Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities

Article in Computers in Human Behavior · March 2003

DOI: 10.1016/S0747-5632(02)00015-8

Jan L. Plass
New York University

Detlev Leutner
University of Duisburg-Essen

PERGAMON

Computers in Human Behavior 19 (2003) 221–243

Computers in Human Behavior

www.elsevier.com/locate/comphumbeh

Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities

Jan L. Plass[a,*], Dorothy M. Chun[b], Richard E. Mayer[b], Detlev Leutner[c]

[a]*Educational Communication and Technology Program, Steinhardt School of Education, New York University, 239 Greene Street, New York, NY 10003, USA*
[b]*University of California, Santa Barbara, Santa Barbara, CA 93106, USA*
[c]*Institut für Psychologie, Erfurt University, PF 900 221, 99105 Erfurt, Germany*

Abstract

When do multiple representations of information in second-language learning help and when do they hinder learning? English-speaking college students ($N=152$), enrolled in a second-year German course, read a 762-word German story presented by a multimedia computer program. Students received no annotations, verbal annotations, visual annotations, or both for 35 key words in the story. Recall of word translations was worse for low-verbal and low-spatial ability students than for high-verbal and high-spatial ability students, respectively, when they received visual annotations for vocabulary words, but did not differ when they received verbal annotations. Text comprehension was worst for all learners when they received visual annotations. Results are consistent with a generative theory of multimedia learning and with cognitive load theory which assume that multimedia learning processes are executed under the constraints of limited working memory.

Keywords: Cognitive load; Verbal ability; Spatial ability; Multimedia learning; Individual differences; Second language acquisition

1. Introduction

This paper is concerned with the question of what role cognitive load plays in multimedia learning, and, in particular, how cognitive load affects the way learners

* Corresponding author. Tel.: +1-212-998-5658; fax: +1-212-995-4047.
E-mail address: jan.plass@nyu.edu (J.L. Plass).

PII: S0747-5632(02)00015-8

with different cognitive abilities process verbal and visual information. The learning environment used for this study is a second-language multimedia program that requires learners to read a text in a foreign language. Vocabulary words in the text are annotated with visual and verbal information in the form of pictures, video clips, and text translations of the word in English. The learning task is to comprehend the foreign language text, which involves two levels of learning: understanding individual vocabulary items and overall comprehension of the text. This paper focuses on the question of how learners with different cognitive abilities, namely verbal and spatial abilities, are able to acquire new vocabulary items from the text and how they are able to comprehend the text depending on the type of annotations they select while reading.

Instructional multimedia software and online materials with multimedia elements enjoy increasing popularity on all levels of education. Our theoretical understanding of the processes of multimedia learning, however, lags behind. When do multiple representations of information in a second-language learning environment help and when do they hinder learning? Recent research has identified conditions under which multimedia instruction can be effective (Mayer, 1997, 2001; Mayer & Moreno, 1998; Mayer & Sims, 1994; Moreno & Mayer, 1999, 2000; Plass, Chun, Mayer, & Leutner, 1998), informing the development of Mayer's (1997, 2001) generative theory of multimedia learning, which focuses on the processing of verbal and visual information. These and other studies have found that multimedia instruction can indeed lead to improved learning outcomes, although multimedia materials might, under certain circumstances, produce detrimental learning effects for some learners (Kalyuga, Chandler, & Sweller, 2000; Mayer, 1997; Mayer & Sims, 1994). The effectiveness of learning in multimedia environments depends on differences in the way learners gather and process information. The dimensions of individual differences that so far were found to have moderating effects on the learning outcome in multimedia learning include visualizer/verbalizer learning preferences (Jonassen & Grabowski, 1993; Leutner & Plass, 1998; Plass et al., 1998) and spatial and verbal ability (Jonassen & Grabowski, 1993; Mayer, 1997; Mayer & Sims, 1994; Pellegrino, Alderton, & Shute, 1984). One probable cause for the detrimental effects that were found in some studies is the cognitive load which is imposed on the learner when using multimedia information for learning and the limited processing capacity of the human working memory (Baddeley, 1986, 1992; Chandler & Sweller, 1991; Miller, 1956; Sweller, 1988, 1994, 1999).

1.1. Generative theory of multimedia learning

Mayer (1997, 2001) considers the learner to be an active knowledge constructor who selects and processes visual and verbal information from what is presented. Following Clark and Paivio (1991), we define the presentation mode of information as *visual* or *verbal* on the level of information-processing. *Verbal* information is coded in modality-specific verbal codes using the linguistic symbol system, even when it is presented visually (e.g. as printed text). *Visual* information, then, is coded in the form of pictures, static or moving, which are imaginal representations that are analogous or perceptually similar to the events they denote, using the iconic symbol system (Clark & Paivio, 1991).

In order for meaningful learning to occur, according to the generative theory of multimedia learning, the learner must first select relevant verbal information from a text and visual information from an illustration and then construct a text base and image base, respectively. The learner must then organize the verbal information in the text base in a coherent verbal mental representation and the visual information in the image base into a coherent visual mental representation. Then, the learner must integrate the newly constructed verbal and visual representations by creating connections between the corresponding visual and verbal information. These processes are summarized in Fig. 1 (Mayer, 1997).

In cases where the learning material includes visual and verbal information that illustrate a specific idea and where the learner is able to make connections between the verbal and visual representation of these materials, we predict an enhancement of meaningful learning based on this theory of multimedia learning.

An extension of the generative theory of multimedia learning to second-language learning was proposed by Plass et al. (1998). This extension focuses on the learning of individual vocabulary words and on the overall comprehension of a foreign language text and, thus, mainly addresses the linguistic competency of reading

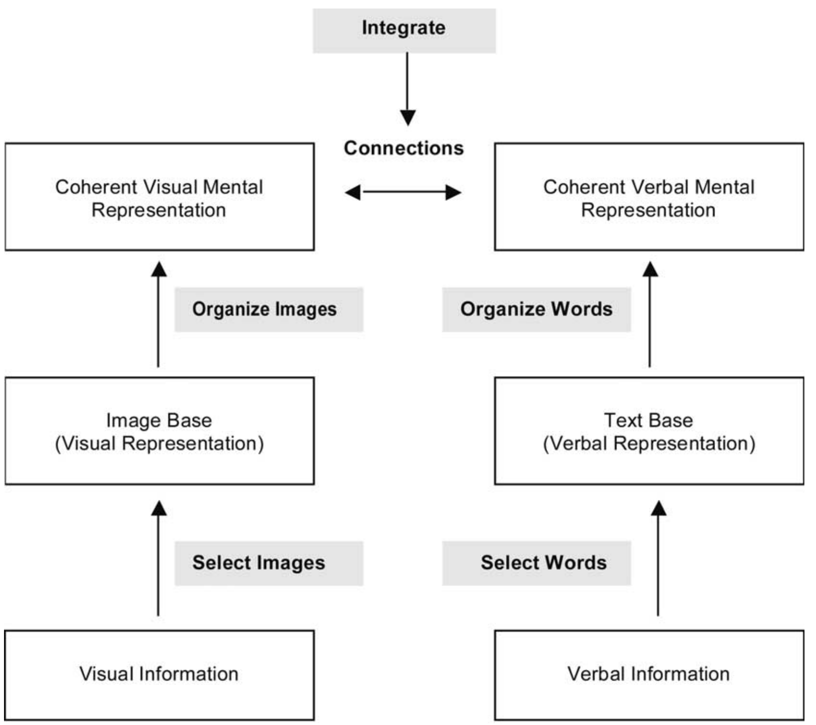

Fig. 1. Generative Theory of Multimedia Learning (Mayer, 1997, 2001).

