The Modality Principle in Multimedia Learning

An Open Question: When Speech Fails to Foster Learning?*

Klaus D. Stiller

University of Regensburg
D-93040 Regensburg
klaus.stiller@psychologie.uni-regensburg.de

Abstract
Learning is fostered when spoken rather than written texts are used concurrently with dynamic or static pictures in instructional settings. This is mainly caused by split attention and can be explained by the cognitive theory of multimedia learning (Mayer, 2001) or the cognitive load theory (Sweller, 1999). Numerous studies have provided evidence for modality effects (Ginns, 2005), but recently, studies also revealed a vanishing or reversal of modality effects in favor of visual texts. This can be attributed to learner characteristics or instructional design, but knowledge about relevant features is still scarce.

1 Multimedia learning

In general, learning is defined as a relatively sustainable change of a subject's behavior or potential for a behavior related to a certain situation, which is created by repeated experiences of subjects in particular situations. Changes can occur as regards cognitive, psychomotor, affective-motivational or social aspects (Lukesch, 1998). On the other hand, teaching refers to an intended and planned initiation and fostering of learning in order to achieve teaching goals. Mostly, it is focused on the cognitive goals of learning and instruction.

Learning and instruction can be realized by using media. The most prominent modes for presenting information are written and spoken texts as well as static and dynamic pictures. Most medial presentations involve at least one of the verbal and

* Published in: Osswald, Achim; Stempfhuber, Maximilian; Wolff, Christian (eds.) (2007). Open Innovation. Proc. 10. International Symposium for Information Science. Constance: UVK, 129-144.

one of the pictorial modes. In Anglo-American literature, learning concurrently with texts and pictures is known as multimedia learning (Mayer, 2001). That is important to note, because in most German literature the word *multimedia* is linked to computer programs, indicating that these programs feature at least a minimum of interactivity and a presentation of information using verbal and pictorial presentation codes, partly in a dynamic way, for the auditory and visual senses (Issing & Klimsa, 2002).

One of the prominent theories which describe and explain meaningful learning with texts and pictures is the cognitive theory of multimedia learning (CTML; Mayer, 2001). It is based on the dual channel assumption, the limited capacity assumption and the active learning assumption. The dual channel assumption proposes that humans process pictorial and verbal information in two separate systems, a visual/pictorial and an auditory/verbal system (Fig. 1). The systems are functionally independent of each other, but may interact. The processing of pictures occurs mainly in the visual/pictorial channel and that of spoken words in the auditory/verbal channel. Written words are initially processed in the visual/pictorial channel before they move to the auditory/verbal channel in the working memory. Furthermore, it is possible for verbal and pictorial information to cross the channel barriers. Verbal information that can be (easily) imagined or pictorial information that can be (easily) verbalized may find access to the other channel.

