (4.1 s/turn). The remote participants have 137 speaker turns, lasting 426 seconds (3.1 s/turn). This shows that the remote participants do indeed on average speak less than the local participants and even have shorter turns.

Hypothesis 1b: The remote "video" participants have a total of 80 turns, lasting 225 seconds (2.8 s/turn). The iCat participants have 57 turns, lasting 201 seconds (3.5 s/turn). In terms of the time speaking, this is only a very small difference, but the remote video participant does manage to grab more turns than the iCat participant.

6.8.3 Gaze Statistics

Tables 13 and 14 shows for meeting series 1 and 2 respectively who looks at who in a matrix form. Row 'A', column 'Other' contains the percentage that Speaker A is looking at Other (e.g. the table). The 1b-~~icat~~ table (with strikethrough 'icat') contains the information for that meeting when not considering the gaze behaviour of the iCat. That

makes, for example, 1b-~~icat~~ and 1a-remote comparable (because they only consider gaze behaviour of local participants). These extra entries have been added because the iCat's gaze behaviour is very unnatural.

Table 13: iCat Corpus General Gaze Statistics (Series 1)

Meeting	Speaker	A	B	C	Other	Units	Avg Unittime
1a-remote	A	-	32%	25%	43%	226	3.1
	B	-	-	-	-	-	-
	C	30%	40%	-	30%	205	3.4
Total time:		210	503	172	511		
Total units:		57	175	60	129		
Avg unit time:		3.7	2.9	2.9	4.0		
1b-icat	A	-	26%	14%	60%	101	4.8
	B	53%	-	46%	1%	16	30.1
	C	27%	25%	-	48%	121	4.0
Total time:		383	243	287	534		
Total units:		41	83	33	81		
Avg unit time:		9.3	2.9	8.7	6.6		
1b-~~icat~~	A	-	26%	14%	60%	101	4.8
	~~B~~	~~53%~~	-	~~46%~~	~~1%~~	~~16~~	~~30.1~~
	C	27%	25%	-	48%	121	4.0
Total time:		128	243	67	526		
Total units:		33	83	26	80		
Avg unit time:		3.9	2.9	2.6	6.6		

Some remarks concerning these numbers:

1. People look more at the screen (remote setting) than at the iCat. For speaker A: 32% vs 26%, for speaker C: 40% vs 25%.
2. The time spend *not looking* at the iCat is seemingly spent looking at the surroundings. For speaker A: 43% vs 60%, for speaker C: 30% vs 48%.
3. The iCat's gaze behaviour is very different than that of the local participants. It is characterized by very long stares at either one of the other speakers, and few focus of attention shifts. This makes it very unnatural.

Hypothesis 3a: The local participants do indeed look more at the video screen than at the iCat.

Looking at the iCat video's, it feels as though the local participants only look at the iCat whenever it's performing one of his actions, like looking to the left or right, nodding, smiling, etc. In order to see if the statistics can back up this idea, we looked at the times that the participants look at the iCat and looked if the iCat was taking an action right before that. The results can be seen in Figures 24 and 25.

Table 14: iCat Corpus General Gaze Statistics (Series 2)

Meeting	Speaker	A	B	C	Other	Units	Avg Unittime
2a-icat	A	-	7%	29%	64%	65	7.9
	B	36%	-	58%	6%	12	42.8
	C	35%	8%	-	57%	94	5.5
Total time:		360	77	448	653		
Total units:		38	31	30	72		
Avg unit time:		9.4	2.4	14.9	9.1		
2a-~~icat~~	A	-	7%	29%	64%	65	7.9
	~~B~~	~~36%~~	-	~~58%~~	~~6%~~	~~12~~	~~42.8~~
	C	35%	8%	-	57%	94	5.5
Total time:		177	77	150	622		
Total units:		34	31	24	70		
Avg unit time:		5.2	2.4	6.3	8.9		
2b-remote	A	-	36%	35%	29%	123	4.8
	B	-	-	-	-	-	-
	C	26%	27%	-	47%	162	3.6
Total time:		152	365	205	449		
Total units:		51	95	40	99		
Avg unit time:		3.0	3.8	5.1	4.5		

In meeting 2a-icat, 25% of the glances happen right after an iCat action (1s) and 50% within 4 seconds. This shows that in this meeting, the local participants are reasonably "distracted" by these iCat movements. In meeting 1b-icat, the numbers do not really show such behaviour.

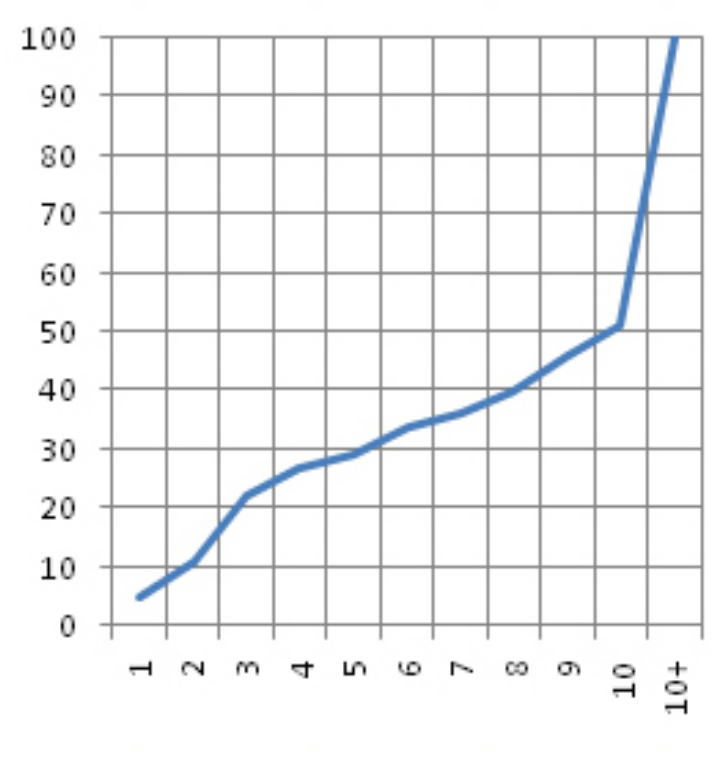


Figure 24: The percentage of glances at the iCat (y-axis) within a number of seconds after an iCat action (x-axis) for meeting 1b-icat.

Table 15 show how much people look at the one who is speaking. For 1a-remote: Row A, column C (107 (52%)) means that C has been looking at speaker A while he was talking for 107 seconds, which is 52% of the total time that speaker A has been speaking (202 seconds total, see Table 12).

Table 15: iCat Corpus Speaker/Gaze Statistics

Meeting	**Speaking**	**A-Looking**	**B-Looking**	**C-Looking**
1a-remote	A	-	-	107 (52%)
	B	80 (55%)	-	79 (54%)
	C	89 (34%)	-	-
1b-icat	A	-	96 (78%)	50 (41%)
	B	44 (35%)	-	17 (13%)
	C	19 (13%)	136 (97%)	-
2a-icat	A	-	94 (58%)	91 (56%)
	B	3 (4%)	-	3 (4%)
	C	91 (56%)	104 (65%)	-
2b-remote	A	-	-	61 (38%)
	B	64 (80%)	-	53 (66%)
	C	93 (52%)	-	-

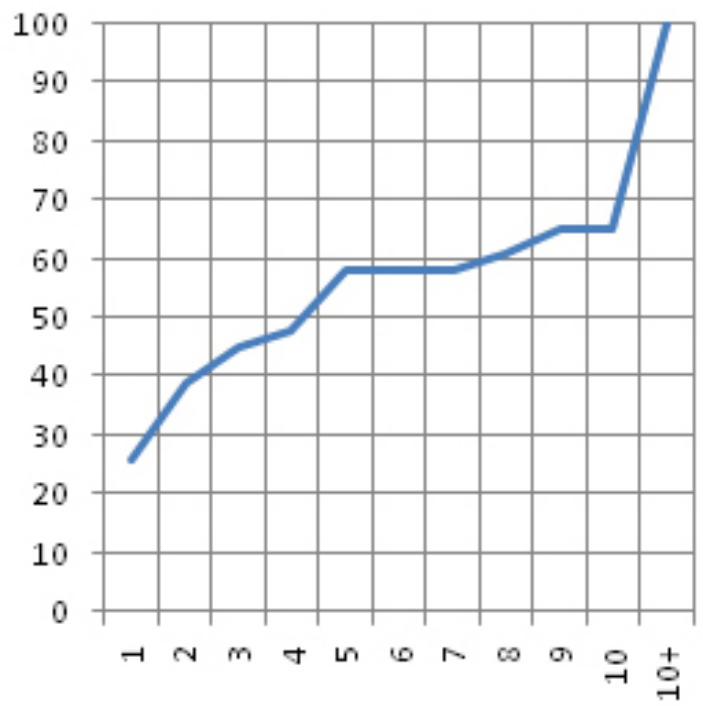


Figure 25: The percentage of glances at the iCat (y-axis) within a number of seconds after an iCat action (x-axis) for meeting 2a-icat.

6.8.4 Turn Taking Behaviour

Table 16 show the frequency of occuring Bi-Grams for the 4 meetings. The last three columns show how often, for that Bi-Gram an overlap occurs, how much time those overlaps cover and what the average time per overlap is.

The bigrams “AA” in the table above means that speaker A had had a turn, then there was a pause in which other speakers had time to take over, but speaker A eventually decided to continue talking.

Figures 26 and 27 show the distribution of pauses between bigrams of speaker-pairs. The blue bars show the AB and BA bigrams (left), the red bars show the AC and CA bigrams (center) and the green bars show the BC and CB bigrams (right). The horizontal axis is divided as: pause between -3 and -2 seconds ... pause between 6 and 7 seconds. A negative pause means that there is that much overlap between the two corresponding speaker turns.

In the second series (2a and 2b), the red bars tend to lean more the left, showing that there are shorter pauses and more overlap between the speach of the two local participants. In the first series (1a and 1b) this tendency is less clear. Table 17 shows the total amount of AB/BA (etc...) bigrams and the average pause length between them for all four meetings.

Table 16: iCat Corpus Speaker Statistics

Meeting	Bi-Gram	Frequency	Percentage	Overlaps	Time	Avg
1a-remote	AB	24	20%	5	3.8	0.76
	BA	19	16%	5	3.1	0.62
	BC	18	15%	3	0.5	0.17
	CA	16	13%	3	2.1	0.70
	CB	13	11%	6	4.2	0.70
	AC	10	8%	4	1.4	0.35
	BB	10	8%	-	-	-
	AA	8	7%	-	-	-
	CC	3	2%	-	-	-
1b-icat	AB	16	22%	4	1.4	0.35
	BC	14	19%	2	0.5	0.25
	BA	10	14%	3	2.7	0.90
	CA	9	12%	-	-	-
	CB	8	11%	2	0.9	0.45
	AA	7	10%	-	-	-
	BB	6	8%	-	-	-
	AC	3	4%	-	-	-
2a-icat	CC	20	18%	5	1.7	0.34
	CB	20	18%	1	0.2	0.17
	BC	19	17%	1	0.1	0.14
	AC	18	16%	-	-	-
	CA	18	16%	2	0.7	0.35
	AA	6	5%	-	-	-
	BB	5	5%	-	-	-
	AB	2	2%	-	-	-
	BA	2	2%	-	-	-
2b-remote	CA	30	18%	11	3.5	0.32
	AC	28	17%	7	2.1	0.30
	AA	26	16%	-	-	-
	BC	23	14%	4	1.9	0.48
	CB	22	13%	3	0.4	0.13
	CC	20	12%	-	-	-
	AB	8	5%	2	1.0	0.50
	BA	6	4%	-	-	-
	BB	3	2%	-	-	-

Hypothesis 2: The last three meetings here show that communication between the two local participants (A and C) goes quicker than when the remote participant is involved. It is strange to notice that in the first meeting (1a-remote), it is exactly the other way around. Overall the hypothesis can be considered confirmed.

Table 18 shows the average pause duration before and after each participant turns. The last two columns show the average pause before and after the speaker turns disregarding

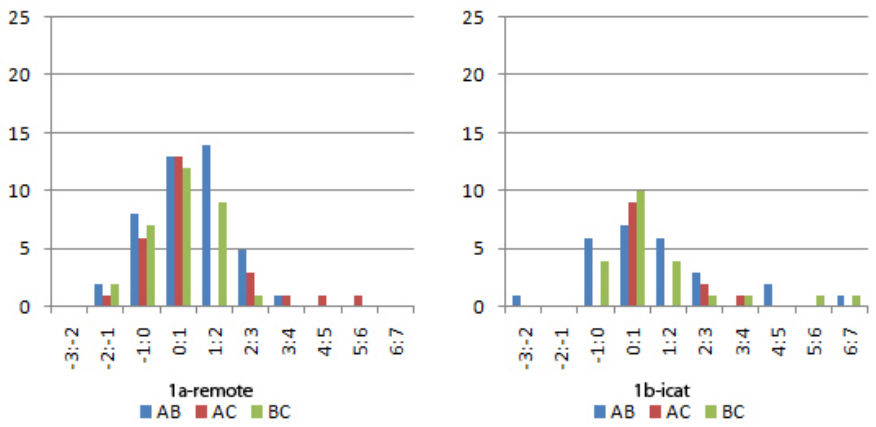


Figure 26: Pause duration between speaker-pairs for meetings 1a-remote and 1b-icat.

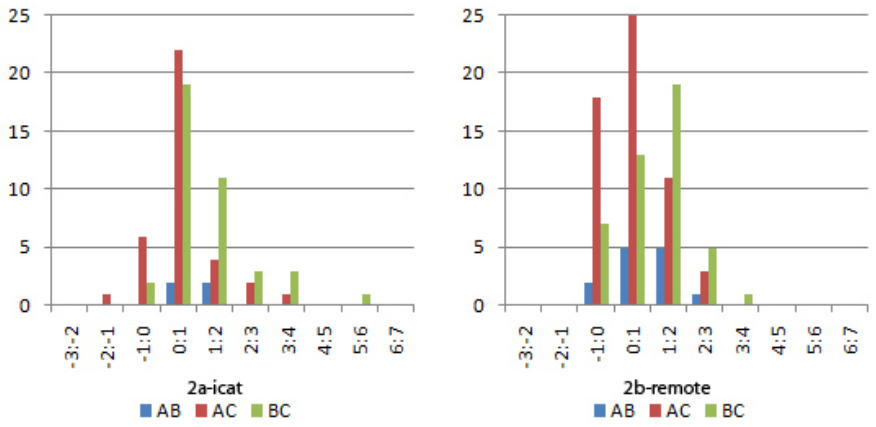


Figure 27: Pause duration between speaker-pairs for meeting 2a-icat and 2b-remote.

Table 17: iCat corpus average pause durations between two participants (including overlaps)

Meeting	**Bi-Grams**	**Frequency**	**Avg. Pause**
1a-remote	A+B	43	0.758
	A+C	26	0.834
	B+C	31	0.580
1b-icat	A+B	26	1.240
	A+C	12	1.008
	B+C	22	1.186
2a-icat	A+B	4	1.122
	A+C	36	0.651
	B+C	39	1.245
2b-remote	A+B	14	1.954
	A+C	58	0.665
	B+C	45	1.008

overlaps in the speech (NO = No Overlaps).

Table 18: iCat corpus average pause durations before and after each participant.

Meeting	Participant	Before	After	Before (NO)	After (NO)
1a-remote	A	1.225	1.265	1.655	1.765
	B	1.932	1.795	2.743	2.257
	C	0.900	1.059	1.241	1.747
1b-icat	A	1.197	2.113	1.470	2.561
	B	1.940	1.357	2.519	1.756
	C	1.310	0.936	1.518	1.119
2a-icat	A	1.907	1.832	2.095	2.350
	B	1.547	1.091	1.613	1.141
	C	1.712	1.947	1.950	2.069
2b-remote	A	1.683	1.255	2.114	1.525
	B	1.219	1.388	1.487	1.655
	C	1.371	1.662	1.688	2.131

Hypothesis 2a: The data in Table 18 does not show that there is on average a longer pause after the remote participant (B) has spoken.

Hypothesis 2b: The data in Table 18 does not show that there is on average a longer pause before the remote participant (B) starts speaking.

6.8.5 Questionnaire results for the iCat corpus

Both de Vries [2008] and van der Veeken [2008] have conducted questionnaires with the participants of the recorded meetings in the iCat corpus. The results are summarized below, starting with the main results from de Vries [2008]:

- The local participants prefer the *video/remote* situation; the remote participant had no preference between *iCat* or *remote*.
- The participants in general had a larger feeling of copresence in the *video/remote* situation compared to the *icat* situation.

The next results are taken from van der Veeken [2008]:

- In terms of involvement, the local participants preferred the *video/remote* setting, the remote participant preferred the *iCat* setting.
- In terms of satisfaction, pleasure, helpfulness of the group members, and information sharing, the local participants preferred the *video/remote* setting as well. The remote participants once again preferred the *iCat* setting.
- All participants preferred the *video/remote* settings for turn-taking purposes.
- For adressing, the local participants preffered the *video/remote* setting. The remote participants preferred the *iCat* setting.
- For grounding the preferences where turned around. The local participants found this easier in the *iCat* setting whereas the remote participant preferred the *video/remote* setting.

- In general, the local participants preferred the *video/remote* setting, whereas the remote participants where neutral.

Hypothesis 3: The results of the questionnaires confirm that the meeting participants feel in general more comfortable with the *video/remote* setting. The reason for this remains unclear.

6.9 Comparison with the AMIDA Corpus

We are interested in knowing how the findings of the experiments with the iCat and Remote video settings apply to a more mature corpus, the AMIDA corpus. Keeping in mind the hypotheses that appear to hold truth according to the analysis of the iCat corpus, the recordings of several meetings in the AMIDA corpus have been thoroughly examined. In particular, the meetings that have 3 local participants and 1 remote participants are considered. The reason for this is that this setting most closely resembles the iCat corpus settings. The sections below contain lists of notes where there is something noticable about the communication between the remote participant and the local participants for a series of AMIDA meetings.

6.9.1 ED1005b

- During the first presentation of one of the local participants, the other participant's spend a lot of time looking at the RP, even though the presentation is addressed to the group.
- During the second presentation, all local participants spend long stretches of time looking at the RP's screen, including the current speaker, even though the presentation is addressed to the group.
- The transition from a local participant's presentation to the presentation of the RP goes very smoothly through direct addressing of the RP by the Project Manager (who can be seen as discussion leader).
  A: So we're now going on to your section Lawrence, if you wanna take control here...
- A short repair made by one of the local participants to one of the other local participants's speech is confusing for the RP, and his attempt to understand it is being ignored.
  D: ... so uh we could um um
  A: Incorporate the feature to find the remote control.
  D: Yeah that's right.
  RP: Can you repeat that?
  D: And then uh fifty percent ...
- During the discussion sections (non-presentation), the RP's contributions are only there when specificially asked for, and are usually very short.
  A: Lawrence, you okay with that?
  RP: Yes.

