Integrating Modality Theory in Educational Multimedia Design

Lorna Uden and **Russell Campion**
School of Computing,
Staffordshire University, UK
l.uden@staffs.ac.uk

Abstract

Multimedia technology is increasingly being used as an alternative way of delivering instruction. According to researchers, the success of multimedia is due to the dual-coding aspect of the information processing theory. In applying dual-coding principles, different information has to be coded in different media in order for people to learn effectively. Designers of interactive multimedia applications are faced with thousands of different combinations of input and/or output modalities of information representation. Each single modality or multimodal combination has its own specific capabilities and limitations for representing or conveying information. It is important to be able to select the right combination of modalities for a given application. This paper describes Modality Theory and how it can be used to guide designers developing interactive multimedia applications. Given any particular set of information which needs to be exchanged between the user and system during task performance in context, it is essential to identify the input/output modalities which constitute an optimal solution to the representation and exchange of that information. This paper shows how Modality Theory can be applied successfully to design interactive multimedia learning applications.

Keywords

Multimedia, Education, Learning, Modality theory

Introduction

The use of multimedia technology has offered an alternative way of delivering instruction. Interactive Multimedia has the potential to revolutionise the way we work, learn and communicate (Stemler 1997). With interactive multimedia, the learning process becomes active, not passive and it ensures that users are doing, not simply watching. Learning with multimedia is a process, rather than a technology that places new learning potential into the hands of users.

Gagne's nine events of learning serve well as a framework for the successful development of educational multimedia applications. However, with all the additional capabilities of the growing number of multimedia applications, the design of these applications is not simple. Although interactive multimedia design is related to traditional and computer-aided learning systems, many of its aspects are arguably different from sequential media and computer-based instruction, as well as from hypertext (Park & Hannafin 1993). Because of the large numbers of different media options available to choose from, it is important that designers select the right combinations of modalities for a given application. How do we know what combination of input/output modalities to use for a given application? We believe that the right combination of modalities for an application must be based on a theoretical foundation, i.e., Modality Theory (Bernsen 1994).

This paper argues that principles from modality theory have important implications for the design of interactive multimedia learning systems. Modality theory can be used to guide designers to choose the optimal unimodal modality from the vast array of alternatives. The paper begins with a brief review of multimedia for learning, followed by a brief review of methods for modality theory in section 3. Section 4 presents a case study describing how principles from modality theory are used in the design of an interactive learning system. The paper concludes with farther recommendations for research in the use of modality theory for designing multimedia learning systems.

Benefits of Multimedia for Learning

The use of computer-based multimedia in learning has been growing rapidly. The explosion of this use has been attributed to the assumption that interactive multimedia helps people learn (Najjar 1996). Studies have shown that computer-based multimedia can help people learn more information better than traditional classroom lectures (Bagui 1998).

Several factors have been attributed to the success of multimedia in helping people to learn. Firstly, there is a parallel between multimedia and the 'natural' way people learn, as explained by the Information Processing Theory. The similarities between the structure of multimedia and the information processing theory account for a large part of the success of learning with multimedia. This is due mainly to the dual coding aspect of the information processing theory. Dual coding refers to using more than one code in the learning process. According to Najjar (1996), dual coding contributed much to the increase in learning through multimedia. Several studies have shown that two media improve learning better than one medium (Mayer & Anderson 1991; Shih & Alessi 1996; Parlangeli et al 1999). Dual coding not only helps in terms of allowing a person to absorb information from the environment using two channels, it also helps in reducing cognitive load in a person's working memory.

Secondly, information in computer-based multimedia is presented in a non-linear hypermedia format. The nature of hypermedia allows learners to view things from different perspectives. Hypermedia systems also allow users to choose information freely. In addition, information in hypermedia can be grouped in chunks. Thirdly, computer-based multimedia is more interactive than traditional classroom lectures. Interacting appears to have a strong positive effect on learning (Najjar 1996). Fourthly, another feature of multimedia-based learning is that of flexibility. Multimedia programs are flexible in terms of how they may be used in classrooms, by individuals or small groups.