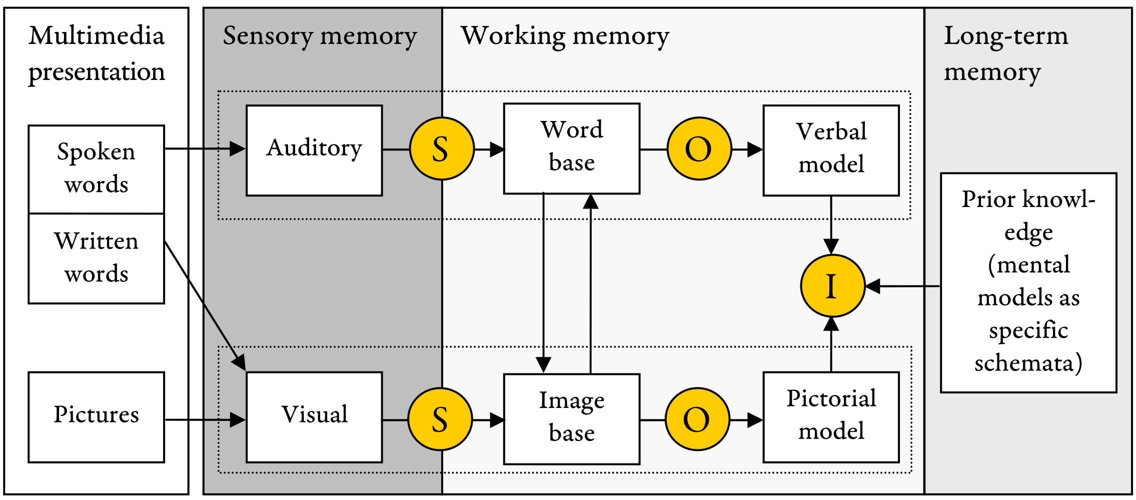

Figure 1: Framework for the CTML (Mayer, 2001). S stands for selecting information, O for organizing information into pictorial and verbal models, and I for integrating mental models and prior knowledge.

The limited capacity assumption suggests that the information processing capacity in each of the two channels is limited. Within capacity limits the learner has to accomplish all the information processing necessary for (meaningful) learning. Especially the cognitive load theory (CLT) emphasizes the role of the working memory in learning, and the problem of occupying working memory by useful und unnecessary processes (Sweller, 1999). A working memory with limited capacity only allows processing of 2 to 4 elements (an estimation by Sweller, 1999), which is less than the magical number 7 plus/minus 2 according to Miller (1956). This is due to the fact that processing information requires more load than merely retaining items in working memory (Sweller, 1999). Additionally, working memory is also limited in time. If rehearsal of memory items is not possible, items fade within seconds. Due to these severe limitations of our working memory, learning is often difficult, because tasks are complex. Learners have to retain a large number of elements in their minds and relate them or search for actions that bring them nearer to solving a problem. This often exceeds the working memory capacity. Working memory resources are divided into various parts that are covered by special kinds of cognitive processes which in turn exert loads on working memory. These loads can be detrimental to or beneficial for learning. The components of cognitive load are extraneous, germane and intrinsic load, which are intrinsically linked to the source of the load. Intrinsic load is created by task complexity and is mainly needed for understanding. Germane load goes beyond understanding and is bound to the learning processes in a real sense – processing information in order to build mental representations and storing these representations. Extraneous load results from the presentation manner of the material and is often the main source that is detrimental to learning. As long as extraneous, intrinsic and germane load do not occupy the entire working memory capacity, some memory capacity that could be used for processing is still available. Mayer and Moreno (2002, p. 108) regard the CLT as "an important component in any theory of learning that is intended to guide the design of multimedia learning environments". Accordingly, they regard the CLT as an important component of the CTML.

The active learning assumption states that learners have to process information actively in order to achieve meaningful learning. Active cognitive processing includes the selection of relevant information from words and pictures, the mental organization of this information to form coherent verbal and pictorial representations, and the mental integration of these representations with one another and with prior knowledge from long-term memory. Selection and organization take place in separate channels (see the dual channel assumption) with integrative processing bridging the channels. The SOI processes "generally do not occur in a rigid linear order, but rather in an iterative fashion" (Mayer, 2001, p. 130).

To a great extent, learning depends on presentational features, which determine the amount of extraneous load. As extraneous load is detrimental to learning, designers should strive towards minimizing it. Modality effect is one of the effects connected with and caused by extraneous load. The modality principle in instructional design is directly linked to modality effect and the associated empirical evidence, which are dealt with in the next section.

2 The modality principle

The modality principle recommends using spoken rather than written texts in conjunction with pictures, because learners learn more effectively in computer-based and book-based environments when spoken rather than visual texts are used concurrently with dynamic or static pictures. The superiority of spoken over written texts was demonstrated in multiple studies, largely using low prior knowledge learners (Ginns, 2005). Modality effects were observed in various measures: (1) less mental effort (a measure for cognitive load) with spoken words while learning (e. g. Tabbers, 2002; Tindall-Ford, Chandler, & Sweller, 1997), (2) less time on problem-solving tasks (e. g. Jeung, Chandler, & Sweller, 1997; Mousavi, Low, & Sweller, 1995), (3) higher scores on various tasks of retention, matching verbal and pictorial information, transfer and practice (e. g. Kalyuga, Chandler, & Sweller, 1999, 2000; Moreno & Mayer, 1999), and (4) shorter reaction time in a secondary task (e. g. Brünken, Plass, & Leutner, 2004; Brünken, Steinbacher, Plass, & Leutner, 2002).

