

16. APPENDIX F: STUDENT PACK

Extracted from:

J. K. Hyde (2002) Multi-Modal Usability Evaluation. PhD Thesis. Middlesex University.

Student Pack for COM 3210

Modality Information

Contents

Introduction

Modality Definition

Taxonomy

Modality Examples

Modality Types and Properties

Modality Clashes

Modality Methodology

Coursework Details

Appendices

Cover Sheets

Introduction

This pack is designed to give you all the information you need about modalities. It starts off by talking about what a modality is and how it is defined. It then talks about how to identify them, the various types and properties of modalities, and the kinds of clashes that can occur. Then the methodology is explained in full, with examples. The coursework instructions are included, and an appendix gives the information in an easily accessible format.

This work is part of my research into modality, and by doing the coursework using this methodology you will be directly contributing to my studies.

It may be that you require clarification and amplification of the material contained in this pack over the course of the next few weeks. If so, please email me on j.k.hyde@mdx.ac.uk, and I will reply as soon as possible.

Modality Definition

The word modality means different things to different disciplines. In psychology it refers to the human sensory channels, mainly the auditory (hearing), visual (seeing) and haptic (touching and feeling) senses. In computer science it refers to the input and output devices of a computer, such as a keyboard, mouse or monitor. Other disciplines have other definitions. When these areas converge in an area like human-computer interaction, it is difficult to gain a clear understanding of what is being discussed, because the definitions cannot be reconciled. Only when modalities can be usefully defined can their usability properties can be established, and potential usability problems identified.

I have defined modality as **a temporally based instance of information perceived by a particular sensory channel.**

This definition has three parts to it: time, information form, and sensory channel.

The temporal nature of the information is important in terms of processing by either the user or the system. Different temporal forms will require different resources; for example a repetitive or continuous form of information is going to require different resources to a dynamic stream of information which is constantly changing. Time is also important in the ordering of modalities.

By stating the form of the information, we are beginning to develop an understanding of what form the information takes. A word is processed differently to a picture. In addition to this, we have to allow for information which may be expressed but in a real-world rather than symbolic manner, for instance a picture or the sound of voices.

Lastly, the use of particular sensory channels makes the definition user-dependent. This is because the form of the user is relatively static in comparison to the form of the system. System input and output possibilities are constantly changing and developing as new technologies are implemented, whereas the sensory channels have been standardised for the past few thousand years and are less likely to change suddenly.

Taxonomy

By defining modality in this way, it is now possible to create a way of identifying modalities, by means of a taxonomy, or classification system. This taxonomy has three dimensions.

Sensory dimension

The first dimension takes the three human senses, audio, visual and haptic, through which humans interact with computers. They are the three main channels through which information is perceived and communicated, and thus any taxonomy of modality should include them. The three categories in this dimension are **acoustic**, **visual** and **haptic**.

An **acoustic** modality is one that utilises the acoustic sensory channel. An example of an acoustic modality could be speech or music.

A **visual** modality is one that utilises the visual sensory channel. An example of a visual modality is a display or a picture.

A **haptic** modality is one that involves touching or movement. An example of a haptic modality could be sign language or pressing buttons.

Temporal dimension

The second dimension refers to the temporal nature of the modality, and whether a modality is **discrete**, **continuous**, or **dynamic**.

A **discrete** occurrence of a modality is one that is unchanging within its occurrence, which is brief, and thus needs only preliminary processing. An example of a discrete occurrence could be the press of a button, the single ring of a doorbell, or the appearance and quick disappearance of a message on a screen.

A **continuous** occurrence of a modality is one that is repeated exactly the same one or more times. It is distinguished from the discrete occurrence, because it allows for some degree of inspection. An example of a continuous occurrence could be the continuous pressing of a button, the continuous ringing of a telephone, or the continuous display of a photo on a screen.

A **dynamic** modality is one that changes in content within its occurrence, which may last for some time, and thus requires continuous processing. An example of a dynamic modality could be a song being played, or a video.

Information Form

The third dimension of the taxonomy relates to how the information is presented. This can be divided into three: **lexical**, **concrete** and **symbolic**.

Lexical is where the information is in the form of text. An example of a lexical modality could be sign language, speech, or text.

Concrete is where the information is in the form of reproduction of a real life object. An example of a concrete modality could be actual noise made by a dog, photos or videos, or physical vibrations as picked up by a lie-detector.

