Motivation
Multimodal systems that support the user by a combination of speech, gesture and graphical-driven interaction are already part of our everyday life. They rely on a static, predefined multimodal interaction setup, where the interaction devices, paradigms and the possibilities of controlling their applications are predefined.

Modeling multimodal systems that support various multimodal setups is an open research issue.

Challenges
1. How to model real multimodal interaction that includes fusion and fission?
2. Have different modalities more in common than they differ?
   - What is a suitable dialogue modeling abstraction?
   - Can design models be executed?
3. What is an appropriate design process?
   - Graceful Degradation vs. Abstract-to-Concrete

Related Work
The CARE Properties
- Complementary denotes several modalities that convey complementary chunks of information.
- Assignment implies that a modality is fixed in the way that the user has no choice in for performing a task with another modality.
- Redundancy indicates that the same piece of information is conveyed by several modalities.
- Equivalence of modalities implies that the user can perform a task using a modality chosen amongst a set of equivalent modalities.

Cameleon Reference Framework and UsiXML
Our work considers the user interface development abstraction levels of the Cameleon Framework. Different to UsiXML, which offers a standardized markup language for several modalities, we add a behavior specification of interactors and focus on designing multimodal interaction.

Evaluation in E-Learning
Multimodal approaches to learning have been proven to be extremely effective since information introduced aurally, visually and kinesthetically can significantly increase the possibility of understand and remembering information. By a case study we want to prove that our approach can be applied to model a multimodal learning application that can be run in different multimodal setups.