Modality effect is thought to be composed of two parts: visual/pictorial processing channel overload and split-attention effect. Firstly, when words and pictures are presented visually, the words are processed at least initially in the visual/pictorial channel, allowing this channel to become overloaded, the auditory/verbal channel remaining unused. Thus, written texts exert extra load on working memory, because they initially use the visual processing channel before the information is transmitted to the verbal processing channel. On the other hand, auditory texts enter the auditory/verbal processing channel directly, thus releasing the visual/pictorial channel from unnecessary load. The working memory resources gained might be allocated to germane load. Overall, the interchannel transition of visual text information is often believed to be an automatic process and that it therefore does not occupy many resources (e. g. Tabbers, 2002; Schmidt-Weigand, 2006).

Secondly, a split-attention effect occurs, because attention has to be spread between visual text and pictures, which results in less relevant material being selected for further processing (especially with system-paced instructions including dynamic presentations). The split-attention effect always occurs when learners have to divide their attention between different sources of information that must be integrated mentally to achieve comprehension (Sweller, 1999). A central characteristic of the separated sources of information is that all sources are needed in order to reach full understanding. The multiple sources of information are related to one another and are unintelligible in isolation (Sweller, 1999). The sources of information can be separated in time or space. If learners have to split their attention between such sources of information, extraneous load on working memory will increase. This means, for example, that learners must search for corresponding information from the disparate sources and keep it concurrently in working memory in order to enable integration processes. This exerts a heavy extraneous load on working memory. If the increased extraneous load results in working memory being overloaded, this in turn hinders learning. An auditory presentation of text reduces these unnecessary processes; no splitting of visual attention between pictures and visual text is necessary in order to select relevant information and subsequently to organize and integrate it in working memory. A simultaneous audio-visual presentation increases the likelihood of learners being able to retain corresponding words and pictures in working memory at the same time. Thus it is also more likely that the learners engage in active meaningful learning, in particular in integrative processing of verbal and pictorial information.

Optimal temporal and spatial contiguity is achieved with spoken words. Verbal and pictorial information can be selected in parallel and integrated immediately. Spoken texts release the learner from extraneous load e. g. from extended visual search and accompanying processes as well as from actively retaining information in working memory longer than necessary.

2.1 Overview of studies confirming modality effects on performance

According to Tabbers (2002), older studies which support modality effect are limited concerning subject matter, length of study time, settings and pacing of presentation. The implemented instructions contained contents taken from the exact sciences such as geometry, mathematics, physics and electrical sciences, and they mostly did not last longer than 5 minutes. They were conducted in laboratory settings and the pacing of presentations was determined by the computer system or the procedure in non-computer settings. Since then, however, empirical evidence of modality effect has also been found with learner-paced or learner-sequenced instructions consisting of up to 90 minutes' learning time in classroom and laboratory settings, sometimes using non-scientific contents (e. g. Brünken, Seufert, & Zander, 2005; Moreno & Mayer, 2002; Mayer, Dow, & Mayer, 2003; Moreno, Mayer, Spires, & Lester, 2001).

Mousavi et al. (1995) confirmed the modality effect on the basis of worked-out examples concerning geometry problems in a book-based environment. Students looked at a printed diagram and listened to or read an explanation. Geometry problems which were similar to the ones used in the instructional phase were solved faster when students had listened to the narration; geometry transfer problems were also solved faster, but the statistical test was not significant. Kalyuga et al. (1999, 2000) found equivalent results as regards transfer performance using diagrams and verbal explanations in computerized environments on the topic of soldering and manufacturing.

