DIPLOMARBEIT

Titel der Diplomarbeit

“Audio-visual comprehension in theory and foreign language teaching practice: How AVC can be established as a communicative skill”

Verfasserin

Petra Ederer

angestrebter akademischer Grad

Magistra der Philosophie (Mag. phil.)

Wien, im Mai 2010

Studienkennzahl lt. Studienblatt: A 190 344 313
Studienrichtung lt. Studienblatt: UF Englisch, UF Geschichte
Betreuerin: Mag. Dr. Susanne Reichl
Acknowledgements

First and foremost, my special thanks go to my supervisor, Mag. Dr. Susanne Reichl, who has given me the best guidance and support. Her helpful advices, constructive criticism and great encouragement were crucial to the outcome of this thesis.

I would like to dedicate this work to my parents, Harald and Eva, who have always supported me during my studies.

Finally, I would like to thank my sister Ines and my boyfriend Maximilian for their patience, encouragement and help of any kind.

Danksagung


Die Diplomarbeit widme ich meiner Familie, allen voran meinen Eltern Eva und Harald, die mir das Studium ermöglichten und mich immer förderten.

Ein besonderer Dank gilt meiner Schwester Ines und meinem Freund Maximilian für ihre moralische und anderweitige Unterstützung.
INTRODUCTION.................................................................................................................. 1

1. AVC IN RESEARCH........................................................................................................... 5
   AVC in the 1970ies ........................................................................................................... 5
   AVC in the 1980ies ........................................................................................................... 7
       Schwerdtfeger’s pioneer work ..................................................................................... 7
       Biechele’s “complex comprehension” ...................................................................... 9
       Bufe’s and Scherer’s concept of interchannel redundancy .................................. 12
       Janßen-Holldiek’s views on image/sound-relations ........................................... 13
       Paivio’s dual coding theory ..................................................................................... 20
   AVC in the 1990ies ........................................................................................................... 24
       Mayer: Cognitive theory of multimedia learning (CTML) .................................. 24
       Baddeley and Hitch: Working memory .................................................................... 27
       Sweller’s cognitive load theory .................................................................................. 29
   AVC in current literature ................................................................................................. 34
       Schnotz and Bannert: Integrated model of text and picture comprehension (IMTPC) ............................................................................................................................................................................ 34
       As I see AVC .............................................................................................................. 41

2. AVC IN AUSTRIAN CURRICULA FOR FOREIGN LANGUAGES .................................. 43
   Curriculum for foreign languages in lower AHS-secondary school ........................ 43
   Curriculum for foreign languages in upper AHS-secondary school ..................... 47
   Summary: The role of AVC in both curricula ............................................................ 50
   Why is AVC not an officially established skill? ......................................................... 52

3. AVC IN TEACHING PRACTICE ....................................................................................... 55
   Why should films be used in foreign language teaching? .......................................... 57
   What is AVC from a pedagogic viewpoint? ................................................................. 60
   Four different teaching designs ................................................................................... 62
       1. AIW: AVC as an instrument for teaching beginners ...................................... 62
       2. AIW: AVC as a learning goal ............................................................................ 73
       3. Seeing, hearing and feeling South African history: Teaching power relations during the 1970ies through AVC ................................................................. 83
       4. How to get students to speak: Using a video clip ........................................... 102

CONCLUSION..................................................................................................................... 109
**Introduction**

In the external world objects and events produce multi-sensory signals which are perceived in more than one sense. If stimuli of different sensory modalities such as auditory and visual are transmitted simultaneously, their occurrence is defined as “audio-visual”. As most - if not all - stimuli in the external world are received and processed by its environment, i.e. living beings, the process assigned to the reception and processing of audio-visual signals is regarded as “audio-visual comprehension” (henceforth “AVC”). AVC happens in real life, as when a person simultaneously receives and processes, for instance, the image and sound of a singing bird or playing children. Regarding these real-life examples as a set of signals transmitting a planned or unplanned message - in short, a text - AVC is not restricted to natural events and objects, but facilitates the comprehension of any audio-visual text such as a filmic text.

If the idea of a text as a set of signals with a message is transferred into the context of foreign language teaching, it becomes evident that texts are often exploited for their communicative nature, which is due to today’s predominant “Communicative Language Teaching” approach, which utilizes communication as both the means and the ultimate goal of learning. In most cases, texts of just one modality - either auditory (listening text) or visual (reading text and two or three-dimensional images) - are an integral part of foreign language teaching; multi-sensory texts, however, play a rather minor role. Among various existing explanations for the underestimation of the capacities of the audio-visual text, this thesis has uncovered one major reason. The undervaluation of films is closely linked to the undervaluation of AVC. Next to the officially established communicative skills listening, reading, speaking and writing, all of which can be assigned to mono-sensory texts, AVC is not considered a skill of its own. Striking evidence therefor derives from a close analysis of the Austrian curricula for foreign languages in lower and upper AHS-secondary schools, constituting the theoretical framework for foreign language teaching in Austria, which shows that AVC is not acknowledged as a skill to comprehend audio-visual texts, but

---

1 In fact, the curricula redivide the skill “speaking” into two distinct skills “spoken production” and “spoken interaction”, which yields altogether five communicative skills (along with listening, reading and writing). This classification is discussed in more detail in chapter 2.
the understanding of films is consigned to listening, which - from my viewpoint and that of many researchers - is simply not true. A consequence of the “missing skill”, as I call it, is that teachers are and stay uninformed about the potentials of AVC, missing out on a vast number of meaningful activities involving AVC. Therefore, it has not and cannot become an integral part of teaching practice. This thesis, then, is the result of my motivation to establish AVC as a communicative skill next to the other skills in order to fully integrate it into foreign language teaching practice.

This endeavor requires various steps, starting with a definition of AVC (see below) and the theoretical examination of AVC in multiple research disciplines such as foreign language teaching research, linguistics and cognitive psychology, which provide different models of audio-visual signal processing. From this review, which ranges from the 1970ies up to the current state of research, the most significant empirical findings and ideas are collated, adapted and finally integrated into a model of AVC as I see it (chapter 1). Moving from theory to practice, a small intermediate step, however, must be taken towards the planning and guidance of all teaching, the curriculum. The two curricula for foreign languages in lower and upper AHS-secondary school, laying the groundwork for every single teaching design, will be systematically analyzed in regard to their view of AVC (chapter 2) in order to legitimize - together with the established model - the didactic implications of AVC (chapter 3). Constituting the heart of the thesis, four different practical applications of AVC realized as explicitly detailed teaching designs will be presented. Each of the planned lessons involves AVC in a different way to a different degree with different target groups, i.e. students of different school grades and language proficiencies. This deliberate choice shall demonstrate the high potential and the various possibilities of AVC in foreign language teaching contexts and thus contribute to an appreciation of AVC in theory and its implementation in practice.
What is AVC?

In the following definition, I consider AVC only in terms of the comprehension of audio-visual texts such as films and disregard the comprehension of real world texts. For me - and as it is understood in this thesis - AVC is the simultaneous receiving and processing of auditory and visual stimuli/signals. The latter term “visual” thereby denotes the dynamic pictures; the term “auditory” refers to the narration either through a narrator or character of the film. The external stimuli appear as codes which are sensory-specific. They are transmitted via two sensory channels, the auditory and the visual channel, into the human brain. Within the processing of the stimuli, their representations are no longer sensory-specific, but they interact with each other and eventually form a complex mental construction integrating information from both types of stimuli (“bottom-up processes”) and already existing schemata stored in the human brain (“top-down processes”). AVC is thus a complex and active cognitive process involving multiple sub-processes.

2 To me the term “audio-visual text” denotes motion pictures, video clips as found on the website “utube”, TV shows, short films and other audio-visual productions, all referred to by the term “film”. I will use the terms “audio-visual text” and “film” interchangeably.

3 Clearly, films do not only comprise narration, but consist of various other sounds such as music, sound effects, live and recorded sounds; however, for my theoretical discussion of AVC in the first chapter I want to consider only narration as auditory stimuli as they provide the most information in regard to the message of the whole text. The pedagogical applications of AVC in the third chapter will take all auditory stimuli into account and thus also consider aspects such as a scene’s atmosphere created through music or its portrayal of real life situations expressed through live sounds. Audio-visual texts have different image/sound relations. For example, in regard to their interactions, images and sounds can complement, potentiate and modify each other just as they can diverge. Considering the type of presentation, auditory stimuli can be on-screen and off-screen presentations. These ideas will be discussed in much more detail in chapters 1 and 3.
1. AVC in research

AVC in the 1970ies

[Film] - what more appropriate way to study the language than to be exposed to a situation in which the student can see and hear native speakers in natural and presumably entertaining situations? (Markey 1976: 731)

This hymn of praise is given by an early advocate named Markey and represents the euphoria many researchers showed when film was discovered by foreign language teaching research. In the 1970ies the audio-visual text was appreciated by teaching experts for various reasonable aspects such as the authenticity of language and the presentation of entertaining situations involving “natural characters”, among which the most striking argument was the simultaneous transmission of both auditory and visual signals.

Markey, for example, based his arguments on the multi-sensory nature of the audio-visual text and argued that it comprised the best way to teach authentic language use, as it was linguistically closer to reality than any other text could be, especially when compared to the language laboratory, a popular and common audio-visual installation for language learning in the 1970ies.

---

4 In this chapter, research is drawn from multiple disciplines such as foreign language teaching research, linguistics and cognitive psychology because AVC comprises an interdisciplinary field of research. Before I start with the outline of different theoretical concepts of AVC for the comprehension of film, I want to make clear that the models I am going to present do not exclusively deal with the definition of AVC as given before. In fact, only few models of AVC as film comprehension exist. As it will turn out in chapter 1, I had to take studies into account which deal with other interpretations of AVC than the processing of dynamic pictures and narration, such as the comprehension of illustrated texts and the processing of static pictures accompanied by sounds, in order to explore how AVC works. A deeper reflection on this problem is given at the end of chapter 1.

5 At this point, I want to briefly discuss the term “natural situation” when applied to films. On the one hand, films are not natural or authentic inasmuch as they are not images of the real world captured on screen, but a theoretical conceptual design of what might be the world. On the other hand, it is legitimate to call film realistic and natural because the presented plot, characters and situational circumstances try to picture real people, actions and situations. This discussion, however, goes beyond the scope of this thesis. See Borstnar, et al. 2002: 29ff.

6 In the 1970ies film entered first language acquisition research as well. In contrast to foreign language learning research, the discussion centered the question on how media education could be defined and put into practice.

Considering Markey’s argument and that of other researchers, I would argue that the link between the audio-visual text and AVC was not yet known in the 1970ies. Furthermore, film was considered mostly in terms of how it could benefit reading, speaking, writing and especially listening. Markey, for example, suggested that films should be used in order to enhance already developed listening skills. This implied that audio-visual texts could only be used with more advanced students, i.e. of higher grades. Markey also regarded audio-visual texts (he primarily argued in favor of literary adaptations) as a fruitful source for the development of speaking and writing skills, which could be achieved by activities involving reactions to and impressions of audio-visual realizations of books.

Markey’s claims demonstrate that he did not consider AVC as a skill of its own, but that he promoted it as a stimulus and material for the development of the four communicative skills: “The combination of text and film seems like an ideal one for the development of the four skills” (Markey 1976: 731). Despite his unawareness of AVC as a skill, I appreciate his ideas insofar as he spoke out for the audio-visual text and its significance in foreign language teaching for the first time.\(^8\)

From Markey’s and other writings it can be concluded that AVC as a concept did not exist in foreign language research in the 1970ies and that film was discussed for other reasons than AVC.\(^9\) Therefore I want to move on to a historical examination of AVC in the 1980ies.

AVC in the 1980ies

The 1980ies are characterized by numerous publications on film in foreign language teaching. It turns out, however, that most of them are far from properly researched as they are field reports, teaching documentations and handbooks offering prefabricated “recipes” and guidelines without any substantial theoretical background. The other part of the vast field of publications stays on a more theoretical level; however, most of them provide the same lists of the advantages and disadvantages of film. However, within the discussion of film, the debate on AVC as a way to comprehend film falls relatively short. Also, only few hard facts are known about it. Agreeing with this observation, a researcher named Tomalin lamented that most publications from the 1980ies would provide only “proto-theoretical knowledge” (1988: 8) about the comprehension of audio-visual texts.

This lack of empirical research seems to be dramatic. Nevertheless, I think that knowledge about AVC as it existed in the 1980ies is not less valuable, because it is mainly based on experience and good theoretical thinking. What I can interpret from it, however, is that AVC was a complex research effort that only few researchers made. Their ideas are presented in the following.

Schwerdtfeger’s pioneer work

ACV fell short until Schwerdtfeger raised her voice in 1989. Of all her writings, her greatest contribution to the discussion of AVC was the coining of the term “Seh-Verstehen” (“visual comprehension”, Schwerdtfeger 1989: 24). Moreover, she promoted it as the fifth skill (“5. Fertigkeit”, Schwerdtfeger 1989: 24) in foreign language teaching, which should equally exist alongside listening, reading speaking and writing.

---

10 In the course of my research, I experienced that a large number of pseudo-empirical and half-theoretical discussions circulates among the many publications on the audio-visual text and its use in foreign language teaching - from the 1980ies onwards until now. This has made my research on AVC quite challenging.


12 This subchapter is based on Schwerdtfeger: 1989.
Another achievement which greatly contributed to a serious discussion of the audio-visual text was the stressing of the differences between audio-visual and purely visual texts, by which Schwerdtfeger understood written texts. She said, “[a] film is not a book” (1989: 13). Although this statement seems to be logical as well as simple, it constituted an important tactical move in the discussion of film, because some researchers, most of all Ehnert, argued that the audio-visual text should be banned from foreign language teaching. According to him, 65 to 95 per cent of 30 to 45-minute feature movies were “allein von der Form her nicht geeignet Informationen zu transportieren” (qtd. in Schwerdtfeger 1989: 11-2). He argued that the pictures within the audio-visual text would move too fast and could not be processed due to film techniques such as camera angles and lighting. Film should therefore adopt more “book-like” characteristics, i.e. be static. Schwerdtfeger accused Ehnert of trying to turn films back into books and claimed that the two media represented “two halves of a ball” (Schwerdtfeger 1989: 12). With each of the two halves, certain teaching goals could be reached which would not be possible with the other. Therefore, the very essence of the audio-visual text, the simultaneous sending of auditory and visual information, would facilitate activities which would not be possible with written texts and thus provide new and different teaching objectives.\(^{13}\)

I would argue that Schwerdtfeger laid the foundation for the development of the concept of AVC by coining the term “visual comprehension” and promoting it as the fifth skill in foreign language teaching. For the first time, the audio-visual text was considered in correlation with a skill enabling the processing of the text. Although Schwerdtfeger did not consider AVC a holistic process, as she disregarded the linkage of visual to auditory signals, i.e. the nature of the audio-visual text, she greatly contributed to the appreciation of the visual component of AVC.\(^{14}\)

---

\(^{13}\) Schwerdtfeger compared films to course books (“visual media”) and film language to print language and pointed out differences between the two media. Cf. 1989: 14-6.

\(^{14}\) Whenever AVC is discussed, Schwerdtfeger’s name should be mentioned, because she is considered the “pioneer” of AVC.
Contrary to Schwerdtfeger, who stressed the visual aspect of AVC, but ignored the auditory side, Biechele centered her discussion exclusively on the multi-sensory character of AVC. She argued that except for telephone conversations, radio programs, speaker announcements and tape recordings, all stimuli are received via both the auditory and the visual channel. Most real-world listening processes are thus audio-visual and require AVC.

The interaction of visual signals with auditory ones is defined by Biechele as “complex comprehension” (“komplexes Verstehen”, Biechele 1988: 87), which she explained as follows:

…das adäquate Aufnehmen und Verarbeiten der kommunikativen Situation in ihrer Gesamtheit, d.h. das Erfassen der übermittelten Sprachzeichen und der nichtsprachlichen, sowie situativen, kommunikativen Funktion tragenden Informationen und deren intentionsgerechte, partnerbezogene und situationsgerechte Dekodierung und Interpretation (Biechele 1988: 87).

Biechele - in this and some other publications (e.g. 1989) - did not use the term “AVC” (“Hör-Seh-Verstehen” in German), but she preferred the term “complex comprehension”. The process which is described by complex comprehension is very similar to the concept of AVC (as I defined at the beginning of this thesis). This fact becomes even more evident in one of her recent writings, in which Biechele used the term “Hör-Seh-Verstehen” and promoted it as one of the communicative skills (Biechele 2006: 310).

I would argue that Biechele used to prefer the term “complex comprehension”, because it adequately described the complexity of AVC. She argued that not only auditory information is transmitted, but also other signals such as gestures or facial expressions, which together constitute a complex entwinement of various stimuli (see above: “…das adäquate Aufnehmen und Verarbeiten der kommunikativen Situation in ihrer Gesamtheit…”). However, in the end, Biechele may have renamed the concept as “Hör-Seh-Verstehen” because it

constitutes the more clear-cut definition for the processing of auditory and visual signals.

Another claim by Biechele, which I support, concerns one of her pedagogic conclusions. She argued that communication outside the classroom, “real” communication of the external world, is a complex process involving more than one sense. Most of this communication, she argued, happens audio-visually. As foreign language teaching should prepare students for the real world, it should give them all the skills they need to communicate successfully. Thus, AVC has to be an integral part of teaching if students should decode the audio-visual signals they receive in communication with other people.

**How does AVC work according to Biechele?**

The audio-visual text exists as a synthesis of images and sounds (i.e. visual and auditory stimuli). This combination consistently differs in the process of transmission in regard to the quantity of information (Which component - the visual or the auditory - plays the greater role in a sequence of delivery for the overall understanding?) and quality (How do both components work together in the process of meaning-making? Are they redundant, complementary, parallel or contrary?). In addition to these two aspects, two more levels are identified, “die Ebene temporaler Relationen” and “die Lage der auditiven Quelle” (both: Biechele 1989: 90). The first concept corresponds to the temporal correlation of visual and auditory components (Are they simultaneous or gradual?) and the latter refers to the tone being switched on or off and the image being transcendent or -immanent. The term “image-immanent” refers to the sounds appearing within a film scene (e.g. the voices of the characters). Here the auditory component is captured in the transmitted picture itself and represented through the visual component. As a contrast to that, the “image-transcendent tone” is not located in the moving picture, but beyond it. This can be, for example, the narrator’s voice, which comments on the events of the scene.

---


In summary, Biechele identified four different types of image/sound-relations:

- type 1: semantic relation
- type 2: temporal relation
- type 3: quantitative relation
- type 4: type of auditory source

Biechele reinforced her hypotheses of the interaction of visual and auditory signals by referring to a study by Scherer and Marino (1980). In their attempt to isolate the visual from the auditory information, they failed because – as they claimed – the visual and auditory signals do not exist in a parallel way. Thus, the visual components do not only function as a support to the understanding of the audio-visual text.\(^\text{18}\) Biechele agreed with Scherer and Marino and accentuated that it is not the primary function of the audio-visual text to develop listening skills by providing visual aid, but that both types of signals are equally involved in the process of comprehension. Also, AVC does not equal listening skills; therefore, knowledge about listening comprehension cannot be transferred to AVC.

Another study, from which Biechele drew in her argumentation, attempted at explaining type 1, semantic relation, in more detail. Bufe and Scherer developed the concept of “interkanalige Redundanz” (1981: 99) (henceforth “interchannel redundancy”) to refer to the redundancy between auditory and visual information, which leads to a better understanding of the overall audio-visual text.\(^\text{19}\)

At this point I also want to mention how Biechele explains the existence of negative judgments on film in teaching. She argues that these are due to the assumption that films must work the same way as “verbal media” (i.e. listening texts), which automatically causes frustration, because the two texts differ in many ways due to their very nature. Some of those negative judgments were: “Images cause a sensory overload”, “The image distracts students from

\(^{18}\) Qtd. in Biechele 1989: 89.
listening properly” and “Students are not used to having two channels at their disposal”.20

**Bufe’s and Scherer’s concept of interchannel redundancy**21

Before I present the concept of interchannel redundancy, I want to mention that it was drawn from a research project called “Hörverstehen im fernsehunterstützten Fremdsprachenunterricht Französisch”. The empiric study was based on the assumption that visual information (i.e. illustrations) influences the comprehension of the overall text in a positive way. The results confirmed this assumption by showing that both narrative and descriptive written texts were received better in terms of global and detailed understanding if they were accompanied by illustrations. It was concluded that the visual signals would influence the understanding of the written text. Another outcome of the study was that the audio-visual text contributed positively to the long-term storage of information. In short, the research project strongly argued for a combination of both visual and auditory stimuli.

For their argumentation for interchannel redundancy, Bufe and Scherer regarded the audio-visual text as having two sensory channels (the visual and the auditory channel) and two semiotic systems, containing iconic (henceforth: “visual”) and the linguistic signals (henceforth: “verbal”). The semantic relation between the two different semiotic codes is most important for the comprehension of the overall information. Therefore, Bufe and Scherer strongly argue in favor of the semantic image/sound-relation (identified as “type 1” by Biechele), as having a key position in AVC. By using Fillmore’s “Case Grammar” (“Kasusgrammatik”), Bufe and Scherer categorized four different semantic image/sound-relations, which constitute the four different levels of interchannel redundancy. For example, on level 1, called “audio-visual synonymy”, both the visual and the auditory channel carry identical information. For instance, the visual signal and the auditory signal “chair” are transmitted at the same time. This means that a video clip would show the image of a chair together with the pronunciation of the word chair (e.g. spoken by a narrator).

---

21 This chapter is based on Bufe and Scherer: 1981 and Batz and Ohler: 1984.
The semantic relations get less intensive from level 1 onwards and show the least intensity at level 4, on which the transmitted information can only be decoded through cognitive inferences based on data that have already been stored in the brain before.

Generally, Bufe and Scherer regard interchannel redundancy as being beneficial for comprehension, because it secures the transmission of signals. The four different levels of redundancy describe how close the semantic relation between the visual and the auditory stimuli is, and, accordingly, how much inference is needed. Based on their findings, they conclude that visual stimuli yield a better understanding of auditory information.

Before my examination of research literature returns to Biechele, I want to discuss interchannel redundancy of image/sound-relations in more detail, because they might give useful insights into how AVC works. For this, I want to discuss Janßen-Holldiek’s ideas because she elaborated Bufe’s and Scherer’s theories and, hence, took a closer look at ACV.

