Multimodal Interfaces make use of different modalities to interact with computers or machines. The field includes both analysis and synthesis of multimodal utterances. This course focuses on analysis, i.e., processing input from, for example, speech, gestures, touch, gaze direction, or even biosensors. The goal here is to determine the intent of the interactor from multiple channels and signals in order to perform desired (inter-) actions. In this course, students will learn about examples of multimodal interfaces, their advantages, the underlying terminology and theoretical background. In addition, students will learn the steps necessary for processing both unimodal and multimodal input. As core content, building on this, the fusion of multimodal signals is taught using the example of synergistic speech-gesture interfaces as well as its integration into an interactive real-time system. This includes on the one hand typical aspects of multimodal dependencies, e.g. temporal and semantic entanglements, and on the other hand prominent approaches to perform multimodal fusion on decision level. In the accompanying exercise, the theoretical contents are deepened by a practical examination of the development of a synergistic speech-gesture interface for a virtual environment.

Intended learning outcomes
After participating in the module courses, students are able to recognize basic application scenarios for multimodal interfaces. They remember subject-specific approaches and can apply them to adequate problems. They can summarize, compare and explain different approaches. They can apply available tools to typically occurring tasks and know their advantages and disadvantages. Furthermore, you can independently familiarize yourself with complex technical systems as well as independently develop problem-solving proposals, communicate these in a team and integrate them in a prototype.