Tindall-Ford et al. (1997) have shown with book-based instructions about electric circuits that modality effect only holds for material of high, but not of low complexity. With highly complex material, an overload of working memory was forced back within the limits of capacity by using auditory texts; with material of simple or medium complexity, although the degree of cognitive load changed when using visual or auditory texts, the load stayed within the limits of working memory capacities. Leahy, Chandler, and Sweller (2003) found equivalent effects. They used questions tapping low or high element interactivity knowledge and only revealed a modality effect with the latter.

Modality effect was also evidenced with computerized instructions lasting between 10 and 20 minutes on scientific (human circulatory system) as well as non-scientific contents (the historic city of Florence) that were system-paced in laboratory settings (Brünken & Leutner, 2001; Brünken et al., 2002, 2004) or learner-paced in classroom settings involving pupils and university students (Brünken et al., 2005). Learners using spoken texts showed a faster reaction time in a secondary task, understood texts and pictures better and had more success at solving transfer problems.

Craig, Gholson, and Driscoll (2002), Mayer and Moreno (1998), and Moreno and Mayer (1999) have shown that modality effect holds for retention, matching (learners had to place verbal labels on the corresponding parts of pictures) and transfer performance in computer-based environments which used explanations about lightning or brakes, using animation and narration or on-screen text. The on-screen text was placed below the animation (Mayer & Moreno, 1998; Moreno & Mayer, 1999) or when a pedagogical agent was used in a speech bubble placed near the top left-hand corner of the screen (Craig et al., 2002). Learners with narrated animations could remember or recognize more of the key ideas and were better at matching pictorial and verbal information and using the knowledge gained for solving new problems.

Schmidt-Weigand (2006) conducted four experiments using animation with spoken or visual texts about lightning formation. In his first experiment (Exp. 1 in Ch. 2), he found modality effects for retention and transfer both when visual text was placed below the animation or near to the pictorial part in question. Furthermore, in two experiments (Exp. 2 in Ch. 2 and Exp. 1 in Ch. 3) using animation and static pictures, only weak evidence was found for modality effects on verbal retention, transfer and visual recall (drawing task). Transfer performance was not influenced by modality in either of the two experiments, retention was better with narrated animation and visual recall was mainly better with narrated static pictures in Experiment 2 in Chapter 2. In the fourth experiment (Exp. 2 in Ch. 3), he found modality effects only on visual knowledge; no effects on retention and transfer were observed. Measures of overall load revealed no modality effect throughout all experiments.

Mayer et al. (2003, Exp. 1) used an interactive, self-explanatory environment using a cartoon agent. The instruction showed a picture of an electric motor; the five parts of the electric motor were clickable. When a part was clicked on, it was highlighted and a list of thirteen questions was shown at the top right-hand corner of the screen. After clicking on a question, the answer was presented either as spoken text or written text, which was also shown in the top right-hand corner of the screen. Sometimes, the verbal answer was accompanied by an animation. The information units could be revisited or replayed. Study time took about 20 minutes. Modality effect was shown to hold for transfer in an interactive computer-based environment.

Moreno et al. (2001) as well as Moreno and Mayer (2002) presented an agent-based multimedia game on botany. Learners had to design eight and five plants respectively, whereby each plant should flourish on a different planet characterized by its own climate. An agent introduced task and planet and gave advice concerning the learners' choices of roots, stems and leaves. The agent spoke or showed written messages. The game could be paced by learners; therefore depending on the number of plants, it took between 24 and 28 or between 10 and 16 minutes, respectively, to complete the game. In five experiments, a modality effect of retention and transfer was found.

Using a computerized simulation environment, O'Neil et al. (2000) staged a virtual training episode on an aircraft fuel system. Learners had to grasp the structure of a fuel system and how the system functions. Therefore, learners were placed in a virtual environment via head-mounted display, and the instructional system provided a guided exploration of the fuel system in a virtual world. Trainees could navigate around and through the internal components of the fuel system by using a 3D-mouse and could even conduct a series of tests. Auditory and visual text groups learned for approximately 26 to 28 minutes on average. A modality effect on retention was not found, but was supported as regards matching and transfer performance. The knowledge structure assessed by concept maps tended also to be better with spoken texts.