**Janßen-Holldiek’s views on image/sound-relations**

Janßen-Holldiek (formerly Marino) promoted a crucial principle for her argumentation: “Information is not conveyed, but created” (1984: 68). According to her, the receiver of the information delivered through the audio-visual text chooses from a range of data according to his/her previous knowledge and current expectations. This phenomenon is considered both a process of “perceptional production” and “productive perception”. In this sense, Janßen-Holldiek promoted the active role of the recipient in the triangular relation “sender-medium-recipient”, of which each party communicates with the other two interactively. Consequently, manipulation or guidance by the teacher would be possible. The choice of sensory data could be one option to induce a certain way of interpretation with the recipient. Another way could be the monitoring of the recipient’s attentiveness by stressing important information (also: repeating it) or reducing irrelevant data from the transmission.

---

22 The outline of Janßen-Holldiek’s views is based on her publication from 1984.
A crucial factor of AVC is the process of recoding. All sensory data that is received goes through transformation processes in order to be linked to already existing mental structures. In audio-visual transmissions, this process works on various levels such as phonology, morphology or syntax and also in the field of semantics. How well or fast semantic recoding processes work, depends on already developed recoding skills and the recipient’s previous knowledge. Thus, comprehension processes do not only work one way, from the sign upwards, which is called “bottom-up”, but also “top-down”, which refers to the recipient’s preformed hypotheses and experiences, in order to interpret the received signals.

By combining the concept of recoding with the model of interchannel redundancy, Janßen-Holldiek came to the conclusion that redundancy is the degree of congruency between mental representations of both sensory systems, which are formed through recoding processes. Janßen-Holldiek also argued that the level of semantic relation between auditory and visual representations, hence interchannel redundancy, is primarily controlled by those recoding processes which link visual stimuli to their mental representations. Recoding processes, however, do not work the same way for all semantic relations. In the following, I want to give examples of Janßen-Holldiek’s categorization system.

**How do different recoding processes influence the levels of redundancy?**

Comprehension not only depends on the recipient’s recoding skills, but also on how extensive and fast his/her recoding processes work. Janßen-Holldiek claimed that comprehension is best when there is only little or no recoding effort involved, as when the mental representation in the recipient’s mind is identical (or at least quite similar) to the stimuli. This case is given when the presented visual signal serves as a repetition of the auditory one. She called this case “direct image/sound-relation” (which is similar to Bufe’s and Scherer’s definition of “audio-visual synonymy”). Only individual concrete objects can be presented in such a redundant way, as when an image of a chair and its linguistic expression “chair” are simultaneously delivered on screen.
I would argue that direct image/sound-relations are not very likely with most video and film formats, simply due to the medium’s entertaining and informative character, but they may rather occur in "traditional" and old-fashioned foreign language teaching videos, in which vocabulary was taught by presenting a visualization of the word (e.g. an image of a chair) together with its pronunciation (e.g. “/tfɛθ(ʊ)/”).

Apart from direct image/sound-relations, all other types of semantic relations need certain abstraction processes on the visual level in order to link the image to the auditory representation. This happens with almost all signals delivered through film. The image “chair” serves again as an example. It is not comprehended as “a specific chair” (if the picture does not show any exceptional characteristics that would make the overall image shown a specific kind of chair), but is recognized as “the category chair”, a representative of the category of chairs. The farther, semantically, the represented signal is from the prototype of the category, the closer the image/sound-relation is to the rhetoric figure of “synecdoche”. This term refers to the act of substituting a concept by another concept from the same semantic field, which has either a broader or narrower meaning. The substituted concept is therefore replaced by a superordinate or subordinate term. With both cases - a stimulus representing a category and a stimulus being substituted by a super-/subordinate concept - more complex recoding processes are involved, because the signals have to be recognized not only as an object, but also as being part of a certain semantic field. These presumably complicated recoding strategies have a positive influence on the process of AVC.

Classifying image/sound-relations according to the levels discussed above can be problematic in regard to “Rekodierritual[e]” (“recoding rituals”, Janßen-Holldiek 1984: 75). Recoding rituals are the result of image/sound-relations that have become semantically close through socio-cultural constructions. They can also be regarded as “social clichés” (“soziale Klischees”, Janßen-Holldiek 1984: 75). Certain abstract concepts are bound to their visual representations so much that a recoding loop between the visual stimuli and its mental representation is not needed anymore. The image is no longer interpreted
according to its features, but is received as a symbol, which results in a direct relation between the visual and the auditory representation. This process is elaborated by the following example.

A nuclear power plant is shown on screen. Janßen-Holldiek argues that its very characteristic shape has become a symbol of atomic energy for us (with “us” she means representatives of cultures which are familiar with the concept of atomic energy). The recipient no longer sees a tapered construction, but recognizes it as a symbol of atomic energy. The decoding process, which is normally involved with abstract concepts such as atomic energy, is redundant, because the image automatically evokes, in fact, stands for, the symbol of atomic energy. Thus, the assumingly associative relation between the visual and the auditory representations becomes a direct one. This means that the recipient’s mind recognizes the symbol “atomic energy”; at the same time he/she sees the image of the nuclear plant and hears the linguistic expression “atomic energy”.

The opposite way is also possible. Objects that are usually comprehended in a larger semantic context can - if they are presented isolated from their usual context - be perceived as strange, because they are no longer surrounded by their additional semantic information. Consequently, the recoding process gets more complex and the overall comprehension takes longer. I would say that such a case is given when, for instance, only a single part of an object is shown without seeing the rest of the object, as when the screw head of a screwdriver is presented without giving more information from where it is taken. In this case, it is hard for the recipient to recognize the screw head immediately as the top part of a screwdriver.

Recoding rituals do not work the same way for every recipient, but are bound to personal, spatial and temporal contexts. As I would additionally claim, they are also dependent on cultural contexts. For example, a recipient from a culture that has never had access to real nuclear atomic plants or pictures of them, will not be able to decode the link between the tapered shape of a nuclear power plant
Despite her reasoned argumentation, Janßen-Holldiek admits that if her ideas are put into practice in a foreign language teaching context, the focus of analysis has to shift to the learners, their language proficiencies and their individual living environments, which are socio-culturally determined. Also, the question of what recoding processes the learners will most likely apply to deal with specific audio-visual signals becomes crucial. It is then the teacher’s task to analyze the processes involved in decoding the visual information and linking it to the auditory representations. As a next step, adequate redundancy levels have to be assigned. With the knowledge gained, the teacher can then design appropriate and meaningful activities and tasks.

The following visualization represents my personal interpretation of Janßen-Holldiek’s ideas:

The semantic relation between image and sound of the audio-visual transmission is most decisive for the comprehension of the text. The type of semantic relation is thereby heavily influenced by the mental representations of
the visual signals. A direct image/sound-relation involves no recoding process, because both types of sensory signals are interpreted in their denotative forms. Other semantic relations require complex recoding processes, because the information of the stimuli can only be presented through associations. However, if these semantic relations have been automatized by recoding rituals, which depend on personal, cultural, spatial and temporal conditions, they become again direct relations.

Criticism

Obviously, Janßen-Holldiek regards the recoding process not only as a crucial factor, but also a sub-skill of AVC. However, she admits that there are no general or universal criteria for AVC, but that the type of learner and his/her socio-cultural background should be taken into consideration when the audio-visual text is used in teaching contexts. This means that the students and their most probable way of dealing with the audio-visual text have to be analyzed first, in order to assign their interpretations to the appropriate redundancy levels, which then should explain comprehension. In other words, it is the teacher’s responsibility to put him-/herself into every individual student in order to make didactic decisions about how to use the audio-visual text meaningfully in the classroom.

Janßen-Holldiek’s model has proved unsatisfactory for foreign language teaching, as it confronts teachers with too much interpretative work of what their students can or cannot do with the audio-visual text. Moreover, it raises critical questions in regard to its completeness. The above-presented illustration shows that only the relation between visual stimuli and their mental representations are discussed, triggering the following questions: How are auditory stimuli linked to their mental representations? Are these relations also processed through recoding strategies? If so, how are representations of visual and auditory stimuli explicitly inked to each other? The model is now depicted with question marks:
The processes on the visual side of AVC can be explained by Janßen-Holldiek’s model. Regarding the processes on the auditory side, however, it gives no answers to the questions asked before. Before Janßen-Holldiek’s ideas are dismissed as being “incomplete” for my discussion of AVC, I briefly want to consider another text (Janßen-Holldiek: 1984) by her.


Janßen-Holldiek’s model is insufficient not only in regard to the question of how auditory stimuli are represented, but also in regard to how recoding skills originate. For me, some questions arise such as “How do recoding skills evolve, e.g. from experience?” and “Can they be learned and taught?”.
The first text passage says that both visual and auditory signals (“verbal[er] Kontext”) stimulate mental representations. This is confirmed by the second quote, which discusses interchannel redundancy (referred to as “der subjective Eindruck”) and argues that the audio-visual presentation (which means both visual and auditory stimuli) can trigger mental representations.

The quotes show that both sensory signals, auditory and visual stimuli, activate mental representations; the questions as to how mental representations of auditory stimuli exactly evolve, how they interact with visual representations and to what extent they are involved in AVC remain unresolved.

**Paivio’s dual coding theory**

Digging deeper into the research literature, it is particularly striking that the works of Paivio, who developed the dual-coding theory (henceforth: “DCT”), are very often quoted.25

Paivio posits that cognition in general involves two distinct systems: a “verbal system” (henceforth: “auditory system”), which is specialized for dealing directly with verbal codes (auditory stimuli) and a nonverbal system (henceforth: “visual system”), which is specialized for dealing with nonverbal stimuli (visual stimuli). Both the auditory system and the visual system are composed of mental representations, called “logogens” and “imagens”. They are assumed to be modality-specific, which means that logogens respond to auditory signals and imagens correspond to visual ones. While the auditory system only processes information sequentially (i.e. one piece after another), the visual system can deal with more than one signal at a time and is thus a parallel system. Logogens are organized in terms of associations and hierarchies, whereas imagens are arranged in terms of part-whole relationships.26

Logogens and imagens are connected by three processes. “Representational processing” as the first type is the direct activation of logogens by auditory stimuli and the direct activation of imagens by visual stimuli. For example, the

---

25 The information in this chapter is drawn from Paivio: 1986 and 2006.
word “chair” can activate an association with the logogen “chair”. Paivio claims that all cognitive tasks involve representational processing. The second type, “referential processing”, runs via both representational systems. An example is the auditory signal “chair”, which triggers a visual representation. The third type, “associative processing”, refers to the activation within one representational system. For example, the logogen “chair” generates an association with the logogen “table” and the imagen “chair” stimulates, for example, the imagen “floor”.

The following figure visualizes Paivio’s DCT more explicitly:

The auditory and visual system work independently from each other, but they are interconnected to varying degrees. Moreover, the interaction of the auditory and the visual system mainly depends on the audio-visual text:

The theory means that both systems are generally involved even in language phenomena. The verbal system is a necessary player in all “language games” but it is sufficient in only a few. In the most
interesting and meaningful ones, the verbal system draws on the rich knowledge base and gamesmanship of the nonverbal system. Conversely, the nonverbal system cannot play language games on its own, but it can play complex nonverbal “solitaire”. The verbal system dominates in some tasks (crosswords is a simple example) and the nonverbal imagery system in others (e.g. jigsaw puzzles). Cognition is this variable pattern of the interplay of the two systems according to the degree to which they have developed. (Paivio 2006: 3)

Paivio’s DCT postulates that the two systems can be developed, which implies that AVC can be learned and taught and hence be regarded as a skill.

The heart of DCT, as applied in foreign language teaching (DCT has also been applied to a wide range of disciplines such as psychology), is that learning is enhanced if visual information is presented with auditory signals (called “verbal information” by Paivio). This is due to the activation of two systems (either automatically or through deliberate effort) and thus the construction of separate but linked memory traces - one in each system. In short, Paivio claimed that double storage is better than single storage. He further argued that visual representations are stored in both systems, which facilitates easier recalling. Thus, visual stimuli are more memorable than auditory ones.

Criticism

Paivio’s model is problematic insofar as that it provides no superior central construct which combines the information from both sensory stimuli. In short, it misses an independent system to which auditory and visual representations can be connected. Another problem arises with Paivio’s definition of “learning”. Paivio applied his ideas mainly to one aspect of learning, the recalling of vocabulary. DCT thus proposes a very limited view on AVC.

A large number of empirical studies were conducted to argue against the model of dual-coding. Also, various studies were directed at testing DCT against its opponents. However, in cognitive research, in which study outcomes are often a matter of interpretation, both sides, DCT-advocates and DCT-opponents, interpreted the findings in favor of their own arguments. Because there are no
hard facts, non-experts are forced to decide on one side of the argument. I do not want to judge either side, but I feel that Paivio’s model is not efficient enough to explain AVC processes for my purposes. Nevertheless, DCT provides a useful approach - especially in regard to the idea that the two channels are linked and that auditory representations evoke visual ones and vice versa. These cornerstones taken from DCT will now be elaborated by other researchers, for example, Mayer, which leads me now into the 1990ies.\footnote{In an early version of my thesis I discussed Engelkamp’s and Zimmer’s multimodal model at this point of the research outline. However, the more closely I investigated their model, the more inappropriate it seemed to me to fit into my discussion of AVC. I therefore dropped the discussion altogether. For more information on the multimodal model see: Engelkamp, Johannes. 1998. \textit{Memory for Actions}. Hove: Psychology Press and Engelkamp, Johannes and Zimmer, Hubert D. 2006. \textit{Lehrbuch der kognitiven Psychologie}. Göttingen: Hogrefe Verlag.}
AVC in the 1990ies

At the beginning of the 1990ies research from neuropsychology suggested the existence of multimodal representational systems, all of which promoted a central conceptual system - one that Paivio’s DCT lacked. Soon the idea of a multimodal representational system was also taken up by cognitive psychology and foreign language learning research. At the same time, the expression “multimedia learning” gained importance in these disciplines. Among various explicitly formulated definitions, the major understanding of the term is “the [presentation of information] in more than one mode, such as in pictures and words” as proposed by Mayer (1997: 1). In the following Mayer’s views on multimedia learning/AVC are presented.28

Mayer: Cognitive theory of multimedia learning (CTML)

Being in line with numerous other researchers, Mayer elaborated Paivio’s DCT, in order to explain the effects of pictures for the better understanding of written texts. Basically, Mayer based his theory on the idea that the human brain has separate cognitive systems for the processing of auditory and visual stimuli. These two systems can process only a limited amount of material at a time, thus have limited capacity.29 The reception and processing of stimuli of two modalities result in the construction of two different mental models. This involves the three following cognitive processes.

In the first sub-progress, “selecting”, incoming auditory information results in the creation of a “text base”. Similarly, visual information yields a “picture base”. In the second process, “organizing”, the word base is organized into a coherent system of representations; thus a text-based mental model is created. Likewise, the reception of visual stimuli leads to the construction of a picture-based mental model. In the third process, “integrating”, the two mental models are combined in a one-to-one mapping process. Elements of the text-based model are mapped on elements of the picture based model and vice versa. Moreover,

28 This chapter is based on Mayer: 1997 and Mayer and Moreno: 1998 (http://www.unm.edu/~moreno/PDFS/chi.pdf.)
29 More information on the limited capacity-assumption will be given later in the paper.
relations within the text-based model are transferred to relations within the picture based model. This process requires the particular components to be activated in working memory at the same time.

**Five principles of multimedia learning**

Mayer tested his theory in a number of experiments which resulted in a five-principle-program for his theory of multimedia learning. The most significant principle, the “Multiple Representation Principle”, says that an explanation is better understood if two modes of representation (both auditory and visual stimuli) are used rather than one. Although Mayer primarily dealt with scientific explanations of cause and effect-systems such as the development of lightning storms or the functions of a pump, his principles can also be applied to domains which do not deal with physics. An example for the first principle is the following: Listening to an auditory stimulus (Mayer preferred the term “narration”; an example of a narration is, for instance, a text about how a tire pump works) alone was not as effective as when a corresponding visual stimulus in the form of an “animation”, i.e. moving pictures, was viewed together with the text. In fact, human subjects who were exposed to both types of stimuli generated twice as many useful solutions to subsequent problem solving tasks than those who had to rely on the auditory stimulus alone.

The “Temporal Contiguity Principle” postulates that simultaneous transmission of auditory and visual signals causes better output results than successive transmission of the stimuli. The following experiment outcome serves as an example. Subjects who simultaneously perceived a text and an animation about how a bicycle tire pump works generated 50% more useful answers to subsequent questions than subjects who viewed the animation before or after listening to the text.

The third principle, “Split Attention Principle”, resulted from experiments about the role of visual presentations in audio-visual texts. It is interesting, because it says that subtitles, visualized narration, do not benefit the understanding of the audio-visual text. Mayer argued that the simultaneous processing of two texts of the same modality, the subtitles and the dynamic pictures, can be overloading
for the visual cognitive system. The attention is thereby split between the two sub-processes and thus cannot be used effectively. In other words, films with subtitles are not better understood, but rather hinder comprehension, due to the larger number of processes involved.

The “Individual Differences Principle” constitutes individual differences in regard to the three foregoing principles. Generally, subjects with low prior knowledge about the subject domain but high spatial cognitive abilities tend to show stronger multimedia and contiguity effects than subjects who possess high levels of prior knowledge and low spatial cognitive abilities. Prior knowledge causes an automatic generation of mental pictures while listening to or reading a text; visual contiguous presentation is therefore not needed. Similarly, high cognitive abilities are usually accompanied by the ability to hold visual mental images in visual working memory.

The fifth principle is called “Coherence” and states that only little extraneous information should be presented with audio-visual texts, because every extra detail, whether visual or auditory, causes extra processing. As outlined in the previous principle, the processing of too many stimuli can lead to an overload. As a general rule Mayer stated that the less irrelevant material accompanies the actual text, the better is the overall understanding. For example, subjects who read a picture-supported text about lightning produced 50% more useful solutions in a subsequent test than did subjects whose texts came with additional information.

Mayer stressed that all multimedia texts deliver information to the learner; however, they are not equally successful when it comes to understanding. The five principles of multimedia learning thereby ensure comprehension. Some of the principles of CTML have been criticized, tested against and expanded (e.g. Muller et. al: 2008 and Schnotz and Bannert: 2003). Nevertheless, the key message of CTML, which is that combined processing of auditory and visual information is most effective, is widely accepted. However, like in any other
model, this hypothesis is not universally valid - not every combination of sound and image is beneficial for every learner in every learning situation.\(^{30}\)

Before I turn to the restrictions of CTML, its criticisms and elaborations, I want to discuss one crucial concept of cognitive psychology, which is “working memory”. This construct accounts for the structures and processes, in what is known under the concept of “short-term memory” and is thus crucial for any explanation of AVC. Embedded into the ideas of working memory, “Cognitive Load Theory” will be discussed afterwards. It gives the framework not only for CTML, but also for various later models that deal with multimedia learning and thus AVC.

**Baddeley and Hitch: Working memory\(^{31}\)**

Based on the pioneer work of Baddeley and Hitch in 1974, the concept of “working memory” aims at explaining how short-term memory works. It refers to a brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning, and reasoning (Baddeley 1992: 556).

The model requires the simultaneous processing and storage of information. It is divided into three components, all of which have limited capacity: the “central executive”, a supervisory system, and two slave systems, the “visuo-spatial sketchpad” and the “phonological loop”. The central executive controls all processes in the working memory. Examples of control processes are monitoring and correcting mistakes, retrieving information from long-term memory, coordinating activities in the two slave systems and selecting attention to certain stimuli while ignoring others. The visuo-spatial sketchpad is assumed

\(^{30}\) For example, Muller et al. found in their experiments that the coherence principle did not apply to natural learning environments outside a laboratory setting. They showed that inserting about 50% additional extraneous material did not cause fewer solutions, but resulted in the same learner performance. See Muller et al.: 2008.

Schnottz and Bannert criticize Mayer’s model for not differentiating between the different sign systems texts and pictures use, which results in different forms of representations. Schnottz’s and Bannert’s elaborations of CTML will be discussed later.

\(^{31}\) The outlines of this chapter are based on Baddeley: 1992, Baddeley and Logie: 1999 and Baddeley: 2000.
to retain visual information and is thus considered the visual system. In later publications Baddeley and Logie parted the visuo-spatial sketchpad into two subsystems, the “visual cache”, in which visual information about color and form is stored, and the “inner scribe”, which deals with spatial and movement information. The phonological loop is assumed to be the auditory system dealing with all speech-related information. It is also subdivided into a storage system that temporarily stores memory traces, called “phonological store” and a rehearsal mechanism that revives memory traces, named “articulacy loop”.

The original model, as proposed by Baddeley and Hitch, was very much accepted among cognitive psychologists. However, over the years a number of phenomena emerged for which the model could not give explanations. Thus, a fourth component was added to the model - the “episodic buffer”. It comprises a system that integrates information from a variety of sources, thereby binding information from the slave systems and from long-term memory into an episodic representation.\(^{32}\) The elaborated model is presented as follows:

![Diagram of Working Memory Model](image)

Baddeley 2000: 421

Within this theoretical framework, working memory is exposed to different cognitive loads, depending on the tasks that it deals with. Cognitive load thereby refers to the amount of resources needed to carry out the required task. Considering these ideas, Sweller developed a theory of how cognitive load

\(^{32}\) The model of working memory - whether the original outline from 1974 or the elaborated version from 2000 - is much more complex than presented in the scope of this thesis. Baddeley and his co-workers give plenty of empirical evidence by testing their model in numerous experiments, especially on speech performance of Alzheimer patients.
affects the efficiency of working memory and applied his findings to teaching, particularly to the aspect of instruction-giving.