comprehension. The extension of the generative theory to the acquisition of vocabulary words suggests the existence of two separate verbal systems which are linked by the translations of individual words and a common visual system. There is improved learning for those words coded visually and verbally as compared to words only coded verbally (Paivio & Desrochers, 1980; Paivio & Lambert, 1981). In a study with 103 participants, Plass et al. (1998) found that students' recall of individual vocabulary items is best when they used both visual and verbal annotations for these words, moderate when they used only one type of annotation, and worst when they used neither. These findings support Paivio's (1971, 1990) dual coding theory by establishing a direct connection between words in the native language, a corresponding picture or video, and the translation of the word in the foreign language, a learner establishes two types of retrieval cues for the word in memory.

Reading comprehension, the second focus of the extension of the generative theory of multimedia learning to second-language acquisition, involves the construction of meaning on a higher level than the low-level rote learning of vocabulary (Kintsch, 1998; Mayer, 1984; Wittrock, 1974, 1990). According to Mayer's (1984) and Wittrock's (1990) theories of comprehension, the reading process involves building referential connections between the mental representations of ideas or propositions that have been presented in different modes (e.g. in a verbal mode and a visual mode). Comprehension depends on the successful storage of these connections along with the two forms of mental representation (verbal and visual) of these propositions or ideas in long-term memory. In the study by Plass et al. (1998) cited previously, for instance, learners recalled propositions from the text better when they had both visual and verbal modes of instruction in the text than those who had instruction in only one mode. If learners cannot build the connections between the verbal and visual mental representations of ideas or propositions, then comprehension is hindered and the storage of these propositions might not occur.

In the Plass et al. (1998) study, students were allowed to choose which type of instruction they preferred to receive. However, in many learning environments, visual and verbal information is presented on the screen by default, without giving the learner a choice. Therefore, in the present study, we employed four different treatments in which all learners read the same foreign-language text. Controls received no additional instruction for unknown vocabulary words, one group of students received additional instruction in the visual mode only, one in the verbal mode only, and a fourth group received instruction for unknown vocabulary words both in the visual and verbal mode. The instruction, i.e. the annotations, were especially designed to assist the learners in the process of selecting information, which requires an understanding of the meaning of an individual vocabulary word. While the choice of visual or verbal modes of instruction supported learners' individual differences in their preference to learn with visual information (visualizer) vs. verbal information (verbalizer), the treatment of the present study, where participants have to use the type of annotation available in their treatment condition, might negatively affect students with low cognitive abilities.

1.2. Cognitive load

While the limited capacity of human working memory may have advantages for case-based reasoning and for making inferences based on small samples of information (Bower, 1997), it might impact learning negatively under some conditions for some learners in multimedia learning. These limitations have been known at least since Miller (1956), but have gained a new importance with the increasing use of multimedia information in instructional materials. The cognitive processes in multimedia learning as described by Mayer's (1997, 2001) generative theory of multimedia learning include the processes of selecting visual and verbal information, organizing this information into coherent representations in working memory, and integrating the resulting visual and verbal representations of information by building one-to-one correspondences between them. In order for this integration to be successful, the visual information has to be held in visual working memory at the same time as the verbal information is held in verbal working memory, a condition which Mayer and Anderson (1992) refer to as *contiguity effect*. In addition, the process of integrating the verbal and the visual representations requires cognitive resources (Sweller, 1994, 1999). Since the capacity of working memory is limited, the integration of the visual and verbal representations of the material is constrained by memory load (Baddeley, 1992; Chandler & Sweller, 1991; Mayer, 1997; Sweller, 1994; Sweller, Chandler, Tierney, & Cooper, 1990; Yeung, Jin, & Sweller, 1998). Yeung et al. (1998) found that explanatory notes that were designed to facilitate vocabulary learning led to reduced overall text comprehension, while explanatory notes designed to facilitate text comprehension resulted in reduced vocabulary learning, and that these effects depended on learners' expertise.

Low-ability students, in particular, may not be able to process verbal and visual annotations and build referential connections between them because of the high cognitive load imposed by this processing. Because multimedia learning involves the processing of verbally and visually/spatially encoded information, the corresponding verbal and spatial abilities are especially of interest (Ernst, 1991; Kirby, 1993; Mayer, 1997; Mayer & Sims, 1994).

In order to measure spatial ability, which we define for the purpose of this study as spatial relations ability (Carroll, 1993), we chose the Card Rotation Test from the *Manual for Kit of Factor-Referenced Cognitive Tests* (Educational Testing Service, 1976), a paper-and-pencil test in which the learner must determine whether two figures are the same (i.e. one can be rotated to be identical to the other) or different (i.e. one must be flipped over and rotated to be identical to the other). We chose the Card Rotation Test because (1) it taps an important component of spatial ability that is related to the kind of processing of pictorial material required in this study, (2) it is a speed test that appears to tap aspects of cognitive load relevant in this study, (3) it has been used successfully in previous studies (Mayer, 2001; Mayer & Sims, 1994), and (4) it is a simple and quick test. Carroll (1993) analyzed spatial ability into distinct factors including spatial relations and visualization, consistent with Richardson (1983, 1994) who demonstrated that spatial measures and imagery measures are different. We chose to focus on a measure of spatial relations (rather than a measure of imagery ability) because it is most closely related to the form of spatial ability required in our multimedia learning task—namely, to mentally connect a pictorial representation with a verbal representation (rather than to form a vivid image or manipulate parts of an image). In the multimedia learning task, learners must manipulate a meaningful pictorial representation and make connections between pictorial and verbal representations—skills that seem more closely related to tests of spatial relations than to tests that measure the ability to visualize or analyze the parts of an image. In summary, the Card Rotation Test depends on speeded performance in processing of a meaningful pictorial representation, a skill that is consistent with our interest in cognitive load in multimedia learning.

To measure verbal ability, we chose the Vocabulary Test from *The Nelson–Denny Reading Test (NDRT)* (Riverside Publishing, 1981), a paper-and-pencil test in which the learner must determine the meaning of words printed in the left column on the page. To the right of each of these words, five other words are printed from which the learner must choose the one that has the same meaning or nearly the same meaning as the word on the left. We chose this Vocabulary Test from the NDRT because (1) it taps an important component of verbal ability that is related to the kind of processing of verbal material required in this study, (2) it is a speed test that appears to tap aspects of cognitive load relevant in this study, (3) it has been used successfully in previous studies, and (4) it is a simple and quick test. The aspect of verbal ability measured by this speed test is most closely related to the processing required in multimedia learning task—namely, to mentally connect a verbal representation with a pictorial representation. Vocabulary tests, like the one from the NDRT, correlate highly with the capacity of verbal working memory (Avons, Wragg, Cupples, & Lovegrove, 1998; Gathercole & Baddeley, 1993; Gathercole, Service, Hitch, Adams, & Martin, 1999; Masoura & Gathercole, 1999), and relationships exist between the capacity of verbal working memory and vocabulary learning as well as vocabulary knowledge in a second language (Atkinson & Baddeley, 1998; Papagno & Vallar, 1995). Thus, we expected that the verbal ability test for the English language used in this study would be good predictor for vocabulary learning in a second language (Masoura & Gathercole, 1999). In summary, the NDRT Vocabulary Test depends on speeded performance in processing of a meaningful verbal representation, a skill that is consistent with our interest in cognitive load in multimedia learning.