Stiller (2007, Exp. 1) investigated modality effects and their dependence on prior knowledge with learner-paced and learner-sequenced instructions under fixed study time. Instructions were hierarchically structured hypermedia systems on the physiology of the visual system. Texts were accessed by clicking on parts of pictures. University students learned with visual or auditory texts for 25 minutes. A modality effect on structural knowledge as measured by completing pictures by drawing and a verbal description task were found in favor of auditory texts; retention, transfer and labeling pictures were not affected. Modality effects did not depend on prior knowledge.

Stiller (2007, Exp. 4) investigated modality effects and their dependence on attitude towards computers with hierarchically structured hypermedia systems concerning the physiology of the perceptional systems. Subjects studied three learner-paced and learner-sequenced instructions under learner-controlled study time (86 min. on average), whereby detailed texts were presented visually, auditorily and bimodally. Introductions, summaries and glossary were always presented visually. Concerning performance at verbal retention, labeling of representational pictures and a line diagram, there is evidence for modality effect with negative-attitude learners but not with positive-attitude learners. Additionally, a pure modality effect occurred as regards labeling mnemonic pictures. Spoken text prolonged learning time, this mainly being due to an intensive use of the glossary as regards the auditory group. In a follow-up study (Stiller, 2007, Exp. 5), modality effects and their dependence on prior knowledge were investigated. The same instructions as in Exp. 4 were used, but the glossary was eliminated. University students studied with two instructions for 69 minutes on average, whereby detailed texts were presented visually or auditorily. A modality effect was found for the labeling of mnemonic and representational pictures as well as transfer; verbal retention and labeling of a line diagram were not affected. Concerning transfer performance, low-prior-knowledge learners benefited from auditory texts and high-prior-knowledge learners could learn equally well with visual or auditory texts. Instructional usage was only marginally influenced by text modality and prior knowledge.

Rinck and Glowalla (1996) found no modality effect on performance with longer-lasting and learner-paced instructions (5 lessons each lasting approx. 2 hours), but they found advantages in the speed of answering performance tasks by learners who had been instructed with auditory texts.

2.2 Overview of studies showing contradictory evidence

It seems very easy to generalize the findings to other learning environments like interactive media, but this is more complicated than one would suppose, in view of the empirical evidence. Tabbers (2002) and Stiller (2007) realized a number of experiments in order to investigate modality effect in newly applied situations. But they could only partly support a modality effect; moreover, they revealed some contradictory results. Schmidt-Weigand (2006) also showed that modality effects might even vanish when using the most demanding presentation modes consisting of animation and verbal explanation. Modality effects also proved to be dependent on learner features such as prior knowledge and computer attitude (Stiller, 2007).

Schmidt-Weigand (2006; Ch. 4) showed that the modality effect disappeared when learners were able to control the pacing of a presentation. Schmidt-Weigand (2006) used explanations by animation and narration or visual text that consisted of 16 sequences presented in succession; the learners could decide when to start each sequence. It was predicted that modality effects would vanish with learner-paced instructions. This was confirmed for retention, visual knowledge and transfer; performance differences were insignificant. Mental effort measures also showed no influence of text modality. Study time did not influence performance. Learner-paced presentation durations did not differ when using auditory or visual texts (on average 183 seconds, SD = 42.86).

Tabbers' (2002) aim was to generalize modality effect to longer instructions from a different content domain in an ecologically valid classroom setting. He used web-based multimedia lessons which lasted from 26 to 70 minutes and dealt with a didactical model of instructional design, focusing especially on learner-pacing. The results of the four experiments can be summarized as follows. (1) Modality effect could be widely replicated for mental effort. Spoken words put less load on working memory with both system-paced and learner-paced instructions. (2) Modality effect of retention and transfer could be partly replicated with system-paced instructions. (3) Performance increased and modality effect vanished with more time on task when using system-paced instructions. (4) Modality effect vanished or more often reversed for retention and transfer with learner-paced instructions.