6.9.2 ED1005c

- The first time the RP is addressed, five consecutive explicit address terms are being used before he (Lawrence) notices that it is his turn to speak:
  B: ... Uh I'll pass over to my uh user interface designer if he's here.
  B: Lawrence?
  A: Lawrence?
  B: Knock, knock.
  A: Lawrence.

- The end of the RP's "presentation" ends with a long list of partly overlapping confirmations that everything was clear:
  B: Okay.
  A: Alright, thank you.
  RP: Any questions?
  A: O uh it's good.
  RP: Clear as crystal?
  A: That was good.
  RP: Okay.
  A: Okay?
  RP: Okay.
- On the first attempt of the RP to break into a conversation between the local participants to give his opinion about one of the evaluation criteria, every attempt gets interrupted with overlapping speech by one of the locals. Eventually the RP seems to give up on that points and says "*One for the very good.*" (meaning he votes perfect for the criteria being discussed).
- On a number of occassions, when the opinion of the RP is wanted, the RP needs a quick recap of what was going on, taking the speed out of the conversation:
  A: Lawrence?
  RP: For what?
  D: For theme.
  A: Pick a number. [laugh]
  B: [laugh]
  D: [laugh]
  RP: A number for what? [laugh] Which section?
  A: For the theme.
  D: [laugh]
  RP: Theme, mm. I'd say we should go for two ...
- When the RP is paying more attention to the meeting at hand, i.e. when he is looking at the screen displaying the meeting room, his responds to the conversation are quick and uninterrupted.
- Short feedback respones like "*good*" and "*looks good to me*" happens frequently within the meeting room, but are used rarely by the remote participant.
- When the RP makes a short comment during a discussion, these comments are usually left uncommented on by the local participants, leaving the RP unsure whether he is being heard, leading to comments like:
  RP: Do you happen to hear me?
  – ...
  RP: Tsveta, do you hear me? Oh, have you heard me a few moments ago?

6.9.3 ED1007b

- During the first presentation in the meeting, most of the visual focus of attention (gaze) is in the direction of the projector screen and the RP-screen.
- During the first presentation, multiple short discussions take place between local participants. The remote participant is completely left out here. During these discussions, al gaze is directed from and to the local participants.

- The first change of floor to the RP is done via the project manager, addressing the RP by its role; this goes very smoothly.
- The moment the project manager initiates a short discussion with the RP, a lot of speaker overlap occurs, where both participants try to take the ground, and both don't give up to easily. This leads to a seemingly natural discussion, with quick reactions to each other, be it with much more overlap than discussions between local participants.
- During the course of the meeting, the RP comments on the discussion a few times. This is often a quite rude interruption, where the RP raises his voice to get the attention of the local participants. Because of the sudden grab of floor, there is no immediate feedback from the local participants, leading to confusion on the side of the RP:
  RP: ... Can you hear me?

6.9.4 ED1007c

- The remote participant gets called away to some important business after a couple of minutes in the meeting, so this meeting does not contain too much interaction with the RP.
- Halway in the local participants need the attention of the RP, so they use a special signal (beep) to notify the RP. This works good, and a temporary discussion between the local participants and RP goes on without too much difficulty.
- A thing to note about visual focus of attention to the RP in this meeting is that the RP is never really in the camera picture, so there is no real reason to look at him.

6.9.5 Generalization

- In general, communication between RP and local participants seem to go more natural in the 1007 meeting series. A reason for this could be that the RP in 1007b and c is a little more dominant in interrupting the meeting and not letting go of the floor too easily.
- The following two tables (19 and 20) taken from UTwente [2008] show the number of dialogue acts uttered by every speaker in the meetings ED1005b and ED1005c, including to whom that dialogue act is addressed. The "**Total**" column that is added here shows the total number of dialogue acts for every speaker. Strangely, these numbers are distributed quite evenly among the speakers. Because both meeting contain quite long presentations where a single speaker is uninterruptedly speaking for several minutes, a last column is added where the amount of uttered dialogue acts that do not fall within a "*presentation*" are counted. These numbers are in line with **Hypothesis 1** that local participants are more dominant in a conversation.
- Regarding **Hypothesis 2a** and **Hypothesis 2b** the following can be said: because the network delay in the AMIDA meetings is extremely low, the longer pause theory does not really apply. What does happen in discussions between RP and local participants is that a lot of overlapping speech occurs. These overlaps however, do not compare well to the natural occuring overlaps that cause for fast floor swithces

	PM	ID	UI(R)	ME	ALL	UNK	Total Ind	Total	Total-PR
PM		7	20	28	116	55	55	226	226
ID	21		1	2	80	62	24	166	64
UI(R)	17			12	96	29	29	154	64
ME	20	1	10		51	70	31	152	98
Totals:	58	8	31	42					

Table 19: Amida Remote Conceptual Design Meeting: ED1005b; Who talks to whom table of a Conceptual Design Meeting. Roles: PM=Project Manager; ID=Industrial Design; UI(R)= User Interface Design (the Remote Participant); ME=Marketing Expert.

	PM	ID	UI(R)	ME	ALL	UNK	Total Ind	Total	Total-PR
PM		16	33	44	83	61	93	237	235
ID	39		22	31	71	96	92	259	215
UI(R)	19	3		23	138	65	45	248	149
ME	46	21	27		56	75	94	225	225
Totals:	104	40	82	98					

Table 20: Amida Remote Detailed Design Meeting: ED1005c; Who talks to whom table. Roles: PM=Project Manager; ID=Industrial Design; UI(R)= User Interface Design (the Remote Participant); ME=Marketing Expert.

in natural face-to-face conversation. They do not just happen at the end of utterances, and it often leads to confusion (both speakers don't understand what the other one is saying). So although speech overlaps a lot, the actual speed of the discussion is not high, because people have to repeat everything they said during the long overlaps or it simple does not come across to the addressee. That means that **Hypothesis 2** is partly true, the speech between local participants **is** faster, but not because of more overlaps, but because of more natural overlaps in speech.

6.10 Conclusion

The conversational analysis performed in this work managed to confirm some of the Hypotheses made in Section 6.7. In some of the cases, the sparse statistical data alone is not enough to confirm or deny a statement entirely, but in these cases, looking at hours of multi-party conversation recordings helps to get an idea of where certain conversational aspects go wrong. The problem of proving them then becomes one of getting more data, and looking at the right type of data. Because large annotated corpora are not available at this moment, this can not be done now.

Below follows a summary of the results of the research presented by the list of hypotheses:

1. **The local participants will be more dominant in the conversation, because the remote participant will have problems grabbing the floor.** This has been proven correct by the analysis of speaker turns. Local participants take more and longer turns than remote participants (see Table 12).

1b. Because the iCat and Remote user lack the same abilities for grabbing the floor, there will be no difference between the dominance of the iCat-user and the Remote-user. The same speaker turn analysis shows that the iCat-users have quite a few less turns, but only a little less time speaking than the video-users. You could say that the video-users are a bit more dominant, but the differences are small.

2. The conversation between the local participants will be faster, with more overlaps in speech. By looking at the pauses between speaker turns in Table 17 and Figures 26 and 27 we can conclude that in general, conversation between local participants goes faster. The data shows that in three of the four iCat meetings, the pause between local participant turns is on average lower than the pause between a local and the remote participant. However, in one of the meetings (1a-remote) this is exactly the other way around.

2a. Pauses in the conversation will be bigger after the remote participant has spoken. Table 18 shows mixed results in terms of pause duration after a remote participant turn. This hypothesis can not be confirmed.

2b. The remote participant only starts speaking after a relatively long pause. Table 18 shows mixed results in terms of pause duration before a remote participant turn. This hypothesis can also not be confirmed.

3. The local participants will feel more comfortable communicating with the remote participant in the Remote setting, because the remote participant adheres more to the *Self-determination* and *Self-expression* characteristics of conversation. The results of the questionnaires in Section 6.8.5 confirm that the meeting participants feel in general more comfortable with the *video/remote* setting, although the reason for this remains unclear.

3a. For this reason, the local participants look more often to the Video screen than to the iCat. The local participants do indeed look more at the video screen than at the iCat (see Section 6.8.3). A clear reason for this can not be derived here, altough the participants have been found more comfortable with the video-setting (Hypothesis 3).

Figure 28: Overview of the Meeting Room in Edinburgh where the scenario based Amida remote meetings were recorded. At the right the screen where the interface designer is visible for the three local participants. This is meeting ED1002B: the Project Manager is seated in the middle front position, closest to the RP screen.

7 Remote Meetings

The section contains some results of our analysis of remote meetings. Based on our findings we present what meeting support technology that aims to improve user engagement in remote meetings should offer its users.

7.1 Addressing in Amida Remote Design Meetings

Does addressing behavior differ between remote and face-to-face meetings?

In the scenario based Amida meetings ED1005b and ED1005c, three members of the group are sitting in the same room, one member is remote. The remote participant (RP) has the role of User Interface Designer (UI). The RP has a video and audio contact with the meeting room (MR). The people in the MR can see the RP on a tv screen. Figure 28 shows an overview of the meeting room in Edinburgh where the Amida meetings were recorded.

The RP could not adjust the camera view.

Tables 21 and table 22 show who talks to whom in these two meetings. The table entries count the number of dialogue acts (proper acts as well as backchannels, stalls and fragments) spoken. For each speaker there is a row. The ALL column contains the numbers of acts addressed to the Group. The number in the UNKNown column is the sum of the number of improper acts and the number of acts for which it was not clear to the annotator whether it was I(ndividually)-addressed or G(roup)-addressed; the right most column contains the total number of I-addressed acts. Thus table 21 shows that in Amida meeting ED1005b the Project Manager is the most talkative; and she is also most I-addressed by the others[11]

11. This we see often in the AMI and Amida meetings: PMs G-address more than others and they are most I-addressed by others. In general: the order of talkativity of participants is often the same as the order in frequency of being addressed by others (see Bales and Gibson).

	PM	ID	UI(R)	ME	ALL	UNK	Total Ind
PM		7	20	28	116	55	55
ID	21		1	2	80	62	24
UI(R)	17			12	96	29	29
ME	20	1	10		51	70	31
totals	58	8	31	42			

Table 21: Amida Remote Conceptual Design Meeting: ED1005b; Who talks to whom table of a Conceptual Design Meeting. Roles: PM=Project Manager; ID=Industrial Design; UI(R)= User Interface Design (the Remote Participant); ME=Marketing Expert.

	PM	ID	UI(R)	ME	ALL	UNK	Total Ind
PM		16	33	44	83	61	93
ID	39		22	31	71	96	92
UI(R)	19	3		23	138	65	45
ME	46	21	27		56	75	94
totals	104	40	82	98			

Table 22: Amida Remote Detailed Design Meeting: ED1005c; Who talks to whom table. Roles: PM=Project Manager; ID=Industrial Design; UI(R)= User Interface Design (the Remote Participant); ME=Marketing Expert.

How does these figures compare to the figures of the face-to-face design meetings? Tables 23 and table 24 show who talks to whom in the AMI face-to-face meetings IS1001b and IS1001c, the co-located counterparts of the two remote Amida design meetings.

	PM	UI	ID	ME	ALL	UNK	Total Ind
PM		42	21	18	132	114	81
UI	35		16	6	100	82	57
ID	7	12		5	68	91	24
ME	15	4	3		50	66	22
totals	57	58	40	29			

Table 23: AMI Face-to-face Meeting: IS1001b; Who talks to whom table. Roles: PM=Project Manager; ID=Industrial Design; UI= User Interface Design; ME=Marketing Expert.

Not all Amida meetings have been annotated with the Dialogue Act Types as in the Ami corpus. [12] Most of them however do have addressees labeled. Contrary to in the Ami procedure and annotation scheme for dialogue acts, where addressees were attributes of dialogue act elements, in this corpus addressees labels are assigned to spurts. These mostly coincide with speaker turns, but when the speaker changes addressee midway a turn to Group or to some (other individual) there is a segment boundary. Only those segments are labeled when the speaker's act is I-addressed. All other acts are supposed to be G-addressed.[13]

12. meeting ED1002C has no DA annotation

13. UNC is used when the annotator was not certain about I-addressing of a segment. It rarely occurs in the data.

	PM	ID	UI	ME	ALL	UNK	Total Ind
PM		11	28	6	104	61	45
ID	8		8		39	52	16
UI	21	15			91	65	36
ME	7				24	25	7
totals	36	26	36	6			

Table 24: AMI Face-to-face Meeting: IS1001c; Who talks to whom table. Roles: PM=Project Manager; ID=Industrial Design; UI= User Interface Design; ME=Marketing Expert.

Table 25 shows for four of the Amida meetings how often speakers I-address others. It includes -between brackets- how often of these the RP is I-addressed. In the row of the remote participants -UI(RP)- the figure between brackets denotes the number of times the RP addresses the Project Manager.

The numbers show that the RP hardly I-addresses others, and when they do they address the PM. RPs are not easily I-addressed by others either, and mostly by the PM.

There is a general rule that says that when A I-addresses B often then B I-addresses A also often. There is an exception for Project Managers: they are often I-addressed but they talk more to the group. More detailed analyses will be carried out to see in what indices of context (activities of group, types of talk) influence this behavior.

	ED1002b	ED1002c	ED1005b	ED1005c
PM	57(20)	34(17)	35(17)	35(18)
ID	13(2)	12(3)	6(1)	22(10)
UI(R)	5(PM=5)	3(PM=3)	6(PM=5)	5(PM=2)
ME	7 24(3)	24(3)	7(5)	34(22)

Table 25: I-addressing in four of the AMIDA remote meetings: Roles: PM=Project Manager; ID=Industrial Design; UI= User Interface Design; ME=Marketing Expert. In meeting ED1002b the PM I-addresses 57 times of which 20 times the remote participant.

7.2 The back channel in remote meetings

The "back-and-forth nature of speaking and listening between two or more interactants" (Goffman [1974], p. 213), is maintained and regulated by non-verbal gestures like head nods, and other small acts. They happen on the "back channel".

> In fact, both the person who has the turn and his partner are simultaneously engaged in both speaking and listening. This is because of the existence of what I call the back channel, over which the person who has the turn receives short messages such as "yes" and "uh-huh" without relinquishing the turn. The partner, of course, is not only listening, but speaking occasionally as he sends the short messages in the back channel. The back channel appears to be very important in providing for monitoring of the quality of communication. ("On Getting a Word in Edgewise", Yngve [1970], p. 568)

We observed that the backchannel between the RP and the meeting room is much less used than in face-to-face meetings. In the AMI meetings 11% of all dialogue acts are back channel acts. Although not all of the four members equally often backchannel, we don't see large differences in their behavior. We noticed that backchannel behavior is elicited by speaker gaze, so partners that are often gazed at by speakers tend to backchannel more than partners that are less gazed at by speakers. As Table 26 shows the number of back channel acts in the 2 remote meetings is much less than in the face-to-face AMI meetings. The RP hardly backchannels, which could be a token of less engagement.

	All	All-RP	BC	BC-RP
ED1005C	969	248	29	2
ED1005B	698	154	26	0
ED1002B	1278	227	131	7

Table 26: Backchannel counts in three Amida Meetings with four participants of which one remote (RP). Columns show All: all dialogue acts; All-RP: the number of dialogue acts by the remote partner; BC: the total number of back channels act in the meeting. BC-RP: the number of back channel acts by the remote partner.

The timing of backchanneling is quite critical. In case of a delayed audio channel a vocalized backchannel act will arrive too late at the site of the main speaker, and this may lead to confusion. Once people have experienced that untimely received backchannel acts easily become a source of trouble they will refrain from using the back channel and remain quiet while someone else is speaking. And this results in the phenomenon that speakers feel they are "talking into a void"; speakers miss the direct feedback signals that tell them that their message is heard and how it is received.

8 Meeting Support Technology

Technologies for remote group collaboration are in common use, and improving all the time. However, current video and telephone conferencing technology is an impoverished medium of communication compared with face-to-face meetings. The AMIDA project builds on the AMI project (IST FP6-506811, http: //www.amiproject.org) to design and implement technologies that recognize and overcome the communication difficulties that groups face.

This section tries to draw some conclusions from our analysis of face-to-face and remote design meetings. It presents some ideas about what improvements remote meeting support technology could bring as well as what the implications of this technology are. We start with some general remarks about the term "mediated" in the context of mediated communication.

The difference between mediated communication and face-to-face communication resembles in a way the difference between a conversation of two peoples that speak the same language and a conversation of two peoples that do not share the same language. Being able to talk in your own language means that the ideas or messages come with their wordings. There is no real distinction between the content of the expression and the form of the expression. Whereas if you have to talk in a strange language you have to "translate" what you want to say into some form in order to bring it as a message understandable for others. (We say: this is a bread and the French call it un "pain") Only if we reflect on language - as linguists do professionally - we distinguish between the words as tokens and their meaning, not if we speak. If we use the words, we don't use them as we use a tool, instead we are immediately present in what they bring about, in the intentional act. As far as the listener or observer of our expression shares the idea that is intended there is communication. But this is never complete, and we can never be sure that there is complete mutual understanding. We can learn a language and so we can learn to use tools to bring about ends. The more we have become used to a tool through practice the more immediate we can handle and bring about what we want. The tool then becomes a part of our body, the immediate observable form of ourselves, that what immediate, whether we want it or not, expresses how we stand, what our mood is. The less familiar we are with the tools we use, the more we have to concentrate on the working with the tool, the less attention we can pay to the immediate effects it brings about: the technology stands in between, it is not a transparent means as the words we speak are transparent means to express meaning. The intentional dialogue act is not something that exists as a complete ideal thing (in a Platonic or mathematical world) and that becomes real in a joint proces between a speaker and a listener (as the relation between a mathematics sequence of numbers and its limit) the content itself is what is changing and gets shape in a unpredicted way through the proces of interaction. That is the interaction creates the act, it doesn't realize the act as if it was already completely specified before. It is only in retrospect that we create the idea of an intentional act as if it was already what is has become in reality.

Mediated communication is communication where you have to do something in order to achieve something else. Immediate communication is communication where no distinction is felt between the act that you are actually involved in and that what you are doing. Thus, the difference is really a difference in stance towards the act that you are involved in. In a technological view in which language can only be seen as a tool used to achieve

communication[14] this distinction between immediate and mediated communication cannot be made. Mediated communication is always based on immediate communication. Where every form of communication has this mediated aspect, sometimes, as in remote meetings and when we have to sue a strange code system it is more apparent and forces us to make a transformation from the way we immediately express what we think, what to do into something that serves as a means to do that.

Face-to-face meetings differ from mediated meetings in a number of ways. In one way or another differences have to do with the fact that participants in a remote meeting are present in a physical space that they do not share immediately through direct contact. Having a live meeting means communication, and that means sharing ideas, participating in joint synchronized activities with others.

In Yankelovich et al. [2007] the authors give a list of problems that people experience with communication in hybrid meetings, i.e. meetings where some people are local and some are remote.

