

Within the broader goal to generate a presentation that conveys effectively the answer on a multimedia query to the user my accent falls on selecting the optimal modality or combination of modalities for presenting a particular type of information.

Within this goal I formulated the following research question:

Given any particular set of multimedia items retrieved as a result of user query identify the modalities which constitute an optimal multimedia presentation of the retrieved information.

In general, designing which information is to be displayed in which manner is a difficult problem. The solution depends on finding modalities with characteristics especially suited to pertinent characteristics of the information to be displayed and then using them to best advantage.

The approach I adopted so far is the following:

1. Establish sound conceptual and taxonomic foundations for describing and analyzing any particular type of modality and any possible combination of modalities
2. Analyse the combinations of different modalities from a human cognitive processing perspective in order to identify those combinations of modalities that can lead to information overload.
3. Based on the findings in (1) and (2) establish mapping mechanisms between the type of information the presentation will hold (what is to be represented) and the respective modalities which will best convey the information to be presented (the representation itself)

Different modalities have different properties which makes them suitable for representing different types of information.

Each single modality or multimodal combination has its own specific capabilities of representing or conveying information and it is obviously important to be able to select the right combination of modalities for a given application. The question is how this might be done in a principal manner so as to generate multimedia presentations that can optimally convey the information the user requested through a query.

Factors that can influence on the choice of particular modality or combination of modalities

- task goals
- the information to be presented
- nature of human cognition

These are also the factors against which a presentation should be optimal.

Multimodal Discourse structure

The discourse structure determines the ordering and interrelating of the information to be presented . We need additional mechanism to determine which is the optimal modality or combination of modalities for the delivery of each portion of information
Some first questions to be answered:

How are discourse generation and modality allocation interrelated. Will the two processes happen simultaneously or sequentially

It is interesting whether methodologies from natural language processing can be extended to a computational model of multimodal communication and how. One aspect is to study whether and how rhetorical structure can be used to generate multimodal presentations. A starting assumption in this case will be that there is a common representation of what is to be conveyed and for each constituent in this common representation decisions has to be taken on the best modality or even combination of modalities for final presentation. Furthermore, in contrast with only text generation, in multimedia generation the different modalities employed in the whole presentation refer and depend upon each other and that is what creates the richness of a multimodal communication but is at the same time the most difficult to capture and formalize for automatic generation.

Establishing taxonomic foundations

Modality theory developed by Niels Ole Bernsen is an attempt to provide science base of HCI from the needs of design practice.

It addresses the general problem of optimally allocating different modalities to different types of information. Modality theory provides taxonomization of output modalities.

(Hovy, Arens, 2000) talk about internal semantic structure of modalities and list a number of characteristics of modalities that can be matched against the type of information that needs to be presented to allocate modalities in a multimodal presentation.

These taxonomizations can be used as a departure point for annotating media items in a way that can allow the identification of the right modality or combination of modalities