1.3. Hypotheses

With regard to *vocabulary learning*, cognitive load theory suggests that low-ability learners have to allocate more of their cognitive resources to processing visual and verbal information than high-ability learners. Their cognitive resources might be insufficient to process the verbal information they select from the text when they are also required to process visual information, such as a pictorial annotation for a vocabulary word in the text. Pictorial annotations cause a high cognitive load because they require learners to translate the picture into a meaning (i.e. into words). The generative theory of multimedia learning suggests as an implication of this high cognitive load caused by the visual annotations that low-ability learners may not be able to build the same number of connections between the information in working memory as high-ability learners. Low-ability learners in this study were therefore expected to be less likely to learn the translation of German words than high-ability learners when they were required to select and process both the verbal information of a foreign-language reading text and visual annotations for unknown vocabulary words. This pattern was expected to be present for learners with low-spatial ability and for learners with low-verbal ability. This first hypothesis is based on the idea that low-ability learners experience high cognitive load that prevents them from building connections between the verbal information of a foreign-language reading text and the visual information of the annotations for unknown vocabulary words in working memory and from integrating these mental representations into their mental models. Verbal annotations, which by design impose less cognitive load than visual annotations, were not expected to have this effect on vocabulary learning.

Unlike vocabulary acquisition, which resembles a low level of rote learning, *reading comprehension* involves the construction of meaning and represents a higher-level learning activity. The annotations for the present reading text were designed to help the learners understand the meaning of a particular vocabulary word (i.e. help them *select* relevant information, not to aid the processes of *organizing* mental representations in working memory into propositions or *integrating* these propositions with existing knowledge). Therefore, by forcing learners to process annotations that are designed to support the lexical level of text processing, the process of comprehending the overall meaning of the text might have been hindered since more cognitive resources needed to be expended on the process of vocabulary selection and processing of the annotations. In fact, learners had to expend cognitive resources on processing lexical information even when they would not have chosen to look up the annotation of a word, either because they knew the word before or because they were able to comprehend the text passage without knowing the exact meaning of the word. If the processing of visual annotations indeed caused a high cognitive load, then it could be expected that insufficient resources would be available for organizing and integrating the text information, resulting in lower performance on the comprehension test. A similar effect was found by Yeung et al. (1998) using explanatory notes for readers of different ages and abilities. Therefore, as a second hypothesis, learners were expected to perform worse on comprehension in treatments where they were required to process visual annotations as compared to verbal annotations or no annotations. This effect is expected to be stronger in the visual annotation only treatment than in the combined visual and verbal annotation treatment as the presence of a verbal annotation, a text that provides the unambiguous meaning of the vocabulary word, should result in a reduced reliance of the learners on the processing of the visual annotation to understand the meaning of the word. Because we expect that the type of the annotations provided will focus the attention of the learners on vocabulary acquisition and not on text comprehension, we expected this to be a general effect for all learners that is present across different levels of verbal and spatial abilities.

2. Method

2.1. Participants

The participants were 152 college students who were enrolled in second-year German language courses at three universities in the USA. None of the students were native speakers of German, and all were fluent in English. Students participated in the study as a regular class activity. The mean grade point average of the students was 3.25, SD = 0.51 (with A = 4.0, B = 3.0, C = 2.0, D = 1.0, and F = 0).

2.2. Materials and apparatus

The materials consisted of a Macintosh software program that presented an interactive multimedia version of Heinrich Böll's (1986) short story, "Anekdote zur Senkung der Arbeitsmoral" ("Anecdote Concerning the Lowering of Productivity"), developed by Chun and Plass (1997). The story consisted of 762 words in German, presented in 11 pages. Each page consisted of approximately 50–100 words of text in German presented on the right side of the screen. Several of the words on each page were marked by a degree symbol (°), with 35 words marked overall (as shown in Fig. 2). For each of these words, different types of multimedia annotations were available. Verbal annotations consisted of a text translation of the word. Visual annotations consisted of either a still photograph or a short video clip illustrating

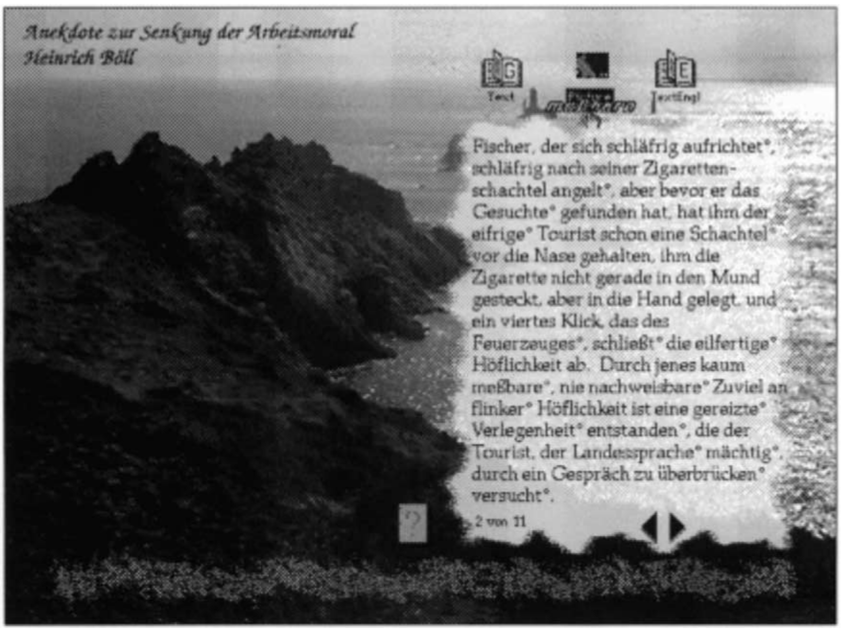

Fig. 2. A screen shot of a page from the program *CyberBuch*.

the word. To access annotations, students moved the mouse over the marked word and pressed the mouse button. One or more symbols appeared at the top of the screen (i.e. icons representing text annotations, pictures, and videos). Students were instructed to drag the word to the icon in order to receive the annotation. For example, when the text translation for the German word *Fischschwärme* was selected, students would see the translation: "schools of fish" on the left side of the screen and would hear the word spoken by a native speaker of German. When the picture option for *Fischschwärme* was selected, the computer presented a picture showing a school of fish and the word spoken by a native speaker of German. When the video option for the German word *Hubschrauber* was selected, the computer presented a 7-s video clip of a helicopter and the word spoken by a native speaker of German.

