Eindhoven University of Technology

MASTER

Real-time indexation of meeting recordings

Schindler, E.

Award date:
2008

Link to publication

Disclaimer
This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Eindhoven, February 12, 2008

Real-time Indexation of Meeting Recordings

by
Eugen Schindler

Final version: 1.0.7 (pending review)

in partial fulfilment of the requirements for the degree of

Master of Science
in Computer Science

Supervisors:
R. van Liere, TU/e, Department of Mathematics and Computer Science
J.M.B. Terken, TU/e, Department of Industrial Design

Eindhoven University, department of Mathematics and Computer Science

The work described in this thesis has been conducted at the department of Industrial Design, subdepartment User Centered Engineering, of the Eindhoven University of Technology in the framework of the EU sponsored P6 Integrated Project *Computers in the Human Interaction Loop* (CHIL).

Abstract

Taking notes during meetings has two important benefits: the *process* of note-taking helps memorization and review of the *product* (the notes) helps recall. Unfortunately, the activity of personal note-taking often consumes attention and reduces a participant's active participation in the meeting.

Aiding or replacing personal notes with audio recordings that are indexed for quick browsing may reduce the demand of personal note-taking on the attention of the meeting attendee (attention demand). Fully automatic indexing of audio recordings after a meeting is not only technologically challenging but also takes away the process benefits of note-taking.

We are interested in reducing the attention demand, while preserving some of the important properties of note-taking that may yield encoding benefits. An indexing process that is driven by the meeting attendee (real-time indexation) gives the attendee the most power over the indexing process: during the meeting in real time, instead of or together with, but similar to taking notes.

A working prototype for usability-based evaluation and refinement of real-time indexation concepts has been built. Experimenting with different interaction concepts lead to insights into realistic preferences of users. Consistent feedback from users in evaluations of the prototype yields design guidelines (do's and dont's) for real-time indexation systems, within the context of semi-formal, semi-structured meetings.

The idea of real-time indexing turns out to be worth pursuing. Positive user feedback in initial experiments with interaction concepts for real-time indexation indicates that it may be practically usable to reduce cognitive load. The most important lessons learnt from the initial studies have been summarised to serve as rules of thumb and design recommendations.

Acknowledgements

I would like to thank the people who provided me with guidance, rendered assistance and supported me morally. Without them, this work would not have become what it is.

Many thanks go out to my daily supervisor *Jacques Terken* for his great input, his friendly guidance and support throughout the whole practical work. Moreover, thanks go out to Jacques for his countless useful feedback during the writing of this text. Jacques was the first one to give me feedback on the infant versions of this text and he didn't say "what is this mess" but rather helped me patiently to get it straight content-wise.

My supervisor *Robert van Liere* receives my thanks for his patience (delivering this thesis did take quite some time) and his tips as well as review feedback on the pre-final version of this text.

My current manager at OOTI, *Harold Weffers*, deserves special thanks for giving me lots of tips on writing and for guiding me in the restructuring of this text.

My brother *Klemens Schindler* is entitled to a big "thank you" for his help with some of the tedious scripting work in the browser prototype.

I would also like to thank *Martin Boschman* and *Jan-Roeloph de Pijper* for their technical assistance with the hardware and software in the Kidslab usability laboratory at the TU/e *Institute for Perception Research* building.

Moreover, I would like to thank *my parents* and all the good colleagues and friends everywhere (specifically at *OOTI* and *JCES Kinjin*) for their moral support and motivation.

Finally, I would like to thank *you* for taking the time to read this discourse on Real-time Indexation of Meeting Recordings.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Note-taking has always been an essential part of meetings. Its major beneficial effects are twofold (discussed in more detail in Section 2.2):

- the process of note-taking helps to memorise (the *encoding* benefit), and
- reviewing the product of note-taking (the notes) helps to recall (the *storage* benefit).

Unfortunately, taking notes also consumes quite some attention (*i.e. it imposes high cognitive load*), making it difficult to simultaneously think about or respond to what has been said. Often, one is forced to make a trade-off between "making perfect notes" and "participating in the meeting very actively".

In a formally structured meeting with an agenda and a practised secretary who can devote all his[1] attention to taking notes, one can choose to concentrate fully on active participation in the meeting (losing the benefit of encoding and completely relying on storage outsourced to the secretary). Small and unstructured meetings (e.g. regular student-supervisor meetings at a university) are usually the most challenging: participants must follow the discussion with their full attention and the information density is high, so that it is often necessary to take a substantial amount of notes. This forces participants to do many things at the same time: listening to the discussion, thinking about what to write down, recalling the last few sentences of the discussion, and reacting to the discussion by thinking what to say, and speaking. A step towards rendering assistance in such a situation is to make an audio (and possibly also video) recording of the conversation. However, an audio recording in itself is unstructured and tedious to handle, leaving little storage benefit, let alone encoding benefit.

1.1 Problem Statement

Browsing through or searching in an indexed (e.g. by keywords) audio recording is much faster and offers more storage benefit because the recording is more accessible for review. Most existing indexing research efforts (e.g. the *Augmented Multiparty Interaction* (AMI) project [AMI], and the *eClass* project [BA04]) focus on post-indexing an audio

[1] Throughout the entire text, *he* should be read as *his or her*, *he or she*, and such.

stream by using text that is related to the meeting (e.g. meeting minutes, pre-meeting e-mails, or other texts that are in relation to the meeting). Such indexing processes are fully automatic and the meeting attendee who uses the indexing system has no control over the indexing process, whatsoever. In contrast, a system for real-time[2] indexing of meeting recordings gives a meeting attendee the most power over the indexing process when he can influence it, which is usually during the meeting, similar to taking notes (instead of or in combination with manual note-taking). An example of a real-time indexing concept is to present keywords derived from speech (so called *candidate keywords*) to a meeting attendee during a meeting, which he can choose to either discard or to keep for later reference.

The effort described in this text focused mainly on the following issues:

- Is it feasible to use speech-recognition technology to implement real-time indexing concepts?
- Given that speech-recognition technology works flawlessly, can we use concepts of real-time indexing to alleviate the cognitive load in demanding situations during note-taking?
- Is it possible to mix the encoding benefits of note-taking and the storage benefits of automatic post-indexing in real-time indexing concepts, as illustrated in Figure 1.1?

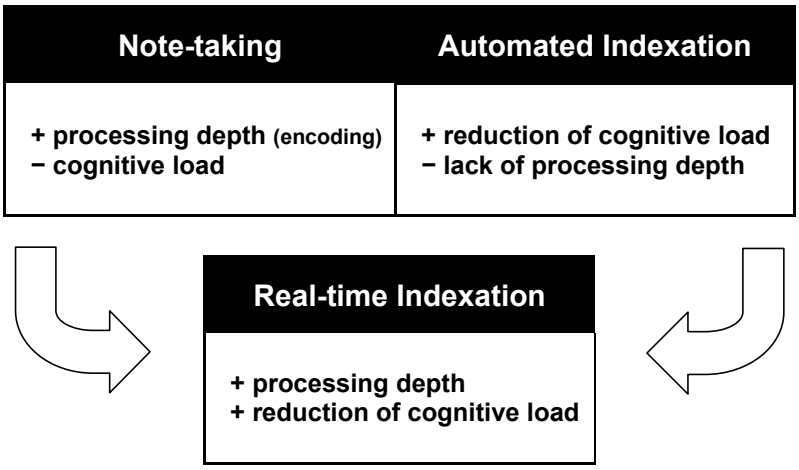

Figure 1.1: Is it possible to combine the benefits of note-taking and automated indexing while reducing the drawbacks?

To investigate the issues described above, we are interested in developing a system prototype for real-time indexing of meeting recordings (as put forward in the original problem statement). Involving users directly by means of user evaluations in the development of the prototype adheres to the user centred nature of the *User Centered Engineering*

[2] In this context, real-time is defined as being within the time scales of the human perception domain. In terms of Butazzo [But97], this would be similar to *firm real-time*, i.e. missing some deadlines is not critical.

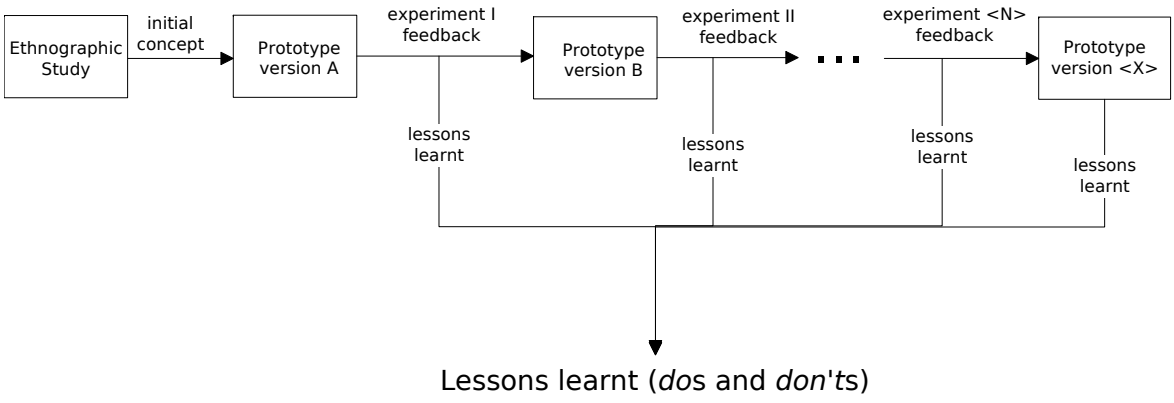

Figure 1.2: Steps and intermediate products in the realisation of the prototype.

subdepartment at the department of Industrial Design. Feasibility of speech technology was of special interest to the *CHIL* [CHI] project.

1.2 First Results and Approach

At the time of creating the prototype and conducting the user evaluations, a practical application of speech recognition technology turned out to be infeasible, mainly due to unavailability of adequate enabling speech recognition technology.

The idea of real-time indexing, however, turns out to be definitely worth pursuing. Positive user feedback in initial experiments with interaction concepts for real-time indexation indicates that it may be practically usable to reduce cognitive load. Consequently, a paper has been written on the subject [TS07] and a derivative bachelor project has been done on the topic. Also, a research proposal in this direction has been written.

The most important lessons learnt from the initial studies (*dos and don'ts* have been summarised to serve as rules of thumb and design recommendations.

Though certain real-time indexation concepts certainly realise reduction of cognitive load, no clear conclusions can be drawn yet on the encoding benefits of the applied real-time indexation concepts, compared to note-taking. Consequently, more longitudinal studies conducted by usability experts are required to draw such conclusions.

The following paragraphs describe the approach taken to reach the above summarised results.

The main idea throughout the project is that the user should be involved in the entire design and research process. In this manner, we do not only learn through qualitative feedback what can best *not* be done, but also what *can* best be done when designing a real-time indexation system.

A working prototype that makes it possible to have a usability-based evaluation and refinement of real-time indexation concepts is necessary for finding realistic preferences of users. Such evaluations can lead to the exploration of different interaction design concepts.

The steps and intermediate products in the realisation of the prototype (as shown in Figure 1.2) were performed in an incremental approach, defining tracks (e.g. requirements gathering, design, implementation). Through system design and implementation as well as user tests, requirements, design, and prototype have been refined to the final version in a number of test-requirements adaptation-design-implementation increments. During the refinement steps, user preferences were collected to derive guidelines for the design of real-time indexation systems.

It should be mentioned that the final version is still *work in progress*, although it has a structure that makes it possible to add new functionality with a relatively small amount of effort.

The first version of the prototype was based on a small subset of the issues that were found as observations/problems in the orientation phase that straightforwardly addressed the missing features of regular note-taking. Before exposing real users to the first version of the prototype, the prototype was first evaluated by a usability expert and me. After a few tweaks, it was used for first user tests.

In order to get some practical insight into note-taking during meetings in addition to the theory that exists on the topic, the preliminary orientation investigation involved collecting ethnographic data (observation reports and survey results) on note-taking practice.

The next versions of the prototype were obtained through incremental refinement based on new insights into technology and use as well as problems with the existing version of the prototype.

The next subsection gives an overview of the realisation of the above described approach.

1.3 Outline

Chapter 2 covers the modus operandi for tackling the posed problem and describes the most important background information that was necessary to perform the work described in the chapters thereafter. First, we sketch the steps that were taken to come from the problem statement to the final results. Next, we position the work done amongst existing work in this area. Then, we summarise the results of a domain analysis of theory and ethnography on human note-taking. Subsequently, we briefly consider feasibility of speech technology application for real-time indexation. Finally, we give a state-of-the-art overview of enabling technologies that may be useful for building a real-time indexation prototype.

Chapter 4 describes the setup and results of the in-the-field experiments that have been done on the real-time indexation prototype which is described in Chapter 3. First, we illustrate the experimentation plan and setup. Then, we describe the results for each of the four experiments.

Chapter 3 elaborates on the current real-time indexation prototype and its development that went hand in hand with the in-the-field experiments described in Chapter 4. First, we briefly characterise the prototype as it is now. Next, we give insight into the different steps that led to the current version of the prototype. We sketch the intermediate major versions of the prototype and show what has changed in time.

In conclusion, Chapter 5 sums up achieved results, further work, and recommendations.

Chapter 2

Analysis

The beginning of knowledge is the discovery of something we do not understand.
- Frank Herbert (1920 - 1986)

Abstract

This chapter covers the modus operandi for tackling the problem described in Section 1.1 and describes the most important background information that was necessary to perform the work described in Chapters 4 and 3.

First, we sketch the steps that were taken to come from the problem statement to the final results. Next, we position the work done amongst existing work in this area. Then, we summarise the results of a domain analysis of theory and ethnography on human note-taking. Subsequently, we briefly consider feasibility of speech technology application for real-time indexation. Finally, we give a state-of-the-art overview of enabling technologies that may be useful for building a real-time indexation prototype.

2.1 Related Work

In the past and present, quite some research has been done on improving pen-and-paper note-taking by means of recording and computing technology. This section describes a number of such efforts and the relation to the work done here to place it in perspective and to give a general idea of feasibility in this line of research. Complementary, Tucker and Whittaker [TW04] have given an extensive taxonomy of meeting browsers. We should stress that most of the work done in this area was focusing on review aspects and technology aspects but not particularly on reducing cognitive load or integrating/considering encoding benefits.

The eClass [BA04] project was concerned with the design and study of an automated capture and access system that helped to capture the materials presented in college lectures for later review by students. Although note-taking in meetings is somewhat different from note-taking during lectures, still many base principles of note-taking apply in the same way. In the controlled environment of lectures and exams, it was possible to test the access of the system at certain periods (e.g. during lecture periods or shortly before the exam) and the performance of students who used the system versus students who did not.

Other work done within the scope of the EU sponsored P6 Integrated Project *Computers in the Human Interaction Loop* (CHIL) has been concerned with topics such as detecting which speaker in a meeting is currently speaking or how much attention the central speaker is getting from other attendees. Integration of the work done here with one of the other CHIL efforts is beyond the scope of this text.

Similarly to the CHIL project, but more specifically aimed at meeting content, the Augmented Multiparty Interaction (AMI) Project [AMI] is concerned with the development of meeting browsers and remote meeting assistants for instrumented meeting rooms - and the required component technologies. For research purposes, an instrumented meeting room has been created that can be used for experiments and development purposes by parties that are working together on the AMI project. Currently, the project has its own prototype meeting browser: the Ferret meeting browser, which can be found on the MultiModal Media (MMM) File server [MMM]. Similarly, the Distributed Meetings [CRG+02] system concentrates on recording and distributing meeting data, whereas LiteMinutes [CBGK01] concentrates on integrating meeting data in multimedia-rich web pages. Whereas research in the scope of the AMI project and the other systems is mainly concerned with making working instrumented rooms for facilitating the meeting and the recording of its content, the work done here is more directed towards interaction concepts for personalising the capture of content and the benefits of this to users (attendees who try the interaction concepts).

More compact systems that do not require instrumented meeting rooms have been made, based on recording audio and timestamping the audio with the writing of notes. One of the earlier attempts was Dynomite [WSS97], a piece of software on a regular PC that can record handwriting and audio. Another, more paper-oriented attempt is the Audio Notebook [SAS01] which records audio and indexes the audio with the handwriting on paper, using a pressure-sensitive sensor to detect writing activity. Filochat [WHW94] provides a functionality similar to Dynomite but providing an LCD tablet for writing by hand. The kind of indexing as described above can be referred to as “ink indexing”. The work done here may at some point use the principle of ink indexing, albeit in another form thanks to the existence of digital pens (see next section), but only in combination with (emulated) speech recognition to possibly find a correspondence between sentence fragments generated by the speech recogniser and manual writing.

NotePals [DLC+99] concentrated on the easy sharing of handwritten notes. Using PDAs as medium for handwriting, digitisation was made implicit. The main point was on combining notes in a “whiteboard” fashion, so the use of collaborative note-taking would be maximised. Handwritten notes of multiple users would be shown in many small screens at the same time in a meeting browser. The combination of notes is not the main point of attention of the work done here. If recordings from different attendees are to be combined, this may be done by overlaying and color-coding different markers.

2.2 Human Factors in Note-taking

Literature on note-taking consists of theories on benefits and cognitive processes as well as ethnographical insights into human note-taking. An important part that we consider ethnographical are note-taking practices (what tools are used, what are notes used for,

when are personal notes taken?). Since personal note-taking practices are not described in literature, we did an investigation in the form of a questionnaire and complementary in-the-field observations of meetings.

2.2.1 Theory on Note-taking

Benefits and problems of note-taking as well as the theory of cognitive processes underlying the activity have been studied for a long time. Most studies on this topic have been done in educational environments, observing students who take notes from lectures [EMS85, Har78, Mee91, Kie89, Kie85]. Some studies are of a more general nature [POK04, CT75, CL72, Dun86].