**Sweller’s cognitive load theory**

The theory of cognitive load can be traced back to educational and psychological literature from the 1950ies. At that time, short-term memory was already assumed to have limited capacity and be able to process only a limited amount of information at a time. In 1976 Sweller’s proposals concerning problem-solving experiments yielded the “Cognitive Load Theory”, which he refined in the 1990ies.\(^{33}\)

Cognitive load theory is concerned with the ease of how stimuli are processed in working memory. If the amount of information is too much or if the information itself is too complex, working memory becomes “overloaded”. Consequently an impairment of schema representation in long-term memory and an overall lower cognitive performance are caused.\(^{34}\) The cognitive load imposed mainly depends on the number of elements that must be processed simultaneously in working memory. This, in turn, depends on the extent of “element interactivity”. Element interactivity can be defined as the extent to which the elements of a certain task can be learned without studying their relations to other elements. In short, element interactivity refers to how isolated certain things can be learned or not. One example for low-element interactivity and, hence, low cognitive load is the learning of individual words in a foreign language.\(^{35}\)

The task may be difficult because there may be a large number of vocabulary items that require learning. Nevertheless, it does not impose a heavy cognitive load. Each element of the task can be learned without reference to any of the other elements. […] When noninteracting elements can be learned in isolation, intrinsic cognitive load is low because working memory load due to the intrinsic nature of the task is low. Low-element interactivity tasks

---


\(^{34}\) A schema is any mental structure which is produced in response to a stimulus. See for example Paivio 1986: 27.

\(^{35}\) Working memory is limited in terms of new information, whereas “old” information (i.e. which has already been stored in long-term memory) can be brought back to working memory and processed there in large amounts.
allow elements to be learned serially rather than simultaneously. The tasks can be fully understood and learned without holding more than a few elements in working memory at a time. (Sweller 1998: 260)

In contrast to individual words, whole sentences and grammatical constructions show high-element interactivity and thus impose a high cognitive load on working memory.

Working memory can be affected by the nature of the material. The cognitive load is then characterized as “intrinsic cognitive load”. Working memory can also be influenced by the type of presentation of the material or by the subsequent task, which is called “extraneous cognitive load”. An example of intrinsic cognitive load is given in the above quote, extrinsic cognitive load is described as follows:

[…] instruction includes multiple sources of information such as a combination of mutually referring diagrams and text. In order to understand the diagram or the text, it may be necessary to mentally integrate them. Such mental integration likely imposes a heavy, extraneous cognitive load. The load is extraneous because it is caused entirely by the format of the instruction rather than by the intrinsic characteristics of the material. (Sweller 1998: 263)

In a pedagogical framework, the differentiation between intrinsic and extraneous cognitive load is important insofar as that the first cannot be altered by “instructional intervention” (Sweller 1998: 259), whereas the latter is considered unnecessary and can be reduced by the alternation of instructions or task requirements through the teacher.

A combination of a high intrinsic and high extraneous cognitive load might be problematic to the learner because it could cause a major overload in the brain and thus cause learning problems. In order to avoid this, one of the two loads has to be reduced. One way is to change the material itself to cut the cognitive load. A second way is to reduce the extraneous cognitive load by altering the design instructions. Whereas both high intrinsic and high extraneous cognitive load are regarded as fatal to learning, low intrinsic and high extraneous cognitive load are considered less harmful, because the total cognitive load is
not high and stays within working memory limits. The same is true for the combination of high intrinsic and low extrinsic cognitive load.

At this point I want to give a practical example to explicate Sweller’s theory in more detail. Consider AHS first graders, who have just started learning English as their first foreign language. These students are now exposed to a short sequence of an English-speaking film, which has rich and complex language and operates on a quite advanced cognitive level. Such a film sequence could be the *Unbirthday Party* scene from Disney’s *Alice in Wonderland*. Clearly, the film sequence, the material itself, would be too difficult for the students, because it not only uses rich and complex native-speaker-like English, but also riddles or nonsense sentences. Moreover, the topic of an unbirthday party might be challenging for some students. Therefore, the film itself would place a high intrinsic cognitive load on working memory. A high extraneous cognitive load would be caused if the teacher was to instruct, “Now watch the film sequence, try to understand it and then write a 100-word summary about what happened”. This instructional design would cause an overload in the students’ working memory. 11-year old beginners of English would not be able to understand the plot of a film scene, which was originally made for grown-up native speakers. Not only would the language be too difficult, but the riddles and stories told in the sequence would exceed the students’ cognitive abilities. Aside from severe comprehension problems, the students would not be able to translate the plot of the sequence into a summary of 100 words. Of course, this example is very much exaggerated; however, the intrinsic cognitive load that the material places on the working memory could be balanced by reducing the extraneous load, hence, changing the instructional design. Keeping the original material, the *Unbirthday Party* scene, a possible appropriate instruction which would not overload the working memory would be: “Watch the film sequence and try to remember at least three words you already know and three more words whose meaning and spelling you want to ask me later”. In this way, the students would not be overextended with the task, although the difficulty of the material stays the same. For the chosen scene, words English beginners could recognize

---

36 I assume that 11-year old students of a lower AHS-secondary school have only had one year of English in primary school and are thus considered beginners.
37 Cf. chapter 17, 39:57- 43:05.
would be very basic ones such as “birthday”, “tea”, “very” and “you”. Of course, the learning objective of the task would be to reinforce already known vocabulary and get to know new words which fit into the overall context, which could be, for example, “How to throw a party”. I would argue that the latter instruction does not overstrain the cognitive abilities of 11-year old students and hence legitimizes the use of this film sequence in first grade.

With the example, I wanted to demonstrate that when dealing with high element interactivity material such as film sequences, it is necessary to reduce the extraneous cognitive load in order to cut the total cognitive load to manageable proportions which are suitable for the age and language level of the students. Firstly I would claim that not the material itself is the source of unmanageable, so-called “hard” tasks, but it is very often the instruction that puts a high cognitive load on working memory. Secondly, almost all intrinsically difficult material can be eased by an appropriate instructional design. This means that there is no difficult material, but there are just too difficult tasks. One major pedagogical conclusion that can be drawn from Sweller’s theory and be applied to all teaching contexts is that cognitive load should always be a pre-eminent consideration when devising instructions.

Sweller mainly developed his ideas for the task of instruction-giving. For my specific research topic, I now want to elaborate his ideas further. Most audio-visual texts cause a high cognitive load due to their inherent difficulty. In other words, they are naturally difficult. A false conclusion would now be not to use them for language beginners, because they seem unmanageable for them. An alternative conclusion, as I would argue, is that films can be used for all learners, even from first grade onwards. This is possible because of Sweller’s differentiation between intrinsic and extrinsic load and the possibility to balance the overall cognitive load in working memory by tasks appropriate to the age and language level of the students. Audio-visual texts themselves are suitable for any learner type. It is all about what you make of the material - as a teacher and an instruction designer.38

38 My claim is based on the assumption that audio-visual texts are not considered just as a way of presenting a specific kind of information, for instance, a grammar point (audio-visual texts would then generate extrinsic cognitive load), but that the two-sensory transmission is regarded
The modality effect

Relating his research to Baddeley’s model of working memory, Sweller argued that audio-visual instructions are superior to purely visual instructions, as they show better results, which is called the “modality effect”. Working memory capacity can be increased by using both visual and auditory working memory (in Baddeley’s words: “visuo-spatial sketchpad” and “phonological loop”, respectively) rather than either memory stream alone.

It has been shown that Sweller’s definition of the modality effect strongly corresponds to what Mayer establishes as the “multiple representation principle”. This overlapping view is not a matter of coincidence, but rather a result of a collaboration of experts from cognitive and educational psychology in the 1990ies. Mayer, Baddeley and Sweller (as well as their respective colleagues such as Chandler, Tindall-Ford, Hitch and others) built on each other’s works by revising, expanding and elaborating them. It is therefore not surprising that if literature from 2000 onwards is reviewed, the same works are being quoted, as they establish a theoretical basis for subsequent models of AVC in the 21st century.

on a holistic level. Only then can audio-visual texts be associated with high inherent difficulty. This difficulty could then be, for instance, “understanding the plot of a complex film scene”, which causes a high intrinsic load on the brain.
AVC in current literature

As reported before, Mayer’s CTML, which has had a major influence on current models of AVC, was not excluded from criticism. Schnotz and his colleague Bannert identified some restrictions of the model. Referring to Mayer’s idea that visual and auditory mental models (“picture-based model” and “text-based model”) are connected in a one-to-one mapping process, they raised the question of how the elements of these models could be linked since they originate from fundamentally different materials - pictures and texts. In the following I want to outline Schnotz’s and Bannert’s criticism on and modifications of CTML.

Schnotz and Bannert: Integrated model of text and picture comprehension (IMTPC)

Schnotz and Bannert doubt that one-to-one connections between visual and auditory representations can be formed. They argue that due to the different sign systems pictures and texts use, their respective representations have to be fundamentally different from each other. Schnotz and Bannert refer to them as “depictive” and “descriptive” representations.39

Before I outline Schnotz’s and Bannert’s differentiation between depictive and descriptive representations, I think it is essential to discuss how they explain the formation of mental representations at all.

How are mental representations formed?

According to Schnotz and Bannert, the processing of (written or spoken) texts includes three different steps. As soon as auditory stimuli enter through the auditory channel, they are organized into a surface representation, called “text-surface representation”. This kind of representation cannot yet be regarded as understanding; however, a text-surface representation allows the recipient to

39 This chapter is based on Schnotz and Bannert: 2003 and Schnotz: 2005.
repeat the linguistic structure (i.e. of what has been read or heard). A text-surface representation of a sentence such as “Some migrant birds fly to the South of Europe for wintering” (2005: 53) elicits an echo of the sentence. Based on the text-surface model, a “propositional representation” is formed through semantic processing, which involves top-down and bottom-up processes. The propositional model expresses the concept that underlies the auditory text, but it is not dependent on its syntax and vocabulary. In the specific case of the above sentence, a propositional representation generates the idea that European migrant birds fly to the Mediterranean area as soon as the winter season starts in Middle and Northern Europe. The underlying proposition is then: “FLY (agent: MIGRANT BIRDS, location: EUROPE, aim: SOUTH, time: SEPTEMBER)” (2005: 53). In a final step, the creation of a “mental model” is triggered by the propositional representation. In the case of European migrant birds, this is a mental map of Europe and the movement of the migrating birds flying southwards. Schnotz and Bannert assume a constant interaction between the mental model and its propositional representation, which controls output processes through schemata-guided processes.

The creation of mental representations is not restricted to texts. In visual comprehension, the recipient first creates a “visual image” through perceptual processing of the visual text. This visual reception, a surface representation similar to that of that of auditory stimuli, is sensory-specific and holds the structural characteristics of the visual text. It is later transformed into a “mental model” as well as a propositional representation. In the case of the mental model, as Schnotz and Bannert outline, schema-guided processes map the features of the actual picture (i.e. graphic entities as well as spatio-visual relations) onto the mental model, this being the result of bottom-up processes.

The process is exemplified via the following example given by Schnotz. If a visual entity such as a graph is being looked at, first, a corresponding surface representational model, a reflection of the graph, is produced. This would then

---

40 According to Mayer understanding occurs when “[…] high element interactivity material can be held simultaneously in working memory” (1998: 261). Another definition says “[it] is the dynamic process of constructing coherent representations and inferences at multiple levels of text […] within the bottleneck of a limited-capacity working memory.” (Britton and Graesser 1996: 350).
be the visual image of the graph. This could be, for instance, vertical bars on a horizontal line in the case of a bar chart or a circular chart divided into sectors, i.e. pie chart. Based on this visual image, a mental model is constructed. In case of the example given above, a chart showing the numbers of different bird species in Middle Europe during the year elicits a corresponding mental model, which contains information such as what bird species leave for the South the latest and where they stay during summer.

A mental model can also be created from “above” by top-down processes, as when the picture is interpreted based on already existing mental models. In the case of the graph “migrant birds”, top-down processing starts at already existing cognitive schemata about migrant birds, which then influences the interpretation of the graph.

In regard to the differences between mental models and the visual images in visual comprehension, Schnotz and Bannert claim that mental models are not sensory-specific; they are constructed by auditory and visual stimuli, just as they are created by kinesthetic and haptic signals. Thus, they are more abstract than their corresponding surface representations. Furthermore, mental models contain more information than surface representations because they include prior knowledge, to which the latter have no access. However, mental models can also hold less information, because unnecessary information and irrelevant details are omitted. For example, a picture of a blue pump may yield a mental model, in which the information “BLUE” is omitted because it is irrelevant to the understanding of how a pump works, whereas additional information about causal relationships coming from prior knowledge may be included in the model, although it is not seen in the picture of the pump.

**Descriptive and depictive representations**

Schnotz and Bannert distinguish between descriptive and depictive representations according to what relation they have with their represented objects. Text-surface representations and propositional representations are descriptive, because they express the represented objects through symbols.
These symbols are signs, which have no similarity with the objects and are only related to them by conventions. Mental models and perceptual representations, on the contrary, are considered depictive due to their inherent structural characteristics, through which they are associated with the represented objects.

Based on the differentiation between descriptive and depictive representations, Schnotz and Bannert established the “Integrated Model of Text and Picture Comprehension” (“IMTPC”):

![Diagram of IMTPC model](image)

Fig. 2. An integrated model of text and picture comprehension.

Schnotz and Bannert 2003:145

The left side of the illustration (including the external auditory text, the internal surface representation and the propositional representation) comprises the
descriptive branch of representations. The right side (including the external visual text, the internal surface representation and the mental model) reflects the sum of depictive representations. The interaction between the descriptive representational models is based on symbol processing, the relations between depictive representations follow analog structure mapping due to structural correspondences.

Schnotz and Bannert identify the type “mental model” as being on the depictive side. This means that all mental models, whether they have been created from visual or auditory stimuli, are always internal depictive representations, because “they have inherent structural features in common with the depicted object” (2003: 147). The comprehension of auditory and thus descriptive texts, which is done in a three-step process (first: text surface representation, second: propositional representation, third: mental model), actually involves a transformation from descriptive into depictive. In contrast to the auditory comprehension, the transformation on the visual side of the model comprises only two steps (first: visual image, second: mental model and propositional representation). Although Schnotz and Bannert do not explicitly argue that auditory comprehension might take longer than visual comprehension due to the longer stream of comprehension of auditory stimuli, I would claim that this is a possible conclusion of the model.

Another interesting aspect of the model is that interactions between text-surface representations (of auditory stimuli) and mental models are possible, as well as interactions between the surface representations of visual stimuli and propositional representations. In this complex system of processes, in which all different types of representations complement each other, one-to-one mappings do not exist. Visual stimuli can yield to both depictive and descriptive representations. Likewise, auditory stimuli may result in depictive as well as descriptive representations.

As argued by Schnotz and Bannert, the processing of auditory and visual stimuli results in the production of a single mental model that contains information from both auditory and visual bottom-up processes as well as top-down processes.
This mental model is in constant interaction with both types of surface representations, the propositional representation of auditory stimuli and already existing cognitive schemata stored in the human brain.

Criticisms

Schnotz’s and Bannert’s IMTPC has both strengths and shortcomings. One of the latter is that the authors do not have empirical proof for their most important argument (which should distinguish their model from all others), the interactions between the different representations. They just assume that these exist. Another point of criticism concerns the left side of IMTPC, which explains auditory comprehension. Schnotz’s and Bannert’s terminology is very similar to that of various models of reading comprehension, which leads to the assumption that they might have taken some ideas from reading comprehension into IMTPC. This claim would also be supported by the fact that Schnotz and Bannert mainly regard written texts as the source of auditory stimuli, which, of course, are read rather than heard. Also, they mainly focus on investigating the question of how the process of mental model construction is influenced by adding pictures (visual stimuli) to a text (auditory stimuli). Thus, visual comprehension is seen as “helping” the understanding of written texts, which is inconsistent with the idea of AVC, which considers both auditory and visual comprehension equally important. These are the downsides of the model.

A major strength of IMTPC is the idea that although the simultaneous transmission of auditory and visual stimuli leads to the construction of different representations, they result in one complex mental model, which is not specific to either modality, but combines information constructed from both types of stimuli.

Some final thoughts on AVC in research

After careful search of more recent models, I found IMTPC to be the most up-to-date model which explains AVC. In fact, studies on the processing of simultaneous auditory and visual stimuli according to my definition of AVC are
relatively few in current literature. What is available is a large number of various loose empirical studies, which investigate individual aspects of AVC, such as the “synchrony and asynchrony discrimination” of audio-visual signals or the timing of audio-visual interactions during reaction time tasks. In most of these and other studies, visual comprehension is seen as having a supporting function for auditory (written text) comprehension. Also, AVC is very often regarded as the perceptual processing of facial and vocal information in the process of speech comprehension, which is not in line with my definition of AVC. Furthermore, I noticed that research has very much moved into the field of neuroscience, measuring electrophysiological responses of the brain (called “event-related potentials studies” or “ERP-studies”), which I do not want to consider here, because they are outside my area of expertise.\footnote{A very recent example of such interdisciplinary research on AVC is Liu: 2009.}

One major problem that I encountered during my search of models of AVC was the many understandings of the concept of AVC, among which I identified the following:

- speech processing (the processing of physical aspects such as facial gestures and lip movements and auditory aspects such as vocal utterances)
- the comprehension of illustrated reading texts
- the processing of static visual texts (pictures) with corresponding sounds and
- the comprehension of simultaneous dynamic pictures and sounds (narration),

the latter being the only definition which explains the transmission of signals by film.

Another problem was that AVC comprises a field of research which lacks empirical evidence. This becomes clear if the most progressive model, Schnotz’s and Bannert’s IMTPC, is considered, which has not been empirically proved yet. A future step is to test IMTPC and, if necessary, develop more accurate models in order to understand how AVC works, along with clear distinctions between the different concepts of AVC.
As I see AVC

In this first part of my thesis, I have given an overview of AVC research from its beginnings in the 1970ies until the year 2010. From the various models and theories encountered, I want to take Schnotz’s and Bannert’s IMPTC as a basis for my personal understanding of AVC. Additionally, from Mayer’s and Sweller’s theory, I want to take the idea that auditory and visual stimuli yield the best results in terms of processing efficiency when they are presented together as a simultaneous transmission of both stimuli, rather than presented in isolation (called “multiple representational principle” and “modality effect” respectively). This claim is also widely supported by recent studies from neuroscience and experimental psychology.  

From Sweller, and hence Baddeley and Hitch, who first studied the cognitive load on working memory, I want to take the idea that, when dealing with films in a pedagogical context, the cognitive load which is placed on the working memory by the given material and the required task should always be an important preliminary consideration - as well as Sweller’s distinction between intrinsic and extrinsic load, which allows the teacher to balance the cognitive load on working memory through appropriate tasks with regard to age and language level.

Considering these ideas as a basis for a theoretical knowledge of AVC, I now want to leave theory and move into practice, in which I will investigate how AVC is achieved in an educational context and how - with its help - other learning objectives can be reached. Before I can do this, I need to inspect the framework for any formal teaching activity, the curriculum.

---

42 See, for example, Molhom et.al.: 2002.
2. AVC in Austrian curricula for foreign languages

In this part of my thesis, I want to investigate pedagogic implications of AVC in Austrian AHS-schools, which could be done in two possible ways. Firstly, I could enter real classrooms and observe in what forms AVC is taught. However, this approach, a field study, would go beyond the scope of my paper, because it requests intensive procedures and analyses. The second possible approach suggests surveying the theoretical background of all teaching - the curriculum. For foreign languages such as English, two curricula exist, one for upper and one for lower secondary school. In the following, I will examine both of them with regard to the following questions: Firstly, is AVC an established skill next to the other communicative skills? Secondly, if AVC is regarded a skill worth pursuing, what correlation is there between AVC and the different educational levels ranging from first to eighth grade in AHS secondary school?

Curriculum for foreign languages in lower AHS-secondary school

The overall aim of foreign language teaching, as stated in the curriculum for lower secondary school, is the development of communicative competences through the teaching of the following skills: listening, reading, spoken interaction, spoken production and writing. A proper calculation of this list yields five skills, which are developed from the traditional classification of four skills (listening, reading, speaking and writing). The questions as to whether it is legitimate to divide speaking into two individual skills, which are both connected to the production of auditory units of meaning and whether spoken interaction should be established as a skill in its own right, even though it is a combination of various skills such as listening, speaking, social networking and the communication of non-auditory signs, are not answerable within the scope of this thesis. Despite the risk of being disregardful towards existing teaching

43 For the curriculum for foreign languages in lower secondary school see: http://www.bmukk.gv.at/medienpool/782/ahs8.pdf
For the curriculum for foreign languages in upper secondary school see: http://www.bmukk.gv.at/medienpool/11854/lebendefremdsprache_ost_neu0.pdf
44 The terminology has been taken from the European Framework of Reference for Languages (CEFR). See http://www.coe.int/T/DG4/Portfolio/documents/Framework_EN.pdf.
expertise, I want to stick to the traditional classification of four communicative skills, because I think that it constitutes a more logical system.

The question that concerns me to a much greater extent, regards the completeness of the curriculum. I asked myself if the pool of skills as presented by the curriculum is complete, inasmuch as it covers all necessary skills for L2 communication which includes AVC. I have to answer the question negatively. The curriculum ignores AVC by not officially establishing it as a skill next to the other skills; the German term for AVC, “Hör-Seh-Verstehen”, is stated nowhere. Even though AVC is not an official skill, the concept of the comprehension of audio-visual signals might be indirectly expressed, e.g. through another term. In foreign language learning research and other fields such as cognitive psychology, the expressions “viewing comprehension”, “multimedia learning” or “auditory picture comprehension” are often used as synonyms for AVC. However, these terms are not stated in the curriculum either. Looking for other expressions which might indirectly refer to AVC, the word “multisensorisch” sticks out. The term appears in the chapter “Didaktische Grundsätze” under the headline “Ganzheitlich-kreatives Lernen” in the following context:

Der Einsatz von spielerischen und musischen Elementen und ganzheitlich-kreativen Methoden ist auch im Fremdsprachenunterricht notwendig […] Multisensorisches Lernen vermag die Aufnahmebereitschaft, Erinnerungsleistung und Motivation zu aktivieren und bringt daher vielschichtigen lernpsychologischen Gewinn.

The use of the term “multisensorisch” can be interpreted in the light of its umbrella term “ganzheitlich-kreatives Lernen”, assuming that the multiple senses involved in the learning process constitute a whole. Holistic models suggest a cooperation of the different senses; multi-sensory learning might therefore argue for AVC. However, the text passage of the curriculum does not suggest that AVC should be regarded as a teaching aim, but rather as a technique to motivate students, boost their memory and increase their receptiveness. In this regard, multi-sensory learning does not refer to AVC as a skill, but promotes it as a teaching technique. Another phrase in the curriculum that might refer to AVC is “die Nutzung von audiovisuellen Medien und neuen
Technologien”, which is found under the heading “Förderung authentischer Begegnungen”. As the heading already suggests, audio-visual media are not argued for in favor of AVC as a skill, but for the sake of authenticity. They can provide opportunities to get in touch with authentic English, for example, through the watching of TV news: “die Nutzung von audiovisuellen Medien und neuen Technologien wie E-Mail und Internet sind im Sinne möglichst großer Authentizität zu empfehlen”.