Audio problems.
- Poor quality speaker phones
- Too much background noise
- Multiple speaker speaking at the same time can be difficult to understand
- People speaking too far from microphones

Remote attendee problems
- Inability to conduct side conversations.
- In-room attendees forget about remote people
- Challenging to brake into lively conversation
- Difficult to detect in-room speaker changes
- Hard to identify people currently in the meeting room
- Hard to identify the current speaker
- Difficult to participate in brain-storming sessions
- Cannot see in-room demonstrations or artifacts

Conference room problems
- Local people more emotionally salient than remote participants.
- Easy to forget about remote participants
- Often do not know who is still connected

We elaborate on some of these problems in remote meetings.

The impact of delay on communication. Since communication in what ever form rests on the physical contact layer, the most critical requirement for mediated communication is the capacity and speed of the audio and visual channels. Audio delay has great impact on the flow of conversation. In Ruhleder and Jordan [1999] and in O'Conaill et al. [1993] the impact of audio and video delay on distributed communication is studied. Delay generated by communication technology affects trust and confidence between communicators because it disturbs natural turn taking process, a key element in social interaction (Duncan [1972]). The authors argue that in face to face communication people can signal problems in communication that can not easily be repaired with in remote conversations because the source of the problem is not easy to identify.

Transmission delay may cause:

14. recall the book title "How to do things with words", Austin [1962]

- Unnecessary rephrasings, because when the speaker stops and expects that someone will take turn, it takes some time before she receives an answer, so she thinks it was not clear.
- Misapplied feedback. Listeners give feedback to the speaker to make clear that he can go on, that he is understood, or that they agree with what the speaker says, without intending to take the floor. Time delay may cause that these signals are received at times they are not expected by the speaker, so that they are misinterpreted and hence may cause disruption of the interaction.

Interactions with more formal turn-taking rules such as formal meetings are less likely to suffer from the delay caused by the mediating technology. In general people show a great capacity to adapt their behavior to the affordances of the mediating technology. It is therefor important to realize that people need to get familiar with the technology. It is also important that participants know what the partners they are interacting with can hear and can see, to avoid the problem that Ruhleder et al. identified with the identification and agreement about the cause of the communication problem. Therefore it makes a difference if partners share the same technology and know how information is received at the other side of the channel. This may be also the reason why hybrid meetings are more problematic than meetings that are symmetric, and where every participant is remote (cf. Yankelovich et al. [2007]).

It is important that participants in a conversation know who is talking and know if they are being referred to or addressed by the speaker. Our analyses are in line with the earlier findings: it shows that in remote meetings partners refrain from using verbal backchannel signal, which makes that speakers often experience that they talk in a void. Visual contact may help here. This is one of the conclusion in Donath [2001], a study that investigates the various functions of the face in communication and the pros and cons of mediating the face, a real picture, a video or some make up in remote conversations. Addressing in remote meetings is more explicit, since speakers are often not convinced that the remote partner is paying attention. We also see more structured meetings, where the project manager has a leading role in announcing how the meeting or a discussion is organized and who will talk next (see the transcript of the remote meetings in the appendices of this report)[15].

The visual channel is important when people discuss objects, or documents. Moreover, it helps to identify who is speaking and to signal focus of attention of the speaker, which helps understanding verbal referring expressions. Remote meetings have the advantage that it is easier to do time-sharing, and only pay attention if something interesting happens. But this requires technology that on the one hand helps remote participants to point at interesting parts, on the other hand helps to inform participants the status of attentiveness in the meeting.

15. Automatic identification and analyses of these type of organizing, reflexive, utterances, could help segmentation, as well as addressing recognition

9 Future Work: WP1 research in year 2009

In the last year of Amida (2009) WP1 will continue the research on the following topics:

1. Physicality in scenario-based remote and natural design meetings (UT).
2. Deictic references and framing in face-to-face and remote design meetings.
3. Communication failures in remote meetings.

The results of this research will be reported in the final deliverable of WP1. Besides that we will work on journal publications about:

- Eliciting acts: matching questions and answers in conversations (UT).
- Voting classifiers: exploiting annotated corpora that show systematic contextual dependencies in the agreement between annotators. (UT)
- Subjectivity: agreement and disagreement in conversations. (UEDIN)

9.1 Physicality

We will analyze the naturalistic design meetings captured during year 2007-08. A main focus of this analysis would be to understand the role of material artefacts in supporting communication and collaboration amongst the designers. A detailed taxonomy of these artefacts will be developed. This work will be used to inform WP2 and WP6 in order to develop an application to support remote communication between designers. The application will be about supporting mediated awareness amongst distant designers. The analysis in WP1 will be focused on how artefacts help in mediating awareness.

9.2 Deictic References and Framing

We plan to do a qualitative analysis of the argument that conversational language technology can benefit from a closer consideration of participant deixis within an activity-oriented view of language. We will investigate how the words ”I”, ”you”, and ”we” play a role in the expression and organization of important meanings in spoken and written language. We hope to show that its use reflects how participants frame their individual and group identities, social relationships, and their participation in the conversational activity. The principal goal will be to show how technologies like discourse summarization, segmentation, structuring, and indexing can be improved by the use of such knowledge. Concentrating on the above forms (first- and second-person pronouns) as well as verb tense and aspect, modality, and subjective and meta-discourse, we will outline a simple analytical framework for exploring how meaning and action are organized and expressed across different conversational genres. Using the general framework, we then plan to show how it can be employed specifically for small group organizational meetings. We will explore the role of deixis in summarizing/characterizing whole conversations and episodes, and its pragmatic meaning within individual speech acts and clauses, and in phrases and words.

9.3 Communication Failures in Remote Meetings

We will continue our analysis of remote meetings with a special focus on how remote partners jointly coordinate and manage their activities in addressing, and turn-taking. We

will look how failures in addressing and turn-taking are related to the audible and visual affordances of the communication technology. We hope that we can come to improved insights in the constraints that communication delay has on the possibility that partners in remote conversations share visual and audio space in synchrony. The research may lead to a number of detailed requirements for technology that supports engagement and control in remote meetings.

References

Jean C. Carletta. Unleashing the killer corpus: experiences in creating the multi-everything AMI meeting corpus. Language Resources and Evaluation, 41(2):181–190, May 2007. ISSN 1574-020X. doi: 10.1007/s10579-007-9040-x.

W.M. Post, A.H. Cremers, and O.B. Henkemans. A research environment for meeting behavior. In A. Nijholt, T. Nishida, R. Fruchter, and D. Rosenberg, editors, Social Intelligence Design, Enschede, The Netherlands, 2004.

H.H. Clark. Using language. Cambridge University Press, The Edinburgh Building, Cambridge CB2 2RU, UK, 1996.

Janet E. Cahn and Susan E. Brennan. A psychological model of grounding and repair in dialog. In Proceedings, AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems, pages 25–33. American Association for Artificial Intelligence., 1999.

Emanuel A. Schegloff. Discourse as interactional achievement: Some uses of "uh huh" and other things that come between sentences. In D. Tannen, editor, Analyzing discourse, text, and talk, pages 71–93. Georgetown University Press, Washington, DC, 1982.

D. R. Traum and P. A. Heeman. Utterance units and grounding in spoken dialogue. In Proc. ICSLP'96, volume 3, pages 1884–1887, Philadelphia, PA, 1996.

David R. Traum. A computational theory of grounding in natural language conversation, 1994.

T. Paek and E. Horvitz. Uncertainty, utility, and misunderstanding: A decision-theoretic perspective on grounding in conversational systems. In AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems, November 5-7 1999.

Erving Goffman. Footing. In Forms of Talk, pages 124–159. Philadelphia: University of Pennsylvania Press, 1981.

K. Krippendorff. Content analysis: An Introduction to Its Methodology. Thousand Oaks, CA: Sage, 2nd edition, 2004.

J. Carletta, A. Isard, S. Isard, J.C. Kowtko, G. Doherty-Sneddon, and A.H. Anderson. The reliability of a dialogue structure coding scheme. Computational Linguistics, 23 (1):13–31, 1997a.

AMI Consortium. Guidelines for dialogue act and addressee. Technical report, AMI Consortium, 2005.

H. H. Clark and F. E. Schaefer. Dealing with overhearers. In Arenas of language use. Chicago: University of Chicago Press, 1992.

Dennis Reidsma and Jean C. Carletta. Reliability measurement without limits. Computational Linguistics, 34(3):319–326, September 2008. ISSN 0891-2017. doi: 10.1162/coli.2008.34.3.319.

Dennis Reidsma. Annotations and Subjective Machines - Of Annotators, Embodied Agents, Users, and Other Humans. Center for Telematics and Information Technology (CTIT), Enschede, the Netherlands, 2008.

Dennis Reidsma, Dirk Heylen, and H. J. A. op den Akker. On the contextual analysis of agreement scores. In Jean-Claude Martin, Patrizia Paggio, Michael Kipp, and Dirk Heylen, editors, Proceedings of the LREC Workshop on Multimodal Corpora, pages 52–55. ELRA, ELRA, May 2008. ISBN 2-9517408-4-0.

Dennis Reidsma and H.J.A. op den Akker. Exploiting 'subjective' annotations. In Ron Artstein, Gemma Boleda, Frank Keller, and Sabine Schulte im Walde, editors, Proceedings of the Coling Workshop on Human Judgments in Computational Linguistics, August 2008. to appear.

T. Wilson. Annotating subjective content in meetings. In Proceedings of LREC, 2008.

John J. Godfrey, Edward Holliman, and J. McDaniel. SWITCHBOARD: Telephone speech corpus for research and development. In Proceedings of IEEE ICASSP-92, pages 517–520, San Francisco, CA, 1992.

I. McCowan, J. Carletta, W. Kraaij, S. Ashby, S. Bourban, M. Flynn, M. Guillemot, T. Hain, J. Kadlec, V. Karaiskos, M. Kronenthal, G. Lathoud, M. Lincoln, A. Lisowska, W. Post, D. Reidsma, and P. Wellner. The AMI Meeting Corpus. In Proceedings of Measuring Behavior 2005, the 5th International Conference on Methods and Techniques in Behavioral Research, Wageningen, Netherlands, 2005.

Lou Burnard. Reference Guide for the British National Corpus (World Edition), 2007.

Thorsten Brants and Alex Franz. Web 1T 5-gram Version 1. Linguistic Data Consortium, 2006.

Alexander Gruenstein, John Niekrasz, and Matthew Purver. Meeting structure annotation: data and tools. In Proceedings of the 6th SIGdial Workshop on Discourse and Dialogue, Lisbon, Portugal, September 2005.

Matthew Purver, Patrick Ehlen, and John Niekrasz. Detecting action items in multi-party meetings: Annotation and initial experiments. In Steve Renals, Samy Bengio, and Jonathan Fiscus, editors, Machine Learning for Multimodal Interaction: Third International Workshop, MLMI 2006, Bethesda, M volume 4299 of Lecture Notes in Computer Science, pages 200–211. Springer, 2006a.

Matthew Purver, Patrick Ehlen, and John Niekrasz. Shallow discourse structure for action item detection. In Proceedings of the HLT-NAACL Workshop: Analyzing Conversations in Text and Speech (ACTS), pages 31–34, 2006b.

Matthew Purver, John Dowding, John Niekrasz, Patrick Ehlen, Sharareh Noorbaloochi, and Stanley Peters. Detecting and summarizing action items in multi-party dialogue. In Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue, Antwerp, Belgium, September 2007.

Jean Carletta, Stephen Isard, Anne H. Anderson, Gwyneth Doherty-Sneddon, Amy Isard, and Jacqueline C. Kowtko. The reliability of a dialogue structure coding scheme. Computational Linguistics, 23(1):13–31, 1997b.

Surabhi Gupta, John Niekrasz, Matthew Purver, and Daniel Jurafsky. Resolving "you" in multi-party dialog. In Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue, Antwerp, Belgium, September 2007.

J. Searle. Speech Acts: An Essay in the Philosophy of Language. Cambridge: Cambridge University Press, 1969.

Emanuel A. Schegloff and Harvey Sacks. Opening up closings. Semiotica, 7(4):289–327, 1973.

N. Jovanovic. To Whom It May Concern: Addressee Identification in Face-to-Face Meetings. Center for Telematics and Information Technology (CTIT), Enschede, the Netherlands, 2007.

David R. Gibson. Participation shifts: Order and differentiation in group conversation. Social forces, 81(4):1335–1381, 2003.

H. Sacks, E.A. Schegloff, and G. Jefferson. A simplest systematics for the organization of turn-taking for conversation. Language, 50:696–735, 1974.

Daniel C. O'Connell, Sabine Kowal, and Erika Kaltenbacher. Turn-taking: A critical analysis of the research tradition. Journal of Psycholinguistic Research, 19, No. 6: 345–373, 1990.

S Wortham. Deictic mapping. Journal of Pragmatics, 25:331–348, 1996.

R. op den Akker and M. Theune. How do i address you? - modelling addressing behavior based on an analysis of a multi-modal corpus of conversational discourse. In Proceedings of the AISB 2008 Symposium on Multimodal Output Generation (MOG 2008), Aberdeen, UK, pages 10–17, April 2008.

Gene H. Lerner. Selecting next speaker: The context-sensitive operation of a context-free organization. Language in Society, 32:177–201, 2003.

Gene H. Lerner. On the place of linguistic resources in the organization of talk-in interaction: "Second person" reference in multi-party conversation. Pragmatics, 6(3): 281–294, 1996.

Emanuel A. Schegloff. Some practices for referring to persons in talk-in-interaction: A partial sketch of a systematics. In B. Fox, editor, Studies in Anaphora, pages 437–485. Amsterdam: John Benjamins, 1996.

C. Heath and P. Luff. Collaboration and control: Crisis management and multimedia technology in london underground line control rooms. Computer Supported Cooperative Work, 1(1):24–48, 1992.

T. Robertson. Cooperative work and lived cognition: A taxonomy of embodied actions. In Proceedings of ECSCW'97, pages 205–220. Kluwer Academic Publishers, 1997.

F. van der Veeken. Effects of using an icat in remote meetings. In Twente Student Conference on Information Technology, 2008.

F.L. de Vries. The effect of a philips icat on the perception of copresence in remote meetings. In Twente Student Conference on Information Technology, 2008.

UTwente. Deliverable d1.3. Technical report, University of Twente, 2008.

E. Goffman. Frame Analysis: An Essay on the Organization of Experience. New York: Harper and Row, 1974.

V.H. Yngve. On getting a word in edgewise. In Papers from the sixth regional meeting of the Chicago Linguistic Society, pages 567–77, Chicago: Chicago Linguistic Society, 1970.

J. L. Austin. How to do Things with Words. Oxford University Press, 1962.

Nicole Yankelovich, Jonathan Kaplan, Nigel Simpson, and Joe Provino. Porta-person: telepresence for the connected meeting room. In Proceedings of CHI 2007, pages 2789–2794, 2007.

Karen Ruhleder and Brigitte Jordan. Meaning-making across remote sites: how delays in transmission affect interaction. In Proceedings of the sixth European Conference on Computer-Supported Cooperative Work, pages 411–429, 1999.

B. O'Conaill, S. Whittaker, and S. Wilbur. Conversations over videoconferences: an evaluation of the spoken aspects of video-mediated communication. Human Computer Interaction, 8:389–428, 1993.

S.D. Duncan. Some signals and rules for taking speaking turns in conversations. Journal of Personality and Social Psychology, 23:283–92, 1972.

Judith S. Donath. Mediated faces. In CT '01: Proceedings of the 4th International

Conference on Cognitive Technology, pages 373–390, London, UK, 2001. Springer-Verlag. ISBN 3-540-42406-7.

Appendix A

This appendix contains the transcripts of two Amida meetings. In the main text of this deliverable we refer to these transcripts.

`C > G(BEP): Hello .`

Means speaker is C who addresses G (means the Group); the dialogue act type is BEP (=Be-Positive) and the text of the utterance is: “Hello” In both meetings A has the role of Project Manager and C is the Remote Partner (the user interface designer)

The shortnames of the dialogue act types used in these listings are the following:

BEP - Be-Positive
BEN - Be-Negative
ELI - Elicit-Inform
ELO - Elicit-Offer-or-Suggest
ELA - Elicit-Assess
CAU - Comment-about-Understanding
STL - Stall
BCK - Backchannel
INF - Inform
SUG - Suggest
OFF - Offer
FRG - Fragment
OTH - Other

ED1005b

C > G(BEP): Hello .
C > G(ELI): Can you hear me ?
A > C(ELI): Lawrence , can you hear us ?
D > G(CAU): What does he say ?
C > A(INF): Yeah I can hear you ,
C > A(ELC): can you ?
A > C(CAU): Yeah I can hear you .
D > C(ELI): Can you hear me ?
C > D(SUG): One moment ,
C > D(INF): let me try to make it full screen .
C > G(STL): Um
A > C(INF): If you right-click then there should be a full screen option .
D > G(FRG): Who is he able to ?
C > G(INF): I did it , I did it .
A > C(BEP): Well done .
C > G(ELI): Can you see me ?
D > G(ELI): Is he seeing all of us or
D > G(CAU): Ah . Okay .
A > G(FRG):
C > G(SUG): If you could speak a little bit louder it would be very very helpful .
A > C(ASS): Okay .
D > A(SUG): Can you turn it up ?
D > C(SUG): Lawrence , say something .
A > G(ELA): How's that ?
C > G(OTH): What ?
C > G(STL): Okay ,
C > G(INF): a new email has arrived .
C > G(ELI): How can I see myself in it ?
A > C(OFF): We'll wait 'til you're on your own again .
A > C(ELI): Hey Lawrence , are you ready ?
C > A(INF): Yes I am ready .
A > C(ASS): Alright .
A > G(STL): Okay , so
A > G(INF): this is our second meeting .
A > G(INF): Just start brief opening , just we'll go over again what we deci talked about before .
A > G(INF): And then we're gonna move into three presentations that you've already submitted ,
A > G(INF): then we'll make our decisions
A > G(INF): and we'll close .
A > G(INF): We only have forty minutes
A > G(SUG): so we'll try and get through this this as quickly as we can .
A > G(STL): Um
A > G(INF): already been decided , just a quick reminder , we're gonna have a p production cost of no more than twelve point five euros ,
A > G(INF): I'm sure we're all well aware of this by now .
A > G(INF): We're aiming this remote at under forties .
A > G(INF): And it's already been decided by the previous um design team that infrared is the best way to get data to the receiver ,
A > G(INF): so we're gonna move forward on the assu assumption that we're using infrared .
A > G(ELC): Alright ?
A > B(SUG): Pass over to the industrial designer .
B > A(ASS): Okay ,
B > G(STL): uh
B > G(INF): just load up my V.N.C. .
B > G(STL): So um
B > G(INF): the method that I'm using in this presentation is to uh consider the requirements of the project ,
B > G(INF): this is pretty simple ,
B > G(STL): um
B > G(INF): we have to get something that's going to be innovative if we want to uh breach the market ,
B > G(INF): I mean I don't think it's often a thought that runs through people's heads that oh I need to get a new remote control .
B > G(INF): So if we want to have reasonable sales we need to have a project that's innovative and it's gonna ha uh get its name out .
B > G(STL): Um
B > G(INF): first I came up with an ideal world solution that is you know what I uh what would be ideal if all resources were available and cost was not an issue .
B > G(INF): And just really brainstorming ,
B > G(STL): so I mean
B > G(INF): the features that we discussed in the previous meeting , uh l such as a touch screen L.C.D. display ,
B > G(STL): um
B > G(INF): and a backlight so it's usable in the dark ,
B > G(FRG): some sort of location
B > G(STL): um
B > G(BEP): sorry ,
B > G(INF): some sort of method for locating it
B > G(INF): so that we can find a lost remote ,
B > G(INF): and uh all these sorts of things incorporated into a fashionable design .
B > G(STL): So um
B > G(INF): something that's appealing to the eye but also comfortable to hold ,
B > G(STL): uh
B > G(INF): quality materials so it's not gonna break ,
B > G(INF): all this um came into my mind .
B > G(STL): Um
B > G(INF): then I had to look at the available hardware um , what we could use and what is being provided to us ,
B > G(INF): and then modify the ideal world solution with uh respect to what is available to us and also our costs .
B > G(STL): Um
B > G(INF): we have yeah , various options t available to us in terms of what we can use ,
B > G(STL): um
B > G(INF): there are several different energy sources available to us
B > G(INF): there's a basic battery ,
B > G(STL): um
B > G(INF): but then there's also the more innovative uh solutions of having something like a a kinetic battery ,
B > G(INF): the sort of thing that you have in wrist watches um that'll charge upon movement .
B > G(STL): Um
B > G(ASS): whether or not this will be ideal for a remote , not really sure .
B > G(STL): Um
B > G(INF): they they're not like a watch ,
B > G(INF): they're not on your wrist at all times ,