There is empirical evidence (Najjar 1996) that interactive multimedia information helps people learn. Multimedia information is most effective when (a) it encourages dual-coding of information, (b) when the media support one another, and (c) when the media are presented to learners with low prior knowledge or aptitude in the domain being learned.

Modality Theory

Although interactive multimedia systems are now an essential part of most computer-based learning environment, very little is known about how to design such system. Some pragmatic guidelines are known from previous reports (Faraday & Sutcliffe). We believe that guidelines are required to cover selection of media resources for representing different types of information and presentation design. In addition, they must address the key issues of selective attention, persistence of information, concurrency and preventing information from over loading.

According to Bernsen (1994), Modality Theory addresses the following general information-mapping problem:

> "Given any particular set of information which needs to be exchanged between the user and system during task performance in context, identify the input/output modalities which constitute an optimal solution to the representation and exchange of the information."

There are believed to be thousands of modalities in existence, both input and output, that can be incorporated into interface designs (Bernsen and Bertels 1993). However, realistically it could be argued that there are not that many media in existence, as there are only really a few major media prototypes. It is only when these major media prototypes are examined at the 'sub-prototype' level that thousands of media types arise, such as in the case of graphs; scatter-plot, categorical, line, stacked bar, bar, pie, box, and histogram etc. To select an optimal unimodal modality from this vast array of alternatives is difficult, due to each modality having a set of information representation characteristics, making it good for the representation of certain information types, and bad for others. The combination of two or more of these modalities exacerbates the problem, as, when several modalities (both input and/or output) are involved, media interference needs to be taken into consideration.

Current multimedia research attempts to address modality design by creating methods that remove the ad-hoc nature of solutions by providing theoretical frameworks for developers to follow. The problem these methods must solve, if viewed in the most basic terms, can be regarded as the information mapping problem. This problem requires that a mapping exist between task requirements and a set of usable modalities. Within this problem the task must be extensively considered; otherwise mappings will be crude and insensitive to the scenario of the task. Work must concentrate on viewing the information in the context of task goals. Hence, in a sense, the information-media pairing is linked to the objectives of the task, and not considered in isolation from it. Evidently modality theory that profiles informational scenarios in relation to media representation abilities is essential.

The Modality Theory of Bernsen (1994) provides us a basis for analysing arbitrary input/output modality types and combinations as to their capabilities of information representation and exchange. Various methodologies based on the Modality Theory have been developed. A brief review of each is described.

The Roskilde method

The main objective of the method is to provide a theoretical and methodological basis for tracking the information mapping problem by finding optimum modalities for information exchange between a user and computer. The method consists of five sequential steps that are addressed in two phases. The method’s first step is used to elicit details of information exchange between the user and computer during task performance. This is used to form information requirements to guide the process of modality selection. As output, the step produces two sets of results: firstly high level information required to solve the information-mapping problem, and secondly a small set of representative tasks.

The second step of the method analyses the representative tasks gathered at an individual level in as much detail as is possible in order to identify goals, initial states, activities and procedures involved. Information on the task environment (work), the intended users and their experience etc. is also gathered at this step. Therefore the step identifies the input/output information representation needs (and constraints) of the task.

Step three forms a representation of the information acquired through steps one and two, by using a representation such as a DSD (Design Space Development) diagram. The diagram is separated into sections to structure the information gathered (e.g. general constraints and criteria, and hypothetical issues

Step Four of the method considers and applies the theoretically developed framework for representing the basic components of interactive unimodal or multimodal interfaces, i.e., the results of modality theory.

The final step of the method, five, uses high level filtering to perform a trade-off between potential solutions. The result of the step is a solution to the task domain/interface mapping problem together with its Design Rationale. It is however still possible to have many possible solutions as several options may emerge with identical scores from the trade-off process.