Hartley and Davies categorised 80 studies on note-taking that were done over a period of 55 years [Har78]. Results of most of these studies are remarkably consistent in that note-taking leads to enhanced performance in subsequent recall of information. In fact, all experimental results agree that taking notes on a point doesn't guarantee its being recalled, but failure to take notes of it very greatly decreases its chances of being recalled, following some delay and no opportunity to review [Kie89, Har78]. Moreover, even if it appears to make little difference to immediate recall whether notes are taken or not, note-taking is useful in the long run, particularly if revision is carried out.

DiVesta and Gray (cited in [EMS85, Kie89]) distinguished between two functions of note-taking: *storage* and *encoding*. Storage means that note-taking facilitates retention of information by providing a form of external storage that can be used for review. Encoding suggests that note-taking leads to activities during learning that are beneficial to memorisation and/or transfer (the influence of prior learning on subsequent learning).

Similarly, Hartley and Davies [Har78] describe this dichotomy as the *product* and *process* benefits. The process benefits are in the fact that writing down things in your own words somehow helps their subsequent recall. Product benefits lie in the created notes: these are useful for revision at a later date. Notes not only provide a record of what was spoken but also enable to recall and reconstruct originally discussed material.

Noted information gets processed more thoroughly than non-noted information, thus recall is improved. In studies resembling classroom conditions, both the encoding and review functions are supported [Mee91]. However, achievement of encoding is usually higher when notes are reviewed (storage is more beneficial).

Storage facilitates retention, i.e. [Kie85] information gets consolidated, previously unrecorded information is reconstructed, the natural process of forgetting staved off and forgotten information is relearned. Twenty-four studies indicate that review yields a higher performance than non-review. 8 studies indicate that there is no difference and no studies indicated a negative effect of review [Har78, Kie89].

The influence of note-taking on encoding of information is two-fold: *quantitative* and *qualitative* [EMS85]. The quantitative effect of note-taking is about *how much* information is encoded (e.g. note-taking may be effective because it increases overall level of attention or effort of subjects). The qualitative effect implies that note-taking *affects the nature* of information processing, increasing recall of information, i.e. note-takers engage in processing that is different from normal listening (e.g. relating ideas to one

another). The effect of note-taking on encoding appears to be larger for far-transfer tasks (e.g. problem solving) than near-transfer tasks (e.g. fact retention), since far-transfer problems require much deeper comprehension of the general concepts of the incoming information [Kie89, EMS85]. The encoding function aids in learning and retention by activating attentional mechanisms and by engaging the learner's cognitive processes of coding, synthesising and transforming aurally received information into a personally meaningful form [Mee91].

Particularly, the benefits of note-taking result from heightened activation of several cognitive processes [Mee91]:

1. selection: actively attend to the message/discussion and select important ideas to retain in the notes
2. relating: paraphrasing and adding own comments, relate own prior knowledge to the new information
3. deep processing: elaborating on content by paraphrasing, indicating relationships among ideas, and developing own examples leads to a more deep processing of content

Note-taking actually requires comprehension similar to original composition [POK04].

Taking notes involves juggling comprehension and production processes under, at times, severe time pressure [POK04]. The note-taker is restricted by the speed of speech: average writing speed is 0.2-0.3 words per second, whereas average speaking speed is 2-3 words per second.

In some domains (e.g. structured meetings with a fixed agenda, progress meetings, design meetings [Kah04], or pre-planned lectures) it is possible to employ good preparation and pre-planning which helps to reduce cognitive effort greatly during discussion of planned points. However, when confronted with discussions that have a high information density, in general pre-planning doesn't reduce cognitive effort significantly [POK04].

Instruction on note-taking has no significant effect on achievement as measured in post-lecture criterion tests, but produces notes that are qualitatively and/or structurally superior [Mee91].

2.2.2 Ethnographical Data on Note-taking in Meetings

Whittaker et al [WLT05] followed two teams in different companies during their meetings and observed as well as interviewed the participants. They categorised the four main functions of public meeting records:

- track group process
- serve as a public record of past actions and decisions
- remind people about their commitments
- resolve disputes about commitments

In contrast, they categorised the main functions of personal meeting records as follows:

- as personal reminders
- provide enough contextual information to carry out personal actions
- check the accuracy of the minutes/public records
- brief others about what went on

Some of the shortcomings of public records named by Whittaker at al were as follows:

- occasionally inaccurate
- lack sufficient detail to allow participants to carry out personal actions
- do not provide enough information for non-attendees to determine what went on in the meeting
- not timely
- laborious to produce

Some of the shortcomings of personal records named were:

- taking notes reduces one's ability to contribute to the discussion
- personal notes sometimes lack both accuracy and comprehensibility
- their esoteric nature made them difficult for non-attendees to understand

These observations confirm many of the conclusions from studies performed within the context of this work (see further on in this section). Moreover, the studies revealed additional shortcomings of personal records:

- apart from esoteric keywords, handwriting cannot be re-read sometimes
- chaotic/hard to organise and hard to archive

The previous subsection gives mainly insights into group meeting minutes. However, a number of questions (e.g. how elaborate are personal notes taken during a meeting, or which other tools may nowadays be used besides pen and paper?) are not answered. In order to find answers to such questions, we disseminated a questionnaire on current note-taking practice and conducted a series of in-the-field observations of meetings.

The questionnaire was disseminated during a period of 3 months. The questions covered a broad spectrum of aspects of meetings and note-taking, such as *meeting structure*, *method of personal note-taking*, *tools used for note-taking*, *actual use of the notes*, *archiving/organization of notes*, *information retrieval from the notes*, *problems with own current note-taking practice*, and *sharing of notes*.

From 106 respondents, most ($n = 83$) were students or employees of the TU/e. The following text will give a short overview of main points that pertain to note-taking and

the use of notes. The exact questions, procedure, and detailed results can be found in Appendix A. The contexts in which respondants held meetings, varied broadly. Main categories were: (middle) school education, hospital, association/club, high tech company, museum, bank/trust fund, hotels, sports, government, research, care/nursing, consultancy, I(C)T, and study.

More than 66% of the respondents held a formal, structured meeting (with agenda, chairman and secretary) 1-3 times per month, while almost half of them held an informal meeting (such as a student/supervisor meeting or informal group meetings with two or more people) almost every day. Spoken information during meetings is preserved mainly by writing down keywords which are worked out to full notes after the meeting or simply kept as-is to facilitate recall. Audio- or video recordings are made only extremely rarely. Pen and paper are the prevalent tools used for note-taking, used often to always by 75% of the respondents. From the commonly available digital tools -such as laptops, hendhelds, and dictation devices- laptops (typing in notes) are used the most.

Making notes "just because the note-taking itself helps to remember" is practised a lot less than really using the notes for review (e.g. as a to-do list) in between meetings or during follow-up meetings. Organization of notes differs widely, from not organising notes systematically at all to ordering notes by date, topic or splitting them into to-do points/journal items. Accordingly, searching through notes for specific information varies from simply skimming through to recalling pages/places or using a digital search function. Custom answers indicated that organization and search depends very much on the size, complexity and digital availability of the notes. Mainly, people like to use digital search, but do not want to create digital versions of the notes.

Strong points of note-taking were mainly things like chronological ordering, help to recall, fast and easy to do (though there was quite some disagreement as to which is more efficient - laptop, or pen and paper), and unobtrusive. In some cases doodling was named as a strong point because it helps to concentrate.

Weak points were mainly the sparseness of notes (lack of context), chaotic (difficult to search), difficult to archive, and demanding lots of attention (making it difficult to participate actively in the discussion). More than half of the respondents experienced problems with their notes now and then. Although illegibility sometimes plays a role in these problems, lack of context in recall (i.e. what was discussed at one point or another during the meeting) plays a much larger role: about half of the respondents experienced context/recall problems.

The open question "In what way would you like to have digital assistance for your note-taking?" produced a variety of responses. Around 65% of all respondents really liked the idea of digital support for their current note-taking. Of the people who saw some use in digital support for note-taking, about one half was interested in the easy digitisation of their own notes (such as digital paper, digital pens, handwriting/character recognition, or tablets) during or after the meeting. The other half was interested in transcripts, audio/video recordings, photos, markers in recordings, or searching for keywords in recordings. About 5% of all respondents were really opposed to the idea of digital support for note-taking, being strong supporters of pen and paper and having bad experiences with obtrusive devices such as laptops or handhelds. 30%, give or take, did not provide any response.

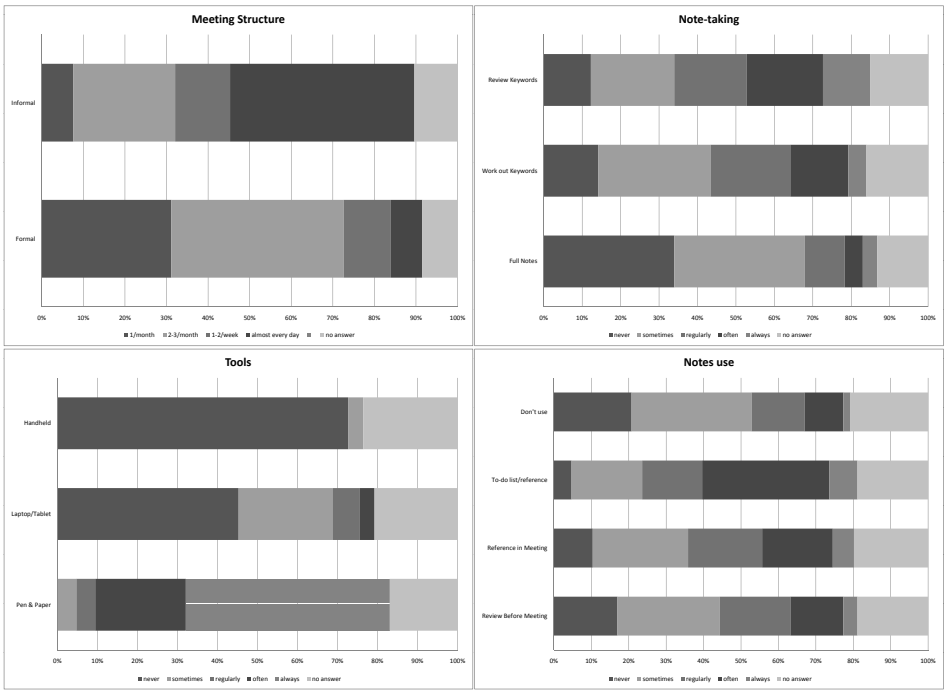

Figure 2.1: Overview of four important survey results

Figure 2.1 summarises the most important results (as summarised in Appendix A. We can see that informal meetings (and thus the need for personal note-taking) are much more frequent than formal meetings with a secretary. Moreover, we can see that keywords for review are taken more frequently than full notes or notes worked out from keywords. As expected, the dominant tools for taking notes are pen and paper. We can also see that the notes are mainly used as reference material during and between meetings. The main strong point of note-taking was ease of use. The main weak points were high cognitive load and sparseness (lack of context in keywords degenerates recall from review quickly).

Although some reference material existed from personal experience during meetings, the reference frame has been widened a bit by doing a number of observations on different groups. Formal and informal meetings have been observed to study the differences. Formal meetings were held in groups of 3-7 people and informal meetings were usually held with two people. The text below describes very briefly some important results, whereas the complete observation reports can be found in Appendix B.

Pen and paper were note-taking tools that were used in abundance. Sometimes a laptop was used for typing notes. In formal meetings there was usually a dedicated secretary. In some cases, the secretary was perfectly able to keep up with the note-taking and even participate well in the meeting. In other cases, the secretary even had trouble with the note-taking itself and had to slow down or stop the meeting several times to complete note-taking tasks on a certain point. Since the secretary was dedicated, in formal meetings, normally the discussion could continue when the secretary was able to keep up or didn't stop the meeting if he was not able to keep up. Consistently with the results from the questionnaire, there were always people who took personal notes, even during formal meetings, regardless of the fact that there was a secretary who took

notes for the group. During informal meetings, sometimes the entire dialogue stopped whenever notes had to be taken, which indicates that the note-taker has a very low active participation while taking notes and the other participants acknowledge this.

Summary When the information density in a discussion increases, note-taking can become very demanding and hinder active participation in the discussion. Apart from structural shortcomings of note-taking (such as inaccuracy and difficulty to organise), big problems were lack of context in notes (people were forced to omit information in demanding situations) and demand on the attention (people had to choose between taking notes and participating actively in the meeting). Yet, the questionnaire suggests that review of the written keywords (during and in between meetings) is the widest use of personal notes. Also, people tend to take personal notes even in the presence of a dedicated secretary (reasons for this have been found above in the literature on ethnography of meetings). If this note-taking reduces the active participation of attendees, real-time indexation might be a very good option.

2.3 Feasibility of Spontaneous Speech Recognition

Searching and/or indexing audio is referred to as spoken document retrieval. A more general name that not only entails the retrieval of information from spoken documents such as newscasts or dictations is *speech-based information retrieval* (despite the suggestion of the name, this is not about using the speech modality in a user interface to query a system). The NIST Speech Group [NIS] evaluates and keeps track of the latest speech-based technology. Though speech recognition makes great progress and there are a number of success reports for spoken document retrieval [GAV00], problems with spontaneous speech (as encountered during meetings) are a special challenge to speech technology [FUR]. One of the major fundamental reasons for the drastic decrease in performance for spontaneous speech is that most currently used acoustic and language models have been built using written language and speech from texts (e.g. newscasts). Where speech from a text is very “clean”, spontaneous speech includes filled pauses (uh, uhm), repairs (revision of what has just been said), false starts (a speaker interrupts the flow of speech to restart an utterance), hesitations, repetitions, partial words and possibly other kinds of disfluencies. Some other fundamental challenges are the recovery of hidden punctuation, realistic turn-taking (in case more than one speaker is involved, with implications such as overlapping speakers or prosody used to find the end of one speaker’s turn), and detection of sub verbal information such as a speaker’s emotion [Shr05].

Although speech is almost always spontaneous, the research area of spontaneous speech acoustic models that take the mentioned disfluencies into account and the obtainment of canonised training data for practical application is relatively very young. Hence, performance for practical use of spontaneous speech recognition is not yet as acceptable as speech recognition of dictation. If speech recognition is used offline (not in realtime), techniques such as word spotting or vocabulary optimization can be used to improve performance [Lan]. Such indexation, however, takes around ten times the length actual

recording. Therefore, such methods are infeasible for real-time recognition of spontaneous speech. However, recent developments in large vocabulary dictation systems (such as Dragon 9, which was not available at the time the practical part of this work was conducted) are very promising at providing at least a rudimentary form of real-time speech recognition that might be used for real-time indexation.

2.4 Enabling Technologies

Apart from mobile devices (PDAs and TabletPCs) with complementary wireless communication technologies (Bluetooth and WiFi), and input devices (touchscreens, digital pens - see Section2.4.2) that may be useful for unobtrusive and efficient interaction mechanisms, the most important enabling technology for keyword generation is that of speech recognition engines (see Section 2.4.1).

2.4.1 Large Vocabulary Continuous Speech Recognition)

Nowadays, continuous speech recognition is prevalent in most speech-related applications, since it is much more natural to speak whole sentences instead of making a pause after each word. Large vocabulary continuous speech recognition (LVCSR) systems are not to be confused with the acronym for some projects that work on spontaneous speech: Large vocabulary conversational/spontaneous speech recognition systems. A large vocabulary nowadays typically consists of more than 50.000 words.

The two prominent commercial speech recognition software packages currently on the market are Nuance Dragon Naturally Speaking 9 [Drab] and Nuance/IBM ViaVoice 10 [Via] (IBM recently transferred or sublicensed ViaVoice to Nuance but kept IBM Embedded ViaVoice [EVi]). On a side note, Nuance also provides a special high-performance transcription software package called Dragon Audiomining [Draa] which makes it possible to quickly index speech in XML. At the beginning of the practical work described in this text, development kits for the commercial LVCSR systems were very expensive and difficult to obtain. Currently, all of the Dragon products from Nuance have complementary software development kits, which make it possible to embed the Dragon engine in any PC application and tailor some parameters to one's own needs. Recently, Dragon Mobile APIs [NMS] make it possible to utilize Nuance speech recognition in PocketPC compatible devices.

In the past, the two commercial systems were already quite comparable [Bro02]. Currently, performance in terms of word error rates is converging for most LVCSR systems and with ViaVoice under Nuance, the systems may converge even more. Nuance reports accuracy rates of 98-99% for Dragon Naturally Speaking 9. Some quick dictation experimentation done while writing this text, using an English version of Dragon 9 confirms accuracy of 95% and more, even for non-native English speakers. An experiment on spontaneous speech, using English versions of Dragon on recorded speech data from real meetings of natively Dutch speakers who spoke English with a minimum amount of training time, revealed that Dragon 8 had an accuracy of less than 20%, whereas Dragon 9 had a much more promising accuracy of more than 60% (pure count, not taking into

account disfluencies). Using the software development kit, it may be possible to improve this a lot, tuning the engine to partially take account with more natural speech.

Recently, Microsoft Windows speech recognition has improved dramatically. Users report accuracy rates of easily more than 90%. The great advantage of the Microsoft engine is that the Microsoft Speech software development kit 5.1 is free and can be obtained very easy by download. This makes it a very good candidate for experimental development and may be useful in further work.

Finally, many domain-specific products exist. A few examples are Philips Speechmagic (mainly targeted at legal and healthcare), Nuance SpeechWorks (used telephony services mainly by AOL) and MacSpeech iListen (targeted at Macintosh users only).

In addition to commercial speech recognition software, there are a number of free and experimental software packages, such as CMU Sphinx [Sph], CSLU Sonic [Son] and Julius [Jul]. Most of these systems are useless without a speech corpus on which the engine can be trained, such as the free VoxForge corpus [VFo]. However, such corpora are mostly in a very immature state. Classical corpora can be bought for a relatively low price from research foundations such as CSLU, but these are hardly usable for spontaneous speech. The University of Twente together with TNO have collected a relatively large corpus mainly from the Dutch broadcast network NOS and use this corpus in part for keyword spotting in meeting recordings.

