Verbal language can be presented in two modes: visual as legible (written) texts, and audible as (spoken) speech. Legible texts have widely been used in cartographic publications. Until now, the use of speech in combination with maps for the public has almost only been used in car navigation systems. This use is intended for short-term orientation.

However, this paper examines the combination of text and speech with thematic maps in learning applications designed for a sustainable cognition.

The use of verbal language will be analysed to find out if and how this combination would improve the acquisition of long-term knowledge for spatial patterns. These spatial patterns are for instance the distribution of thematic data such as population variables for a country’s districts. This research deals with small scale thematic maps to make sure that the users can only retrieve the spatial patterns out of the map, the text/speech, and their combinations.

Based on the assumptions and theories of cognitive sciences, the users of interactive map applications in combinations with verbal language information should have an advantage for the acquisition of spatial knowledge over the single map users.

The combinations imply certain aspects:
- Codes for information transfer are verbal, pictoral, or numerical. A verbal code (e.g. text) combined with a pictoral code (e.g. map) results in “multicodality”.
- Modes for codes are visual or audible. A combination of a visible code (e.g. map) with an audible code (e.g. speech) represents “multimodality”.
- Users have different preferences for choosing a code or a mode. This is often a subjective choice, and is not equal with the sustainable learning effect. Also the user’s level of experience in map and text work has to be considered.

The external representation of information and knowledge is different from our internal representation.

To compare verbal language with map language one has to analyse the principles of linguistic and cartographic semiotics. The definition of equivalence of map elements with verbal elements has to be made by syntactic and semantic structures. Maps are representations which always contain a spatial structure, while words have to build a spatial structure externally.

There are various functions of texts with reference to maps as well as vice versa. The reading strategy can be detached or embedded. The temporal coincidence of map and text presentations can be synchronous or not.

Results of the ongoing research will be presented and discussed.