Symbolic is where the information is a representation of something rather than an actual reproduction of it. An example of a symbolic modality could be a fire alarm, gestures with meaning other than sign language, icons on a windows display, or the use of colour to denote danger.

Modality Examples

Any modality can now be identified in terms of its sensory form, information form, and temporal form. Here I will use the example of a telephone in order to show how modalities are identified.

Interaction with a telephone makes use of various modalities:

The bell ringing

The user lifting the receiver

The caller speaking

The bell ringing uses the acoustic sensory channel, is symbolic in its information form, in that the bell sound means “telephone”, and is continuous, in that it will keep making the same noise until someone answers the call. The modality can therefore be written as acoustic-symbolic-continuous (written aco-sym-cont for short).

The user lifting the receiver uses the haptic sensory channel, is symbolic in its information form, in that the user is lifting something and thereby signalling acceptance of the call, and is discrete, in that it is one action that is not prolonged or repeated. The modality can therefore be written as haptic-concrete-discrete (written hap-con-dis for short).

The caller speaking uses the acoustic sensory channel, is lexical in its information form, because the caller is (hopefully!) using words, and is dynamic, in that the words change and are not repeated. The modality can therefore be written as acoustic-lexical-dynamic (written aco-lex-dyn for short).

However, here I have not distinguished between what the system is doing and what the user is doing. It is important to be able to show both the modality that is expressed by either the system or user, and the modality which is received by either the system or user. For example, using this idea, the modalities used in the interaction with the telephone increases:

The bell ringing (expressed by the system)

The bell ringing (received by the user)

The user lifting the receiver (expressed by the user)

The user lifting the receiver (received by the system)

The caller speaking (expressed by the system)

The caller speaking (received by the user)

There is not always a direct pairing between expressive and receptive modalities in this way, but most often there is. For example, I might shout at my computer, thus using the modality UE aco-lex-dyn, but the computer would remain blissfully unaware of this. Unfortunately!

In order to denote user, the letter U is used. In order to denote system, the letter S is used. In order to denote if a modality is expressive, the letter E is used. In order to denote if a modality is receptive, the letter R is used.

So the modalities listed above could be written as:

SE aco-sym-cont (the bell ringing, expressed by the system)

UR aco-sym-cont (the bell ringing, received by the user)

UE hap-sym-dis (the user lifting the receiver, expressed by the user)

SR hap-sym-dis (the user lifting the receiver, received by the system)

SE aco-lex-dyn (the caller speaking expressed, by the system)

UR aco-lex-dyn (the caller speaking received, by the user)

Modality Types and Properties

Having identified the modalities in an interaction, it is now possible to talk about the types and properties of those modalities.

Modalities are not always at the same level of **granularity**, or abstraction. For example, imagine a telephone conversation, where a person is talking. That modality would be written as **UE aco-lex-dyn**.

However, when a person is talking they can often add more information to what they are saying by varying their tone, the volume, and even the accent in which they are speaking. These are all extra modalities, because they are different layers of information on the original modality of speech. By taking it down to a lower level of granularity, we can add in further

modalities. Obviously this depends on what level of analysis is appropriate: normally, when analysing a telephone conversation, the UR aco-lex-dyn modality would be sufficient.

Imagine a Scotsman talking on the telephone. He is irritated, so he has raised his voice. He is in a hurry, so he is talking fast. The modalities would be as follows:

UE aco-lex-dyn: the user talking. This would be the overall, **composite** modality. It is made up of other **atomic** modalities:

UE aco-lex-dyn: the words the Scotsman is using

UE aco-sym-dyn: the volume of his words

UE aco-sym-dyn: the speed of his words

UE aco-con-dyn: his accent

The volume and speed are symbolic, because they represent something in the interaction: they show that the user is irritated and in a hurry. The accent is not important in this case, it just is, therefore it is concrete.

The **atomic** modalities are those which make up a composite modality. They are at the lowest level of analysis at that stage. An atomic modality may in turn be made up of other modalities if the analysis was taken down to a lower level of analysis. This atomic modality would therefore become a composite modality, and those modalities which it was comprised of would then be known as atomic modalities.

The volume, speed and accent modalities are known as **dependent** modalities, because they are dependent on the aco-lex-dyn modality, or the actual words of the conversation. They are properties that can only exist if there is a modality originally.