2.4.2 Digital Pens

Digital pens are very useful unobtrusive input devices that enable people to simply take notes as they are always used to. The main use of digital pens is to have written notes available in a digital format without having to scan paper. Another useful feature of digital pens is that it is possible to capture at which time a certain note has been taken. This can be used as an implicit way by people to index an audio- or video recording of a meeting. Similar projects have been performed before common-off-the-shelf digital pens have existed (see earlier in this chapter). Combination with speech technology may present interesting possibilities.

Digital pens such as the Nokia SU-1B [NOK], Logitech Io2 [LOG] or the Mobile NoteTaker [PEG] are available as affordable off-the-shelf components. One disadvantage of these pens is that note-taking is mostly done offline, so indexing cannot be done synchronously online. However, the recently released PC NoteTaker pen [APC, PEG] can capture and transfer notes to a PC in real-time, which brings indexing that is synchronous with speech recognition within reach.

2.5 Summary

This chapter reported on the approach that has been taken to solve the problem described in Section 1.1 and listed investigation results which are necessary for carrying out the work described in Chapters 4 and 3.

Related work mainly deals with specific technical challenges, public records, and transcribing or post-indexing. The main issues addressed in Section 1.1 (i.e. real-time indexation and encoding benefits) are not addressed in related work.

The existing body of ethnography on human note-taking has been extended with quantitative data (mainly on note-taking tools used) and concrete insight into personal note-taking behaviour.

Feasibility of applying speech recognition engines for real-time indexation has been investigated and the practical use of speech recognition software has been found infeasible for use in in-the-field experiments. However, new developments in (commercial) speech technology may currently (or in the near future) justify a new feasibility study for real-time indexation.

Because a recent state-of-the-art overview on the given topics was not only of interest for this work but also for other projects within the CHIL framework, this chapter has been updated to the current state-of-the-art. Consequently, investigative work has not only been done at the beginning, but also while writing this thesis. Therefore, some of the newest material described here was not used or mentioned in the description of the (practical) work done.

Chapter 3

Prototype

"Innovation is 1% inspiration and 99% perspiration."

adapted from: "Genius is 1% inspiration and 99% perspiration."
–Thomas Alva Edison (1847 - 1931)

Abstract

This chapter describes the current real-time indexation prototype and its development that went hand in hand with the in-the-field experiments described in Chapter 4.

First, we briefly characterise the prototype as it is now. Next, we give insight into the different steps that led to the current version of the prototype. We sketch the intermediate major versions of the prototype and show what has changed in time.

3.1 Overview Current Version

Figure 3.1 gives a global picture of how the current prototype is built. The key point is the co functionality of the *Recording Orchestrator* and the *Meeting Browser*. The Recording Orchestrator together with the devices connected to it (such as the Audio Recorder, one or more Wizards or Speech Engines, and one or more User Interfaces such as a PDA running the annotation software constitutes the *annotation subsystem*.

To get complete feedback from users, both inseparable concepts of capture and review were developed. Thus, the current prototype consists mainly of two subsystems:

- the annotation subsystem which is used by a meeting participant during the meeting to indicate which information may be important for preserving;
- the browsing subsystem which is used by a meeting participant to review the information that he has preserved.

When the Audio Recorder (assuming we have one recorder or a group of recorders working as one) is set to start recording, it sends a synchronisation signal to the Recording Orchestrator, which starts to count time. After this, the Recording Server starts accepting messages from User Interface devices and from Wizard/Speech Engine devices.

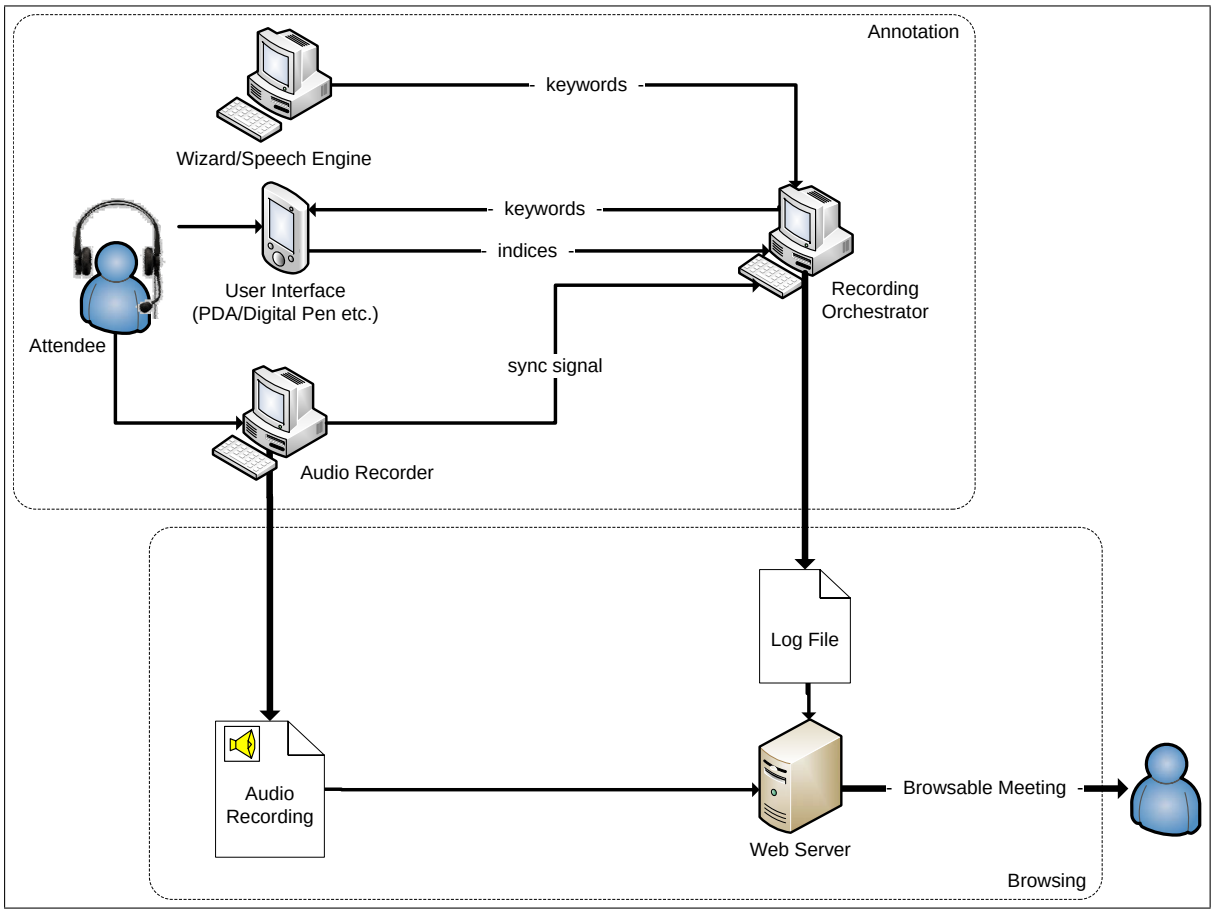

Figure 3.1: System overview of the current prototype.

Every time the server receives a message, it stores the message together with a time stamp.

The total product created by the annotation subsystem consists of one or more audio files (one audio file for each speaker) and a log file that holds all received messages together with their timestamps.

The audio file(s) and log file can simply be placed on a web server in a directory together with the browser software. The browser software reads both files and becomes a browsable version of the meeting in the form of a web page.

A more technically detailed description of the current prototype can be found in Appendix C.

3.2 Development Process

This chapter describes the iterative development of the prototype that has been used to evaluate concepts of real-time indexation.

The following sections describe successive iterations in which the prototype has been refined. The description of each of the iterations has the following structure:

- the (new/changed) *user interface* of the prototype is described
- the (new/changed) *functionality* of the prototype is described

Before releasing the first version of the prototype to users, the user interface was audited by dr. Jacques Terken and reviewed by the author. After each evaluation session, the understanding of the interaction concepts was adjusted and the prototype was changed accordingly.

3.2.1 Annotation Subsystem

3.2.2 Prototype A

The first two iterations were to get a rough feel for a system that allows selection of keywords on simple everyday hardware. The user interface was very simple without any

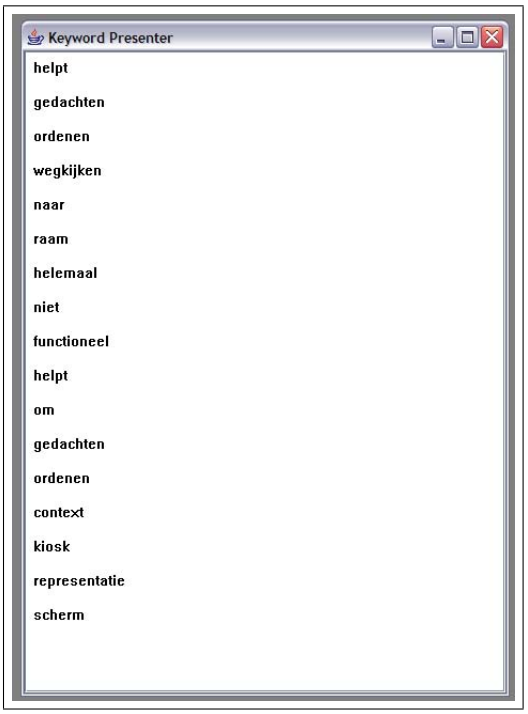

Figure 3.2: First working prototype of the indexation interface: the user can choose the keywords he wants to preserve from the list

functionality, like a mockup (see Figure 3.2), run on a laptop. A list box window shows candidate keywords that can be clicked in order to create an index that corresponds with the pertaining keyword.

The first audit revealed that it may be useful to visualise the processing of a selected keyword. It was proposed to remove the keyword when it has been processed. Furthermore, the list should always scroll to the newest added keyword.

Immediately after the creation of the mockup interface, functionality was added in the form of a backend for the wizard that made it possible to simulate speech recognition.

Whenever the wizard finishes typing a word, that word is sent to the mockup. The wizard back end and the mockup were used together to evaluate the first crude approximation of the concept of presenting candidate keywords from which the user could choose.

3.2.3 Prototype B

Figure 3.3: Indexation interface implemented on a PDA device: the user can use his stylus to choose keywords from the list for preservation or he can write/type a keyword in the text box for later reference

To make the user interface less obtrusive and reduce the cognitive load from the use of the mouse, a PDA was used instead of a laptop (see Figure 3.3). The mockup user interface was rebuilt on the PDA and an audit indicated that it may be useful to present candidate phrases instead of candidate keywords, since keywords were too fine-grained and flooded the user with input. Phrases also provide more context (similar to the Adobe Reader search), which might make it easier to choose desired content for preservation.

3.2.4 Prototype C

The “Index” button (previously used as an alternative to tapping a phrase) has become the “Record” button, which makes a manual time stamp when pressed. The presentation window for key phrases was split into two columns so that the list box would not move items anymore and the candidate phrases could be browsed like pages in a book. The direction to which phrases would be added (top or bottom) could be adjusted as desired by the user. During the evaluations, it was suggested to users that they may also use a handwriting recognition facility to add their own keywords, if desired.

Figure 3.4: On the right side, a more refined version of the indexation interface: the list of keywords is spread over two pages, so that the user doesn't get disturbed by the system adding new keywords. On the left side (to compare), the one-page version.

3.2.5 Prototype D

A "simple mode" has been added with only one big "Record" button that creates a manual time stamp when pressed and every important action has a feedback in the grey status bar (see Figure 3.5. All of the evaluations have been done in simple mode with some variations (see Appendix C for more detail).

3.2.6 Browsing Subsystem

Browsing has been introduced in prototype D. The browser has been made in such a way that it is easily accessible from everywhere (a website). All audio files and timestamps are loaded from the same web server. When a phrase is clicked, the audio starts playing from the corresponding times tamp for a duration corresponding with the phrase duration (see Figure 3.6). The find function makes it possible to search for a certain phrase and the black triangles are manual indices. The audio stream of every attendee can be turned on or off separately and the *Stop Int.* check box plays from the start of a clicked phrase time stamp until the end of the track or until stopped/paused, if unchecked.

The major improvements suggested by the audit of the first version were:

- add "padding" around the start- and stop points (timestamps) of a phrase, so that there is an offset to both sides (start earlier and end later)
- make a second seek bar that zooms in on a part of the first seek bar

The Stop Int. check box has been inverted and called *Continuous Play*. The second seek bar below the first zooms in on the part that is bounded by the black box on the first seek

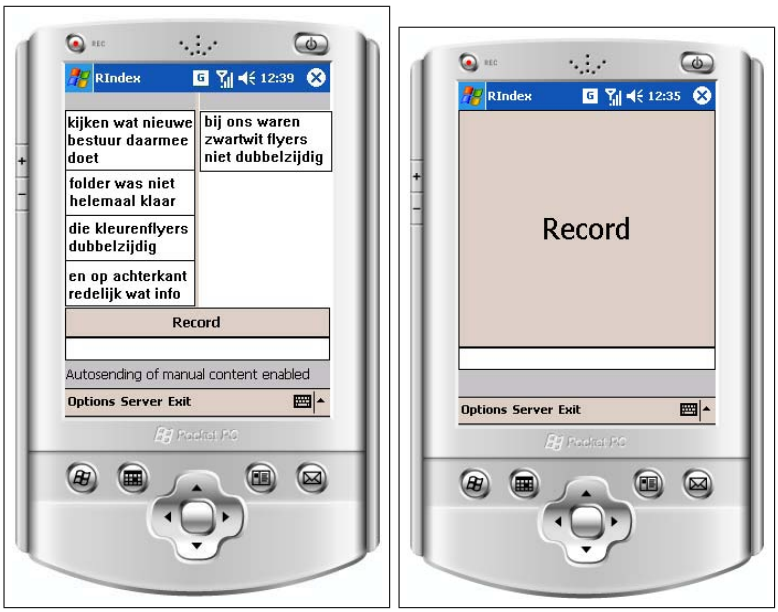

Figure 3.5: Feedback has been added to all actions (grey bar). A “simple mode” with only a “Record” button exists.

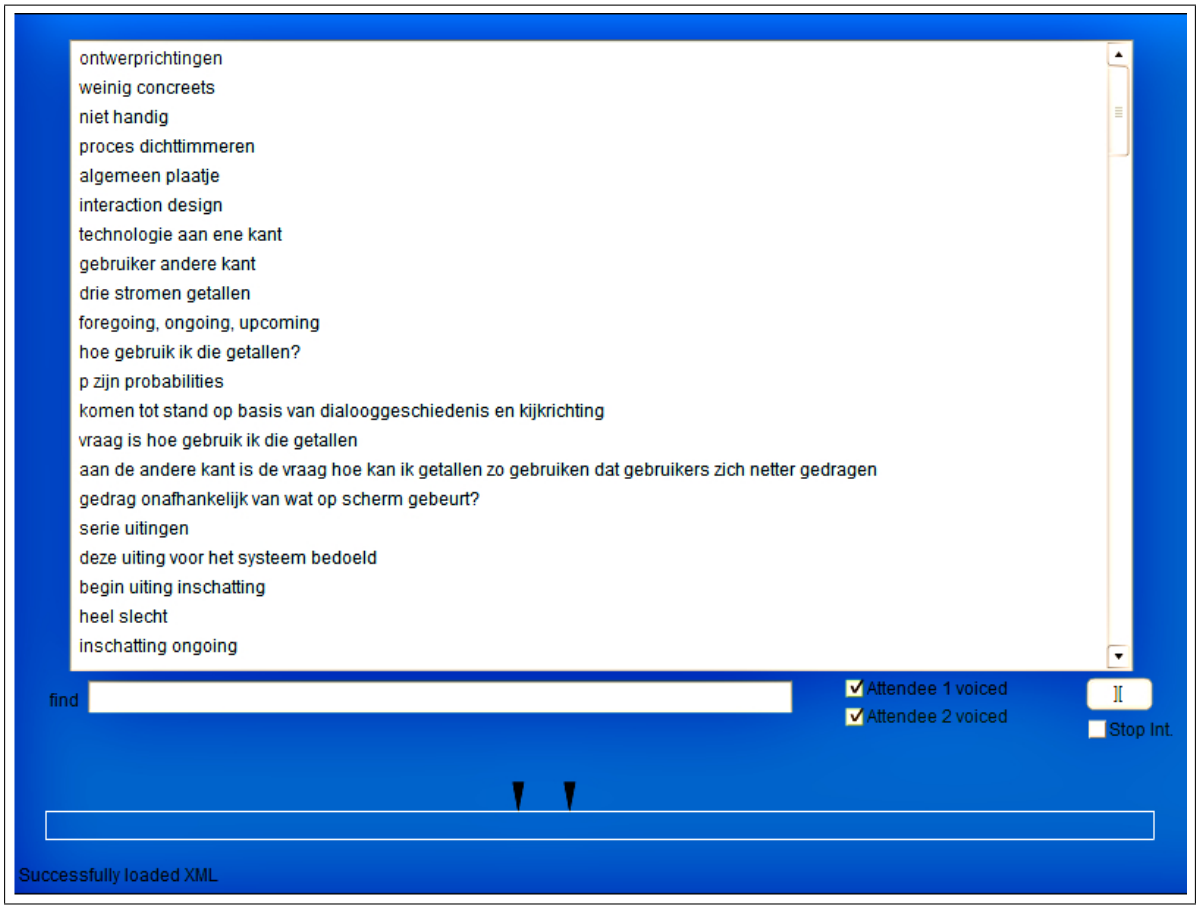

Figure 3.6: The first prototype of the RIMR meeting browser. The time in the audio track is scaled to the fixed-width seek bar in the lower part.

