Pedagogic Videos as a Foreign Language Learning Resource in Textbooks used in the German Studies Section of a South African University: A Digital Multimodal Discourse Perspective

A thesis submitted in
fulfilment of the
requirements for the
degree of
MASTER OF ARTS
of
RHODES
UNIVERSITY

2020

Sasha-Lee Schafli
Student number: g15S3552
Supervisors: Dr. Ian Siebörger & Dr. Undine S. Weber
Department: Linguistics and Applied Language Studies

Abstract

Video is a prominent teaching and learning tool within foreign language (FL) textbook media packages in the 21st century. While studies undertaken in the Global North highlight that video materials in the FL classroom have the potential to influence learning and cultural knowledge acquisition, there is a lack of research on the manner in which pedagogically designed videos influence adult FL learning and cultural knowledge acquisition in a South African context. In this study, I explore the opportunities and challenges in terms of language learning and cultural knowledge acquisition that arise from three pedagogic videos in the *Menschen A1* textbook which is used in teaching students registered for the German Studies 1 course at Rhodes University. I compare and contrast two sets of data to examine the relationship between pedagogic video and student knowledge acquisition: the results of a digital multimodal discourse analysis (DMDA) of these videos, and questionnaires and transcriptions collected from semi-structured group interviews with German Studies 1 students. These questionnaires and transcriptions were analysed thematically. Findings in terms of the language learning experience indicate that actor over-exaggeration and visual aids assist students when learning German at this level with this type of video. However, these visual aids can be distracting and confusing without balanced representation and contextual information. Differences between videos in terms of speech rate seem to affect students' perceptions of their ability to comprehend the videos. Students reported forming only positive impressions of German culture(s) on the basis of the videos. In general, students find Germans portrayed as friendly and helpful in the videos. The results of this investigation provide recommendations for the optimal use of this type of teaching and learning resource, for example, teachers/ lecturers/ facilitators should allow for focus group discussions on cultural discourse to occur in order to balance stereotype formation and should consider the speech rate of videos for language learning.

Table of Contents

List of Abbreviations

PCSTIF – The Parallel Constraint Satisfaction Theory of Impression Formation
SFL – Systemic Functional Linguistics
SF-MDA – Systemic Functional Multimodal Discourse Analysis
DMDA – Digital Multimodal Discourse Analysis
GFL – German as a foreign language
FL – Foreign language
CALL – Computer-assisted language learning
STM – State Transition Machine
CEFR - Common European Framework of Reference for Languages
EFL – English as a foreign language

List of Figures

Acknowledgements

I would above all like to thank my supervisors, Dr. Ian Siebörger and Dr. Undine Weber for their support, their knowledge and their patience, and for allowing me the opportunity to create the research I have waited so long to produce (for four years). The dream team we were and both supervisors assisted me at every meeting, putting up with my "Sashaisms" and helping me graft this finished product. My respect for them cannot be described in words and this dissertation would not be what it is today without their assistance. I am immensely grateful for their time and effort and it has truly been an honour to work with them. Your support and prayers got me through this thesis. I am indebted to you both.

I was blessed to get the outside assistance of academics that are experts in the field of multimodality. Prof. Kay O'Halloran, a thousand thanks to you for all your time and effort over email to assist me throughout the annotation stages of this project. I would also like to thank Dr. Fei Lim for taking the time to assist me on intersemiotic texture. You both were invaluable to this research and it has been amazing to work with you both. Dr. Zsuzsanna Abrams, thank you ever so much for your assistance in providing me with detail about your paper that was an invaluable model for my own research. Prof. Arlene Archer, such a pleasure to correspond with a fellow South African carrying the torch for multimodality and I thank you for all of your assistance and input. Prof. Uwe Petersen, vielen herzlichen Dank für Ihre Hilfe und Ihre Bemühungen, ich weiß es wirklich zu schätzen. It has been an honour to correspond with you all and I truly hope I am blessed to work with you all again in future.

I would further like to extend my deepest gratitude to the entire Linguistics and Applied Language Studies Department for their guidance and taking my research ideas seriously from the very beginning (2016). There are too many to mention by name, but I wish to thank them all for their advice and for allowing me to use the staff and postgraduate tearoom in my third year. The discussions that took place in that room provided me with a research space, in which I could grow as a researcher and this dissertation would also not be here without it. Similar thanks are extended to the German Studies section for the conversations we shared that allowed this budding flower of research to fully bloom.

To Malcolm Hacksley, my heartfelt appreciation for taking the time out to assist with editing this thesis. Erin Alexander, my friend, this thesis would not be where it is without you. You were one of my day one people – listening to the idea, watching it take fruit. In the hardest year that was 2020, you were there to be a cheerleader, editor, supporter and I am honoured to have you as a friend. This thesis would not stand without you and you have my eternal gratitude. Reverend Boitumelo Gaborone, thank you for being a lighthouse, for your prayers, for your support and for listening to many parts of this thesis being read aloud! Again, an invaluable person to ensuring this thesis sees the light of day. Many other friends have left their marks on this thesis – too many to fit on this page, but you all know who you are, and I thank you from the bottom of my heart for your support.

Last but not least, my undying love and gratitude goes to my husband and rock, Tyler Schafli, for his support, for his never-ending belief in me, for his love, for lending an ear to all my complaints and also for listening to this thesis being read aloud! A big thank you also goes to my parents and to my Heavenly Father for giving me the strength to survive through this degree and produce this dissertation. All the glory goes to Him.

Chapter 1 – Introduction

Formal language learning contexts in the foreign language (FL) classroom in South Africa is often not authentic enough (Ferreira-Meyers and Horne 2017:27-28). Furthermore, "more emphasis on the experience of learning a language is required in [FL] teaching" in South Africa (Horne 2019). This research seeks to describe and analyse the opportunities and challenges of students' experiences when learning German as a Foreign Language (GFL) at beginner level with pedagogic video clips at Rhodes University in South Africa. Additionally, the way in which culture(s) of German-speaking countries are represented by these videos is analysed, in order to ascertain how a video designed for the classroom portrays such a complex entity to beginner learners of GFL. The impressions that are gained by students and assumptions that students make about these cultures based on their viewing of the video clips are also examined qualitatively in this research. This thesis is consequently situated within the research areas of linguistics, German Studies and teaching and learning of a foreign language in South Africa in Higher Education. It results in recommendations to teachers/ lecturers/ facilitators as to how to fully realise the potential of these videos for adult students within the GFL classroom in South Africa.

This chapter outlines the research and its rationale. I introduce the motivation for the study in 1.1., which is informed by the context of this research in 1.2. By examining the relationship between the German language and culture(s) and South Africa, 1.2. seeks to concretise the rationale for this research, as well as to understand the context of teaching the language and cultures in South Africa, which is an important focus of this research. Firstly, I discuss the decline of German Studies in South African Higher Education. I report on the manner GFL is taught in South African Higher Education in order to understand the pedagogic side of the GFL learning context. This is done with a concise examination of South African education policies regarding the teaching of GFL to ascertain the recommendations for teaching German in South Africa, even if it is not at a level of Higher Education. This provides an understanding of government policy and incorporates discussions of these policies with discussions by academics on actual practice in the GFL classroom. Finally, a brief context of FL teaching methods concerning video is provided. This is followed by a discussion of key concepts for this thesis in 1.3., specifically pedagogic videos versus naturalistic or authentic videos and multimodality. The research goals and questions are discussed in 1.4. The last section concludes this chapter and provides a structure of the thesis.

Section 1.1. Motivation for the study

The motivation for this study can be described in terms of three gaps in the research:

1. There remains a lack of research into the opportunities and challenges of learning this language and acquiring knowledge on culture(s) of German-speaking countries at this level of education. Of all the literature reviewed in 2.1., none examine German cultural representation in pedagogic videos in a South African context.
2. While video has been an important part of the development of language learning and teaching worldwide, it is unknown as to the opportunities and challenges specifically with this language learning resource in South African Higher Education. For example, Diab et al. (2016:11) examine medical students' perceptions of pedagogic video in teaching language skills and cultural competence in isiZulu, one of South Africa's indigenous languages. This research was not undertaken with a multimodal approach such as DMDA, nor does it look at GFL. The closest study found from an African context that examines GFL and pedagogic videos was by Wahba and Zappen-Thompson (1994). This study, however, was undertaken in Namibia, not South Africa and it did not analyse the videos with a multimodal analysis. Further discussions of the literature that clarify this gap are to be found in 2.1. and 2.2.
3. Despite the fact that video is a multimodal resource, it has not been examined multimodally with regard to GFL in South African Higher Education before, nor examined in this manner and triangulated with data from student experiences in a South African context. For example, Abrams (2016) examines videos designed for a mother-tongue audience and uses a multimodal framework to discuss videos and to interview students. However, this is in an American context and not in a South African context. Further literature highlighting this gap is discussed in 2.3. and multimodality is contextualised in 1.3.

Section 1.2. Research Context

This section contextualises the first two gaps in 1.1. concerning GFL teaching policy and practice. Despite German investments in South African Higher Education, German Studies sections at South African universities (which usually include the study of the German language, Cultural Studies or *Landeskunde* and Literary Studies) specifically are on the

decline with further cuts anticipated. In 2012, 13 universities offered a German Studies course (Bauer 2016:617). Only eight universities offering a German Studies course remain (Auswärtiges Amt 2020:17), indicating a steady decline. Only two universities discuss foreign languages in their language policy, namely The University of the Witwatersrand and Rhodes University. The University of the Witwatersrand language policy was adopted in 2003 and has not been changed since. They claim to support multilingualism by offering a university major in an FL, French, as stated in their policy:

> As the provision of foreign language instruction is costly, the University will negotiate the provision of foreign language instruction in the region, with Wits negotiating for French as its focus. The choice of French is based on the language research conducted at the University. In addition, French is an official language of the African Union and it is a language widely required for qualifications in science degrees internationally (University of Witwatersrand Johannesburg 2003:4).

Rhodes University, the university at which this research is situated in South Africa, states in its language policy that it will promote the study of foreign languages (Rhodes University Language Policy 2019: n.p).

According to Bauer (2016:621), German academics need to be convinced of the opportunities afforded by Global Southern Research contributions towards a language that is inherently from Germany. Furthermore, prejudices against the Global South in relation to research exist in the manifestation that former financial investment in this geographical area is always seen as financial aid or assistance. This is rather than as a facilitation for intercultural relations as South Africa in particular wishes to see it today (Bauer 2016:621). The DAAD and the British Council note a dramatic shift of focus on applied sciences for the labour market in South African Higher Education, with government aiming for the majority of doctoral theses to be in these areas by 2030 (Herman and Sehoole 2014:10). Financial need has decreased the popularity of humanities research by students, due to the increasing need for education to secure a particular profession (Lange 2017:37-38).

The South African 2002 Language Policy recognised that "the study of foreign languages is also under threat, with declining enrolments in most language programmes" (Ministry of Education 2002:8), but this was omitted in the revised 2017 policy (Ministry of Education 2017:17). Another reason for the declining enrolment in foreign language programmes is that South Africa has 11 official languages (The Constitution of South Africa 1996:4), which has resulted in an "'overpopulated' language curriculum" (Ferreira-Meyers and Horne 2017:25). The 2017 South African policy for language in Higher Education, however, states that

institutions are to implement their own programmes, policies and strategies to encourage foreign language study. It focuses particularly on "languages that are important for the promotion of the country's cultural, trade and diplomatic relations" (Ministry of Education 2017:17). While a list of foreign languages is not mentioned in this revised version of the policy, the former policy of 2002 had listed German as one of these languages (Ministry of Education 2002:14). The 2002 policy stated that the approach in developing foreign language would be similar to that taken to develop South African languages (Ministry of Education 2002:14). This has also, however, been omitted in the new policy which has shifted its multilingual focus from the old policy, which they describe as having failed, to the development of a multilingualism that focuses on indigenous language development (Ministry of Education 2017:2; 11). The 2017 policy thus de-emphasises foreign languages in order to give more emphasis to indigenous languages.

The omission of foreign languages within the revised language policy indicates the reality of foreign languages in South Africa. FL teaching and learning in South African universities is side lined, as poor language competencies in local languages and the desire to address the underdevelopment of the local languages after apartheid has caused priority to be given to the teaching of these languages (von Maltzan 2009:208-210). Von Maltzan (2009:210) examines the challenges GFL in universities faced regarding curriculum, stating that course contents and learning objectives had to be redefined within the context of an outcomes-based education. This had stemmed from a general curriculum for GFL by the Ministry of Education (2002) originally put together for schools but adopted by universities. There is an ambiguity over areas of responsibility regarding curriculum planning for teachers and lecturers. Furthermore, there is an overall lack of communication between the Department of Education and examining authorities for GFL in a South African Higher Education context (von Maltzan 2009:210). Thus, there is still ambiguity in GFL in South African Higher Education. In addition, the lack of communication between the Department of Education and South African Higher Education institutions has not been remedied. Auswärtiges Amt (2020:17) indicates that there were 10 668 German learners in 2020, with 909 of them studying at universities. This represents a 38% decline in German university students since 2015 (Auswärtiges Amt 2020:17). Annas (2016), while being optimistic about the continued hiring of German lecturers and steady numbers of German students at universities, states the main challenges of German sections at university are to encourage black students to study German, as well as students to continue studying German through to a postgraduate degree.

One can conclude that there are systemic problems maintaining and integrating the teaching and learning of GFL within educational policies in Higher Education.

German courses are assessed and usually designed in Europe and internationally based on the *Common European Framework of Reference for Languages* (CEFR). It is "a framework of reference [...] designed to provide a transparent, coherent and comprehensive basis for the elaboration of language syllabuses and curriculum guidelines, the design of teaching and learning materials, and the assessment of foreign language proficiency" (Council of Europe 2019). Since there is an international CEFR for learning and teaching this language, German curricula in South African Higher Education institutions generally follow the CEFR structure (see, for example, University of Pretoria 2019). This, however, works slightly differently in schools that have *Curriculum Assessment Policy Statements* (CAPS) for German as a Second Additional Language (SAL) for the primary level of education. CAPS have an extremely detailed description of how German as an SAL i.e. a language that students may not have knowledge of when arriving at school (Department of Basic Education 2012:10) must be taught in South Africa. It follows four language skills similar to the CEFR (Department of Basic Education 2012:10):

- Listening and Speaking.
- Reading and Viewing.
- Writing and Presenting.
- Language Structures and Conventions.

Since the focus of this research is not on German taught in schools, the CAPS document is not discussed in great detail. However, there are a few points to note in the CAPS document pertaining to curricula that is worth noting for this research, and key concepts within it are discussed in 1.3.

This outline of the South African context for the teaching of GFL has illustrated the current crisis that faces foreign languages and GFL in particular. There are problems such as, finding a space in an "'overpopulated' language curriculum" (Ferreira-Meyers and Horne 2017:25), government policies, and the lack of opportunities in the country to learn GFL at tertiary education institutions. Furthermore, as stated earlier regarding the CEFR, a clearly defined framework allows students' linguistic competence to be compared internationally. Government policies regarding curriculum are only being addressed at a level of education

for South African schools. The 13 universities do not share their curricula online in the form of what materials they use to teach on public websites in relation to the contents of their individually designed language courses. This research focuses on the German Studies 1 course at Rhodes University, a course which corresponds to the A1 level of the CEFR. A student at A1 level competency is described as able to:

> understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help. (Council of Europe 2019)

Language learning difficulties for students of German in South Africa also remain relatively unexplored. While research in a South African context illustrates that literary texts can allow students to critically engage with cultural topics (Weber and Domingo 2011:183; Collins 2017:13**;** Weber 2018:79), there is a further lack of research focusing on student acquisition of cultural knowledge via pedagogic video materials in a South African context. This is dealt with in 2.1. In this thesis, I refer to culture as "the enduring yet evolving intergenerational attitudes, values, beliefs, rituals/customs, and behavioural patterns into which people are born but that are created and maintained by people's ongoing actions" (Spitzberg and Changnon 2009:6-7). Cultural representation, in this thesis, is thus the representation or portrayal of culture, as defined above, in media such as textbooks and video. Theories regarding the acquisition of cultural knowledge regarding impression formation are discussed in 3.1.

Despite the lack of research on pedagogic videos in the FL classroom in South Africa, contexts outside South Africa show that video materials positively influence communication, listening comprehension, interest, motivation and cultural knowledge of students (for example, Herron et al. 1999; Çakir 2006; Abrams 2016). These scholars also discuss the challenges of implementing videos in FL classrooms. This literature is discussed in 2.2 of this thesis in order to illustrate the lack of research on the use of pedagogic videos in a South African context, as well as in a South African context of Higher Education. It is also used to examine the research around this topic that exists from other contexts for comparison. Learning of language in this research is briefly discussed in connection with Bandura's (1989) Social Cognitive Theory of learning. This recognises that learning is a social and cognitive experiential process and consequently provides a definition of learning for this thesis. This and Mayer's (2005) cognitive theory of multimedia learning are discussed in 3.2.,

the latter providing principles for understanding how one can best learn with multimedia resources such as video.

Video as a resource for FL teaching has been used since the early 1960s, and first in the US (Gottschalk 1965:86; Salaberry 2001:41). Video as a medium in the classroom had a minor role in formal grammar-translation didactic methods, mainly assisting in reproducing formal language production in the classroom environment (Byram 2000:398). As the audio-visual teaching method, which was developed in the 1950s became popular; video was more frequently used as a supplementary teaching tool. This again increased as the focus of teaching shifted to building students' communicative competence in the 1970s (Byram and Hu 2013:134). Video and language learning software influenced the shift from didactic approaches that were teacher-centred, to introducing further learner autonomy in the language learning process as media centres and language laboratories were established (Byram 2000:398). Media centres provided more choices in ways of FL learning to FL learners than earlier language laboratories which had formed as a product of the audio-lingual teaching method (Rösler 2010:1202). As computer assisted instruction (CAI) took root in the 1960s in the US, it later developed into computer-assisted language learning (CALL), where students worked with video discs in computer laboratories (Salaberry 2001:44-46). It was in these computer laboratories and media centres that media such as video, language learning software and audio recordings were first packaged together as a unit within study materials to allow students a vast array of options in their individual learning process (Byram 2000:399-400). Despite the integration of technology and the internet in language learning, textbooks with their inclusion of multimedia packages remain the preferred teaching tool in the language classroom (Byram and Hu 2013:721).

Within CAPS, the focus on SALs in education is on building communicative competence (Department of Basic Education 2012:9). Communicative competence may be regarded as a combination of grammatical knowledge, "knowledge of how language is used in social contexts to perform communicative functions, and knowledge of how utterances and communicative functions can be combined according to the principles of discourse" (Canale and Swain 1980:20). Communicative language teaching (CLT) focuses on how to use language as a communicative resource (Byram and Hu 2013:134). CLT has become an umbrella term in which other methods are classified, including humanistic methods, content-based instruction, the natural approach, bilingual methods, task-based language learning

(Byram and Hu 2013:675;711), as well as "interkulturelle Didaktik" (intercultural didactics) (Byram and Hu 2013:352). Consequently, CLT encompasses a variety of instructional methods that, while focusing on communicative competence improvement, does not forsake grammatical knowledge. Dal (2010:4) states that CLT does not always include videos in teaching, but his study argues that when they are in the material, the CLT approach would solve the previously stated problem, by making students active in video production, and not just passive video viewers. CLT includes teaching intercultural aspects, and intercultural learning within GFL teaching has always been present, but it is the responsibility of the teacher/ lecturer/facilitator to present intercultural information effectively (Königs 2003:245). The focus on interculturality in German language learning thus increased from the early 1990s onwards (Königs 2003:243). Video can be an effective tool in enriching intercultural contact and knowledge (Herron et al. 1999:518; Çiftçi 2015:322). Engagement with the videos by students and lecturers in Higher Education, as well as the way in which videos are presented are important aspects of consideration for the implementation of videos for learning in classrooms (Mitra et al. 2010:413). More on this is discussed in 2.1.

Section 1.3. Key Concepts

According to CAPS, learners must "communicate effectively using visual, symbolic and/or language skills in various modes" (Department of Basic Education 2012:4). A mode is defined in CAPS as "method, a way or manner in which something is presented; a way of communicating (e.g. the written mode, the spoken or oral mode, the visual mode [...]" (Department of Basic Education 2012:75). Multimodal texts are defined as texts "mak[ing] use of visual and written material in a single text [,] e.g. advertisements[or] cartoons. They can also combine this with spoken language and gesture" (Department of Basic Education 2012:20). Teachers and learners should have access to audio or visual aids, different authentic texts from different genres and an approved textbook (Department of Basic Education 2012:13). Audio-visual texts mentioned in the syllabus for listening comprehension include "[f]ilms, television programmes and documentaries, slide shows, recordings, radio programmes, photographs, music videos" (Department of Basic Education 2012:15). The aids mentioned above are all based on naturalistic recordings of speech taken from the German-speaking context, such as advertisements and films (Department of Basic Education 2012:14). Multimodal teaching and learning tools that are designed for language pedagogy, such as constructed video scenarios within textbook material or online that are

scripted and recorded especially for FL teaching and learning, are not discussed within the CAPS document. Uhlen (2018:1) categorises video into three categories: film, videos without artistic aims such as interviews and news reports and *Lernvideos*, i.e. pedagogic videos. Therefore, I refer to this category of videos in this research as 'pedagogic videos,' as opposed to naturalistic recordings of speech or excerpts of already existing German-language videos, such as television dramas or documentaries, often labelled as authentic texts (see 2.1.).