Suppose it is you that the Scotsman is talking to. You would be receiving the modalities:

UR aco-lex-dyn: the words the Scotsman is using

UR aco-sym-dyn: the volume of his words

UR aco-sym-dyn: the speed of his words

UR aco-con-dyn: his accent

However, imagine that you don't realise that he is speaking fast because he is in a hurry. You might think that that is how he normally talks. To you, then, that modality would be UR aco-con-dyn, with the speed represented as concrete, because it does not represent anything of importance to you.

This is known as a **mis-match**, where information expressed in a modality is not correctly received. It most often occurs because the recipient is not aware that the modality exists in this

context (if you knew that the Scotsman usually spoke slowly, you would probably realise that he was speaking faster than usual because he was in a rush).

Another type of mis-match is when information is received that was not intended. For example, imagine that the Scotsman usually spoke quickly, but that you were not aware of this. Hearing him speak, you might think that he was in a rush, when actually it was his normal speed. He would be giving the modality SE aco-con-dyn, and you would be receiving UR aco-sym-dyn.

Imagine that both volume and speed increased if you were speaking in a hurry. You would be expressing the following modalities: UE aco-sym-dyn and UE aco-sym-dyn. However, the system would only need to identify one of those modalities in order to realise that you were in a hurry. It would not matter if it was volume or speed, either would do. In this case we have a **redundant** modality. This is where exactly the same information is being given in an extra modality. In some cases this is important in order to emphasise a point. In other cases it can show that there is overload, and that the modality is not needed.

Modality Clashes

Modalities can often interfere with each other, and cause processing problems both for humans and systems. Various types of clashes can be identified, and some examples are given below. It should be borne in mind that there can be more than one clash occurring at any given time. The clashes, with the exception of the first type, are user-dependent in the examples given; however, the same principles can be applied to a system once the system configuration and its limitations are known.

Physical clashes

This is where the modalities clash in terms of the physical characteristics of the user or system. For example, a computer may not be able to process two streams of voice input simultaneously. Similarly, a user cannot say two different things at the same time. It is not possible to detail exactly when system clashes will occur, because this will depend on the configuration of each particular system. It is possible to identify human physical clashes, although users with special needs will have different problems.

One very important point is that a user cannot look in two different directions at the same time. Only if things are within a field of vision can they be visually processed. This is obviously important if a system is expressing two modalities, in two different spatial locations, where the user would not be able to see them both at once.

Lexical clashes

There are various restrictions on what a user can and cannot process lexically. A user cannot process two different lexical inputs or outputs at the same time; for example, reading a book and listening to a sports commentary, writing an address and reciting the alphabet, or writing a letter and listening to the news on the radio. Only if the lexical information is the same would a user be able to do this.

Temporal Clashes

The temporal nature of the modalities can have an effect on how well they are processed. It is difficult to combine two dynamic modalities unless the modalities relate closely to each other, and can be thought of as a composite modality. An example of this is reading someone's lips and listening to what they say. If the two modalities (lip-movement and speech) are coordinated, with no time delay or synchronisation problems, then these two modalities are combined and do not interfere with each other. However, if there is a time delay between the two, it becomes extremely difficult to process the modalities, and problems can occur.

The same problem can occur with combinations of dynamic and discrete modalities, and discrete and discrete modalities. This is due to the user focusing on one of the modalities, and

thus losing the information given in the other modality. For example, imagine that I am listening to a radio commentary, and the doorbell rings. For a moment, my attention switches from the commentary to the doorbell, and I do not process what the commentary says for that brief instant.

There is no problem when combining continuous modalities with other continuous modalities, or with discrete or dynamic modalities. This is because continuous modalities allow a certain degree of freedom of inspection.

The following table clarifies this:

	Dynamic	Continuous	Discrete
Dynamic	possible interference	fine	possible interference
Continuous	fine	fine	fine
Discrete	possible interference	fine	possible interference

Semantic clashes

These occur when information is provided in two modalities, which do not in themselves clash, but where the content of the modality causes problems. For example, having the word butterfly under the picture of a dog, or someone whispering conspiratorially about the weather.

Clash unless expert

This is when the combination of modalities is such that only an experienced person would be able to either receive or express them properly. This type of clash is one that cannot be explicitly defined, and must be identified by the analyst using good judgement as to what the user may or may not be capable of. An example of such a clash is trying to change gears in a car whilst going round a bend and talking to a passenger. An expert driver would have no problem with this, whilst a learner driver might find that combination impossible.