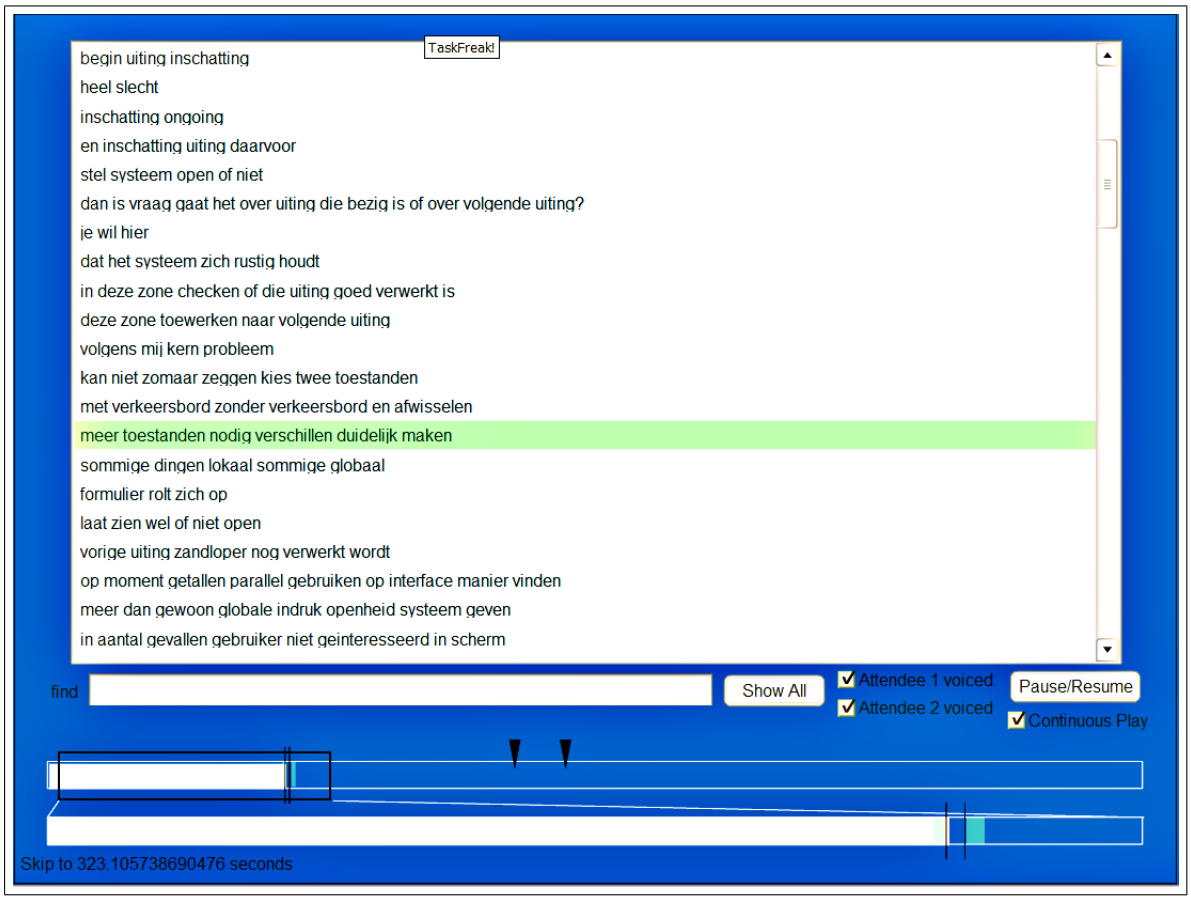

Figure 3.7: The second prototype of the RIMR meeting browser. The lower seek bar is a 'zoomed-in' representation of the part that has been marked with the black box in the upper seek bar. Events and status are more clearly reported to the user in the lower left status bar.

bar. The padding on the phrases is visualised and names of the buttons are made clearer. Also, keys have been added for playback controls.

3.3 Summary

The current prototype is a simple medium-fidelity demonstrator for initial in-the-field experiments in the area of real-time indexation. However, the structure of the prototype is not just a throw-away draft. It is a rough infrastructure that enables creation or adaptation of prototypes with different hardware and software component configurations with relatively little effort for performing experiments on speech-based user-interaction concepts such as real-time indexation of meeting recordings.

Chapter 4

Experiments

You cannot acquire experience by making experiments. You cannot create experience. You must undergo it.
- Albert Camus (1913 - 1960)

Abstract

This chapter describes the setup and results of the in-the-field experiments that have been done on the real-time indexation prototype which is described in Chapter 3.

First, we illustrate the experimentation plan and setup. Then, we describe the results for each of the four experiments.

4.1 Method

Meetings studied during the experiments are in-the-field (and thus not very controlled). Therefore, the evaluations are mainly restricted to qualitative feedback from users.

The users who evaluated the system consisted of students, PhD students and professors from the TU/e. In total, there were 6 groups (of about 2-3 people) who did evaluations. The meetings were mainly (PhD) student-supervisor meetings. All meetings were held in the *Kids Lab* Usability Laboratory of the Institute for Perception Research (IPO) building at the TU/e. Before the beginning of each experiment, consent for audio recording was given verbally by each participant and it was explained that the audio recording would only be used for research purposes within the scope of the current work. During the meetings, users were not forbidden to write notes manually in addition to using the prototype, so people could take notes at all times.

Figure 4.1 depicts a typical formation during the experiments. Every meeting participant wears a special microphone that records only *his* voice. At least one of the participants uses the annotation interface (depicted by a handheld device) of the real-time indexation system.

We have four different experiments (the results of which are described in Section 4.2, with four corresponding major prototype versions (described in Section 3.2).

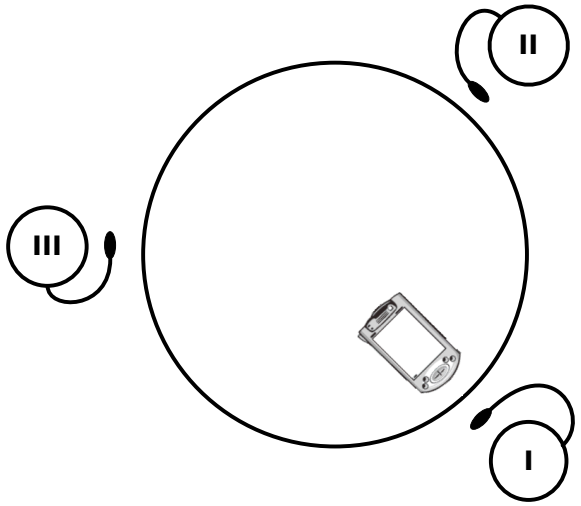

Figure 4.1: A typical experimental setup. See also Figure C.5 in Appendix C.

Experiment	Prototype	Changes
I	A	Smaller view screen and less attention demanding input device
II	B	Reduce load of scrolling candidate keywords and create a browser
III	C	Another concept: index button; finish creating browser
IV	D (current)	-

Table 4.1: The experiments and major versions of the prototype

For each different experiment, we indicate on which version of the prototype it has been conducted and we describe relevant feedback that was given by the users of the prototype. It should be mentioned that the experiments have been conducted mainly in the context of semi-formal student-supervisor meetings. Therefore, although carefully applicable to different situations, the context of the lessons learnt is dictated by the nature of the small, informal meetings without minutes or an agenda.

Table 4.1 gives an overview of the major prototype versions and the major changes for the following prototype version.

Two types of evaluation sessions for every version of the prototype were intended to collect feedback from users and a usability expert:

- Expert audit: a usability expert assesses the prototype and indicates different problems or possible improvements (without really using it).
- User test: users test the prototype in the field. This means that two or more users have a real life meeting and at least one of them is asked to use the prototype during the meeting. After trying out the prototype, the user gives (qualitative) feedback in a face-to-face interview.

The prototype has two aspects of design and implementation. The first and most directly important aspect for the usability evaluations is the system, as users experience it. The second aspect is the structure of the system. Both aspects have been developed in each iteration. Properties of the system as users experience it (especially user interfaces) have been changed by user feedback, whereas properties of the structure have been changed due to better insight of the developer/programmer into technology, possibilities and infrastructure as well as identification of problems with extension or bottlenecks in system performance.

In some cases, a user gave feedback about a certain system property that conflicted with the feedback of another user about the same system property. This kind of feedback has been incorporated into the next iteration by adding an option for each corresponding wish to that particular system property.

4.2 Results

4.2.1 Experiment I

This experiment has been performed on the prototype version described in Section 3.2.2.

User feedback indicated that a laptop or PC monitor is way too big and too obtrusive. Using the mouse imposes a high cognitive load, diverting a lot of attention from the meeting to operating the prototype.

4.2.2 Experiment II

This experiment has been performed on the prototype version described in Section 3.2.3.

The most important remarks from the user evaluation were:

- when used in addition to normal note-taking, the constant switching between pen & paper, and selecting keywords on the PDA imposes a high load and reduces attention
- using the PDA exclusively (abandoning manual note-taking completely) does not completely leverage the problem of high cognitive load: especially talking and choosing words imposes too high a load; however, users themselves indicated that this may be because it takes getting used to using the PDA, just like it takes effort to learn taking good notes, so results may be different when the use of the PDA is better practised
- most users did not dare to abandon manual note-taking, since they did not know how, when or whether the recording would be made available to them; this makes it very important to have a corresponding browser that makes the meeting data available almost immediately, since otherwise users do not trust a digital assistant
- keyphrases were moving on the PDA screen too fast, depriving the user of having a good look at the phrases. An idea to reduce this problem is to use a “two page” system that can be leafed through, like a book

- the direction from which phrases were added (from bottom or from top) was a point of dispute: some users found the top-to-bottom direction good and others found this direction to complicate their cognition, wishing the direction would be inverted
- it was suggested to add a “simply index” button that would make a manual time stamp whenever pressed, making indexation more flexible

4.2.3 Experiment III

This experiment has been performed on the prototype version described in Section 3.2.4.

As the annotation interface of the prototype became more high-fidelity, more useful feedback could be obtained from users:

- although the presentation area for the candidate phrases has been constructed such that phrases are easily browsable and the addition of phrases is very unobtrusive, users experienced cognitive load that was too high above expectation; there is a strong indication that the high load is not imposed by the user interface, but by the reading of the candidate phrases; comments from some users about difficulty in reading the phrases and actively participating at the same time may confirm this
- some users complained about the amount of writing freedom on the PDA: on paper one can easily draw diagrams
- most users who tried to use the handwriting recognition complained that the recognition was very inaccurate
- the visualisation of processing candidate phrases by removing them is confusing; one would think that the item disappears due to selecting it

One user indicated as social implication of using such as system: he felt closed off from his environment, just like if he would be working on a laptop or leafing through a pile of paper all the time, while the meeting partners were waiting. Users also expressed a number of wishes:

- a bigger screen or two PDAs with a shared view for collaborative note-taking(correcting the speech recogniser together with others) - this would also prevent yourself closing off from your environment
- integration with standard packages such as Outlook

An audit revealed that the user interface should give feedback whenever an action has been done, e.g. “manual index saved” or “candidate saved”. Moreover a number of inconsistencies in the interface have been pointed out.

4.2.4 Experiment IV

This experiment has been performed on the prototype version described in Section 3.2.5.

Feedback from users was as follows:

- one has to be aware that the button is there and should be used. When not used to, this imposes extra cognitive load
- the button was sometimes pressed "just in case". This has in part to do with the ambiguous understanding of the "Recording" button. Different interpretations may be for example: record past, record present, record future, recording on/off
- sometimes the button was pushed pushed before something was said (preventive/proactive)
- some users were prepared to do editing in the browser after the meeting if it was not too much effort, such as re-listening the whole meeting

A wish expressed by some users was to have a more refined interface than simply one button, e.g.: a before, important, after and less important button. Other users were very pleased with just one button and did not want more complexity.

The simplicity of the "one button" has another social implication: everybody can see what one is recording. One user wanted therefore to keep the PDA below the table for his own privacy.

People who used the browser after a meeting generally had a good impression of the tool. Conceptually, the browser is quite useful, but there are a lot of points for improvement:

- the position in the audio stream should also be shown in the list of keywords (when a position is clicked, the corresponding key phrase should be marked)
- manual bookmarks (indicated by the black arrow markers that point down) should also be indicated in the list of keywords somehow (e.g. other coloring)
- sometimes users were disoriented because of the jumps on the seekbar
- it would be useful to see which speaker uttered which phrase (e.g. by coloring phrases), since that makes the discussion structure more clear
- the search should be more Google-like, allowing multiple keywords instead of exact matching
- there is no time scale, so no idea how long a record is or how fast one goes forward or back when seeking
- a hierarchical structure division would be very useful (clearly requires advanced text analysis in addition to speech recognition and therefore beyond the scope of this type of research)
- it should be possible to drag the black boundary window on the above seek bar
- it is not clear that the black triangles represent manual bookmarks
- some remarks about colors and layout improvements have been made (mainly that black text on dark blue is not good, that the seek bar and the text box look too much alike and that the listbox with the phrases is too wide; making it narrower would leave more space for the buttons)

- currently the audio is simply played on clicking a phrase; it would be much better to highlight corresponding phrases with the current playback or to scroll phrases according to playback

4.3 Summary

As mentioned in Section 2.4, using speech technology in the in-the-field experiments was infeasible at the time of experimenting. To nevertheless be able to experiment with interaction mechanisms during a meeting, a *Wizard of Oz* approach was chosen in which a person simulated a speech recognition engine.

The most important results from the experiments are:

- Writing keywords down is not the activity that causes the most cognitive load during note-taking. When users had to select candidate keywords from a screen, the visual load of reading the keywords was comparable to the load of writing down keywords.
- Users did not dare to reduce their manual note-taking, since they were uncertain about the availability of the recording and efficient review ability.
- Some users were prepared to do post-processing of the recording after meeting by means of a meeting browser, if this wouldn't take too much time.

Chapter 5

Conclusion

5.1 Problem Revisited

The effort described in this text focused mainly on the following issues:

- Feasibility of using speech-recognition technology for implementation of real-time indexing concepts
- The use of real-time indexing concepts to alleviate the cognitive load in demanding situations during note-taking
- The possibility of mixing the encoding benefits of note-taking and the storage benefits of automatic post-indexing in real-time indexing concepts

To investigate these issues, we are interested in developing a system prototype for real-time indexing of meeting recordings.

5.2 Overview Results

An infrastructure has been created for prototypes for experiments on speech-based user-interaction and a medium-fidelity prototype (see Chapter 3) has been created for an initial study of meeting recording real-time indexation concepts.

The question on the general usefulness of the real-time indexation concept has been answered mostly affirmative. Users of the final version of the prototype have generally reacted very positively to the novel concept, e.g. to the simple question “Would you like to use such a system on a regular basis?” most people answered along the lines of “Yes, please!”, even though many points of improvement have been brought forward and one of the most important problems is that the current user interface in the annotation subsystem of the real-time indexer (involving displaying and selecting of words) imposes a cognitive load that is comparable to the load of manual note-taking. This makes it clear that real-time computer-aided meeting recording is not “just a good idea”, but a useful concept that is worth investigating in more detail.

Lessons learnt from the experiments in the initial study are:

- The user interface for an annotation subsystem (the part that is used by an attendee during the meeting to control the indexing process) should be unobtrusive and preferably small. A big screen i.e. a regular PC or laptop is not a very good candidate for such a user interface since the screen is big and input devices such as a mouse or keyboard impose a lot of cognitive load on the attendee, whereas the goal is to reduce cognitive load.
- It is ideal if the user interface of the annotation component requires little or no visual cognitive load at all. A successful example is the "one button" interface: people can simply keep one finger on the button and don't even have to look at the button to operate it.
- It is important to give users feedback of the success of an indexation action. In case of a user interface that doesn't require any visual load, tactile feedback might be interesting, since sound feedback is not desirable during a meeting. A successful example would be a pushbutton which makes sure you feel a response.
- When reasoning about meeting indexation, the entire process, ranging from annotation to review should be considered. It is very difficult to do research on only one part if one wants real in-the-field circumstances. For instance, users will never use the system seriously if they don't get a browser for the system to annotate their meetings.
- It is important to explain to users that the annotation component is not necessarily a replacement of their current note-taking practice, but rather an augmentation or addition to it.
- Using an annotation user interface in addition to or instead of regular note-taking is quite different from regular note-taking and takes getting used to. The user has to find a balance between writing notes and completely automatic indexing.

Although these recommendations may be carefully applied to different situations, the restrictive context is dictated by the nature of the experiments: mainly research-oriented student-supervisor or researcher meetings.

5.3 Future Work and Recommendations

The initial study compared the concept of real-time indexation roughly with note-taking and only on one refined interface. To draw conclusions about encoding and storage benefits of real-time indexation, more longitudinal studies and subsequent refinement of the prototype are necessary.

Moreover, I recommend to try the following interesting possibilities of exploration:

- Technology: the development pace of speech recognition technology is currently quite fast. It might be interesting to really push the limits of speech recognition technology and see what results are possible. Moreover, I didn't have full access to tools that enable the creation of speech-based applications (such as the Dragon

API) due to long negotiations for licenses. This problem will soon be solved, which opens up entirely new possibilities.

- Improvement/Extension: different input devices could be added to the annotation component and different architectures to automate or combine features should be tried. Also, it may be very interesting to look into collaborative note-taking possibilities (making experiments with the prototype as a multi-user system). Finally, there is the interesting technological aspect of integrating the RIMR system with other systems from the CHIL project (such as speaker activity detection systems) and looking into the interesting possibilities that arise (such as using speaker activity as additional contextual information for the annotation or browsing subsystem).
- Usage and improvement: how to find the best balance for cognitive processing and how can such a balance be found for the whole process (indexing as well as browsing)? The RIMR system indexing user interface may make the processing depth more shallow during the meeting since note-taking may be minimised or replaced completely, but it may not be a problem because of the (cognitive) post-processing after the meeting.
- What is the result of the cognitive postprocessing that occurs in a browser with "editing capabilities" which makes it possible for the user to create a condensed version of the meeting recording? This could be found through more usability studies.