The Department of Basic Education has a very basic definition of a mode and multimodal texts, assuming that visual and written are the only modes within a multimodal text, without discussing audio, for example, and does not discuss how modes can contain other modes. Within this thesis, a mode "refer[s] to a socially organised set of semiotic resources for making meaning" (Jewitt et al. 2016:221). Some examples of modes include language (with written or spoken modes within the mode of language) and image (Stöckl 2004:10). According to Jewitt et al., "video is a medium for representing the world in [the modes of] sound and moving image" (2016:267). Multimodal analysis thus illuminates how "people draw on distinctly different sets of resources for meaning making (e.g. gaze, speech and gesture)" (Jewitt et al. 2016:218). FL learning practices have naturally included multimodality, although this has only comparatively recently been recognised as an object of study in educational research (Farias et al. 2011). The ways in which these semiotic choices are presented through the modes within the medium of pedagogic videos, and the influence they have on the students' experience of learning language and on their knowledge of the target culture are not obvious. According to O'Halloran et al. "[f]rom a multimodal perspective, [...] videos fulfil their respective communicative aims and functions through various combinations of semiotic choices in their organisational structure [...,] that is, through the ways in which authors organise, present and orientate the information to their readers" (2015:6).

Due to the study of multimodality arising from multiple disciplines, clarification is needed on exactly the type of discourse analysis I am undertaking in the research. Jewitt (2009:28-31) outlines three main theoretical frameworks used to analyse multimodal semiosis: Systemic Functional Multimodal Discourse Analysis (SF-MDA), social semiotics and Multimodal Critical Discourse Analysis (MCDA). MCDA focuses on ideologies, and social semiotics discusses "macro-political and social/institutional interests" (Jewitt 2009:31), while SF-MDA "approaches 'discourse' at the micro-textual level" (Jewitt 2009:31). SF-MDA thus focuses

on the multimodal artefact itself (Jewitt 2009:35). While the visual components of this research are analysed using SF-MDA, the use of computer software allows SF-MDA to be combined with components of the other systems that are mode-specific, and consequently allows "the micro-level of discourse" to "be related to the macro-level" (Tan et al. 2015:560). Therefore, I am not using an SF-MDA approach alone, but a combination of systems for analysis that Tan et al. call "a digital multimodal discourse perspective" (DMDA) (2015:559). The combination of analytic frameworks deployed in the software used in this thesis allows all the modes to be examined as a whole, while still measuring their contributions and representing them graphically (Jewitt et al.. 2016:88). Examining videos at the micro-textual level allows us a deeper understanding of necessary features for effective FL learning video design. The methods in many studies on multimodal videos in language learning do not take a digital multimodal perspective (cf. Kruger and Doherty 2016:26-28). This research, however, seeks to rectify this by combining the analysis of students with an annotated analysis to track how the modes work together, rather than studying them in isolation. DMDA allows for a smoother integration of modal systems in analysis, thus allowing more modal systems to be analysed with less methodological challenges (see 2.3 later in this thesis). The videos in this research are examined alongside multimodality and literature on this is discussed in 2.3. in greater detail, with the theory used to analyse videos being described in 3.3.

Currently, the German Studies section at Rhodes University in South Africa uses the *Menschen A1* textbooks published by the German publishing company Hueber, in teaching their first-year language course. When video clips are included in FL textbooks, the publisher packages them with other multimedia resources as optional and additional extras (Condray 2012:88). Video clips in *Menschen A1* are linked to every module, yet only the teacher/lecturer/facilitator has access to these video clips on a CD-ROM, while the tasks that accompany them are available in the students' course books. The videos are supposed to offer insight into the everyday life of Germans, Austrians and Swiss people. They are described by the publishing company Hueber as containing entertaining films for the training of auditory and visual comprehension and cultural topics from these countries, with a focus on fostering interactive learning (Kalender and Pude 2012:3; Hueber.de 2019). It is these video clips that this research examines in terms of a multimodal analysis (Digital Multimodal Discourse Analysis or DMDA), which, is discussed in 3.3 and 4.3. of this research.

Kaltenbacher argues that pedagogic videos on English FL CD-ROMs "often deliver 'multimodality' which actually disrupts learning" (2004:119). A lack of gestures with unanimated speech result in FL learners incorrectly assuming that one gesture applies to all words in the FL, and incorrect lip movements could cause misinterpretation, resulting in poor to no language learning (Kaltenbacher 2004:133-134). These negative findings emphasise the need for discovering the influences of multimodal resources on students' experiences of learning with a fine-grained method that investigates how the different modes in videos influence student learning in the areas of language and culture. This research seeks to examine if the underutilisation and misuse of modes could lead to a disruption of learning. The possible challenges within the design of these videos has proved to be the inspiration for this research, along with the lack of research as to pedagogic videos' benefits and challenges in a South African context.

Section 1.4. Goals and Research questions

This first main goal of this research, therefore, is to explore the opportunities and challenges of using video clips from *Menschen A1* as multimodal learning aids when learning GFL at a South African university, at beginner level. By examining the interplay of the semiotic choices and each mode to create meaning and examining this in combination with data about the students' experiences, one can determine how modes can work together to create effective learning in pedagogic video clips. Often learner tasks are based on video resources of naturalistic speech, and previous research on videos has consequently focused on analysing such tasks (Abrams 2016:344). In this study, the focus is *off* the task construction and *on* the video construction, illustrating how video design affects FL learning and giving us insight as to which features in this type of video assist in effective learning. When the modes work with one another to create meaning, this can create opportunities for effective language learning and cultural knowledge acquisition. The types of challenges that are examined in this thesis include finding out whether students find the combination of semiotic choices in various modes that occur simultaneously overwhelming and/ or distracting (i.e. too much is happening in the video for them to keep track of linguistic information). It also looks at whether there is synchronicity in these areas, allowing meaning to be derived (image-and-language correspondence, for example, that may reinforce words students picked up in the video).

The second main goal is to investigate how culture(s) of German-speaking countries are conveyed by these videos in the form of an analysis. Additionally, as it looks at the impressions and assumptions that students make about these cultures based on their viewing of the videos, in order to mediate between my perspective and student perspectives. To achieve these goals, this research seeks to describe and analyse the various modes in pedagogic videos for FL German textbooks, as well as student engagement at university with them. This thesis aims to answer six research questions:

1. What learning opportunities and challenges do Rhodes University students from diverse backgrounds experience in using pedagogic videos as a multimodal learning resource in German Studies 1?
2. What is the relationship between the various modes within these videos? (i.e. speech, as well as modes within the visual elements such as gaze and gesture)
3. How do the relationships between the various modes in these pedagogic videos influence student learning?
4. How are the German-speaking characters represented in these videos with regard to culture(s) of German-speaking countries?
5. What cultural assumptions and impressions do students make based on the above representations?

And therefore:

6. How can the videos be used in a manner that maximises the benefits and minimises the challenges outlined in Question 1?

Section 1.5. Conclusion and Structure of Thesis

Thus, this dissertation fills the gaps in knowledge mentioned in 1.2. regarding the opportunities and challenges of pedagogic videos as a resource in the field of FL teaching and learning, with the implementation of a digital multimodal approach. In this chapter, key concepts in 1.3. and the research goals and questions in 1.4. have been outlined that address these gaps. This research seeks to further elaborate on these gaps in research by contributing to the body of knowledge that is discussed in Chapter 2.

Chapter 2 discusses relevant literature that further highlights the existence of these gaps and contextualises this research within the areas of cultural representation (2.1.), language

learning (2.2.) and multimodality (2.3.). The theory used is provided in a separate chapter, with theories for learning about culture (3.1.), the student learning experience (3.2.) and digital multimodal discourse analysis (3.3.). Theory has been separated from the literature review in order to highlight the differences between these two sections.

The methods and methodology of this research are discussed in Chapter 4 in order to understand how findings were obtained – a research outline (4.1.) and detailed descriptions of data collection (4.2.) and analysis (4.3.) are provided and concluded in (4.4.). Findings are reported in Chapter 5 in terms of the three videos selected (5.1., 5.2. and 5.3.), questionnaires (5.4.) and transcriptions (5.5.). Chapter 5 allows for an analysis in Chapter 6. Chapter 6 integrates the various kinds of data and literature in order to provide an answer to the research questions and is organised according to the three videos (6.1., 6.2. and 6.3.). A conclusion and the answers to the research questions derived from the analysis in Chapter 6 are described in Chapter 7.

Chapter 2 - Literature Review

This chapter organises relevant literature thematically and often chronologically when this perspective is required. As is demonstrated, pedagogic videos seem to be rarely examined on their own and seem to usually be grouped with analyses done on either FL textbooks or CALL software in the GFL and FL classroom. Thus, I firstly review studies undertaken from the 20th century onwards, first focusing on cultural representations in textbooks, secondly on cultural representations that can be found in CALL software and finally cultural representations within the genre of pedagogic videos. I look at the historical development of cultural representation in these areas to allow this research to position the cultural representations within the video clips in the *Menschen A1* textbook alongside those within other pedagogic materials. This kind of review also determines similar or different aspects in the way in which culture is represented for later comparison in Chapter 6. It further seeks a contextual understanding of why culture is represented in this way. This assists in answering research question 4 and understand answers from student data for research question 5. Section 2.2 discusses relevant GFL literature thematically. This is done in order to ascertain the already established opportunities and challenges that video has for GFL learners in other contexts. This is compared to the findings in response to question 1 of this study. It also delves into the chronological development of film didactics for GFL over time to understand the context of GFL material design, as well as choices made by GFL material authors. This should answer research question 2. Section 2.3 discusses the development of DMDA and work that has been undertaken within the medium of video. Due to the lack of DMDA studies on video, the focus of this section references notable literature on multimodality and studies on pedagogic technology. Various alternative frameworks used to analyse the medium of video in language learning contexts are also examined. This builds on the rationale established in Chapter 1 for the use of this theory in this research, as well as to highlight its suitability to discover the answers to question 2.

Section 2.1. Cultural Representations

This section chronologically reviews literature on cultural representation in FL textbooks, then in GFL textbooks specifically. As pedagogic video also often appears in CALL software in GFL and FL classrooms, literature on this is examined, before studies are reviewed that focus solely on video in an FL context.

The earliest example of an analysis of cultural representation in textbooks is that of first-year French textbooks published between 1972 and 1978 (Levno and Pfister 1980:47). Levno and Pfister (1980:47) created a model to examine the surface culture these textbooks depicted and discovered that only two of the twenty-one textbooks examined were considered effective regarding surface culture content and presentation. Levno and Pfister (1980:47) thus considered that many French FL textbooks depicted an unsuitable representation of surface culture. This is not only restricted to French FL textbooks, but other FLs as well. Ramirez and Hall (1990:48) conducted a sociolinguistic and sociocultural study of Spanish linguistic and cultural representation in Spanish FL textbooks used in a Spanish language classroom in the United States. They found Mexico and Spain were given cultural priority as well as certain cultural themes, mostly with regard to social themes and nationality in first-year Spanish textbooks (Ramirez and Hall 1990:49-50). The authors state that readers of the textbook remain disorientated as to what cultural interpretation they should take from it, as photographs of cultural information are present without explanation (Ramirez and Hall 1990:52). The photographs depict the small population of middle- to upper-class Spanish-speaking society, while omitting topics such as everyday cultural events, as well as poverty, which is a reality for the majority of the Spanish-speaking population in the US (Ramirez and Hall 1990:63). Ramirez and Hall (1990:63) postulate that the reasons for these omissions include culture-specific information being unavailable to textbook publishers, as well as to appeal to a wide-ranging and a conservative an international market as possible. Thus, the design of Spanish FL textbooks not to deal with a specific representation of lived reality for a global audience was already uncovered in the 1990s.

Regarding similar trends in how the cultures of Europeans in general are represented, especially of German cultural representation in the 20th century, Kramsch states: “For pedagogic reasons, textbooks and materials focus exclusively on national characteristics, showing the target culture as more monolithic than it really is” (1993:350). R sler (1994:75) accentuates the differences between GFL textbooks that are produced outside of Germany in comparison to textbooks produced in Germany. Textbooks produced outside of Germany usually consist of a contrastive component between the cultures of the type of FL student they are targeting. This is in comparison to textbooks produced in Germany that do not have this component, as they target a broader, international audience (R sler 1994:75). Kramsch

discusses the separation of culture from language and how it appears within language textbooks:

> Within this national perspective, not only did language teaching get separated from the teaching of literature, it got separated from the teaching of culture as well. Subjects like French "civilisation", German "Landeskunde", English "culture" have developed separately from language instruction, enclosed in textbooks within culture capsules, cultural notes, glossy photographs and more recently a [*sic*] array of so-called authentic texts (Kramsch 1995:4).

Thus, it would appear that during this time, representations of German-speaking cultures in the language learning textbook was restricted to texts, photographs and sections that refer to culture and *Landeskunde* (cultural studies), and that German culture was restricted to a nationalistic depiction of itself. According to this quote by Kramsch (1995:4), this clearly was not only restricted to representations of German cultures within textbooks but was a trend in many FL textbooks of the time. A discussion of video in this literature is absent.

Culture is complex and dynamic (Kramsch 2008:5), and this feature of culture creates the impossibility for a static representation of a culture to ever exist without inaccuracies, yet material writers continue to attempt this feat. The question then remains as to how one navigates 'authentic' information, i.e. based on reality, different from an 'inauthentic' or constructed source in the classroom. Wehage (1997:523) discusses the use of videos in the German language learning classroom to assist in FL learning, describing *Goethe-Institut* video materials, textbook video and video series used for German language learning as authentic materials. Yet, he states these videos provide culturally authentic information:

> Videos bereichern den Lehrstoff mit authentischem Material. […] Videos dieser Art wecken bei ausländischen Studenten das Interesse an einer anderen Kultur. Die in Zusammenarbeit mit Langenscheidt und Inter Nationes konzipierte Videoserie *Alles Gute* vermittelt besonders landeskundlich interessante Szenen, die sich als Segment gut in den Unterricht eingliedern lassen[1] (Wehage 1997:523).

It would appear for Wehage (1997:523) that cultural information is authentic insofar as it awakens interest in a culture. This is not the stance of this research, as any information, perhaps even false information and stereotypical information about a culture is then, according to Wehage (1997:523) authentic simply because the student's interest is engaged with it. This thesis views the most culturally authentic resource as non-scripted footage of native speaker conversations and practices, however, as discussed earlier, a complete,

[1] Videos enrich the curriculum with authentic material. […] Videos of this kind arouse foreign students' interest in another culture. The video series Alles Gute, conceived in collaboration with Langenscheidt and Inter Nationes, conveys scenes that are particularly interesting from a cultural studies perspective, which can be easily integrated into the classroom as a segment.

culturally-authentic representation is virtually impossible due to the limited capacity media have to portray the entire complex nature of culture, as discussed by Kramsch (1993:350), referred to earlier in this section. The videos in Wehage's (1997) study are pedagogic videos designed for classroom use. Nevertheless, he claims that another series, *Zielpunkt Deutsch*, which is video material that accompanies the textbook *Wie geht's? Lehrbuch für das erste Jahr Deutsch*, provides 'authentic' situations with university students that have characters rooted in reality (Wehage 1997:524). He claims these videos mainly depict cultural practices or conventions (Wehage 1997:524). He further concludes that the option to repeatedly stop and play these videos allows for more information on cultural practices or conventions to be perceived by students than what reality could offer (Wehage 1997:525). Teachers should always consider the goals and aims of the videos they are playing, and which materials can be used both before and after the video, especially for stimulating group work (Wehage 1997:525). Therefore, it would appear, the responsibility lies with teachers to navigate through pedagogic videos based on the goals and aims of the videos, if provided. The question, however, remains whether the goals and aims of these videos are clear in all materials.

Allen (2008) conducted a study interviewing teaching assistants for Italian and French programmes in a state university in the United States to determine the role of textbooks and how they are used in the FL classrooms. Although the teaching materials that the teaching assistants used contained videos, they were not deemed as useful as audio recordings and readings (Allen 2008:21). The teaching assistants found limitations regarding the cultural content that textbooks could offer, stating that cultural information was a "side product" in textbooks (Allen 2008:19). Additionally, they disregarded video texts and other textbook materials that contained inauthentic language, preferring authentic texts (Allen 2008:20), i.e. texts that were not artificially composed for teaching and learning purposes. It would appear in this context that pedagogic videos are seen as unhelpful tools in imparting cultural information. Regarding 21st century studies, Weninger and Kiss (2015:19) state that, in order to obtain deeper insight into the impact of textbooks in terms of cultural awareness of FL learners, a guided semiotic approach is necessary, which examines how students are guided in order to interpret them. This explains why a semiotic approach in terms of a DMDA is undertaken in this research (see 4.1.). In addition, Drewelow (2012:283) finds that despite the emphasis on an intercultural approach (discussed in 1.2.) in recent years, an interview with 22 students enrolled in a French FL course revealed there remained a separation identified by

students of language and culture, rather than seeing language as part of culture. This shows that despite the changes in how culture is taught (discussed in 1.2.), the perception of culture from a pedagogical perspective does not always equate to the perception learners have.

McConachy (2018:80) collected data from eight participants of an intermediate to advanced level of English as a foreign language at a Japanese university in the form of a writing task, asking students to identify limitations of their textbook in representing culture and cultural differences. Ethnocentric and stereotypical cultural representations were identified as the most problematic by students, especially how these textbooks portrayed the Japanese, illustrating an essentialist portrayal of culture (McConachy 2018:86). McConachy (2018:86) states that teachers need to assist learners in reflecting on cultural representations, but at the same time to maintain caution that they do not influence students with their own ideologies:

> When learners are less proficient in the target language or have less experience in reflecting on cultural phenomena (particularly one's own national culture), there is an even more important role for the teacher in gradually helping learners become able to make sense of cultural representations for themselves (McConachy 2018:86)

In this way, learners are building on intercultural communication skills (McConachy 2018:87). His study shows that national and stereotypical representations of culture still exist in textbook materials for FL learning, but that there have been teaching recommendations in order to guide students through them.

Since the video clips analysed in this research are additional textbook resources, it is necessary to examine the discourse(s) portrayed in these textbooks in order to gather a comprehensive image of the context. Ashby examines a GFL textbook used in American schools and cultural representations of German "Othering" (2003:11) with a traditional semiotic and discourse approach. These school textbooks were found to continue to maintain the German myths and stereotypes that stemmed from German national identity discourse (Ashby 2003:205). Positive and negative properties can stem from inherent prejudices in textbook designers, which are then filtered through to instructors who were themselves taught with materials similarly entrenched in nationalist ideologies and/or themselves have biased and personal target culture interactions (Ashby 2003:207). These properties from inherent prejudices, when read by the GFL learner, combine with their lack of knowledge or preconceived ideas that they possess (Ashby 2003:207). Ashby (2003:210) claims there is a significant difference between German culture as it exists in society and the imagined

German culture for Americans. These claims are not, however, backed up by interviews or a data collection from the GFL learners themselves.

Neuner (2003:17) states: "[i]n foreign language teaching and learning we do not deal directly with the 'real world' of the target language." He calls the presentation of culture in classrooms "interim worlds" (Neuner 2003:17-18), stating this presentation by teachers is further influenced by media, including the textbook and audio-visual media, but also by the selection of information due to authority interpretation (Neuner 2003:17-18). This is then further filtered by learner perceptions and assumptions learners, possibly already in possession of preconceived notions, make when encountering the text (Neuner 2003:18). Examining examples from GFL textbooks, Neuner (2003:20) provides examples of how a shift in socio-cultural presentation occurred from big C to little c methods of teaching culture (mentioned in 3.1.) and from the grammar translation method of language teaching to the audio-visual method (mentioned in 1.2.): "socio-culture as a topic of foreign language teaching almost entirely vanishes into the background as part of the – often visually presented - framework of situations and settings of the dialogues in the foreign country" (Neuner 2003:20). With the turn to a CLT approach and intercultural teaching approach (mentioned in 1.2.), there was more criticism of the authentic texts used in the classroom, as well as of the portrayals of the target culture in FL textbooks (Neuner 2003:21).

Neuner further elaborates the "interim world" (2003:35) of cultural presentation tends to be influenced by politics, guidelines for education or the predominant grammatical features. The traditional dominance of the grammatical component in textbooks causes a distortion of socio-cultural aspects: "In this connection, it is worth remembering the 'bloodless' characters, which populate the foreign language textbooks and as 'puppets of the grammatical progression' are only allowed to use phrases that the grammar syllabus of the current lesson 'puts in their mouths'" (Neuner 2003:35-36). He suggests that learners' perceptions should not be prohibited from FL teaching, but can be used as a point of reference for the topic selection and task design (Neuner 2003:50). Furthermore, one of the essential features of socio-cultural learning is the emphasis on developing procedural knowledge (Neuner 2003:50).

Ulrich (2004:175) examines four FL textbooks used in UK schools, two of these textbooks are German. In these textbooks, cultural material also seems to be additional, and there is a

further lack of authentic materials and controversial aspects of German culture portrayed (Ulrich 2004:176). Maijala (2007) examines cultural portrayals in GFL textbooks produced by Finnish and Swedish publishers. In her analysis, she determines: 'Die Untersuchung hat gezeigt, dass die regionalen Lehrwerke jedoch nach wie vor eine klare Trennung zwischen Ausgangs- und Zielkultur machen. [...] Die neueren Lehrwerke zeigen das multikulturelle Leben in Deutschland häufiger, die Tendenz ist steigend"[2] (Maijala 2007:12). She thus concludes GFL learning materials have more possibilities and can be used more frequently in intercultural learning, which involves the acquisition of knowledge about relationships between one's own and foreign culture, as well as a sensitisation for an intercultural change of perspective (Maijala 2007:2-13). This is in stark contrast to previous studies, which usually discuss a rather negative or 'inauthentic' representation portrayed by textbooks. Ruthner (2013:209) discusses how GFL textbooks are limited in that they are designed to approach students of all sociocultural origins, a trend recognised in 1990 (Ramirez and Hall 1990:63). She claims texts for non-mother-tongue speakers allow for the adaptation of its contents to the sociocultural context, in which the text is being taught (Ruthner 2013:209).