The annotations were designed to aid the learners in understanding the meaning of individual vocabulary items, not to organize mental representations or integrate them with one another. In other words, the function of these annotations, based on the generative theory of multimedia learning, was to aid the selecting of relevant information rather than organizing or integrating mental representations (see Fig. 1). The multimedia software was designed following the Instructional Systems Design process (see, for instance, Smith & Ragan, 1999) and has been evaluated and revised to reduce the possibility of a material effect. Fig. 2 is a page from the story program. The English translation of this portion of the story is: "...[awakens the dozing] fisherman, who sleepily sits up, sleepily gropes for his cigarettes, but before he has found what he is looking for the eager tourist is already holding a pack under his nose, not exactly sticking a cigarette between his lips but putting one into his hand, and a fourth click, that of the lighter, completes the overeager courtesy. As a result of that excess of nimble courtesy—scarcely measurable, never verifiable—a certain awkwardness has arisen that the tourist, who speaks the language of the country, tries to bridge by striking up a conversation."

The paper-and-pencil materials consisted of a questionnaire, German vocabulary pre-test and post-test, a test of spatial ability, and a test of verbal ability, each typed on 8.5×11 inch sheets of paper. A text comprehension post-test was typed by students on the computer. The questionnaire solicited information concerning the participants' gender, native language, grade point average, and grades in previous German classes. The vocabulary pre-test and post-test were identical and included a list of the 35 annotated German words (see Appendix A). The test of spatial ability was the Card Rotations Test from the *Manual for Kit of Factor-Referenced Cognitive Tests* (Educational Testing Service, 1976), and the test of verbal ability in the learners' native language, English, consisted of the Vocabulary Test from *The Nelson–Denny Reading Test* (Riverside Publishing, 1981). For the text comprehension post-test, students typed a summary of the story in English.

2.3. Procedure

Students were tested in their intact German language classes in groups of 15–20 per session during their normally scheduled class hour. Classes met for 50 min per day, and the entire procedure required two 50-min class periods on two consecutive days. On the first day, all students first filled out the paper-and-pencil questionnaire at their own rates. Students were then given 7 min to do the German vocabulary pre-test. This test was a multiple choice test in which students were given an English word and had to choose the correct German synonym from a group of six words.

Each student was seated in front of a separate Macintosh computer system in a language computer lab. Following collection of the questionnaires and the German vocabulary pre-test, students were given a brief demonstration of the program on a large TV monitor at the front of the lab. Students were told how to turn pages and how to look up a marked word. The instructor demonstrated how to click on a marked word, hold the mouse button down, and, when the icons appeared indicating which type of annotation was available, how to drag the word to the desired icon and release the mouse button. They were instructed that upon selecting an annotation they would hear the word pronounced in German and either see the translation of the word, a picture, or a video clip depicting the word. Students were told that they had to look up all of the marked words before they would be allowed to go to the next page. They were told that the learning task was to comprehend the story. When they were finished reading the story, they were asked to quit the program.

Students were randomly divided into four groups, 38 students each. Students in the control group did not have any words marked in their story; they were instructed simply to read the story and did not receive any annotations. Students in the three treatment groups were instructed to "look up" all of the marked words. Depending upon which treatment group the student was in, different icons appeared and different types of annotations were available for the marked words. Students in the "verbal-only" group saw a dictionary icon appear above the text; when they dragged the word to the icon, they received a text translation of the word on the left side of the screen and heard the word spoken by a native speaker of German via the headphones. Students in the "visual-only" group saw either a picture or a video icon appear above the text; when they dragged the word to the icon, they received either a still image or a video depicting the word on the left side of the screen and heard the word spoken by a native speaker of German via the headphones. Students in the "verbal and visual" group saw both a dictionary icon and a picture or video icon appear above the text. They were instructed to look up both the translation of the word as well as the visual annotation by dragging the word to each of the icons, and they received both a text translation of the word and an image or videoclip, in addition to hearing the word spoken by a native speaker of German via the headphones. Students were instructed that the software would let them advance to the next page only when they had looked up all available annotations.

On the second day, students were given 7 min to complete the German vocabulary post-test, followed by the text comprehension post-test, for which they were given 10 min to type a summary of the story or *recall protocol* in English on the computer. They were then given 3 min for the spatial ability test. Lastly, they were given 4 min for the verbal ability test.

2.4. Scoring

On the German vocabulary pre-test and post-test, the data for each student consisted of whether or not the correct German synonym was selected for each of the 35 marked words. In constructing the text comprehension post-test, two raters independently listed the main idea units or propositions of the text, identifying the central ideas and occurrences in the passage. Thirty propositions were identified (see Appendix B). Each recall protocol was then scored in terms of these 30 propositions: A point was given for each of the propositions that was mentioned, and totals for each student were tallied (cf. Lee & Ballman, 1987, regarding determining idea units and scoring, and Deville & Chalhoub-Deville, 1993, who found no difference in scores whether recalls are scored dichotomously or are weighted). The spatial ability and verbal ability tests were scored for the number of correct answers given.

3. Results

Hypothesis 1. Low-ability learners will recall fewer translations of German words than high-ability learners when they are required to select and process visual annotations while reading.

Using the SPSS procedure GLM (type one sum of squares), an analysis of covariance (ANCOVA) was computed with the number of correct answers on the vocabulary test (German vocabulary post-test scores, treated as residual gain scores with German vocabulary pre-test scores partialed out) as dependent measure, types of treatment (verbal annotations: present vs. absent; visual annotations: present vs. absent) as between-subjects factors, and verbal ability and spatial ability as covariates. As the hypotheses are dealing with aptitude-by-treatment-interactions (ATI), ATI-effects were specified as interactions of ability (standard *z*-scores) and treatment variables (see Leutner & Rammsayer, 1995).

The analysis showed a significant two-way interaction (ATI) of verbal ability and the availability of visual annotations, $F(1, 140)=4.46$, $P=0.036$, $\eta^2=0.031$, as well as a significant three-way interaction (ATI) of spatial ability, the availability of visual annotations, and the availability of verbal annotations, $F(1, 140)=12.76$, $P<0.001$, $\eta^2=0.084$. In addition, there was a statistically significant (positive) main effect of verbal ability, $F(1, 140)=7.30$, $P=0.008$, $\eta^2=0.050$, but not for spatial ability, $F(1, 140)<1$. There were also statistically significant main effects of the availability of verbal annotations, $F(1, 140)=10.92$, $P=0.001$, $\eta^2=0.072$, and the availability of visual annotations, $F(1, 140)=20.80$, $P<0.001$, $\eta^2=0.129$, as well as a significant two-way interaction of the availability of verbal and visual annotations, $F(1, 140)=5.59$, $P=0.036$, $\eta^2=0.038$; $MSE=13.40$ for all significance tests.

The overall pattern of combined treatment effects (main effects and the interaction of the availability of visual and/or verbal annotations) represents a multimedia effect (Mayer, 2001): Students in the group that did not receive any annotations performed worst ($M=20.5$, $SD=3.7$), those who received only one type of annotation performed moderately well ($M=24.1$, $SD=3.4$ for visual annotations; $M=24.6$, SD = 3.3 for verbal annotations), and learners who received both types of annotations performed best (M = 25.4, SD = 4.9). This multimedia effect accounts for 23.9% of the variance of the German post-test vocabulary knowledge. While not all differences between means were statistically significant, the rank order of means exactly replicates that found in previous research by Plass et al. (1998). Note that the two ability variables were found to be statistically nearly independent: the Pearson correlation between the scores for verbal and spatial ability was $r = 0.183$.