Bibliography

[AMI] Augmented Multiparty Interaction Project. http://www.amiproject.org/.

[APC] Apcom Ltd. homepage. http://www.apcom.ltd.uk/.

[BA04] Jason A. Brotherton and Gregory D. Abowd. Lessons learned from eclass: Assessing automated capture and access in the classroom. *ACM Trans. Comput.-Hum. Interact.*, 11(2):121–155, 2004.

[BNC] Britsh National Corpus Homepage. http://www.natcorp.ox.ac.uk/.

[Bro02] M. Broughton. Measuring the accuracy of commercial automated speech recognition systems during conversational speech. In *Virtual Conversational Characters (VCC) Workshop, Human Factors Conference December 2002, Melbourne, Australia*, 2002.

[But97] G.C. Butazzo. *Hard real-time computing systems, predictable scheduling Ű algorithms and applications*. Kluwer Academic Publishers, 1997. ISBN 0-7923-9994-3.

[CBGK01] Patrick Chiu, John S. Boreczky, Andreas Girgensohn, and Don Kimber. Liteminutes: an internet-based system for multimedia meeting minutes. In *World Wide Web*, pages 140–149, 2001.

[CHI] CHIL Project Information Page. http://chil.server.de/.

[CL72] F.I. Craik and R. S. Lockhart. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 12:599–607, 1972.

[CRG+02] Ross Cutler, Yong Rui, Anoop Gupta, JJ Cadiz, Ivan Tashev, Li wei He, Alex Colburn, Zhengyou Zhang, Zicheng Liu, and Steve Silverberg. Distributed meetings: a meeting capture and broadcasting system. In *MULTIMEDIA '02: Proceedings of the tenth ACM international conference on Multimedia*, pages 503–512, New York, NY, USA, 2002. ACM Press.

[CT75] Fergus I. M. Craik and Endel Tulving. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology*, 3, Sep 1975.

[DLC+99] Richard C. Davis, James A. Landay, Victor Chen, Jonathan Huang, and Rebecca B. Lee. Notepals: Light weight note sharing by the group, for the group. In *CHI*, pages 338–345, 1999.

[Draa] Nuance Dragon AudioMining. http://www.nuance.com/audiomining/.

[Drab] Nuance Dragon Naturally Speaking. http://www.nuance.com/naturallyspeaking/.

[Dun86] Patricia A. Dunkel. Developing listening fluency in l2: Theoretical principles and pedagogical considerations. *The Modern Language Journal*, 70(2):99–106, 1986.

[EMS85] G.O. Einstein, J. Morris, and S. Smith. Note-taking, individual differences, and memory for lecture information. *Journal of Educational Psychology*, 77:522–532, 1985.

[EVi] IBM Embedded ViaVoice. http://www-306.ibm.com/software/pervasive/embedded_viavoice/.

[FUR] Sadaoki FURUI. Recent progress in corpus-based spontaneous speech recognition.

[GAV00] J. Garofolo, G. Auzanne, and E. Voorhees. The trec spoken document retrieval track: A success story, 2000.

[Har78] J. Hartley. Note-taking: A critical review. *Programmed Learning and Educational Technology*, 15:207–224, 1978.

[Jul] Julius Open Source LVCSR engine. http://julius.sourceforge.jp/en_index.php?q=en/index.html.

[Kah04] Tuomo Kahkonen. Agile methods for large organizations - building communities of practice. *adc*, 00:2–11, 2004.

[Kie85] K. A. Kiewra. Providing the instructor's notes: An effective addition to student notetaking. *Educational Psychologist*, 20:33–39, 1985.

[Kie89] K. Kiewra. A review of notetaking: The encoding-storage paradigm and beyond. *Educational Psychology Review*, 1:147–172, 1989.

[Lan] Roeland Ordelman Language. Speech-based information retrieval for dutch.

[LOG] Logitech io_2 Digital Writing System. http://www.logitech.com/index.cfm/mice_pointers/digital_pen/devices/408&cl=us,en.

[Mee91] Mac Yin Mee. Notetaking - an overview. *The English Teacher*, XX, 1991.

[MMM] MultiModal Media (MMM) Fileserver. http://mmm.idiap.ch/.

[NIS] NIST Speech Group. http://www.nist.gov/speech/.

[NMS]

[NOK] Nokia SU-1B Digital Pen. http://www.nokiausa.com/A4410909.

[PEG] Pegasus Technologies Ltd. homepage. http://www.pegatech.com/.

[POK04] Piolat, Olive, and Kellogg. Cognitive effort during note taking. *Applied Cognitive Psychology*, 19:291–312, 2004.

[SAS01] Lisa Stifelman, Barry Arons, and Chris Schmandt. The audio notebook: paper and pen interaction with structured speech. In *CHI 2001*, volume 3, pages 182–189, 2001.

[Shr05] E. Shriberg. Spontaneous speech: how people really talk and why engineers should care. In *Eurospeech 2005, Lisbon, Portugal*, 2005.

[Son] CSLR Colorado Sonic. `http://cslr.colorado.edu/beginweb/speech_recognition/sonic.html`.

[Sph] CMU Sphinx. `http://www.speech.cs.cmu.edu/sphinx/Sphinx.html`.

[TS07] J.M.B. Terken and E. Schindler. Automatic indexation for meeting browsers: A human-centered approach. In *Interspeech 2007 - Eurospeech, Antwerp, Belgium August 27-31*, 2007. Submitted.

[TW04] Simon Tucker and Steve Whittaker. Accessing multimodal meeting data: Systems, problems and possibilities. In Samy Bengio and Hervé Bourlard, editors, *Proc. MLMI'04*, Martigny, Switzerland, June 2004. Springer-Verlag LNCS. AMI-35.

[VFo] VoxForge GPL speech corpus collection. `http://www.voxforge.org/`.

[Via] Nuance - IBM ViaVoice. `http://www.nuance.com/viavoice/`.

[WHW94] Steve Whittaker, Patrick Hyland, and Myrtle Wiley. Filochat: handwritten notes provide access to recorded conversations. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 271–277, New York, NY, USA, 1994. ACM Press.

[WLT05] Steve Whittaker, Rachel Laban, and Simon Tucker. Analysing meeting records: An ethnographic study and technological implications. In *2nd Joint Workshop on Multimodal Interaction and Machine Learning Algorithms*, 2005.

[WSS97] Lynn D. Wilcox, Bill N. Schilit, and Nitin "Nick" Sawhney. Dynomite: A dynamically organized ink and audio notebook. In *CHI*, pages 186–193, 1997.

Appendix A

Questionnaire

Abstract

This appendix contains the results from the quantitative ethnography survey by means of a questionnaire, mentioned in Section 2.2.

Section A.1 is made up of the paper version of the questionnaire. Section A.2 summarises responses from a first trial run of the questionnaire and Section A.3 is a condensed enumeration of the responses on a large-scale electronically deployed version of the questionnaire.

A.1 Questions

The next pages contain the questions, as they have been posed to respondents. Nine of the ten questions of the survey were multiple choice questions with a Likert-scale and a possibility to give custom answers. The tenth question was an open question. The pilot version of the questionnaire was disseminated amongst people who took part in the observations (mainly students of the TU/e who were doing group work, such as OGO).

Questionnaire of Current Note-taking Practice

The **goal** of this questionnaire is to **poll for current note-taking practice during meetings**. The questionnaire is **anonymous** and the results will **solely** be **used for research purposes and requirements gathering**.

The questionnaire has **10** questions. In order for the results to be useful, the questionnaire should be **completed without leaving any answers blank**. Please fill out this questionnaire either in **English or in Dutch**.

Each of the first 9 questions poses a number of propositions that can be answered with “never”, “sometimes”, “regularly”, “often” or “always”. You can give your answer by circling or underlining your choice. If the choices do not fit your answer or if you think you might have more to say than can be done with the standard answers, please **feel free to write your own proposition** (if you do that, please indicate how often that proposition holds true for you) or a comment at the end of the relevant question. The 10th question is an open question. Please make sure to answer it **completely** and not only part of it. If you would like to write some additional comment, you can do so after the 10th question.

1. How would you assess the structure of the meetings that you attend(ed)?

- Formal meetings with an explicit structure (e.g. with an agenda, a chairman and a minutes secretary).

  once a month | 2-3x per month | 1-2x per week | almost every day

- Informal meetings with no or an implicit structure (e.g. student/supervisor meetings or informal (project) group meetings/discussions with 2 or more people).

  once a month | 2-3x per month | 1-2x per week | almost every day

- Other, namely:_____

The rest of the questions is about all the different kinds of meetings (e.g. formal and informal) that you attended.

2. What do you normally do during a meeting (how do you preserve information that is of personal relevance for you)?

- I write down full notes[1], so that I don't have to work out notes later on.

  never | sometimes | regularly | often | always

- I write down keywords and work out full notes later on.

  never | sometimes | regularly | often | always

- I write down keywords only and use those to remember what was relevant.

  never | sometimes | regularly | often | always

- I make an audio recording.

  never | sometimes | regularly | often | always

- I make a video recording.

  never | sometimes | regularly | often | always

- I don't make any notes whatsoever.

  never | sometimes | regularly | often | always

- Other, namely (please indicate how often):_____

  _____

  _____

- **Which tools do you use for taking personal notes during a meeting?**
  - Pen and paper.

    never | sometimes | regularly | often | always

  - Laptop/Tablet (manual input).

    never | sometimes | regularly | often | always

  - Palmtop/Hand held (manual input).

    never | sometimes | regularly | often | always

  - (Digital) dictation device/Memory stick with microphone.

    never | sometimes | regularly | often | always

  - Other, namely (please indicate how often):_____

    _____

    _____

1 Notes that can be understood by someone who did not attend the meeting.

4. **How do you make use of your (either original or worked out) personal notes/recordings?**

- I read them before the next meeting.

  never | sometimes | regularly | often | always

- I use them for reference during the next meeting.

  never | sometimes | regularly | often | always

- I use them as a to do list and for reference between meetings.

  never | sometimes | regularly | often | always

- I don't really use them, the note-taking itself is just helping me remembering things.

  never | sometimes | regularly | often | always

- Other, namely (please indicate how often):_____

5. **How do you organize your notes?**

- I just keep them somewhere on paper or in a digital file without organizing them.

  never | sometimes | regularly | often | always

- I order them by meeting/date.

  never | sometimes | regularly | often | always

- I order them by topic.

  never | sometimes | regularly | often | always

- I split them into to-do list and/or journal.

  never | sometimes | regularly | often | always

- I order them by (please indicate how often):_____

- **How do you search your notes for specific information?**

➢ I skim through them until I find what I need.

never | sometimes | regularly | often | always

➢ I recall a page/place in the notes.

never | sometimes | regularly | often | always

➢ I use a digital search function (text search, file search, etc.).

never | sometimes | regularly | often | always

➢ Other, namely:

- **Do you recall situations in which your normal note-taking practice was lacking? If so, what was missing?**

➢ I experienced problems with my notes.

never | sometimes | regularly | often | always

➢ The writing was illegible.

never | sometimes | regularly | often | always

➢ I had trouble recalling what was discussed during the meeting, even after consulting my notes.

never | sometimes | regularly | often | always

➢ Other, namely (please indicate how often):

- **Do you compare or share notes with other people (if applicable)?**
  - I compare or share notes with other people.

    never | sometimes | regularly | often | always
  - I use the notes of other people, no personal notes of mine.

    never | sometimes | regularly | often | always
  - I compare/share notes with other people (e.g. by copying or discussing) for enriching my (and their) personal notes.

    never | sometimes | regularly | often | always
  - I share my notes with other people at the next meeting for reference.

    never | sometimes | regularly | often | always
  - I do collaborative note-taking with other people (document editing by multiple people, (digital) bulletin board, wiki-page, etc.).

    never | sometimes | regularly | often | always
  - Other, namely:

9. **Do you make reference to documents, such as minutes, agendas, handouts, etc.? Do you use documents as a guide for structuring your note-taking? If you use/refer to other documents, which ones (i.e. what kind) are those?**
   - I make reference to or use internal documents for structuring my notes.

     never | sometimes | regularly | often | always
   - If possible, I use an agenda to structure my notes.

     never | sometimes | regularly | often | always
   - If possible, I refer to minutes.

     never | sometimes | regularly | often | always
   - Other documents, namely:

10. a. What are the strong and weak sides of your current note-taking practice?

10. b. If you would have the opportunity to have your note-taking digitally supported, how would you like to have this done? In case you would not like any digital supported way of note-taking, why not?

Thank you very much for filling out this questionnaire! Your contribution makes it possible to get a tangible idea of current note-taking practice. This knowledge will be used for research and development of a prototype for a digital system that supports note-taking.

If desired, the next pages have space for additional comments.

Comments (whatever general or other comments you like to place)

A.2 Pilot Results

Number of respondents: 46

Preservation of information

- 15% never writes out full notes, 60% does it sometimes, 20% often and 5% always.
- 35% never writes down keywords and works out full notes later, 50% does it sometimes, 10% often and 5% always.
- 15% never writes down keywords only for remembering what was relevant, 40% does it sometimes, 40% often and 5% always.
- Nobody ever makes an audio recording.
- Video recordings are also never made by anybody.
- 50% sometimes makes no notes at all, 10% does it often, 5% never makes any notes and 35% always makes some kind of notes.

Custom answers

- One person (minutes secretary) takes notes.
- Headnotes[2].
- Read minutes
- I do not make full notes, but I always make minutes.

Tools

- 35% often uses pen and paper. 65% always uses pen and paper.
- 55% never uses a laptop or tablet, 40% does it sometimes and 5% often.
- Nobody uses a hand-held device or
- a dictation device.

Use of notes/recordings

- 15% never reads their notes before the next meeting, 55% does it sometimes, 25% often and 5% always.
- 30% never uses their notes for reference during the next meeting, 35% does it sometimes and 30% often.
- 15% never uses their notes as a to do list and for reference between meetings, 30% does it sometimes, 40% often and 15% never.
- 20% never uses their notes for remembering only (i.e. they don't re-use them for reference), 40% does this often, 30% sometimes and 5% always.

2 Means that he/she “always” remembers “everything” :)

Custom answers

- It is both for me to use and remember things.

Organization of notes

- 20% never keeps their notes just somewhere on paper or in a digital file without organizing them, 25% does this sometimes, 40% often and 15% always.
- 30% never orders them by meeting/date, 40% does this sometimes, 20% often and 10% always.
- 65% never orders their notes by topic, 20% does it sometimes and 20% often.
- 80% never splits their notes into to do list or journal, 10% does this sometimes and 10% often.

Custom answers

- I write to what meeting a certain note belongs and keep this note in a clipper or dictation book.
- Sometimes I write to-do actions on my general to-do list.
- I have one paper notebook that I always keep notes of meetings then I make electronic files of minutes.

Notes search

- 30% sometimes skims through their notes until they find what they need, 60% does this often and 10% always.
- 35% never just recalls a page/place in the notes, 40% does this sometimes and 20% often.
- 60% never uses a digital search function, 30% does this sometimes and 10% often.

Custom answers

- My notes are just keywords, never more than a few lines per meeting.
- I do digital search in case of official minutes (“notulen”).
- For lengthy notes, usually the notes are ordered by topic (agenda) so I can use thes structure in the notes to guide the search process.
- Also, in digital notes, to-do actions are marked by AP[3] and/or different colors.

Lack of regular note-taking practice

- 25% never experienced any problems with their notes, 55% sometimes did and 5% often did.
- 55% had never problems with illegible writing, 65% sometimes did and 5% often did.

3 Action Point?

- 30% never had any trouble recalling what was discussed during the meeting even when using their notes, 65% sometimes did and 10% often did.

Custom answers

- I could not remember or did not write down names (20% of the time).
- Sometimes it is difficult to reconstruct things from notes (especially keywords) when the context is forgotten.

Comparing/sharing of notes

- 40% never compares or shares notes with other people, 55% sometimes does and 5% often does.
- 30% never uses the notes of other people and no personal notes, 40% sometimes does and 30% often does.
- 55% never compares or shares notes with other people for enriching their personal notes, 35% sometimes does and 10% often does.
- 50% never shares their notes with other people at the next meeting for reference, 35% sometimes does and 15% often does.
- 85% never does collaborative note-taking with other people and 15% sometimes does.

Custom answers

- I use notes of other people only in case of a formal meeting (I use the minutes).

Reference to documents

- 40% never makes any reference to or uses documents for structuring their notes, 45% sometimes does, 5% often does and 10% always does.
- 35% never uses an agenda (if possible) to structure their notes, 25% sometimes does, 20% often does and 20% always does.
- 65% never refers to minutes (if possible), 30% sometimes does and 5% often does.

Custom answers

- In projects, I refer to project docs. At board meetings to internal docs.

Meeting structure*

- 7% sometimes attends formal meetings, 50% often does and 33% always does.
- 8% never attends informal meetings, 58% sometimes does and 33% often does.

* This question was not included in all pilots, so there is less data (about 2/3 only) for this question.

Current note-taking versus digitally supported note-taking

Pros and cons of current note-taking practice

Paper and pen

Pros

- Easily organized (e.g. all notes in one notebook): 5%
- Helps me remember things: 10%
- Fast: 15%
- Easy: 20%
  - Easy to correct and expand during the meeting: 10%
  - Easier than digital systems, especially for making drawings and non-text annotations: 10%
- Coverage is high; I write down everything: 10%

Cons

- Sloppy; organization is a pain: 20%
- Not fast enough, I cannot keep up: 15%
  - I have to choose between talking/listening or writing: 10%
- Post-processing needed: 5%

Laptop / PC

Pros

- Fast: 10%

Cons

- Slow: 5%

Other

- A minutes secretary is helpful, but I can remember the most important issues/items without taking any notes.