Modality Methodology

The methodology is quite straightforward and consists of several stages. Each stage will be examined in turn, using an example interaction in order to clarify the procedure. It is expected that the analyst will repeat some of the stages as their knowledge of the task and interaction develops. It is not expected that the analysis will follow this methodology in a completely linear fashion. This methodology aims rather to support the analyst by making apparent the

various stages necessary for a complete analysis of the modalities used in an interaction and the potential usability consequences of such modalities.

The example analysis will be of the modalities used in the following task:

to answer a telephone and have a conversation with the person on the other end.

The telephone in this case will have a flashing light as well as a ringing noise in order to attract the user's attention to the incoming call.

Stage 1. Define the task that is to be analysed

Before the task can be analysed it has to be defined. In this case it is defined as:

To answer a telephone and have a conversation with the person on the other end.

The selection of this particular task as opposed to any other task now has to be justified and explicitly stated. For example, is this a standard task for this device? Is the task complicated enough to show a wide selection of potential problems?

Since the telephone is primarily a communication device, being able to receive incoming calls successfully is a standard task. Problems with this task would show that the device interface has been badly developed.

Stage 2. Modality lists

At this stage all the modalities used by the user and the system should be listed, both descriptively, i.e. "phone ringing", and notationally, i.e. "SE aco-sym-cont". This list will probably grow as the analysis is iterated. The level of granularity should be discussed at this point, in terms of the level of detail appropriate to the analysis. Again, this may change as the analysis is iterated and further considerations become apparent.

User modalities:

(both expressive and receptive modalities should be listed here)

Hear phone ringing: U R aco-sym-cont

See phone ringing: U R vis-sym-cont

Pick up receiver: U E hap-sym-dis

Speak to voice on end of phone: U E aco-lex-dyn

Listen to voice on end of phone: U R aco-lex-dyn

System modalities:

(both expressive and receptive modalities should be listed here)

Phone ringing: S E aco-sym-cont

Phone ringing: S E vis-sym-cont

Voice on end of phone: S E aco-lex-dyn

Note that the buttons on the telephone have been ignored for the purposes of this analysis since they are not necessary for the completion of this task.

Granularity:

The analysis was done at a reasonably high level of granularity. For example, it was not considered necessary to break down the speech modalities (UE aco-lex-dyn, SE aco-lex-dyn) into further atomic modalities, since this would not have any bearing on the usability of the device.

Stage 3. Define the user, system and environment profiles

At this point, the analyst should explicitly state the characteristics of the user, system and environment that can reasonably be expected to have a bearing on the usability of the interaction.

User Profile:

The user is able-bodied adult human, who speaks the same language as the caller. There are no physical or cognitive impairments.

Various assumptions have been made about what the user knows:

The user has had experience of using a telephone before

The user knows the sound of a telephone

The user knows what the flashing light means

The user knows to pick up the receiver to answer it

The user knows to say something to alert the person on the other end that the phone has been answered

System Profile:

The telephone is a simple handset, which needs to be picked up to be answered. A call coming in is indicated by a continuous ringing noise and by a continuous flashing light. The phone will ring at a certain volume in a certain manner, the light will flash at a certain rate, the handset has to be held at a certain distance from the mouth of the user for effective communication, and the handset has to be held at a certain distance from the ear of the user for effective communication.

Environment Profile:

The environment is an office setting, with subdued noise levels, and good lighting. There is only one system in the immediate vicinity of the user.

Stage 4. Profiles compared to modality listings

The profiles of the user, system and environment are compared against the lists of modalities in order to check for obvious discrepancies.

User profile compared with System modalities: there are no obvious discrepancies.

System profile compared with User modalities: there are no obvious discrepancies.

Environment profile compared with User and System modalities: there are no obvious discrepancies.

Stage 5. Interaction modality listing

At this point the interaction should be written out in sequence and in terms of the modalities used. Check that for every expressive modality, there is a corresponding receptive modality. There will not always be a pairing, but in most cases there will be.

It would probably be easiest for the analyst to start by writing out the interaction sequence in terms of the actions, then breaking it down into a description of the modalities, and finally in terms of the notation. For this example, the writing has gone directly to the notational stage.