3.1. Verbal ability

The statistically significant two-way interaction of verbal ability and the availability of visual annotations indicates that low-ability learners recalled fewer translations of German words than high-ability learners when they were required to select and process visual annotations while reading: When visual annotations were not available, the within-group Pearson correlations between verbal ability and German vocabulary post-test knowledge were −0.07 (no verbal, no visual annotations group) and 0.15 (verbal annotations only group), whereas when visual annotations were available, the correlations were 0.22 (visual annotations only group) and 0.35 (verbal and visual annotations group). In other words, the effects of verbal ability on vocabulary learning were stronger when visual annotations were available than when they were not available.

For the sole purpose of depicting the results graphically (see Fig. 3), 33% extreme groups were created based on verbal ability. Computing planned comparisons (ANOVA with linear contrasts) between high-verbal ability learners and low-verbal ability learners for each of the four treatment groups revealed that low-verbal ability learners performed significantly worse on the vocabulary test than high-verbal ability learners when only visual information was available, $t(84) = 2.00$, $P < 0.05$, $d = 0.97$, and when both visual and verbal information were available, $t(84) = 2.32$,

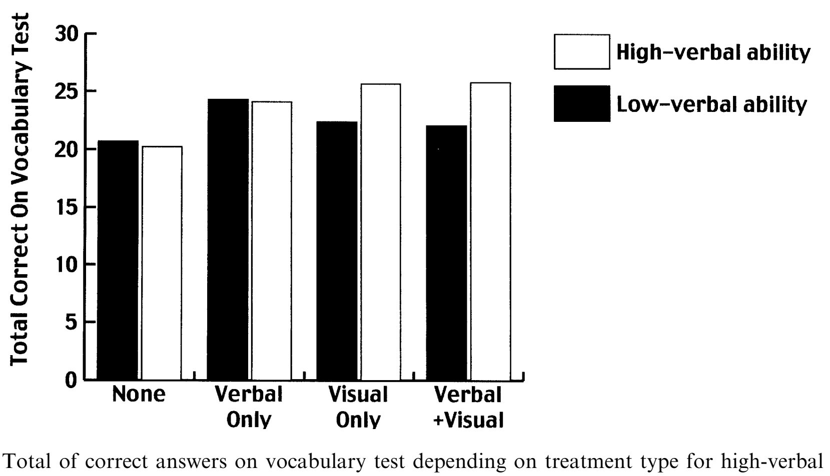

Fig. 3. Total of correct answers on vocabulary test depending on treatment type for high-verbal ability learners and low-verbal ability learners.

$P<0.05$, $d=0.74$. For the visual-only treatment, the descriptive statistics are $M=22.4$ (SD = 4.2) for low-verbal ability and $M=25.7$ (SD = 2.7) for high-verbal ability learners; for the visual and verbal treatment, the descriptives are $M=22.1$ (SD = 5.1) for low-verbal ability and $M=25.8$ (SD = 4.3) for high-verbal ability learners. There was no difference between these groups when no information or only verbal information was available. As to be expected, the results of this additional analysis correspond perfectly to the results of the ANCOVA and correlation analyses reported above.

Thus, as predicted by the generative theory of multimedia learning and cognitive load theory, low-verbal ability learners performed worse than high-verbal ability learners when visual information was available and had to be looked up and processed. Low-verbal ability learners do not have sufficient cognitive resources to process the visual information and build referential connections and therefore experience deleterious effects on vocabulary learning.

3.2. Spatial ability

Similar to low-verbal ability learners, it was expected that low-spatial ability learners might have insufficient cognitive resources to build connections between the visual and verbal representations of the information they select and hold in working memory. When the treatment condition forces these learners to look up all available visual annotations for vocabulary words, their low-spatial ability might result in insufficient cognitive resources to build these connections, and consequently the likelihood for these learners to have two kinds of retrieval routes when presented with the target word in a vocabulary test is expected to be lower than for high-spatial ability learners.

The statistically significant three-way interaction of spatial ability and the availability of visual and verbal annotations provided empirical evidence for this hypothesis: when visual annotations were not available, the within-group Pearson correlations between spatial ability and vocabulary post-test knowledge were −0.27 (group with no annotations) and 0.21 (verbal annotations only group), whereas when visual annotations were available, the correlations were 0.37 (visual annotations only group) and −0.25 (verbal and visual annotations group). In other words, the effect of spatial ability on vocabulary learning was negative when no annotations or both visual and verbal annotations were available and positive when only verbal or only visual annotations were available.

For the sole purpose of depicting the results graphically (see Fig. 4), again 33% extreme groups were created based on spatial ability. Computing planned comparisons (ANOVA with linear contrasts) between high-spatial ability learners and low-spatial ability learners for each of the four treatment groups revealed that low-spatial ability learners ($M=22.9$, SD = 5.3) performed significantly worse on the vocabulary test than high-spatial ability learners ($M=25.6$, SD = 2.9) when only visual information was available, $t(84)=1.65$, $P=0.051$, $d=0.67$. There was no difference between these groups when only verbal information was available. However, high-spatial ability learners performed worse on the vocabulary test than low-spatial

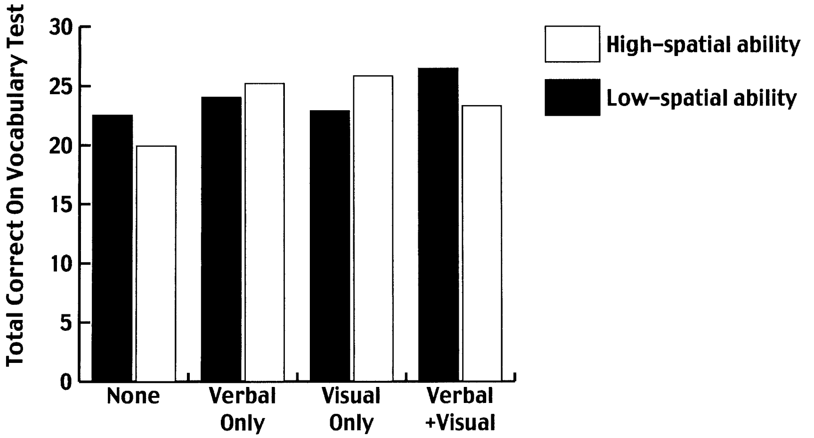

Fig. 4. Total of correct answers on vocabulary test depending on treatment type for high-spatial ability learners and low-spatial ability learners.

ability learners when no information was available, $t(84) = -1.78$, $P < 0.05$, $d = 0.84$, and when both visual and verbal information were available, $t(84) = -1.80$, $P < 0.05$, $d = .60$. For the treatment without annotations, the descriptives are $M = 22.5$ (SD = 3.0) for low-spatial ability learners and $M = 19.9$ (SD = 3.2) for high-spatial ability learners; for the visual and verbal treatment, the descriptives are $M = 26.4$ (SD = 4.3) for low-spatial ability learners and $M = 23.8$ (SD = 5.1) for high-spatial ability learners. Again, and as would be expected, the results of this additional analysis correspond perfectly to the results of the ANCOVA and correlation analyses reported above.