Suggested ways of digital note-taking support

- Complete automatic transcript: 15%
  - A system that highlights the relevant parts in a complete automatic transcript: 5%
- Structuring of notes/recordings (timestamps, keyword linking): 5%
- Use a laptop (I don't use one, because I don't have one): 5%
- Digital pen (it helps me bind the gap between the physical and digital world): 5%
- No need for digital support (very few notes): 5%

- If there would be digital support that gives high note-coverage with minimal effort, I would make more notes: 5%

➢ Digital support not preferred/desired: 20%

- I am content with my manual way of note-taking: 5%
- Digital note-taking is impersonal: 5%

Desired features of digital note-taking support

➢ More note coverage with less writing effort: 15%

- Faster way of taking detailed notes: 5%
- Link more content to the keywords I write down: 5%

➢ An easy way to organize notes: 5%

➢ Visualization of complex items/issues that were discussed: 5%

➢ Automatic warning when I forget to take a note of something: 5%

➢ System must work “in the background” (without me noticing/controlling it) as much as possible: 5%

➢ Detailed preservation (audio, video, text): 5%

➢ Easy way to search: 5%

A.3 Final Results

Following are the quantitative results of the electronically disseminated actual (final) survey .In total, the final questionnaire had 106 respondents. The first summary (Dutch & English TU/e respondants) shows only the “internal” TU/e respondants, divided into people who responded in English, respectively in Dutch. The second summary (TU/e & External respondants) shows the total number of respondants, divided into answers from the TU/e and answers not from the TU/e.

A.3.1 Dutch & English TU/e respondants

1. How would you assess the structure of the meetings that you attend(ed)?

1a. Formal meetings with an explicit structure (e.g. with an agenda, a chairman and a minutes secretary)

	1xmonth	2/3xmonth	1/2xweek	almost every day	no answer
D (n=71)	22	31	5	5	8
E (n=12)	3	6	3	0	0
SUM	25	37	8	5	8
Perc	30	45	10	6	10

1b.Informal meetings with no or an implicit structure (e.g. student/supervisor meetings of informal (project) group meetings with two or more people)

	1xmonth	2/3xmonth	1/2xweek	almost every day	no answer
D (n=71)	3	21	5	37	5
E (n=12)	1	0	6	5	0
SUM	4	21	11	42	5
Perc	5	25	13	51	6

2. What do you normally do during a meeting (how do you preserve information that is of personal relevance to you)?

2a. I write down full notes (i.e. notes that can be understood by someone who did not attend the meeting), so that I don't have to work out notes later on.

	never	sometimes	regularly	often	always	no answer
D (n=71)	23	23	8	5	2	9
E (n=12)	9	1	0	0	0	2
SUM	32	24	8	5	2	11
Perc	39	29	10	6	2	13

2b. I write down keywords and work out full notes later on

	never	sometimes	regularly	often	always	no answer
D (n=71)	11	26	15	8	1	0
E (n=12)	0	3	5	1	1	2
SUM	11	29	20	9	2	2
Perc	13	35	24	11	2	2

2c. I write down keywords only and use those to remember what was relevant

	never	sometimes	regularly	often	always	no answer
D (n=71)	11	16	12	12	10	10
E (n=12)	1	1	3	4	1	2
SUM	12	17	15	16	11	12
Perc	14	20	18	19	13	14

2d. I make an audio recording

	never	sometimes	regularly	often	always	no answer
D (n=71)	54	5	0	0	0	12
E (n=12)	10	0	0	0	0	2
SUM	64	5	0	0	0	14
Perc	77	6	0	0	0	17

2e. I make a video recording

	never	sometimes	regularly	often	always	no answer
D (n=71)	53	6	0	0	0	12
E (n=12)	8	2	0	0	0	2
SUM	61	8	0	0	0	14
Perc	73	10	0	0	0	17

2f. I make no notes whatsoever

	never	sometimes	regularly	often	always	no answer
D (n=71)	22	28	3	4	1	13
E (n=12)	5	3	0	1	0	3
SUM	27	31	3	5	1	16
Perc	33	37	4	6	1	19

3. Which tools do you use for taking personal notes during a meeting?

3a. Pen and paper

	never	sometimes	regularly	often	always	no answer
D (n=71)	0	2	5	18	33	13
E (n=12)	0	0	0	3	7	2
SUM	0	2	5	21	40	15
Perc	0	2	6	25	48	18

3b. Laptop/table (manual input)

	never	sometimes	regularly	often	always	no answer
D (n=71)	27	20	6	4	0	14
E (n=12)	6	3	1	0	0	2
SUM	33	23	7	4	0	16
Perc	40	28	8	5	0	19

3c. Palmtop/handheld (manual input)

	never	sometimes	regularly	often	always	no answer
D (n=71)	50	4	0	0	0	17
E (n=12)	10	0	0	0	0	2
SUM	60	4	0	0	0	19
Perc	72	5	0	0	0	23

3d. (Digital) dictation device/memory stick with microphone

	never	sometimes	regularly	often	always	no answer
D (n=71)	51	3	0	0	0	17
E (n=12)	10	0	0	0	0	2
SUM	61	3	0	0	0	19
Perc	73	4	0	0	0	23

4. How often do you make use of your (either original or worked out) personal notes/recordings?

4a. I read them before the next meeting

	never	sometimes	regularly	often	always	no answer
D (n=71)	16	17	15	7	2	14
E (n=12)	1	4	4	1	0	2
SUM	17	21	19	8	2	16
Perc	20	25	23	10	2	19

4b. I use them for reference during the next meeting

	never	sometimes	regularly	often	always	no answer
D (n=71)	10	16	15	14	1	15
E (n=12)	1	5	2	2	0	2
SUM	11	21	17	16	1	17
Perc	13	25	20	19	1	20

4c. I use them as a to-do list and for reference between meetings

	never	sometimes	regularly	often	always	no answer
D (n=71)	2	12	11	26	5	15
E (n=12)	1	2	2	5	0	2
SUM	3	14	13	31	5	17
Perc	4	17	16	37	6	20

4d. I don't really use them, the note-taking itself is just helping me remembering things

	never	sometimes	regularly	often	always	no answer
D (n=71)	11	21	14	7	2	16
E (n=12)	4	4	0	2	0	2
SUM	15	25	14	9	2	18
Perc	18	30	17	11	2	22

5. How do you organize your notes?

5a. I just keep them somewhere on paper or in digital file without organizing them

	never	sometimes	regularly	often	always	no answer
D (n=71)	9	12	11	11	12	16
E (n=12)	2	3	1	2	2	2
SUM	11	15	12	13	14	18
Perc	13	18	14	16	17	22

5b. I order them by meeting/date

	never	sometimes	regularly	often	always	no answer
D (n=71)	20	9	10	8	5	19
E (n=12)	3	2	1	2	2	2
SUM	23	11	11	10	7	21
Perc	28	13	13	12	8	25

5c. I order them by topic

	never	sometimes	regularly	often	always	no answer
D (n=71)	28	14	5	4	1	19
E (n=12)	3	3	1	3	0	2
SUM	31	17	6	7	1	21
Perc	37	20	7	8	1	25

5d. I split them into to-do list and/or journal

	never	sometimes	regularly	often	always	no answer
D (n=71)	19	15	11	6	2	18
E (n=12)	6	1	0	3	0	2
SUM	25	16	11	9	2	20
Perc	30	19	13	11	2	24

6. How do you search through notes for specific information?

6a. I skim through them until I find what I need

	never	sometimes	regularly	often	always	no answer
D (n=71)	3	12	17	21	2	16
E (n=12)	0	3	2	3	1	3
SUM	3	15	19	24	3	19
Perc	4	18	23	29	4	23

6b. I recall a page/place in the notes

	never	sometimes	regularly	often	always	no answer
D (n=71)	3	19	12	19	1	17
E (n=12)	1	2	2	3	1	3
SUM	4	21	14	22	2	20
Perc	5	25	17	27	2	24

6c. I use a digital search function (text search, fiel search, etc.)

	never	sometimes	regularly	often	always	no answer
D (n=71)	32	13	2	4	0	20
E (n=12)	6	3	0	0	0	3
SUM	38	16	2	4	0	23
Perc	46	19	2	5	0	28

7. Do you recall situations in which your normal note-taking practive was lacking? If so , what was missing?

7a. I experienced problems with my notes

	never	sometimes	regularly	often	always	no answer
D (n=71)	13	36	5	0	0	17
E (n=12)	1	8	0	0	0	3
SUM	14	44	5	0	0	20
Perc	17	53	6	0	0	24

7b. The writing was illegible

	never	sometimes	regularly	often	always	no answer
D (n=71)	33	13	5	2	0	18
E (n=12)	7	2	0	0	0	3
SUM	40	15	5	2	0	21
Perc	48	18	6	2	0	25

7c. I had trouble recalling what was discussed during the meeting, even after consulting my notes

	never	sometimes	regularly	often	always	no answer
D (n=71)	14	31	8	1	0	17
E (n=12)	3	5	1	0	0	3
SUM	17	36	9	1	0	20
Perc	20	43	11	1	0	24

8. Do you compare or share notes with other people (if possible or applicable)?

8a. I compare or share notes with other people

	never	sometimes	regularly	often	always	no answer
D (n=71)	12	24	9	7	2	17
E (n=12)	1	5	2	1	0	3
SUM	13	29	11	8	2	20
Perc	16	35	13	10	2	24

8b. I use the notes of other people, no personal notes of mine

	never	sometimes	regularly	often	always	no answer
D (n=71)	17	23	11	3	0	17
E (n=12)	4	4	1	0	0	3
SUM	21	27	12	3	0	20
Perc	25	33	14	4	0	24

8c. I compare/share notes with other people (e.g. by copying or discussing) for enriching my (and their) personal notes

	never	sometimes	regularly	often	always	no answer
D (n=71)	5	19	9	1	0	17
E (n=12)	3	4	2	0	0	3
SUM	8	23	11	1	0	20
Perc	10	28	13	1	0	24

8d. I share my notes with other people at the next meeting for reference

	never	sometimes	regularly	often	always	no answer
D (n=71)	23	18	9	2	1	18
E (n=12)	2	4	2	1	0	3
SUM	25	22	11	3	1	21
Perc	30	27	13	4	1	25

8e. I do collaborative note-taking with other people (document editing by multiple people, (digital) bulletin board, wiki-page, etc.)

	never	sometimes	regularly	often	always	no answer
D (n=71)	29	15	7	1	0	19
E (n=12)	7	1	1	0	0	3
SUM	36	16	8	1	0	22
Perc	43	19	10	1	0	27

9. Do you make reference to documents such as minutes, agendas, handouts etc?
Do you use documents as a guide for structuring your note-taking?
If you use/refer to other documents, which ones (i.e. what kind) are those?

9a. I make reference to or use internal documents for structuring my notes

	never	sometimes	regularly	often	always	no answer
D (n=71)	19	25	8	0	0	19
E (n=12)	5	3	1	0	0	3
SUM	24	28	9	0	0	22
Perc	29	34	11	0	0	27

9b. If possible, I use an agenda to structure my notes

	never	sometimes	regularly	often	always	no answer
D (n=71)	21	12	7	10	2	19
E (n=12)	6	0	3	0	0	3
SUM	27	12	10	10	2	22
Perc	33	14	12	12	2	27

9c. If possible, I refer to minutes

	never	sometimes	regularly	often	always	no answer
D (n=71)	23	18	7	1	3	19
E (n=12)	3	2	3	0	0	4
SUM	26	20	10	1	3	23
Perc	31	24	12	1	4	28

A.3.2 TU/e & External respondants

TU/e: respondents on TU/e	n = 83
External: respondents outside TU/e	n = 23
Total respondents: TU/e and Ext. together	n = 106

1. How would you assess the structure of the meetings that you attend(ed)?

1a. Formal meetings with an explicit structure (e.g. with an agenda, a chairman and a minutes secretary).

	1/month	2-3/month	1-2/week	almost every day	no answer
TU/e	25	37	8	5	8
External	8	7	4	3	1
Sum	33	44	12	8	9
Percentage	31	42	11	8	8

1b.Informal meetings with no or an implicit structure (e.g. student/supervisor meetings of informal (project) group meetings with two or more people)

	1/month	2-3/month	1-2/week	almost every day	no answer
TU/e	4	21	11	42	5
External	4	5	3	5	6
Sum	8	26	14	47	11
Percentage	8	25	13	44	10

2. What do you normally do during a meeting (how do you preserve information that is of personal relevance to you)?

2a. I write down full notes (i.e. notes that can be understood by someone who did not attend the meeting), so that I don't have to work out notes later on.

	never	sometimes	regularly	often	always	no answer
TU/e	32	25	8	5	2	11
External	4	11	3	0	2	3
Sum	36	36	11	5	4	14
Percentage	34	34	10	5	4	13

2b. I write down keywords and work out full notes later on

	never	sometimes	regularly	often	always	no answer
TU/e	11	29	20	9	2	12
External	4	2	2	7	3	5
Sum	15	31	22	16	5	17
Percentage	14	29	21	15	5	16

2c. I write down keywords only and use those to remember what was relevant

	never	sometimes	regularly	often	always	no answer
TU/e	12	17	15	16	11	12
External	1	6	5	5	2	4
Sum	13	23	20	21	13	16
Percentage	12	22	19	20	12	15

2d. I make an audio recording

	never	sometimes	regularly	often	always	no answer
TU/e	64	5	0	0	0	14
External	18	0	0	0	0	5
Sum	82	5	0	0	0	19
Percentage	77	5	0	0	0	18

2e. I make a video recording

	never	sometimes	regularly	often	always	no answer
TU/e	61	8	0	0	0	14
External	18	0	0	0	0	5
Sum	79	8	0	0	0	19
Percentage	75	8	0	0	0	18

2f. I make no notes whatsoever

	never	sometimes	regularly	often	always	no answer
TU/e	27	31	3	5	1	16
External	7	7	1	3	0	5
Sum	34	38	4	8	1	21
Percentage	32	36	4	8	1	20

3. Which tools do you use for taking personal notes during a meeting?

3a. Pen and paper

	never	sometimes	regularly	often	always	no answer
TU/e	0	2	5	21	40	15
External	0	3	0	3	14	3
Sum	0	5	5	24	54	18
Percentage	0	5	5	23	51	17

3b. Laptop/table (manual input)

	never	sometimes	regularly	often	always	no answer
TU/e	33	23	7	4	0	16
External	15	2	0	0	0	6
Sum	48	25	7	4	0	22
Percentage	45	24	7	4	0	21

3c. Palmtop/handheld (manual input)

	never	sometimes	regularly	often	always	no answer
TU/e	60	4	0	0	0	19
External	17	0	0	0	0	6
Sum	77	4	0	0	0	25
Percentage	73	4	0	0	0	24

3d. (Digital) dictation device/memory stick with microphone

	never	sometimes	regularly	often	always	no answer
TU/e	61	3	0	0	0	19
External	17	0	0	0	0	6
Sum	78	3	0	0	0	25
Percentage	74	3	0	0	0	24

4. How often do you make use of your (either original or worked out) personal notes/recordings?

4a. I read them before the next meeting

	never	sometimes	regularly	often	always	no answer
TU/e	17	21	19	8	2	16
External	1	8	1	7	2	4
Sum	18	29	20	15	4	20
Percentage	17	27	19	14	4	19

4b. I use them for reference during the next meeting

	never	sometimes	regularly	often	always	no answer
TU/e	11	21	17	16	1	17
External	0	6	4	4	5	4
Sum	11	27	21	20	6	21
Percentage	10	25	20	19	6	20

4c. I use them as a to-do list and for reference between meetings

	never	sometimes	regularly	often	always	no answer
TU/e	3	14	13	31	5	17
External	2	6	4	5	3	3
Sum	5	20	17	36	8	20
Percentage	5	19	16	34	8	19

4d. I don't really use them, the note-taking itself is just helping me remebering things

	never	sometimes	regularly	often	always	no answer
TU/e	15	25	14	9	2	18
External	7	9	1	2	0	4
Sum	22	34	15	11	2	22
Percentage	21	32	14	10	2	21

5. How do you organize your notes?

5a. I just keep them somewhere on paper or in digital file without organizing them

	never	sometimes	regularly	often	always	no answer
TU/e	11	15	12	13	14	18
External	8	3	2	1	4	5
Sum	19	18	14	14	18	23
Percentage	18	17	13	13	17	22

5b. I order them by meeting/date

	never	sometimes	regularly	often	always	no answer
TU/e	23	11	11	10	7	21
External	2	3	5	2	7	4
Sum	25	14	16	12	14	25
Percentage	24	13	15	11	13	24

5c. I order them by topic

	never	sometimes	regularly	often	always	no answer
TU/e	31	17	6	7	1	21
External	9	3	2	3	2	4
Sum	40	20	8	10	3	25
Percentage	38	19	8	9	3	24

5d. I split them into to-do list and/or journal

	never	sometimes	regularly	often	always	no answer
TU/e	25	16	11	9	2	20
External	10	3	2	2	2	4
Sum	35	19	13	11	4	24
Percentage	33	18	12	10	4	23

6. How do you search through notes for specific information?

6a. I skim through them until I find what I need

	never	sometimes	regularly	often	always	no answer
TU/e	3	15	19	24	3	19
External	0	8	2	6	4	3
Sum	3	23	21	30	7	22
Percentage	3	22	20	28	7	21

6b. I recall a page/place in the notes

	never	sometimes	regularly	often	always	no answer
TU/e	4	21	14	22	2	20
External	2	8	4	4	2	3
Sum	6	29	18	26	4	23
Percentage	6	27	17	25	4	22

6c. I use a digital search function (text search, fiel search, etc.)

	never	sometimes	regularly	often	always	no answer
TU/e	3	8	16	24	0	23
External	10	5	4	0	0	4
Sum	13	13	20	24	0	27
Percentage	12	12	19	23	0	25

7. Do you recall situations in which your normal note-taking practive was lacking? If so , what was missing?

7a. I experienced problems with my notes

	never	sometimes	regularly	often	always	no answer
TU/e	14	44	5	0	0	20
External	8	8	2	1	0	4
Sum	22	52	7	1	0	24
Percentage	21	49	7	1	0	23

7b. The writing was illegible

	never	sometimes	regularly	often	always	no answer
TU/e	40	15	5	2	0	21
External	14	5	1	0	0	13
Sum	54	20	6	2	0	34
Percentage	51	19	6	2	0	32

7c. I had trouble recalling what was discussed during the meeting, even after consulting my notes

	never	sometimes	regularly	often	always	no answer
TU/e	17	36	9	1	0	20
External	6	12	0	0	1	4
Sum	23	48	9	1	1	24
Percentage	22	45	8	1	1	23

8. Do you compare or share notes with other people (if possible or applicable)?

8a. I compare or share notes with other people

	never	sometimes	regularly	often	always	no answer
TU/e	13	29	11	8	2	20
External	6	8	3	2	0	4
Sum	19	37	14	10	2	24
Percentage	18	35	13	9	2	23

8b. I use the notes of other people, no personal notes of mine

	never	sometimes	regularly	often	always	no answer
TU/e	21	27	12	3	0	20
External	7	9	1	2	0	4
Sum	28	36	13	5	0	24
Percentage	26	34	12	5	0	23

8c. I compare.share notes with other people (e.g. by copying or discussing) for enriching my (and their) personal notes

	never	sometimes	regularly	often	always	no answer
TU/e	28	23	11	1	0	20
External	11	7	1	0	0	4
Sum	39	30	12	1	0	24
Percentage	37	28	11	1	0	23

8d. I share moy notes with other people at the next meeting for reference

	never	sometimes	regularly	often	always	no answer
TU/e	25	22	11	3	1	21
External	7	7	2	1	2	4
Sum	32	29	13	4	3	25
Percentage	30	27	12	4	3	24

8e. I do collaborative note-taking with other people (document editing by multiple people, (digital) bulletin board, wiki-page, etc.)

	never	sometimes	regularly	often	always	no answer
TU/e	36	16	8	1	0	22
External	14	4	1	0	0	4
Sum	50	20	9	1	0	26
Percentage	47	19	8	1	0	25

9. Do you make reference to documents such as minutes, agendas, handouts etc?
Do you use documents as a guide for structuring you rnote-taking?
If you use/refer to other documents, which ones (i.e. what kind) are those?