In Experiment 1, static diagrams were explained by auditory or visual texts. Referenced diagram parts were highlighted in red. Cueing of pictorial parts was introduced to minimize visual search and thus extraneous load. Learners controlled the pace of instruction. Information units were put in linear order and learners could move forward and backward along this line. Mental effort tended to be lower with spoken words during instruction, was significantly higher for the visual text group when solving retention tasks and similar for treatment groups as regards solving transfer tasks. Retention and transfer performance was better after students had received the instruction with visual texts. Thus a modality reversal effect for retention and transfer was found, which suggests a visual presentation of texts with learner-paced instructions.

In Experiment 2, static diagrams were accompanied by narrations or on-screen texts, but this time instructions were system-paced. Referenced diagram parts were cued to minimize visual search. Study time was 26.2 minutes. The participants, students from a teacher training college, were considered to have no prior knowledge. Study time, mental effort during instruction and testing as well as retention and transfer performance were assessed. With spoken words, less mental effort was reported, but no differences on retention or transfer were found between groups. The audio-text group achieved the same performance level as the visual-text group, but with less mental effort.

In Experiment 3, modality effects were investigated in dependence on instructional pacing. Instructions consisted of static diagrams accompanied by narrations or visual texts and paced by system or learner. Referenced diagram parts were colored in red to minimize visual search. The system-paced instruction took 26.2 minutes; the learner-paced instruction was not restricted in time. Information units were arranged in linear order and learners could only move forward in line with the sequence. The university students participating were considered to have no prior knowledge of the topic. Study time, mental effort during instruction and testing as well as retention and transfer performance were assessed. No modality effect on study time was found. On average, the self-paced audio group learned only 2.6 minutes less (28.3 min.) than the self-paced visual group (30.9 min.). The average study times of the self-paced groups were not much longer than the duration of the system-paced instructions (26.2 min.). Mental effort was not influenced by text modality. Modality effects for retention and transfer occurred with system-paced instructions; with learner-paced instructions no effects were observed at all. Additionally, transfer performance was influenced by pacing. The learner-paced groups outperformed the system-paced groups, mainly as a result of the lower performance of the system-paced group that had to learn with visual texts. The transfer performances of the other three groups had the same level.

The aim of Tabbers' (2002) fourth experiment was to test the idea of time on task as the cause for the vanishing of modality effect with learner-paced instructions. Six conditions were used, resulting from the combination of two modality and three pacing conditions. The modality conditions consisted of static diagrams accompanied by (1) narrations or (2) visual texts, which were paced (1) by system as every narration part was played once, (2) by system as every narration part was played twice, or (3) by learner. Information units were put in linear order. Study time was 19.3 minutes for single system-paced instructions and 38.6 minutes for double system-paced instructions. The university students participating were considered to have no prior knowledge of the topic. Study time, mental effort during instruction and testing as well as retention and transfer performance were assessed. The average study time duration of the learner-paced instruction was 24.1 minutes. There was no difference in average study time between learner-paced instructions using auditory or those using visual texts. Mental effort was lower for students learning with spoken texts, no effects on mental effort were observed for solving retention and transfer tasks. Retention and transfer performance were affected by modality and pacing. A modality reversal effect of retention and transfer occurred with learner-paced instructions, whereas no modality effects appeared for system-paced conditions. Moreover, the double system-paced groups performed better at retention tasks than the single system-paced groups.

Stiller (2007, Exp. 2) investigated the effects of prior knowledge and computer attitude on modality effect. 10th grade German grammar school pupils learned with visual or auditory texts on the human eye in learner-paced and learner-sequenced instructions for 13 minutes. Modality effect was found to depend on prior knowledge and computer attitude. Overall, the low-prior-knowledge and negative-attitude learners in the auditory and visual group, respectively, did not differ significantly, but the high-prior-knowledge and positive-attitude learners benefited most from visual texts and were handicapped by spoken texts. Transfer and drawing performance were not affected by modality, but a superiority of visual texts over auditory texts was found as regards verbal retention and labeling pictures. The instructional usage was only marginally influenced by prior knowledge and text modality.