The didactic principles described by the curriculum are based on the guidelines of the “Common European Framework of Reference for Languages” (henceforth “CEFR”).45 This framework assigns the communicative skills (listening, reading, spoken interaction, spoken production and writing) to various levels of competences, called “Common Reference Levels”, ranging from A1 to C2-level. The result is a detailed description of can-do’s at six different competence levels for each skill. The Austrian curriculum for second languages in lower secondary school basically adopts the prescribed can-do’s for levels A1, A2 and B1, which are then considered the minimum achievement for students in lower secondary school. Thereby, only slight differences in the can-do description between the CEFR and the curriculum can be detected: In the CEFR the information is presented in the format of a self-assessment sheet, with which students can check their individual can-do’s at the various levels. For example, for “listening” on A1-level the CEFR prescribes:

I can recognise familiar words and very basic phrases concerning myself, my family and immediate concrete surroundings when people speak slowly and clearly.

This personal connection by the use of the pronoun “I” is not implemented by the Austrian curriculum, which uses the phrase “Schülerinnen und Schüler” instead.

Under “Lehrstoff/Kernbereich” the different can-do’s for each competence from level A1 to B1 are prescribed. As already pinpointed, AVC is not explicitly established as a skill; however, different hints can be detected which allude to

---

AVC as a skill, and in fact AVC is indirectly referred to. The following evidence appears in the section “listening” at B1-level:

Sie [Die Schülerinnen und Schüler] können vielen Radio- und Fernsehsendungen über aktuelle Ereignisse und über Themen aus ihrem (Berufs- und) Interessengebiet die Hauptinformationen entnehmen, wenn relativ langsam und deutlich gesprochen wird.

Understanding the main points of TV programs inevitably requires AVC, as audio-visual texts can only be comprehended through the use of AVC. In order to comply with the requirement to understand TV programs, I would argue that AVC must become a teaching goal.

I now want to turn to the discussion of what school level the can-do description of B1-listening is actually prescribed for. At the end of third and fourth grade, B1 can-do’s come into effect. However, they are only partially requested.46 The curriculum says:

Die Schülerinnen und Schüler können die Hauptpunkte verstehen, wenn klare Standardsprache verwendet wird und wenn es um vertraute Dinge aus Arbeit, Schule, Freizeit usw. geht.

The above quote shows that the understanding of TV programs is omitted for third and fourth grade. The overall assignment of skills to secondary school reflects one major idea: Students from first to fourth grade are not supposed to understand audio-visual texts. This means that the curriculum for lower secondary school prescribes not a single lesson which uses film and that students from first to fourth grade are not exposed to any audio-visual teaching material - at least not if teachers scrupulously follow the standards proposed by the curriculum. Teaching AVC thus becomes the teacher’s choice, but is not a standard set up by teaching experts - neither by those in Austria nor in Brussels.

46 The can-do description of B1 listening as defined by the curriculum is not wholly applied for third and fourth grade (it comes into effect in the curriculum for upper secondary school, as we will see later).
This conclusion is even more dramatic with regard to the fact that communication in the real world involves a great deal of AVC. Considering communication as the goal of foreign language teaching, as is done by the “Communicative Language Teaching” approach, the prescriptions of the curriculum inevitably lead to communicative failure in the real world because students have never learned how to deal with audio-visual texts.

Curriculum for foreign languages in upper AHS-secondary school

The curriculum for upper secondary school does not explicitly refer to AVC as a communicative skill either. It neither uses the term AVC itself, nor expresses it through other words such as “viewing comprehension”. Interestingly, however, the authors of the curriculum are not consistent in their discussion of the skills. For example, at the very beginning of the curriculum, they mention four communicative skills (listening, reading, speaking, and writing). Later they name five skills (listening, reading, spoken interaction, spoken production and writing).

Similar to the curriculum for lower secondary level, the curriculum for upper secondary level implicitly declares AVC a skill through hidden hints. One of these can be found in the section “Didaktische Grundsätze” under the heading “Vielfalt an Lehrmethoden, Arbeitsformen und Lernstrategien”, which declares “eine breite Streuung an schülerzentrierten, prozess- und produktoorientierten Lehrmethoden, Arbeitsformen und Lernstrategien” as crucial. The phrase can refer to a number of teaching techniques, among which I would also count the use of audio-visual texts, because they fulfill the above mentioned aspects. If so, they are considered a technique used to develop certain other skills, but not a skill of its own, which is confirmed by the subsequent text passage: “[die Vielfalt...] ist sowohl dem Fremdsprachenerwerb als auch der Entwicklung dynamischer Fähigkeiten (Schlüsselkompetenzen) dienlich”

The next paragraph contains a more explicit reference to AVC:

---

Im Rahmen der Lehrmethoden und Arbeitsformen sind verschiedene Wahrnehmungs- und Verarbeitungskanäle zu nutzen und entsprechend vielfältige Angebote an Lernstrategien in den Unterricht zu integrieren.

The phrase “verschiedene Wahrnehmungs- und Verarbeitungskanäle” could be interpreted as a clear promotion of AVC, as it considers the receiving and processing of stimuli of two modalities. The whole quote, however, rather refers to the heterogeneity of learner groups, which requires different techniques for different learner types. For example, visual learners profit more from a written text, whereas other types learn better through listening tasks or tasks that involve some physical effort. Audio-visual texts can thereby please two types of learners, visual and auditory ones. Again, AVC is not considered as a skill, but rather as a teaching technique to please different learner types.

Another evidence for AVC is found in the paragraph “Erwerb linguistischer Kompetenzen”, in which the training of linguistic aspects such as sound recognition, pronunciation and intonation is prescribed. Auditory texts such as radio shows and tape recordings are a good source to develop these linguistic aspects. The same is true for the audio-visual text. Considering AVC this way, however, it is seen as a component or part-skill of speaking and listening, but it is not promoted for its own sake.

The paragraph “Erwerb sozio-linguistischer Kompetenzen” suggests that different national varieties and registers of the foreign language should be taught through listening tasks. I would claim that film can be an adequate source for teaching both the different varieties of a language and its various registers, because it provides authentic language, however, it is not considered in regard to the development of AVC.48

Digging deeper into the curriculum and becoming something of a detective, I would claim that the term ”media” could also serve as evidence which points at AVC. The curriculum for upper secondary level deals with “media” to a much greater extent than the curriculum for lower secondary level. For example, in the

48 In chapter 3, I will reflect more deeply on the authenticity of presentation film claims.
paragraph “Vielfältige Themenbereiche und Textsorten” media are regarded as an appropriate topical framework for the teaching of vocabulary (“Zur Erlangung eines möglichst umfassenden lexikalischen Repertoires sind verschiedenste Themenbereiche zu bearbeiten (wie z.B. ... [die] Rolle der Medien”). Under “Länder und Kulturen” the curriculum prescribes: „[...] für grundlegende Einblicke in Gesellschaft, Zivilisation, Politik, Medien [...] des betreffenden Sprachraumes [ist] zu sorgen”, which refers to the teaching of cultural aspects of the foreign language. However, the development of AVC cannot be inferred from the use of the term “media”.

I found the most evident reference to AVC I found in the can-do description of B1 and B2 listening. Can-do’s at level B2 are expressed as follows (B1 has already been presented before):

Die Schülerinnen und Schüler können längere Redebeiträge und Vorträge verstehen und auch komplexer Argumentation folgen, wenn ihnen das Thema einigermaßen vertraut ist. Sie können im Fernsehen die meisten Nachrichtensendungen und aktuellen Reportagen verstehen. Sie können die meisten Spielfilme verstehen, sofern Standardsprache gesprochen wird.

Understanding TV news and films in the standard variety of the foreign language is listed under B2 listening. As this request can only be fulfilled by AVC, it has to become a skill.

At the end of fifth and sixth grade, students should be able to employ B1 listening skills. “[Die] gleichzeitig[e] Erweiterung und Vertiefung der kommunikativen Situationen, Themenbereiche und Textsorten” is additionally prescribed for sixth graders. This somewhat vague phrase might function to differentiate sixth grade skills from fifth grade skills. From seventh grade onwards, students are supposed to be at level B2.
Summary: The role of AVC in both curricula

Both curricula have been examined with reference to the question in what way they deal with AVC. The results drawn from a close reading of the two curricula are quite similar.

In both of my examinations, I aimed at finding AVC in the listing of the communicative skills, hoping that it would be defined as such. However, AVC is not considered a skill; the term is found nowhere in the curricula. Also, AVC is not explicitly expressed in the CEFR, which serves as a basis for the Austrian curricula for foreign languages. I also searched for potential references to AVC in the form of synonyms, paraphrases and some other expressions (e.g. “viewing comprehension”). It turned out that these were not existent either. Thus, the curricula do not officially regard AVC as a skill.

Despite this “setback” I still believed that AVC was part of a serious pedagogic discussion. I therefore concluded that the curricula provide some hints that point to AVC. In both texts, I found plenty of phrases which looked very promising for an interpretation in favor of AVC. However, most of them did not suggest it as a skill, but regard it as a teaching technique or authentic teaching source. AVC is therefore not considered a teaching goal in its own right, but rather an instrument and intermediate step to achieve other goals.

AVC is considered a skill at only two instances. In the can-do description of B1 and B2 listening (which are based on the guidelines set by the CEFR), the understanding of audio-visual texts such as films and videos is demanded. This is only possible through AVC. Although AVC is not explicitly stated and recognized as a skill, the formulations of B1 and B2 listening implicitly request it.

The can-do description of B1 and B2 listening only applies to upper secondary school: B1 and B2 can-do’s have to be fulfilled by fifth/sixth and seventh/eighth graders, respectively. This means that in lower secondary school - if the curriculum is closely followed - students do not work with audio-visual texts at all.
What do the findings mean for an overall interpretation of the curricula?

AVC is inherently claimed for through the can-do descriptions for B1 and B2 listening skills in the curriculum for upper secondary school. However, I want to note that it is the result of interpretative work and not an explicitly stated claim. Bearing this in mind, I would draw the following disturbing picture of the curriculum for upper secondary school: the pool of skills that is suggested falls short, because it cannot fulfill the can-do’s postulated by the very same framework. If the comprehension of TV programs and films is requested, AVC has to be regarded as a skill crucial to the fulfillment of the request. In this sense, the curriculum is incomplete because it lacks AVC as a skill.

The curriculum for lower secondary school does not consider AVC at all. Here, the problem goes even deeper, inasmuch as the dealing with the audio-visual text, and thus AVC, is not prescribed, because neither B1 nor B2 skills are required from first to fourth grade. Here, a rethinking of the can-do descriptions has to take place - not only because students might not be able to fulfill B1 and B2 listening skills in upper secondary school without having already gained some experience with AVC in lower secondary school, but also because students would miss out on some meaningful activities which involve AVC.

The analyses have demonstrated that both curricula lack AVC as a skill. I claim that this gap has to be filled by integrating AVC into the pool of skills and by distinguishing it from listening. This has to be done by professional curriculum designers. However, I am not sure whether this is a clear objective for the near future, in regard to the fact that AVC is not considered a teaching standard by European teaching experts. The CEFR, which serves as the basis for all language teaching frameworks in the European Union, dedicates only half a page to AVC (out of 264 pages) and - what is even more alarming - does not define AVC as a communicative skill (see below). Nevertheless, I hope that my discussion has pointed out the gap in the curriculum as well as in the CEFR, which might have some positive changes on the curricula for foreign languages in Austrian upper and lower AHS-secondary school.

Why is AVC not an officially established skill?

One reason for the non-consideration of AVC as a skill might be ignorance of the relation between the watching of film and AVC. It could be possible that the authors of both the CEFR and the Austrian curricula do not know that the comprehension of audio-visual texts takes more than listening. This assumption is supported by the fact that AVC is discussed under the heading “listening”. However, evidence is found which argues that the concept of AVC was not unknown - at least not to the authors of CEFR, as illustrated by the following text passage:

In audio-visual reception the user simultaneously receives an auditory and a visual input. Such activities include following a text as it is read aloud, watching TV, video or a film with subtitles and using new technologies (multi-media, CD ROM, etc.).

Apart from the fact that the above-quoted lines constitute the only (!) theoretical and research-based information about AVC, I want to stress that my previously stated assumption on the missing link between the audio-visual text and AVC is not very likely, because the authors acknowledge the processing of multi-sensory signals: “The user simultaneously receives an auditory and visual input”. Nevertheless, a closer analysis of the quote accentuates the phrase “with subtitles” in “watching TV, video or a film with subtitles”, which suggests that film should be watched with subtitles. Which subtitles, L1 or L2, are referred to here? Why are they recommended in the first place? Do the authors believe that the content of audio-visual texts cannot be conveyed without the support of subtitles - not even with intermediate learners at level B1? If that was really the reason for the suggestion of subtitles (whether in L1 or L2), I would claim that students’ skills and competences in dealing with film are very much underestimated.

The mentioning of subtitles in the same breath as audio-visual perception very much influences my interpretation of the word “visual”. Being stated in this context, “visual” might not refer to the dynamic pictures of the audio-visual text, but rather to its subtitles. This assumption seems quite absurd; however, it becomes reasonable if the previously stated text is considered, “[audio-visual]
activities include following a text as when it is read aloud”. I am shocked by the information given by the text. It equates the act of watching someone read out a text with the comprehension of film, which - as I would argue - is an illegitimate comparison. It shows that AVC and its very essence, the simultaneous transmission of visual and auditory signals, is either not rightly understood or ignored by the authors of the CEFR.

The short analysis has shown that there are plausible and very possible explanations for the missing link between the requirement to understand audio-visual texts and the necessity of AVC as a skill. With it, I intended to show that the CEFR and the Austrian curricula are incomplete, because they disregard AVC as a skill. Moreover, these frameworks reflect that AVC has not yet achieved recognition in foreign language teaching.
3. AVC in teaching practice

In this part of my thesis, I want to illustrate practical implications of AVC in foreign language teaching. When AVC is applied in an educational context, it can be treated in two ways: firstly, it can be regarded as a learning objective in itself, as when the development of AVC as a skill (next to the other communicative skills) is the aim of the teaching sequence. Secondly, AVC can also be considered an instrument by which means other objectives can be achieved. In this case AVC is still regarded as a skill; however, it is used in order to fulfill the required task and reach a superordinate learning goal.

In my discussion and illustration of the use of AVC in foreign language teaching, I incorporate both treatments of AVC - as a learning objective and an instrument/sub-goal of another objective. They are presented in fictional, nonetheless possible, teaching contexts, together with lesson plans and teaching materials. These didactic applications of AVC are embedded into the “Communicative Language Teaching” approach (henceforth: “CLT”) and its principles. They are also based on the theoretical model of AVC as outlined at the end of chapter 1. One of its ideas is that AVC is a skill which can be learned and used by students of any age or language level, since the cognitive load of audio-visual texts can be balanced by tasks appropriate to the age and language level of the students through the teacher. This I will exemplify in my first didactic application of AVC - a teaching session, in which beginners work with the supposedly inappropriate film “Alice in Wonderland” (as it is a film made for native speakers and not for language beginners) and in which AVC is seen as an instrument. By designing tasks and instructions in order to reduce the cognitive load caused by the intrinsically difficult material, I want to demonstrate that - contrary to what is prescribed in the Austrian curricula and the CEFR - dealing with film and the development of AVC can and should be incorporated from the very beginning of language learning.

Another message I want to convey concerns the wrong assumption that only films that correspond with the learners’ language level should be dealt with; hence films can only be used in upper secondary school because only there do
most students appear to be able to handle the “difficult” material. However, with the chosen teaching design, I will show that almost every film - in regard to the difficulty of its language - is suitable teaching material, provided that the task and the instruction design are appropriate for the learners. The very same film AIW will be discussed together with its source, the novel “Alice’s Adventures in Wonderland” (henceforth: “AAIW”) by Lewis Carroll, in order to illustrate how AVC as a learning goal can be achieved with a group of more advanced learners. This second teaching outline will be the only design which aims at reaching AVC as a goal at the end of the lesson. A third example of AVC in foreign language teaching demonstrates how historical learning is achieved. By analyzing scenes from the film “Cry Freedom” (henceforth “CF”), students of an eighth grade will learn about a specific aspect of South African history: the different types of power relations between black and white South Africans in the 1970ies, established as positions of inferiority/superiority. The lesson also requires the discussion of film techniques, such as camera angles, sound and lighting. Thus, AVC is not only employed as a skill in order to comprehend audio-visual texts, but also considered on a meta-level as audio-visual techniques are discussed in regard to their effects on the power relations portrayed in the film. A fourth teaching design, which incorporates a short video clip from the movie “Notting Hill”, exemplifies how the communicative skill “speaking” can be trained. The clip, being used as a discussion prompt and teaching material for succeeding speaking activities, is comprehended through AVC, which is considered a tool and a subgoal of the lesson, as it facilitates the achievement of the superordinate aim “Getting the students to talk”.

Before I present my ideas of how AVC is used in foreign language teaching, I want to discuss two more points. Firstly, I want to survey the various reasons - apart from the development of AVC - why audio-visual texts constitute a great teaching material and thus should be regarded as crucial in foreign language teaching. I want to present these reasons, because I think that every serious pedagogic discussion of AVC and film has to include all its benefits, especially with regard to the fact that AVC as a skill and the various potentials of the audio-visual text have not been fully acknowledged by the Austrian curricula and their authors yet. As a second point, I want to discuss the concept of AVC
once more. This time it is considered from a more pedagogic viewpoint, which aims at explaining what AVC means in a foreign language teaching context.

**Why should films be used in foreign language teaching?**

**Authentic content? - Authentic discourse!**

A lot of texts present authentic content inasmuch as they thematically reflect real life with real people. For the audio-visual text, this means that it shows content which could happen in the viewers’ own lives. However, most audio-visual texts do not provide such an authentic representation of life. AIW, for example, is not authentic, as its content, i.e. the storyline, is not taken from real life. What is authentic about AIW and other fictional texts is their discourse. In order to outline the authentic discourse of films, the concept of discourse needs to be clarified.

In a very narrow sense, discourse means communication. Based on the significant works of Michael Foucault from the 1960ies and 70ies, the concept was later extended by a social component: language acquires meaning only through its social usage. At the same time the users of that language are given an identity by being in the discourse. Discourse thus creates social interactional processes between the users of the language. In the words of the sociolinguist and discourse analyst Gee discourse is “[an] identity tool kit replete with socially shared ways of acting, talking and believing” (quoted in Hicks 1996: 53). The discourse of a group of language users, which includes speech as well as behavior, expresses the belief system, customs and ideologies of that certain group, in short, its identity. If this idea is now transferred to audio-visual texts, it means that the presented communication between film characters, their discourse, reflects their similar (or different) identities. In short, film presents authentic discourse with authentic communication strategies, facial expressions,

---

50 This chapter is drawn from Lowe: 2008 and Stempleski: 1993.
51 Please see also my discussion on page 4.
language codes and other signals which are reflections of the identity of the discourse group.\textsuperscript{52}

The authenticity of discourse is not restricted to the communication among film characters within the storyline of the audio-visual text. It is also existent between the audio-visual text and its audience. Language is thereby a social construct, whose meaning is created relationally between the text and the recipient/viewer. Opposed to those (audio-visual) texts which are specially created for language learners (the discourse between the text and its recipient is then a “learner discourse”), films such as AIW are authentic as they are intended for native speakers who have no intention of profiting from the film language-wise, but they deal with it for other than learning reasons.

**Authentic language use**

As discussed before, audio-visual texts constitute a source of authentic discourse, i.e. language use, between film characters. In this sense, audio-visual texts provide discourse which presents

\[\ldots\text{ungraded and unsimplified English, spoken at normal pace and in typical accents. Such language is also real in that it is current; that is, makes use of idioms and expressions common in contemporary English-speaking environments. (Stempleski 1993: 9)}\]

Although some science fiction films, such as *Lord of the Rings*, with imaginary, non-existent languages represent an exception to the idea of “real language” reflected in film, most audio-visual texts provide authentic discourse and are thus an appropriate teaching material in CLT, in which authentic and meaningful communication is the highest goal.\textsuperscript{53}

In regard to the argument that film offers language in typical accents (see the above quote); I need to note that they provide an unprecedented easy and

\textsuperscript{52} Cf. Hicks: 1996. For further elaboration on the authenticity of discourse and language use see the subchapter below.

\textsuperscript{53} Although science fiction movies with fictional languages are not suited as sources of authentic language use, they can be used for creative fun activities in which, for example, students decipher the language or create their own secret language.
comfortable access to different language varieties and accents. I cannot think of any other medium or material, which is as easily available as films (where should teachers get, for example, a non-didactic, authentic tape of a Texan speaker?). At the same time, the different accents are visualized in audio-visual presentations of the medium, which eases understanding as the multiple representation principle applies.

Using film in foreign language teaching has another positive consequence. Generally, most learners figure out quickly that teaching material is inauthentic, in that it is specially created for their needs (film as “learner discourse”). Students as authentic L1 users might respond negatively to the imposed restraint and be demotivated by the material. Films - as they are not specially made for L2 learners - can cause the exact opposite effect: films can be regarded as a challenge by students, because they are not simplified and thus motivate them.

**Audio-visual texts as motivators**

The authenticity of film is called “the real thing” by Stempleski. She argues that when students realize that the real thing, i.e. real English made for native speakers, is not necessarily beyond their comprehension, but can be achieved with some effort; it has a motivating effect not only for the concrete text, but also for all other learning processes. The comprehension of authentic material can lead to the enthusiasm to understand more authentic material and thus the target language itself.

**Audio-visual texts develop the “media skill”**

Most Western cultures are audio-visual environments, in which films are viewed a great deal. It is therefore crucial to have certain skills which help viewers to cope with these texts. There is, firstly, the processing of audio-visual signals, a skill which I have defined as AVC. Secondly, the content of an audio-visual text has to be critically analyzed on a more abstract level of comprehension, in which the text is put in relation to other texts. For this process, another skill is
needed, which I call “media skill”. This skill needs to be developed as any other skill - film-based teaching can thereby turn students into critical and effective viewers who are able to deal with audio-visual texts outside the classroom.