9a. I make reference to or use internal documents for structuring my notes

	never	sometimes	regularly	often	always	no answer
TU/e	24	28	9	0	0	22
External	9	3	2	4	0	5
Sum	33	31	11	4	0	27
Percentage	31	29	10	4	0	25

9b. If possible, I use an agenda to structure my notes

	never	sometimes	regularly	often	always	no answer
TU/e	27	12	10	10	2	22
External	8	2	2	6	1	4
Sum	35	14	12	16	3	26
Percentage	33	13	11	15	3	25

9c. If possible, I refer to minutes

	never	sometimes	regularly	often	always	no answer
TU/e	26	20	10	1	3	23
External	4	6	5	3	1	4
Sum	30	26	15	4	4	27
Percentage	28	25	14	4	4	25

A.4 Summary

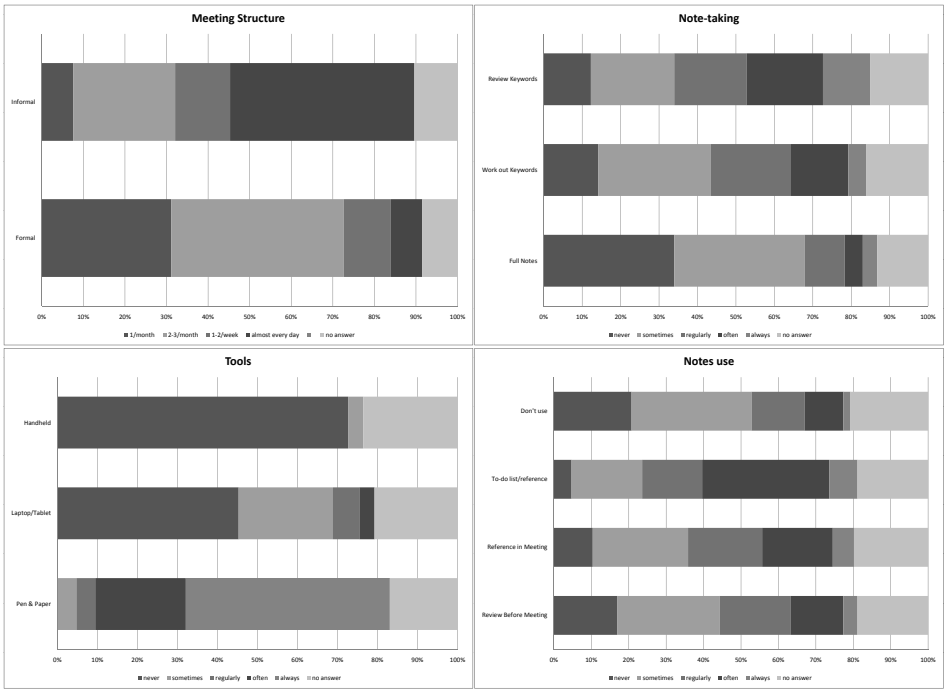

Figure A.1: Four important survey results (overview)

Figure A.1 summarises the most important results.

- Informal meetings (and thus the need for personal note-taking) are much more frequent than formal meetings with a secretary.
- Keywords for review are taken more frequently than full notes or notes worked out from keywords.
- The dominant tools for taking notes are pen and paper.
- Notes are mainly used as reference material during and between meetings.
- The main strong point of note-taking was ease of use.
- The main weak points were high cognitive load and sparseness (lack of context in keywords degenerates recall from review quickly).

Appendix B

Observations

Abstract

This appendix lists the results from the human note-taking observations mentioned in Section 2.2.

The pages of this appendix cover the detailed reports of observations of meeting groups at the TU/e. The goal of the observations was to study the note-taking behavior of people in the field, answering a number of research questions that are essential to the RIMR project. Before an observation started, participants were presented a video recording consent form in which they could determine for which purposes a video recording, if made, could be used (research for this work only or internal communication). The following questions were explored:

- What does an attendee do during a meeting? How is the information produced/presented during the meeting preserved (note-taking, recording, remembering and the like)? If notes are taken, which tools are used (paper and pen, laptop, palmtop, voice/video recording, etc.)?
- How do people make use of their notes (read before a next meeting, use them as a to do list, discard them and so on)?
- Do people compare or share notes with other people?

Initially, the following Observation Guidelines were used to keep some structure in the observation results.

1. How do people make (personal) notes (paper, computer, etc.)?
2. Do people use notes from previous meetings? If so, what for?
3. Can people keep up with note-taking and follow the meeting properly at the same time?
4. How much cognitive strain does the note-taking cost (how much can people follow of what is said and how much are they distracted by note-taking)? This is hard to assess, but may be discovered to some degree by asking people.
5. Do people share their notes before/during/after the meeting?
6. How do people store their notes at the end of the meeting (loose sheets, notebook, discard them)?

1. OGO meeting department BMT 09/16/05

Number of attendees: 7

Expected type: formal

Types of peers: students + tutor

Guided Observations

1. **Note-taking means:** Everybody used pen and paper only. Everybody has a pen and a paper ready, but only a small part actually writes something down.
2. **Use of previous notes:** One person made use of his own notes (from a previous meeting or especially prepared for this meeting) for looking up open issues that had to be resolved.
3. **Note upkeep:** very few notes were taken; most of the attention went to discussions about the contents of the OGO-project, so people did not have any problems keeping up the note-taking and following the meeting.
4. **Cognitive strain:** The notes that were taken were very brief (keyword-like), so there was none to little trouble in keeping up with the notes. However, for the time the notes were taken, no participation came from the secretary.
5. **Note sharing:** in principle no notes were shared; some inherent sharing came forth of one person using his notes to remind the others of open issues in the project.
6. **Note storing:** All people who made notes, did so in a notebook. The notes were kept in the notebook afterwards.

Additional Observations

- Something that really comes to attention is the fact that the meeting is very content-rich. It is mainly not a progress-report meeting, but a technical discussion. Because of this, the note-taking is very poor: first many technical issues are discussed and then the results and/or decisions are written down very concisely.
- One person wrote down the technical results/solutions, another one (the minutes secretary) wrote down the decisions that had been made.
- Some people used their paper for demonstration purposes (like a white board) instead of note-taking: something is written or drawn on the paper so that the others can see it. Others then added to the picture or writing.
- The to do list for the entire group has been drawn up on the white board (collective note-taking of to do items).
- Something else that caught the eye was that in the more relaxed parts of the meeting, almost everybody started clicking with their pens.
- People did not often talk at the same time. When this happened, it was for one of the following reasons:
    - Interruption/Draw attention.
    - Funny/Loose remarks (do not really matter to the contents, but do loosen the atmosphere).
    - Two or more people who appear not to be interested in the central discussion start talking with each other on the background.

2. OGO meeting department BMT 09/19/05

Number of attendees: 7

Expected type: formal

Types of peers: students + tutor

Guided Observations

1. **Note-taking means:** Nobody, except for the minutes secretary had any special tools for taking notes. One person used the agenda for writing to do items and decisions on it.
2. **Use of previous notes:** No notes from previous meetings were used. Instead, people used their laptops for discussing technical topics and bringing up new issues.
3. **Note upkeep:** Nobody took personal notes. Even the minutes secretary wrote down almost nothing. So there was no problem in keeping up note taking.
4. **Cognitive strain:** The minutes secretary who took notes did so in pretty short style, so he had no noticeable cognitive strain for taking the notes themselves. He also indicated that he had no trouble following the meeting. He did not participate much in the meeting, though when writing one of the few things down.
5. **Note sharing:** No previous notes were used and none were shared. However, a beamer was used for sharing technical data about which the discussion went.
6. **Note storing:** The minutes secretary stored his notes in a notebook. The person who made notes on his agenda, stored the notes on a loose sheet (the agenda itself).

Additional Observations

- People sometimes used the agenda creatively as a white board.
- One of the agendas was used creatively for some kind of collective note-taking (the person who had to write the project report let the agenda go around and people wrote their affiliation, name and student number down on it.
- The beamer was directed at a white board and people were able to draw on the projected picture by drawing on the white board.

3. OGO meeting department Computer Science 09/21/05 (11:45)

Number of attendees: 6

Expected type: formal

Types of peers: students + tutor

Guided Observations

1. **Note-taking means:** The minutes secretary used a laptop. Three other people used pen and paper for writing during the meeting.
2. **Use of previous notes:** The tutor used previous notes (probably from a tutor-meeting) for looking up tips and tricks for the group.
3. **Note upkeep:** The minutes secretary was perfectly able to keep up with the discussions. The tutor wrote down keywords (probably for the next tutor meeting). One other person wrote down to do items on a loose sheet (agenda).
4. **Cognitive strain:** The minutes secretary indicated that he had absolutely no problem typing the minutes in real time and following all of the discussion. He also had some input to the discussion, but by far not as much as others had.
5. **Note sharing:** No note sharing whatsoever was done.
6. **Note storing:** The minutes secretary stored the minutes in an MS Word document. The tutor stored his notes in a note block. One person stored his notes on a loose sheet (agenda).

Additional Observations

- The chairman wrote on his copy of the agenda, but these were no notes that were intended for keeping; annotating the agenda was only used for tracking the progress of the meeting. At the end of the meeting, the annotated agenda was discarded.
- The meeting had the structure of a normal project meeting, in which progress was discussed and new tasks were distributed between the attendees. No technical issues were discussed in detail.

4. Board meeting student association 09/21/05 (12:45)

Number of attendees: 3

Expected type: formal

Types of peers: students (board members)

Guided Observations

1. **Note-taking means:** All notes were taken using pen and paper only.
2. **Use of previous notes:** The chairman partly used minutes from the previous meeting and partly prepared notes (especially prepared for this meeting).
3. **Note upkeep:** The chairman and one other person wrote keywords, so upkeep was easy. The minutes secretary wrote clearly and slowly: the meeting was regularly paused by the secretary in order to be able to keep up with minute-taking.
4. **Cognitive strain:** The secretary indicated that in some cases it was possible to write and follow the meeting, but in other cases he had to choose between either note-taking or following the meeting. Whenever the secretary considered information that he could not follow to be important, he paused the meeting.
5. **Note sharing:** No note sharing whatsoever.
6. **Note storing:** Everybody stored his notes in a note block.

Additional Observations

- Comes to attention: the agenda was not printed on paper (as is usually done), but written on a white board.
- The minute-taking was actually divided between chairman and minutes secretary. The minutes secretary wrote down details about every issue and the chairman made a global to do list as well as a list of all main decisions that were made.

5. OGO meeting department Computer Science 09/21/05 (15:30)

Number of attendees: 6

Expected type: formal

Types of peers: students + tutor

Guided Observations

1. **Note-taking means:** The minutes secretary used a laptop. Three other people used pen and paper. The others didn't take any notes. One person made notes on his copy of the agenda. The chairman wrote on the agenda for keeping track, but discarded the annotated agenda at the end.
2. **Use of previous notes:** One person used some notes from a brainstorm-session for illustration purposes.
3. **Note upkeep:** The minutes secretary appeared to have no problems at all with keeping up. The others only wrote down keywords in order to reduce the overhead of note-taking.
4. **Cognitive strain:** The minutes secretary indicated that usually the cognitive strain is very low, making it possible to type and still follow everything that is said. Mumbled or unclear pronunciations or utterances, however contribute to a substantial increase in cognitive strain. In such cases, the secretary asks to repeat what was said.
5. **Note sharing:** No note sharing whatsoever.
6. **Note storing:** The minutes secretary typed and stored the minutes in LaTeX markup. Some of the more elaborate markup was not done during the note-taking itself, but added at the end of the meeting. The tutor stored his notes in a note block. One person stored his notes on a loose sheet (agenda).

6. OGO meeting department Computer Science 09/28/05 (12:00)

Number of attendees: 6

Expected type: formal

Types of peers: students + tutor

Guided Observations

1. **Note-taking means:** The minutes secretary used a laptop computer and the tutor used pen & paper. The other attendees didn't take any notes at all.
2. **Use of previous notes:** No previous notes were used at all.
3. **Note upkeep:** The tutor only wrote essential information, often in keywords, which made the upkeep relatively easy for him. The minutes secretary was generally able to keep up with the discussion, but he had to pause the meeting a few times, because people started to raise to many issues and handling them too fast. Also, he had to ask for certain data that was vague or that wasn't mentioned (but should be) at all during the meeting.
4. **Cognitive strain:** The minutes secretary explained that he did not really have problems with cognitive strain. In some rare cases, he had trouble choosing between listening/talking/thinking along and writing/typing. He explained that he had a method for reducing the cognitive strain: he just wrote stuff down as quickly and dirty as possible and only later beautified the minutes. This helped him reducing the effort of thinking how to write things down in a neat way during the meeting.
5. **Note sharing:** No notes were shared at all.
6. **Note storing:** The tutor stored his notes in a notebook. The minutes secretary created his notes in word and made a PDF file afterwards which he mailed around. Some people kept the previous minutes in their mailbox, others just read them and then discarded them. Nobody kept the notes in an organized way on his system or in some repository.

Additional Observations

- Only the minutes secretary and the tutor made notes. Upon asking the others why didn't (need to) make notes, the general answer was that they trusted that their minutes secretary would write down **everything** relevant.
- The tutor did not use any (previous) notes from the tutor meeting.

7. OGO meeting department Computer Science 09/28/05 (13:30)

Number of attendees: 5

Expected type: formal

Types of peers: students + tutor

Guided Observations

1. **Note-taking means:** The tutor as well as the minutes secretary used pen & paper.
2. **Use of previous notes:** In principle, no previous notes were used, but the group used their project schedule for reference to progress, problems, to do items etc.
3. **Note upkeep:** Everybody was perfectly able to keep up with the discussion in a relaxed way (no strained note-taking).
4. **Cognitive strain:** The minutes secretary -who had to write the most detailed notes- claimed that he had no problems at all keeping up with an OGO meeting. However, he said that he often had trouble keeping up with other, less structured or more decision-rich/name-rich meetings. Little input came from the secretary to the meeting.
5. **Note sharing:** No sharing at all.
6. **Note storing:** The tutor stored his notes in a notebook. The minutes secretary stored his notes temporary in a notebook. He explained that he would type the minutes into a computer. The digital version of the meeting would be mailed and archived in a central repository that is accessible to all attendees.

Additional Observations

- Only the minutes secretary and the tutor made notes. When asked why they did not make notes, the most people answered that everything would be in the minutes which they normally obtained relatively soon after the meeting. One person told that he normally took personal notes during OGO meetings. He also told that before a meeting, he normally constructed a list of issues that should be addressed during the meeting.
- Again, the tutor did not use his (previous) notes from the tutor meeting.

8. Ad hoc meeting 09/29/05 (11:00)

Number of attendees: 2

Expected type: informal

Types of peers: both PhD students

Guided Observations

1. **Note-taking means:** Pen & paper
2. **Use of previous notes:** One student used own writings (not from a previous meeting, but from own ideas, prepared for this meeting)
3. **Note upkeep:** Relatively good. Only keywords.
4. **Cognitive strain:** No problems according to the participants. It was interesting to observe though, that the discussion often stopped completely when some important keywords were written down.
5. **Note sharing:** Some of the self-prepared notes were shown to the other. One student often uses the notes of the other one after and in between meetings.
6. **Note storing:** Loose sheet; the student who took notes indicated that he would keep this paper for a week. Some notes may be digitized for longer keeping and easier search/structuring.