B > G(INF): so maybe they're not gonna get as much movement as they require to keep up a reasonable charge .
B > G(STL): Um
B > G(INF): the other options were solar cells ,
B > G(ASS): which I thought were entirely unsuitable
B > G(INF): because almost all the uh usage is going to be indoors
B > G(INF): and uh indoor lighting isn't going to provide enough enough energy .
B > G(STL): Um
B > G(INF): or a hand dynamo ,
B > G(INF): but I don't think anyone's gonna particularly want to wind up their remote control before using it .
B > G(SUG): So I think either the basic battery or the kinetic battery is what we should settle upon .
B > G(STL): Um
B > G(INF): next would be the case design ,
B > G(STL): um
B > G(INF): there is the general case which is a flat uncurved case , uh or single or even double curved cases .
B > G(STL): Um
B > G(SUG): I think you know we need to go at least for a single curved if not double curved case ,
B > G(INF): but which design we uh decide upon has impacts on what material we can use .
B > G(STL): Um m
B > G(INF): materials we can use are plastic ,
B > G(STL): um
B > G(ASS): which I think might come across as a little cheap ,
B > G(STL): uh
B > G(INF): we could have rubber , such as in anti-R.S.I. stress balls , uh wood and titanium .
B > G(ASS): I think the uh wooden design probably isn't suitable for our uh target audience of the under-forties ,
B > G(STL): um so yeah
B > G(SUG): my preference in that would be the titanium .
B > G(STL): Uh
B > G(INF): though the titanium cases can't be used with the double curved designs .
B > G(STL): Um
B > G(INF): there are L.C.D. displays available for the interface ,
B > G(STL): um
B > G(SUG): which I think is without doubt the uh way we want to go if we want to have a truly innovative design .
B > G(STL): Um
B > G(FRG): unless
B > G(INF): if we're going to choose the L.C.D. design then a uh we need to use an advanced chip .
B > G(INF): This is gonna be more expensive than the other chips ,
B > G(SUG): but uh I think it's necessary for our product .
B > G(STL): Um
B > G(INF): there is the option of a uh microphone and speaker ,
B > G(ASS): but at least in uh my opinion I think that voice recognition would be more of a gimmick than a geniune function ,
B > G(INF): and um a unless you had a very sensitive microphone then uh it wouldn't be much use ,
B > G(ASS): and if you do have a very sensitive microphone , then why not put it in the television rather than a remote .
B > G(FRG): I don't think there's any
B > G(STL): uh
B > G(ASS): it doesn't make much sense to include that in our product .
B > G(STL): Um so
B > G(INF): this is pretty much what I've covered .
B > G(STL): Um
B > G(FRG): the
B > G(INF): looking at the other products from the company , it seems that a futuristic design um seems to be the modus operandi that's what um the company is known for .
B > G(STL): Um
B > G(INF): and so this just outlines what I've just said , my the uh choices ,
B > G(STL): and that uh
B > G(INF): this is just a brief schematic of um how the how the remote would work , with the chip being central to the operation , um and in control of everything other than the power supply .
B > G(STL): Um
B > G(INF): that is the end of the slideshow .
B > G(INF): I'll pass back to our project manager .
A > G(STL): Alright .
A > G(STL): So
A > C(SUG): we're now going on to your section Lawrence , if you wanna take control here .
C > A(ELO): How much time do I have ?
A > G(STL): Um
A > C(SUG): if you could try and keep it about nine or ten minutes at most .
C > A(BEP): Sorry ,
C > A(CAU): how much ?
A > C(INF): About nine or ten minutes at most .
C > A(CAU): Nine or ten minutes , okay .
C > G(STL): So
C > G(INF): my interface concept would um would make us do some some quite of drastic decisions I might say .
C > G(FRG): For example
C > G(STL): um
C > G(OFF): let's consider a little bit the methods that I've used .
A > G(ASS): Oops .
C > G(STL): Um
C > G(INF): I do have something on my screen .
A > G(BEP): Sorry
A > G(BEN): that's my fault .
C > A(SUG): Could you go back ?
C > A(ASS): Okay ,
C > G(STL): so
C > G(INF): methods . What I have done and what we should maybe de uh dig some more in it .
C > G(SUG): Survey the market
C > G(SUG): and decide upon wh what we want to do , something traditional or something new .
C > G(SUG): Decide upon the target public , or how to enlarge it .
C > G(INF): I know that we already have a decision on what kind of public we are going to address our our product to ,
C > G(ASS): but then again I think it will be kind of a very good idea to try to find ways , let's say gimmicks uh to enlarge to enlarge our target audience ,
C > G(INF): because this is going to give us uh bigger revenues
C > G(INF): and after all this is what we aim for .
C > G(STL): Uh
C > G(SUG): consider the advantages and disadvantages that uh one or the other um products that we're going to build is going to to to have .
C > G(SUG): For example I think that we should take into consideration the marketing point of view , the manufacturing point of view , and the economic one .

184 C > G(ASS): For example it will it would give us no uh no joy whatsoever if we would come up with something that would be very hard to manufacture , or that um it wouldn't be very very appealing for the for our audience .
185 C > G(INF): And the best way of finding all this kind of information , which is quite valuable I might say , is from getting feedback from a diverse array of users .
186 C > G(SUG): So this is where I would like uh to emphasize the fact that we need to do extensive testing with it .
187 C > G(OFF): Now , findings , what did I found out .
188 C > G(INF): The idea that novelty give us a headstart ,
189 C > G(INF): I received an email from from one of our departments which says something like we that we are able to incorporate in our future products a voice recognition mechanism ,
190 C > G(SUG): and I think that if we are able to do it in our and still remain in our budget , it would be a good idea to do that
191 C > G(INF): because then again we would have a head start in front of our competition .
192 C > G(INF): Now our public as we've decided until now is under forty and technology curious I might add , plus all the others that we can we can try to mm to attract with our new product .
193 C > G(SUG): What should we do with the product in order to be attractive for all the others , I think we should we should get something quite unique and quite attractive , attractive from different point of view points of view , but I would mainly emphasize from the economical and uh functional point of view .
194 C > G(FRG): And then again , least but not
195 C > G(STL): um yeah I uh
196 C > G(ASS): I wouldn't like to put it on the last place but still
197 C > G(STL): uh
198 C > G(SUG): we have to take into consideration the competition but not so much .
199 C > G(SUG): If we're going to to concentrate on whatever we have to I think that that is going to be enough for for us getting a good produ project .
200 C > G(SUG): Another very important finding is um that we should try to avoid misleading signs .
201 C > G(INF): As you can see on my on my picture there , you have something like a remote control which has two uh volume buttons ,
202 C > G(FRG): and of course when you'd like to
203 C > G(INF): let's say you're a teenager
204 C > G(INF): and you're you do want to get the T.V. at a really really low volume ,
205 C > G(INF): 'cause your parents are sleeping
206 C > G(INF): and you kind of press the button that you that you think would do that but actually that button would would make it louder .
207 C > G(ELI): Why it's so ,
208 C > G(INF): because if you see there there are two V.s , and both of them gives you the impression that you're you have an arrow down rather than having a volume up button .
209 C > G(ELO): So how are we going to avoid this kind of misleading signs or whatsoever ,
210 C > G(SUG): I think we should make heavy use of colours and L.E.D.s .
211 C > G(OFF): Now I'm going to present you some some of uh the products which are already available on the market ,
212 C > G(INF): in order for you to make to make an idea of what we can do or how we can differentiate us from them .
213 C > G(INF): The latest model would be something uh which resembles a lot like an iPod ,
214 C > G(INF): and it has the L.C.D. screen that we've decided that we should incorporate .
215 C > G(STL): Now
216 C > G(INF): on the other hand we all agreed with the fact that we want something to be uh extremely user friendly ,
217 C > G(INF): and that one would be the remote control for elderly people ,
218 C > G(INF): which has um let's let's call it a jumbo jet remote control ,
219 C > G(INF): because it's a little bit of a big big size , bigger than normal anyway .
220 C > G(STL): Now
221 C > G(SUG): then again if we have such a narrow target as we have , um we have to differentiate between what do we actually want want to do .
222 C > G(INF): A complete beginner's remote control is the one on the left ,
223 C > G(INF): I think it's your left ,
224 C > G(INF): and then advanced , yours is the one on the right .
225 C > G(STL): Now
226 C > G(SUG): I think that we should be somewhere in between ,
227 C > G(SUG): that's why I suggest that let's say if indeed we're going to have the L.C.D. screen , we should have something like a pointer or a function that is going to to uh to acc let you permit you access to higher-order functions , rather than already displaying them .
228 C > G(SUG): So I think that for when you're going to just to take it in your hand you should have just the very basics on it .
229 C > G(INF): Personal preferences ,
230 C > G(SUG): the general public personal preferences should be taken into account a lot ,
231 C > G(SUG): and I think that we should try to find ways of quantifying those preferences and take advantage of them , and incorporate them in our products ,
232 C > G(STL): that why
233 C > G(SUG): that's why I would like to suggest us to use uh let's say different colours ,
234 C > G(INF): maybe females will would like a pink remote control rather than a black one ,
235 C > G(STL): uh let's say uh
236 C > G(SUG): we could do uh vary a little bit in in in sizes ,
237 C > G(INF): because elderly people probably are going to uh be more biased towards a bigger remote control rather than a really thin and slim one as the youngers would .
238 C > G(STL): Now , uh
239 C > G(ASS): another thing uh very important I could say is thinking at all that might come up . Uh that is , different situations ,
240 C > G(INF): for example you would probably like if you would be uh quite young , in your young let's say twenties , you'd you'd uh rather prefer to have a remote control that would that's going to be advanced user , rather than having something um a complete beginner ,
241 C > G(INF): but then again you have to think of the idea that the situation is going to change at some point
242 C > G(INF): and you're going to need that that kind of a remote control let's say for your kids or fo for whoever else you want to restrict access to your to your services .
243 C > G(SUG): And of course we have to come up with interesting ways of suggesting that .
244 C > G(SUG): About my general preferences , I would say uh that uh I would really like to have basic functions ,
245 C > G(SUG): which would be like blindly reachable ,
246 C > G(INF): that means that I don't have to do any kind of effort
247 C > G(INF): and they could be they could be guessed from let's say ten miles away .
248 C > G(STL): Now ,
249 C > G(INF): special situations require special functions ,
250 C > G(INF): so let's say for example that I would like to record a a show that I I I am going to miss ,
251 C > G(INF): and of course I would like to do that in the easiest way possible ,
252 C > G(INF): but then again um that would that would mean that um we need to add lots and lots of new interesting high-tech functions ,
253 C > G(SUG): which as I said before , should be reachable in a little bit in a more complicated way .
254 C > G(STL): Now
255 C > G(SUG): anothing things that should be taken into consideration is the idea of having a cool look ,
256 C > G(SUG): using also materials , let's say with interesting uh interesting aspects ,
257 C > G(SUG): and of course to produce a long-lasting project product which should be adaptable in different situations and for different users .
258 C > G(INF): This concludes up my presentation ,
259 C > G(SUG): for any other kind of information to look in the shared folder for my word document .
260 C > G(INF): That gives you more information about what I thought until now when with what I could come up with .
261 A > C(BEP): Thank you .
262 A > G(STL): Alright .
263 A > D(SUG): Now pass over .
264 D > G(STL): Yeah , so

D > G(FRG): What did
D > G(STL): yeah ,
D > G(INF): the methods I used ,
D > G(STL): um
D > G(INF): we had an observa uh lab experiment ,
D > G(INF): and observed the opinion of one hundred people ,
D > G(INF): and they filled in a questionnaire as well .
D > G(INF): And the findings ,
D > G(STL): uh
D > G(INF): the most important thing for them is fancy look and feel , both .
D > G(STL): So uh
D > G(INF): seventy percent of them uh find most remote controls ugly .
D > G(STL): So um
D > G(SUG): ours shouldn't be .
D > G(STL): Uh
D > G(INF): eighty percent of them are willing to spend more on a remote control if it has a fancy look .
D > G(STL): So
D > G(SUG): if um making the remote control more fancy costs us more money , we we can afford some of this ,
D > G(ASS): if it's not too much of course .
D > G(STL): Uh
D > G(INF): the theme that um is going to be for the next year for uh clothes , shoes and furniture , is fruit and vegetables ,
D > G(SUG): and maybe we could implement this in our remote control as well .
D > G(INF): And the feel of material is supposed to be spongy , sponge for the next year .
D > G(ASS): So maybe not s titanic
D > B(SUG): you can take care of this .
D > G(STL): Um
D > G(INF): the second most important thing for them is technological innovative ,
D > G(STL): um
D > G(INF): people between um sixteen and thirty five years prefer L.C.D.
D > G(INF): and uh above thirty five years old prefer speech recognition .
D > G(SUG): So as we are going to to people under forty , we should uh go for um L.C.D. ,
D > G(INF): and according to my research twenty six percent of the people think that remote controls are bad for R.S.I. ,
D > G(SUG): so we should come up with a shape that um is comfortable and not causing any harm .
D > G(INF): And third uh important thing c uh according to the people is to be easy to use ,
D > G(INF): because thirty four percent of them think that it's uh it cause them too much time to learn how to use it .
A > G(FRG):
B > A(BEP): Oh sorry .
D > G(FRG): What's
A > D(SUG): You're good .
D > G(INF): And fifty percent um told us that uh most of the time the remote control is lost somewhere in the room ,
D > G(STL): so uh we could um um
A > D(SUG): Incorporate the feature to find the remote control .
D > A(ASS): Yeah that's right .
C > A(CAU): Can you repeat that ?
D > G(INF): And then uh fifty percent of them told us that ten percent of the buttons aren't used ,
D > G(SUG): so we should be careful about the functionality .
D > G(STL): And um
D > G(INF): my personal preferences ,
D > G(FRG): as you can see maybe
D > G(STL): uh w
D > G(SUG): we should go for the L.C.D. one
D > G(INF): as the public we are referring to prefer it more ,
D > G(INF): and then um as they said , they had too many buttons they don't use .
D > G(SUG): So we can put the most usable buttons on one layer ,
D > G(SUG): and those we don't use so much on the second layer ,
D > G(INF): and when they want something more they could just switch them off .
D > G(SUG): And to be careful with the ha uh the shape of the remote control as well .
D > G(SUG): And uh the labels should not go off when using it for a while .
D > G(SUG): So maybe if we are using colours , we can have something to to lit , the number to be in light , not to be written on the button ,
B > G(BCK): Yeah .
D > G(INF): and of course to be able to find it in the dark , this would help as well .
D > G(STL): So
D > G(INF): this is my contribution to
A > D(BEP): Alright .
A > G(INF): So just quickly before we go into the decision stage , I thought I'd add this in ,
A > G(FRG): so we could look at
A > G(INF): these are some of the points we came up with at the last meeting ,
A > G(ELA): just to see if they're still relevant
A > G(ELA): or should we still be discussing these ,
A > G(ELA): and if so what decisions should we be reaching .
A > G(STL): So
A > G(INF): interfaces , I think we've really covered it I think ,
A > G(INF): all three of them sort of came out and said the interface we want is an L.C.D. ,
D > A(INF): L.C.D. .
A > G(ELA): and I also think what came out I mean feel free to disagree with me here , is that what we should be going for is a basic interface first and foremost , as ma as little controls as possible . And then maybe an option somewhere in the remote control to switch it to an advanced view of the controls .
A > G(SUG): So sort of when they first pick it up , there would be just the controls needed to use it ,
A > G(SUG): and then the could be a button somewhere on the remote control which would be we could do it with an L.C.D. to then switch it so more advanced features appear .
A > G(ELA): I mean is that sort of
B > A(ASS): Yeah well I I agree with that , I think that's a good way to approach it .
D > A(ASS): Yeah the was my idea for the two la layers .
A > C(ELA): Lawrence , you okay with that ?
C > A(ASS): Yes .
A > G(STL): Okay
A > G(INF): and I think that covers the second one as well ,
A > G(INF): and um I think the marketing covered what is really needed .
A > G(STL): So
A > G(SUG): we'll move on to
A > G(INF): these are the decisions we have to make just now ,
A > G(INF): we need to have these done and dusted by the end of this meeting .
A > G(STL): So
A > G(ELO): looking at um the components concept , I think the on the only one really there that we didn't seem to come to was on the how we power it , uh if we're gonna go for the kinetic or a standard battery .