Bock (2015:15) examines how German culture is presented and which stereotypes exist for Danish-produced GFL textbooks for the 7th Grade, *Der Sprung* (2008) and *Ach So!* (2012). She concludes her analysis of the former: "Vermittelt wird im Lehrwerk ein naives, „kitschifiziertes", problemfreies Bild von Alltagskultur in Deutschland"[3] (Bock 2015:22). Idealised images of *Ampelmännchen* (little male figure on a German traffic light) and *Lübecker Marzipanschweinchen* (marzipan pig from Lübeck) do not contribute to the development of cultural competence (Bock 2015:22). *Ach So!* (2012) on the other hand depicts "ein deutlich vielseitiges und auf Authentizität und Glaubwürdigkeit angelegtes Bild von Deutschland"[4] (Bock 2015:22). This illustrates the differences that still exist between GFL learning materials in regard to cultural representation and that some textbooks contain problematic cultural representations. Leitner (2015:72) found similar differences when analysing German A1 textbooks for representations of diversity. She states: "Die Lehrwerke orientieren sich also an der gesellschaftlich vorgegeben Realität, induzieren allerdings keine Impulse zeitgenössischen Fortschritts sondern reproduzieren teilweise die aktuellen

[2] The investigation has shown that the regional textbooks still make a clear distinction between the original and target culture. [...] The more recent textbooks show the multicultural life in Germany more frequently, the trend is increasing.
[3] A naive, "kitschified", problem-free picture of everyday culture in Germany is conveyed in the textbook.
[4] A clearly diverse image of Germany based on authenticity and credibility.

Gegebenheiten"[5] (Leitner 2015:73). This depicts that despite the current more comprehensive cultural representations in textbooks, gaps still remain.

Brunsing (2016:501) examines cultural representation in images with a corpus of 17 GFL textbooks in order to ascertain whether representations have changed between 2000 and 2010. She concludes that the number of cultural images in textbooks has not changed during these years (Brunsing 2016:510). This contradicts previous studies, illustrating there may not be as great a change in cultural representation overall as previously thought. Karpinnen (2016:7) examine cultural representations of the German countries in Finnish textbooks for GFL. She concludes: "Es gab vielseitige Kulturthemen in diesen Büchern und sie waren zum n Teil alltäglich"[6] (Karpinnen 2016:25). Proving her hypothesis of a Germany-dominant representation false, the textbooks analysed equally represented Germany and Austria, although it seems not much was said of the other German-speaking countries (Karpinnen 2016:25). This illustrates the DACH (Germany, Austria and Switzerland) approach to equally representing culture(s) of German-speaking countries is not being adhered to in some GFL learning materials. 21st century portrayals of culture(s) of German-speaking countries in learning materials seem to still be essentialist in nature, as it is impossible to provide a comprehensive representation of all the facets of a culture. Various learning materials represent culture(s) of German-speaking countries differently.

In order to prevent students from interpreting learning materials through their own internal bias and stereotypes, the question is raised in Garrett's (1991:93) article whether it would be best to first introduce pedagogic videos before videos of more natural or authentic conversations or simply adhere to presenting students with naturalistic videos. Shaughnessy (2003:254-255) examines GFL CALL software and claims that there is a discrepancy between educational content and design, as well as general multimodal resources used for a wide variety of languages. This can often lead to inaccurate representations of culture, particularly in the visual elements of the programme (Shaughnessy 2003:254-255). Problems in programme design can lead to an incomplete or inaccurate representation of the target culture, or, as he terms it, "unreality" (Shaughnessy 2003:259). For example, an inaccurate visual, such as map of Nazi Germany and not an updated map of Germany for a text that does not discuss Nazi Germany, provides an "unreality" (Shaughnessy 2003:259). There is thus a

[5] The textbooks are based on the reality given by society, but do not induce any impulses for contemporary progress, rather in some cases reproduce current conditions.
[6] There were eclectic cultural themes in these books, and for the most part they were commonplace.

mismatch between language and visuals. Furthermore, discrepancies in design can also lead to linguistic bias, gender stereotypes, a lack of diversity represented and racist/sexist discourse in content (Shaughnessy 2003:260-261). This all contributes to inauthentic and problematic visuals such as within video that may be used for GFL learning.

This review deals with work which has solely concentrated on cultural representations in FL classrooms, and the various approaches that have been used to examine them. None of the below video studies discuss GFL but are important to this research as they examine cultural representations in video specifically in a language learning context. Herron et al. (1999:518) examine whether students learn about culture represented by video-based second language programmes. Students in their study were exposed to 10 videos in a textbook series described by Herron et al. (1999:518) as authentic (according to this thesis, they are pedagogic videos) with pre-tests and post-tests to assess student gains long-term in big "C" culture (material aspects) and little "c" culture (non-material aspects) (Herron et al. 1999:518). Their findings support the use of video to teach cultural knowledge (Herron et al. 1999:518). "Students perceived that the videos contained more little "c" than big "C," and that they learned more little "c" information" (Herron et al. 1999:518). Marefat (2002:143) explains that videos in an English as a foreign language (EFL) classroom expose students to authentic voices and dialects other than the teacher's, also providing a cultural context that the classroom alone cannot. Wilcox (2010:91) finds video an excellent tool for developing cross-cultural competence in Japanese learners of EFL in the US. A wide variety of video materials should be used in the classroom, with clear objectives of the videos made explicit to students (Wilcox 2010:97). Furthermore, reflective and social practices should be encouraged and maintained around the video (Wilcox 2010:97).

Tschirner's study discusses emotional, cognitive, linguistic and cultural benefits of using DVD and digital video as a teaching tool in FL classrooms: "Through digital video -- and through other features of digital media such as easy communication around the world -- teaching and learning conditions in FL classrooms may become similar to conditions that apply when living in the target culture" (2001:318). He describes a safe and an emotionally secure learning environment that is provided for FL learners with digital video and multimedia (Tschirner 2001:318). Thus, videos have the opportunity to create a similar environment for the target culture. Herron et al. (2002:41) examine the cultural knowledge retention and maintenance of 51 students studying intermediate-level French at an American

university. They claim that the video series used in the study is authentic for cultural purposes (Herron et al. 2002:37), but these videos are not seen as authentic from the perspective of this thesis as they are pedagogically designed. With pre-testing and post-testing, it was shown that indeed cultural knowledge was improved through the use of the series (Herron et al.2002:48). It showed that there was no difference in gain between cultural practice knowledge and cultural product knowledge, although they did retain cultural practices easier than cultural product knowledge (Herron et al. 2002:9-50). Additionally, intermediate students tend to learn more concerning cultural practices when compared with a beginner level (Herron et al. 2002:9-50). When using video to organise textual information as a means of retention, mixed results were given, where support was made for this type of video to be more successful with beginner French learners (Herron et al. 2002:50). Yet, intermediate students with more French cultural knowledge did better in these kinds of tasks than intermediate students that did not (Herron et al. 2002:50). Students did perceive that they had learned more about cultural practices in the video series than about cultural products (Herron et al. 2002:51). This study illustrates that cultural knowledge can be acquired through pedagogic videos.

Representations of culture in FL learning are thus determined to be mainly essentialist and stereotypical, but there has been research which has revealed that many teaching materials have sought to change this, particularly for GFL. When stereotypical portrayals exist, many of the studies above have illustrated that the responsibility falls on the teacher/facilitator to ensure they are a useful tool to students for the acquisition of cultural knowledge. There is, however, a vast difference in portrayal of materials versus the perceptions of learners and this has to also be kept in mind when attempting to diffuse materials that deal with such portrayals, and that learning about culture(s) of German-speaking countries is a possibility with these materials.

Section 2.2. FL learning and teaching with video technology

This section introduces video as a technological tool for teaching and learning, before examining specific factors and skills that video has been shown to benefit. These are affective factors such as interest and motivation, recall and comprehension (both visual and auditory), and vocabulary regarding visual aids and gestures. The discussion of language and gestures relates to multimodality and consequently studies that address video as multimodal resources, directly and indirectly are unpacked. Disadvantages of FL learning with video are then

discussed, before this section concludes with a discussion of learning with video as a cognitive process and an exploration of the only GFL study found which examines pedagogic videos in an African context.

Videos, and pedagogic videos in particular, have been shown to provide both advantages and disadvantages in FL learning. Salaberry (2001:51), in his discussion of L2 language pedagogy and technology, claims that, rather than using a technology-driven approach, teachers should select specific technologies according to their teaching aims. Since FL learning is a category within L2 learning (Dixon et al. 2012), some studies of L2 learning such as these may shed light on this research. Through careful observation of the influence on learning that tools have in the classroom context, one can establish the adequate pedagogic possibilities of the technological tools at the teacher's disposal (Salaberry 2001:51). He further claims that the success of a pedagogical activity is in the type of processing the task demands of the learner (Salaberry 2001:51). Stempleski states: "It is worth emphasizing that the teacher, not the video, is responsible for making any video-based lesson a rewarding language learning experience" (2002:367). Stempleski's discussion of the role of the teacher in video-based lessons highlights important recommendations for the lesson design for teachers (2002:365-366). Some recommendations include, the use of short video sequences that are integral to the lesson, the treatment of video as both visual and aural texts, the allotment of repeated viewings and guidance toward students recognizing videos as language learning tools (Stempleski 2002:365-366).

Vanderplank discusses literature from the past ten years of language laboratories and DVDs, stating: "they [television and video] are valuable for developing cultural knowledge as well as language, provided their use is structured, well-prepared, selected and graded" (2009:31). He states that there is variation in how learners use video by themselves, thus guidance and training by teachers is helpful (Vanderplank 2009:31). Teachers should be conscientious of the vocabulary required by students when exposed to video (Vanderplank 2009:31-32). Although television programmes may be considered an activity for leisure, they are not taken seriously by students, and video still seems to prove a valuable tool in the language learning classroom (Vanderplank 2009:31-32). Thus, the above has depicted a general shift of scholarly attitude to the function and uses of video in the classroom over time. The major functions of video material in language learning that appeared in this literature were to improve interest and

motivation (affective factors), speaking and comprehension, and vocabulary (also in relation to gestures).

The study of White et al. (2000:174), which examines Spanish as an FL, describes positive affective attitudinal aspects of the videos that contribute to language learning. Students enjoyed the videos and felt less anxious when watching them in comparison to other activities in their study (White et al. 2000:174). A positive attitude toward learning and relaxation are also claimed as benefits of video in language learning by Berk (2009:2). He also discusses possible opportunities to generate interest, motivation, anticipation, energisation and concentration, as well as to strengthen the social relationship between instructor and student, memory, understanding, creativity and flow of ideas (Berk 2009:2). Video can further serve as a vehicle for expression, collaboration and the creation of memorable visual input (Berk 2009:2). Pisarenko (2017:16) likewise illustrates memory and motivational advantages, in addition to communicative and attention advantages. Pisarenko (2017:16) recommends encouraging students to personalise these materials. She believes that videos carry a lasting impact on FL learning (Pisarenko 2017:18). Nova et al. (2017:69-77) conducted action research on EFL students, in which affective factors are proposed as the main contributor to language learning, and educational videos, discovering that students feel motivated to learn due to these technological tools and their vocabulary was subsequently strengthened. From a German perspective and an earlier work, this is also stated by Schwerdtfeger (1989:23), who explains humanity's interest in the audio-visual medium, elaborating that video promotes curiosity and sympathy of the viewer with the characters and moves one's emotions (Schwerdtfeger 1989:23). The selection of interesting audio-visual media in combination with specific exercises to achieve specific goals allows students to foster a type of concentration that is un-arbitrary and focused on specific aspects of the video, thus learning (Sass 2007:6-7). Successful learning through video takes place when stories and protagonists collide and connect with students' emotions (Sass 2007:7). Sass (2007:8) states that film creates an exciting story and this is what causes student curiosity, as the new and unexpected allow for a look into another world and culture, and an opportunity for discussion in the target language is created. The discussions elicited from the *Menschen A1* videos in this thesis and how this contributes to learning is discussed in 6.4.

Altay and Ünal (2017:62) observe English and German FL classes that initially do not use video then later use video and they also indicate that learners do not resist video in the

classroom, on the contrary, it creates motivation and interest amongst students. Mekheimer (2011:26) states there were no measurements of affective factors of learning in his study of EFL student reading and writing skills at a university in Saudi Arabia. However, he noticed that the students in the class using video were more interested, attentive and enthusiastic than the class that was not based around the use of video (Mekheimer 2011:26). He also notes that more discussion and writing prompts occurred in these classes (Mekheimer 2011:27). His study explains how video dramatisations of literary works improve listening, speaking, writing and reading skills (Mekheimer 2011:26).

EFL students' recall and comprehension skills are also found to be improved when they were exposed to pedagogic videos (MacWilliam 1986:133), and thus it is suggested that FL students in general receive the same benefits. Multimedia video has positive influences on pronunciation, and video assists in the acquisition of listening and speaking skills, orienting students to new material by providing background to subsequent material, as well as recall due to visual and contextual features (White et al. 2000:167). Marefat's (2002:143) study of EFL videos in Iran further claims that the visual aids in videos, rather than just the videos themselves, can improve comprehension and retention of information and can also relieve potential student boredom. Furthermore, useful linguistic material, interesting thematic and multi-layered content, good quality visuals, a high correlation between sound and visuals and appropriate length seem to be features that separate a good pedagogic video from a bad one (Garza 1996:5-6).

According to GFL literature specifically, film and video offer *Sehverstehen* (visual comprehension), a fifth skill in addition to the idea of four main language skills: speaking, listening, writing and reading (Schwerdtfeger 1989:24), and this is important to this research with regards to understanding the current design of video tasks in GFL textbooks. Schwerdtfeger (1989:24) further states that visual comprehension is not only important for listening comprehension, but is central to the development of individual language learning skills and motivation.[7] The dual skill concept of *Hör-Sehverstehen* (auditory and visual comprehension) was established in consequence – a term prevalent in the German literature on film and video in the classroom, as well as pedagogic video goals (as mentioned in 1.3., the goals of the *Menschen A1* videos include the training of auditory and visual comprehension). Schwerdtfeger (1989:14) also describes how the language textbook is

[7] Chudak 2012:87 states that Schwerdtfeger (1989) as the first to identify this fifth skill.

different to film, but that they do not work in opposition, rather they complement one another as one can do what the other cannot. Literature topicalising *Hör-Sehverstehen* (auditory and visual comprehension) and *Sehverstehen* (visual comprehension) is now reviewed.

Löschmann (1989:26) also mentions *Hör-Sehverstehen* (auditory and visual comprehension) as a benefit of video, expanding on this in her later work in 1990. She states: "Das fremdsprachige Hör-Sehverstehen, verstanden als Prozeß und Resultat von miteinander in Beziehungsgemeinschaft realisierten Hör- und Sehhandlungen bzw. -operationen, verdient unsere besondere Aufmerksamkeit"[8] (Löschmann 1990:52). Löschmann (1990:52) further describes how video is a mediating process, not simply a process to achieve an objective. Furthermore, images create anticipation that, if successfully realised, acts as an efficient language learning aid, allowing video, especially as part of a series, to be understood more easily by students as they adjust their cognitive schemas (Löschmann 1990:60-62). She concludes, stating that university students in Poland displayed better comprehension results and performance with audio-visual comprehension than when being tested for audio comprehension alone (Löschmann 1990:62).

In a study by Biechele et al. (2001:362), students thought that videos assisted with semantics, syntax and visual comprehension and many also thought they were assisted with pronunciation and cultural knowledge. Biechele (2006:1) introduces pedagogic videos from the textbook *geni@l A1*, stating it offers many exercises, "die das Verstehen der Filmsequenzen initiieren, stützen und begleiten, die Lernende zum Hör-Sehverstehen, Sprechen und Schreiben motivieren und dabei entdeckendes Lernen als Tun in verschiedenen Sozialformen ich machen"[9] (Biechele 2006:2). She claims that aside from supporting listening and visual comprehension, students have an opportunity to experience people and dialogue in their authentic context (again the term 'authentic' here is debatable and complex), film scenes can be understood relatively well on their own without speech (Biechele 2006:2). In addition, they offer opportunities for students to speak and write and that images can assist students in understanding when mother-tongue speech is too fast (Biechele 2006:2-3).

[8] The foreign-language listening-visual comprehension, as a process and result of listening and visual actions or operations carried out in relation to one another , deserves our special attention

[9] which initiate, support and accompany the understanding of the film sequences which motivate the learners regarding listening and visual comprehension, speaking and writing and thereby make inquiry-based learning possible as actions in different social forms

Sass (2007:6) discusses ways in which one can improve students' media competence and *Sehverstehen* (visual comprehension). She discusses pedagogic videos in teaching material, stating " -Sehverstehen wird zwar nicht mehr übersehen, aber es ist noch längst keine integrierte Fertigkeit"[10] (Sass 2007:7). This is from the 1980s when films in the teaching and learning context were used for their information value and students were simply expected to react and repeat the content in exercises that followed (Sass 2007:9). The 1990s showed an emphasis on film as a process of learning information, where students have more of an opportunity to express their interpretations and feelings on the videos before, during and after the screening (Sass 2007:9). Before the viewing, students have an opportunity to be introduced to the historical context, key terms and the title of the film, as well as visual shots or a written dialogue from the film as an activity to elicit hypotheses from students as to what they expect of the film (Sass 2007:10). Sass (2007:10) recommends multiple-choice exercises, and text and picture puzzles for students to engage with the content and narrative structure, but also gives ideas of first presenting students with the film transcriptions, or images without sound. Additionally, students should be given characters on which to concentrate before the film, to discuss and describe them after the film (Sass 2007:10). She further discusses a plethora of various writing and speaking exercises, such as writing e-mails, letters and role plays that one could construct around the film, as well as asking students to interpret the ending, or construct their own ending for the film (Sass 2007:10). Sass (2007) not only offers an historical overview of how video was used in the GFL classroom, but informs this research as to the preferred tasks by textbook publishers for using video, some of which appear with the videos analysed in this thesis.

Sigurðardóttir (2016:62-63) states that films are not used to their full potential because universities do not offer film courses separately as part of FL learning courses and often video is used for *Lückenfüller* (gap filler) exercises, thus contradicting Biechele (2006) and Sass (2007), who argue these types of tasks are an effective way of learning the five skills (*Sehverstehen*, listening, speaking, reading and writing) for GFL. Sigurðardóttir states: "Der Film sollte nicht nur als Lückenfüller verwendet werden, sondern sollte durch viele Aufgaben mit dem Film als Unterrichtsmaterial bearbeitet werden: vor, während und nach der Aufführung[11]" (2016:62). She also elaborates on *Hör-Sehverstehen* (auditory and visual comprehension) as a vital component for FL learning, similar to many opinions of previous

[10] Listening and visual comprehension is no longer overlooked, but it is far from being an integrated skill.
[11] The film should not only be used as a gap filler but should be worked on through many tasks with the film as teaching material: before, during and after the viewing.

authors (Schwerdtfeger 1989, Löschmann 1989, Löschmann 1990, Biechele 2006, Sass 2007), stating that the visual experience of the behaviour of people and visual aspects of culture combined with hearing the language cause all cognitive channels to be utilised (Sigurðardóttir 2016:62). Chudak (2012:87-91) provides an overview of the literature of *Hör-Sehverstehen* (auditory and visual comprehension). Like Sass (2007:6), he describes strategies to achieve this in the classroom (Chudak 2012:91-92). Students need to be made aware and reflect on their *Hör-Seh-Gewohnheiten* (i.e. what listening and visual components assist them in their comprehension of the video) and the genre of the video (Chudak 2012:91-92). Chudak (2012:98) calls *Sehverstehen* (visual comprehension) in this context "film literacy" (2012:98) and elaborates that it should be integrated into and included in the FL classroom.

The comprehension of the various modes that film has to offer is described by Biechele as "film literacy" (2017:200), similar to Chudak (2012:98). She lists multiple choice exercises, true or false questions, sentences out of sequence, *Lückenfiller* and vocabulary lists as exercises that accompany film in teaching materials to build listening and visual comprehension (Biechele 2017:201). This thesis analyses to what extent tasks such as these paired with video are indeed helpful for a South African context. Often film and video are taken as natural tools for assisting language comprehension, but film as an art form is often missing in the GFL classroom and thus film literacy is recommended in order for students to accurately understand what is being presented to them (Biechele 2017:209). Uhlen states: "Videos n genutzt werden, um Vokabular, Dialogstrukturen, grammatikalische Strukturen oder die verschiedenen Dialekte der deutschen Sprache in einem authentischen, gesprochenen Kontext zu zeigen"[12] (Uhlen 2018:2). Videos can also be used as good "Sprechanlässe" (speaking opportunities), in which they provide an opportunity for students to speak and to complete spoken exercises about a particular topic, so they are not only being used with regards to *Hör-Sehverstehen* (auditory and visual comprehension) (2018:3). The learning of vocabulary in the form of gestures and the multimodal nature of videos are now discussed.

Studies on kinesic behaviour may assist this research in determining the potential of learning through gestures. Kellerman (1992:239) reviews literature that examines all types of body

[12] Videos can be used to show vocabulary, dialogue structures, grammatical structures or the various dialects of the German language in an authentic, spoken context.

movement in video material in EFL language learning. Iconic gestures are further linked to the semantic level and assist with recall and decoding (1992:242). In her review of other studies, Kellerman (1992:242) states that gestures are processed along the same channels as visuals and auditory information, paired with one or the other. For example, Riseborough (1981), cited in Kellerman (1992:242), claims that participants claimed they were not aware of the gestures in the video. Without gesture, the listener could lose part of the meaning (Kellerman 1992:243). Kinesic behaviour also signals specific discourse information and regulates interaction (Kellerman 1992:243-245). When this kinesic information is not present, the quality of information may be reduced, and the character of communication may be altered; often communication may also fail (Kellerman 1992:245-246). In terms of comprehension, body language can lead to a reduction of ambiguity and carry more advantages for learning than audio material as the visual material and gestures provide comprehension support, and thus reduce learner cognitive load (Kellerman 1992:249-250). The implications are that raising teachers', and consequently learners', awareness of kinesic behaviour could assist in L2 and FL learning (Kellerman 1992:253). Gestures in this research are discussed in 5.1. and examined in 6.1.

Kelly et al. (2009:313;322) use video to examine whether gestures that co-occur with speech assist in word learning with adult students of Japanese in the US. They show that a possible explanation for gestures assisting in vocabulary learning is that co-speech gestures capture "the imagistic memory trace for a new word's meaning in the brain" (Kelly et al. 2009:330). This provides possible insight as to why visuals are effective as a learning tool when paired with spoken language stimuli, which is applicable to this research that focuses on how learning is achieved multimodally.