**A window into the culture**

Films constitute a source which allows experiencing the target culture, which is very often a new culture for the learners. These cultural explorations are very special, as they are both visual and auditory. This also means that audio-visual interactions between representatives of that culture can be experienced only with film. It also provides a window into the cultural performances, such as that of customs and values, which would not be possible with other teaching material (see my discussion about the authentic discourse of film). With this exploration, awareness of the target culture and its people is heightened, as well as the students’ awareness of their own cultures.54

**What is AVC from a pedagogic viewpoint?**

Research on the cognitive processes of AVC has proved that a mental construct is created from both auditory and visual signals, which interact on various levels. This mental construct is created not only from the text signals upwards (“bottom-up process”), but also through knowledge in form of schemata which are stored in the brain (“top-down process”). From a pedagogic and didactic viewpoint, this idea can be interpreted insofar as AVC is an active and complex process. Although AVC - together with listening and reading - is a so-called “receptive process” in that it receives information and does not produce it, it is dynamic and thus requires attention and concentration. Furthermore, the process of AVC connects known with unknown information, which needs to be considered when planning and executing tasks involving or aiming at AVC. A preliminary step must therefore be to activate already stored knowledge before viewing the audio-visual text, in order to link it to the new

---

54 At this point I want to clearly state that the various arguments for the use of film in foreign language teaching have been listed for the sake of completeness concerning the advantages of the audio-visual text. The listing is thus a general collection of arguments for the pedagogic use of film in L2 contexts. This does not automatically mean that the films used in the four teaching design follow all these arguments.
information during the process of AVC. Thus, so called “pre-viewing activities” become crucial in every serious and effective teaching design which involves AVC.

Another pedagogic principle of AVC comes from reading and listening comprehension research, which states that recipients only receive and understand what they actually want and/or have to. This means that students and human beings in general cannot and do not comprehend a text to its full extent, but only parts of it or a particular piece of information. A conclusion of this idea must be the realization that an instruction such as “Watch the film scene” cannot work in teaching practice, as students do not know on what exactly they should concentrate, i.e. what they should actually see and hear. Therefore, different “comprehension styles” need to be determined in regard to AVC. From the receptive skills listening and reading, various strategies and technical aids are known, depending on what purposes they serve. Based on those, I have determined the following categorization system for AVC:

1. Selective AVC = A specific piece of information of the text is searched for, such as a number or a certain word.

2. Detailed AVC = A text passage is understood in its very detail.

3. Orientating AVC = Here comprehension aims at answering the question: “What is the text about?”. It gives an impression of the content and message of the text, in short, the gist of the text should be understood.

Another principle I want to determine for the use of AVC in foreign language teaching has already been discussed at various points of the thesis. The level of difficulty of an audio-visual text mainly depends on the given instructions. Thus, there are no unteachable films (in regard to their difficulty), but just inappropriate tasks.

55 Considering reading, the most important strategies are “skimming” and “scanning”. See Nuttall 1996: 48-50 and footnotes 51-2.
The last principle concerns the interdependence of text and recipient. Solmecke says, "Zur Mitteilungsabsicht des Sprechers gehört die Verstehensabsicht des Hörers." (2001: 5). Solmecke relates this idea to the skill of listening; however, he does not exclude its relevance in regard to the other skills. I would claim that any text, also an audio-visual text, has the intention to send a certain message to a recipient or a group of recipients. This I have already defined as authentic discourse between text and viewer. In order to establish successful communication, the recipient needs an intention to deal with the text; otherwise communication fails and the text itself becomes irrelevant. Solmecke argues that the intention to understand a text and its message is either conscious or subconscious; it can exist before or during reception, but in any case it is decisive for the recipient’s dealing with the text and thus AVC.

A didactic consequence of Solmecke’s idea is that it cannot be assumed that students automatically have or develop the intention to understand the text or that the intention is given because the text is used in a learning context. In case the intention does not exist, the teacher has to try to induce it. This can be achieved by offering the prospect of success in form of a good grade and praise or by providing challenging and exciting tasks, the accomplishment of which is based on the wish for self-fulfillment or the interestingness of the task itself. However, there is also the possibility that students might not develop an intention to deal with a text at all - even if they are given various reasons and incentives by the teacher. In this case, the teacher has to accept that not every student is interested in every text at all times.

**Four different teaching designs**

1. **AIW: AVC as an instrument for teaching beginners**

In this first teaching design, I want to exemplify that film can be used in any foreign language teaching context, even with L2 beginners. In the following I will outline the teaching design in regard to the following points (which are applied to the other teaching designs as well):
**Target group:** Who is the target group and what is its general language proficiency level according to the CEFR/Austrian curriculum? How is the target group described by the latter?

**The audio-visual text:** What audio-visual text is used in the teaching design? What is shown - the whole film or just parts of it? What are the difficulties involved with the audio-visual text? Is it, for example, appropriate for the students in regard to their age and proficiency level? Is the text considered intrinsically difficult? How is then the cognitive load balanced out by the subsequent activities and tasks?

**Learning aims:** What are the aims of the teaching sequence? What are the general aims and the specific aims? How do the aims relate to the can-do descriptions of the curriculum?

**The role of AVC:** What role does AVC play in the teaching sequence? Is it film-based teaching (in which AVC is seen as an instrument and its development is thus not the ultimate goal of the teaching)? Or is the teaching sequence AVC-oriented because its general goal is the development of AVC? At what point of the teaching sequence is AVC used and how does it contribute to the achievement of the learning aims?

**General outline of the activities:** At what time and for how long should what students do which activity, with what kind of teaching material? What are the pre-, while- and post-viewing activities? (These many points will be clearly presented as lesson plans found in the appendix.)

For the following teaching design, the questions are answered as follows:

**Target group:** I assume first graders of an Austrian lower AHS-secondary school, who are at the very first stage of learning the foreign language - in this case English. Furthermore, I assume that the students have just got familiar with the alphabet, although they have not fully mastered it yet. According to the CEFR and the curriculum for lower AHS-secondary school, the chosen target
group is regarded as being at level A1 (considered “breakthrough” by the CEFR).

**The audio-visual text:** The students participate in a film-based teaching sequence, in which AVC is seen as an instrument and is thus not the ultimate aim of the teaching sequence. The Disney animation AIW, based on the novel AAIW by Lewis Carroll, is used. The material, the audio-visual text itself, can be considered intrinsically difficult, as it presents English intended for native speakers, which is not adapted for use in foreign language teaching. One particular sequence, the “smoking caterpillar scene” (chapter 14, 30:33 - 31:25 min.) will be used. It will serve as an instrument and medium to achieve the following goals, differentiated into general and specific aims:

**Learning aims:**

**General aim 1:** The students should get familiar with introducing themselves, asking other people their names and talking about other people’s identities.

**Specific aim 1a:** The students should correctly use the sentence “Who are you?” when they ask fellow students their names. Furthermore, they should be able to answer the question “Who are you?” correctly, i.e. “I am____ (the student’s name)”, when they are asked by their fellow students.

**Specific aim 1b:** The students should be able to ask their fellow students about their identity (depending on what letter card they get, e.g. the display of letter D). Additionally, they should also be able to give correct answers about the identity of their letter when they are asked by fellow students.

**Specific aim 1c:** The students should be able to understand the question “Who is ____ (the student’s name)?” correctly and answer it with the phrase “He/she is ____ (the letter on the back of the card)”.

---

59 In the formulation of this and the other learning aims, the phrase “the correct use of “Who are you?” refers to its grammatical, phonetic and semantic accuracy. Concerning one aspect of its phonetic correctness, the right stress of the phrase, it shall not be regarded as there is no right or wrong stress. All three monosyllabic words can be stressed according to the intended meaning. As the students are at the beginning of learning the language, they shall not be confronted with the different possibilities of stressing the phrase. However, when the teacher gives the example of the right use of the phrase, he/she shall put the stress on the last syllable “you”. The focused word accentuates the fact that the main interest is on the person and his/her identity.
General aim 2: The students should practice the letters of the English alphabet in regard to production and perception.

Specific aim 2a: The students should be able to recognize the different letters from the film scene by using AVC.

Specific aim 2b: The students should be able to recognize letters when they are presented orally by another student (oral perception) and be able to note them down (written production).

Specific aim 2c: The students are able to correctly verbalize the information on the work sheet by putting it into a sentence like this: “He/she is letter _____ ”.

The chosen general and specific aims are in line with the prescriptions of the curriculum for lower AHS-secondary school. Under the skill “A1 listening”, which applies to first graders, students should be able to understand easy and familiar words and sentences and can relate them to themselves. Under the skill “A1 speaking” the curriculum requires first graders to use easy phrases and sentences in order to describe other people. I would argue that the previously stated aims are in line with the requirements of the curriculum. In fact, the learning aims demand that first graders are able to orally describe not only other people, but they enable them to talk about themselves (e.g. saying their own names) - which is not specifically requested by the curriculum. The overall topic of the teaching sequence is “asking names and introducing”. This topic is not explicitly mentioned under “Vielfältige Kommunikationssituationen und Themenbereiche” in the curriculum; however, I would argue that it is crucial to any communication and hence a legitimate subject matter in the teaching of foreign languages.

The role of AVC: In this teaching design, AVC is not regarded as the desired overall goal of the teaching sequence, but plays a role in the achievement of the defined general and specific aims. Specific aim 1a says that students should correctly use the phrase “Who are you?” and correctly answer it when they are asked by fellow students. In the constructed teaching design, these specific aims can only be achieved by understanding the film sequence in which the phrase “Who are you?” is presented. (In the film sequence, Alice curiously

encounters the smoking caterpillar after having wandered through the woods. The caterpillar asks Alice her name by using the line "Whooo ... are ... you?", which is visualized as exhalations of smoke in the shapes "O", "R" and "U").\textsuperscript{61} Obviously, the film sequence serves as a medium, which presents the content of an encounter between two different characters. In order to comprehend this audio-visual transmission, AVC is needed; hence AVC is seen as an instrument to comprehend the audio-visual text and based on this comprehension, the specific aims, the understanding and correct use (including the correct answering) of the phrase "Who are you?" is achieved.

It could be argued that other media could also be used to teach the use of the phrase "Who are you?" For instance, the teacher could act a scene out together with one of the students or present the phrase in written form, which could be far less time-consuming and complicated in terms of preparation of the material. However, in both cases the teacher would have to explain the use of the phrase, i.e. in what context it is used. These explanations would be quite boring - especially for such young learners. With the use of the film scene of Alice's encounter with the smoking caterpillar, the explanations are automatically omitted, because the phrase is presented in a clear context. This context - a curious girl meeting a strange, smoking caterpillar - makes clear that the film scene is about the first encounter of two strangers who want to clarify each other's identity.

Another advantage of the audio-visual text over other media such as a purely visual presentation in written form or a teacher's ad-lib performance is that it is well suitable for learners of the age of 10 or 11. This is logical, as the Disney animation AIW is specially made for children (the DVD blurb defines the film as "Kinderfilm"). The scene from AIW presents colorful moving pictures, accompanied by off-sound music and on-sound humming noise (the smoking caterpillar accompanies the background music). Furthermore, AIW is suitable for children, as the main character, Alice, is depicted as a young girl with typical

\textsuperscript{61} The smoking caterpillar stresses all three words equally and - at the same time - unusually long. This might create the impression that it is the universally correct use of the phrase. In order to avoid this, the teacher shall stress the three words equally short, with the main stress on the third syllable (see my discussion before).
child-like qualities such as curiosity, spontaneity and thirst for adventure - character traits which also appeal to children who are the same age as Alice.

Scientific research has shown that AVC yields the best results in terms of the processing of new information and memorization (multiple representation principle). I would therefore claim that in teaching contexts, audio-visual texts stand out inasmuch as they constitute the most efficient way to learn. For the particular case of learning how to ask a name and answer such a question, I would also argue that the chosen teaching design is effective, as the chances that students actually reach the learning goals are very high.

**General outline of the activities:** The teaching sequence is assumed to comprise 50 minutes, as most (English) lessons in Austrian AHS schools last that long. The lesson starts with a pre-viewing activity, in order to meet the principle concerning the interdependence of old and new information. This pre-viewing task activates students’ previous knowledge about the letters of the alphabet, in order to either connect new information to it or to further consolidate existing knowledge. The students have already become familiar with the English alphabet in regard to reception and production; however, apart from recognizing and using individual letters of the alphabet, they cannot form sentences or express themselves otherwise within the foreign language. Also, they have not mastered the alphabet yet and the existence of knowledge gaps cannot be excluded. The teacher elicits individual letters from the alphabet by asking the students what English letters they can remember. In order to avoid chaos, a ball is used, which is passed from one student to the next. Every student should name a letter by saying it out loud and simultaneously drawing it into the air. In this way, knowledge is activated, interest is raised and especially young learners can exhaust themselves a little through kinesthetic participation.

As a next activity, which is a while-viewing activity, the students are instructed to catch all the letters they see and hear in the film scene, which is presented twice. They are instructed to just memorize the letters and not write them down (which would put too much cognitive load on the students). Before the film scene is shown, the teacher briefly explains from where the film scene is taken
and who the characters are. Depending on the type of target group, in regard to the willingness to participate actively, its homogeneity and internal dynamics, the while-viewing activity can also be designed as a group activity or even competition in which the occurrences of the letters are counted. The while-viewing activity involves selective comprehension, as specific information is searched for.

The after-viewing activity is constructed as an open conversation between students and the teacher (henceforth “teacher-student talk”: “TST”). They discuss the types of letters which occur in the scene, which are mostly vowels. An appropriate list consisting of a “vowel” and a “consonant column” is drawn on the blackboard by the teacher. I assume that the concept of vowels and consonants is not unknown to the students - in case they do not know, the teacher has to discuss the difference between vowels and consonants, since it has already been mentioned. The discussion should happen in German.

As a next activity, the film scene is shown again. This time the instructions are the following: “Listen and see. What does the caterpillar (the vocabulary is already known) say to the girl? Try to catch the phrase and memorize it”. This time the film scene is shown just once. Again, the task is designed as selective AVC inasmuch as the search for particular information, the caterpillar’s question, is requested. At the same time, this search requires full attention in order to hear/see every single syllable and to understand every single word. Thus, detailed comprehension is needed in order to do the activity. After having watched the scene, the teacher asks the whole class what they have understood. I expect a chorus saying “Who are you?”, which is spoken at a usual, native speaker-like pace, with no stress, or the stress being on the third syllable. The teacher reacts adequately and says his/her name with the phrase “I am…”. Then the teacher picks a student, approaches him/her and asks again “Who are you?” (The teacher stresses the phrase in a usual, native speaker-like way). The student is expected to answer by using the phrase “I am …”. The teacher then turns around the sequence by introducing him/herself first and asking the student his/her name. The student gives the answer and asks the teacher his/her name. The teacher repeats the same activity with two more
students and then gives the instruction to introduce oneself and ask the follow
student his/her name (this has to be performed five times).

In a subsequent activity, the teacher picks a student, introduces him/herself and
asks for the student’s name. After receiving the answer, the teacher points at
the student’s neighbor and asks, “Who is he/she?”. The teacher helps with the
answer “He/she is ….”. After three demonstrations with female and male
students (in order to use both personal pronouns he and she), the students
should form circles of five people (of course, the desks and chairs are moved
beforehand). They introduce themselves first, ask their left-hand neighbors their
names, wait for the answer and then form the phrase: “He/she is…. (the
student’s name)”. The activity ends with the completion of the circle, i.e. when
the first student of the circle is asked his/her name.

As a second post-viewing activity, the following task is designed. The English
alphabet consisting of 26 letters has been printed on 26 laminated cards,
assuming that the group consists of about the same number of students; (if not,
certain letter cards need to be doubled or students get more than one letter
card). Every student gets a card, but does not show it to the other students.
Also, every student gets a work sheet which looks like this:

<table>
<thead>
<tr>
<th>Who are you???</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valentin: He is _________</td>
</tr>
<tr>
<td>Christina: She is _________</td>
</tr>
<tr>
<td>Emil: He is _________</td>
</tr>
<tr>
<td>Patrick: He is _________  Michael: He is _________</td>
</tr>
<tr>
<td>Gül: She is _________</td>
</tr>
</tbody>
</table>
The students approach their fellow students, introduce themselves and ask what letter they represent by employing the phrase “Who are you?”. The answer is then noted down on the sheet. The activity is assumed to be fun and loud, as well as to a certain degree not fully controllable by the teacher, inasmuch as the students run around the classroom trying to be the first to finish. The teacher thus passes the power of control on to the students, which does not mean that the teacher cannot intervene; however, the task is considered to be student-centered. After having completed the task and moved the chairs and desks back into their previous positions, the work sheet is compared in a TST. The teacher asks, for instance “Who is Patrick?”, a chorus of students should then answer, for instance, “Patrick is letter C”. The teacher also notes down the answers on a copy of the work sheet printed on an overhead transparency. In this way, the teacher makes sure that every student gets the right answers and that the spelling of the letters can be checked.\(^\text{62}\)

The highest goals of teaching, as requested by CLT, are authenticity and meaningfulness of communication. The teaching design aims at the authentic use of language – at least parts of it. In these, authenticity and meaningfulness are achieved, as students express their own and other people’s identities. However, the very last task centers around the question of which English letters are assigned to which students. Students’ names being letters of the English alphabet is not very realistic; however, the task does not claim to call for authenticity - it is designed to be a fun activity, in which the students realize that they are involved in a game with letters. Apart from this, names being expressed as letters are not as inauthentic as it seems at the very first glance. Some stage names consist of individual letters, such as a member of a panel from a German talent show named “Q”. Also, there are nicknames such as J.D. (e.g. for Jonathan Michael) or D.J. (e.g. for Donna Jo) or names of magazines, bars and films (e.g. the magazine “V”, the Viennese bar “Q” or the German movie “M”) that use English letters.

\(^{62}\) The whole teaching outline is put into a time frame of a lesson plan, which is found in the appendix of this thesis.
Another aspect that concerns the authenticity of the teaching sequence is the use of the phrase “Who are you?”. Transferring it into the communicative context of English native speaker grown-ups, the phrase is considered to be rude and inappropriate. This aspect definitely needs to be considered in CLT contexts. In order to do that, the teacher has to explain to the students that the use of “Who are you?” is not always appropriate - at least not in formal settings. Furthermore, the teacher can introduce and practice the usual polite English introduction, which goes like:

A: Hello, I am ____.
B: Hello, I am____.
A: It is nice to meet you, ____.
B: It is nice to meet you, too.

in a subsequent teaching sequence.

An argument, however, that justifies the use of the phrase “Who are you?” in CLT contexts is that the phrase is not necessarily out of place in settings which involve children. For example, a conversation between 10-year-olds could certainly consist of the phrase “Who are you?”. Furthermore - as I know from experience - adults tend to use the phrase with children. And even in some settings involving grown-up speakers, I can image that the phrase is appropriate. The situation in which an adult is approached by the words “Hello. I am sorry to bother you. But I wanted to ask you if you held that excellent speech at the English department last Monday?” serves as an example. In such a case, the addressed person - as I would argue - is definitely legitimized to answer, “Yes, I was that person. And may I ask: Who are you?” With a smile, this question is definitely appropriate. The examples given are surely not the normal case of an English introduction, however it can be said that “Who are you?” is not as inauthentic as it might be assumed at first glance and that it should have a (small) place in English language teaching.

Another argument for the use of the phrase in CLT concerns the students’ average age and language level. Requesting from the target group, who has not
mastered the English alphabet yet, the correct formation and use of the conventional English introduction most probably causes too much cognitive load on the brain. Yet, if the phrase “Who are you?” and its answer “I am …..” are introduced first and later expanded to include the phrase “Nice to meet you” and drop the expression “Who are you?” most students - as I would argue - will master the English introduction.

The teaching design has shown that an intrinsically difficult film scene can be used for language beginners. What is important is that the requested tasks are not too difficult in regard to the students' cognitive abilities as well as to their age and language level, because the material itself already puts a high cognitive load on working memory, as it is made for native speakers. I would claim that within the presented teaching sequence, all activities are appropriate, nonetheless challenging, as they are not too “easy”. At the same time, the designed tasks are interesting and motivating, which can induce students’ intention to understand the film scene and to successfully accomplish the tasks.
2. AIW: AVC as a learning goal

The following teaching sequence focuses on AVC as a learning goal. Advanced students will learn how a text passage from the novel AAIW is translated into another medium, the filmic medium, and what changes and advantages such an audio-visual realization brings.

Target group: The target group consists of 16 to 17-year-old students of an AHS seventh grade. According to the Austrian curriculum, the target group should be in-between level B1 and B2 (the text says that after the sixth grade students should have reached B1 and after seventh grade B2 level for all communicative skills). In reference to the CEFR, B1 and B2 speakers are considered “independent” language users, the first being defined as “threshold”, the latter as “vantage”.

The audio-visual text: The students participate in a teaching sequence, in which the discussion centers on the audio-visual realization of a text, the source of which appeared as written text, which can thus be read or recited. I therefore call the original text “verbal” as this term combines both ways of comprehension: written comprehension (reading) and oral comprehension (listening). This verbal text is presented below (see “The other text”). From the Disney animation AIW the “Smoking caterpillar” scene (chapter 14, 30:33 - 33:40 min.) is used (as opposed to teaching design 1, a longer scene is shown). The film as well as this specific chapter are quite challenging for the target group, as they use unsimplified English, intended for native speakers.

The other text: The verbal text, which is taken from Carroll’s novel AAIW, is a poem titled “How doth the little crocodile”. It appears in chapter 2 when Alice tries to recite the moralistic poem “How doth the Little Busy Bee” by Isaac Watts, but fails. The poem goes.63

How doth the little crocodile
Improve his shining tail,
And pour the waters of the Nile

On every golden scale!
How cheerfully he seems to grin,
How neatly spreads his claws,
And welcomes little fishes in
With gently smiling jaws! (Carroll 1865: 23)

Learning aims:

**General aim 1:** The students should become familiar with AVC by using it as a skill in order to analyze the “Smoking caterpillar” scene in regard to the audio-visual realization of the verbal text “How doth the little crocodile”.

**Specific aim 1a:** The students should realize that AVC differs from listening by first listening to a recitation of the poem “How doth the little crocodile” and then viewing its filmic realization.

**Specific aim 1b:** The students should realize that the audio-visual realization creates puns which do not exist in the verbal text and detect at least one of these puns (e.g. a pun is created through the rapture of the signifier-signified bond of the word “scale” and the mental concept “biological scales” through the image “musical scale”).