1. [SE aco-sym-cont] **and** [SE vis-sym-cont]
 phone ringing *phone flashing*

[UR aco-sym-cont] **or** [UR vis-sym-cont]
user hears phone ringing *user sees phone flashing*

precon: [SE aco-sym-cont] precon:[SE vis-sym-cont]
phone ringing *phone flashing*

 precon: user looking at phone

The interaction starts with the phone ringing and flashing, and the user becoming aware of this.

Each part of the interaction sequence is numbered to allow for easy discussion of the sequence.

The modalities are written both notationally, with square brackets round them, in order that their constituent parts can be examined, and descriptively, within the star symbols (*) in order that the analyst can keep track of which modalities are being discussed.

Both of the system modalities occur at the same time, so they are written as joined with a logical **and**.

Only one of the user modalities needs to occur at this stage, so they are written as joined with a logical **or**.

There are ordering considerations at this point to take into account. The telephone has to ring or flash before the user can become aware of the ringing or flashing. Therefore the user modalities have the system modalities as **pre-conditions** (shortened to “precon” in the example given).

There is another pre-condition that the user should be looking at the phone in order to see if it is flashing. All visual modalities have this pre-condition.

2. [SE aco-sym-cont] **and** [SE vis-sym-cont]

 phone ringing *phone flashing*

[UR aco-sym-cont] **or** [UR vis-sym-cont]

user hears phone ringing *user sees phone flashing*

precon: [SE aco-sym-cont] precon:[SE vis-sym-cont]

phone ringing *phone flashing*

 precon: user sees phone flashing

[UE hap-con-dis]

user picks up phone

precon:[UR aco-sym-cont] **or** [UR vis-sym-cont]

 user hears phone *user sees phone flashing*

The phone is ringing and flashing, the user is aware of this, and has picked up the receiver. The pre-condition to the user picking up the receiver is that the user is aware of the phone ringing and flashing.

3. [UE hap-con-cont]

 user holds phone

 precon:[UE hap-con-dis]

 user picks up phone

[SE aco-sym-cont]

line quiet

precon:[UE hap-con-dis]

user picks up phone

[UR aco-sym-cont]

user hears line quiet

precon:[UE hap-con-dis]

user picks up phone

The user is now holding the phone, which is different to picking up the phone. The line is giving no tone, indicating there is probably someone on the other end, and the user is aware of this. The line being quiet is dependent upon the user picking up the phone.

4. [UE aco-lex-dyn]

user talks

precon:[UE hap-con-cont]

user holds phone

[SR aco-lex-dyn]

user talks

precon:[UE aco-lex-dyn]

user talks

[UE hap-con-cont]

user holds phone

precon:[UE hap-con-dis]

user picks up phone

The user is still holding the phone, and is now talking to the person on the other end. In order for the user to start talking, they must be holding the phone. There is an ordering condition in that the person on the other end of the phone cannot hear the user until the user has started talking.

5. [SE aco-lex-dyn]

caller talks

precon:[UE aco-lex-dyn]

user talks

[UR aco-lex-dyn]

caller talks

precon:[SE aco-lex-dyn]

caller talks

precon:[UE hap-con-cont]

user holds phone

[UE hap-con-cont]

user holds phone

precon:[UE hap-con-dis]

user picks up phone

The user is now listening to the caller speaking. This is conditional upon the caller talking and the user holding the phone.

Stage 6. Add in clashes, etc.

Now that the modalities have been established, and the sequence and pre-conditions explicitly stated, it is possible to look for the properties of the modalities and the potential clashes.

1. [SE aco-sym-cons] **and** [SE vis-sym-cons]

phone ringing

phone flashing

[UR aco-sym-cons] **or**

[UR vis-sym-cons]

user hears phone ringing

user sees phone flashing

precon: [SE aco-sym-cons] precon:[SE vis-sym-cons]
phone ringing *phone flashing
precon: user looking at phone

At this point it can be discerned that the two system modalities are **redundant**, that is, exactly the same information (incoming call) is being given in two different ways. There are no clashes, and the only other property to be aware of is that the user will have to be looking at the telephone in order to see it flashing.

Because this was such a simple example, there were no further properties or clashes found in the rest of the sequence.

Stage 7. Assess the use of modalities

The analyst should assess the modalities used at this point in order to see if any are under-utilised, and if any of the properties flag issues that should be addressed.