C > G(SUG): I'll go for the lithium-ion one .
C > G(ASS): I think it would be a good one , a good idea .
D > G(STL): Okay
D > A(ELI): which one is uh long lasting ?
A > D(INF): But we wouldn't really know until we began to produce it and have it tested .
A > D(INF): The kinetic could be long-lasting in that if it moves around a lot ,
A > D(INF): but then the battery would be long-lasting if the remote did not move very much .
D > A(ASS): But then it loses its point .
D > A(INF): It should be removed and stay in one place , so
A > G(STL): But I mean
A > G(SUG): personally I think for some people the kinetic is gonna be better ,
A > G(INF): I think they will move around enough ,
A > G(INF): but there will be a lot of people who will pick up the remote , change the channel , put it back down again ,
A > G(FRG): which I don't think will provide enough
A > G(ASS): I would not want to take the risk of it providing enough power .
B > G(STL): Yeah I think um I mean
B > G(INF): a watch if which is where this technology is already implemented , I mean if it's on your wrist all day you move a lot ,
B > G(INF): and it's gonna get a lot of energy from that ,
B > G(INF): and a watch is gonna require less power than uh an L.C.D. remote ,
B > G(INF): I mean it's not gonna have to send out any infrared signals
B > G(INF): and the screen is probably gonna be significantly smaller ,
B > G(STL): so um
B > G(SUG): given that I think we might as well stick to the uh basic battery if it's gonna reduce production costs .
A > G(STL): Right ,
A > G(ELA): everybody okay with that ?
C > G(OTH): Oh .
D > A(ASS): Yeah .
D > G(STL): Okay
A > G(STL): Um
D > G(FRG): I'll .
A > G(SUG): chip on print .
A > G(INF): I think we're fairly certain it has to be an advanced chip for the L.C.D. .
A > B(ELA): I am right in saying that , yeah ?
B > G(STL): Yeah , yeah um for
B > A(BEP): sorry ,
B > A(OFF): let me just double check on this .
B > A(ASS): Yeah , the uh L.C.D. di display requires an advanced chip , so
A > B(CAU): Okay
A > G(INF): so that's uh
A > B(INF): in terms of case , I mean you said your personal preference was titanium ,
B > A(ASS): Mm ,
B > G(FRG): but
A > G(INF): but from the market research it sounds as if we're going for more spongy .
A > G(SUG): But that was that we go for the rubber ,
A > B(INF): I'm just noticing from reading over your shoulder there ,
B > G(BCK): Hmm .
A > G(INF): it suggests that rubber could be anti-R. R.S.I. as well .
A > G(INF): So it would tick two boxes as far as in terms of the marketing ,
A > G(INF): 'cause we can say we both have we have you know a spongy feeling
A > G(ASS): well we didn't say that ,
A > G(INF): but we'd have the most likely popular feel to the remote , as well as having the anti R.S.I.
A > G(ASS): which would be a good feature to market if that's what people are worried about .
A > G(SUG): So I think personally I would be in favour of the rubber
A > G(ELO): but what's other people's opinions ?
B > A(ASS): Yeah .
B > G(STL): Well ,
B > C(SUG): Lawrence uh .
C > G(SUG): Let's do something like a titanium that we're going to dress in something . Let's say in something mm spongy or who knows what ,
C > G(SUG): so the still the the remote control is going to be manufactured out of titanium ,
C > G(ASS): the case would be out of titanium ,
C > G(SUG): but it will be let's say wrapped in in some spongy material or in something like that .
A > C(CAU): titanium with like a rubber casing round it .
A > G(SUG): So you would sort of see the titanium but sort of where the person would hold the remote could be rubber ,
B > G(BCK): Yeah .
A > B(ELI): is that possible ?
B > G(FRG): That could I think
B > G(STL): yeah um
B > A(INF): it doesn't say specifically
B > A(ASS): but I think that would probably be a good compromise , if you had the actual the casing made of the of uh material like titanium and then just the points at which it's being held being rubber ,
B > A(INF): 'cause I mean it would also encourage um holding the remote in the correct like way , uh such as to reduce the risk of R.S.I. .
D > G(ASS): Yeah , .
A > G(STL): So
A > G(INF): we're moving on to more uh of Lawrence's user interface concept ,
A > G(FRG): the
A > G(STL): so
A > G(INF): just have to decide exactly what we're gonna go for for the interface here ,
A > G(ELA): are we going for the basic ,
A > G(ELA): are we going for the advanced ,
A > G(ELA): are we going to try and get somewhere in between .
C > A(ASS): I think we we already decided on the idea of having a really basic uh basic approach for starters , and then through in some way to d to come up with with some let's say some more complicated functions ,
C > G(SUG): so basically we'd have just the basic buttons for on off , volume and uh channel switching , and then again the extra button which would take you let's say to the more complicated functions that you might need occasionally .
A > C(CAU): Alright .
D > G(FRG): 'Cause o w we
A > G(FRG):
D > A(ELA): we should have two designs , right , one basic and one more or
A > D(ASS): There'd be two designs for the actual user interface ,
A > D(ASS): but that would be both user interfaces would be available from the same device .
A > D(INF): So you'd be sitting there , you'd have the remote in your hand , and you would have you know your channel one , channel two , change channel , change volume .
D > G(BCK): Mm-hmm .
A > D(INF): Then there would be a button somewhere on it that said go to advanced , or a button that said advanced or whatever ,
D > G(BCK): Mm-hmm .

447 A > D(INF): you'd press that
448 A > D(INF): and then the user interface because it's L.C.D. we could then change it to one that had all the advanced controls .
449 D > G(BCK): Mm-hmm .
450 A > D(INF): And then a button that would be basic and that would take you back .
451 A > D(INF): So we'd have this same m same remote control but a different design of the actual features that are offered .
452 D > A(CAU): Mm-hmm , okay .
453 A > G(ELA): I think .
454 A > G(OFF): I just wanted to do the closing .
455 A > G(OFF): So I'll just go over this quickly .
456 A > G(STL): So
457 A > G(INF): for the next meeting , the two of you will be working together on creating some type of prototype by using the videoconference , using the modeling clay , I think
458 A > G(INF): I presume that's the stuff there .
459 B > A(ASS): Yeah .
460 A > C(ELI): Lawrence do you have modeling clay on your side ?
461 C > A(ELI): I do not .
462 A > G(INF): So you're gonna have fun with modeling clay in the next forty minutes .
463 A > G(STL): Um
464 A > G(FRG): as Tsveta
465 D > A(INF): Tsveta .
466 A > D(CAU): Tsveta ,
467 A > D(INF): you're gonna be working on the product evalutaion , um sort of l looking at decisions we've made and how they're gonna affect the actual marketing of the product .
468 D > G(BCK): Mm-hmm .
469 A > G(OFF): So I'll put the minutes in I'll put the minutes as minutes two , into the share folder , about two or three minutes after we finish the meeting ,
470 D > G(BCK): Mm-hmm .
471 A > G(OFF): and just sort of give a list of all the decisions we've made
472 A > G(INF): so that both you two can then get down to actually working on the prototype ,
473 A > D(INF): and then you can actually look at how that's gonna affect the marketing of it .
474 D > G(BCK): Mm-hmm , okay .
475 A > G(INF): And you'll get obviously specific instructions sent to you after the meeting's finished .
476 A > G(STL): So
477 A > G(ELI): with the decisions in mind , is there anything else that people have noticed since the last meeting ,
478 A > G(ELI): something that we don't think has been covered ,
479 A > G(ELO): or you know just any idea that you have ,
480 A > G(ELI): just is there anything that like you'd like to be added on ?
481 C > A(INF): I think we cover up the most the most important facts already .
482 B > G(ASS): Yeah , I think uh that's just about everything that's been covered .
483 A > G(BCK): Mm-hmm .
484 A > G(BCK): Mm-kay .
485 A > G(STL): So we're going for
486 A > G(INF): we're going for a remote control using the L.C.D. display with advanced chip ,
487 A > G(INF): we're gonna have a titanium casing , with rubber padding and rubber casing round it .
488 A > G(INF): We'll then have a basic user interface , with the option to switch to an advanced user interface ,
489 A > G(INF): and it's gonna be powered by battery .
490 D > G(BCK): Mm-hmm .
491 A > G(STL): That is our
492 A > G(INF): that is our final decision on exactly how this remote will look and work .
493 B > A(ASS): Sounds good to me .
494 D > A(ASS): Yeah , to me too .
495 A > G(STL): Alright .
496 D > G(STL): Yeah .
497 A > G(SUG): We can just sort of sit and stare .
498 D > G(ELI): Was it forty minutes ?
499 D > G(INF): I think we have ten more .
500 A > D(INF): Forty minutes .
501 A > D(ASS): Yeah .
502 D > G(INF): Because we started on ten to four .
503 A > G(INF): We started just before ten to four ,
504 A > G(INF): so we've got a good ten minutes .
505 D > G(ELI): This blue one which Lawrence offered us , it was and it had some buttons over it , was it uh like our model ,
506 D > C(ELI): or it was already produced ?
507 D > G(STL): Uh
508 A > G(STL): The the
509 A > D(CAU): the one on the screen ?
510 D > A(INF): on screen yeah ,
511 D > C(INF): it was one , one little blue , Lawrence ,
512 D > C(ELI): where did you get it from ?
513 D > G(FRG): Did you
514 C > G(STL): Oh .
515 A > G(STL): Uh
516 A > C(OFF): I'll get back to you .
517 C > G(FRG): I think it's
518 C > G(STL): uh yeah ,
519 C > A(SUG): would you go to my
520 C > A(ASS): okay ,
521 C > G(STL): let me see .
522 C > D(ELI): You mean uh this one ?
523 D > C(INF): Next one ,
524 D > C(INF): no no
525 D > C(INF): next .
526 A > C(INF): Next .
527 C > D(ELI): This one ?
528 D > C(INF): This one , yeah .
529 C > G(STL): Well
530 C > D(INF): that's basically a complete beginner remote control ,
531 C > D(INF): and it's basically aimed for for children
532 C > D(INF): so you would be able to programme it somehow in order for your kids to access just a limited number of of channels .
533 C > D(INF): And you'd have some protection ,
534 C > G(STL): you know you you see that kind of um how would you call it ,
535 C > D(INF): it's composed out of two parts ,
536 C > D(INF): so one of them would would automatically go yep , that part will automatically cover the other buttons that are not needed there but are needed only only by you when you make different settings I presume .
537 D > G(BCK): Mm-hmm .

C > D(INF): So it's already in use ,
C > D(INF): it's not a new idea .
A > G(STL): Uh so
C > G(INF): I put it j there , just to to realise , to see different shapes and different approaches to a narrow target complete beginner type remote control .
A > G(FRG): if y
D > C(CAU): Mm-hmm .
A > G(STL): So basically
D > G(ASS): Because this is fancy looking
D > G(INF): and um it has basic the basic things above and the more advanced features in it , so
A > G(STL): Mm . So
A > G(FRG): that's sort of
A > G(SUG): the one on the left , the blue one , is sort of what we'll be um that will be sort of what's the basic one ,
A > G(SUG): then the one on the right is gonna be how the advanced will look
A > G(INF): so basically we'll have a remote switching back and forth between these complete beginner and advanced user interfaces .
B > G(SUG): I imagine the final product will probably be more if you uh return to the previous slide , um yeah , probably be more like uh what 's on the left here ,
B > G(STL): I mean ,
B > G(SUG): it's not gonna be exactly the same ,
B > G(INF): I mean you can see it's got the basic um functions as push button push buttons and an L.C.D. displaying the information ,
B > G(SUG): and I imagine our final product because the buttons will need to be able to change , it'll have to be a touch screen thing ,
B > G(SUG): but to begin with it will have like the uh you know the very basic functions and the advanced ones available at the touch of a button .
A > G(STL): Yeah ,
A > G(BEP): and obviously there won't be an M.P. three player .
B > A(BEP): True . Uh no no headphones .
A > G(STL): Um just , well
A > G(INF): we've got the time .
A > G(STL): Um
A > G(ELI): in the design , as I remember , remember that the Real Reaction logo must be included somewhere in it ?
B > A(INF): Mm , yeah okay .
A > G(SUG): So we just need those two R.s inside the yellow background .
A > G(STL): Uh
A > G(INF): I don't know where
A > G(INF): that's obviously up to you ,
A > G(INF): but it needs to be somewhere ,
A > G(INF): I don't know how big they really want it .
B > G(BCK): Mm-hmm .
A > G(INF): Obviously it needs to be noticeable .
A > G(INF): That sounded like an email .
A > G(INF): We have five minutes to finish the meeting .
B > G(STL): Uh well uh
D > G(BCK): Mm mm mm .
B > G(BEP): push for time or possibly not .
A > G(BEP): I know , we'll never get this done in time .
D > A(ELI): And ho how am I going to do the product evaluation when I have only only our decisions now ,
A > D(INF): You'll get an email .
D > G(FRG): but you you made
D > G(STL): When
D > G(ELI): when you're done with uh the model of it , can I then do that , or
A > G(STL): You
A > D(INF): you do the product evaluation I assume based on the decisions that we've made .
A > G(STL): So
A > D(INF): you know that we're gonna come out with a product with a titanium case that's got rubber around it .
D > G(BCK): Mm-hmm .
A > D(INF): You'll then look at how that you know how the idea of that case appeals to the market .
D > G(BCK): Mm-hmm .
A > G(FRG):
B > D(INF): I imagine uh more specific instructions will come through the email after the meeting's finished .
D > G(BCK): Yeah .
B > D(INF): Make things clearer .
D > G(BCK): Okay .
A > G(OFF): I'll just finish this minutes
A > G(FRG): and then I can .
B > G(SUG): I think maybe um if we actually combined the emails that we've received that have you know , factual information in them , um so that they're available to everyone that might help the decision making process ,
B > G(INF): so I mean I've got an email here that tells you what limitations there are on using particular um particular parts of the particular materials together , etcetera ,
B > G(INF): and uh I'm sure that I mean you each have individual emails that give you factual information on either , you know , the market research side of things or the user interface side of things .
B > G(STL): Um so
B > G(BCK): Yeah .
B > A(ASS): Useful .
A > G(STL): And um
A > G(INF): voice recognition , I mean it sort of came out as it could be a good gimmick for selling the thing .
D > G(BCK): Mm-hmm .
A > G(ELA): Do we wanna approach it as an actual feature in the product ?
A > G(FRG): I mean you were saying it
D > A(CAU): Voice recognition .
A > G(INF): I mean the market research says it could be a good gimmick ,
D > G(STL): Mm .
B > G(BCK): Mm .
A > G(INF): whereas it might not actually be used
A > G(ELA): but then do we really to be used once they've bought the product ?
D > G(STL): Well
D > A(ASS): the market says that it won't be it will be more used from people um who are older than thirty-five years ,
A > G(FRG): So we just sort of
D > A(INF): but our public is below that .
A > D(CAU): Yeah ,
D > A(INF): And they say that they're more willing to use .
A > G(SUG): so we just drop it out .
A > G(ASS): Alright .
A > G(SUG): And remember fruit and veg when you're designing it .
B > A(ASS): Yeah , we'll uh keep that in mind .
A > G(ELA): And do we go for a single-curved case ?

B > A(ASS): Yes
B > G(STL): um
B > A(INF): the double curve isn't feasible with the rubber or the titanium , so
A > G(CAU): Single curved .
B > A(INF): Yeah .
B > G(FRG): So the
B > G(INF): not what exactly that means I'm not quite sure ,
B > G(STL): uh
A > G(ELA): And are we going for a backlight and the feature to find it ?
B > G(STL): Uh
B > A(ASS): yeah
B > G(INF): I think they were quite in demand
D > G(FRG): Why's
B > G(ELA): weren't they ,
B > G(INF): I mean with half of the people saying that uh fifty percent saying that they think the remote spends a significant time lost in the room somewhere ,
B > G(ASS): then a feature to find it will probably be quite sought after .
A > B(ASS): Alright .
A > G(INF): That's the minutes in the shared document , if anybody wants to go and look them up .
A > G(SUG): Make sure I haven't messed up somewhere ,
A > G(BEN): as I probably have .
D > A(CAU): So we're going to make them mm documents and put it in the sharing folder .
A > D(INF): Just the emails that you think would be useful for everybody to know , if you just copy the contents and put it into a word document ,
D > G(BCK): Yeah .
A > D(INF): that way w people can look it up and say you know , oh this is what the market research says or oh this is what the features available are .
D > G(FRG): Why's
D > G(STL): Okay ,
D > G(INF): I'm trying to save it for a second time .
C > G(SUG): Should we put them all in the emails folder , rather than having them all over in the same place ,
C > G(INF): 'cause afterwards we're not going to realise what and where we have .
C > G(INF): So there is a folder that they did , add emails ,
C > G(ASS): and it would be good if everyone would put his emails there .
A > G(STL): Mm oh we'll put
A > G(OFF): I'll put them in a separate word document
D > G(ELI): Are you able to make it a document and put it in the
A > G(INF): so we don't overlap .
A > G(INF): I've been told to finish the meeting now though .
A > G(STL): So
B > G(FRG): Alright , yep if you yeah , if you just
A > G(FRG): say
A > G(ELI): If everybody knows what they're gonna do ,
D > A(INF): Yes .
A > G(INF): we're all set .
D > G(INF): If this works .
D > G(INF): I'm trying to save it for a second time ,
A > G(FRG): Then I'm not su
A > C(INF): I think we're leaving you there Lawrence
D > G(FRG): and it's
C > A(CAU): Sorry ?
A > G(FRG): for you to do if
A > C(INF): I think you're staying on the videoconference in for Ben .
A > G(INF): But that's the meeting officially closed
D > G(BCK): Yeah .
A > G(INF): and we'll speak again in forty minutes .
D > G(STL): Okay ,
D > G(INF): I give up .
D > G(FRG): This is
D > G(BEP): hi ,
D > G(INF): I'm trying to save a document
A > G(OTH): Yep .
D > G(FRG): and for a second time it gi
D > G(INF): not responding ,
D > G(INF): and then again ,
D > G(INF): and now it's just not responding .
D > G(INF): I was just trying to decide where
D > G(INF): and it's not responding .
D > G(OTH): Okay .
D > G(FRG): So if I
D > G(BCK): Word documents ,
D > G(BCK): and save as
D > G(ELI): Can I save it on the shared folder? .
D > G(STL): Mm-hmm ,
D > G(ELI): is it because I wanted to copy paste from the email ?