3 Discussion

Modality effect was found and amply evidenced by using system-paced instructions for experiments. For system-paced instructions which determine and limit time on presented information, spoken texts are regarded to be more effective because texts and pictures are perceived simultaneously; there is no danger of missing information or increased cognitive load due to split-attention. Learners studying a system-paced instruction based on visual texts suffer from dividing attention between texts and pictures, thus experiencing more extraneous load and learning disruption. Moreover, the visual entry channel might also be overloaded.

Tabbers (2002) argued that the standard rationale for explaining modality effect which is offered by the CLT and the CTML cannot explain the vanishing or the reversal of modality effect. As modality effect is assumed to result from split-attention and visual processing channel overload, these factors must be essentially influenced in order to obtain contradictory results. Tabbers (2002) rejects the assumption that presenting visual text and picture at the same time easily overloads the visual processing channel. Learners must split their attention between visual text and picture, but visual text is immediately transformed to phonological code and does not hinder the subsequent processing of pictorial information in the visual processing system. So, what are instances of factors influencing modality effect?

To a large extent, modality effect depends on medial features. One of them is the pace of the presentation, which determines the time available for meaningful information processing. Tabbers (2002) and Schmidt-Weigand (2006) have shown that extended time on task compensates mainly for split-attention effects, thus preventing working memory from being overloaded or at least relieving working memory from extraneous load. Extended time on task enables learners to spend more processing time on the verbal and pictorial information presented. Hence learners are more likely to select relevant information and to integrate it adequately within extended time. Thereby compensation of mainly split attention is balanced out by increased mental effort (Tabbers, 2002). To sum up, time on task can account for the vanishing of modality effect with system-paced and learner-paced instructions.

The argument of time on task cannot explain modality reversal effects that only occurred with interactive instructions. Time on task can only explain a vanishing of modality effect due to a compensation of split attention. The reversal of modality effect can be explained by an increased flexibility of information processing. Therefore Tabbers (2002) proposed that reading visual texts is a more active process that allows regulation of information processing. For this reason, learner-paced instructions with visual texts might be more appropriate in certain situations, suiting the learners' individual, cognitive needs. It is much easier, for example, to jump back and forth in written text in order to access or revisit pieces of information; contrarily, it is quite difficult to do this within audio files. A more flexible work-through might lead to a greater chance of retaining corresponding verbal and pictorial information in working memory, of improving mental representations by noticing hitherto unnoticed, but essential information, of mapping verbal and pictorial models adequately and of using prior knowledge for knowledge construction. In this context, listening is a rather passive process which makes it especially suitable for linear presentations with minimal interactivity, as in system-controlled instructions. Learner-pacing sometimes does not go along with normal listening behavior. Therefore spoken texts are sometimes not suitable for learner-pacing. To sum up, learner-pacing in the sense of guaranteeing sufficient time on the task in question only suggests that spoken and visual texts should be equally effective, because the factors detrimental to learning with visual texts can be compensated; learner-pacing in the sense of guaranteeing more flexible processing can account for the reversal of modality effect.

Modality effects also depend on learner characteristics. This was shown by Stiller (2007, Exp. 2) who found a vanishing of modality effect with pupils who have low prior knowledge and negative computer attitude, and a reversal of modality effect with high-prior-knowledge and positive-attitude pupils. But Stiller (2007, Exp. 1) also found modality effects with university students, independent of prior knowledge, using comparable learner-paced instructions as in Exp. 2. Hence, it seems to depend on the learners as to how they make use of learner-pacing as regards information processing in order to benefit from or be hindered by visual texts. Alternatively, these contradictory results might be attributed to text difficulty. Kalyuga et al. (1999) assume that speech would not be an effective instruction mode if texts were too long or complex, as this might also overload working memory. The possibility of overload is attributed to the fleeting character of auditory information that makes it more difficult to retain information in the memory. Visual text is permanent and can therefore be referred to repeatedly. The texts might have been more difficult for pupils than for students, hence pupils might have experienced overload with auditory texts. Overall, text difficulty might not be a cause of vanishing or reversal of modality effects in Tabbers' (2002) experiments, because the texts used were kept concise.