This interface makes use of all three senses. However, the use of a flashing light is redundant, since the user would hear the phone ringing (the environment is quiet and the user has no special physical constraints). Also, the user has to be looking at the phone in order to see the flashing, which makes it less effective than using sound.

Stage 8. Final report

Finally, write up into a report, assessing the overall usability of the system as a result of analysing the interaction in this way. Include conclusions and recommendations.

The system seems very usable. No problems on a usability front were found, which is to be expected from such a simple system. However, the use of a flashing light seems unnecessary given the user and environment profiles.

Further Examples

The previous example was simple and straightforward, in order to show how the methodology works. However, interactions can often be more complex, and involve many modalities. Here is an example of the methodology for a task using the wordprocessing package Word 6 for a Mac. Only the most relevant sections of the methodology have been included.

Stage 1. Define the task that is to be analysed

To change a paragraph of an open document from Times font to Courier font.

Stage 5. Interaction modality listing

The Word 6 wordprocessing package for the Mac has a lot of information on the screen at a time. There is the text space, where text is entered. There are the scroll bars which border this space. There is the menu at the top of the window giving options such as File, Edit, View; and there is a row of icons and text information at both the top and bottom of the text space. There are thus many visual modalities being expressed by the computer at once. This can make it quite complicated to write out the interaction modality listing.

Effectively what will happen is that user will only be looking at certain of the modalities displayed, but the system will be expressing all of them. Therefore, the ones that user is not looking at can be expressed at a very high level of abstraction, in order to show that they exist, but without unnecessary detail.

A portion of the interaction sequence below explains this:

1.

[SE vis-sym-cont]

scroll bars

and

[SE vis-sym-cont]

menu options

and

[SE vis-sym-cont]

icons at top of text space

and

[SE vis-sym-cont]

information strip at bottom

and

[SE vis-sym-dyn]

mouse cursor moving

precon: [UE hap-con-dyn]

user moving mouse

and

[SE vis-lex-cont]

text

[UE hap-con-dyn]

moves mouse

and

[UR vis-lex-cont]

scanning text

and

[UR vis-sym-dyn]

watching mouse cursor

precon: [UE hap-con-dyn]

user moving mouse

The system is displaying information on the screen. This information is represented at a very high level of abstraction. The user is only looking at two of these modalities, and is moving the mouse.

The list of modalities used here might seem lengthy, but imagine if all the composite modalities were written in terms of their atomic modalities. The list would be incredibly long, and it would be almost impossible to effectively analyse.

As the user switches attention, the high-level modalities can be described at a lower level of detail. For example, the section of the screen giving the menu options [SE vis-sym-cont] would then be written as follows:

[SE vis-sym-cont] *area containing menu options*

[SE vis-lex-cont] *File*

[SE vis-lex-cont] *Edit*

[SE vis-lex-cont] *View*

[SE vis-lex-cont] *Insert*

[SE vis-lex-cont] *Format*

[SE vis-lex-cont] *Tools*

[SE vis-lex-cont] *Table*

[SE vis-lex-cont] *Window*

[SE vis-lex-cont] *time display*

[SE vis-sym-cont] *system option*

[SE vis-sym-cont] *help option*

[SE vis-sym-cont] *current packages option*

and the rest of the screen could be described as:

[SE vis-sym-cont] *scroll bars*

[SE vis-sym-cont] *icons at top of text space*

[SE vis-sym-cont] *information strip at bottom*

[SE vis-sym-cont] *mouse cursor*

[SE vis-lex-cont] *text*

Coursework Instructions

This coursework is worth 40% of your final coursework mark (i.e. 20% of your overall mark for this module).

The coursework is to be handed in on Thursday 12th November before close of Student Office. Marks will be deducted for late submission. Please use the cover sheet supplied at the back of this pack.

The report is to take the form of an analysis of two interactive devices, as detailed below. You are to work in pairs on the analysis of the devices, but the final written report is to be completed individually.

You are also required to hand in any notes made whilst working on the coursework at the same time as you hand in the report. These notes will **not** be assessed, and will be used only as information towards my research. Please use the cover sheet supplied at the back of this pack.

The report is to consist of:

An analysis of the modalities used in the following task:

to buy an adult return ticket to St Pancras at 10am on a week-day from Bounds Green tube station, using a ten pound note. You may assume that there is sufficient change in the machine.