ED1005c

B > G(OTH): Uh yeah .
D > G(INF): He can't see me .
B > G(FRG):
D > C(ELI): Lawrence , can you see me ?
C > D(INF): No I cannot .
D > G(FRG): I really cannot
C > G(FRG): I can't I don't know how to
C > G(OFF): let me see if I can work a little bit with my
B > G(OTH): Uh W. .
D > C(ELI): Can you see us all ?
C > D(INF): Almost , yep .
D > C(CAU): Almost , what does almost mean ?
B > D(BEP): Missing some limbs .
D > C(CAU): Okay , okay .
A > G(STL): Okay
A > G(INF): we're gonna get rid of them .
A > G(INF): 'Cause I love them ,
A > G(FRG): we need to actually um
D > G(ELI): What was what was the last one ?
C > G(STL): So
A > D(INF): ah well ,
C > G(INF): I have to do something now ,
A > D(INF): there's enough funny quotes .
D > A(CAU): Okay .
C > G(ELI): do you have to do anything ?
B > C(BEP): Sorry .
C > G(STL): I have
C > G(INF): I have to do something like a report for the meeting ,
C > G(ELI): do you have to do anything
C > G(ELI): or you're just all waiting for me and my report ?
B > G(FRG): I I don't kn
B > C(INF): I think we're supposed to be giving a kind of join report aren't we , on uh how how the design went .
B > C(SUG): So shall I just show the design to everyone else and you can kind of walk them through it ,
A > B(OTH): Go down .
B > C(ELI): or is that not what we're supposed to be doing ?
A > B(OFF): Let me get this up first .
B > A(CAU): Uh
B > A(ASS): okay .
A > G(STL): Okay
A > G(INF): this is our detailed design meeting .
A > C(ASS): Nope ,
A > C(SUG): stop playing with the mouse .
C > A(CAU): Sorry ?
A > G(INF): This is our detailed design meeting .
A > G(STL): Um
A > G(INF): this is our a bit of opening here ,
A > G(INF): just gonna briefly talk about what we've all been doing .
A > G(INF): Then we're going to have the presentation of the prototype by the industrial designer ,
B > A(ASS): Oh ,
B > G(INF): that's me .
A > G(INF): then evaluation of the design by our marketing expert .
A > G(INF): And then I'm making a finance presentation ,
A > G(INF): which is basically what we had to cut out because it was too expensive .
D > G(FRG):
A > G(ASS): Kind of .
B > G(FRG):
A > G(INF): And then we'll work on the product refin refinement if necessary .
A > G(INF): And we've got to get this all done in forty minutes .
A > G(ASS): Actually less than forty minutes
A > G(INF): but we'll get it done .
B > A(ASS): I'm fairly sure ,
A > G(STL): So
B > G(FRG): just .
A > B(SUG): I'll pass over to our industrial designer expert .
B > G(STL): Well , um
C > A(SUG): Wait a moment
C > A(ELI): when do I have to do my presentation about usability testing ?
A > G(FRG): You would do that
C > G(FRG): That's not
A > C(SUG): you would better well work together to do this presentation of prototype ,
A > C(SUG): so when he sort of says I'm fed up of talking , you can continue .
B > G(FRG):
D > G(FRG):
A > C(SUG): You can just chime in .
C > G(FRG): I cannot
C > G(INF): I have no time to look over my usability testings ,
C > G(INF): I just got them .
B > G(STL): Um ,
B > C(CAU): okay .
A > G(STL): Well
A > C(SUG): you can look over while Ben's talking .
A > C(ELA): Okay ?
C > A(CAU): Sorry ?
A > C(INF): You look over it while Ben is talking .
C > A(CAU): Okay .
B > G(STL): Right . So
B > G(INF): here we have the uh famous T-mote
B > G(INF): which we've been working on so hard over the past few hours .
B > G(STL): Um
B > G(INF): as you can see ,
B > G(INF): it's made of PlayDoh ,
B > G(ASS): which may be a problem in production ,
B > G(SUG): I think maybe rubber or plastic would be a better option .

B > G(STL): Um
B > G(INF): we shaped it like this just because it's uh ergonomic to hold ,
B > G(INF): it fits in the hand quite nicely
B > G(STL): um ,
B > G(INF): easy access to the buttons ,
B > G(INF): but it's also uh quite a unique shape ,
B > G(INF): it gives us a unique uh spot in the market .
B > G(STL): Um
B > G(INF): the blue is the outline of the shape of the actual thing ,
B > G(INF): whereas the red is the L.C.D. screen .
B > G(STL): Um
B > G(INF): and the yellow parts are the uh touch sensitive wh sorry , where the buttons would appear on the screen .
B > G(STL): Um
B > G(INF): we have the power button ,
B > G(INF): which is located centrally ,
B > G(INF): just um the most important button probably on the thing ,
B > G(STL): um
B > G(INF): and we've got two scroll uh scroll options here for volume and channels ,
B > G(INF): which you just operate by sliding your thumb up and down .
B > G(STL): Um
B > G(INF): and then at the bottom uh slightly out of the way you have the advanced uh sorry , switch to advanced view ,
B > G(INF): which would uh you know upon pressing it would changing change the layout of all these buttons
B > G(INF): and you get the control over things like you know tuning , contrast , all this sort of stuff .
B > G(STL): Um
B > G(INF): along the bottom there's another very small L.C.D. display
B > G(INF): which is just used for the uh power .
B > G(INF): It shows you how much battery there is left in the thing .
B > G(STL): Um
B > G(INF): which is separate from the rest
B > G(INF): because it will never be turned off ,
B > G(STL): uh
B > G(INF): whereas the main L.C.D. will uh go off after not being used uh to save battery life .
B > G(STL): Um
B > G(INF): further than that , I don't think there's really that much to say .
B > G(STL): Uh
B > G(SUG): I'll pass over to my uh user interface designer if he's here .
B > C(OTH): Lawrence ?
A > C(OTH): Lawrence .
B > C(BEP): Knock , knock .
A > C(OTH): Lawrence .
B > G(FRG):
D > C(OTH): Ooh .
B > C(ELA): Would you like to uh do your bit now or
C > G(ELI): I already have to do it , so fast ?
C > G(BEP): Sorry
C > G(STL): I I I
C > G(INF): I've muted the system in order for me to work on whatever I had to .
C > G(STL): So um
C > G(INF): my bet is something like I should give you a word report ,
C > G(OTH): oh my god ,
C > G(INF): I can't actually give you the word report
C > G(OFF): but then again I can speak about what should uh what the report should contain to a certain extent .
C > G(OFF): So I should give you a full description of what usability testing is .
C > G(INF): Usability testing as I could gather until now , is the idea of investing in in um user centred design ,
C > G(INF): which would uh return your investment in a quite rapid amount of time .
C > G(STL): Now then again um
C > G(OFF): the importance of usability testing ,
C > G(INF): it is important because we could get our money back I would say .
C > G(STL): Um
C > G(INF): Another thing that I have to touch what would be the estimated costs and duration for each operation ,
C > G(FRG): which
C > G(SUG): let me see where they would be
C > G(STL): mm
C > G(ELI): in the expert review ,
C > G(INF): no , it's not in here .
C > G(STL): Mm .
C > G(INF): Cost .
C > G(INF): A typical user walk through would take three two six weeks
C > G(INF): and is performed by at least two or of our usability experts .
C > G(INF): Costs are usually between fifteen thou thousand euros and thirty thousand euros , depending on the size of the application or website . The detail the required detail of the output and the availability and number of subjects .
C > G(STL): Okay . So let me
C > G(OFF): let me make a short a brief um how do you call it , put it in a in fewer words .
C > G(INF): It would take from three to six weeks
C > G(INF): and it would cost us between fifteen and thirty thousand euros .
C > G(ELI): What else do I have to cover ?
C > G(OFF): What I recommend to be done ,
C > G(INF): because I do consider your usability testing to be very important ,
C > G(SUG): I say that we should we should spend our money for it .
C > G(FRG): Rough estimation of the return on investment , it says in one of my documents that the investment would
C > G(STL): uh
C > G(OTH): Oh
C > G(SUG): wait ,
C > G(INF): usability walk through , usability test , I found something else .
C > G(INF): Description of the usability method .
B > A(SUG): Can you try and turn it up a little ?
C > G(OTH): Four to eight weeks and it costs around depending on the size and
C > G(OTH): oh my god .
C > G(INF): I kind of found two different two different costs .
A > C(SUG): Lawrence , can you speak up a bit ?
C > G(INF): So I need a little bit of time to see which one would apply to us .
C > G(STL): Um
D > A(ELI): What was it made of again ?
A > D(INF): I had rubber .
D > A(CAU): Rubber .

188 C > (OTH): So which is the difference between user walk through and and usability test ?
189 A > C(SUG): Can you speak up a bit Lawrence ?
190 C > A(ASS): Of course of course ,
191 C > G(SUG): one minute
192 C > G(INF): I'm just mumbling in order to to convince myself and to understand whatever is the difference between user walk through and usability test .
193 D > A(SUG): Can I hold it for a while ?
194 D > A(SUG): Can I hold it for a while ?
195 C > G(CAU): Oh right ,
196 C > G(INF): so there we go .
197 C > G(STL): We have two ways
198 C > G(INF): we have two ways of of doing our our test ,
199 C > G(INF): one would be user walk through ,
200 C > G(INF): which would mean that some users would would would go through the programme with with an expert
201 C > G(INF): or we can go in the usability test ,
202 C > G(INF): which would mean probably that a bunch of people would just test our our product ,
203 C > G(INF): and we have two different types of costs
204 C > G(OFF): which I am going to state in a moment
205 C > G(INF): for the user walk through , the cost as I have told you would be between fifteen thousand and thirty thousand euros
206 C > G(STL): and it would take around um from
207 C > G(INF): it would take from three to six weeks ,
208 C > G(INF): and for the usability test it would take from four to eight weeks
209 C > G(INF): and it would cost us around thirty thousand to fifty thousand euros .
210 C > G(STL): Now
211 C > G(SUG): I would say that it would be better for us to go through to take maybe uh to choose the user walk-through
212 C > G(INF): because then I think that an exp if an expert would go with you through the product , he has better chances of pointing some parts that you might actually miss otherwise .
213 C > G(STL): What else ,
214 C > G(FRG): return on the investment , I've read that let me see here
215 C > G(OTH): Techno ta tum tum tum tum tum tum tum .
216 C > G(INF): Oh expert review ,
217 C > G(STL): wait
218 C > G(INF): we have actually three ways of doing our usability test .
219 C > G(INF): The expert review would take from one to four weeks
220 C > G(INF): and it would cost us between five thousand and twenty thousand euros .
221 C > G(ASS): Interesting .
222 C > G(OTH): Let me see , can I go back home , yep I can .
223 C > G(STL): So
224 C > G(INF): recent research shows that investing in user centred design has a return of investments in about fifty days .
225 C > G(INF): Usability helps to extend the user group and to realize accessibility for different users .
226 C > G(STL): Right , so
227 C > G(SUG): I'd say that we should go still for user walk through where where normal people would would just test our our application , our remote control .
228 C > G(INF): And the rev the investment as they say it would go back to us in fifty days ,
229 C > G(STL): um
230 C > G(ELI): why to hire T.N.O. ,
231 C > G(INF): I say that we should hire them because they have a showcase uh an interesting showcase I might say .
232 C > G(STL): Um
233 C > G(INF): clients we have worked with for the last few years include government agency , public public safety , space education , mobile operators , I.T. companies , research and standardisation ,
234 C > G(STL): so
235 C > G(INF): because they they worked with a large pool of people
236 C > G(SUG): I think that we they would be suitable for us .
237 C > G(INF): Because after all , we are a very diverse company who produces coffee machines and remote controls and lots of things in the same time .
238 A > G(BCK): Alright .
239 C > G(INF): This would be it I would say .
240 A > C(CAU): Okay .
241 C > G(ELC): Is the information clear ,
242 C > G(STL): so let me
243 C > A(OFF): let me make a short review out of it .
244 C > G(INF): We have three ways of doing our testing ,
245 C > G(INF): which would be expert review , user walk through or usability test .
246 C > G(INF): Both have increased the both increasing costs in the same uh order as I have stated them ,
247 C > G(INF): and they both increase in in length of the time in the same order that I stated them .
248 C > G(INF): The investment would go back to us in fifty days ,
249 C > G(FRG): usability testing is um
250 C > G(INF): the usability in general is defined as the effectiveness , efficiency and satisfaction with which specific users achieve specific goals in um specific context .
251 C > G(ASS): I don't think that I worry to hire T.N.O. ,
252 C > G(STL): yeah
253 C > G(INF): I've touched that why we should hire this company
254 C > G(INF): because it has a large portfolio .
255 B > C(CAU): Okay .
256 A > C(CAU): Alright ,
257 A > C(BEP): thank you .
258 C > G(ELC): Any questions ?
259 A > G(STL): O uh
260 C > G(ELC): Clear as crystal ?
261 A > C(CAU): it's good .
262 A > C(BEP): That was good .
263 C > A(CAU): Okay .
264 A > C(ELI): Okay ?
265 C > A(INF): Okay .
266 D > G(STL): Um
267 B > G(STL): So
268 B > G(INF): evaluation by the marketing expert .
269 D > B(ASS): Yep ,
270 D > A(ELI): and where is my presentation ?
271 A > D(INF): Is it saved in here ?
272 A > D(SUG): Give me two seconds .
273 D > G(STL): So
274 D > G(OTH): Oh .
275 A > D(ELI): What's it called ?
276 D > A(INF): Evaluation .

277 D > G(STL): Okay so um
278 D > G(INF): I picked out the most important features
279 D > G(INF): and we all should uh give evaluation according to them .
280 D > G(INF): And now the criterium .
281 D > G(INF): Fancy look and feel , we have to kind of vote for this feature of our product ,
282 D > G(OFF): and maybe I'll I'll type
283 D > G(STL): So
284 D > G(ELA): fancy look and feel .
285 A > G(FRG): I would say
286 D > G(ASS): I'm seven .
287 A > G(ASS): I'll be a bit more optimistic and say a two out of three .
288 A > G(ASS): I think it does have a fancy but we're not as fancy as we would have liked .
289 D > G(CAU): Oh
290 D > G(STL): it was ,
291 D > G(STL): okay ,
292 D > G(ASS): one .
293 B > D(CAU): Right .
294 D > G(INF): I mixed them up .
295 A > G(STL): No
296 A > G(FRG): I do
297 A > G(ASS): I think it's not as fancy or as sophisticated as we would have liked it to be , but for the cost I still it's come out fairly fancy ,
298 A > G(ASS): I would give it a two or a three .
299 B > G(STL): Yeah
300 B > G(ASS): I think um I'd go for a three ,
301 B > G(INF): and the only reason we didn't achieve higher was you know due to financial constraints .
302 A > G(BCK): Mm-hmm .
303 C > G(ASS): I agree .
304 C > G(STL):
305 A > G(STL): Alright .
306 B > C(INF): You agree .
307 B > G(INF): A three as well then .
308 D > G(STL): So ,
309 D > G(INF): three .
310 A > G(STL): Um mov
311 A > D(SUG): moving swiftly on ,
312 A > D(INF): 'cause we've got ten minutes .
313 D > A(ASS): Okay
314 D > G(FRG): y you was ten ,
315 D > G(STL): uh
316 D > A(INF): you was two ,
317 D > A(ELA): right ?
318 A > D(ASS): I say two or three .
319 D > A(CAU): Two , okay .
320 B > G(FRG):
321 D > G(STL): So ,
322 D > G(ELA): shape .
323 A > G(FRG): I think this is gonna
324 A > G(INF): I think the shape will get it a lot of attention , whether it's positive or negative , I think it will definitely make it stand out .
325 B > A(ASS): Yeah .
326 A > G(ASS): I'd go for a one .
327 D > A(ASS): Yeah .
328 B > G(INF): I think we're all agreed ,
329 D > B(ELA): You ?
330 B > G(ASS): yeah , one as well .
331 D > C(ELA): Lawrence ?
332 C > D(ELC): Sorry ?
333 C > G(INF): I'm working on my report .
334 D > C(INF): The shape .
335 C > D(CAU): Shape ,
336 C > D(ELI): what about the shape ?
337 A > G(FRG): There's a good
338 D > C(ELA): How will you evaluate it ?
339 C > G(BEP): Sorry ,
340 C > G(CAU): I couldn't understand any of you .
341 D > C(ELA): How will you evaluate the shape ?
342 D > C(INF): True false false , you have scale from one to seven , one is true .
343 D > G(INF): You can see it actually .
344 C > G(STL): Wait wait wait ,
345 C > D(CAU): I can't understand the idea .
346 C > D(ELI): What do I have to do now to evaluate ?
347 A > G(FRG): It is the
348 D > C(INF): We are evaluating the project ,
349 D > C(INF): and everyone of us should give his evaluation according to the criterias I have given .
350 C > G(STL): Well right ,
351 C > D(INF): I kind of missed the c criterias .
352 D > C(ELA): So what's yours ?
353 A > C(ELA): Do you think the shape was good ?
354 D > G(BCK): Shape , yeah .
355 A > C(INF): One for very good , seven for very bad .
356 C > A(CAU): One for the very good .
357 D > G(BCK): Yeah .
358 B > G(CAU): Okay .
359 A > C(BEP): Thank you .
360 D > G(STL): Okay ,
361 D > G(INF): next one .
362 D > G(ELA): Theme
363 A > G(FRG):
364 D > G(FRG): is yeah is the theme
365 B > G(STL): Oh
366 B > G(INF): it didn't really follow onto the fruit and vegetable theme very much ,
367 B > G(INF): they say it might look vaguely organic
368 B > G(ASS): so I'd say a four at best .
369 A > B(ASS): Yeah , four .

D > G(SUG): And ca can we just model the outside to be with fruit and vegetables ?
B > D(ASS): Yeah we could have the casing you know adorned with pictures of fruit or vegetables .
D > G(STL): Yeah
D > G(FRG): maybe d different different uh
A > G(STL): We
A > D(SUG): we'll get to that in the the design ,
D > G(FRG):
A > D(INF): but at the moment I think we have to go for a four .
D > G(STL): Okay
D > G(FRG): at the moment ,
D > B(ELA): so you're four ?
B > D(ASS): Yeah .
A > D(ASS): I'm a four .
D > G(OTH): Four four .
A > C(ELA): Lawrence ?
C > A(ELI): For what ?
D > C(INF): For theme .
A > C(ELA): Pick a number .
B > G(FRG):
D > G(FRG):
C > G(ELI): A number for what ? Which section ?
A > C(INF): For the theme .
D > G(STL):
C > A(CAU): Theme ,
C > G(STL): mm . I
C > G(ASS): I'd say we should go for two .
A > G(BCK): Two .
C > G(ASS): It's a very good theme . I like it .
D > G(BCK): Two .
B > G(FRG):
D > G(STL): Okay ,
D > G(INF): next one is material .
D > G(ELI): Do you think rubber will satisfy the public's feeling of sponge ?
A > D(INF): Yeah .
B > G(STL): Uh
B > D(INF): yeah I think it'll satisfy what you know the trend is supposed to be ,
B > D(ASS): I think it's gonna be a a one in that respect .
D > G(BCK): Yeah .
C > G(ASS): A one indeed .
D > G(ASS): One for all of us .
D > G(ELA): Then easy to use ?
D > G(ASS): Definitely .
B > D(ASS): I think yeah , an another one ,
D > C(ELA): Lawrence ?
C > D(ASS): One indeed .
D > C(CAU): Okay .
D > G(ELA): Less time to learn how to use it ?
A > D(ASS): Yeah .
B > G(STL): Um
B > G(ASS): I think just because it has two different displays , it might put people off a little bit ,
B > G(INF): but uh it's quite basic ,
D > B(ASS): Yeah
D > B(ASS): but for the basic one it's
B > D(ASS): yeah ,
B > G(STL): I think
B > G(ASS): I think a two I'd say .
C > G(STL): Um
C > G(SUG): we should stop for a moment ,
D > G(BCK): Yeah .
C > G(INF): I'm having a big workload .
C > G(SUG): Wait ,
C > G(INF): I won't be able to follow you .
C > G(SUG): So if you can carry on without me it's okay ,
C > G(INF): if not I need to go through my my new email .
A > C(ASS): Alright .
B > C(ASS): Okay .
C > G(STL): So um say the t
D > G(STL): Okay
D > G(FRG): we need your
C > G(SUG): say uh fast what else I have to choose
C > G(OFF): and I'm going to say them fast
C > G(SUG): and you can discuss them between you .
D > C(ASS): Okay
D > G(STL): w
D > G(OFF): we can do this faster ,
D > C(ELA): so less time to learn how to use it , what's your opinion ?
C > G(STL): Okay
C > D(ASS): one .
D > C(CAU): One .
D > C(ELA): And then accessible in the dark and if you have lost it .
C > D(ASS): Two .
A > D(ASS): Go for a four .
D > G(STL): Okay .
D > A(ELA): You ?
A > D(ASS): Four .
D > A(INF): Four .
B > G(STL): Uh
B > D(ASS): three .
D > G(STL): Um
D > G(ELA): appeal to people below forty and cost , twelve point five .
A > D(ASS): Two .
B > G(STL): Uh
B > D(ASS): I'd go for one .
C > D(CAU): Appealing to people to do what ?
D > C(ELA): Lawrence ?