Macedonia and von Kriegstein (2012:407-408) elaborate on this memory trace, stating that multimodal learning and multi-sensory input assist in vocabulary learning, and gestures have the potential to link words to the body, which further enhances memory. Talaván (2007:7) discusses authentic video use in assisting EFL students with learning vocabulary: "Foreign language words are learned better when directly associated with appropriate nonverbal referents (objects, events, emotions, context, etc.); this efficient association can only be found in real life or in authentic video." Talaván (2007:5) describes 'authentic' video, for example film and television, as using non-graded language. It is illustrated in a GFL context later in this section, however, that vocabulary learning is not exclusive to 'authentic' videos but is

also found to be an opportunity when viewing pedagogic videos. Within a South African context, for example, Diab et al. (2016:11) examine medical students' perceptions of pedagogic video in teaching language skills and cultural competence in isiZulu, one of South Africa's indigenous languages. Students found that video assisted in vocabulary learning, skills in communication, pronunciation and recall (Diab et al. 2016: 12-13). Pedagogic suggestions included having the videos available on students' cellphones and having students watch them in small groups similar to a tutorial space where groups can discuss the videos and activities (Diab et al. 2016:13). With regard to the German context, Biechele (2006:4) discusses particular examples from *geni@l A1*, where there are opportunities to increase vocabulary and improve skills associated with this, such as phonetic, syntactic and semantic skills. Video has also been shown to be used successfully in teaching reading, writing and grammar in the GFL classroom (Altay and Ünal 2017:62-64).

Jewitt (2006:2) uses a multimodal approach to provide a better understanding of the relationship between multimodal meaning-making resources and school knowledge, learning and literacy in England. She examines how new technologies reshape knowledge as curriculum, learning, writing and how they impact teaching (Jewitt 2006:2-4). While her study focused on school education and not tertiary education or FL education, her findings on modal preferences and learning remain relevant to this study. In an interactive English novel *Of Mice and Men* software, her analysis of character portrayal exhibits that multimodality allows for student engagement at the mode and narrative level. Modal learning serves to emphasise particular objects, characters and moments in the story and offer students several readings of the novel (Jewitt 2006:71-74). Some students chose to 'read' the novel solely through the use of video clips (Jewitt 2006:88). She states that how students learn is shaped by engagement (interest), and students do not read the novel as designed by the CD-ROM designers (Jewitt 2006:93). Students privilege visuals and colour over movement and writing, illustrating the difference between design and perception (Jewitt 2006:93). She claims modal preference differs from learning style theories, in that learning styles are fixed, whereas "modal preference is socially shaped, fluid, changing and contextual" (Jewitt 2006:100). This thesis takes a similar stance on learning style theories (see 3.2.). Her analysis examines student transcriptions and the modes, metafunctions and semiotic resources without the use of software such as used in this thesis (see 4.1.), rather with a typed transcription (Jewitt 2006:38-41). Her analysis reveals the visual mode is often the dominant mode, which may cause the image to distract students from reading. Student modal preferences and reading

principles, when differing from teacher and designer's principles, may impact learning negatively (Jewitt 2006:135-136). She describes learning as involving "internalising the representational and communicative means of the subject discourse. Put more directly, learning involves taking on the identity and discourse of the [school] subject" (Jewitt 2006:25). The first part of this quote does not take into account the outward practice of learning, and her second sentence assumes learners can initially just assume an identity of a specialist, with which they are perhaps too unfamiliar. While she describes learning with less of a traditional focus, which often solely focuses on language, she often leans too much into the semiotic, without including enough linguistic content in her analysis. To avoid this, this thesis places an equal focus on discussing linguistic content and semiotic (see Chapter 4 and 5), and whether modes produce similarity or differences in the meanings they convey, is further analysed (see Chapter 6).

In terms of a GFL context, Plass et al. hypothesised that English-speaking students in the US are more likely to learn German words when they are provided with both visual and verbal modes containing information "than when they select information in one mode or none" (1998:25). This was proven the case, stating as a recommendation that visual and verbal aids should be an option for students and their students learned best if they had a preview video with the relevant vocabulary (Plass et al. 1998:33-34). Students were more likely to produce a word correctly in a written vocabulary test when they reported using a visual aid (such as video) in comparison to an audio aid (Plass et al. 1998:31). Plass et al. (1998:31) draw on the work of Mayer and his cognitive theories of multimedia learning, which is discussed in Chapter 3 of this thesis. Mayer's cognitive theory of multimedia learning is also drawn upon by Roche (2010:1247) in his discussion of audiovisual media and cognitive aspects of visualisation, stating that there is a lack of perception studies in the area of language learning, but that it is clear, suitable visuals allow for the activation and selection of semantic concepts. This is similar to the cognitive process that occurs when language is observed in its real-life context (Roche 2010:1247).

Vysockaja (2013:92-93) discusses pedagogic films (*Sprachlehrfilme*), quoting the broad learning outcomes of *Menschen A1* as an example of the promises made by publishers regarding this type of material. She states that their poor cinematography and the artificial attempt at authenticity would lead to boredom in the classroom (Vysockaja 2013:92-93), yet the student participants in this thesis have a different opinion to this, stating videos are more

memorable due to their poor cinematography and the artificial attempt at authenticity in 5.4. and 5.5. Scherpinski (2014:171) compares listening comprehension of audiovisual material with audio material in a course on GFL in South Korea with a quantitative approach:

> Die Bewegtbilder verhelfen durch ihre konkrete Darstellungsweise und aufmerksamkeitsfördernde Wirkung zu einem höheren Globalverstehen als reine Hörtexte. Dadurch unterstützen sie den Rezipienten auf der einen Seite dabei, sprachlich Verstandenes effektiver in einen kohärenten Gesamtkontext einzuordnen und fördern auf der anderen Seite durch die visuellen Stimuli gleichzeitig die Aktivierung und Generierung sprachlichen Wissens[13] (Scherpinski 2014:183).

Scherpinski (2014:171) thus illustrates the importance of visual aids, which are an important part of cognitive processing discussed in Chapter 3, and are also discussed in the findings and analysis of the videos in Chapters 5 and 6. The results indicate that the portrayal of audiovisual communication and the presence of visuals are an additional advantage for listening comprehension in comparison to audio recordings.

In contrast, Abrams (2016:343) explores task-based language learning with a German video series, *Rosenheim-Cops,* which was not designed for language students, but rather for a mother-tongue audience. She examines multimodal learning resources used in the lesson, which is rare in this area of research, and since this study aims to do the same, this article is one of great importance to bear in mind going forward. Her focus, rather than solely on an analysis of the multimodal resources used in the lesson, includes tasks that informed students on multimodal aspects of the films' design (Abrams 2016:351). The results indicate that "the tasks encouraged semiotic awareness, helped activate referential knowledge useful for accessing multimodal resources, and elicited a positive response to authentic L2 use in context" (Abrams 2016:343). Challenges included lack of linguistic knowledge that hindered learning for some students (Abrams 2016:343). The context is extremely different from the context of this study, however, with students in the first-year course at an American university already able to communicate and talk in German in class, as well as watch a resource as linguistically advanced as *Rosenheim-Cops*.

While there are many advantages to FL learning with video, disadvantages to learning with this type of technology have also been discovered. Video is costly and time-consuming to produce and purchase (Willis 1983:18). If there is a mismatch between linguistic and visual

[13]The moving images help through their concrete presentation and attention-promoting effect toward a higher global understanding than pure audio texts. In this way, they support the recipient on the one hand in classifying what is linguistically understood more effectively in a coherent overall context and, on the other hand, promote the activation and generation of linguistic knowledge through the visual stimuli.

elements, it could lead students to confusion or distraction (Gruba 1999:275). Gruba (1999:275-277) recommends that teachers should show students, particularly salient visuals for their learning. In support of this, Kaltenbacher argues that pedagogic videos on English FL CD-ROMs "often deliver 'multimodality' which actually disrupts learning" (2004:119). In his analysis of CALL English software, the presence of inauthentic gestures (inanimate speech and out-of-sync lip movements) possibly result in FL learners forming incorrect assumptions of paralinguistic practices, as well as misinterpretations that may result in poor to no language learning (Kaltenbacher 2004:133-134). MacWilliam (1986:133) further warns that one should use them with caution and that these kinds of material assist in recall and comprehension, but that they relay more content than grammatical information in comparison to materials designed for a native-speaking audience. This could prevent students from learning vocabulary and grammatical structures from the videos (MacWilliam 1986:133). Dal (2012:3) claims video texts such as pedagogic videos are "passive" and do not generate enough communicative activities in the classroom. He highlights the communicative benefits of video production in the FL classroom (Dal 2012:4), concluding that: "more research is needed on the effectiveness of the interplay between content, language, and media" (Dal 2012:12). Students should also be made aware of digital literacy skills, which depends on linguistic literacy skills (Dal 2012:12). Tschirner (2001:318) agrees, stating that video can provide an opportunity to elicit spoken output by GFL students in activities, and that video lessons must not solely focus on the input. Löschmann (1989:25) provides a limitation of video as a learning material in the GFL learning classroom, as the entire classroom process cannot consist of a video alone without other types of learning resources and exercises. She describes two main factors to consider when determining the goals of FL learning with video: how close the video is in representing reality (Löschmann 1989:25-26), as well as the learner as user of this material (Löschmann 1989:25-27).

In her review of the TV language course *Einblicke*, which is aimed at learners who have completed the basic level knowledge of German, Biechele (2000:166) states that while the chance of learning a FL through video alone is still considered extremely low, video is a good tool for the FL classroom. In their study of media usage by 44 foreign students learning German in Germany, Biechele et al. (2001:348) illustrates that video was not the least or most used within the variety of media resources students had at their disposal, such as books, television and the internet. Students, however, had the most experience with video in comparison to other media (Biechele et al. 2001:360). Lay (2009:107) finds a similar lack of

usage of video in his investigation of how seven Taiwanese Higher Education institutions use film and audio-visual learning materials in their German language classes. Most students watch films and television on a regular basis in their spare time (in English), yet films and videos do not play a dominant role in the German language classroom, rather the textbook does (Lay 2009:125-129). Many students wish films would be integrated more frequently into their lessons (Lay 2009:129), and find that films and video in the classroom assist in their acquisition of communication skills and intercultural knowledge more than the textbook does (Lay 2009:142)

The perception and comprehension of film is discussed by GFL authors as a cognitive process. Biechele (2007:194), for example, discusses film, but also film designed for GFL learners, and its learning potential in the classroom. She describes film as follows:

> Film ist nicht Abbild der Wirklichkeit, sondern ist, bezogen auf das Kriterium der Abbildung von Realität, gedankliche Konzeption von Welt, auf die sich Mitglieder einer Kultur verständigen und durch die über ein Arsenal von Annahmen und Erkenntnisse geregelt ist, welche Aussagen über die Welt gültig sind, welche nicht [14](Biechele 2007:195).

Images in film show people and objects in relation to time and place and these evoke primary processing of identification from the audience, as well as secondary processing that leads to interpretation, perception and subjective understanding of meaning (Biechele 2007:196). She discusses how interpretation can be both accurate and inaccurate, depending on the students' knowledge: if there is a lack of student knowledge in one of these three areas, misinterpretation can occur (Biechele 2007:200). Biechele's work is important for this thesis, as it supports a cognitive approach outlined in Chapter 3.2 to examining student learning with video.

Schwerdtfeger (1980:28) connects video to perception-orientated learning, stating that "practising perception-oriented FL teaching without technical and non-technical media is very difficult, because only the use of media allows the teacher to introduce the most varied speech options into the classroom situation." She defines perception-orientated FL teaching as a teaching strategy which consists of two social situations: one of instruction and one outside instruction for target language use (Schwerdtfeger 1980:26). She further claims that videos of authentic social events in the FL are vital for acquiring and practising perception

[14] Film is not a mirroring reality, but, with regard to the criterion of reality depiction, is a mental conception of the world on which members of a culture agree and which regulate, through an arsenal of assumptions and insights, which statements about the world are valid and which are not. (own translation)

skills in the FL, and that they need not be of great length, 5-7 minutes in total, (Schwerdtfeger 1980:28). She does not attempt to define authenticity; thus, it cannot be assumed that pedagogic videos automatically fall under her description of 'authentic' videos. Social roles and choice of speech and non-verbal behaviour should be analysed by students in either their L1 or the target language, thus mapping between cultures is advised by Schwerdtfeger (1980:28), yet this may not be the best option if the L1 culture is vastly different from the L2 culture. This could lead to attitudes being formulated that create an unrealistic representation based on the differences in cultural representation between the students' culture and the culture represented in the video. 6.4. shows how the students in this study corroborate this. Tschirner (2001:317-318) states that video allows for situated learning (i.e. the learning context is similar to contexts in which students would apply learned content) and it also provides linguistic input that is situated at all language levels, which are perceived both emotionally and cognitively by students.

The only GFL study conducted in Africa in this literature is by Wahba and Zappen-Thomson (1994). Wahba and Zappen-Thomson (1994:29) state that they do not agree with the application of video in the FL classroom as a passive activity for teachers and students. Rather, as their experience with their students in Namibia has illustrated, a great deal of preparation and didactic revision is needed in order to use video in the FL classroom (Wahba and Zappen-Thomson 1994:29). They claim there is a shortage of films suitable for beginner learners and films that do exist, often have to omit language structures that are too difficult, but appear in the situations depicted (Wahba and Zappen-Thomson 1994:31). They recommend showing photographs and short scenes beforehand to prepare the learner for the video, but this of course requires instructor preparation, and students should be allowed to identify different perspectives or alternative situations to show their awareness that the film is not a true representation of reality (Wahba and Zappen-Thomson 1994:33). They examine exposing Namibian students to a pedagogic video that discusses ecology (Wahba and Zappen-Thomson 1994:36). They explain that many concepts such as air pollution and nuclear power, of which the accompanying book expects a middle-class learning base to be aware, are foreign to Namibian students that may not fit into a middle-class mould that the material expects of them (Wahba and Zappen-Thomson 1994:36). Furthermore, the exercises that accompanied the video were not relevant for their goals and would have caused their students to be demotivated and bored (Wahba and Zappen-Thomson 1994:37), although they did not discuss students' answers of these exercises to confirm their claims. Thus, they

exhibit their adapted lesson to allow students to work through videos step-by-step with teachers choosing the information transmitted and the language imparted (Wahba and Zappen-Thomson 1994:37-38). This highlights the difficulties in adapting materials that were designed for learners familiar with the target culture and country to students in an African context with no previous exposure.

In examining studies and literature that examines learning in a variety of FLs including German, one can observe the overall benefits of video implementation in the classroom that are common for all, particularly for speaking, listening skills and comprehension, vocabulary and recall, interest and motivation. While a few challenges were reported, these included student distraction and confusion due to a mismatch of modes, as well as the dominant modes causing students to focus on one over the other, and time and financial factors. When examining literature on video in the GFL classroom, the changes of film pedagogy were observed, from the pivotal point of the addition of a fifth language skill (*Sehverstehen*) to the use of different teaching methods in different stages of the viewing. Cognitive theories of learning drawn on by the aforementioned literature are further elaborated upon in Chapter 3 as a further understanding of the stance of this research concerning learning through video. The multimodal theory Digital Multimodal Discourse Analysis used to analyse the video clips is also described in Chapter 3, however I review literature that discusses it in the next section in order to examine how various multimodal theories were used in teaching and learning and thus elaborating on the gaps, towards which this research wishes to contribute.

Section 2.3 Multimodality, pedagogy and technology

Since videos combine multiple modes, frameworks of multimodal analysis are needed to analyse them. This section elaborates on discussions of education and multimodality, and describes the connection of their theory and method to DMDA, in order to ascertain how DMDA can provide additional assistance in data analysis for this research that previously could not be attained.

Video multimodal analyses in a pedagogic context are virtually non-existent in the FL context of South African Higher Education, yet multimodal research is being conducted in South Africa. Multimodal research at the University of the Western Cape ranges from HIV/AIDS discourse in Kenyan communities (Banda and Oketch 2011) and discourse of

South African University home pages (Mafofo and Banda 2014) to the analysis of toilet graffiti at the University of the Western Cape (Ferris and Banda 2015). There has, in addition, been a substantial amount of multimodal research in pedagogy from South Africa. As Archer and Newfield discuss: "The University of the Witwatersrand initiated the project of multimodality in South Africa through expanding the semiotic consciousness of teachers and learners in its multilingual, diverse classrooms" (2014:1). This culminated in a publication on the multimodal landscape of the English classroom in South Africa by Stein and Newfield (2006). The focus has been on English as a Second Language (ESL) classroom learning and literacy, due to the present literacy crisis in South Africa (Pretorius and Mokhwesana 2009:56). While studies in Archer and Newfield (2014) are analysed with multimodal social semiotic analyses, none of them are on video, or occur in an FL classroom, thus I do not discuss them here.

There has been a significant amount of multimodal research undertaken in the school classroom space, with a focus on multimodal pedagogy (e.g. Prinsloo & Stein 2010; Newfield & Stein 2000), but this has also infiltrated into the South African Higher Education context. Thesen (2001) reflected on multimodal choices in a critical literacy course called "Texts in Contexts" (Thesen 2001:136-137) for an Engineering course at the University of Cape Town to benefit the previously disadvantaged majority of students. Video was not analysed in this study, thus I do not discuss her article in great detail here, except to mention that upon a written assessment by students of their visual analysis of photographs, a large majority of students that failed were those from a working-class background having English as an additional language, and not a home language (Thesen 2001:138). Thesen (2001:138) thus indicates how schooling background influences the choices of multimodal resources in the university curriculum, as students were unable to effectively master the written mode of assessment, nor analyse interpretations of the visual content including plays and images taught in the course.

Yet Archer (2004:41) indicates that the Engineering students at the University of Cape Town may have resource-poor backgrounds, thus limiting access to printed material access, but not to television and video. She analyses textual productions of students with Kress and van Leeuwen's (1996) Visual Grammar, in order to examine what representational academic discourses should be kept or discarded in relation to student learning. Later, she examines writing in South African Higher Education using multimodal approaches such as mind-

mapping, online writing and talking, in order to explore different affordances of multimodal pedagogical approaches used in a student writing centre (Archer 2017).

Archer's (1997) research also discussed in 2.2, reflects on student video reception and production in an ESL context at the University of Witwatersrand, stating "The focus is on authentic texts on an authentic task for an authentic audience" (1997:24). She discusses Halliday and Kress and Van Leeuwen when reflecting on the classroom practice and video production (Archer 1997:51). She also uses a semiotic analysis in order to plan her lesson (Archer 1997:197-198), which uses the video from the British Sitcom *Mind Your Language* to teach students about stereotypes in the English classroom (Archer 1997:172). The only transcriptions discuss sound/music and camera movement (Archer 1997:210-216). There is no fine-grained analysis of the video itself triangulated with student experience, rather it was as admitted, teacher-dictated, with the teacher and researcher's semiotic analysis becoming curriculum content (Archer 1997:170-171). Kruger and Doherty (2014) examine educational videos using cognitive methods such as eye movements, to track whether students can cope with the cognitive load of videos and create a framework to adapt to various pedagogic videos. Cognitive load was not the primary focus of this research. Thus, I did not utilize their approach when analysing the videos in this study.

Gachago et al. (2014:29) examine digital counterstory videos, i.e. stories that challenge social and racial injustice, created by students in South African Higher Education. Their multimodal analysis does not take metafunctional meaning into account. Their research reveals how the complex relationships of students with privilege and identity impact their selection of semiotic resources, with more privileged students in the South African classroom eliciting less straight-forward stories and utilizing distance and point-of-view in order to elicit more emotional responses from the audience. This is in comparison to less-privileged students (Gachago et al. 2014:39-40). Thus, while Gachago et al. (2014:29) recognize important pedagogic opportunities from these types of videos and provide recommendations for their usage in the classroom, the primary focus of the article is on video production by students, and not the reception, which differs greatly from the goals of this research.

Due to the lack of South African research that use DMDA and examine pedagogic videos in an FL learning context, I return once again to examine research at a global level, in order to review the literature on pedagogic videos, and research approaches for non-pedagogic video

that has been found that may be pertinent to this study. I have already reviewed these studies in 2.2., but here I focus on their multimodal approaches, rather than their findings on pedagogic video analysis. In terms of CALL research, there is evidence of more use of multimodal social semiotic approaches than SF-MDA in video analysis. It would appear, although it is not explicitly stated, Kaltenbacher (2004) takes on a multimodal social semiotic approach in his analysis. The same can be said for Jewitt (2006), except she does attempt to order her social semiotic analysis of the salient modes according to their metafunctional meaning. Kaltenbacher (2004) includes no viewer interpretations in this study to support, whether English-language learners would indeed notice these discrepancies of modes mentioned in 2.2. or that it would indeed influence their language learning. Jewitt (2006) attempts to include student opinions to illustrate their reception of the multimodal software, but unfortunately does not on the video clips within that software.

While Lim (2011:45) wrote his thesis on examining classroom interaction, and not pedagogic videos, with an SF-MDA approach and an early prototype of the software that would later become "Multimodal Analysis Video", pedagogic video is briefly discussed regarding the analysis of a single classroom interaction and its discourse. His analysis of a teacher using video in a lesson for students' General Paper in Singapore depicts, however, that the teacher did not take the video seriously, although watching the video clips took up the majority of the classroom time, and did not elaborate enough on the video's information and its relevance to the students (Lim 2011:238-239). He states video and other technological tools "present semiotic selections for the teacher in the construction of the lesson experience for students" (Lim 2011:56). This illustrates the potential of video-screening as a micro-genre of classroom discourse and its potential, but also when it is not used effectively. Amundrud (2017:xiii) also used "Multimodal Analysis Video" to examine classroom interaction for two EFL courses at a tertiary institution in Japan using SF-MDA as well as APPRAISAL analysis. While he examines video in a similar method to this thesis, it is recordings of classroom interactions and not EFL video material which are examined. Nevertheless, his research examines video with a digital multimodal approach in a foreign language learning context, similar to this thesis. It differs significantly to this thesis in terms of the type of video analysed and the context and FL in question, thus it is only briefly described here.