As predicted, low-spatial ability learners performed worse than high-spatial ability learners when visual information was available and had to be looked up and processed. No such difference was found for verbal annotations, which is not surprising, as their processing is not expected to be affected by spatial ability. A finding of theoretical interest is that high-spatial ability learners performed worse than low-spatial ability learners when no annotations were available and when both verbal and visual annotations were available. An interpretation for these findings will be provided in the discussion section.

Hypothesis 2. The number of propositions learners recall from the text will be lower in treatments where they are required to process visual annotations as compared to treatments with verbal annotations or no annotations.

An analysis of covariance (ANCOVA) was computed with the number of propositions reported in the comprehension test (text comprehension post-test scores, treated as residual gain scores with German vocabulary pre-test scores partialed out) as dependent measure, types of treatment (verbal annotations: present vs. absent; visual annotations: present vs. absent) as between-subjects factors, and verbal ability and spatial ability (standard z-scores) as covariates. Aptitude-treatment-interaction effects (ATI) were specified as interactions of ability and treatment variables.

Statistically significant (positive) main effects were found for spatial ability, $F(1, 137) = 7.14$, $P = 0.008$, $\eta^2 = 0.050$, and for verbal ability, $F(1, 137) = 7.40$, $P = 0.007$, $\eta^2 = 0.051$. In addition, there was a statistically significant two-way interaction of the availability of verbal and visual annotations, $F(1, 137) = 5.72$, $P = 0.018$, $\eta^2 = 0.040$, but no main effects for visual annotations, $F(1, 137) = 1.45$, $P = 0.230$, or for verbal annotations, $F(1, 137) < 1$. There were no significant interactions (ATI) of verbal or spatial ability and treatment conditions, $F(1, 137) < 1$ for spatial ability×visual annotation; $F(1, 137) < 1$ for verbal ability×visual annotation; $F(1, 137) < 1$ for spatial ability×verbal annotation; $F(1, 137) < 1$ for verbal ability×verbal annotation; $F(1, 137) = 1.16$, $P = 0.284$ for spatial ability×visual annotation×verbal annotation; $F(1, 137) = 3.55$, $P = 0.062$ for verbal ability×visual annotation×verbal annotation. For all significance tests, MSE = 15.83.

The pattern of the obtained disordinal two-way interaction of the availability of verbal and visual annotations was consistent with our hypothesis. In order to test the hypothesis more specifically concerning the difference between visual and verbal annotations, planned comparisons were computed (ANOVA with linear contrasts) between the comprehension scores for the four treatment groups; see Fig. 5. These planned comparisons revealed that, compared to no annotations ($M = 9.3$, SD = 3.3), the processing of visual annotations ($M = 7.2$, SD = 4.3) led to a significantly lower comprehension score, $t(145) = 2.17$, $P = 0.031$, $d = 0.55$, while no such significant difference was found in comparison to the processing of verbal annotations ($M = 7.9$, SD = 4.1), $t(145) = 1.51$, $P = 0.134$, and to the processing of verbal and visual annotations ($M = 9.3$, SD = 4.7), $t(145) = 0.01$, $P = 0.994$. The difference between the results for verbal annotations and visual and verbal annotations did not reach statistical significance, $t(145) = 1.50$, $P = 0.136$, whereas the difference between the results for visual annotations and visual and verbal annotations reached statistical significance, $t(145) = 2.17$, $P = 0.032$, $d = 0.47$.

Learners, when forced to look up annotations designed to support the lexical level of text processing, are hindered in reading comprehension, especially when only

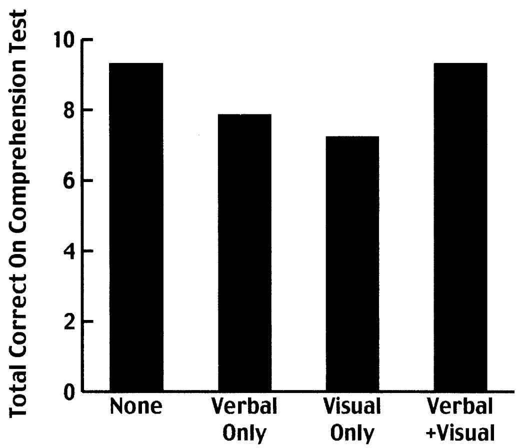

Fig. 5. Total of correct answers on comprehension test depending on treatment type.

visual annotations are available. This effect is not present when, in addition to the visual annotations, verbal annotations are available, and it is not moderated by either verbal or spatial ability.

4. Discussion

The basic theme of this study was that multimedia information, which is generally expected to lead to more effective learning, can potentially have deleterious effects depending on the learning conditions and the individual differences of the learners. When low-ability learners select and process both verbal and visual information, the resulting high cognitive load may lead to insufficient cognitive resources for the construction of referential connections between the verbal and visual information, which is, according to Mayer's (1997, 2001) generative theory of multimedia instruction and cognitive load theory, expected to result in decreased learning. In order to gather empirical evidence for this idea, students in this study were assigned to treatments that required them to select and process certain information. Indeed, instructional materials, such as web pages, often force the learner to process all available visual and verbal information by presenting this information on the screen instead of allowing the learner to request the information that is relevant to him or her.

4.1. Vocabulary acquisition

First of all, the results of this study are in line with the generative theory of multimedia learning in conjunction with cognitive load theory. Replicating previous research, we found a general multimedia effect for vocabulary acquisition: participants learned more vocabulary words when both visual and verbal annotations were present than when only one type of annotation or no annotations were present, a result that is in line with the dual coding assumption of dual coding theory and the generative theory of multimedia learning (Mayer, 1997, 2001; Paivio, 1971, 1990; Plass et al., 1998). In addition, we found that verbal ability had a positive effect on vocabulary learning but spatial ability did not, which is consistent with findings by Atkinson and Baddeley (1998) and Papagno and Vallar (1995).

The main focus of this study, however, was on the differential effects of specific instructional treatments for high- and low-ability learners. We found that when low-ability students processed verbal information from a reading text and had to process additional visual annotations for vocabulary words in the text, they learned fewer vocabulary words than high-ability students. In contrast, no such difference was found between high- and low-ability learners when they only had to process additional verbal annotations for vocabulary words in the text. This effect, which was found both for low-verbal and for low-spatial ability students, may be caused by the different amount of cognitive load imposed by verbal and visual annotations. While visual annotations have to be processed and their meaning interpreted, verbal annotations provide the clear and unambiguous meaning of the annotated words. For example, learners need to decide if the picture of a school of fish refers to a school of fish, a fish, water, the sea, sea-life or to a similar meaning of the vocabulary word. This, in effect, can be described as the translation of an ambiguous pictorial representation into an unambiguous verbal representation. The verbal statement "school of fish," on the other hand, does not require such a decision. The high cognitive load imposed by the visual annotations makes the construction of referential connections between the verbal and the visual information less likely, which results in fewer retrieval cues for the vocabulary and, hence, fewer vocabulary words that could be retrieved during the vocabulary post-test.