Limitations of the Roskilde Method

- Although the Roskilde method provides a framework for modality theory for multimedia design, it has certain limitations:
- there is a lack of detailed information for identification of representative tasks;
- the DSD diagrams are solely textual and do not provide a visual representation of the task;
- the rules of the matching technique are open to interpretation and are not specific enough;
- the rules are cumbersome to use in selecting media;
- it is also questionable whether a complete set of rules can ever be generated.

The Sutcliffe and Faraday method

Sutcliffe and Faraday (1993) describe a method which focuses on the analysis and design of multimedia based interfaces. The method provides simple tools and techniques that interface well together with an emphasis of a methodology, thus making the stages easy to interpret and achieve. It has the aim of creating a single solution of an optimal interface design from the outset. There is no trade-off to be applied. The first phase of the method is to create a task model, and highlight information requirements for each task step. In order for the task step to be successfully completed by the user, information concerning the step must be represented as a TKS. Desired communication effects for the task step are described by a set of information and dialogue acts. The 'effect' of the acts shapes what information is provided to support the task step (the 'base' media resource used) and the presentation techniques which are used to draw the user's attention to the information (presentation analysis- attention controlling media or technique choice). Presentation design draws upon a media resource model (containing the resources available to the system and the information types the resources contain). Validation is then carried out in presentation design by cognitive and design heuristics to ensure that it is appropriate for the demand made of the users' attention and working memory. The final step of the method integrates media on a timeline (Faraday & Sutcliffe 1993).

Limitations of the Sutcliffe and Faraday method

The Sutcliffe and Faraday method provides a simple framework that is logical for designers to follow. However, it has certain limitations. Among theses are:

- it is questionable whether there are enough acts to be expressive enough to cover all multimedia system contexts;
- the guidelines provided for choosing media (selection guidelines and general heuristics) are very simple, lacking specific focus, and practical implementation guidance.

In order to overcome the above limitations, we have developed a Multimedia method known as Multimedia design method (MDM) for designers to use. We will use a case study to show how the method can be used to develop an interactive learning system for learners to learn about the British country garden.

Case Study

We have adopted the Courseware Engineering Methodology (Uden 1999) for the development of our interactive multimedia learning system. We start with the identification of the task information to be learned by the students. The subject domain is represented in the task model using Task Knowledge Structures (TKSs) of Johnson (1992). The task statement for our example is an interactive multimedia package showing the students a British country garden.

A - identifiers and recommended media used for their representation		
1. descriptive	Gesture, speech input, speech output, text input, text output, & video	
2. definition	Gesture, speech input, speech output, text input, text output, & video	
3. dialogue	Gesture, speech output, text output & video	
4. opinion	Gesture, speech input, speech output, text input, text output, & video	
5. summary	Gesture, speech input, speech output, text input, text output, & video	
6. index	Diagrams, realistic drawings, icons, lists, network charts, photo images, & sketches	
7. schedule, timetable, timeline	Charts, lists, speech output, & tables	
8. instruction	Gesture, lists, speech input, speech output, text input, text output, & video	
9. sequence, activity	Animation, diagrams, realistic drawings, photo images, sketches, & video	
10. motion	Animation, diagrams, realistic drawings, photo images, sketches, & video	
11. spatial	Animation, diagrams, realistic drawings, n-models, maps, photo images, sketches, & video	
12. hierarchy, association	Network charts	
13. construction	Diagrams, realistic drawings, n-models, photo images, sketches, & video	
14. dimension	Diagrams, n-models, & maps	
15. object	Diagrams, realistic drawings, gesture, n-models, photo images, sketches, speech input, speech output, text input, text output, & video	
16. characteristic	Realistic drawings, gesture, n-models, natural sounds, photo images, sketches, speech input, speech output, text input, text output, & video	
17. appearance	Realistic drawings, n-models, photo images, sketches, & video	
18. situation	Animation, diagrams, realistic drawings, gesture, photo images, sketches, speech input, speech output, text input, text output, & video	
19. concept	Animation, diagrams, gesture, n-models, photo images, sketches, speech input, speech output, text input, text output, & video	
20. requirements	Lists, tables, text input, & text output	
21. rationale	Charts, diagrams, graphs, lists, network charts, sketches, tables, text input, & text output	
22. trend	Charts, & graphs	
23. mathematics	Charts, graphs, tables, & text output	
24. temporal	Animation, diagrams, realistic drawings, photo images, sketches, & video	
25. intangible	Gesture, speech input, speech output, text input, & text output	
B-identifiers (message detail identifiers)		
B1i - is the message a 'main message'		B1ii - is the message a 'support message'
B2i - is a direct / precise representation required		B2ii - is a representation required that stimulates the imagination
B3i - is background information required to be shown (focus on background and object)		B3ii - is foreground information required to be shown (focus on object only)
C-identifiers (presentation effect requirement identifiers)		**D-identifiers (screen design identifiers)**
C1 - is an add attention / illuminate act required		D1 - title
C2 - is a comparison act required		D2 - interactor
C3 - is a re-statement act required		D3 - element for design purpose (e.g. background)