Specifically, in terms of GFL, the focus on multimodal video analysis is usually on video texts not designed for the classroom (e.g. Abrams 2016). I have yet to find a GFL study that

uses a DMDA approach, or even simply an SF-MDA approach, for video analysis. Abrams (2016) used another multimodal analysis framework titled Multimodal Interactional Analysis (MIA), which was developed from the fields of interactional sociolinguistics and intercultural communication (Jewitt et al. 2016:164). This type of framework is more focused on studying social actor interaction with and through multimodal mediational means, thus focusing more on interaction than interpretation (Jewitt et al. 2016:164). Since the participants in the video are not the only salient semiotic resource in my study, I have decided to broaden my analysis so as to not just focus on the social actor, thus selecting DMDA over an MIA.

Abrams' (2016:344) data collection was extensive, having been collected in a four-week study. She used a content analysis of all written tasks and assessments that students answered, whereas I use a thematic analysis. Vaismoradi et al. (2013:400) differentiate these two analyses: content analysis which can quantify qualitative data, whereas thematic analysis is a qualitative approach. Due to student experience being unquantifiable, particularly in a smaller sample that has reached saturation, and the focus not being on task performance, I chose to use a thematic analysis for this study. The size of the task in the preset curriculum was also a factor in that the tasks used specifically target a few linguistic skills, but may not present the entirety of what students learn from these videos, nor the challenges they have, in our context. Thus, I chose to focus primarily on qualitative data, as well as describe experiences from the students in detail, rather than seeking data that would have reflected their performance.

Iedema (2001:200-201) identifies the limitations of social semiotic analysis for video and lists them as follows:

1. This type of analysis is laborious, technical and extremely interpretative, causing the issue that an analyst's own subjectivity may skew the analysis. With annotation software, this has become less labour intensive (see O'Halloran *et al.* 2011:113-114).
2. Due to the reason that social semiotics is primarily concerned with textual structures, it does not examine categories of viewers and their readings of the text. This trend continues with SF-MDA and DMDA, however, this is the reason this study seeks to combine DMDA with a thematic analysis of viewer interpretations of the text.
3. The social circumstances and context of the creation of the text is not included in a social semiotic analysis, and particularly script choices linked back to context would further concretise interpretations. Contextual script choices are often difficult to

obtain for analysts, but many analysts do attempt to examine the context as much as is possible in their analyses.

Despite all of this, Iedema (2001:202) argues for the use of social semiotics as a means of analysis, particularly outlining the need for images and sound to be analysed as well as spoken language, in order to fully understand a text. Within DMDA, image and sound analysis can be integrated and undertaken with one piece of software, which is why this study utilizes this type of analysis, rather than multimodal social semiotic analyses such as these. DMDA is less laborious than previous transcription methods due to its annotation software tool, and the combination in this study of this analysis with viewer interpretations of the video allows for its integration with context, as well as the alleviation of the highly interpretative nature of a textual analysis. The analysis is no longer limited to the analyst's interpretation, as it is now triangulated with the audience members' interpretations in this context.

As one can see from 2.2. and 2.3., there has been an international shift from the analysis and use of specific pedagogic media to the co-opting of already existing media, often media seen as more 'authentic.' Stickler and Shi (2016:124) discuss this shift in a CALL context specifically, but this also applies to general multimodal research in pedagogy as well. Perhaps this is due to previous technological limitations, as Iedema (2001:201-202) highlighted, but these technological limitations are no longer as pertinent as they were years ago. With the help of software such as "Multimodal Analysis Video", analyses can be conducted that examine a large number of modes, leading to a fine-grained analysis of complex media.

This overview shows the literature on multimodality and the medium of video that has influenced this study. When video has been examined within the language classroom in South Africa, studies examining video have not used an SF-MDA approach, nor have they focused on pedagogic videos within textbooks. This section highlighted the many gaps, to which this research seeks to contribute, from multimodal approaches used in South Africa and for video, to multimodal approaches dealing with designed pedagogic videos and videos not designed for the classroom. Noteworthy, this study is unique in three aspects - a DMDA approach to analysing pedagogic videos for a GFL course in South Africa.

Section 2.4. Conclusion

Since this research focuses on language within the domain of multimodality and the aspects of culture and its representations, a section of this chapter is dedicated to examining each of these topics. The teacher or facilitator of the videos seems to play a major role in assisting students to navigate the nationalistic and often stereotypical representations of culture both in terms of textbooks and CALL software. This literature is not only used as a guide to discuss theories on how these impressions are formulated in 3.1. of Chapter 3 but also is drawn on when discussing impression formation on culture in Chapter 6, in triangulation with findings from Chapter 5. Cognition has appeared briefly in the literature discussed in 2.2., and 3.2. in Chapter 3 this is taken a step further drawing on Bandura's (1989) and Mayer's (2005) theories of cognition for learning. Speaking, listening skills and comprehension, vocabulary and recall, interest and motivation and the development of the fifth skill of *Sehverstehen* (visual comprehension or film literacy) are all advantages of language learning that are discussed in the findings in Chapter 5. These findings are triangulated and reported on in terms of what they show about how the videos can be used more effectively in learning in Chapter 6, as well as disadvantages for language learning such as the dominant modes over others, student distraction and confusion. Finally, the overview of research on multimodality in South Africa has illustrated the lack of studies of this kind in the country, and thus provides a rationale for using this type of analysis in this research.

Chapter 3 - Theory

This chapter comprises three sections. 3.1. describes theories for learning about culture with regards to Hall's (1959) iceberg analogy and discussions of culture in the FL classroom. The Parallel Constraint Satisfaction Theory of Impression Formation (PCSTIF) is then described. This is to understand how research question 4 is answered, as well as to define student impression formation of culture. 3.2. deals with Social Cognitive Learning Theory and the cognitive theory of multimedia learning in order to describe how learning takes place, both in terms of language learning, as well as cultural awareness, in order to understand the learning experience of students when viewing the videos and how they learn (research question 1). Risager (2018) describes three ways of analysis, in regard to cultural representation in textbooks: content analysis, discourse analysis and semiotic analysis. DMDA used in this thesis is a combination of discourse and semiotic analyses. Therefore, this theoretical framework and accompanying methodology are used in order to determine the cultural representations in the pedagogic videos, and linguistic features that can be analysed to confirm opportunities and challenges concerning student language learning. Its formation and description are discussed in 3.3. The sections in this chapter elaborate on theories and manners of analysis for learning about culture, the student learning experience mentioned in research question 1 in terms of language and the illustration of DMDA as a theory.

Section 3.1. Impressions of Culture in the Language Learning Classroom and PCSTIF

Hall (1959:85) presents the analogy of an iceberg to distinguish between overt culture (visible and simple to identify) and covert culture (not visible and complex). He states that, while many have refuted the iceberg model as being inadequate, he believes it is valuable to maintain the separation of cultural aspects, comparing it with Freud's theory of the conscious and subconscious levels of the mind (Hall 1959:85-87). While he then also continues to describe a theory of formal, informal and technical aspects of culture, these differ considerably from the simpler differences described above. The concept of two aspects, however, one simple and visible and the other complex and invisible, remains in his view, important. A distinction between these two aspects, which I retain as overt and covert culture and which were used by him and anthropologist Ralph Linton (Hall 1959:86), are thus used in this thesis. Figure 1 illustrates the difference between overt and covert culture with my own examples of cultural aspects that are visible or invisible and that can be spoken about in

terms of observable cultural aspects in the videos versus any intrinsic and invisible aspects, with which they may be associated.

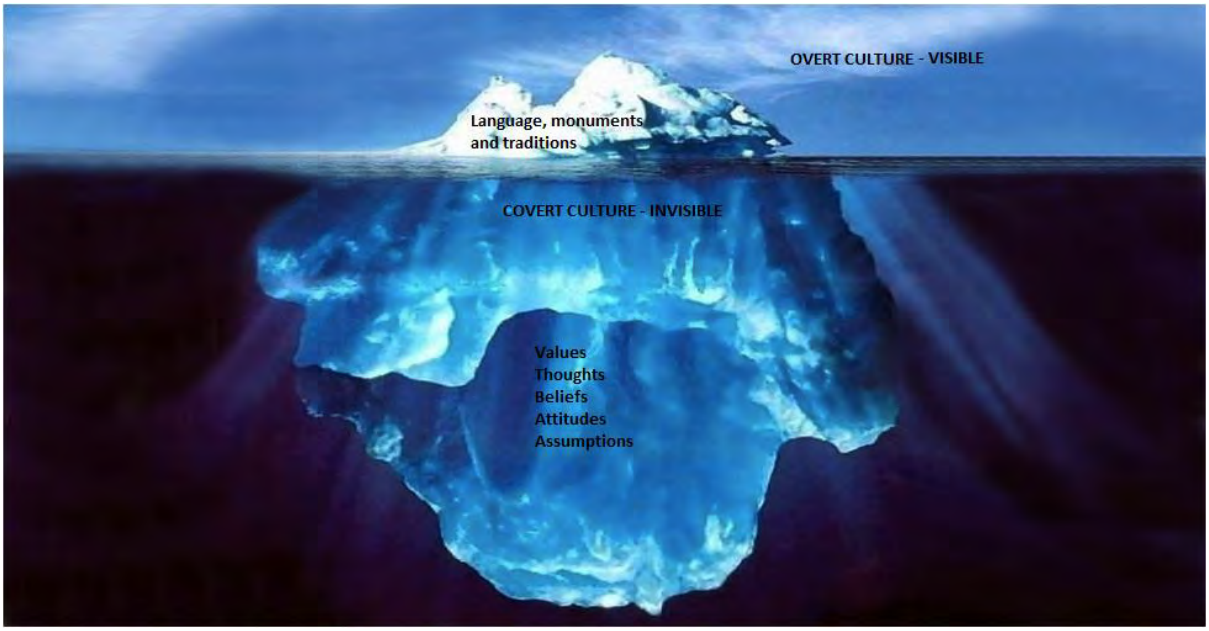

Figure 1: two-part iceberg model (adapted from Hall (1959:85-87)). Background image source: https://oerresearchhub.files.wordpress.com/2014/07/iceberg.png.

Kramsch (2013:65) states that, in the 1970s language classroom, the teaching of language, when implementing discussion on culture, was focused on the big "C" of culture, which means the aspects of culture, such as art and literature that are traditionally taught with a language. The little "c" of culture that encompasses behaviour, communication and non-material culture, only came into focus in language teaching in the 1980s (Kramsch 2013:65-66). Nevertheless, this shows that when the teaching of cultural knowledge was integrated with language teaching, aspects of culture were also viewed in this visible/invisible dichotomic manner.

Kramsch and Andersen (1999:32) examine the representations and manifestations of culture in videos and CALL software through a social semiotic approach and assert that student comprehension of the relationship between language and other sign systems is vital for their linguistic and cultural learning.

> In a video or multimedia program, however, [signs] are inevitably filtered through the film-maker, the camera, and its lens, in other words, through the semiotic system of the video itself. Interpreting that semiotic system means understanding as much as possible why certain events might have been selected, others ignored, why certain people were focused on, others left in the background, and so forth (Kramsch and Andersen 1999:32).

They conclude that multimodal texts are not straightforward regarding meaning and understanding and that multimedia technology has caused mediation which can only act as a substitute and illusion of the context it attempts to recreate (Kramsch and Andersen 1999:32).

Kramsch and Andersen (1999:32) describe the limitations of multimedia and video, one of which is that the computer can both imitate and represent life, thus one does not always know the difference between a true representation and an imitation. They stress that students need to be made aware of a gap between, text or representation and context or the lived experience, which also causes the gap between participant-attributed meanings to a scene versus the meaning understood by the teacher or learner or analyst (Kramsch and Andersen 1999:32). This gap is what they term "the pedagogic challenge par excellence," informing teachers that they need to contextualise the texts used (Kramsch and Andersen 1999:32).

As discussed in 2.1., there has been a tendency to portray German culture as a national culture in GFL textbooks. Language learners find the navigation through representations of national culture in comparison to their own cultures difficult, which often leads to resorting to stereotypes (Kramsch 2013:65). Brinitzer et al. (2019:98) denounce the idea of a national culture, stating:

> Kultur wird in Gruppen durch soziale Interaktion und Anpassung erworben. Welche Musik oder Freizeitaktivitäten wir mögen, hängt mehr vom direkten Umfeld ab als von der Nationalität. Es gibt keine homogenen National- oder ethnischen Kulturen, aber es gibt nationale Stereotype[15] (Brinitzer et al. 2019:98).

They explain that stereotypes are useful tools because they act as reference points to orientate contact, comprehension of and behaviour with other cultures, providing as an example the knowledge about the fact that Germans often greet one another with a handshake (Brinitzer et al. 2019:98). Rösler (2012:85), however, describes the dangers for language learners who are confronted with these stereotypes -- dangers that include language learners losing sight of the heterogeneity of the target culture and dealing with these condensed cultural images as truth, which leads to stereotype reinforcement. Thus, one may observe that cultural representation does not always characterise a group's deep culture, and that stereotypes presume homogeneity where it does not exist. This has both advantages and disadvantages for

[15] Culture is acquired through social interaction and adapting within groups. Which music or recreational activities we like depends more on the direct environment than on nationality. There are no homogenous national or ethnic cultures, but there are national stereotypes.

language learners. Due to building on learners' exposure to stereotypes (as Chapter 2 has illustrated) and using them as orientation mechanisms in learning, this thesis draws on Kunda and Thagard's PCSTIF.

Kunda and Thagard (1996:284) refer to Asch's (1946) based in Gestalt psychology, stating that he believed one's impression of a person comprised a diverse blend of information derived from the characteristics of a single individual. PCSTIF "assumes that social stereotypes and individuating information such as traits or behaviours constrain each other's meanings and jointly influence impressions of individuals" (Kunda and Thagard 1996:284). This model has been extremely useful in illustrating how different cultural influences or cultural representations stimulates or suppresses the formation of stereotypes and therefore this is discussed in detail below. The theory presents this information within nodes in a connected network that is undirected, and the starting point, whether a stereotype or individualistic behaviour, depends on factors such as saliency within the context, prior knowledge and the goals of the perceiver (Kunda and Thagard 1996:287). Specific to these factors, specific observed information is activated and activated connections spread between the observed information and a fixed number of inferred associations based upon these observations (Kunda and Thagard 1996:287). Activation levels of these associations increase or decrease depending on whether the link is excitatory (positive) or inhibitory (negative) (Kunda and Thagard 1996:287). The strength of connections between associations is determined by prior knowledge, and which connections are active or not (Kunda and Thagard 1996:287). This network goes through repeated cycles of observation and association (Kunda and Thagard 1996:287). These processes are described as automatic, although the theory acknowledges that additional and controlled inferences can still be made after network settling (Kunda and Thagard 1996:287). Factors that affect controlled processes include causal reasoning and logical steps for task completion, motivation or when individuals fail to comprehend information successfully (Kunda and Thagard 1996:288). Once any additional information has been successfully integrated, a final impression is then formed (Kunda and Thagard 1996:287).

This theory can be represented schematically. Based on a figure in Kunda and Thagard (1996:286), I provide my own example of observed behaviour and possible impressions. There is a stereotype that Germans do not possess the capacity for humour (Nielsen 2016:234). An advertisement presents a perceiver of non-German nationality with a serious

German man who does not understand a joke told to him. The perceiver's observed and associated information is represented by Figure 2 in the below schema:

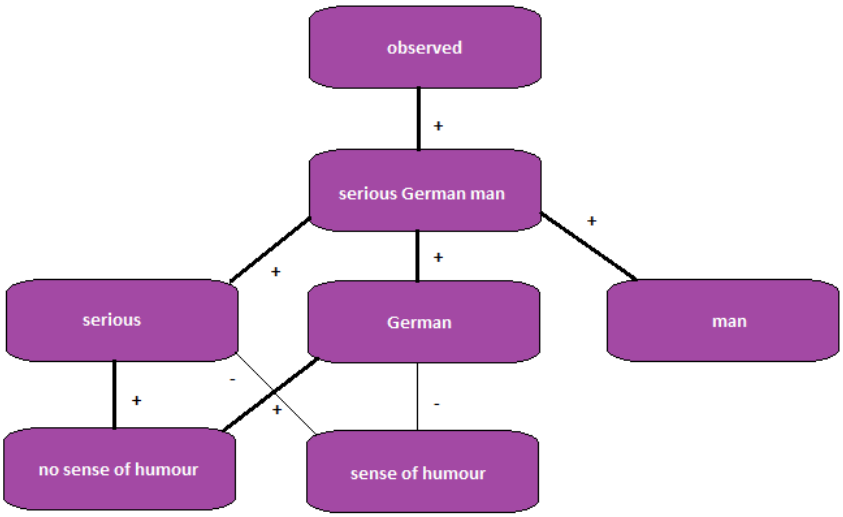

Figure 2: Schema representing how PCSTIF operates, adapted from Kunda and Thagard (1996)

Schemas such as Figure 2 allow us to understand how impression formation operates (Kunda and Thagard 1996:286). Boxes depict nodes that represent the observed person and behaviour, possible interpretations and associated stereotypes (Kunda and Thagard 1996:286). Bold line connections indicate excitatory connections and thin lines inhibitory (Kunda and Thagard 1996:286). The number of plusses and minuses indicates connection strength and does not refer to positive or negative semanticity of associations (Kunda and Thagard 1996:286). Any of the associations, whether stereotypical or behavioural, can be a starting point for the network from what has been observed. The schema in Figure 2 attempts to depict how perceivers see a serious German man according to this theory. If they are aware of the stereotype that Germans supposedly do not have a sense of humour, then this node would be as illustrated (excitatory and active). Furthermore, the individuating information of 'serious' would further strengthen the connections toward reinforcing this stereotype, or in the case of the first observed behaviour, would lead initially to the stereotype association rather than to the nationality. Due to the 'no sense of humour' node being activated by both prior knowledge of the stereotype and 'serious', the stereotype's opposite, 'a sense of humour', is automatically inhibited. Thus, stereotypes influence the meaning of behaviour and traits (Kunda and Thagard 1996:289). Individuating information can determine which

stereotype's subtype is used and in the absence of individuating information, stereotypes influence impressions (Kunda and Thagard 1966:290-291). Thus, if the advertisement shows a German joking and laughing, the activation of the stereotype may not occur, and it would certainly not occur at all if the perceiver was not aware of the existence of the stereotype. Stereotypes sometimes affect impressions of uncertain and irrelevant information (Kunda and Thagard 1996:295-296). "Pseudo-relevant information" such as a biography of a student when forming an impression of self-control can eliminate or weaken stereotype effects (Kunda and Thagard 1996:297). Furthermore, stereotypes can activate subtypes of another stereotype, creating a complex impression network influenced by multiple stereotypes (Kunda and Thagard 1996:298). It is demonstrated how this is relevant regarding cultural knowledge acquisition learning in the videos in Chapter 6.

Section 3.2. Theories of Student Learning Experience

This section outlines the student learning experience, proposing a synthesis of social and cognitive approaches, based on the view that a learning experience is not only a social and sociocultural process, but also a cognitive one. The presence of affective factors acts as proof of learning taking place. Without a positive attitude towards learning, students are not motivated to engage in the active cognitive processing required for learning input during the receiving experience (Bandura 1989:24). The affective component of the learning experience is not examined in great detail. Rather, due to the focus of this research on the learning experience itself, applicable social and cognitive theories of the learning experience are examined to determine the relation of analysed meaning to the process of learning via experience.

According to Bandura's (1989) Social Cognitive Theory of learning (an adaptation of his 1971 social learning theory), a learning experience is both created and learned from, as people "affect the nature of their experienced environment through selection and creation of situations" (1989:4). This is termed reciprocal determinism (Bandura 1989:2) and can explain how learners in a group can influence one another's opinions according to the manner in which they express their opinions, particularly as this study has shown. In this way, learners' cognition and behaviour influence the environment in which they are while the environment influences their cognition and subsequent behaviour (Bandura 1989:2). The notion that behaviour is an output based on input that has been cognitively processed (see Figure 1)

means that the extraction of information and decision-making with regard to behaviour choice often results in learning occurring without an observable change in behaviour (Bandura 1989:24). This is the theoretical underpinning for the analysis of how learning took place in the interview focus groups when watching the videos in this study.

Some of the capabilities which are highlighted by the Social Cognitive theory of learning include: symbolizing and self-reflection (Bandura 1989:9-70). The theory states:

> [p]eople process and transform passing experiences by means of verbal, imaginal and other symbols into cognitive models of reality that serve as guides for judgment and action. It is through symbols that people give meaning, form, and continuity to the experiences they have had. Symbols serve as the vehicle of thought (Bandura 1989:9)

Thus, Bandura (1989:9-13) identifies many symbolic systems of capability that people possess, including language, which can be processed and transformed by people into reality and which further motivates material and non-material processes. He states that children develop symbolic capabilities in language development when they begin connecting symbolic form to meaning in specific experienced situations, as they receive feedback from adults and peers (Bandura 1989:17). Often adults will speak to children more slowly and with exaggerated intonations in an attempt to make learning easier (Bandura 1989:17). This shows how exaggerated intonations, and a slower speech rate are important for learning a new language, even in adulthood. Once syntactic rules and symbolic form are learned from the environment, learners can generate new sentences they have never encountered before (Bandura 1989:16). Language is not only learned through direct experience, but also through indirect events (Bandura 1989:14).

Until children know to which linguistic forms to relate in the world, learning via the spoken mode alone is insufficient (Bandura 1989:16), which shows that learning is a perceptual process requiring more than one mode. While adults have more knowledge of abstract and non-abstract concepts than children, they have to undergo, according to the aforementioned descriptions, another learning process of this nature. For example, the mapping of new linguistic forms, both visual and verbal, onto already known concepts. An inherent capacity of humans, according to Bandura (1989:58), is that of self-reflection. Bandura relates the act of self-reflection directly to learning: “By reflecting on their varied experiences and on what they know, they can derive generic knowledge about themselves and the world around them. People not only gain understanding through reflection, they evaluate and alter their own thinking by this means” (Bandura 1989:58).