Additional Observations / Remarks

- The writing was really chaotic. Bulks of keywords were grouped by drawing lines around them. In this way, a lot of “cells” with keywords in them were created.
- When asked, one student remarked that it would be desirable to preserve discussions in more detail, especially when more than two people attend a discussion and when many good ideas are uttered in a spontaneous way and often at high speeds.
- Another desire the student brought forward is that it is important to preserve what others said when you get a good idea. It would be nice to be able to write out your own idea while the others talk, knowing that you can later re-study what they said; so you don't have to interrupt the others and by doing that possibly dampen a creative discussion.

9. Planned progress- and work-discussion 10/12/05 (11:00)

Number of attendees: 2

Expected type: informal

Types of peers: A PhD student and his daily supervisor.

Guided Observations

1. **Note-taking means:** The supervisor used pen & paper only. The student used pen & paper as well as the PC to record input from the supervisor on a PowerPoint presentation that he had created.
2. **Use of previous notes:** The supervisor used notes from previous meetings that were on a few loose sheets. The student searched in a few previous e-mails that he sent/received as well as in a few memo's that he had made.
3. **Note upkeep:** The supervisor took a moderate amount of notes, but could keep it up relatively good, because he listened and then wrote down information that was relevant to him in a very compressed form (not keyboards, but a point-by-point summary). The student once wrote a few keywords in his notebook and occasionally edited the PowerPoint presentation, while the supervisor waited, so there were no problems.
4. **Cognitive strain:** There is some kind of protocol: first something is discussed; if a person thinks that what was discussed is relevant for him, he takes his pen and starts writing; the other person finishes what he said and waits with continuing the discussion until the writer is ready with writing. If the writer needs more information to write down, he asks for clarification which leads to a short interchange of words and the writer writing the additions. Every such note-taking action completely stops the dialogue.
5. **Note sharing:** No personal notes were shared. Editing the PowerPoint presentation may be considered collaborative note-taking.
6. **Note storing:** The student stored his notes in a notebook and in the form of changes to his PowerPoint presentation. The supervisor stored his notes on a few loose sheets.

Additional Observations

- The meeting was twice interrupted by a telephone call. The student -who had to wait- reminded the supervisor -who had the call while he was talking- what he was talking about after a call was finished.
- The student indicated that with a work-meeting -like this one- the cognitive strain is not so high, because he explains what he did to the supervisor and the supervisor gives remarks on the material. With other meetings, however, (especially with his promoter) the student has not only to talk a lot, but also has to record a lot of things that are said by the supervisor/promoter. In such situations the student would like to have the possibility to re-study the things the others said that he missed because of the high cognitive strain.
- The student indicated that he was prepared to take certain hurdles (like putting on a head set or going to a special meeting room) if the meeting is important enough for him and if it would enable him to re-study things that he missed.

10. Ad hoc meeting 10/12/05 (13:50)

Number of attendees: 2

Expected type: informal

Types of peers: Both PhD students.

Guided Observations

1. **Note-taking means:** Pen & paper.
2. **Use of previous notes:** E-mail from the daily supervisor, a draft version of the paper to be produced, one student used an own summary from several papers on one broad topic.
3. **Note upkeep:** The speed is relatively low and the topic is familiar for both attendees, so no new context is introduced quickly; hence the upkeep is pretty good.
4. **Cognitive strain:** No problems at all according to the participants. Since most of the talk was on a very specific topic, very little personal notes were taken, so there was indeed almost no hold-up in the meeting due to note-taking.
5. **Note sharing:** One student writes out some items and sends the outline to the other student.
6. **Note storing:** On a loose sheet (the e-mail of the supervisor).

Additional Observations

- One person forgot the name of a researcher whose work he wanted to reference; he indicated that it would be good to have his notes present.
- Next to text, underlining is used (or words as well as the existing words in the e-mail are underlined); also circling was used.
- Words were also striked through; picture or arrows were also drawn, but only used as a doodle.
- One person took a book to show a few pages to the other. He indicated that it would be handy to make an instant picture of the relevant pages and add this to your notes. Maybe you could also “draw” on a book-page to highlight stuff that should be added digitally to the notes.

B.1 Summary

- Formal as well as informal meetings have been observed.

- Formal meetings were held in groups of 3-7 people and informal meetings were usually held with two people.
- Pen and paper were note-taking tools that were used in abundance. Sometimes a laptop was used for typing notes.
- In formal meetings there was usually a dedicated secretary. Sometimes, the secretary was perfectly able to keep up with the note-taking and even participate well in the meeting.
- Sometimes, the secretary even had trouble with the note-taking itself and had to slow down or stop the meeting several times to complete note-taking tasks on a certain point.
- In formal meetings, normally the discussion could continue when the secretary was able to keep up or didn't stop the meeting if he was not able to keep up.
- There were always people who took personal notes, even during formal meetings, regardless of the fact that there was a secretary who took notes for the group.
- During informal meetings, sometimes the entire dialogue stopped whenever notes had to be taken.
- Note-sharing was hardly observed.
- Previous notes are used sporadically and the storing of notes is highly variable (from throwing away to storing on a laptop and sending the notes around).

Appendix C

System Description

Abstract

This appendix elaborates on the implementation of the prototype addressed in Chapter 3.

C.1 Overview

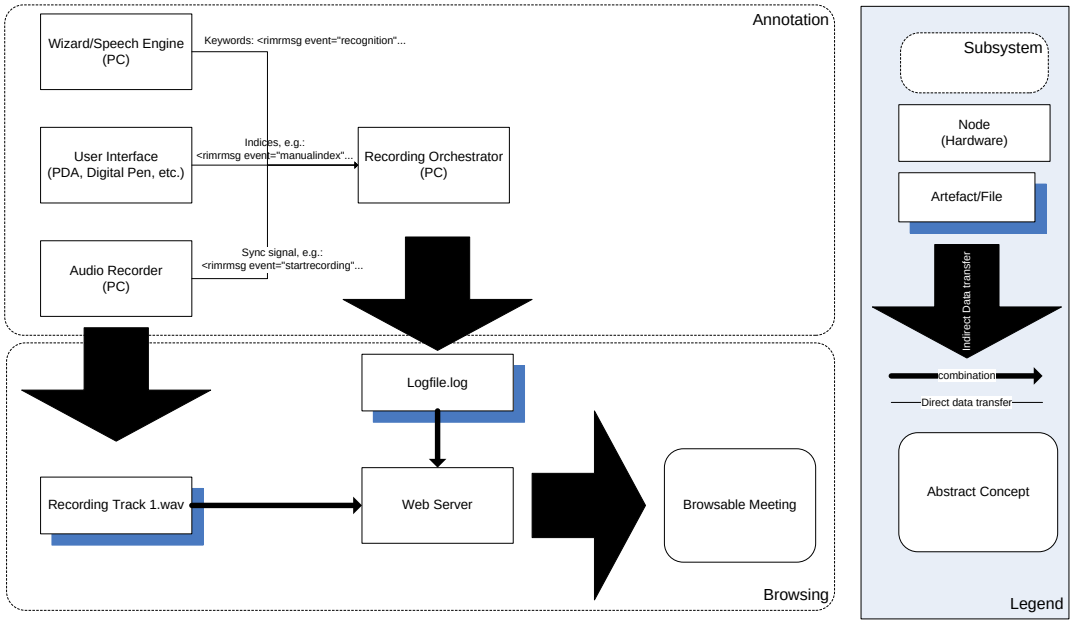

Figure C.1: Detailed system overview of the current prototype. In the setup used for experimentation, Wizard/Speech Engine was realised by means of a “chat”-like program called *keywordentry*, User Interface was realised by means of a Windows Mobile program called *RIndex*, Audio Recorder was realized by means of *Adobe Audition* together with a plug-in for synchronisation, the Recording Orchestrator was realized by means of a program called *recserver*, and the Browsable Meeting on the Web Server was realized by means of a Flash program called *RIMRBrowser*.

Figure C.2: The microphone: this device has a kidney-shaped sensitivity field of about 15 centimetres, which filters out the speaker's voice only and minimises the volume of the speech of others, providing a good way to separate the voices of different meeting participants.

Figure C.3: The recording hardware: this Terratec sound processor can handle 8 simultaneous mono or 4 simultaneous stereo audio inputs.

```
00:04:11:762: <rimrmsg event="recognition" recduration="00:00:02:797">as a first question</rimrmsg>
00:04:15:928: <rimrmsg event="recognition" recduration="00:00:02:031">more an open question</rimrmsg>
00:04:21:516: <rimrmsg event="recognition" recduration="00:00:01:453">added value</rimrmsg>
00:04:24:170: <rimrmsg event="recognition" recduration="00:00:02:406">experience problems</rimrmsg>
00:04:31:742: <rimrmsg event="recognition" recduration="00:00:03:610">assume disruptive in your office</rimrmsg>
00:04:37:260: <rimrmsg event="recognition" recduration="00:00:03:344">should be asked at some point</rimrmsg>
00:04:38:883: <rimrmsg event="recognition" recduration="00:00:01:328">perhaps later</rimrmsg>
00:04:42:558: <rimrmsg event="recognition" recduration="00:00:03:047">second question more</rimrmsg>
00:04:49:829: <rimrmsg event="recognition" recduration="00:00:05:985">could you describe your experience with the system</rimrmsg>
00:04:53:695: <rimrmsg event="recognition" recduration="00:00:02:360">two levels ansering</rimrmsg>
00:04:59:073: <rimrmsg event="recognition" recduration="00:00:04:657">one level interface level</rimrmsg>
00:05:01:427: <rimrmsg event="recognition" recduration="00:00:02:156">other personal level</rimrmsg>
00:05:09:940: <rimrmsg event="recognition" recduration="00:00:02:859">web questionnaire</rimrmsg>
00:05:13:395: <rimrmsg event="recognition" recduration="00:00:02:797">paper questionnaire</rimrmsg>
00:05:20:706: <rimrmsg event="recognition" recduration="00:00:03:875">every day collect this information</rimrmsg>
```

Figure C.4: Snippet from a *Recording Orchestrator* logfile.

Figure C.1 gives an overview of how the different software units in the current prototype work together to provide annotation and browsing functionality. The annotation and browsing components are coupled by means of file exchange (indirect data transfer):

- for a certain meeting session, the annotation component produces one or more audio files in .WAV format together with one accompanying log file (see Figure C.4) that consists

Figure C.5: The environment: on this table in the kids lab of the Institute of Perception Research building, all users held their meetings.

of timestamped, XML formatted descriptions that represent annotation or speech engine events during the meeting;

- the browsing component reads the audio file(s) and the log file and produces a browsable record of the meeting.

C.1.1 Annotation

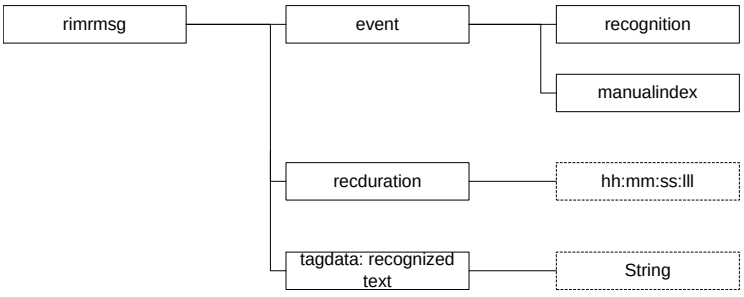

Figure C.6: The structure of the XML files. h is the number of hours, m the number of minutes, s the number of seconds, and l the number of milliseconds; every tag is prefixed with a timestamp which makes it possible to calculate start/stop times and the points in time when an manual index has been made.

The *Recording Orchestrator* is the central part of the annotation component. It consists of a single laptop running a command line C++ program that listens for one or more devices such as audio recorders or speech recognition engines. The different parts of the annotation component (speech engine, audio recorder etc.) are loosely connected to the Recording Orchestrator by means of exchange of XML formatted (as described in Figure C.6) text messages over TCP/IP

sockets. This has been done to keep the component easily reconfigurable (i.e. change parts, such as substituting the current simple audio recorder for a microphone array device or substituting the wizard with an actual speech recognition engine) and extensible (i.e. add parts, such as a speaker activity detection engine which can determine which speaker speaks at which time). Upon receival of a message, the Recording Orchestrator timestamps the message and stores it in a log file.

The choice for XML instead of a binary encoding has been made in anticipation to changing/adding many new messages. XML is human readable and little documentation is needed on the message format. Sending XML messages over the network is feasible. A sentence like this one can already be considered to be very long for a single recognition event (be it wizard or speech engine); yet this sentence is only 219 bytes long (assuming simple character encoding in bytes). On the TU/e network the average latency of 1000 test messages is about 0.14 ms and the average bandwidth upon transferring approximately 300 MiB of data is 1.9 MiB/s. Thus, one such long message takes less than 0.2 ms to transfer and easily fits within the time scales of the human perception domain, accommodating the real-time requirement. For wireless networks, the figures are way less reliable, but from practical use (a WiFi communication channel has been used to send messages from a PDA to the Recording Orchestrator) there were no cases in which the wireless latency or bandwidth was seriously insufficient for the purpose of sending the messages.

The *audio recorder* part consists of a PC with a multitrack audio recording system (see Figure C.3) which can handle multiple microphones. The software used to record the audio was Adobe Audition (formerly known as CoolEdit). Since the users of the prototype who hold a meeting are seated in a classical “round table” configuration (see Figure C.5), special microphones (see Figure C.2) have been used to separate the speech of different speakers. To prevent high network load of multiple audio streams simultaneously, the recording files are written locally to the hard disk of the PC that runs Adobe Audition and at the end of the meeting transferred to the Recording Orchestrator. To automate synchronization of annotation events with the audio recording, a C-program in the form of an Adobe Audition plug in (called *AuditionCV2RIMRController*) has been written, such that if recording is started, the Recording Orchestrator receives a message and starts a timer which is used as a source for timestamps on annotation events.

The *wizard* part consists of a PC running an instant messaging-like program (called *keywordentry*) built with GTK+ which the wizard can use to type in words.

The *user interface* part most often used consists of a PDA (connected through WiFi) running Windows Mobile 5 and a C#.NET program (called *PDAPresenter* or *Smart Indexer*) that enables the user to create indices in the form of bookmarks (simply press a button), candidate phrase selects and self-written phrases. A digital pen system has also been tested, but unfortunately the hardware and drivers did not allow interactive integration with the Recording Orchestrator. The first prototype that emulated the idea of the PDAPresenter was a crude Java-program called *keywordpresenter*.

C.1.2 Browsing

The browsing component’s enabling software is a Flash program whose functionality is implemented in Action Script. A small C++-program (called *xmlindexer*) that functions as XML parser/generator has been made to produce an index file (using the log file from the Recording Orchestrator) that can be read conveniently by the browser software. This index file is placed together with the browser software and the audio files in a directory on a web server. This provides users with a browsable meeting that is accessible from the world wide web.

C.2 Design Iteration Details

This section gives an overview of the development of the prototype's implementation.

C.2.1 Iteration 1.1 and 1.2

A simple pair of java programs. One for sending text manually and one for receiving and displaying test (keywordpresenter).

Functionality: none, just a mockup.

System/Architecture: communication done by sending keywords as bare strings without any meta-information over a TCP stream oriented socket. No information is saved; the only goal is to get an impression of system interaction.

C.2.2 Iteration 1.3

Experimentation with PDAs and development of the main annotation component structure.

Functionality: No successful communication link due to old PDA hardware (missing Bluetooth stack elements).

System/Architecture: Communication over Bluetooth serial port with an incomplete Bluetooth stack is very clumsy and not uniformly usable, like IP sockets.

C.2.3 Iteration 1.4

Integration of a WiFi PDA and realisation of decoupling of various hardware devices.

Functionality:

- well working keyword entry program created with Minimalistic GNU for Windows and GTK– windowing toolkit
- preservation of recognised keywords in SQLite database engine component

System/Architecture: extensible infrastructure due to independent message server (called Recording Orchestrator) created with standard components (C++/STL + Glib– + libXML++) and capable of supporting multiple clients (speech engines and user interfaces), i.e. multiple wizards/speech engines and multiple meeting participants can connect to Recording Orchestrator which combines their annotation information for a certain meeting. Messages are sent over TCP stream oriented sockets. Message types are encoded in XML format.

C.2.4 Iteration 1.5

Automatic synchronisation and preparation of data from database.

Functionality:

- the audio recorder is integrated as a device with the Recording Orchestrator and can send start/stop signals via a C-program plug in for the Adobe Audition recording software.
- a small utility (MinGW, C++) constructs an XML data file from the contents in the SQLite database.

Note: a technical problem during the experiments prevented the storage of all keywords during the meeting; presumed fault: the Recording Orchestrator crashes inexplicably due to a standby in the PDA.

C.2.5 Iteration 1.6

Addition of bookmark possibility.

Functionality: added the message type “manual index” which represents a manually generated bookmark by pressing a button.

Note: a technical problem during the experiments prevented the storage of keywords after a certain point in the beginning of the meeting; real fault found: SQLite database stops saving after a database connection has been open for some time.

C.2.6 Iteration 1.7

Problem fixes.

Functionality: instead of saving keywords in the SQLite database, all events are simply flushed to a log file with timestamps.

C.3 Browsing Component

C.3.1 Iteration 2.1

Functionality: a flash application that can be accessed from anywhere in a web browser. Reads data (annotated recording) from a web server and presents this to the user.

C.3.2 Iteration 2.2

Processed some minor remarks (changed labels/buttons, added padding, etc.).

C.4 Summary

The loose coupling between the various hardware and software components as well as the relatively simple design and compact implementation make it easy to adapt the prototype and to transfer knowledge on further development of the prototype.