D > C(INF): Appeal to people below forty years old , and the cost .
C > D(CAU): Yep
C > D(ASS): one , one .
D > G(STL): Okay , so
D > B(ELA): what was yours about
D > G(STL): um
B > G(FRG): so I was
B > D(ASS): I was saying two .
D > B(ELA): Less time ,
B > D(CAU): Yeah ,
B > D(ASS): two .
D > A(ELA): and you ?
A > G(STL): Uh
A > D(ASS): one .
B > G(BCK): Right .
D > G(STL): Okay so
D > G(FRG): now according to to this , I think we could
C > G(SUG): Wait wait wait .
C > G(INF): According to my email we just discovered something miraculously interesting .
C > G(STL): Uh
C > G(OFF): I am going to briefly read the important part ,
C > G(FRG): the innovative interface component's a colour touch display .
D > G(INF): I have email as well .
C > G(STL): Um
C > G(BEP): Sorry ,
C > G(INF): a touch colour display measures for over six centimetres
C > G(INF): and has a resolutions of which is sufficient for twenty four icons .
C > G(STL): Uh
C > G(INF): the idea is that the production costs are only three euros ,
C > G(FRG): and it m
C > G(INF): you can incorporate it in any of our current projects .
C > G(ELA): Should we go for it ? The new touch colour display screen , which seems to be interesting .
C > G(ELI): How much are we paying for the old one ?
A > C(INF): Three euros .
D > A(CAU): Three euros , for what ?
C > G(STL): Well ,
C > G(FRG): this one at least has
A > D(INF): L.C.D. display .
D > A(CAU): Oh .
C > G(INF): this one probably has more interesting features .
C > G(ASS): So I think we should change it .
B > C(ASS): Yeah .
A > C(ASS): Yep .
B > C(ASS): Let's jump on the bandwagon .
D > G(STL): Um I
D > G(INF): I have a mail for release of new user interface component ,
A > D(CAU): Yeah ,
A > D(INF): that's what we were just talking about .
B > D(CAU): Yeah .
D > G(FRG): he just
D > G(BCK): yeah .
A > G(STL): Okay .
A > D(SUG): Move onto the next
A > D(SUG): keep going .
B > A(ASS): Yep , okay .
A > G(INF): Six minutes .
A > D(SUG): Keep going .
D > G(STL): Uh for uh yeah , so now uh we
D > G(ELO): we might want to change these numbers , I mean to change something in our product so that we could have all ones , or at least two .
B > G(FRG):
A > G(STL): Well
A > G(FRG): we'll get to that at the
A > D(INF): there's a stage for refining that .
A > D(INF): So this is our evaluation of the current one ,
B > G(BCK): Yeah .
A > D(SUG): and when we get to the refining stage we'll think about how to change it .
A > G(STL): So
D > A(ASS): Mm-hmm ,
A > D(ELI): it that the end of your presentation ?
D > A(INF): yeah that's that's all .
B > D(CAU): You've done it all , okay .
A > G(STL): So
A > G(OTH): Uh no what are you doing ? Don't want you , go away .
A > G(OTH): five .
A > G(INF): Project manager's finance presentation .
B > A(INF): That would be you .
A > G(BEP): Ooh , this'll be fun .
A > G(OFF): show you our finances .
A > G(INF): Our finances , using the battery with the advanced chip , the single curved , the rubber , and the L.C.D. display which we'll now replace with the new uh touch colour display , gives us an overall total of twelve euros , keeping us zero point five euros underneath the production the maximum production cost .
A > G(INF): Therefore creating profits of thirteen euros for every unit sold .
D > A(ASS): Good .
B > G(ASS): That looks good to me .
A > G(BCK): Alright .
B > G(INF): We even have half a euro to spare .
A > G(STL): Uh
D > G(ASS): Mm-hmm .
A > G(FRG): it means that
A > G(SUG): if you hold on a second
A > G(OFF): while I and get the calculator ,
A > G(INF): then we can earn fifty thousand fifty million by selling just three million units .
D > G(STL): Mm so
B > A(ASS): Very good .

D > G(FRG): make it better .
A > G(INF): So we're gonna reach our target by selling just under four million units .
D > G(ELA): Yeah .
D > A(CAU): Mm-hmm ,
D > G(ASS): so the profits will be good .
B > D(ASS): Yeah .
A > D(ASS): Hopefully .
D > G(BCK): Yeah .
A > G(STL): Um
A > G(OTH): where am I here ,
A > G(FRG): so it needs .
A > G(OTH): Come on .
D > A(ELI): Is this all ?
A > G(INF): I didn't mean to hit that .
A > G(ELI): Is somebody else on here ?
A > G(ELI): Who closed presentation one ?
B > A(CAU): Oh .
D > A(CAU): Mm-hmm .
B > G(STL): Um
A > G(INF): Somebody's shutting down everything .
B > G(STL): Okay ,
B > G(ASS): that's odd .
B > G(INF): I've not touched it .
D > G(ASS): Mm-hmm , me too .
A > G(INF): I lost all my fancy stuff .
C > G(ELI): Do we need to do anything else ?
A > G(STL): Um
A > G(INF): I think there was something else .
A > G(BEP): Never trust technology , it always fails you .
D > G(FRG):
A > G(STL): Um
A > G(INF): product refinement if necessary , we've already said we're gonna go for the colour touch instead of the L.C.D. display .
D > A(ASS): Yeah
D > G(FRG): and here for these things .
A > B(ELI): Did you press something ?
B > G(FRG): I meant to quit l uh
B > A(INF): I mean to quite the uh the L.C. view ,
B > A(BEP): sorry .
B > G(STL): Um
B > A(INF): I wasn't doing it last time .
A > G(STL): Um so
B > A(INF): Product refinement .
A > G(INF): we got rid of the titanium in the design to save costs ,
A > G(INF): got rid of the coloured buttons to save costs ,
A > G(INF): though with the colour touch display that's now included .
A > G(STL): Uh
A > G(ELO): is any other product refinement you think are necessary to improve the productivity ? To improve the appeal of our product .
D > A(SUG): Maybe the team .
A > D(CAU): Hmm ?
D > A(INF): The theme .
D > G(STL): We can
B > G(BCK): Yeah .
A > G(FRG): If we add on to the .
D > G(SUG): we can make different uh different designs of different colours and vegetables uh fruits , I mean
B > G(BCK): Mm .
A > D(INF): aesthetic look of the product .
C > G(BCK): Fruits .
D > A(ASS): Yeah
D > G(INF): like uh like G.S.M.s , they have different
D > G(ELI): how do you call it ?
A > D(INF): Covers .
D > A(ASS): Yeah ,
D > G(SUG): so ours can have the same so it could suit every taste .
B > G(ASS): Yeah , I think that sounds like a good idea .
A > D(ASS): Yep .
A > G(STL): Okay ,
D > G(ASS): Yeah ,
A > G(ELO): anything else ?
D > G(INF): so they can choose that as well .
D > G(STL): Um
D > G(INF): Accessible in the dark , we have four three two three .
A > G(STL): Um
A > D(ASS): the problem of that is it would cost more to do .
B > G(FRG):
B > G(FRG):
D > G(BCK): Mm-hmm .
A > D(INF): So we've sort of sacrificed that functionality in the attempt to keep the production costs low .
B > G(ASS): I think that given there are constraints we've achieved quite a lot .
A > B(ASS): Yeah .
B > G(STL): Um
D > B(ASS): Yeah .
B > G(INF): good ratings in most the categories .
A > G(STL): Okay .
D > B(ASS): Yeah you're right .
A > G(BEP): just leave to say thank you .
B > G(FRG):
D > G(STL): Yeah .
B > G(FRG):
D > A(ELI): How much time do we have ?
A > G(STL): Well
A > G(FRG): we have
B > D(INF): Two minutes or so .
A > B(ASS): No
A > G(INF): we have twenty seconds .
B > A(CAU): Uh okay .

652 D > A(CAU): Huh . Okay .
653 A > G(INF): That's the meeting .
654 D > G(BEP): So great job , .
655 A > G(STL): Um
656 D > G(ASS): I really buy w will buy this that remote .
657 B > G(BCK): Mm .
658 A > G(INF): I just have to do the product specification report .
659 A > G(INF): So this won't be for you two .
660 A > G(SUG): Specify how it works .
661 B > G(STL): Um
662 C > A(BEP): For ever and ever .
663 B > C(BEP): Sorry .
664 B > C(ELO): Do you wa do you wanna go with this Lawrence ,
665 B > C(ELO): or shall I ?
666 C > G(STL): Well ,
667 C > B(SUG): it's your choice after all .
668 C > B(INF): You're the industrial designer ,
669 C > B(INF): I'm just for user interface .
670 B > C(ASS): Okay .
671 B > G(ELI): How does it work ?
672 B > G(STL): Um yeah it
673 B > G(INF): it's got an infrared uh interface , and uh with sorry inf infrared to interface with the T.V. and a touch screen to interface with the user .
674 A > B(SUG): Continue .
675 B > G(STL): Uh
676 B > A(BEP): sorry
677 B > G(STL): yeah
678 B > G(INF): infrared interface with the T.V. , touchscreen with the user .
679 B > G(STL): Um
680 B > G(INF): and has the ability to change between two displays for both basic and advanced control .
681 D > G(OTH): Digital pages .
682 B > G(INF): I don't think there's any real call for a more technical explanation ,
683 B > G(OFF): and I don't think I would give one anyway .
684 C > G(STL): So
685 C > G(FRG): from from my point of view
686 A > B(INF): Provide argumentation for the decisions that have been made .
687 C > G(ELC): Do you happen to hear me ?
688 B > A(ASS): Okay .
689 A > C(CAU): Yeah we can still hear you .
690 C > A(CAU): Oh okay .
691 C > G(INF): From my point of view what I would add to the idea ideas already mentioned , it would be the fact that you could give it it's a remote control that you could look from two approaches ,
692 C > G(INF): it would be a simple one ,
693 C > G(INF): and just by pressing a button you would get a um advanced user remote control .
694 C > G(INF): So this is how it would work .
695 C > G(STL): You're just you're
696 C > G(INF): you would be able to access your very basic functions without doing anything at all , and without being distracted by any other complicated functions ,
697 C > G(INF): and when needed you can you could access them those as well .
698 B > C(CAU): Okay .
699 B > G(FRG): So how would you s
700 B > A(ELI): what was the next thing ?
701 A > G(STL): Um
702 A > B(ELI): provide argumentation for the decisions that you've made .
703 B > A(CAU): Okay ,
704 B > G(FRG): do you want
705 B > C(SUG): Lawrence do you wanna cover that , provide some arguments for the decisions that were made , or argumentation ?
706 A > B(OFF): I can sort of get this one done ,
707 A > B(INF): I'm just sort of typing it .
708 B > A(ASS): Yeah , okay ,
709 B > G(STL): um
710 D > G(STL): Um
711 B > G(INF): what was the argumentation , I don't know .
712 C > G(STL): Well ,
713 C > G(SUG): let's think a little bit um which are the decisions first ?
714 D > G(STL): Well y y
715 D > A(INF): you said it when you were explaining how to how to use it .
716 B > G(STL): Yeah I mean ,
717 B > G(FRG): the reasoning behind it was
718 C > D(ELC): Tsveta , do you hear me ?
719 D > C(CAU): Yeah .
720 B > C(CAU): Yep .
721 C > D(CAU): Oh ,
722 C > D(ELC): have you heard me a few moments ago ?
723 A > G(STL): Mm
724 D > C(CAU): No .
725 A > C(CAU): no .
726 C > D(ASS): Oddly enough .
727 C > D(INF): I've been speaking until now .
728 C > G(FRG): And probably you couldn't
729 D > G(STL): Well
730 D > C(CAU): I heard that you'll be speaking something but I didn't quite heard it properly .
731 C > D(CAU): Oh okay .
732 C > D(OFF): Let me go back ,
733 C > D(INF): I thought that it might be a problem with the equipment or who knows .
734 B > C(ASS): No it's all working .
735 C > G(INF): So the idea is uh we we need to to justify our decisions
736 C > G(SUG): but let's think at what decision actually we have to justify .
737 C > G(STL): We
738 C > G(INF): we gave up at some bits and pieces
739 C > G(INF): because we had to , the costs said that we we can't allow ourselves to spend more , the budget didn't allow us to spend more .
740 C > G(ELI): What else ?
741 A > G(ASS): I've put down the basic and advanced the basic and advanced views allow the device to cater to a greater variety of people .
742 A > G(INF): The original titanium design was sacrificed for cost as was the voice activation feature .
743 B > G(STL): Okay , yeah um I think

A > G(INF): I think that's as much as we need to say there .
B > G(BCK): Yeah .
A > G(STL): Um
A > G(ELI): components and look and feel . Specify all components , properties and materials here .
B > G(STL): Okay so
B > G(INF): the components we have are um the rubber casing .
B > G(STL): Um
D > G(INF): L.C.D. L.C.D. ?
C > G(INF): L.C.D. screen .
B > D(ASS): Yeah the L.C.D. touch screen ,
A > G(INF): Colour touch screen .
B > G(STL): so
B > A(ASS): yeah the colour touch screen .
D > G(ELI): And what is this um innovative interface component ?
A > D(CAU): Mm ?
D > C(ELI): Lawrence ?
B > G(STL): That's
B > D(INF): that's the colour touch screen .
C > D(INF): That's the colour s touch screen that I said that we should incorporate it rather than our old L.C.D. screen .
D > B(CAU): The colour touch
D > C(CAU): Uh-huh .
B > G(BCK): Yeah .
C > G(INF): So basically we're getting at the same value another another screen
C > G(ASS): which seems to be better than our old choice .
D > G(BCK): Mm-hmm .
B > G(INF): I think with our original L.C.D. disc uh screen it wasn't touch screen
B > G(INF): so we would have needed buttons as well ,
B > G(STL): uh
B > G(ASS): so it's lucky that this thing came along
B > G(INF): 'cause otherwise our product wouldn't have been viable at all .
A > B(ELI): Describe the form , material and colour of the case here .
B > G(STL): Okay so it's a well
B > G(OFF): describe the form of it .
B > G(INF): It's in the shape of a T. .
A > G(INF): A T. .
D > G(INF): T. ?
D > G(INF): Or maybe like a tree ,
B > G(BCK): Mm .
D > G(FRG): because it's
A > G(STL): Well
A > D(INF): we're calling it the T-mote ,
B > A(ASS): Yeah .
A > D(ASS): so we're saying it's in the shape of a T. .
D > A(CAU): T-mote .
B > G(FRG): But with a more with
D > G(SUG): Trimote ,
D > G(FRG): or
B > D(ASS): oh yeah , could go for the theme .
D > G(INF): Because it's like a tree
D > G(FRG): and
A > G(INF): That would involve me going back and changing all the references though .
B > A(ASS): Uh never mind , doesn't matter .
D > G(FRG):
A > G(ASS): I'm too lazy .
D > A(OFF): Let me do it .
B > G(STL): Um yeah , so
B > G(INF): in the shape of a T. but with a more organic more organic feel uh likening it to our theme , our somewhere trees .
B > G(STL): Um
A > G(FRG): Call it changing call it availa
A > G(INF): colour variant depending on user preference .
B > A(ASS): Yeah .
A > G(ELI): Provide argumentation for the decisions that we made .
B > G(STL): Uh
B > A(INF): we were just going by the market research really ,
B > G(ELA): weren't we ,
B > G(STL): and
A > B(ASS): Yeah ,
A > G(INF): following market research and making it unique .
D > G(INF): And evaluation criterias .
A > G(STL): We'll follow market we'll say we'll follow follow
A > G(INF): follow market research with a view of making a unique and
B > G(INF): Innovative product ,
A > B(ASS): Yeah .
B > G(BCK): mm .
C > G(INF): And attractive .
D > G(INF): Maybe marketing strategy , trends and user requirements .
B > G(STL): Well that
B > D(ASS): that is the market research really , I think ,
B > D(ELA): isn't it ?
D > B(ASS): Market okay , yeah .
B > G(FRG):
A > C(ELI): User interface . Describe the function , position , form , material and colour of the interface elements here .
C > G(BCK): Yep .
B > A(INF): Uh we've already done this .
C > G(STL): So
C > G(INF): positions , uh central position for the on and off button , uh convenient positions for the other two buttons , and uh let's see , button positions for the energy level button .
C > G(ELI): What else do we have there ?
C > G(STL): Oh yeah ,
C > G(FRG): top position for the top ,
C > G(ASS): no no no
C > G(INF): it was still bottom ,
C > G(INF): bottom position for the extra functions button as well .
A > G(STL): 'Kay so
A > C(INF): at the interface central position for the power button and uncluttered look for the other f uh uh other buttons available .