While this study does not attempt to measure cognitive processes and their effect on the learning experience, it does not focus solely on a behaviouristic approach that simply examines social aspects of learning. Rather, this study examines the output of learners, which is based on their reflection of their own cognitive filtering and processing. This is further combined with the analysis and reflection of the learning material and context as input structures, which allow us to view how the input has been filtered into output via cognitive structures.

The concept of different learning styles is a topic discussed in a wide range of learning theories (for example, Kolb's Experiential Learning Theory and Gardner's Theory of Multiple Intelligences) (Willingham et al 2015:268; Scott 2010:6). Willingham et al. broadly define learning styles as "differential preferences for processing certain types of information or [...] for processing information in certain ways" (2015:266). For example, there are learners that learn better visually (visual learning type) and learners that learn better when listening (auditory learners) (Willingham et al 2015:266). While the concept of learning styles is popular, previous research on the topic has failed to prove that individuals have preferred and more successful processes of learning that are categorised cognitively (Scott 2010:5). Despite the debate on whether different learning styles indeed exist, the idea that different learners have different individualised styles of learning, and that these need to be catered for in different ways, has become a widespread and popular belief that is further found in pedagogic materials (Willingham et al 2015:268). Thus, while this thesis holds that the multiple modes of cognitive processing play a role in learning, multiple modes are needed together in context in order to be understood. Consequently, theories of learning styles categorise students, which results in specific treatment of them and this has been perpetuated by learning materials as scientific knowledge. Mayer's cognitive theory of multimedia learning is used to explain how multiple modes interact in the context of learning with technology.

To elaborate on the cognitive processes of the language learning experience as discussed by Bandura (1989:15-16), this thesis draws on the cognitive theory of multimedia learning by Mayer (2005). Multimedia instructional messages are "communication containing words and pictures intended to foster learning" (Mayer 2005:31-32). Words can be printed or spoken, and pictures can be static or dynamic, such as those found in video clips (Mayer 2005:32).

His theory focuses on learning in the form of a transfer from words and pictures to comprehension regarding multimedia design (Mayer 2005:32).

The cognitive theory of multimedia learning takes into account three assumptions for the design of multimedia messages: "the dual-channel assumption," "the limited capacity assumption" and "the active processing assumption" (Mayer 2005:33). The assumption of dual channels presupposes that humans possess separate channels for information represented visually and auditorily (Mayer 2005:33). This theory provides understanding in terms of the initial learning process when learners are exposed to the medium of video prior to task completion, social interaction and environmental influence. Since the pedagogic videos analysed in this research are an audiovisual medium, the focus here is on these two channels. For example, if the reader is inexperienced, on-screen text is first processed visually, and an experienced reader could process information visually, while simultaneously converting visual letters and words into sounds, thus further using the auditory channel (Mayer 2005:35). This is the same in terms of writing: ignoring the cognitive processes involved for motor movements, visualizing the writing is either completed via the visual channel for known words and the auditory channel for unknown words, before being written down and processed with the visual channel (see Rapp 2002:47 for a dual-route approach to spelling). If writing and observing take place simultaneously, rather than separately, this might potentially cause a cognitive overload, as the second assumption suggests.

The limited capacity assumption presupposes that each channel is limited in the amount of information that can be processed within it simultaneously (Mayer 2005:35). When the moving images within videos are presented, a learner would only be able to select a small number of visual representations within their working memory (Mayer 2005:35). The same applies for the auditory channel (Mayer 2005:35). This selection process further adds to the reasoning behind why, in terms of Bandura's Social Cognitive Theory, certain observations are selected and processed by the learner. Another assumption by Mayer proposes that "humans actively engage in cognitive processing in order to construct a coherent mental representation of their experiences. These active cognitive processes include paying attention, organizing incoming information and integrating incoming information with prior knowledge" (2005:36). The active processing assumption thus concludes that learners are active rather than passive processors and that the material must have a clear structure and guiding message for comprehension to occur (Mayer 2005:36-37).

According to Mayer (2008:761), “learning is a change in the learner’s knowledge that is attributable to experience.” He states that there are ten principles required for successful implementation of multimedia in the classroom (Mayer 2008:760). The first five are principles to avoid extraneous processing, which is a type of cognitive processing that is not helpful in building a cognitive representation useful for learning (Mayer 2008:763):

- The coherence principle states that the less extraneous material present within multimedia, the greater students’ cognitive capacity will be for important material (Mayer 2008:764).
- The signalling principle implies that when essential words are highlighted, learners are guided in terms of what knowledge to process, and thus learn more efficiently (Mayer 2008:764).
- The redundancy principle states that people learn better with video when on-screen text is not present along with visuals and narration, even if the narration and on-screen text match in content (Mayer 2008:764).
- When corresponding visuals and words are presented closer to one another, learning conditions are better. This is known as the spatial contiguity principle (Mayer 2008:764).
- The last principle for extraneous processing is the temporal contiguity principle, which states that it is also better that the aforementioned visuals and words should be presented simultaneously (Mayer 2008:765).

Three of the ten principles are required in order to successfully manage vital processing for learning (Mayer 2008:765). These are:

- The segmenting principle (narrated animation should be presented at the pace of the learner)
- The pre-training principle (learners comprehend easier having prior knowledge of terms for video content), and
- The modality principle (graphics and spoken language used together are better to learn from than graphics and written language) (Mayer 2008:765-766).

The last two refer to fostering more generative cognitive processes and these are:

- The multimedia principle (when pictures and words are combined, learning is easier than with words alone), and
- The personalisation principle (if the text is conversational and not formal, learning is easier) (Mayer 2008:766).

These principles can be used in this study when analysing the videos' design in combination with DMDA to determine whether the videos' designs are optimal for learning.

According to the cognitive theory of multimedia learning, there are three kinds of memory: sensory, working and long-term (Mayer 2005:37). Pictures and words are processed from outside through a multimedia presentation and enter sensory memory. The holding and incorporating of these raw materials into knowledge take place in the working memory (Mayer 2005:37-38). The long-term memory holds a vast amount of knowledge for a lengthy period of time, but this knowledge must first be processed within one's working memory (Mayer 2005:38). Thus, the five processes in the theory include selecting relevant words from the multimedia presentation, selecting images, organizing selected words, organizing selected images and integrating both representations (Mayer 2005:41). Although these modes are separate, the concept that the meanings they convey have to be integrated in order to be understood corresponds with the Gestalt philosophy of DMDA.

Other cognitive factors, not discussed within Mayer's (2005) cognitive theory of multimedia learning, but which are related to cognitive processes and discussed within DMDA are frequency and salience. Frequency and salience relate directly to memory, attention and learning, with frequency correlating to memory and the manner in which learning elements are presented is a key component for linguists to understand language learning (Divjak 2019:6). However, "experience cannot be reduced to frequency of exposure: information needs to be attended to [...]. The fact that repetition is not the only determinant of encoding in memory makes frequency-based measures very crude approximations of the relation between what is experienced and what is encoded" (Divjak 2019:6). Since frequency of items alone does not create an accurate portrayal of the learning process and what is memorised and salient, the input from learners as to what they have learned not only illustrates what they have not learned, but also represents what they have remembered and learned via attention, while a DMDA of the videos accounts for frequency of items and memory potential, as well as saliency and attentive potential. The next section will elaborate on DMDA.

Section 3.3. Digital Multimodal Discourse Analysis (DMDA)

In order to understand pedagogic videos as multimodal learning resources, it is necessary to examine how the study of multimodality has evolved and how the various methods have been adapted to the complex medium of video as time has passed, which has resulted in the Digital Multimodal Discourse Analysis (DMDA) used in this research. This is graphically represented in Figure 3. The two founding fathers of what became the field of multimodality, and whose work is still relevant to understanding the current DMDA framework, are de Saussure (1974, first published in 1916), the founder of semiotics, and Halliday (1976), the founder of Systemic Functional Linguistics (see Figure 3).

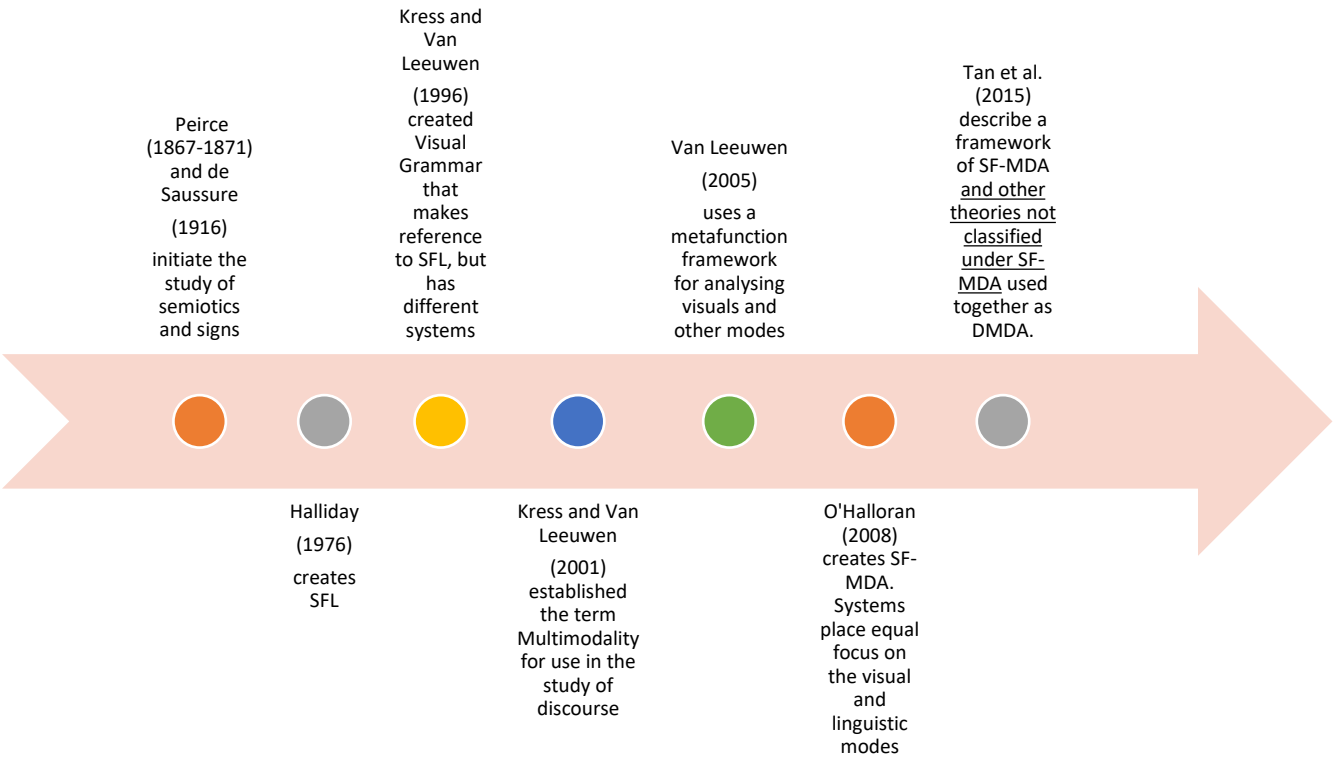

Figure 3: Timeline of important literature in the development of semiotics

The importance of the above authors' work for what became known as the theory of social semiotics is discussed by Hodge and Kress (1988:17). They describe how social semiotics stemmed from "De Saussure's rubbish bin" (1988:17), i.e. other semiotic systems that de Saussure did not examine in as much detail as language or *langage*, such as are outlined more in the work of Peirce (Bateman et al. 2017:57). This included "extrasemiotic phenomena" (Hodge and Kress 1988:18), such as culture and society, and other semiotic systems that were not spoken language, but alongside it and the other more abstract categories within these

other semiotic systems: speaking as an act (*parole*), diachrony, the negotiation between reference structures and their signifier systems, the *signified* or meaning structures behind the form (*signifier)* of a sign and their material nature (Hodge and Kress 1988:18). They further explain how these systems "specify and assume specific relations of power and solidarity between categories of participant, projecting an ideological vision of reality" (Hodge and Kress 1988:46). Hodge and Kress explore many possibilities for semiotic analysis (1988:124), none of which includes the medium of video. This work on social semiotics would, however, become the beginning of Kress and van Leeuwen's (1996) visual grammar of design (see Figure 3).

This visual grammar of design was the resulting attempt of social semiotics to neatly categorize meanings made in the visual mode, as one might when analysing language, and to broaden the field of Critical Discourse Analysis to analyse more than language alone. They say of discourse: "we need to be able to 'read between the lines', in order to get a sense of what discursive/ideological position, what 'interest', may have given rise to a particular text, and maybe to glimpse at least the possibility of an alternative view" (Kress and van Leeuwen 2006:14). Interest in this sense is defined as a "transformative, productive stance towards sign-making [which] is at the same time a transformation of the sign-makers' subjectivity" (Kress and van Leeuwen 2006:12). They use the example of a child drawing circles to represent a car, thus the concept of "car" to the child as a sign-maker, in this case the child, focuses on the characteristic of having wheels and constructs meaning out of circles based on this inference (Kress and van Leeuwen 2006:8). Thus, they fully situate and analyse discourse in terms of social practice. In Kress and Van Leeuwen's (1996) work, they take Halliday's (1976) ideas of metafunctions, and instead of keeping the semiotic plane separate from these as Hodge and Kress (1988) once did, they realize the potential for Halliday's linguistic metafunctional framework to include de Saussure's other semiotic systems that are not language, such as gestural, visual, and auditory modes (see Figure 3). A framework for these other semiotic systems that was originally based on Halliday's (1976) metafunctions, the ideational, interpersonal and textual, was then developed (Kress and van Leeuwen 2006:15). Kress and van Leeuwen (2006) further mention that actions and artefacts that are analysed with the grammar of design, such as the videos used in this study, involve a range of semiotic resources, rather than only comprising the actions and artefacts in themselves, and that the specific choices of these are socially restricted. In order to deliver the intended meaning, this grammar of visual design organizes meaning into three categories:

representation of meaning, interaction of meaning between the sign-maker and the reader of the visual text and the composition of the text as a whole. These mirror the ideational, interpersonal and textual metafunctions of language respectively. The various categories within these systems are defined further below in this section, as all of them are available as resources for analysis within SF-MDA. It is thus important to note that these categories are inherited from Kress and van Leeuwen (1996) for this current framework and are used in SF-MDA.

This medium of video and the entire term multimodality (rather than simply modality) in terms of social semiotic analysis first appears in Kress and van Leeuwen (2001) (see Figure 3). The term distinguishes a modern era of social semiotics when rather than focus on ways to analyse a singular mode (for example, separate analytic linguistic frameworks for the mode of language), Kress and van Leeuwen (2001:20) sought for a framework that could analyse multiple modes. Mode, however, was already defined in terms of register theory: the interest "in the interplay between spoken and written ways of meaning, and in their relation to other modalities of communication (e.g. image, sound, activity)" (Martin and Rose 2007:256). Kress and van Leeuwen (2001) focus mainly on static visuals, (i.e. images) but also refer to more possibilities for analysis within the visual, as music and action within film are mentioned (2001:20). Kress and van Leeuwen (2001) state:

> We move away from the idea that the different modes in multimodal texts have strictly bounded and framed specialist tasks [...]. Instead we move towards a view of multimodality in which common semiotic principles operate in and across different modes, and in which it is therefore quite possible for music to encode action, or images to encode emotion. This move comes, on our part, not because we think we had it all wrong before and have now suddenly seen the light. It is because we want to create a theory of semiotics appropriate to contemporary semiotic practice (Kress and van Leeuwen 2001:2)

Their reasoning for this is due to the editing process whereby with help from technology multimodal text has changed from having been organized in the past by separate specialists of these modes to an often single, multi-skilled producer (Kress and van Leeuwen 2001:2). Subject disciplines have previously used "one language to speak about language (linguistics), another to speak about art (art history), yet another to speak about music (musicology), and so on, each with its own methods, its own assumptions, its own technical vocabulary, its own strengths and its own blind spots" (Kress and van Leeuwen 2001:1). This creates tension between the relevant disciplines as time passes, which is mentioned again below in this section when discussing Kress (2015).

Van Leeuwen's later work in 2005 shows a shift in his way of thinking of discourse as merely a display of interest. He calls his former description of grammar with Kress in 1996 a "systemic-functional grammar of visual design" (van Leeuwen 2005:77). Thus the former visual grammatical categories of representation, interaction and composition are re-focused as their original metafunctional counterparts: the ideational metafunction, the interpersonal metafunction and the textual metafunction, although the semiotic systems within them remain unchanged (see Figure 3). Van Leeuwen describes discourse as vital in determining how "semiotic resources" construct and represent ideas (2005:91). This work extends Halliday's (1976) idea of linguistic grammar to the visual mode, thus being, rather than a fixed collection of regulations for how language should be used, a resource for meaning creation. Therefore, he describes semiotic resources as "the actions and artefacts we use to communicate, whether they are produced physiologically [...] or by means of technologies" (van Leeuwen 2005:3). Van Leeuwen asserts, "In the 1990s I was influenced by my work with members of the critical discourse analysis group" (2005:xii). This can be observed in his 2005 work, as his view of discourse stemming from the area of linguistics, rather than as what he and Kress saw as socio-institutional interest, became the beginning of SF-MDA.

Up until this point, multimodal social semiotic approaches had often not been classified as linguistics, since they focused mainly on the visual or "extra-linguistic" (Kress 2015:53) rather than the linguistic, despite the large linguistic influence within the development of multimodal social semiotics as described above. There is thus a tension between former critical and theoretical disciplines which examine artefacts with a mono-disciplinary lens discussed earlier and the reality of objects that use multiple modes to create an integrated meaning. The challenge or problem of focusing on one mode and ignoring the meanings of others in a multimodal text arose for lack of space within research articles and theses, which can lead to an unbalanced analysis. There is a consequent need for an equitable integration of these disciplines for multimodal analysis to ensure multiple modes can be examined in the same space, which is necessary for a balanced analysis where one mode alone is not dominant. Kress (2015) describes the relationship of multimodality to applied linguistics:

> [...] MM [Multimodality] names and describes a domain for work; it does not name a theory. Nevertheless, once there is a larger domain to be accounted for, there is then the need for an integrating theoretical frame which fully 'accommodates' and accounts for all the entities within it. This becomes important in thinking about points of connection, and possible effects of MM for AL [Applied Linguistics]. (Kress 2015:54)

Bezemer and Jewitt (2010) assert that it is still unclear whether multimodality will eventually form its own field of study, or be integrated into applied linguistics, as certain communities have adopted particular ideas of what multimodality is and has become. Linguistics in particular has a tendency to see multimodality as everything non-linguistic (items that are analysed by the semiotic school of thought), rather than as a framework that seeks to encompass multiple modes, including language (Bezemer and Jewitt 2010). In order to achieve balance and study both the visual and linguistic modes and their semiotic resources and affordances equally, SF-MDA is examined (O'Halloran 2008:443).

The name SF-MDA describes itself: a type of multimodal discourse analysis that stems from systemic functional theory (O'Halloran 2008:443). "SF-MDA provides a transdisciplinary bridge across traditionally distinct fields of study" (O'Halloran 2008:444-445). It thus analyses language with Systemic Functional Linguistic theory, and analyses visuals with a framework adapted from O'Toole (1994) and Kress and van Leeuwen's Visual Grammar (1996), (see Figure 3). It attempts to organize both language and the visual mode on two of the same strata (Content and Expression), yet analysing them differently as different modes, while also attempting to examine all the parts as a whole, rather than interpreting them individually as was done in the past (O'Halloran 2008:444). SF-MDA thus seeks to describe the cumulative impression made by the different modes collectively.

The above attempt to differentiate, review and categorize the aforementioned literature has been assisted by Jewitt (2009) and Jewitt et al. (2014). Jewitt (2009) realizes the complexity of dividing multimodal studies into separate schools of thought, however, she manages to do this based on historical influences, "degree of emphasis each gives to context", internal modal "relational systems" and "the agentive work of the sign-maker" (2009:29). She thus firmly categorizes the different approaches to multimodality in this literature. This categorisation is also present in Jewitt et al. (2014), and while there are many approaches that come from different monomodal disciplines of thought, such as art, linguistics and other areas, all that is important here is that it was Jewitt (2009:29-31) that separated SF-MDA from other work in the social semiotic tradition: "In comparison to Kress's notion of interest, Van Leeuwen's focus on the social production and articulation of meaning gives less emphasis to the individual and the contingent and places more emphasis on the communal and the social" (Jewitt 2009:31). This shows the development of a visual grammar initially begun by Kress and Van Leeuwen and how this influenced the establishment of SF-MDA.

She further describes how SF-MDA, created by O'Halloran, remains focused on Halliday's social systemic functional theory of grammar and rather than describing and analysing 'discourse' at the macro-textual level as Kress did, Van Leeuwen and later O'Halloran's SF-MDA attempt to examine it on the micro-textual level (Jewitt 2009:31). The macro-textual level in this context, examines the text and visuals as a whole work while the micro-textual level is at the level of clauses and shots (lexicogrammar and episode). This means that the focus shifts from Kress' notion of the sign-maker's individual interest to an analytical focus on rules and normative discourse and the social and communal production and articulation of meaning (Jewitt 2009:31). SF-MDA emphasises the metafunctional systems underlying the semiotic resources mentioned by Van Leeuwen (2005) and moves away from the exclusive focus on the social semiotic (Jewitt 2009:32).

DMDA includes SF-MDA and other semiotic resource frameworks for other modes such as sound and music and it was established in order to integrate multiple theoretical frameworks efficiently and to use them in one analysis. Using the "Multimodal Analysis Video" software designed from the creators of SF-MDA described in Tan et al.(2018), O'Halloran and Lim (2014:143) analysed the genre of television advertisements. They integrated SF-MDA with other types of modal analysis frameworks such as intersemiotic texture to elicit various meanings. In O'Halloran and Lim (2018), the method was termed a multimodal semiotic approach, but in order to differentiate this from the multimodal social semiotic approach described above, the term "a digital multimodal discourse approach" has been adopted from Tan et al. (2015:559) to describe my transcription and analysis of the data in this thesis (see Figure 3). Tan et al. (2015:559) coined "digital multimodal discourse analysis," and O'Halloran and Lim (2014), Tan et al. (2015) and Tan et al. (2018) use this methodology and software, but have not yet examined pedagogic videos or videos simply in the context of education. Tan et al. (2016) do examine the context of Computer-assisted Language Learning (CALL) within an English lesson, but they use "Multimodal Analysis Video" to examine 3-D virtual spaces which students use to learn language. I use this methodology, taking its multiple frameworks from the sample project catalogue that comes with the software (illustrated in 4.3.) and adjusting them in order to elicit answers to the research goals of this thesis.