4.2. Text comprehension

A similar effect of high cognitive load was found for overall text comprehension. The presence or absence of the requirement to select and process visual or verbal annotations interacted in a disordinal way, where comprehension was significantly lower when visual annotations had to be processed than when no annotations, verbal annotations or verbal and visual annotations had to be processed. In other words, when only visual annotations had to be processed, learners remembered fewer propositions of the text than for any of the other conditions. Here, the high cognitive load imposed by the visual annotations had a negative effect on comprehension for all learners, independent of their abilities. Since the annotations for vocabulary words were designed, following Mayer's (1997, 2001) generative theory of multimedia learning, to support the process of selecting relevant information, i.e. understanding vocabulary words, learners had to invest cognitive resources in this low-level processing that therefore were unavailable for the construction of a mental model of the text and the integration of visual and verbal information with one another and with the mental model. When no annotations were available, this additional processing of vocabulary-level information was not necessary, and all cognitive resources could be allocated to the comprehension of the text. When verbal and visual annotations were available, the high cognitive load required to process the visual annotations was reduced by the availability of verbal annotations, which translate and explain the meaning of the word. A visual annotation is then an illustration of a verbal explanation, and not the sole source of the meaning of the word. In other words, learners receiving both visual and verbal annotations do not have to invest the higher amount of cognitive resources to process the visual annotations, as it was shown to be the case for learners receiving only visual annotations, but can obtain the meaning of the word from processing the verbal annotations, which requires less cognitive resources. In that sense, the amount of cognitive load imposed by visuals is reduced by the availability of verbal information, which results in a high level of comprehension. This interpretation of the comprehension test scores is in line with our hypothesis that the high cognitive load requirements of processing visual annotations would hinder the comprehension of the text and replicates results from previous research by Yeung et al. (1998). In their study, Yeung et al. found that explanatory notes designed to facilitate vocabulary learning resulted, under certain conditions, in lower comprehension of the overall text. In our study, this effect was only found for the treatment where only visual annotations were available, which by design require more cognitive resources to process than verbal annotations. The results are, therefore, in line with a generative theory of multimedia learning (Mayer, 1997, 2001) in conjunction with cognitive load theory (Sweller, 1994, 1999).

4.3. Effects of individual differences

Some additional findings of this study warrant further discussion because of their potential theoretical implications. First, when high-verbal ability learners and low-verbal ability learners did not receive any annotations during reading, their performance in vocabulary acquisition, which was not the primary learning task and was therefore incidental, did not differ. In the explicit learning task, however, which is comprehension, the main effect of verbal ability shows that high-verbal ability learners performed better than low-verbal ability learners. We argued earlier that this learning condition, where no annotations are available, along with the condition where only verbal annotations are available, does not impose the same high cognitive load that is caused by visual annotations. The finding that high- and low-verbal ability learners did not differ in their vocabulary learning when no visual annotations were present, but indeed differed when visual annotations were present, empirically supports this assumption: in the high cognitive load condition, which exist in treatments with visual annotations, low-verbal ability learners did not have sufficient cognitive resources available to process the text and the visual annotations, while high-spatial ability learners had sufficient resources. Under low cognitive load conditions, which existed in treatments with no or with only verbal annotations, both low- and high-verbal ability learners were able to process the text and the verbal annotations. Similar interaction effects for different levels of cognitive load were found by Yeung et al. (1998) for high- and low-ability ESL readers.

A second finding that can provide further insights into the relationship of spatial ability and cognitive load is the performance of low-spatial ability learners in the acquisition of vocabulary when they received no annotations and when they received both verbal and visual annotations. Under both conditions, low-spatial ability learners remembered more vocabulary words than high-spatial ability learners. However, the main effect of spatial ability on text comprehension shows that high-spatial ability learners performed better than low-spatial ability learners in the comprehension of the reading text, the explicit learning task. In other words, high-spatial ability learners were able to focus on the comprehension of the text and to infer the meaning of unknown vocabulary words without having an explicit knowledge of its translation (Kintsch, 1998). Low-spatial ability learners, in contrast, spent more cognitive resources on the lower-level processing of individual vocabulary words and their meaning and did not have enough resources available for organizing the propositions of the text in working memory and integrating them with one another and with the mental model of the reading text. In the treatments where either verbal or visual annotations were present, learners were required to look up these annotations even if they knew the word or could infer its meaning from the context. Learners therefore had to spend more cognitive resources on the processing of vocabulary, which can be expected to result in lower comprehension. In our study, this effect was statistically significant for the treatment with visual annotations. However, when both visual and verbal annotations had to be looked up, high-spatial ability learners were again able to focus on the main task of comprehending the text and to spend less cognitive resources on the lower-level processing of vocabulary words, while low-spatial ability learners spent more cognitive resources on the lower-level processing of vocabulary words and less on the comprehension of the text. This effect of the condition where both types of annotations had to be looked up can also be interpreted based on the concept of *redundancy*, as it is used by Yeung et al. (1998) and Kalyuga, Chandler, and Sweller (1999, 2000). This line of research showed that processing information from two sources that contain the same information in different representation modes can induce cognitive load in high-ability learners for whom information from one source would have been sufficient. This cognitive load imposed by the requirement to process redundant information is detrimental to learning for these high-ability learners. Low-ability learners, on the other hand, may benefit from the integration of information from two different sources. A similar effect was found for high-spatial and low-spatial ability students in the present study when they had to process verbal as well as visual annotations. This finding may inspire new insights into the function of visual and verbal working memory and in the allocation of resources to the processing of vocabulary information versus text comprehension in second-language multimedia learning environments and warrants further investigation.

4.4. Implications

The present study offers educational and theoretical implications. On the educational side, it provides implications relevant for the design of multimedia instruction and for second language instruction. In particular, our results suggest that learners should have options for using study material in both a visual mode and a verbal mode, but should not be forced to select and process both types of information, as would be the case with a web page or multimedia software that by default displays all available information. In this present study, students learned fewer vocabulary words when they had to process information that led to a high load on their cognitive resources. In other words, providing both options of visual and verbal annotations is only effective in addressing individual differences when students can choose which information they would like to select and process. This could be implemented in practice by providing features that let the student request information instead of presenting it by default to all learners. This is in line with earlier findings on learning preferences (Plass et al., 1998).

On the theoretical side, this study provides evidence for a generative theory of multimedia learning that makes a distinction between visual and verbal working memory in vocabulary learning and text comprehension. In addition, this study provides evidence concerning the moderating effect of individual differences, particularly verbal ability and spatial ability, on the learners' cognitive load in multimedia learning.

When do multiple representations of information in second-language learning help and when do they hinder learning? The findings of this study show that multiple representations of information do not always help learning. Indeed, they may hinder learning in low-ability students when they experience high cognitive load as it is imposed by the requirement to process visual information.