Table 1: Full set of A, B, C and D-identifiers

The TKS model now needs to be taken and divided into related information. This is important so that we can translate/allocate the task model to separate screen designs. Structuring the presentation can then commence. This activity is achieved by carefully examining the task model and using a line technique to clearly separate related information into groups. To highlight information needs of the intended application, a set of identifiers is used. This is a modified version of Sutcliffe and Faraday's method (1994). Table 1 above shows the full set of identifiers. The identifiers are separated into: A, B, C and D types. The A identifiers are used to model specific information requirements that the media selected must be able to represent. This is similar to Sutcliffe and Faraday's information type identifier and subject based dialogue acts. B identifiers are used to highlight specific issues required of any media selected (e.g. whether background or foreground information is required in the information). To represent such identifiers requires a sound knowledge of media resource program facilities and operations. The C identifiers are attached to illuminate specific information effects that are needed (e.g. drawing attention to an aspect of a base medium). This identifier is derived from the presentation-based dialogue acts of Sutcliffe and Faraday (1994). The D identifier addresses additional aspects of interface design such as the inclusion of a title.

Figure 1: The Country Garden

The screen design ‘The country garden’ (figure 1) contains eight media message elements (A, B, C, D, E, F, G and H shown in Table 2 below). The main message, parts A and B of the figure, provide simple introductory text and a visual description about country gardens. To show this information it was decided that; descriptive, or summary identifiers needed to be modeled for part A, and appearance for part B in the chosen media. It was also decided that the messages were main ones, requiring an imaginative representation for part A and a precise representation for part B. The best representations considered to model the identifiers were text and a photograph. No C-identifiers were attached to the two elements. The full set of details related to identifier attachments and media selection is summarized in Table 2 below.

Part C of the figure provides a flat 2D graphic animation that shows the garden growing over time. To model this several possible identifiers were initially attached: schedule, time-table, time-line, sequence, activity, motion, spatial, appearance and temporal, leading to the selection of 2D animation. It was also decided that the information was a supporting message requiring an imaginative representation with a background as well as foreground focus. C1 attention acts are required to show flower and tree growth. The acts are translated into simple arrows. Part D adds interest to the screen design and is purely supplemental. For this informational concept it was thought that an identifier of definition would be suitable. As with part C this element was also considered to be a support message to the main one. The intention was to use an imaginative foreground focusing media, leading to the use of text output. The remaining parts of the interface design, E, F, G and H were deemed to be related to screen furniture, being; a screen title, background, line and interactive button.