**Specific aim 1b:** The students should find at least one faithful audio-visual animation of the verbal text.

I have already outlined in my analysis of the Austrian curricula that the development and understanding of AVC is stated nowhere. It is not part of serious didactic discussions, neither in Austria nor in the European Union. However, some other requirements of the curriculum for upper AHS-secondary school are covered by the teaching design. For example, “[der] reflektierende Umgang mit Sprache” is certainly an aspect which is developed by the designed lesson as the students discuss and reflect the workings of the language in two different media. Regarding the latter term “media”, the teaching design complies with the requirement that considers the incorporation of different media under the headline “Vielfalt von Lehrmethoden, Arbeitsformen und Lernstrategien” and the requirement “Rolle der Medien”. As the latter point is very unspecific, it can be interpreted in favor of the teaching design, because it involves a discussion

---

64 The students should not use the same terminology as I did, but express their findings in their own words.
of the role of the media in regard to the message of the text and its interpretation by the audience.\textsuperscript{65}

Considering the can-do descriptions of the curriculum for B2 listening, students should be able to understand films in which a standard variety of the foreign language is spoken. This expectation is met by the teaching design. The use of lyric for teaching seventh graders is not specifically prescribed by the curriculum, only prose texts are considered (B2 reading: “Sie [The students] können zeitgenössische literarische Prosatexte verstehen”). However, as the poem is rather short and its vocabulary, syntax and grammar are appropriate to the language level of the students, I would argue that it is a suitable text type.

\textbf{The role of AVC:} I have already established that audio-visual texts speak via audio-visual signals, which are decoded through AVC. In order to discuss the message of a text, i.e. what it says and how it talks to its audience, an analysis of communication is essential. At this point AVC comes into play. In a pedagogic framework, this means that students analyze in what ways the miniature film adaptation of the poem “How doth the little crocodile” communicates and how it is different from the verbal text of the novel AAIW. By doing this, the students employ AVC as a skill in order to process the audio-visual transmission. At the same time they develop an understanding of how AVC works, thus deal with it from a critical and analytical point of view, i.e. on a meta-level, which is concerned with the question of how language functions as a system. It involves awareness processes which eventually put the new information in relation to already existing information.

AVC is seen here as the topic of the lesson and its employment and understanding are the ultimate goals. Furthermore, an appreciation of AVC as a skill to process film information shall be gained.\textsuperscript{66}

\textsuperscript{65} Cf. http://www.bmukk.gv.at/medienpool/11854/lebendefremdsprache_ost_neu0.pdf
\textsuperscript{66} The teaching sequence involves a discussion of both the original verbal text and its filmic realization. However, this does not mean that a debate on the relationship between books and films or the faithfulness of the latter to the former shall take place. What shall be discussed is how a miniature adaptation of a mini-text is done and what new strategies are needed. In short, an awareness of AVC and its application to that certain filmic realization shall be achieved.
General outline of the activities: I presume that the students have already read the novel AAIW - partly at home, partly in class. Additionally, the teacher has discussed the topics of the story and its play with logic and nonsense. This means that the target group is familiar with the essence of the novel and is thus not surprised by the puns and mockery which appear in the film scene.

The planned teaching sequence starts out with a TST of what happens in the book in chapter five “Advice from a caterpillar” (Carroll 1865: 53-65). In this previewing activity, knowledge of the storyline and the characters should be activated. In this sense, it serves as revision for those students who need to reactivate stored information and as a chance to fill gaps for those learners who cannot remember the chapter that well. The answers given by the students are noted down on the blackboard.

The teacher then informs the students that the poem “How doth the little crocodile” appears in the second chapter and asks if they can remember it. I expect that the students will not remember the poem, because the teacher did not discuss it in class. If some of the students can remember it, they should be given the chance to express what they know. In either case, the teacher gives the instruction to listen to the poem and try to grasp it globally (“globales Verstehen” according to Buttaroni and Knapp 1988: 15). After listening to it the first time, the students discuss all information they could get from the text with their left-hand neighbors. Then the poem is recited again. This time the students note down two words (or sound-clusters which they identify as English words) of which they want to know the meaning. Afterwards, the students form new two-and-two formations, exchange information and explain the meanings of the new words to each other - without any intervention by the teacher. What follows is an activity named “Gemeinsames lebendes Wörterbuch” by Buttaroni and Knapp. In a TST the still unknown words are discussed. The teacher thereby gives the information in the form of synonyms, gestures or German equivalents – but only if none of the students can give the answer to the class. After this task, the
teacher recites the poem, which is followed by an information exchange in new
groups of two. Then the poem is recited one last time.67

Having acquired a global understanding of the poem and the acquisition of new
vocabulary, the teacher explains the following activity. The smoking caterpillar
scene from the Disney adaptation is presented, which reworks the poem “How
doth the little crocodile”. Since Alice’s encounter with the smoking caterpillar
has been discussed on the basis of the chapter review at the beginning of the
teaching sequence, the context should be clear to the students. The instruction
is to globally understand the scene (Orientating AVC) and take notes regarding
the first impressions of the scene, especially in regard to the appearance of the
poem. The impressions are noted down on a while-viewing commentary sheet
which is passed out by the teacher beforehand (see “Commentary sheet 1” in
the appendix). After having watched the scene, the students pass their sheets
to the left-hand neighbors, who then communicate their first impressions of the
scene to the whole group.

The film scene is watched again. This time the groups exclusively concentrate
on the realization of the poem on screen. They are instructed to note down how
the poem is presented. What the students shall find out is the following:

The poem is recited by the caterpillar, which is accompanied by a comical
cartoon of its content in the form of smoke exhaled by the hookah-smoking
reciter. The filmic realization of the poem can be defined as an interplay of the
verbal narration and its animated adaptation. The accompanying animated
pictures faithfully illustrate, follow and support the words of the rhyme. The
pictures, however, also mistranslate the intended verbal meaning into pictures.
An example of such a mistranslation is the crocodile’s golden tail which
represents scales as musical notes ascending in a scale, rather as scales on

67 This activity is a modified version of Buttaroni’s and Knapp’s model of „authentisches Hören“.
According to them, “authentic listening” does not comprise follow-up comprehension checks in
the form of yes/no-questions, text summaries, answering of questions or the like. Otherwise the
aims of the activity, “[die] globale Sinnkonstruktion”, and the acquisition of new vocabulary
would be destroyed.
the crocodile’s body. The intended and rather obvious meaning “scales on a reptile’s body” is not realized by the smoke images.

So what happens from a linguistic point of view? The signifier, the word “scale”, can be interpreted into two different mental concepts: “biological scales” and “musical scale”. In Saussurean words, for the signifier “scale” more than one signified exist. Viewing the scene, the signifier “scale” is presumed to mean “biological scales” due to the semantic content of the poem. The viewer, however, is confronted with another meaning, which is “musical scale”. At this point I do not want to discuss (and certainly not with the students) the simultaneous processing of the word “scale” and the image “musical scale”; however, I assume that the troublesome creation of the single mental model from the processing of disrupted signifier-signified bonds realized as divergent audio-visual signals must have certain effects. These can be negative in the form of a dilemma (the viewer might not know the two signified of “scale” and thus not understand the filmic realization) or positive in the form of humor and enjoyment. I assume that in the film AIW humor must have been in the foreground, as it is intended for native speakers. Moving back to the didactic discussion of the rupture of signifier-signified ties, the students shall find out that the mistranslation of the intended biological scales into a musical scale works only through the audio-visual realization. The original verbal version does not have the same (comic) effect as the filmic version as there are no smoke images that would cause it (surely, static pictures could have been added, but they were not). The pictures make the pun happen. In a nutshell, the pun only works in its filmic adaptation.

An example of a faithful visual realization of the recitation is the crocodile’s movement of his smoke claws as he helps the fish into his jaws. Here, the animation is faithful not only to the original meaning of Carroll’s text, but also to the recitation by the caterpillar. The signifier “jaws” as written or recited words corresponds with its signified “jaws” as smoke images.

Throughout the whole recitation, the smoke images dissolve as the sounds fade. Elliot, coming from interdisciplinary film/literature studies, states “… [The]
existence [of the smoke images is] dependent on the same breath that produced the words, with little persistence beyond the words." (2003: 225). Pictures depend on words - Elliott says. I would go even further and claim that words and pictures depend on each other. Without the interdependence of animation and recitation, the poem would not work the same way and its fascination would be lost. But what is the fascination of the filmic adaptation? As Elliott argues, it lives from the “ruptures between semantic sense and sensory sense, and within sensory sense between auditory and visual sense” (2003: 226). I completely agree. Additionally, as I would claim, the fascination also emerges from the harmony between semantic sense and sensory sense as well as the harmony between words and pictures. The auditory realization correlating with the semantic meaning is reinforced by its correspondence with the visual signals. As a result, the signifier-signified bond is strengthened.

Returning to the teaching outline, the students analyze the scene in groups of four. A commentary sheet, passed out by the teacher, shall help them to notice the particular characteristics of the audio-visual realization of Carroll’s poem (see “Commentary sheet 2” in the appendix). This activity constitutes a project-oriented task as the students work autonomously in groups. Later, the results are shown to the whole class in a short presentation by a group leader. After all the results have been presented, the teacher reviews the characteristics of the audio-visual adaption together with the students (TST).

What is important for the discussion is that the focus should not be on the faithfulness of the film scene to the verbal text, but rather on what effects the former creates and how they work for the audience. The goal of the analysis of the film scene and the verbal text is to create awareness for the sensory duality of the medium and what it “does” to the text it remolds (as outlined above). Furthermore, its effects on the audience, their comprehension of the audio-visual text, are discussed. AVC is thus employed by the students and at the same time it constitutes the subject-matter of the discussion.
Another interesting aspect of the audio-visual realization, the fact that the animation comes in the medium of smoke, should not be left out of a discussion of how the poem is realized on screen. It is moved to the end of the lesson, as it concerns a particular aspect of the filmic realization.

Smoke has a range of meanings: it can refer to the smoking of tobacco, the burning of material, the disinfection of food for preservation or idle talk. In a TST the teacher tries to trigger these various meanings from the students by giving them hints in the form of pictures presented on an overhead transparency (see “Smoke sheet” in the appendix). As the smoke images are not straightforward, it should not be too easy for the students. Then the teacher elicits various characteristics of smoke (e.g. “vaporous”, “see-through” or “vanishes quickly”) and fills the empty cloud of smoke (on an overhead transparency) with the students’ suggestions, which could then look like this:

Afterwards the teacher gives the students three more words, “unreal”, “unsubstantial” and “transitory” which describe smoke. The students shall now relate these three characteristics of smoke to the film scene and answer the questions: “In what way is smoke unreal and how does unreality fit into the

---

68 In fact, it has many more meanings. The ones listed above are the most important ones.
context of Alice’s story?” The same questions are formulated for “unsubstantial” and “transitory.” The three questions are assigned to groups of three students, in which each student works on one question independently. Then the ideas are discussed within the groups first and finally in a TST.

The answers to the questions above could be that smoke appears unreal as it is often connected with dreams and illusions. In regard to the story, unreality reflects the fact that Alice is in the unreal world of Wonderland, experiencing things that cannot be real. Considering the term “unsubstantial”, smoke is unsubstantial because it cannot be touched and thus has no real substance. Insubstantiality is reflected in the recitation of the poem “How doth the little crocodile”. For Alice the poem has no meaning, as she cannot deal with it. The caterpillar, however, plays with Alice, in fact is delighted by her confusion, when he recites the original poem “How doth the little busy bee” in a wrong way. In this respect, the poem has no real substance, i.e. meaning, neither for Alice nor the caterpillar, but is regarded as a deliberate confusion and an instrument to cause this confusion used as a means of amusement. Concerning the word “transitory”, smoke vaporizes easily and quickly, hence dies fast. This fact could refer to words easily fading together with their meanings. They become insignificant, which alludes to the German saying “Worte sind wie Schall und Rauch.” Of course, any interpretation, as long as it is well argued, is accepted by the teacher.

In a subsequent task, the teacher enters the discussion of the idiom “It is all smoke and mirrors” by asking the students where both mirrors and smoke can be found. I assume that the students come up with the idea that mirrors and smoke were used by professional magicians in order create illusions and thus deceive the audience. Considering smoke as an instrument to distort and blur facts, the teacher instructs the students to relate this idea to the film as a group discussion. Appropriate answers could be that the caterpillar’s recitation is an actual act of deception, because he recites the poem incorrectly. Another idea concerns the content of the poem, as it deals with the art of deceit, which is

---

⁶⁹ “In what way is smoke insubstantial and how does insubstantiality fit into the context of Alice’s story” and “In what way is smoke transitory and how does transience fit into Alice’s story?”
mirrored as smoke images. Generally, the story of Alice lives from fooling people and affecting their comprehension.

As a last step the teacher gives the students the idiom “It is all smoke and mirrors” on an overhead transparency and discusses the fact that it does mean exactly the same as the German phrase “Schall und Rauch”. Whereas the English idiom refers to trickery and the attempt to distract and deceive people, the German phrase relates to the insignificance of, for example, words and actions.

As homework, the groups are supposed to find more examples of signifier/signified-relations realized as word/picture-relations in the film.70 Thereby they have to find both raptures and correspondences. Regarding the raptures, many of those can be detected. Such mobile audio-visual puns are, for instance, the “rocking-horse-fly”, the “bread and butterflies” (both: chapter 11, 25:22 -25:48 min.) or the “half cup of tea” sliced by the March hare (chapter 17, 44:49 min.). The students present their findings to the class. In a TST the results are summed up. If the teacher decides to let the students express their results in written form, each student can pick one audio-visual pun and write a short analysis as a homework assignment.

---

70 My first thought was to provide copies of the DVD for the students. This, however, causes problems due to the copyright law. There are four possible ways to solve this problem. One solution is that the students and the teacher decide together that they buy a DVD for each group from money collected from all the students. Another solution is to provide the DVD in the framework of so-called “course reserves”. Materials from course reserves can be checked out for a limited amount of time (e.g. 5 hours) any time the school library is open. The students just have to leave their names and bring back the DVD reliably. While they borrow the DVD, they can either watch it in school (on the school laptops or in a media room) or at home. To make this system work, the teacher has to set up a timetable which organizes which group borrows the DVD at what time. Another possibility is to arrange group viewings. The teacher thereby offers, for instance, three dates on which the students can watch the movie in school. A final solution is to let students organize private group-viewings for which they have to procure the DVD themselves.
3. Seeing, hearing and feeling South African history: Teaching power relations during the 1970ies through AVC

The history of a country is a broad topic which could fill hundreds of lessons. In this teaching design I therefore concentrate on a specific period of South African history, Apartheid in the 1970ies and a specific aspect of that period, the relations between black and white South Africans, presented as what I call “power relations”.

Before I outline what power relations actually mean and how they can be taught, I want to make clear that the presented teaching design is taken out of a context of many lessons that deal with South Africa, its history and present-day problems. This means that the designed lesson is not a self-contained unit of teaching, which could be taught independently, but has to be preceded and followed by lessons that reflect history in more detail and - what is even more important - consider contemporary South Africa and its current problems, which result from the country’s past.

In this teaching design I regard power relations as a correlation of an inferior person (or a group of people) and a superior person (or a group of people). A power relation is also given if the two parties are on the same level. Power relations are demonstrated as positions of inferiority/superiority or equality, of which different types exist, operating on the following levels.\(^71\)

- Social inferiority/superiority: It refers to the person’s social standing within a social group which influences various factors such as education, job and living standards. In case of Apartheid in the 1970ies, black South Africans are generally regarded as being socially inferior, white South Africans as socially superior. Social power is very much linked to political power.

\(^71\) The following definitions do not constitute general concepts of power relations or power in general which could be applicable to any context, but they refer to the very concrete situation of black and white South Africans in the 1970ies.
- Political inferiority/superiority: It refers to the power to control and regulate the lives of all people living in the country. In the 1970ies white South Africans were politically superior as they regulated the lives of black South Africans.

- Moral inferiority/superiority: It is dependent on a person’s morals, which are either in line with human rights or not. Moral inferiority is given if, for instance, a person morally agrees with the oppression of black South Africans. Moral superiority is given, for example, if a person fights for equality.

- Situational inferiority/superiority: It refers to whether a person is dependent on another person in a particular situation or not, which does not necessarily have to correspond with the social or political power this person otherwise has. For example, a black South African doctor is situationally superior to a white South African who needs urgent medical care despite the political inferiority of the former.

- Educational inferiority/superiority: It concerns a person’s educational and thus his/her job position. A black South African who has received the same education as a white person is educationally equal to a white South African, even though he/she is socially inferior.

The different types of power relations established as positions of inferiority/superiority will be explored through a close analysis of selected film scenes from the movie CF. The scenes will be examined in terms of how the filmic techniques camera angle, sound and lighting (which also includes the study of colors) display these power relations.

**Target group:** The target group consists of 17 to 18-year-old students of an AHS eighth grade. The Austrian curriculum prescribes that the students (considered “vantage” learners by the CEFR) should already have reached B2 level for all communicative skills. Under the can-do description for “B2 listening”, the understanding of feature films using the “standard language” is listed. In the case of English, the definition of the term “standard” is quite challenging as an international standard does not exist. Regarding South African English as the
Standard English of South Africa, the film CF provides different and - as I would claim - authentic social variations of that standard spoken by different groups: the suppressed black majority, the small community of educated blacks, white upper class and South Africans of Afrikaans descent. This variety of South African English complies with the requirement “Nationale Sprachvarietäten sind exemplarisch in den Fertigkeitsbereich Hörverstehen zu integrieren” under the heading “Erwerb soziolinguistischer Kompetenzen”.72 The different variations of South African English cause a relatively high cognitive load in the students’ working memory; however, the designed tasks will reduce the overall load and help students to cope with the different variations of South African English.

The audio-visual text: CF is based on the real-life characters Steve Biko, the banned black opposition leader of the movement “Black Consciousness” and the white editor Donald Woods. In the 1970ies, Biko befriends and converts Woods to his ideas of his movement: black pride, peaceful change and equality of both races. After Biko is murdered by the police, Woods, meanwhile an opponent of the Apartheid regime, becomes banned himself for pursuing an inquest concerning his friend’s death. Woods finally flees into British exile where he fights Apartheid and raises awareness of its cruel nature by publishing a book on Biko’s life.73

The different varieties of South African English realized in the audio-visual text make it an intrinsically difficult material even for eighth graders. Furthermore, the topic itself is emotionally demanding, as it shows human cruelty and suffering. To lessen this effect, appropriate tasks and a careful though critical approach to this sensitive historical topic are needed.

Learning aims:

General aim 1: The student should become familiar with South African history in the 1970ies by learning about the different power relations existing at that time.

Specific aim 1a: The students should be able to understand the teacher’s argumentation in an in-depth analysis of the film scene “De Wet vs. Biko” from

---

CF, in which the different types of positions of inferiority/superiority are discussed through the examination of film techniques.

**Specific aim 1b:** The students should be able to identify at least two examples of inferiority/superiority in the scene they are given as a group, and two corresponding film techniques which illustrate these examples.

**Specific aim 1b:** The students should be able to illustrate their ideas by creating a presentation (Power Point), which includes at least two frames taken from their respective film scene and two edited frames.

**The role of AVC:** AVC as an instrument to comprehend an audio-visual text also constitutes a way to experience the history of the country portrayed in the text. This historical journey into an often unfamiliar country is both visual and auditory and thus provides a multidimensional window into various historical aspects, which would not be possible with other media. In CF, the power relations between the black majority and the white minority of South Africa during the Apartheid era in the 1970ies are portrayed through audio-visual interactions between white and black representatives. In this sense AVC serves as a tool to learn about the historical past of South Africa, raise the students’ awareness of this topic and contribute to peace education for the purpose of the educational mission Austrian schools have.

In the designed teaching sequence, the target group analyzes scenes of CF which display the existence of different power relations between black and white South Africans during the Apartheid regime. The students thereby employ AVC as a skill in order to process the audio-visual text; however, at the same time they consider AVC on a meta-level, as they evaluate how power relations are communicated and highlighted through the use of film techniques. In particular, the students investigate how camera angles, lighting, sound and other techniques give characters a sense of inferiority/superiority and thus power over other characters. This teaching design thus gives the students the chance to learn how to analyze audio-visual texts.

**General outline of the activities:** I presume that the students have already developed an idea of South African history and that they are informed about the
key data from previous English lessons: the indigenous people of South Africa, the Khoikhoi and San, Dutch colonization in the 17th century and their 150-year rule over the country, subsequent British control, the Anglo-Boer War and the following segregation of the white and non-white “races” until the end of Apartheid in 1991, the first democratic elections in 1994 with Nelson Mandela as the president ruling until 1999 and the various problems resulting from the oppression of the non-white majority for centuries. Furthermore, they are familiar with very recent problems in South Africa such as HIV/AIDS, an increase in crime, racism, corruption and the growing gap between rich and poor.74

The teaching design involves a close analysis of three film scenes. In order to interpret them in a detailed way, they have to be first understood globally (Orientating AVC). Therefore, the students are supposed to watch the whole movie before they can have a deeper look at its details. In this way, the students are immersed into the context of segregated South Africa in the late 1970ies and they are prepared for the upcoming discussion. Another argument for watching the whole movie is that dealing with just parts of it is very often unsatisfying for the viewers, because they are not given the chance to see how the story ends or starts. Not always can the whole film be watched (particularly not in the limited time frame of a 50-minute English lesson in school), but in the case of CF the students should be given the chance to get to know the whole story. The watching of CF is a homework task, which is either done at home (the students can borrow the DVD from the course reserve counter of the library for a limited amount of time) or it can be watched by all students at an arranged date outside their usual classes, preferably in the afternoon (the teacher then has to set up a facility, e.g. a media room, in school).

At the beginning of the lesson the teacher presumes that all students have watched CF. In order to reactivate the students’ knowledge of what happens in the film, the following warming-up activity is done. The teacher shows the trailer

74 The upcoming soccer world championship 2010 is hosted by South Africa. This information is not valuable for the teaching of South African history in the 1970ies as it is intended here, however, it can serve as a teaser which arouses the students’ interest for the topic, as the forthcoming event is relevant and up-to-date. For more information on South African history see Davenport: 2001.
(found in the main menu of the DVD) with the sound being switched off. The students, who have formed groups of four people, are instructed to produce a written text which corresponds to the pictures they see. Thereby they shall write two types of texts: the lines spoken by the characters, indicated by their lip movements and the narration spoken by a third-person, summing up the storyline. For this, the groups have to split into two sub-groups. As a start-off aid, the students get the first words of the trailer on an overhead transparency:

This is Stephen Biko (narrator), Oh, Biko, Biko... (a crowd singing)
we stand out in any way he chooses (Biko) ...
(chapter 1, 00:00-02:43 min.)