C > A(CAU): Sorry ?
A > C(INF): The interface was designed with a central position for the power button , and an uncluttered design for the other buttons present on the basic interface .
C > A(ASS): Clearly , yep .
A > G(ELI): Provide argumentation for decisions that we made .
B > G(STL): Um
B > G(FRG): aesthetically appealing , it's s
B > G(STL): yeah okay , so
B > A(INF): the decision or whatever , aesthetic appeal .
A > B(INF): Aesthetically appealing and simple to use .
B > A(ASS): Yeah .
A > G(ELI): Costs . Specify the composition of the production costs here .
C > A(INF): We have them in the table I think .
B > G(BCK): Okay .
D > A(INF): This is the whole specification of the product ,
D > A(ELA): right ?
A > D(ASS): This is the product specification report .
D > A(CAU): Mm-hmm .
B > D(ASS): Yeah .
A > G(INF): total production value , cost of twelve euros .
B > A(ASS): Of twelve , yeah .
B > G(FRG): Tw
B > G(BCK): yeah
B > A(INF): four .
B > A(INF): Twelve euros , leaving thirteen euros profit at the intended selling price or something .
B > G(INF): This'll be our five minute warning .
D > B(ASS): Yeah .
A > G(ELI): How come my one doesn't beep ?
D > G(FRG):
A > G(INF): Everybody else's beeps apart from mine .
B > A(BEP): It's broken .
A > G(BEP): I feel left out .
B > A(OTH): Oh .
C > G(INF): Five minutes left to finish your meeting and your design .
A > G(ELI): Provide argumentation for the decisions that have been made .
A > G(INF): The lower the cost the more profits that are made , therefore the quicker the targets of fifty million is reached .
B > G(BCK): Mm ,
D > A(ASS): Yeah .
B > A(ASS): yeah .
B > A(INF): The higher the profit margin or something like that .
A > G(INF): But the product should also possess innovative innov should also be appealing to potential users .
B > A(ASS): Yeah .
A > G(ELI): The market .
B > G(STL): Uh
B > G(INF): aimed at young people who are trendy and cool and hip .
D > G(FRG):
B > G(FRG):
D > G(INF): For people under forty ,
D > G(FRG): or
B > C(ASS): Yeah , or I I'd say under thirty five , even .
D > B(ASS): Yeah .
A > G(INF): And the T-mote are targeted at young people under forty with advantaged features
A > G(FRG): and
A > G(ELI): I'm trying to think of a word for the design .
A > G(INF): Unusual design .
C > A(INF): Unique maybe ? Almost ?
A > C(ASS): Unique , yes .
A > G(INF): Unique design .
A > G(FRG): This should ape
A > G(INF): this should be found appealing to the generation generations looking for cool and technically technologically
B > G(BCK): Yeah .
C > G(FRG): Frien
D > G(FRG): Innovati
C > A(INF): definitely friendly , technologically curious .
A > G(BCK): Yeah .
A > G(INF): Innovative . Technolog technologically innovative .
B > G(OTH): Product
B > G(STL): um
A > D(INF): Just grab the results of the product evaluation here .
D > G(FRG):
A > G(STL): So
A > D(ELI): was it mostly in the high end ?
D > A(CAU): The project evaluation .
B > D(INF): Yeah .
D > G(STL): Uh
D > A(ELI): do I have to calculate the overall ?
B > D(INF): No , o overall view .
A > D(SUG): Just give us a rough
D > G(STL): Overall .
D > G(STL): Mm
D > G(INF): Two maybe ,
D > G(FRG): or
B > D(INF): It generated consistent you know c reasonably consistently high results in evaluation .
B > G(STL): Um
D > B(INF): Yeah , if we exclude the
B > D(INF): The theme or whatever it was
D > G(STL): uh
D > B(ASS): no no
B > D(ASS): or no
D > G(FRG): th we
B > G(FRG): e
D > B(INF): We decided something for the theme ,
D > B(ELA): right ?
D > B(INF): To be in different design ,
B > D(ASS): Yeah , yeah , yeah .

D > G(INF): so this is this must be one ,
D > G(FRG): and then we have only accessible uh
B > D(INF): It wasn't ability to find it in the dark or when lost , that's the only other weak point .
D > B(ASS): Yeah and if you have lost it ,
D > A(SUG): so you can write that on this is the low the lowest mark
D > G(FRG): but that's because
A > G(INF): Made a sacrifice .
B > G(INF): For practicality . Or practicability .
D > A(ASS): Yeah .
D > B(ASS): Yeah .
B > G(FRG):
A > G(INF): Some had to be ignored due to the high production costs they would incur .
A > G(INF): Where possible the market research was used to define define a better product for our target group .
B > G(STL): Um yeah
B > A(ASS): maybe determine instead of define .
B > A(ASS): I don't know ,
B > A(INF): nitpicking really .
A > G(OTH): I wasn't I didn't like define
D > G(INF): And maybe if we had to produce a product that costs a little bit more , we could add this feature .
B > D(ASS): Yeah . With a higher budget we'd be able to accommodate more of the things .
D > B(ASS): Yeah .
B > G(INF): That must be them telling us to stop .
D > G(ASS): Yeah . Due to time restrictions
D > G(INF): Save everything , yeah save it .
A > G(STL): Uh
A > G(INF): I finished it .
D > A(INF): Project manager , please finish the product specification within ten minutes .
A > D(INF): We just finished it .
D > A(CAU): Okay .
A > G(INF): And I think that's everything saved .
D > G(INF): Save it in the shared project folder .
D > A(SUG): But you can change T. to tri , come on .
A > D(ASS): I can live without changing T-mote to Trimote .
D > A(ELI): Why not ?
C > G(STL): Okay ,
C > G(INF): I'm hanging up .
D > A(OFF): I can do it .
A > G(INF): That's us done .

Appendix B

What follows is an abstract from the AMI Dialogue Act Annotation Guidelines, focussing on the description of eliciting acts. The complete guide gives instructions for the following tasks

- Dialogue act segmentation and dialogue act labeling.
- Adjacency pair annotation: for example question answer relations.
- Addressing: who the speaker is talking to.
- Reflexivity: if the act is on task or about the communication or it's organisation.

The annotation schema distinguishes four different subtypes of elicit acts.

Elicit Acts in the AMI Dialogue Act Annotation Schema

Three classes of acts are distinguished that have their own associated types of eliciting acts:

1. Acts about information exchange: ELICIT-INFORM
2. Acts about possible actions: ELICIT-OFFER-OR-SUGGESTION
3. Acts that comment on the previous discussion: ELICIT-COMMENT-ABOUT-UNDERSTANDING and ELICIT-ASSESSMENT.

Acts about information exchange

Acts in this category can either express information or attempt to elicit information from others.

The ELICIT-INFORM act is used by a speaker to request that someone else give some information. The act doesn't have to make clear who is meant to give the information.

Example 1

UI	any of these things we need to address in the next meeting | or should we at least set some basic things that we need to prepare for
PM	YEP SO MARK ARE THERE ANY CONSTRAINTS OR INFORMATION FROM THE MARKETING SIDE | I mean | WHAT SORT OF WHAT SORTA PRODUCT ARE WE ENVIGIT- ENVISAGING
ME	well I think from the research that we've done so far what we would uh what we would like to have is something that would work for a variety of different devices but still maintaining simplicity

Example 2

PM	...DO YOU HAVE ANY OTHER INFORMATION FOR US AT THIS STAGE
ME	no
PM	no, okay

Acts about possible actions

This group of act classes is about expressing possible actions that the group, some individual in the group, or some person or group in the wider environment could do, such as the person running the recording equipment or people from the organization to which the group belongs (e.g. for the remote control design teams, researchers from the marketing department).

In a SUGGEST, the speaker expresses an intention relating to the actions of another individual, the group as a whole, or a group in the wider environment. Sometimes, SUGGEST can take the form of a question, especially when the speaker is not sure that the group will accept the idea or suggestion.

In an OFFER, the speaker expresses an intention relating to his or her own actions.

In a ELICIT-OFFER-OR-SUGGESTION, the speaker expresses a desire for someone to make an offer or suggestion. This can either be about something specific or a more general attempt to elicit an act about a possible action, such as "What should we do next?"

Example 3

	... the group discuss a range of features and try to see whether there are any further ideas...
ID	Should we maybe make a decision about what features we actually want to include, 'cause we've thrown a lot of features onto the table, but
PM	Yeah.
ID	Do we actually want to incorporate all of them | OR HAVE WE MISSED ANYTHING?

Example 4

UI	YOU NEED TO GIVE ME YOUR IDEAS — and then I need to see whether that would sell in the market place

This last example is a harder case than the others - the idea here is that the PM wants management to make a suggestion to the group about the design.

Commenting on Previous Discussion

In these acts, the participants contribute to the discussion by commenting on what has been said or done so far. As usual, there is an act for attempting to elicit comment on the previous discussion.

An ASSESS is any comment that expresses an evaluation, however tentative or incomplete, of something that the group is discussing, where the something could be another dialogue act or something apparent from the working environment, like slides or, in the remote control design trials, the playdough remote control mock-up. There are many different kinds of assessment; they include, among other things, accepting an offer, expressing agreement/disagreement or any opinion about some information that's been given, expressing uncertainty as to whether a suggestion is a good idea or not, evaluating actions by members of the group, such as drawings.

In an ELICIT-ASSESSMENT, the speaker attempts to elicit an assessment (or assessments) about what has been said or done so far. Sometimes a speaker seems to be making a suggestion and eliciting an assessment about it at the same time. In these cases, look at the information presented: if it is new information then it will be a SUGGEST (see above), if it implies previously known information then it will be ELICIT-ASSESSMENT.

Example 5

	... the group discuss a range of features and try to see whether there are any further ideas...
ID	Should we maybe make a decision about what features we actually want to include, 'cause we've thrown a lot of features onto the table, but
PM	Yeah.
ID	DO WE ACTUALLY WANT TO INCORPORATE ALL OF THEM | or have wee missed anything?

Example 6

	... discussion of several designs...
UI	I WANTED FEEDBACK | I think we need to rate these | BUT WE'LL SEE WHAT YOUR UH PERSONAL PREFERENCES ARE

Example 7

PM	*[...]* so, | MISTER MONEY, WHAT'S YOUR OPINION ON THIS REMOTE CONTROL? |
ME	we gonna try to measure how good it is instead of just talking about it

COMMENT-ABOUT-UNDERSTANDING is for the very specific case of commenting on a previous dialogue act where the speaker indicates something about whether they heard or understood what a previous speaker has said, without doing anything more substantive. In a COMMENT-ABOUT-UNDERSTANDING, the speaker can indicate either that they did understand (or simply hear) what a previous speaker said, or that they didn't. This class is quite limited, since if the speaker does anything beyond comment on whether or not they understood, the act counts as something else. For instance, if they ask for clarification, that would be some type of eliciting act (depending on what they need clarified), or if they express a reaction to something they understood, that would be an ASSESS.

In an ELICIT-COMMENT-ABOUT-UNDERSTANDING, the speaker attempts to elicit a comment about whether or not what has been said or done so far has been understood, without further asking for assessment of anything in the discussion.

Example 8

	B is describing an idea about the remote control functions
B	like three mentals states, yeah, you know what I mean, we can just make it uh
	[other participants acknowledging]
B	controlled by a brain, | HUH?

Adjacency Pairs in the AMI Dialogue Act Annotation Scheme

When people speak, they are often responding to something someone has said, something someone has done, or something around them. As you classify each dialogue act, you should consider whether or not it is a response to something, and if it is, indicate that relationship by adding a link. For instance, consider the case of two dialogue acts, the first of which is a question (an ELICIT-INFORM), and the second of which is the answer (an INFORM). The answer occurs in response to the question, so you need to link the two together, using the question as the source of the link, and the answer as the target. This is the most usual kind of relationship, but there are other possible acts that could be related to the question, too. Refusing to answer the question, asking for it to be repeated, and saying that one doesn't know the answer are all possible responses, and therefore would require annotation.

There will be cases when you may want to mark a relationship, but will not be able to identify a dialogue act as the source. This can happen for several reasons. The source may be something that isn't expressed verbally, such as for assessments of physical objects:

Example 9

	...upon picking up a whiteboard pen and stepping up to the whiteboard for the first time...
C	Okay | VERY NICE | alright

Or the source could be more than one previous dialogue act, as in assessments of entire discussions.

Example 10

C	Should we maybe make a decision about what features we actually want to include, | 'CAUSE WE'RE THROWING A LOT OF FEATURES ONTO THE TABLE.

In these cases, the annotator is instructed to leave the source of the relationship unspecified.

As with ELICIT-INFORMATION and INFORM, we find relationships that have ELICIT-ASSESSMENT or ELICIT-COMMENT-ABOUT-UNDERSTANDING as source and ASSESS or COMMENT-ABOUT-UNDERSTANDING as target, and between ELICIT-OFFER-OR-SUGGESTION and either OFFER or SUGGEST. But the we also find relationships where the target is an ASSESS and the source is just about any other substantive act, and among various OFFERs, SUGGESTs, and INFORMs, as the group discusses what to do.

As well as a source and a target a relationship, a relationship type: POSITIVE, NEGATIVE, PARTIAL, or UNCLEAR.

POSITIVE means the target supports the intention of the source, for instance, by reacting positively to it, accepting or agreeing with it, indicating it has been understood, or providing what the source is attempting to elicit;

NEGATIVE means it rejects the source, for instance, by presenting an objection to it, countering the source with an alternative the speaker prefers, or refusing to provide what the source is attempting to elicit;

Example 11

A	Mm. So, *some kind of idea uh with um um cellular phone with a a screen that will tell you what,* no .
C	NO, NO SCREENS | it's too complex.

PARTIAL means it partially supports the source but rejects it in some aspects, for instance by agreeing with part of a suggestion or providing part of what the source is attempting to elicit; and

Example 12

	[C is drawing on the whiteboard]
D	A kind of snake? *A cobra?*
C	Yeah, uh | NOT REALLY, | a small cobra.

UNCERTAIN means it expresses genuine uncertainty about the source, for instance, by saying that speaker is unsure whether or not a suggestion is a good idea or whether some information is true, or by expressing an inability to provide what the source is attempting to elicit.

Example 13

C	We can adapt only one switch, suppose here like we can make two switches and if I'm left-hander I use this switch to follow the main operations.
B	*I mean if it's less than three uh then we can make it uh like a...*
D	THREE BUTTONS, YOU MEAN?

Addressing

A speaker may address a dialogue act to the whole group of participants present in the meeting, or to a particular subgroup of them, and also to one single participant in particular. Your task is to mark the addressee(s) **for each proper dialogue act** that is not a BACKCHANNEL, STALL, FRAGMENT or OTHER. The last category includes cases of self-addressed speech, where a speaker mumble to himself or thinks aloud without really addressing anybody. Examples of this would be utterances like "Oops!" (after spilling water on the table), "Now, where am I?" (while trying to find the right slide on the laptop), "I have to get up" (while realizing that he cannot give a presentation from his seat). You do not have to make any distinction between addressing the entire group or addressing a subgroup of participants. In both cases, you should mark dialogue acts as addressed to a group.

Group vs. Individual

In a group discussion, many of the dialogue acts will simply be addressed to the group as a whole. However, at times a speaker shows by verbal or non-verbal behavior that he singles out one participant as the intended receiver of the dialogue act he performs. In such cases, only the participant receiving the primary attention of the speaker is the addressee.

There may be several reasons for a speaker to address an utterance to one specific participant:

1. the speaker expects a reaction, response or an action to be performed from the addressed individual (which does not mean that the addressee takes it that way)

Example 14

PM	to	Can you go to the next slide?
ME		

Example 15

ME	to	From from your side uh, you're gonna have to go back the management
PM		and s- be more s- precise

Example 16

PM	to	Yeah but but end of the day, you're the sales guy, so I will come back and
ID		sit on your head because uh you are going to give your sales projection, okay.

Also, the speaker may explicitly announce to the addressee that he intends to ask him a question before the speaker actually asks the question. For example, "I have a question" can be paraphrased as "I have a question for you".

2. the speaker provides a *direct response* to a previous speaker who requested some information or opinion to be provided primarily to him.

Example 17

PM	to	what's mean exactly, advanced chip on print? What's the meaning of
ID		that?
ID	to	I think it's um um a multiple uh chip design um and it's uh maybe
PM		printed on to the circuit board.

Example 18

UI to ID	Why was plastic eliminated as a possible material?
ID to UI	Because um it gets brittle, cracks -

However, A will sometimes continue addressing B with the dialogue acts that follow such a direct response. In these cases it is not always easy to define whether a dialogue act is addressed to an individual or to the group. Other cues like gestures and body language, or knowing what role each participant plays in the meeting may help define the directionality of the dialogue act in such ambivalent cases.
On the other hand, sometimes it is obvious from the content of the request and

from the conversational context that the previous speaker intended that the selected individual or someone from the group should provide a response to the whole group.

Example 19

PM	to	So what do you think about uh the design?
ID		

Example 20

PM	to	Do you have any other information for us at this stage?
ME		

Example 21

		[the group is making decision what type of functions they would like to have.]
ID	to	Do we need a power button at all?
Group		

Responses to these types of questions are mostly addressed to the group.

3. the speaker provides an evaluation of what a previous speaker has said or done.

Example 22

UI	to	I think if you re- if you use really good quality wood, then it might
Group		work, but you can't just use...
ME	to	No y- no no no
UI		

Example 23

PM	to	Okay, tha- that would be great, so if you find out from the technology
ID		background, okay, so that would be good.
ID	to	sounds good
PM		

In other instances, however, such evaluations can be addressed to the whole group: in a group discussion, participants -sometimes simultaneously- may comment on decisions that they took together, or an idea that has just been suggested.

4. the speaker expresses himself in a positive or negative way towards an individual. For example, the speaker may apologize for something that he said or did to the one that is being addressed (e.g. interrupting), or make rude comments to that individual.

If a speaker is talking about another participant in the third person, perhaps indirectly addressing him, that participant should not be marked as the addressee. For example, if PM says to the group *"So, I ll invite uh Christine to discuss about uh the functional design"*, he is addressing the whole group, despite the fact that he is indirectly asking Christine to begin her presentation.

Sources of information

The speaker may explicitly address a single individual by, for example, using vocatives (*Christine, can you tell us about industrial design*). In most cases, the addressing is not explicit and you should use different sources of information regarding speaker verbal and non verbal behavior as well as the conversational situation, to identify who is being addressed.

The following list contains criteria that you should consider in identifying who is being addressed:

- **Content of dialogue act**- Sometimes the content of a dialogue act can provide enough information to identify the addressee(s) of the dialogue act. For example, *"Okay then, uh, let's move to Agnes"* and *"And finally in this meeting we have to decide together about the conceptual design"* are suggestions addressed to the group, whereas *Yeah but but end of the day, you're the sales guy, so I will come back and sit on your head because uh you are going to give your sales projection, okay"* is addressed to the marketing expert. On the other hand, the content of a dialogue act can give a clue that an individual is addressed, without refereing to that individual. For example, *Do you have any other information for us at this stage?* indicates that an individual is addressed, but the other sources of information has to be considered to identify which individual it is.
- **Conversational situation**- In a number of cases, the person addressed will be the person who last spoke, the person whose speech was the stimulus for the present response. It may also happen that two participants start exchanging information about a certain issue, addressing each other. Two participants may also form a side conversation.
- **Meeting context**- the particular activity taking place can provide clues as to the addressee(s)
  During presentations, most of the dialogue acts are addressed to a group, unless there is a distinctive cue, that indicates that an individual is addressed. For example, the speaker may comment or refer to something that an individual has said previously, giving primary attention to that individual, or he makes a side remark, like *"I have added your figure on the slide"*.

- **Gestures and postures**- Looking at the single participant, or pointing at the single participant sometimes may indicate who is being addressed. But, it was not always necessary to believe that because A looks at B while saying something, he is addressing B. In group discussions a speaker's gaze may be directed to the person who last spoke, or to the person whose line of work is related to what is said or the person who coordinates the discussion i.e. PM, or simply at the person who is sitting in front of the speaker. The speaker's gaze can also function to monitor listeners, to see how they take up his words. If a speaker is monitoring one participant, it does not necessarily imply that he is addressing this participant in particular; he can be addressing the whole group. Therefore, you should not rely on gaze information only to identify the addressee. Gaze can be a supporting criterion but usually not a primary one. The same holds for pointing: the gesture can be used as a non-verbal reference to a participant, who is not actually the addressee.

Example 24

P1 to P3 As she *[pointing at P2]* has just said.

- **Domain knowledge**-The project manager will probably not ask the user interface designer how many of the tested people would like to have speech control.