Element	Purpose	Identifiers attached	Media used effects applied
A	Introductory text to country gardens	A1 – descriptive A5 – summary B1i – key message B2ii – imaginative representation B3ii – foreground focus	Text output
B	Introductory visual description of a country garden	A17 – appearance B1i – key message B2i – precise representation B3i – background focus required	Photograph

C	Visual description showing temporal development of the garden throughout the year	A7 - schedule, time-table, time-line A9 - sequence, activity A10 – motion A11 – spatial A17 – appearance A24 – temporal B1ii – support message B2ii – imaginative representation B3i – background focus required C1 – attention acts	2D animation Simple arrows
D	Simple quote	A2 – definition B1ii – support message B2ii – imaginative representation B3ii – foreground focus	Text output
E	Background design	D3 – element for design purpose	Smoothed photograph
F	Title	D1 – title	Simple text embodied in title bar
G	Animation control and subtitle	D1 – title D2 – interactor	Simple text Standard iconic play button
H	Page line feature	D3 – element for design purpose	*Green line design*

Table 2: Summarized table of identifier attachments

Conclusion

Interactive Multimedia Technology has the potential to revolutionise the way we learn. However, the development of these systems is not trivial. Designers are faced with a vast option of media to use. To develop effective interactive learning systems, designers must be provided with principles and guidelines that can help them in the choice of media. We have developed a method based on Modality Theory that can be used to help designers in their development of interactive learning systems. Evaluation of our system by users reveal that they found the application easy to use, fun to learn and stimulating.

References

Bagui, S. (1998), Reasons for Increased learning using multimedia, *Journal of Educational Multimedia and Hypermedia*, 7 (7), 3-18.

Bernsen, N.O. (1994). Modality theory in support of multimedia interface design. *Proceedings of the AAAI spring symposium on intelligent Multi-Media –Modal systems*. Stanford. March. 37-44.

Bernsen, N. O. & Bertels, A. (1993). A methodology for mapping information from task domains to interactive modalities, *working papers in cognitive science and HCI*, University of Roskilde, Denmark, WPCS-93-10.

Faraday, P. & Sutcliffe, A. (1993). A Method for multimedia Interface Design, *People and Computers, HCI'93*, 173-190.

Johnson, P. 1992. Human-Computer Interaction: Psychology, Task Analysis and Software Engineering, Maidenhead: McGraw-Hill

Mayer, R.E. & Anderson, R.B. (1991). Animations need narrations: an experimental test of dual-coding hypothesis. *Journal of Educational Psychology* 83, 484-490.

Najjar, L.J. (1996). Multimedia Information and Learning. *Journal of Educational Multimedia and Hypermedia*, 5(2). 129-150.

Park, I. & Hannafin, M. J. (1993) Empirically based guidelines for the design of interactive multimedia, *Education Technology Research and Development*, 41(3), 65-85.

Parlangeli, O., Marchigiani, E., & Bagnara. S. (1999), Multimedia Systems in distance education: effects of usability on learning. *Interacting with computers*, 12, 37-49.

Shih, Y.F. & Alessi, S.M. (1996). Effects of Text versus Voice on learning in multimedia courseware. *Journal of Educational Multimedia and Hypermedia*, 5(2). 203-218.

Stemler, L.K. (1997). Educational Characteristics of Multimedia: A literature review. *Journal of Educational Multimedia and hypermedia* 6(3/4), 339-359.

Sutcliffe, A. & Faraday, P. (1993). Designing Multimedia Interfaces, *EWHCI'93 3rd Int. Con.* Moscow, 105-114.

Sutcliffe, A. & Faraday, P. (1994). Designing presentation in multi-media interfaces, CHI'94, *Proc ACM Human Factors & Computer systems*, 92-98.

Uden, L. (1999), User-Centered Courseware. *ED-MEDIA 1999: World conference on educational material, hypermedia and telecommunications*. Seattle, Washington, USA, June 19-24, ACCE publication.