While the students work, the trailer is silently played non-stop. The activity is challenging because it requires creativity and a high level of language proficiency. At the same time it is a fun task, which is more exciting than a usual oral plenum discussion of the homework such as “Tell me, what was the text about?” Furthermore, the designed activity comes in handy for the teacher as the students formulate a summary of what they have understood and they express the points they find most important - without even noticing that they are reviewing homework.

The students present the results in their groups. They read or act out the text along with the playback of the trailer. Then the trailer is played with the sound switched on. Depending on the type of class, the activity can be seen as a challenge of what group comes closest to the original text of the trailer.

Having activated the students’ knowledge, they watch the scene “De Wet vs. Biko” (chapter 5, 33:41 - 36:23 min.). For now, they are only instructed to understand the scene globally (Orientating AVC). The scene is watched twice, which is followed by the teacher’s analysis of two selected frames, presented with the help of Power Point. The teacher thereby introduces the students to film analysis techniques. During the TST, the students have to take notes on a sheet distributed beforehand (see “Discussion sheet” in the appendix).
The teacher asks who holds the superior position in this frame. I presume that the students claim that Biko is in the inferior position to De Wet and the other policemen. Before this idea can be discussed in detail (also that it is not completely true), the different types of inferiority/superiority are clarified first, which is done best by establishing them together with the students and noting them down on the blackboard. Afterwards, the different types of inferiority/superiority are applied to this particular frame. The students most likely have already found out that Biko occupies the inferior situational position as he is in police custody, especially at this point of the interrogation. Additionally, he is socially and politically inferior like every other black person in South Africa at that time. Despite the fact that most black South Africans are denied access to good education, Biko is presumably more intelligent than the other three characters in the frame (this can be inferred from the whole movie, which the students have already watched) and thus holds the educational superior position. Furthermore, he is morally more powerful than the other characters, as he fights for the equality of both races, whereas the policemen, with De Wet leading the way, do everything to keep Apartheid alive.

75 The program “Power DVD” is useful as it provides a function, with which screenshots can be easily taken: press “Frame speichern”, change to the Microsoft Word program and insert the shot.
Having discussed the types of inferiority/superiority in the frame, the students now concentrate on how they are reflected on the visual and auditory level, i.e. they interpret the visual and auditory components of the frame. The teacher gives the hint to concentrate on camera angle, lighting and sound. In order to memorize the discussion, the students get a help sheet (see “Discussion sheet” in the appendix), on which they can note down the most important points during the TST.

After the students have expressed their ideas, the teacher illustrates his/her arguments with the following edited frame:

**Camera angle**
The frame shows that all the policemen including De Wet look down on Biko, indicated by the three arrows pointing downwards. The phrase “looking down on somebody” is reflected through the type of visual presentation in this frame, a so-called “high-angle shot” (Konigsberg 1998: 174), a shot made with the camera looking down. This way of filmic expression represents the power of one character over the other and in this particular case the situational superiority of the police over Biko, enabled by the system of Apartheid and its laws.

---

76 The terminology used for describing film techniques is taken from Konigsberg: 1998.
Colors and lighting

Biko’s moral superiority over De Wet and the others is demonstrated through the technique lighting and the play with colors. The frame constitutes the highly contrasted colors black and white. The policemen, the antagonists of the film, are presented in dark colors such as dark grey, checked grey, blue and black. This choice is deliberate; it supports their image of the “bad guys”. Biko, the hero and protagonist, has to be presented in strong contrast to the bad guys, which is achieved by light colors. As Biko’s skin color itself is dark (and in this scene his clothes are blue), the light appearance is achieved through his reflection in the glass window of the dividing wall in the background of the picture. Supposedly, a sunbeam enters the room through a trellised window facing Biko, breaks through him and throws a shadow on the glass window behind him. The image created shows the shadow of Biko’s head in the middle of illuminated bars. It creates the impression as if he was a savior (the religious reference arises from Biko’s reflection, reminding me of church windows).

The very same picture can be interpreted as displaying Biko’s political and situational inferiority, especially in regard to the continuing storyline. Taking the same frame, it can be argued that it foreshadows Biko’s imprisonment and his death associated with it. His head as a shadow on the window glass can be read as a person being captive in prison or, in Biko’s case, trapped in a political inhumane system. The shadow itself is a symbol of transience signifying Biko’s political vulnerability and human mortality.

The lighting technique used to describe Biko’s morally superior position as well as his politically and situationally inferior position is called “low key lighting” (Konigsberg 1998: 223-4), as it creates strongly contrasted areas of light and shadow (it is low because the main light is set low or off). The following edited frame is presented to help the students understand the teacher’s arguments:
The two interpretations of the same frame should show that when it comes to the interpretation of film texts and techniques, no clear-cut answers exist - especially not in regard to positions of inferiority/superiority. Some frames are more obvious and straightforward to interpret, others are more complex.

**Sound**

Before the teacher moves on to the next frame, the auditory aspects of the film scene have to be discussed in terms of how they display power relations. An auditory evidence of the police’s political and situational superiority over Biko is the police siren wailing at the beginning of the scene, as Biko is led into the incident room (chapter 5, ~ 33:47 min.). The siren literally screams (out) the political power of the police. Considering the siren as a technique, it is defined as an “ambient sound” (Konigsberg 1998: 10), which is a seemingly live background sound creating the illusion of experiencing the real world. The wailing of the siren, a symbol of the supremacy of the state with its laws and regulations, dies away as soon as De Wet closes the door. In this way de Wet communicates that law does not exist in this room, which thus becomes an anarchic place isolated from the outside world. Another interesting aspect of the scene is that, apart from the siren and the soft noise of the fan (which I will discuss below), there is no single sound that accompanies the scene, which - I would argue - intensifies the different power relations by drawing full attention to them.
The next frame that is discussed is the following (chapter 5, ~ 36:10 min.):

“Off with his head” - this point of view, defined as a “bird eye’s shot” (Konigsberg 1998: death. 33), foreshadows Biko’s death, as the rotating ceiling fan comes dangerously close to Biko; his head will literally “roll”. This interpretation is supported by the fact that the fan was switched off for most parts of the interrogations; only when Biko starts speaking about the equality of all people does De Wet suddenly turn the fan on. Evidence is found also in De Wet’s utterance “We are going to catch you red-handed one day, and then we will see how human you are”, referring to Biko’s vulnerability and mortality as a human being.

Another indication of Biko’s death is the position of the policemen’s hands. They are very close to his neck and aorta symbolizing that his lifeline is in mortal danger by the hands of white men. Also, in regard to the overall composition of the frame, the policemen’s hands are positioned high up on Biko’s body; the police have literally gained “the upper hand”, i.e. the situational power and control over Biko.

77 The phrase “red-handed” might not be known to all of the students and thus has to be clarified by the teacher.
Sound
Apart from the characters’ speaking, only the quickening of the fan can be heard, indicating the situation is coming to a head for Biko. The general silence of the scene represents a death-like stillness as it exists before hurricanes, storms or other violent forces.

The arguments are visualized through the following edited frame:

Through the in-depth analysis of the two frames and their film techniques, the students should become more aware and develop a better understanding of what different power relations between black and white South Africans existed in the 1970ies. They also explore how power relations and positions of inferiority and superiority are communicated in real life (e.g. by literally looking down on someone) and how they are highlighted through filmic techniques.

A major consequence of the use of film scenes to teach historical or other aspects of the target language is that the learning effect is intensified because the knowledge is acquired audio-visually (“multiple representational principle” or “modality effect”), rather than though just one sense. For the teaching design this means that students will remember better and longer what they have learned.
Next the students should work independently in groups (they stay in groups of four). Presuming that three groups exist, each group is instructed to analyze a film scene in regard to the different types of inferiority/superiority and find evidence for these, which they should outline in a short Power Point presentation. This should contain both original frames taken from the scene and edited frames visualizing their arguments. Additionally, the students get a help sheet, a glossary containing the most important film techniques and their definitions (see “Help sheet” in the appendix), which they can use for their analyses.78 Basically the students follow the argumentation as presented in the TST, but they are confronted with new material. The following three scenes are distributed to the groups, each of which is equipped with a laptop. Ideally, the students can work undisturbed in three separate rooms.79

Scene 1
The first scene is set in Woods’ office as he discusses his political views and that of his newspaper with the black freedom activist Dr. Ramphele (chapter 2, 09:45 - 12:26 min.). The power relations can be identified as follows:

- Dr. Ramphele’s superior situational position, as she verbally attacks Woods for publishing the untruth about Biko; her superiority is also based on the fact that she knows more about the Black Consciousness movement and Biko’s ideas
- Woods’ situational inferiority resulting from Ramphele’s verbal attacks and his lack of knowledge about Biko and his political ideas
- the educational equality between Dr. Ramphele and Woods, as both characters are highly educated
- Woods’ belief that he occupies the morally superior position, because he thinks that his liberal views stand above Biko’s “racial war” (as he calls it), which is in contrast to his true
- moral inferiority, as he thinks that white South Africans are the superior “race” in South Africa, evidence for this is given by Woods’ question to Ramphele “Where are you from?” (he does not think that the educated woman is from

79 The students are neither instructed to use English subtitles, nor are they forbidden to do so, but in case they find them helpful for understanding the scenes better, they can be used.
South Africa) and his comment “I am glad we [the whites] did not waste our money [to give you, black people, education]” (chapter 2, ~11:25 min.)

- the political and social inferiority of black South Africans in the 1970ies: the inferiority of the black people is particularly displayed on the job market: they are house servants, cleaners, tea boys, nannies and the like
- the political and social superiority of South Africans, as displayed by Woods’ secretary, who not only looks down on Ramphele due to her skin color, but treats her in a condescending way; such behavior constitutes her moral inferiority

The following frame (chapter 2, ~10:39 min.) could be an example the group uses in the presentation. It displays Woods’ situational inferior position by a “low angle shot” (Konigsberg 1998: 223), made with the camera filming upwards. Wood has to look up to Ramphele, listening to her accusations and verbal attacks.

Scene 2
The second scene (chapter 3, 14:00 - 16:11 min.) describes Woods’ and Biko’s first encounter. When Woods first hears and sees Biko, an educated, eloquent and clever man, he becomes aware of the fact that black people have the same capacities as white people, and are thus on the same level. For the first time, he

---

80 I want to clarify that Woods’ moral inferiority vanishes as he becomes an anti-Apartheid activist himself. Only then does he occupy a high moral position - equal to that of Biko, because he risks his life for the fight for equality.
is aware of the total equality of all people. However, he still tries to shift the blame of Apartheid by saying:

\[ I \text{ wonder what sort of liberal you would make, Mr. Biko. If you were the one who had the job, the house and the Mercedes...} \]

(chapter 3, ~15:45 min.)

This comment shows that Woods is not yet on the same moral level as Biko, because he cannot fully admit his and his people’s guilt.

The following frame could be chosen by the students (chapter 3, ~16:07 min.):

The frame consists of almost perfect harmony and symmetry in regard to the overall composition. In terms of color the frame is generally dark; however it becomes lighter to its middle. The lightest point is found in the very center (indicated by the second smallest and the smallest green frame), at which is the handshake, a symbol of friendship, equality and teamwork. The illumination stresses the central position of the handshake, hence its importance, and creates the impression of a “radiant”, i.e. deep and warm friendship between Woods and Biko. In terms of symmetry, the frame is structured by four frames diminishing in size. As they are almost entirely symmetrical to each other, they suggest that a camera is zooming on the most important aspect of the composition, the symbolic handshake. The main attention is thus drawn to it by
the interplay of the four frames. Each of the frames alone conveys an additional message by enclosing a certain part of the frame. The second largest frame, for example, shows that Biko and Woods are face to face, neither of them is looking down or up to the other person. This means that the two characters regard each other as being on the same level, which makes them equals. The second smallest frame demonstrates that Biko and Woods join hands, which can also refer to their common fight against Apartheid. The smallest frame accentuates the clasping of the two hands, signifying the mental and emotional union of two people.

The composition of the frame as diminishing frames can also be interpreted in the light of the phrase “to be in the frame”, which means “someone is likely to be awarded” or “to achieve”. As Biko and Woods are indeed “in the frame(s)”, they are likely to be successful in their common endeavour to achieve equality.

Returning to the discussion of the whole scene, I would argue that it lives from the use of light. When Woods meets Biko in the garden of the community church for the first time, the camera shoots directly into the sun through the swaying branches of a willow tree. The fractured light infolds Biko as he approaches Woods from behind the tree, which makes him look like a savior. This and the fact that Woods is literally blinded by the sight of Biko, give him the situational power (chapter 3, ~14:06 min.) over Woods:
Scene 3
In scene 3 (chapter 12, 1:52:49 - 1:54:27 min.), Woods encounters an old black man who helps him flee the country. Woods is situationally inferior as he requires help from another person. The usual political and social power relations are reversed: a white person’s fate is dependent on a black person’s will. The old man, however, does not take advantage of his situational superiority, but treats Woods kindly and decently due to his high morals. The only advantage he takes of the situation is to burst into laughter at the idea of a white person escaping with incriminating evidence against the South African government: "Kruger will shit himself - Boethe will shit himself..." (chapter 12, ~1:53:51). This funny outburst serves as comic relief from the overall emotionally demanding story.

The following frame demonstrates that Woods occupies the inferior position in this situation, as he has to look up (chapter 12, ~1:53:16 min.). His inferiority is highlighted by his desperate look, which speaks volumes. Regarding the overall composition of the frame, it suggests that Woods is sitting in a prison cell, supported by the sunbeam entering the predominantly dark room through the window behind him. The idea of captivity (in fact, it correlates with the story, as the ban imposed on Biko makes him a prisoner of his own banning area), is contrasted with the feeling of hope raised through the sunbeam, which fills the dark room with light.

The dominant dark colors in the background create the impression as if Woods was sitting in a prison cell. The sun, however, fills the room with light and hope.
After the group presentations, a TST follows, which sums up the different positions of inferiority and superiority identified by the groups.\textsuperscript{81}

The last task constitutes a rounding-off activity. The students are instructed to choose the most appropriate movie poster from the following choice of posters.\textsuperscript{82}

\textsuperscript{81} The teacher collects the presentations via email, copies the most important points onto a Word document and hands out a copy to each student in the next lesson.
\textsuperscript{82} All the movie pictures are taken from the following website:
The students discuss the key messages of the posters with their neighbors, decide on the one they find most appropriate and give arguments for their decisions. This way the students have to think of the key message of the story and what aspects they regard as most significant to communicate to potential viewers.

I have already stated at the beginning of this chapter that history does not only refer to what is past but also to what is present. A country’s existing situation is always the result of its past events. Teaching South African history thus means that the teacher cannot stop here, but has to involve a serious discussion of contemporary South Africa, as it is a result of its own past.
4. How to get students to speak: Using a video clip

Originally, I wanted to use a video clip from the website “utube” in order to show that not only films but also short video sequences are suitable as teaching material. However, in the course of my research on utube I stumbled across a funny scene from the movie “Notting Hill”. This short clip shows a stunned half-naked person, named Spike, who is described as "the stupidest person in the world, only doubled" (chapter 3, ~ 20:45 min.) by his roommate, as he proudly poses for reporters on his doorstep. I fell in love with the scene and decided to use it in my fourth teaching design. This time AVC plays only a minor role in the overall teaching design, as it is used as an instrument to understand the short clip, which in turn serves as an impulse to practice the skill “speaking”.

Target group: The lesson is designed for students of an AHS-fourth grade. Regarded as “Waystage” by the CEFR, the students should be on A2 level for all communicative skills (at the end of the fourth grade, the target group should also have developed some of the prescribed B1 can-do’s for the skills listening, writing and reading, e.g.: „Die Schülerinnen und Schüler können die Hauptpunkte verstehen, wenn klare Standardsprache verwendet wird und wenn es um vertraute Dinge aus Arbeit, Schule, Freizeit usw. geht.“ (B1 listening).  

The audio-visual text: The scene, which I named “Posing Spike”, is taken from the movie “Notting Hill” (chapter 11, 1:19:54 - 1:20:22 min.). The content of the scene is the following: A strange, weird-looking guy wearing grey underpants steps out of a blue front door. Suddenly, hundreds of cameras take pictures of him, presuming that the movie star Anna Scott has just come out of the house. Spike, instead of running back into the house or acting frightened, poses athletically for the reporters. After his few seconds of fame, Spike goes back in, closes the door and wanders along to the mirror in the hall, muttering:

Spike: How did I look? [looking in a mirror]
Spike: Not bad, not bad at all. [inspecting himself] Well chosen briefs I must

84 The second part of the scene can also be found on utube: http://www.youtube.com/watch?v=4JFT4lfLpVY.
say. Chicks dig grey. [*clenching his butt*]


(chapter 11, 1:19:54 - 1:20:22 min.)

The scene ends with him leaving the room.

The clip presents a half-dressed man, showing and inspecting his body. It could be argued that this is not the most suitable material to show to 13- to 14-year-old students; however, I would claim that the context (a young man who has been surprised by the press and checking what impression he has just made on the reporters), the presentation (the way Spike talks and behaves) and the language (what Spike says) add a major comical effect to the scene and thus make it suitable for fourth graders. I would say that the focus of the scene, its humor, is recognized, understood and appreciated by its viewers.

The students will not understand every word of the clip, because they are not articulated clearly. Moreover, some of the words such as “briefs”, “chicks”, “dig” and “buttocks” might be unknown. Therefore, and due to the fact that the film is originally created for the amusement of a grown-up English speaking audience, the comprehension of the clip puts a high load on the students' working memory. Consequently, some aid must be given to the students and the tasks must be designed to reduce the overall cognitive load (see “General outline of the activities”). Despite the intrinsic load put on the students, I presume that they know at least 50 per cent of the vocabulary involved in the film scene, such as the phrases “How did I look?” and “I must say” and, for instance, the words “bad”, “at all”, “well chosen”, “grey”, “nice” and “firm”.

**Learning aims**

**General aim 1**: The students should talk as much as possible.

**Specific aim 1a**: The students should be able to give accurate descriptions of the pictures they see (either frame 1 or 2). The description of the person's appearance, his location and the surrounding items should comprise words of locations, i.e. prepositions (e.g. “in front of” or “next to”), words of actions or
conditions, i.e. verbs (e.g. “step out” or “stand”) and vocabulary relating to appearance, i.e. nouns (e.g. “pants”, “hair” or “face”) and adjectives (e.g. “funny”, “awkward” or “surprised”).

**Specific aims 1b**: The students should actively participate in both TSTs and group discussions by expressing at least one idea or presenting the group results orally in front of the class and they contribute actively in the oral discussion of an alternative script by suggesting at least one line.

**Specific aim 2**: In order to reach aim 1a and 1b the students should understand the scene globally (i.e. the content of the scene) and in detail (i.e. the script).

**The role of AVC**: In this teaching design, AVC is needed in order to decode the audio-visual signals transmitted by the video clip. AVC is therefore a tool with which the clip is understood, which in turn is essential to fulfill the subsequent tasks. In contrast to the other teaching designs, here AVC is neither the goal of the lesson, nor is it considered on a meta-level as when film techniques are discussed or an audio-visual realization is compared to its verbal equivalent. AVC is not consciously given attention, which rather lies in the overall goal of the teaching sequence, to get students to talk. However, AVC is important as it facilitates the achievement of the learning aims.

Moving from the role of AVC to the role of film, one reason for the use of audio-visual texts strikes out in this and the previous teaching designs: films as motivators. I have already argued that most students, themselves being authentic L1 users, react positively to the use of audio-visual texts such as films and video clips, because they notice that they are not confronted with material designed for their needs, but with authentic material as authentic language users. This has a motivating effect, as many students want to deal with the same texts native speakers do. In the case of “Posing Spike”, students realize quickly that the scene is taken from a Hollywood movie intended for native speakers, as Spike speaks in a thick Welsh accent. Another motivating effect is
that the text has been ripped out of its context - students want to know what happens before and after the scene.\textsuperscript{85}

**General outline of the activities:** The students are supposed to get together in groups of two people, being equipped with a laptop and the DVD (beforehand, the groups were instructed to procure the DVD, either through purchase, borrowing or otherwise). One student sits in front of the computer facing the other student who cannot see the screen. The student with the laptop opens a word file, which contains the following frame from the scene:

1. “Spike posing in front of the mirror”, (chapter 11, \(~1\text{:20:12}\) min.)

![Frame from the scene](image)

The student facing the laptop describes to his partner what he/she sees, who simultaneously draws what he/she hears. This activity is called “visual dictation”. Describing a person’s appearance and his/her position in a room should be a manageable, though challenging and fun task. When the dictation is finished, the students swap places. This time a different file is opened with the following frame:

2. “Spike being photographed by the press”, (chapter 11, \(~1\text{:20:02}\) min.)

\textsuperscript{85} If the students want to watch the whole movie, they teacher can set up a group viewing outside the regular lessons, e.g. in the afternoon.
The students repeat what they have done before with switched roles. Then the two drawn pictures are discussed in comparison with the frames from the scene.

The teacher’s role in this activity is very passive. Apart from giving instructions and helping students in case of technical or other problems, he/she is silent. In contrast to the teacher, the students are loud and active. They are supposed to talk as much as they can - especially when the drawings are compared with the frames from the scene, which is designed as a fun activity in which students can laugh at their talents for drawing.

The two frames, shown in reverse order from where they appear in the scene, represent the video clip in visual form. They basically sum up what happens in the scene; however, the students do not yet know its context, what is spoken, what the atmosphere is like and in what order the frames appear, which are crucial factors for the overall message of the scene, which is only conveyed through the audio-visual text. In a next activity, the students have to discuss these points by answering the questions, “How are the two picture related?” and “What happens between the two pictures?” in groups of four (two previous groups form a new group).

The results are discussed in a TST. As the focus is on speaking, the students do not have to make notes of their ideas.
The students stay in groups of four. Now they watch the video clip on the laptop with the instruction to understand it globally (Orientating AVC). After having watched the scene twice, the students discuss what happens before and after the scene. As an aid, the teacher provides the following key questions on an overhead transparency:

- Why has Spike left his flat/apartment wearing just his briefs/underpants?
- Why is he awaited by hundreds of reporters?
- What is he doing in front of the mirror?
- What will his next reaction be?