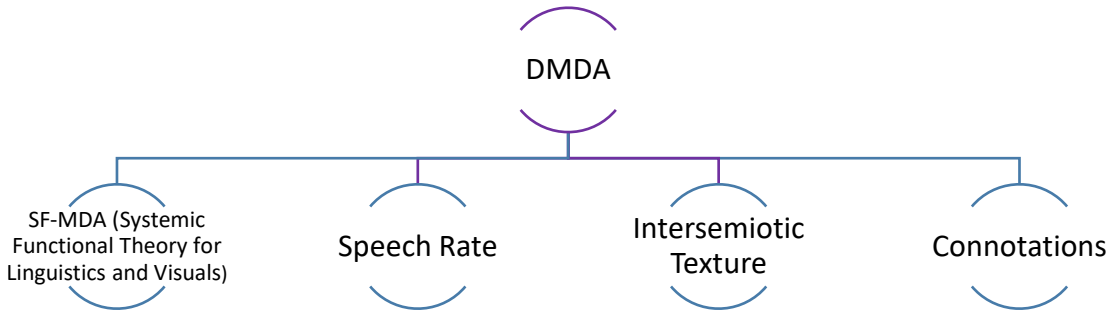

Figure 4: Digital Multimodal Discourse Analysis (DMDA) System Categories adapted from Multimodal Analysis Video*'s General Video Catalogue*

I have selected the DMDA categories presented in Figure 4 based upon the thematic analysis of viewer questionnaires and interview transcriptions. More on how the categories are selected is found in 4.3. Language was analysed in terms of the Systemic Functional Linguistic (SFL) framework and the three metafunctions, namely the ideational, interpersonal and textual, and the speech rate was examined using a formula advocated by de Linde and Kay (2016:45), described further in this section. I adapt English SFL frameworks to German, as Petersen does, stating SFL performed in German Linguistics is a "terra incognita" (unknown land) (2012:1). Elements of language along with the visual elements are thus described in terms of SF-MDA. The visual elements using the three metafunctions, experiential, interpersonal and textual. Penultimately inter-semiotic texture is discussed. This involves inter-relational meaning, which is discussed by Liu and O'Halloran (2009). Inter-relational meaning is the semantic relations between different modes realized through various devices in multimodal discourse (Liu and O'Halloran 2009:369). O'Halloran et al. (2011)'s categories of connotation conclude this section.

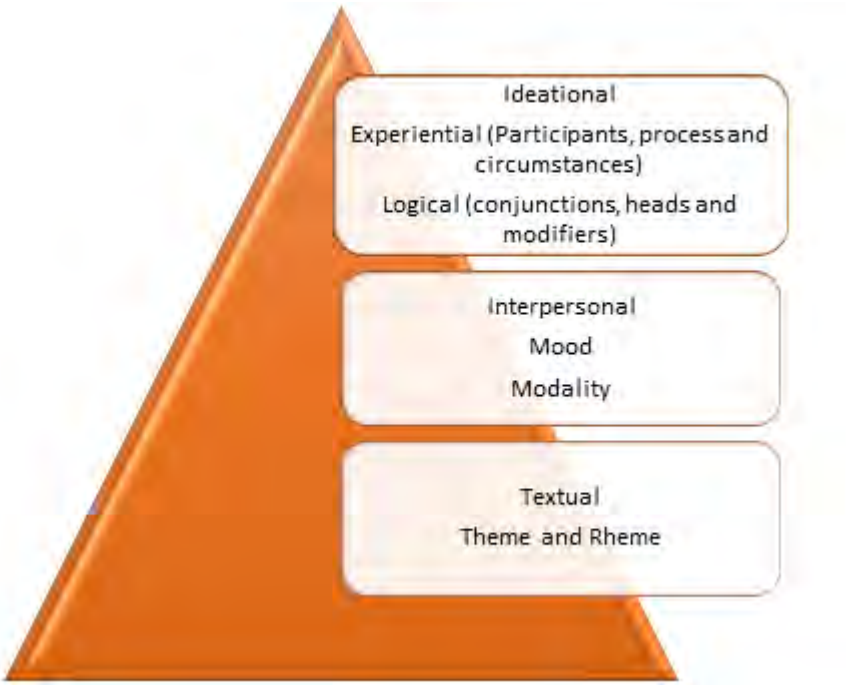

Figure 5: Systemic Functional Linguistic (SFL) framework of metafunctions and systems used with SF-MDA as per O'Halloran et al. *(2018:12-13)*

Halliday (2008:449-451) bases her SF-MDA linguistic framework on Halliday (1976) and takes her visual framework from O'Toole (1994). Language within SF-MDA comprises the ideational, interpersonal and textual metafunction (see Figure 5). The ideational metafunction construes meaning in the form of ideas and world knowledge (experiential metafunction), and the relationships between this knowledge (logical metafunction) (Bloor and Bloor 2004:10). The interpersonal metafunction discusses meaning as a social relationship between the text author and reader (Bloor and Bloor 2004:11). The textual metafunction is concerned with the coherence of the text in terms of its structure and organisation (Bloor and Bloor 2004:11).

In terms of language, one analyses ideational meaning, and experiential meaning specifically, by examining the Participants (people and or abstract, or non-abstract objects), Processes (actions) and Circumstances (e.g. time or place) represented in the linguistic mode, i.e. spoken or written (Bloor and Bloor 2004:111). Processes are sub-categorized by O'Halloran et al. (2018:4) as follows: Material (action), Mental (thought and perception), Relational (possessing an attribute); Verbal (speaking), Behavioural (e.g. actions of human behaviour) and Existential (to exist). Bloor and Bloor (2004:111-126) describe Participant roles based on these categories and I describe them in turn.

The Material Process has either or all the Participants roles of Actor, Goal and/or Beneficiary. An example of Actor and Beneficiary would be “Und jetzt zeig ich Ihnen” (And now I show you) from Video 2 with “ich” (I) as the Actor and “Ihnen” (you) as the Beneficiary of the Material Process “zeig” (show). An example of Goal is “das Geschirr” (the dishes) that Ms Möllemann did not wash in Video 2. The Mental Process has either or both Participants roles of Sensor and Phenomenon. An example of this is “ich mag das Zimmer” (I like the room) from Video 2 where “ich” is the Sensor and “das Zimmer” (the room) is the Phenomenon of the Mental Process “mag” (like). The Relational Process has both Participants roles of Carrier and Attribute or Identifier and Identified Participants. An example of a Relational Process with Carrier and Attribute Participants is “Hier haben Sie alle Läden gleich um die Ecke” (here you have all shops right around the corner), with “Sie” (you) as Carrier and “alle Läden” (all shops) as Attribute for the Relational Process “haben” (have). The Verbal Process has either or all Participant roles of Sayer, Quoted, Reported and Verbiage but only Sayer, Quoted and Verbiage appear in the videos annotated. Examples of this is in Video 3 with “Hier spricht man Bärndütsch” (here one speaks Bernese German) and “Salome hat gesagt: einen lieben Gruß an alle Deutschlerner” (Salome said: a loving greeting to all German learners). In the video, “man” and “Salome” are Sayers, “einen lieben Gruß an alle Deutschlerner” is Quoted and “Bärndütsch” is Verbiage from “sagt” and “spricht” as Verbal Processes. A Behaviour Process has either or both Participant roles of Behaver and Behaviour. Only Behaver appears as a Participant in the videos, for example in Video 2 Ms emann says “Aber ich schlafe sehr gut hier” (but I sleep very well here). “Ich” is the Behaver of the Behavioural Process “schlafe” (sleep). The Existential Process with Existent as a Participant role does not appear in the videos and thus is not elaborated upon here.

Identifying these constituents allows for the understanding of how they relate to each other to construe specific meanings in the linguistic mode. In my analysis I make use of the experiential metafunction, but it has not been pertinent to use the logical metafunction. The logical metafunction delves too deeply into the internal structuring of groups which is not relevant to the research enquiry of this thesis.

The interpersonal metafunction focuses on the Mood realized in each clause, in order to ascertain whether goods and/or services are being provided or requested, and the Modality of the clause that determines the degree of obligation of the offer or request (Bloor and Bloor 2004:283). The three Moods are Declarative, Imperative and Interrogative. An example of

statements in the Declarative Mood includes “Das hier ist der Flur” (This here is the hallway) from Video 2. An example of the Imperative Mood includes “Sehen Sie!” (see/look!) from Video 2. An example of questions in the Interrogative Mood includes “ n Sie mir helfen?” (can you help me?) in Video 1. This Interrogative Mood has an example of Modality in the verb “ n” (can). Due to the relative lack of modals in the video, they are not discussed in great detail in the analysis.

In order to ascertain the textual meaning of the individual clauses, one has to identify and examine the Theme (prioritized information) from the Rheme (remainder of the clause) (Bloor and Bloor 2004:288). An example of Theme and Rheme is found in the sentence “ich zum Goetheplatz” (I would like to go to the Goetheplatz) in Clip 13. “Ich” is the Theme, and the rest of the sentence is the Rheme. Similarly, in unmarked Imperative sentences, the verb becomes the Theme, such as with “Kommen Sie!” (come!) in Clip 14, “Kommen” becomes the Theme.

The visual mode within SF-MDA is also discussed via experiential, interpersonal and textual metafunctions (see Figure 6). This grammar for the visual mode is developed from a study by O’Halloran et al. (2018:12-13) (see Figure 6) and these ranks or levels are based on O’Toole (1994) and his grammar for visual art forms. He has a rank-scale system that examines the semiotics of visual art on the ranks of Work, Episode and Figure (O’Toole 1994:16). Work is the topmost layer of the rank scale and refers to an artistic entity as a whole (O’Toole 1994:225). It is equivalent to the clause in linguistic grammar and comprises multiple episodes (O’Toole 1994:225). Episode can be thought of as a scene represented by the visual artwork and is a rank between Work and Figure, thus it can consist of many Figures (O’Toole 1994:221). Finally, a Figure is an entity that can refer to humans, animals or objects and is the smallest rank of analysis for SF-MDA (O’Toole 1994:221). There is also the rank of Member, which is a unit within Figure, such as tree branches or eye colour (O’Toole 1994:222), but this does not appear in O’Halloran’s SF-MDA, and thus is not used here. These ranks, as with all the ranks in spoken language, are analysed in terms of each of the metafunctions, except the logical metafunction, since equivalents for conjunctions, heads and modifiers in clauses are not easily identifiable in the visual mode (see Bloor and Bloor 2004:140, O’Halloran 2008:458 and Liu and O’Halloran 2009:378). The logical metafunction is rather described in terms of intersemiotic texture as described below.

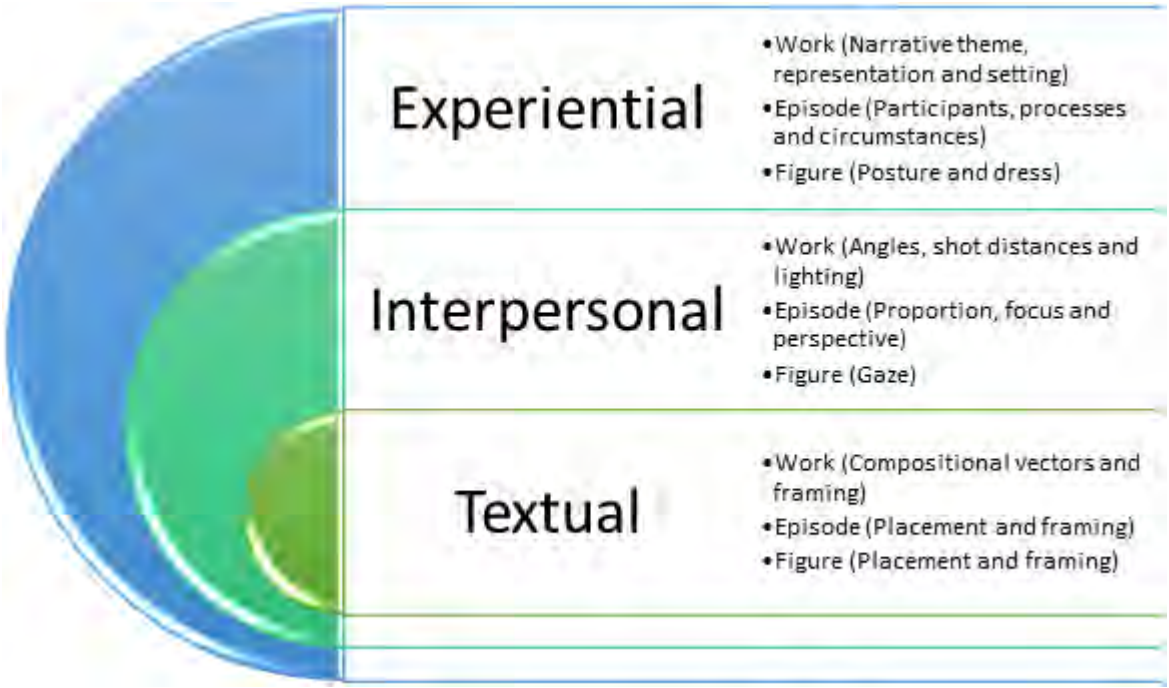

Figure 6: The Visual System in Systemic Functional Multimodal Discourse Analysis (SF-MDA) adapted from O'Halloran et al. (2018:12-13).

At the rank of Work, natural or staged representation and settings are described briefly at the beginning of the analysis, along with the presence or absence of a narrative theme in order to ascertain the "[n]ature of the scene" (O'Halloran et al. 2018:12). I briefly describe each narrative scene and settings for each video in 5.1., 5.2. and 5.3. In order to ascertain the experiential meaning in the visual mode, visual Participants, Processes and Circumstances are described at the rank of the Episode to describe the "[v]isual happenings, actions and relations" (see also O'Halloran et al. 2018:12). The sub-categories of Participants, Processes and Circumstances adapted to visuals derive from the category of Representation from Kress and van Leeuwen (2006:59). The category of Representation comprises Narrative Processes (Action and Reactional Processes) and Conceptual Processes (Kress and van Leeuwen 2006:59). For example, in Video 1 the directions are Concepts being conveyed by Conceptual Processes of arrow representation and gestures. Vectors are oblique lines that indicate relationships and Action Processes are a result of an Actor, from which a Vector emanates (Kress and van Leeuwen 2006:63), while Reactional Processes are formed by glance directions of a Reactor (Kress and van Leeuwen 2006:67). For example, in Video 1, when Oliver is speaking to the tourist, he is an Actor and she is a Reactor, and when she speaks, and he listens the roles are reversed. Due to the transactionality of these Processes, Actors can also often act simultaneously as Reactors, thus they may be classified as Interactors to

illustrate their dual simultaneous roles (Kress and van Leeuwen 2006:63-67), a category that also appears in Video 1 when the tourist both reacts to what Oliver said previously but also acts/speaks herself. The natural or posed posture and collective or individual dress of the Figure rank as described by O'Halloran et al. (2018:12) were not found to be significant in answering the research questions or triangulating with student data and are thus not examined at the level of experiential meaning. Identifying these visual constituents in all three Ranks allows for understanding of how they relate to each other to construe specific meanings in the visual mode and whether these meanings are similar to or different from the meanings conveyed in the linguistic mode.

Interpersonal meaning in the visual mode of SF-MDA is again examined at the ranks of Work, Episode and Figure. At the rank of Work, high and low angles, distance of shots and lighting are examined, which relay information on positionality and power relations between viewer and text creator (O'Halloran et al. 2018:13). At the rank of Episode, proportion and focus (zoom, camera movement [stationary, pan, tilt, pedestal and dolly] and visual saliency [sharpness of focus, colour contrast, foreground and background]), as well as perspective (horizontal [front and angled] and vertical [high angle, eye level or low angle]) are examined. The Figure rank is examined by identifying and analysing the direction of the Figures' gazes (i.e. whether they are focused internally to image or externally to the viewer and directly or indirectly), contrast and the perspective of Figure within the Episode (O'Halloran et al. 2018:13). The sub-categories adapted to Work, Episode and Figure for the interpersonal metafunction derive from the category of Interaction from Kress and van Leeuwen (2006:114-129). Interaction deals with the image as a social act not only between Participants within the image but between the viewers and the image as a whole, with categories of gaze (direct or indirect), social distance (long, medium or close shots) and perspective (horizontal and vertical) (Kress and van Leeuwen 2006:114-129) This provides an examination of the social relationship established between viewer and text creator through the visual mode.

Within the textual metafunction of the visual mode in SF-MDA, Work, Episode and Figure ranks are focused on "the presence or absence of framing devices" such as vectors (Kress and van Leeuwen 2006:59). Furthermore, compositional vectors, when found, are examined in the Work rank. Vectors are opaque lines, such as one portrayed by an arm, and if "participants are connected by a vector, they are represented as doing something to or for each other" (Kress and van Leeuwen 2006:59). In addition, relative placement of Episodes

(information value as described by Kress and van Leeuwen 2006:177), as well as Figures of Episode, are examined. These sub-categories are taken from the Composition category by Kress and van Leeuwen (2006:177). The Composition category has three systems: information value (the placing of Participants and elements), salience (additional focus based on contrast and lighting on particular Participants or Processes) and framing (framing devices and their presence or absence) (Kress and van Leeuwen 2006:177). The software package "Multimodal Analysis Video" places salience and adapts it to the interpersonal metafunction, rather than the textual. I choose to agree that salience is more interpersonal in nature than textual and thus have moved saliency from the textual metafunction to the interpersonal. This allows for the examination and marking of parts as a whole Work, both in the position of "happenings, actions and relations" and "figure" (O'Halloran et al. 2018:13).

Completing a description of visual lexicogrammar, this section will now discuss discourse semantics in terms of intersemiotic texture. Intersemiotic texture is the system for DMDA used to analyse semiotic resources (including language), which is one of the vital systems in answering research questions 2, 3 and 4. Liu and O'Halloran (2009:385) discuss intersemiotic texture in terms of two types of contextualisation relations: co- and re-contextualisation. Co-contextualisation means that intersemiotic meanings converge while re-contextualisation is when intersemiotic meanings diverge from one another (Lim 2004:239, cited in Liu and O'Halloran 2009:385). An example given from the videos of co-contextualisation can be seen in Video 1. Spoken directions are given, while the speakers' gestures and arrows appearing on the screen reflect the information given in these directions (categorised as Comparative relations described below). An example of re-contextualisation can be seen in Video 2 when Ms. emann says "nicht wahr?" (isn't it?)[16] and the image of her in the room does not correspond to the linguistic message conveyed as there is no visual portrayal occurring for the words "nicht wahr?" (isn't it?). It is automatically assumed and analysed in this research that re-contextualisation occurs when co-contextualisation does not occur. The three metafunctions listed which exhibit intersemiotic cohesive devices, or semantic-forming resources that create intersemiotic texture, are logical, experiential and textual (Liu and O'Halloran 2009:375). Micro-components of the logical metafunction, as mentioned previously, are difficult to pinpoint within the visual mode, but they can be mapped between modes.

[16] 'Nicht wahr?" here is the equivalent of the English mood tag "isn't it" (see Halliday and Matthiessen 2014:73).

Implication sequences describe the logical sequences in which a solution to a problem is presented, and these are used to analyse meaning at the level of the logical metafunction (O'Halloran 2005:196-197). Implication sequences are divided into comparison, addition, consequence and time (Liu and O'Halloran 2009:379), with only comparison realising co-contextualisation. The other categories of implication sequences do not realise co- or re-contextualisation (Liu and O'Halloran 2009:385). Due to my interest in how the modes relate to one another in terms of co-contextualisation and temporality, I examine only the intersemiotic Comparative relations and intersemiotic Temporal relations in this analysis. Intersemiotic Comparative relations are a type of meaning organisation in which language and visuals converge in similarity (Liu and O'Halloran 2009:384). Liu and O'Halloran demonstrate this with an image of a dog attacking a woman and how this image conveys the same meaning as the heading of a dog attacking a woman (2009:379).

Video 1 provides a good example, with the arrows used of Comparative relations. Intersemiotic Comparative relations are always accompanied by other intersemiotic cohesive devices (Liu and O'Halloran 2009:379), illustrating that multiple instances of cohesive devices that can be used simultaneously. Intersemiotic Temporal relations map out a successive relationship between semiotic resources of different modes, such as manual instructions being portrayed linguistically, but with a further sequence of images to expand on a single linguistic step (Liu and O'Halloran 2009:383). This is seen in Video 3 when the presenter Martin discusses "der Zytglogge" and then shots of it are shown.

Intersemiotic hyponymy, antonymy, metonymy, polysemy, collocation and correspondence are all categories within the experiential metafunction that can be used to analyse intersemiotic texture (Liu and O'Halloran 2009:371), but are not found pertinent to this study and are thus not used in the analysis. The cohesive devices of reference, Theme-Rheme and Given-New organisation are elaborated within SF-MDA but occur between modes (originating from Kress and van Leeuwen (1996)). Another type of cohesive device used to realise meaning relationships within the textual metafunction is that of parallel structures. These do not appear in the videos, so I do not describe them here. Further details of annotation are presented in Chapter 4.