An analysis of the modalities used in the following task:

using the Alhambra heating control panel, located at <http://www.ndirect.co.uk/~thomas.green/devices/>, set the heating controls for the whole week (Monday through to Sunday), with the heat coming on at 7am each day, and going off at 10pm each day.

The report should be clearly written, with each stage of the methodology detailed, and final usability considerations discussed, for each analysis.

Appendices

A short summary of the information detailed in the pack is given here.

Taxonomy

Modalities are identified according to three dimensions:

Three dimensions: sense, information form, time.

Sense: acoustic (aco), visual (vis) and haptic (hap).

Information form: lexical (lex), concrete (con), symbolic (sym).

Time: discrete (dis), continuous (cont), dynamic (dyn).

They are further distinguished as being user (U) or system (S), and expressive (E) or receptive (R).

Modality Types and Properties

Modalities can be composite, atomic, dependent, mis-matched and redundant. Analyses can be done at differing levels of granularity.

Modality Clashes

There are several types of clashes: physical, lexical, temporal, semantic, and clash unless expert.

Methodology

The methodology is quite straightforward and consists of several stages. It is expected that the analyst will repeat some of the stages as their knowledge of the task and interaction develops. It is not expected that the analysis will follow this methodology in a completely linear fashion. This methodology aims rather to support the analyst by making apparent the various stages necessary for a complete analysis of the modalities used in an interaction and the potential usability consequences of such modalities.

Stage 1. Define the task that is to be analysed

Before the task can be analysed it has to be defined.

The selection of this particular task as opposed to any other task has to be justified and explicitly stated. For example, is this a standard task for this device? Is the task complicated enough to show a wide selection of potential problems?

Stage 2. Modality lists

At this stage all the modalities used by the user and the system should be listed, both descriptively, ie “phone ringing”, and notationally, ie “SE aco-sym-cont”. This list will probably grow as the analysis is iterated. The level of granularity should be discussed at this point, in terms of the level of detail appropriate to the analysis. Again, this may change as the analysis is iterated and further considerations become apparent.

User modalities:

(both expressive and receptive modalities should be listed here)

System modalities:

(both expressive and receptive modalities should be listed here)

Granularity:

Stage 3. Define the user, system and environment profiles

At this point, the analyst should explicitly state the characteristics of the user, system and environment that can reasonably be expected to have a bearing on the usability of the interaction.

User Profile:

System Profile:

Environment Profile:

Stage 4. Profiles compared to modality listings

The profiles of the user, system and environment are compared against the lists of modalities in order to check for obvious discrepancies.

User profile compared with System modalities:

System profile compared with User modalities:

Environment profile compared with User and System modalities:

Stage 5. Interaction modality listing

At this point the interaction should be written out in sequence and in terms of the modalities used. Check that for every expressive modality, there is a corresponding receptive modality. There will not always be a pairing, but in most cases there will be.

It would probably be easiest for the analyst to start by writing out the interaction sequence in terms of the actions, then breaking it down into a description of the modalities, and finally in terms of the notation.

Number each part of the interaction sequence to allow for easy discussion of the sequence.

Write the modalities both notationally, with square brackets round them, in order that their constituent parts can be examined, and descriptively, within the star symbols (*) in order that the analyst can keep track of which modalities are being discussed.

If modalities happen at the same time they should be connected with the logical **and**.

If there is a choice between modalities, they should be connected with the logical **or**.

Take into account any pre-conditions relating to ordering or the use of the visual sense.

Stage 6. Add in clashes, etc.

Now that the modalities have been established, and the sequence and pre-conditions explicitly stated, it is possible to look for the properties of the modalities and the potential clashes.

Stage 7. Assess the use of modalities

The analyst should assess the modalities used at this point in order to see if any are under-utilised, and if any of the properties flag issues that should be addressed.

Stage 8. Final report

Finally, write up into a report, assessing the overall usability of the system as a result of analysing the interaction in this way. Include conclusions and recommendations.

COM 3210 Coursework

Modality analyses of two interactive devices

Module Leader: Ann Blandford

Name:

Student Number:

COM 3210 Coursework Notes

Notes made whilst doing modality analyses of two interactive devices

Module Leader: Ann Blandford

Name:

Student Number:

16. APPENDIX F: STUDENT PACK..... 231