The groups have to come to an agreement concerning the story behind the scene, which is then presented to the class. Supposedly, the class consists of twelve students, which makes three groups; each group presents the story from a different viewpoint: there are a first-person narration from the viewpoint of Spike, a third-person narration from an omniscient narrator’s point of view and a dialogue between two or more reporters. The presentation can be done freely as an improvised role play (e.g. between the reporters), or it can be recited, depending on the type of personalities, outgoing or introverted, the students are.

After the presentations, the teacher gives the students the context of the scene in oral form: The reporters are actually waiting for the film star, Anna, who is intimately involved with Spike’s roommate, William, rather than expecting Spike, who is a hopeless case of dumbness, having been told not to leave the apartment, and who - by the way - is the real cause for the chaos outside, because he has spread the word in a pub that William is having an affair with Anna, which is followed by Anna leaving William.

The next activity involves detailed AVC. The students are instructed to understand the second part of the scene (when Spike poses in front of the mirror) in its details, i.e. they have to note down the script. The groups stay in their previous formations (three groups). They are allowed to watch the scene as often as they need to. Assisting the students in finding the supposedly
unknown words, a help sheet is provided by the teacher (see appendix). Furthermore, the groups can also use the internet to search the meaning of the unknown words (here the students can use their knowledge of English sounds and how they translate into written form). In this way, the overall cognitive load is reduced, as the provided aids ease comprehension.

Later, the script is noted down on an overhead transparency by the teacher in a TST. Finding the exact words does not only have the positive side effect of learning new vocabulary in use, but also lays the groundwork for the subsequent speaking activity, in which the students have to come up with an alternative script in oral form in pair work. Each pair can pick one of the following two choices of how Spike reacts differently to the press than in the film scene.

- first choice: Spike reacts aggressively, he locks horns with one of the reporters
- second choice: Spike reacts frightened and intimidated

If time permits, all of the alternative scripts are presented to the class; in case of limited time, the groups present their results to the neighboring group and vice versa.
Conclusion

I have set myself the goal to establish AVC as an official skill next to listening, reading, speaking and writing, in order to make it an integral part of foreign language teaching. The question as to whether I have been successful can be answered both positively and negatively.

YES, my endeavor has been successful, because I have literally pressed every button that is in my power as an English student to achieve recognition and acceptance of AVC:

Firstly, I have provided an outline of research on AVC from the beginning of its existence in the 1970ies until present day, which shows that knowledge about the processing of audio-visual signals and their subsequent integration into a single mental model is limited, inasmuch as only few studies exist, and insufficient, as the most up-to-date model, Schnotz’s and Bannert’s IMPTC, lacks empirical research. Another major obstacle in AVC research constitutes the various understandings of the concept of AVC. A future step must therefore be to test IMTPC, based on which more accurate models are developed within the framework of empirical research, along with clear distinctions between the different concepts of AVC. Secondly, I have taken the best ideas which I encountered during my study of AVC knowledge and collated and integrated them into a model that represents how I see AVC. Thirdly, I have taken my model as a basis to systematically analyze the Austrian curricula in regard to their views of AVC. This examination has shown that both frameworks do not regard AVC as a skill. Whereas the curriculum for lower AHS-secondary school does not prescribe the use of audio-visual texts at all, the curriculum for upper AHS-secondary postulates AVC in an indirect way, by demanding the understanding of films and TV shows. This assignment constitutes a misinformation of the curriculum which is deficient, as it cannot conform to its own requirements. Thus, both curricula need some modifications which also involve the inclusion of the “missing skill” AVC. Finally, I have provided practical examples of how AVC can be applied in foreign language teaching. The four different teaching designs, in which AVC is either the ultimate goal of the lesson...
or an instrument to achieve other learning objectives, constitute my strongest argument for AVC, as they explicitly show what interesting, meaningful and fun activities AVC and film provide.

NO, my endeavor has not been successful in view of the fact that it is not in my power as an English student to empirically investigate audio-visual processing and rewrite the curricula, which has to be done by specialists, researchers and curriculum designers respectively.

Whether these two major attempts are realized in the near future and eventually succeed in an appreciation and acceptation of AVC, i.e. its establishment within Austrian teaching standards, is out of my control from this moment on. I have taken the first step by drawing the attention to AVC as a skill, discussing it within the framework of an academic discipline and widening the view of the use of film in foreign language teaching, which brings me back to the optimism I expressed before, with which I now want to close.

_I hereby declare that the information on which my work is based has been collected by me personally and has not been plagiarized from any unacknowledged sources. I have properly credited the source of any and all quoted or paraphrased material._
BIBLIOGRAPHY

Primary Sources:


Secondary Sources:


----- 2006. „Film/Video/DVD in Deutsch als Fremd- und Zweitsprache. Bestandsaufnahme und Perspektiven.“ In: *Umbrüche (= Materialien Deutsch als Fremdsprache)* 76. 309-328.


Elliott, Kamilla. 2003. Rethinking the Novel/Film Debate. Cambridge: CUP.


Hicks, Deborah. 1996. „Discourse, Learning and Teaching.“ In: Review of Research in Education 21. 49-95.


Muller, D. A. at el. 2008. "Coherence or interest: Which is most important in online multimedia learning?". In: Australasian Journal of Educational Technology 24 (2). 211–221.


**Internet sources:**

<http://www.bmukk.gv.at/medienpool/11854/lebendefremdsprache_ost_neu0.pdf>


*Common European Framework of Reference*. Council of Europe. 24 June 2009  

Zeitter, Ernst. *Denn auf die Mischung kommt es an - DER TONFILM*. 29 June 2009  
<http://www.mediacultureonline.de/fileadmin/bibliothek/zeitter_tonfilm/zeitter_tonfilm.html>

<http://www.unm.edu/~moreno/PDFS/chi.pdf>

<http://images.google.at/images?hl=de&q=cry%20freedom&um=1&ie=UTF-8&sa=N&tab=wi>


**Films:**

Alice in Wonderland, USA 1951, 75 minutes, Production: Walt Disney

Cry Freedom, Great Britain 1987, 157 minutes, Production: Richard Attenborough

Notting Hill, USA and Great Britain 1999, 119 minutes, Production: Tim Bevan, Richard Curtis, Eric Feller and Duncan Kenworthy

Appendix
# LESSON PLAN 1

**AIW: AVC as an instrument for teaching beginners**

<table>
<thead>
<tr>
<th>Time</th>
<th>Content/Activity</th>
<th>Material</th>
<th>Social form</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.00-8.05</td>
<td>Pre-viewing activity: activating knowledge about the English alphabet by passing round a ball</td>
<td>Ball</td>
<td>Student-student</td>
</tr>
<tr>
<td>8.05-8.10</td>
<td>Watching the scene twice, while-viewing activity: what letters occur?</td>
<td>DVD “smoking caterpillar scene” (chapter 14, 30:33 - 31:25 min.), DVD player, screen</td>
<td>Individual work</td>
</tr>
<tr>
<td>8.10-8.15</td>
<td>Discussion of the results, drawing of a list of consonants and vowels</td>
<td>Blackboard or whiteboard</td>
<td>TST</td>
</tr>
<tr>
<td>8.15-8.20</td>
<td>Showing the film scene, while viewing activity: what is Alice asked? The teacher demonstrates the use of the phrase Students ask each other their names</td>
<td>See above</td>
<td>Individual work</td>
</tr>
<tr>
<td>8.20-8.30</td>
<td>Demonstration of the use of the phrase “He/she is...”, forming of circles of five, task: asking names, answering and talking about the other person's identity (“He/she is...”)</td>
<td>None</td>
<td>Student-student (groups of 5)</td>
</tr>
<tr>
<td>8.30-8.40</td>
<td>Students fill out the missing information on the work sheet by asking the identities of their fellow students (which is on the back of the card)</td>
<td>Laminated cards</td>
<td>Individual work within a group activity</td>
</tr>
<tr>
<td>Time</td>
<td>Activity</td>
<td>Equipment</td>
<td>Location</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------</td>
<td>----------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>8.40-8.50</td>
<td>Oral discussion of the results</td>
<td>Work sheet, overhead projector, overhead transparency</td>
<td>TST</td>
</tr>
</tbody>
</table>
**LESSON PLAN 2:**

**AIW: AVC as learning goal**

<table>
<thead>
<tr>
<th>Time (BLOCK PERIOD: 8.00-10.30)</th>
<th>Content/Activity</th>
<th>Material</th>
<th>Social form</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.00-8.10</td>
<td>Revision of the content of chapter 5, re-activation of knowledge</td>
<td>Book “AAIW”, chapter 5 “Advice from a Caterpillar” (53-65), blackboard</td>
<td>TST</td>
</tr>
<tr>
<td>8.10-8.20</td>
<td>Reading of the poem by the teacher, discussion of its content in groups of two</td>
<td>Poem “How doth the little crocodile” from AAIW, chapter 5</td>
<td>Student-student</td>
</tr>
<tr>
<td>8.20-8.30</td>
<td>Second reading of the poem, students note down two words they want to know, new group formations and discussion</td>
<td>See above, sheet of paper</td>
<td>Student-student (new formation)</td>
</tr>
<tr>
<td>8.30-8.40</td>
<td>“Gemeinsames lebendes Wörterbuch”: discussion of the unknown words</td>
<td>Blackboard</td>
<td>TST</td>
</tr>
<tr>
<td>8.40-8.50</td>
<td>Third reading of the poem, information exchange (groups of two), last reading of the poem</td>
<td>Poem</td>
<td>Student-student (new formation), teacher</td>
</tr>
<tr>
<td>8.50-9.05</td>
<td>Orientating AVC, taking notes (first impressions), presentation of the neighbor’s impressions</td>
<td>DVD “Smoking caterpillar scene”, DVD-player, screen, commentary sheet 1</td>
<td>Individual work, group work (groups of 5)</td>
</tr>
<tr>
<td>Time</td>
<td>Activity</td>
<td>Materials/Tools</td>
<td>Method</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>9.05-9.30</td>
<td>Second watching of the film scene (Selective AVC), group presentation through a group leader</td>
<td>DVD “Smoking caterpillar scene”, DVD-player, screen, commentary sheet 2</td>
<td>Group work</td>
</tr>
<tr>
<td>9.30-9.40</td>
<td>Discussion of the results (optional: discussion of the results in written form -&gt;homework)</td>
<td>none</td>
<td>TST</td>
</tr>
<tr>
<td>9.40-9.50</td>
<td><strong>BREAK</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.50-9.55</td>
<td>Eliciting of the various meanings of smoke</td>
<td>Overhead transparency: “smoke sheet”</td>
<td>TST</td>
</tr>
<tr>
<td>9.55-10.00</td>
<td>Filling the cloud of smoke with characteristics</td>
<td>Overhead transparency displaying an empty cloud of smoke</td>
<td>TST</td>
</tr>
<tr>
<td>10.00-10.10</td>
<td>Answering the questions of how “unreal”, “unsubstantial” and “transitory” are related to smoke and to Alice’s story</td>
<td>Overhead transparency displaying the three questions</td>
<td>Group work (groups of 3)</td>
</tr>
<tr>
<td>10.10-10.15</td>
<td>Discussion of the results</td>
<td>none</td>
<td>TST</td>
</tr>
<tr>
<td>10.15-10.20</td>
<td>Question: where are mirrors and smoke found together?</td>
<td>none</td>
<td>TST</td>
</tr>
<tr>
<td>10.20-10.25</td>
<td>Question: how does deception relate to the use of smoke images in the film?</td>
<td>none</td>
<td>Group work</td>
</tr>
<tr>
<td>10.25-10.30</td>
<td>Discussion of the students’ ideas</td>
<td>none</td>
<td>TST</td>
</tr>
</tbody>
</table>
Teacher explains the differences in meaning between the German phrase “Schall und Rauch” and “It is all smoke and mirrors”

Homework:
The groups are supposed to find three signifier/signified-relations realized as word/picture-relations in the film AIW.
**Commentary sheet 1:**

<table>
<thead>
<tr>
<th>Characters</th>
<th>Text</th>
<th>Filmic realization</th>
<th>Poem</th>
</tr>
</thead>
<tbody>
<tr>
<td>(how are they portrayed? What character traits do they have?...)</td>
<td>(vocabulary, accent, stress … what sounds are used?)</td>
<td>(how is it done? what camera angle is employed?)</td>
<td>(Does/how does the poem change?..)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Commentary sheet 2:

<table>
<thead>
<tr>
<th>Relationship words and pictures</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the text?</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

What effects are created?

How does the filmic realization differ from Carroll’s poem?

Other comments:
Smoke sheet:

SMOKE
# LESSON PLAN 3:

Seeing, hearing and feeling South African history: Teaching power relations during Apartheid through AVC

<table>
<thead>
<tr>
<th>Time (BLOCK PERIOD: 8.00-9.40)</th>
<th>Content/Activity</th>
<th>Material</th>
<th>Social form</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.00-8.15</td>
<td>Warming-up activity: Finding the text to the silent trailer</td>
<td>DVD “Cry Freedom”: trailer (main menu) , DVD-player, screen, overhead transparency</td>
<td>Group work (4 people)</td>
</tr>
<tr>
<td>8.15-8.20</td>
<td>Presentation of the group results</td>
<td>See above</td>
<td>Group work, TST</td>
</tr>
<tr>
<td>8.20-8.25</td>
<td>Watching the scene “De Wet vs. Biko” twice (Orientating AVC)</td>
<td>See above</td>
<td>TST</td>
</tr>
<tr>
<td>8.25-8.45</td>
<td>Analysis of the two frames and presentation of the teacher’s arguments including edited frames (Power Point)</td>
<td>Laptop, screen, Power Point presentation, “Discussion sheet”</td>
<td>TST, Students note down independently</td>
</tr>
<tr>
<td>8.45-9.10</td>
<td>Analysis of the scene in regard to power relations and film techniques, preparation of the Power Point presentation</td>
<td>Laptops, DVD, DVD program “Power DVD”, Power Point</td>
<td>Group work</td>
</tr>
<tr>
<td>9.10-9.20</td>
<td>Presentations of the analyses(^{86})</td>
<td>See above</td>
<td>Group presentations</td>
</tr>
</tbody>
</table>
| 9.20-9.40                      | Discussion of the most | Overhead | Student-

\(^{86}\) The teacher collects the presentations via email, notes down the most important points and hands out a copy (Word document) to each student in the next lesson.
<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
<th>Transparency with the movie posters</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.35-9.40</td>
<td>Comparing students’ decisions</td>
<td>See above</td>
<td>TST</td>
</tr>
</tbody>
</table>
Discussion Sheet (it has to be filled out during the teacher’s explanations):

What different types of inferiority/superiority are reflected in the film scene in regard to the two frames?

- 
- 
- 

How are types of power expressed visually?

(Please draw on the frames to indicate types of inferiority/superiority.)

Remember the film scene. How does the film technique “sound” represent types of power?
Help sheet:

FILM TECHNIQUES

STRUCTURE

protagonist almost always the major character of a drama, the "hero" who seems to initiate the action or is most affected by it

antagonist any character or force opposing the protagonist’s desires, making for difficulties or dramatic conflict; in U.S. films usually, though not always, a "bad guy"

climax that moment or scene near the end of a dramatic action when conflicting characters or forces confront each other

closure a sense of finality felt as the action ends, with the plot problem resolved, other issues adequately settled, and the outcome of the characters’ supposed lives thereafter sufficiently predictable

FILMING

shot a take, the film from a single continuous, uninterrupted run of the camera.

sequence the spliced shots and scenes making up a single significant dramatic unit

cut the spliced place between two frames where one shot ends abruptly and another begins; also the director’s call to the crew to shop shooting.

PHOTOGRAPHY incl. ANGLES, SHOTS

close-up a shot filling the screen with the image of any object the size of a human face or smaller, generating strong viewer attentiveness and feelings of intimacy

medium shot a shot made with the camera seemingly near what it sees but not close to it, showing a human figure from the waist up

long shot a shot made with the camera some distance from the object viewed, showing at least a human form fully visible (head to toe) within the frame but sometimes showing a wide panorama seen by a camera even farther away

high-angle shot a shot made with the camera looking down, as if superior to what it sees.
**low-angle shot** a shot made with the camera looking up, as if the viewer were awed or cowed by what is seen

**zoom** a lens which allows the viewer to seem to move closer or further away from an object without the camera actually moving

**LIGHTING**

**contrast** (of image) grades of light and dark.

**back light** lights illuminating the main image from the rear, sculpting it from the background with highlighted edges, as with haloed hair

**key light** lighting which selectively illuminates from the front various prominent features of the image, such as faces or hands

**high-key lighting** lighting style in which all parts of the set and the screen are relatively evenly lit, suggesting a familiar world containing few surprises or mysteries

**low-key lighting** lighting with strongly contrasted areas of light and shadow, often with one feature of the image lit from one side or below and the rest dark, creating a sense of lurking mystery

**SOUND**

**ambient sound** "live" background sounds creating the illusion that we are seeing and hearing a real world, such as the sounds of distant birds or cars

**wild (live) sound** actually recorded while the shot is made

**background music** off-screen, extra-diegetic (= outside the "reality" of the film story) music heard during the film, not originating within the action but accompanying that action in order to heighten its dramatic power.

**sound effects** the sounds mixed onto the sound track, created to "accompany" various images as if they were originated by those images.
# LESSON PLAN 4:

**How to get students to talk: Using a video clip**

<table>
<thead>
<tr>
<th>Time (BLOCK PERIOD: 8.00-9.40)</th>
<th>Content/Activity</th>
<th>Material</th>
<th>Social form</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.00-8.15</td>
<td>Visual dictation, comparison of the drawings with the frames</td>
<td>laptop with DVD drive, two separate word files with the two frames, blank sheets for drawing</td>
<td>Student - student</td>
</tr>
<tr>
<td>8.15-8.20</td>
<td>Talk: answering the questions, “How are the two pictures related?” and “What happens between the two pictures?”</td>
<td>none</td>
<td>Group work: two pairs from a new group (4 people)</td>
</tr>
<tr>
<td>8.20-8.30</td>
<td>Discussion of the ideas</td>
<td>none</td>
<td>TST</td>
</tr>
</tbody>
</table>
| 8.30-8.50                      | Orientating AVC: Watching the clip twice  
Discussion of the context of the scene: answering the questions on the overhead transparency through telling a narration from a specific viewpoint (3 viewpoints) | DVD scene “Posing Spike”, laptop, overhead transparency | Groups (4 people) |
<p>| 8.50-9.00                      | Presentations: narrations (role plays or recitations) | none | Group presentations |
| 9.00-9.05                      | Teacher gives the context of the video clip | none | Teacher |
| 9.05-9.15                      | Detailed AVC: finding the script | Help sheet, DVD, laptop, if available: | TST |</p>
<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
<th>Methodology</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.15-9.20</td>
<td>Discussion of the script</td>
<td>Overhead transparency</td>
<td>TST</td>
</tr>
<tr>
<td>9.20-9.35</td>
<td>Finding an alternative script (2 choices)</td>
<td>none</td>
<td>Student-student</td>
</tr>
<tr>
<td>9.35-9.40</td>
<td>Class presentation or presentation among the pairs</td>
<td>none</td>
<td>TST or pairs present to each other</td>
</tr>
</tbody>
</table>
Help sheet:

Words which Spike uses at some point of the clip:

- [Image of legs]
- [Image of chicken and egg]
- [Image of underwear]

The images seem to represent words or concepts used in the clip.
German abstract

Die vorliegende Arbeit beschäftigt sich mit der rezeptiven Fertigkeit „Hör-Seh-Verstehen“. Darunter versteht sich die Fähigkeit audio-visuelle Reize, die durch den visuellen Kanal (das Auge) und den auditiven Kanal (das Ohr) aufgenommen werden, zu verarbeiten und die enthaltenen Informationen durch verschiedene kognitive Verarbeitungsabläufe wie Top-Down und Bottom-Up-Prozesse zu entschlüsseln, um schließlich eine innere Repräsentation, ein mentales Modell, zu bilden, mit dessen Hilfe die Quelle der audio-visuellen Reize, der Text, verstanden werden kann. Im Genaueren handelt es sich um filmische Texte im Kontext des Fremdsprachenunterrichts der Unter- und Oberstufe österreichischer Gymnasien. Eine Untersuchung der Lehrpläne, die den strukturellen und thematischen Orientierungsrahmen für den Fremdsprachenunterricht darstellen und diesen somit legitimieren, zeigt, dass Hör-Seh-Verstehen nicht als fünfte Fertigkeit neben Hören, Lesen, Schreiben und Sprechen anerkannt wird. Da die Lehrpläne das Verstehen filmischer Texte und Fernsehprogramme voraussetzen, es jedoch fälschlicherweise der Fertigkeit Hören zuordnen, muss Hör-Seh-Verstehen als fünfte Fertigkeit in die Lehrpläne aufgenommen und folglich im Unterricht umgesetzt werden. Diese Arbeit stellt den Versuch dar, Hör-Seh-Verstehen als offizielle fünfte Fertigkeit zu etablieren und in den Fremdsprachenunterricht zu integrieren.

Curriculum vitae

Petra Ederer
Geb. 13.06.1985
1090 Wien, Österreich
Gussenbauergasse 2/18
+43 (0) 650 561 6464
Petra.Ederer@gmx.net

Ausbildung

Studium English und Geschichte (Lehramt) 03/2004 –
Universität Wien/Ö

Deutsch als Fremd- und Zweitsprache (Modul) 10/2006 – 06/2009
Universität Wien/Ö

Gymnasium 09/1995 – 06/2003
Bundesoberstufenrealgymnasium in Eisenstadt/Ö

Auslandserfahrungen

Praktikum am Institut für Germanistik 09/2009 – 01/2010
Universität Soka/JP
im Rahmen des Deutsch als Fremdsprache-Auslandspraktikums

Studium Anglistik 08/2007 – 01/2008
Universität Maryland/USA
im Rahmen des „Joint Study“-Programms

Au-Pair 07/2006 – 09/2006
in Irvine, Kalifornien/USA

Berufliche Erfahrungen

Deutschlehrerin am Sprachinstitut „Cultura Wien“ 02/2009 –
Einzel- und Gruppenunterricht für alle Niveaus

im Zuge des Moduls Deutsch als Fremdsprache

Englischlehrerin bei „English for Kids“ 04/2008
in Wien und Linz/Ö
07/2008 – 08/2008
04/2009