Jeung et al. (1997) believe that the reduction of working memory load with visual-auditory presentations would only enhance learning if mental resources were not devoted to an extensive, visual-based search in order to coordinate auditory and visual information. They found auditory texts to be helpful when a diagram was easy to process, but not when it was difficult to process. Their results showed that a modality effect was not observed with highly complex pictures; auditory text was only beneficial with complex pictures when visual attention was guided by visual aids like the flashing of pictorial parts concurrently with the corresponding text presentation. As Tabbers (2002) and Stiller (2007) cued the relevant pictorial parts of the pictures, this is not regarded as a cause for a vanishing of modality effects. In addition, the learning materials used were not of low complexity, so this could neither be a cause for a vanishing of modality effects (Tindall-Ford et al., 1997; Leahy et al., 2003).

As we have seen, it is not easy to predict the circumstances under which it is better to use visual texts with interactive instructions, because there is also ample evidence of modality effect with such instructions, with effects seeming to depend on learner characteristics. Hence, future research must specify which characteristics of instructions and learners determine, whether a modality effect will occur, vanish or reverse. Against that background, the decisions between auditory and visual text productions could be made more easily.

4 References

Brünken, R., & Leutner, D. (2001). Aufmerksamkeitsverteilung oder Aufmerksamkeitsfokussierung? Empirische Ergebnisse zur "Split-Attention-Hypothese" beim Lernen mit Multimedia. *Unterrichtswissenschaft*, *29*, 357-366.

Brünken, R., Plass, J. L., & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual task methodology: Auditory load and modality effects. *Instructional Science*, *32*, 115-132.

Brünken, R., Seufert, T., & Zander, S. (2005). Förderung der Kohärenzbildung beim Lernen mit multiplen Repräsentationen. *Zeitschrift für Pädagogische Psychologie*, *19*, 61-75.

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual task methodology. *Experimental Psychology*, *49*, 109-119.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*, 233-246.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, *94*, 428-434.'

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *15*, 313-331.

Issing, L. J., & Klimsa, P. (Hrsg.). (2002). *Information und Lernen mit Multimedia und Internet* (3. vollständig überarbeitete Auflage). Weinheim: Psychologie Verlags Union.

Jeung, H. J., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology*, *17*, 329-343.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*, 351-372.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, *92*, 126-136.

Leahy, W., Chandler, P., & Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. *Applied Cognitive Psychology*, *17*, 401-418.

Lukesch, H. (1998). Einführung in die pädagogisch-psychologische Diagnostik. Regensburg: Roderer.

Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E., Dow, G., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, *95*, 806-813.

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., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*, 107-119.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

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. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, *94*, 598-610.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based multimedia learning: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, *19*, 177-214.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, *87*, 319-334.

O'Neil, H. F., Mayer, R. E., Herl, H. E., Niemi, C., Olin, K., & Thurman, R. A. (2000). Instructional strategies for virtual aviation training environments. In H. F. O'Neil & D. H. Andrew (Eds.), *Aircrew training and assessment* (pp. 105-130). Mahwah, NJ: Erlbaum.

Rinck, M., & Glowalla, U. (1996). Die multimediale Darstellung quantitativer Daten. *Zeitschrift für Psychologie*, *204*, 383-399.

Schmidt-Weigand, F. (2006). Dynamic visualizations in multimedia learning: The influence of verbal explanations on visual attention, cognitive load and learning outcome. Retrieved 1-31-2006, from http://geb.uni-giessen.de/geb/volltexte/2006/2699/.

Stiller, K. D. (2007). Computerised multimedia learning. Modes of text presentation and access to text. Hamburg: Verlag Dr. Kovač.

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

Tabbers, H. K. (2002). The modality of text in multimedia instructions. Refining the design guidelines. Heerlen: Open University of the Netherlands.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When to sensory modes are better than one. *Journal of Experimental Psychology: Applied*, *3*, 257-287.