Acknowledgements

CyberBuch was developed by Dorothy M. Chun and Jan L. Plass with support of several Instructional Improvement Grants from the Office of Instructional Consultation at the University of California, Santa Barbara. We would like to thank Karl Otto at University of Pennsylvania and Katrin Schroeter at University of New Mexico for their support in conducting the present study. More information about the computer materials can be accessed on the world wide web at the following address: http://www.gss.ucsb.edu/faculty/dmchun/cyberbuch/

Appendix A. Thirty-five annotated words

ärmlich (poor, poorly)
auftauen (to thaw)
Begeisterung (enthusiasm)
besorgt (anxious)
betrübt (sad)
dösen (to doze)
drohen (to threaten)
Fang (catch)
feindselig (hostile)
Feuerzeug (lighter)
Fischermütze (fisherman's cap)
Fischschwärme (schools of fish)
gereizt (irritated)
Gesichtsausdruck (facial expression)
Hubschrauber (helicopter)
Hummer (lobster)
Kopfnicken (to nod one's head)
Kopfschütteln (shaking of one's head)
Körbe (baskets)
Kutter (cutter)
meßbar (measurable)
Münder (mouths)
munter (lively, cheerfully)
nachdenklich (pensive)
See (ocean)
sich aufrichten (to sit up)
sich recken (to stretch)
sich verschlucken (to choke)
sprengen (to explode)
Spur (trace)
unterrichten (instruct)
verleihen (to give, grant)
verpaßt (missed)
Wellenkämme (whitecaps)
Zeichensprache (sign language)

Appendix B. Major propositions

Anekdote zur Senkung der Arbeitsmoral, by Heinrich Böll

- Setting: west coast of Europe, marina/harbor
- Setting: green ocean, wave crests, fisherman's hat

- Tourist approaches (nicely dressed, etc.)
- Fisherman, dozing, poorly dressed
- Tourist: taking photos
- Four clicks: camera and lighter
- Tourist wakes fisherman
- Embarrassed situation due to tourist's intruding
- Tourist tries to engage fisherman in conversation
- Tourist asks "Why aren't you fishing?/Why won't you go out again?"
- Fisherman communicates nonverbally at first
- Fisherman finally responds to tourist's questions
- Tourist asks about the catch, the weather
- Fisherman describes the catch
- Tourist asks about fisherman's health
- Fisherman: I'm doing great!
- Tourist is sad, nervous
- Cigarettes are smoked
- Tourist says, "If you would go out more..."
- Tourist describes possible objects to purchase
- Tourist describes possible abstract things to get
- Tourist is excited about his ideas
- Fisherman is reluctant
- Tourist feels sad
- Fisherman is concerned about tourist
- Fisherman asks: Why work more?
- Tourist: You could doze in the sun
- Fisherman: That's what I'm doing already
- Tourist leaves without pity, but envy
- The moral of the story is...

References

Atkinson, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, *19*, 537–552.

Avons, S. E., Wragg, C. A., Cupples, L., & Lovegrove, W. J. (1998). Measures of phonological short-term memory and their relationship to vocabulary development. *Applied Psycholinguistics*, *19*, 583–601.

Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press.

Baddeley, A. D. (1992). Working Memory. *Science*, *255*, 556–559.

Böll, H. (1986). Anecdote concerning the lowering of productivity. In H. Böll (Ed.), *The stories of Heinrich Böll* (pp. 628–630). New York: Alfred A. Knopf.

Bower, B. (1997). The power of limited thinking. *Science News, Nov*. *22*, *152*, 334–335.

Carroll, J. B. (1993). *Human cognitive abilities*. New York: Cambridge University Press.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*, 293–332.

Chun, D. M., & Plass, J. L. (1997). *CyberBuch*. New York: St. Martin's Press [Computer Program, Macintosh].

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*, 149–210.
Deville, C., & Chalhoub-Deville, M. (1993). Modified scoring, traditional item analysis and Sato's caution index used to investigate the reading recall protocol. *Language Testing Journal*, *10*, 117–132.
Educational Testing Service. (1976). *Manual for Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.
Ernst, C. H. (1991). Ability differences in prose learning. *Intelligence*, *15*, 455–477.
Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: a critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*, 259–272.
Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. *Applied Cognitive Psychology*, *13*, 65–77.
Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning, and instruction*. Hillsdale, NJ: Lawrence Erlbaum.
Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*, 351–371.
Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, *92*, 126–136.
Kintsch, W. (1998). *Comprehension: a paradigm for cognition*. New York: Cambridge University Press.
Kirby, J. R. (1993). Collaborative and competitive effects of verbal and spatial processes. *Learning and Instruction*, *3*, 201–214.
Lee, J. F., & Ballman, T. L. (1987). FL learners' ability to recall and rate the important ideas of an expository text. In B. VanPatten, T. R. Dvorak, & J. F. Lee (Eds.), *Foreign language learning: a research perspective* (pp. 108–118). Cambridge: Newbury House Publ.
Leutner, D., & Plass, J. L. (1998). Measuring learning styles with questionnaires versus direct observation of preferential choice behavior in authentic learning situations: the visualizer/verbalizer behavior observation scale (VV-BOS). *Computers in Human Behavior*, *14*, 543–557.
Leutner, D., & Rammsayer, T. (1995). Complex trait-treatment-interaction analysis: a powerful approach for analyzing individual differences in experimental designs. *Personality and Individual Differences*, *19*, 493–511.
Masoura, E. V., & Gathercole, S. E. (1999). Phonological short-term memory and foreign language learning. *International Journal of Psychology*, *34*, 383–388.
Mayer, R. E. (1984). Aids to prose comprehension. *Educational Psychologist*, *19*, 30–42.
Mayer, R. E. (1997). Multimedia learning: are we asking the right questions? *Educational Psychologist*, *32*, 1–19.
Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*, 444–452.
Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*, 312–320.
Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *86*, 389–401.
Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: the role of modality and contiguity. *Journal of Educational Psychology*, *91*, 358–368.
Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: the case for minimizing irrelevant sounds in the design of multimedia messages. *Journal of Educational Psychology*, *92*, 117–125.
Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.
Paivio, A. (1971). *Imagery and verbal processes*. New York, NY: Holt, Rinehart & Winston.
Paivio, A. (1990). *Mental representation: a dual-coding approach*. Oxford, UK: Oxford University Press.
Paivio, A., & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology*, *34*, 388–399.

Paivio, A., & Lambert, W. (1981). Dual coding and bilingual memory. *Journal of Verbal Learning and Verbal Behavior*, *20*, 532–539.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *48A*, 98–107.

Pellegrino, J. W., Alderton, D. L., & Shute, V. J. (1984). Understanding spatial ability. *Educational Psychologist*, *19*, 239–253.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second language multimedia learning environment. *Journal of Educational Psychology*, *90*, 25–36.

Richardson, A. (1983). Imager: definition and types. In A. A. Sheikh (Ed.), *Imagery: current theory, research, and application* (pp. 3–42). New York: Wiley.

Richardson, A. (1994). *Individual differences in imaging: their measurement, origin, and consequences*. Amityville, NY: Baywood.

Riverside Publishing. (1981). *The Nelson Denny Reading Test*. Riverside, CA: Riverside Publishing.

Smith, P. L., & Ragan, T. J. (1999). *Instructional design* (2nd ed.). Upper Saddle River, NJ: Merrill.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295–312.

Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, *119*, 176–192.

Wittrock, M. C. (1974). Learning as a generative activity. *Educational Psychologist*, *11*, 87–95.

Wittrock, M. C. (1990). Generative processes of comprehension. *Educational Psychologist*, *24*, 345–376.

Yeung, A. S., Jin, P., & Sweller, J. (1998). Cognitive load and learner expertise: split-attention and redundancy effects in reading with explanatory notes. *Contemporary Educational Psychology*, *23*, 1–21.