In this research I am particularly interested in analysing co-occurrences of meanings across systems. In general SFL, a coupling is a connection of two or more meanings across potential systems (Zappavigna 2011; Zappavigna et al. 2008:170). It is a co-selection of certain features of functionality, for example, Knox et al. (2011:101) describe human Participants co-occurring with negative Judgement in an APPRAISAL analysis of Thai newspaper texts. Couplings often appear as states in DMDA. A state can be a single representation of a system choice or combination of them, whereas a coupling is only a connection of two or more system choices and cannot represent a single system choice. States and couplings are referred to in the thesis as similar entities when couplings occur as a state (Tan et al. 2016:260), for example, in Video 2, a cross-metafunctional coupling/state occurs when an attributive Relational Process and Manner Circumstance is within the Declarative Mood and has "Der Flur" (the hallway) as Theme, such as in "Er ist nicht sehr groß" (he/it [the hallway] is not very large). Transitions in DMDA depict visually how combinations of system choices occur in terms of the time duration of the video (Tan et al. 2016:260). In 4.3., State Transition Machines are discussed more and how they relate to the software used for a DMDA. When couplings are rhetorically motivated and occur multiple times in texts, this is known as a syndrome (Knox et al. 2011:101). This means "[a] syndrome is the recurrent co-instantiation of patterns of linguistic potential" (Zappavigna et al. 2008:175). In Knox et al. (2011:101), the negative Judgement of Southern bandits aligns readers against the anti-government movement in the conflict in southern Thailand.

According to de Linde and Kay, "Speech rates were measured by taking the mean rate at which information relating to one subtitle was presented on the soundtrack (i.e. number of words ÷ time of utterance). The rate excludes pauses between utterances, thereby relating purely to textual content" (2016:45). This thesis, despite not examining subtitles, has applied the formula: number of spoken words divided by the time of utterance without pauses, in order to work out the average speech rate of the videos. Rather than comparing this to mother tongue speaker speech rates, the videos are compared with one another to see which are the faster. Then, based upon the student data this determines which average speech rates are too fast or not for students as this was related to the students' responses about their perceptions of the speech rate and its usefulness or lack of usefulness for their learning. More on how the speech rate was calculated is found in 4.3.

Connotation is the final system of DMDA. O'Halloran et al. (2011:118) examine connotation and denotation in a short video advertisement, illustrating how semiotic representation in the text invokes cross-cultural stereotypes, a form of meaning derived connotatively in text. They derive their explanation of connotation and denotation from van Leeuwen (2005) and Kress and van Leeuwen (2001), who further describe linguistic development of descriptions of connotation. van Leeuwen (2005) derives his understanding of connotation in a semiotic context described as cultural provenance as "culturally shared meanings" of ideas and values that are expressed through objects, poses and visual techniques (van Leeuwen 2005:37-39). This is thus connotational meaning, which contrasts with denotational meaning that simply describes the objects and people represented (van Leeuwen 2005:37). Kress and van Leeuwen (2001:73) examine cultural provenance, describing how connotation can be evaluated positively or negatively. The framework in "Multimodal Analysis Video" allows one to annotate the video according to separate idea and value systems relating to connotations that are represented in the text. By annotating ideas and values according to connotative meaning through various frames, I illustrate how the categories within the visual and linguistic analysis (SF-MDA) create various types of connotative meaning, categorising them accordingly, to determine how culture(s) of German-speaking countries are represented in the video, and whether these portrayals are then observed and confirmed by the viewers.

Section 3.4. Conclusion

This chapter has examined how culture can be constructed in terms of Hall's (1959) iceberg analogy and Kunda and Thagard's (1996) PCSTIF. This research has defined learning experience as a cognitive and social process in a social context, drawing on Bandura's Social Cognitive Theory and Mayer's cognitive theory of multimedia Learning. Lastly, DMDA and its development have been described in order to understand how the videos were analysed in this research. The videos were analysed separately with DMDA in order to elaborate on how the different modes in the video work together to create the possible opportunities and challenges that students may face when learning via pedagogic videos, as well as possible cultural connotations which students may formulate. Thus, DMDA is used to analyse the meanings which the video produces while Hall's (1959) iceberg analogy and Kunda and Thagard's (1996) PCSTIF allow us to understand the filtering of these meanings by students, as well as the creation of additional meaning. Chapter 4 describes the methodology undertaken in order to achieve the goals of this research.

Chapter 4 - Methodology

Section 4.1. Introduction and research design

This research uses two different qualitative approaches to analysis – a DMDA of three German pedagogic video clips and a thematic analysis of questionnaire and transcription data from two semi-structured focus group interviews. I took the roles of researcher, interviewer and analyst to ensure uniformity in data collection and analysis. This chapter examines the research methodology and methods in this thesis. Firstly, in this section, an outline of my research design is presented to clarify how the above method combination of DMDA, and thematic analysis were integrated. These two lines of investigation were necessary for this research to gain a complete and accurate picture of how language learning and cultural knowledge acquisition from videos occur within the context. As discussed in 3.2., frequency of items alone does not create an accurate portrayal of the learning process (Divjak 2019:6) and what is memorised and salient. Thus, the input from learners as to what they have learned not only emphasises what they have not learned, but also represents what they have remembered and learned via attention (Divjak 2019:6). A DMDA of the videos accounts for frequency of items and memory potential, as well as saliency and attentive potential (Divjak 2019:6). These qualitative approaches were used, as the focus of this research is on individual student experiences of language learning and cultural knowledge acquisition, thus there was no need for a quantitative approach to examine test or task results. Consequently, both sets of data support one another as the collected data from students corroborates the DMDA and vice versa. Then data collection is discussed in detail in 4.2. in order to demonstrate the reasoning behind choices made. Data analysis is discussed in 4.3. Finally, this chapter is summarised in 4.4.

Since the representation of language and culture in the videos, as well as how the videos facilitate student learning and cultural impressions, are the focus of this research, both methodologies are connected to the videos themselves, see Figure 8, which is a diagrammatic overview of my research design.

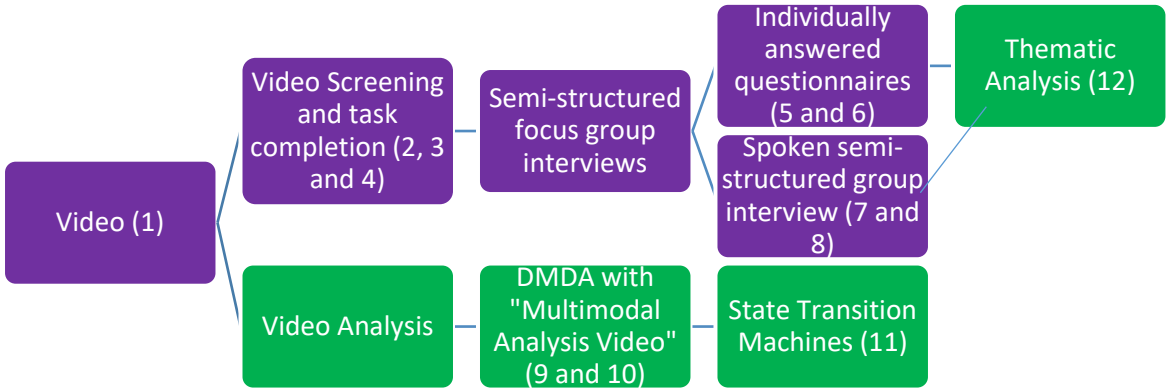

Figure 8: Research Design Outline

Purple indicates steps of data collection and green indicates steps of data analysis. Numbers refer to step numbers of research outline above. The video analysis with DMDA allows for research questions 2 (what is the relationship between the various modes within these videos?) and 4 (how are the German-speaking characters represented in these videos with regard to culture(s) of German-speaking countries?) to be answered by means of State Transition Machines (STMs) while the second methodological process (questionnaires and semi-structured interviews) allows for research questions 1 (what learning opportunities and challenges do Rhodes University students from diverse backgrounds experience in using pedagogic videos as a multimodal learning resource in German Studies 1?), 3 (how do the relationships between the various modes in these pedagogic videos influence student learning?) and 5 (what cultural assumptions and impressions do students make based on the above representations?) to be answered. The answers to the first five research questions in turn, then assist in answering research question 6 (how can the videos be used in a manner that maximises the benefits and minimises the challenges outlined in Question 1?).

An outline of my research design depicted above in Figure 8 is as follows:

1. Three video clips from *Menschen A1* were selected (see 4.2.).
2. The questionnaires and semi-structured focus group interview schedule of questions were drafted according to answers required for research questions (details of this are discussed later in this section).

3. Volunteers were sought for focus-group interviews, which were then arranged (see 4.2. for further elaboration).
4. Videos were screened to students in these interviews and tasks that accompanied the videos from *Menschen A1* were completed either during or after screening depending on student preference (details of video screening are discussed further in 4.2.).
5. Questionnaires were then completed by students (see 4.2.).
6. Semi-structured interviews took place between me and students in focus groups (see 4.2.).
7. Questionnaire data was compiled (see 4.3.).
8. Recorded interviews were transcribed (see 4.3.).
9. DMDA systems were selected based on questionnaire data and transcribed interview content for "Multimodal Analysis Video" (see 4.3. for further details on how DMDA systems were selected).
10. Videos were annotated in "Multimodal Analysis Video" to produce STMs. (see 4.3.)
11. STM findings were described in Chapter 5 and the analysis was reported on in Chapter 6 of this thesis.
12. Themes were identified from questionnaire data and recorded interviews and were described in Chapter 5 and the analysis was reported on in Chapter 6 of this thesis.

This research design allowed for the videos to be the beginning and end of the methodological process before analysis occurred, illustrating the focus of this research to be on the videos and student impressions of them, which is also visible in Figure 8. The decision to have both questionnaire and focus group interviews is due to both types of data elicitation allowing for different types of data to be collected. Questionnaires allowed for Likert scale answers which allowed for data to be categorised and compiled together easily, while the focus group interviews provided more elaboration on descriptive questionnaire answers that students may have felt required too much effort to write out, thus resulting in potential answers being incomplete or lacking detail. The decision was made for interviews to be held in groups as this is usually how tutorials are conducted and students could feel more comfortable to share their opinions on the videos and their experiences learning with the videos in groups than on their own in my presence.

Data collected from focus group interviews were questionnaires and audio recordings, which were later transcribed. The language tasks completed from the *Menschen A1* textbook did not

show significant errors. In Video 2, four students misheard 130 000 for 120 000 “Einwohner” (residents) and three students could not answer 2d, “Was kann ich in dem Viertel gut machen?” (what can I do well in the quarter?). Since this research is focusing on a qualitative analysis rather than a quantitative analysis and a task-based approach, the task results were not analysed in great detail in this thesis and only mentioned in relation to findings and analysis that connected to the DMDA in Chapters 5 and 6. Nevertheless, it was decided that tasks still be completed by students as this adheres to the original format of video screenings and in tutorials as much as possible, thus providing the most similar experience possible to an actual tutorial. The interview setting was reconstructed to be as much like a German Studies 1 tutorial as possible as this is the setting where videos are usually encountered by students at Rhodes University. Since I am attempting to describe student learning experiences, setting these video screenings and interviews in a non-naturalistic context may not have been relatable for students. Thus, they could not have accurately reported on their experience learning with these specific videos and relate it to the other videos with which they have learned in tutorials. Furthermore, the videos as discussed by Hueber (also in this thesis in 1.3.), are supposed to foster interactive learning (Kalender and Pude 2012:3; Hueber.de 2019), suggesting that these videos are indeed meant to be viewed in a group along with social interaction amongst students rather than individually. This further demonstrates my reasoning of reconstructing the interviews to be as close to a tutorial in nature as possible.

Section 4.2 Data collection

The data collected in this thesis were the three video clips from the *Menschen A1* textbook, which were analysed in this research and are used in the German Studies 1 course, as well as written questionnaires and audio recordings from semi-structured focus group interviews of German Studies 1 students mentioned in 4.1. The tasks were not analysed as data, as explained in 4.1. This section firstly examines the collection of videos and how these were selected, and then describes how volunteers for the semi-structured focus group interviews were selected.

The videos were first taken from a CD-ROM in *Menschen A1.* Clip 13, 14 and 15 and their accompanying tasks (titled Video 1, Video 2 and 3 in this thesis) were selected from chapters towards the end of the book in order to ensure students had some previous experience with other videos from their tutorials, but had also not previously been exposed to the three video

clips selected for this study. In this manner, students could comment on this type of video with more experience and learning and this would count as experiential learning that develops over a longer time period as discussed by Bandura's (1989:15) Social Cognitive Learning Theory in 3.2. This was to ensure that data was collected from their initial impressions and experiences of these particular videos, but they could comment on how the type of video helped them learn and their cultural impressions of the videos to which they were exposed in general. Video 3 clearly focus on more overt cultural aspects as it discusses a specific city (Bern) in Switzerland, while the other two videos portray general situations of asking for directions (Video 1) and viewing an apartment (Video 2). Thus, they present a variety in terms of content and their tasks appear altogether on a page in a section after the three chapters to which they relate. Therefore, three video clips were chosen and these three in particular. Questionnaires and semi-structured focus group interview questions were then drafted to gain information which would aid me in responding to research questions 1, 3, 5 and 6.

Volunteer participants were sought, for this research from a class of 48 students who all participated in the German Studies 1 course in the year 2019. This did not include staff members participating in the course, as in most cases, staff have more academic and lived experiences than students, and thus could skew the sample. One Master's student in English Literature between the ages of 25 and 30 participated in the study. However, she still had considerably less academic and lived experience than the staff member enrolled in the course, who was excluded from participating. Students' other characteristics were varied, including age, ethnicity and year within the Higher Education system. These statistics, however, were not collected and noted for each participant due to these statistics not being put forward as reasoning behind certain experiences of the videos in this study, and thus were not a major focus for this research. Age ranged from 18 to 30 years of age and year of study ranged from first year of undergraduate study to Master's level. The 48 students of German Studies 1 at Rhodes University were reached in order to achieve the largest possible sample size to answer the research questions which focus on student language learning at this level. They were contacted first by asking for voluntary participation in the interview before a class started, and then asking that they write down their student numbers on a list that went around the classroom if they were interested in participating. Tutors then further passed the sheet around in tutorials. Of the 48 students in the class, 14 agreed to participate and were emailed with a Google form to pick suitable times they could attend the interview. Eleven participants

arrived and gave consent to participate. When the 11 had shown agreement, I had randomly organised them into groups, except since there were two men, and the rest women, I ensured one man was present in each group to attempt to stratify more heterogeneously. Thus, of the 11 participants, 2 were men and 9 were women in terms of sex, varying in age, ethnicity and year within the Higher Education system.

A minimum of four students (and no more than six students per group) to participate in one focus group discussion were needed in order to have a quality discussion of the questions that relate to their learning experience in Higher Education, as confirmed by Breen (2006). Of the 14 students that agreed to participate via an electronic survey on Google forms, 11 students attended the study with 5 students in the first group and 6 in the second group. The amount of participants in this research is believed to have reached data saturation, despite this being difficult to achieve with qualitative data such as this. As Guest et al. (2006:59) in terms of health research have illustrated, basic themes from a thematic analysis occur after six individual interviews with saturation occurring within the first twelve individual interviews. While this study had one less participant than the number recommended by Guest et al. (2006:59) for saturation, it is accepted in this thesis that the majority of themes that exist were found due to this being reached already at six participants.

The questionnaire was designed to elicit information from students regarding their cultural impressions of the videos, as well as their language learning experiences when being presented with these videos. Figure 9 represents each of the questionnaire questions, along with the focus group interview schedule questions and overall research questions to which they correspond.

Questionnaire	Focus Group Interview Schedule Question	Research Question
1. What were your general impressions of the videos? On a scale of 1 to 4, please rate with a ✓, with (1 = this does not apply at all, 4 = this completely applies to my experience of the videos). If there are any options you wish to add, please do so beside "Other:" below.		
2.a. To what extent do you find the videos to be helpful with your language learning? On a scale of 1 to 4, please rate with a ✓ (with 1 = not at all helpful, 4 = extremely helpful). b. Please explain the reasons for your answers above.		
3. a. To what extent do you find the people and situations in the videos to come across to you as people or situations you might encounter in Germany, Switzerland and Austria? On a scale of 1 to 4, please rate with a ✓ (with 1 = impossible/unrealistic, 4 = possible/realistic)? b. Please explain in one sentence the reason for your answer above. son for your answer above.	11. To what extent would you say the interactions and the depictions of the German-speaking countries in these videos match your idea of what Germany, Austria and Switzerland are like?	5. What cultural assumptions and impressions do students make based on the above representations?
4. a. What do you think the videos are trying to portray about German speakers in German-speaking countries? b. On a scale of 1 to 4, please rate with a ✓ to what extent do you agree or disagree with these portrayals (with 1 = strongly disagree, 4 = strongly agree)?		5. What cultural assumptions and impressions do students make based on the above representations?

Questionnaire	Focus Group Interview	Research Question
5. What did you find difficult to understand in these videos regarding: a. language in the videos? b. the characters' actions or behaviours in the videos?	12. Regarding language in the videos, what did you find difficult to understand?	1. What learning opportunities and challenges do Rhodes University students from diverse backgrounds experience in using pedagogic videos as a multimodal learning resource in German Studies 1? 3. How do the relationships between the various modes in these pedagogic videos influence student learning? 5. What cultural assumptions and impressions do students make based on the above representations?
6. On a scale of 1 to 4, please rate with a ✓ to what extent you would recommend these videos as a study aid? Please place ✓ in one of the boxes below (with 1 = would never recommend, 4 = would definitely recommend).	6. From your personal experience, what are the advantages of language learning from these videos, especially in comparison to learning from books? 7. What are the disadvantages of language learning from these videos from your personal experience, especially in comparison to learning from books?	1. What learning opportunities and challenges do Rhodes University students from diverse backgrounds experience in using pedagogic videos as a multimodal learning resource in German Studies 1? 3. How do the relationships between the various modes in these pedagogic videos influence student learning?
7. To what extent do you think the videos are being used as effectively as possible in the German Studies 1 course? On a scale of 1 to 4, please rate with a ✓ (with 1 = not at all, 4 = they are being used as effectively as possible)	8. What do you think these videos are useful for when learning for German Studies?	6. How can the videos be used in a manner that maximises the benefits and minimises the challenges outlined in Question 1?

Figure 9: Table depicting corresponding questionnaire questions, interview schedule questions and research questions.

Question 1 of the questionnaire (see Figure 9) can be corroborated to the positive or negative nature of the cultural impressions that students made. Likert scales were designed with choices from 1-4 to ascertain whether students specifically agreed or disagreed, providing them with no neutral option. Question 2 of the questionnaires (see Figure 9) was asked in order to answer research questions 1 and 3. Students were given a Likert scale in Question 3 of the questionnaire along with some space to answer and this question of the questionnaire is termed question 2 according to when it was asked of both groups during the interview in 5.5. This was asked in order to provide answers to research question 5 (see Figure 9).

Question 4 of the questionnaire is also to provide answers to research question 5 (see Figure 9), as well as corroborate these answers with findings in the DMDA that sought to answer research question 4. Question 5 of the questionnaire is asked in order to specifically find out the challenges they had in relation to research questions 1, 3 and 5 (see Figure 9). This was also elicited by a question asked of them in the focus-group interview (see Figure 9) in order to get more detail on questionnaire responses. This question of the interview is discussed as question 5 asked of students in 5.5., again due to the order in which they were posed in the interviews. Question 6 of the questionnaire allowed for further elaboration on the uses of video in terms of language learning for research question 1 and 3 answers were asked in questions 6 and 7 of the interview schedule (see Figure 9). The questions in the interview schedule not reflected in the questionnaires asked students what the advantages and disadvantages are in terms of learning from videos in comparison to learning from books. These were discussed as questions 3 and 4 in 5.5. Question 7 of the questionnaire is asked in order to elicit assistance in the formulation of recommendations for research question 6 (see Figure 9). This was also asked in question 8 of the interview schedule in the form of asking students what they thought the videos were useful for when learning for German Studies (see Figure 9). Student responses are discussed in terms of this question as question 5 of 5.5. (see Figure 9). In the beginning of the interview, students were asked if they noticed anything interesting they wish to discuss in order to initiate conversation about the videos and such an open question could elicit results for research questions 1, 3, 5 and/or 6, also providing me with feedback from students to decide with which of the questions within the interview schedule should follow. Due to time constraints for the focus-group interviews, not all the interview questions on the interview schedule were asked and I chose the questions asked in the interview based upon the topics discussed in the interviews themselves, as well as to elicit more detailed responses for questionnaire questions as well as for research questions.

The semi-structured focus group interviews were scheduled for October 2019 at the latest in order for the process of ethical clearance to be completed and thus the videos that were chosen were the videos that appeared in the German Studies 1 syllabus during this time, scheduled for being presented to the students during their tutorials. It also allowed for students to have previous exposure to the videos selected from the *Menschen A1* textbooks. The videos were only shown in Swot week (a week of study and no course work before examinations) for research purposes, but it was made clear that the videos would not be examinable content, and thus were optional to be viewed. If students did not view the videos, they were not missing out on integral syllabus content.

In the data collection sessions, the informed consent form and rules thereof were first explained to the students, after which students were given the opportunity to consent by signing said forms and told they could leave the study at any time. The tasks that accompany the videos in the *Menschen A1* course books as exercises were then handed out to them in a paper-based form, and the tasks were explained to them before the videos were screened. They were allowed to fill in the tasks as the video was being screened or after each clip. Each clip was played once. This process took 10 minutes. After the video screening, I went through the answers in spoken language with students by asking them for the answers, waiting for their responses as a group, and then confirming that their answers were indeed correct, thus attempting to simulate tutorial conditions. Then the questionnaires (see Appendix A) were handed out and they were given 15 minutes to complete this (see Appendix B for student answers). Questionnaires were completed individually before focus group discussions occurred, during which time students were quiet and did not exchange information with one another in order to ascertain individual opinions without prior group influence. We then had a 15-minute semi-structured focus group interview (see Appendix C for interview schedule) take place, and I recorded two audio recordings on two separate devices. The recordings were then transcribed. The transcriptions are separated according to focus groups, which are labelled A and B respectively. The transcription of focus group A can be found in Appendix D and the transcription of focus group B can be found in Appendix E. It is important to note that, as these were semi-structured focus group interviews, not all of the questions in the interview schedule were posed, and the questions were asked in terms of providing detail to questionnaire answers above asking other additional questions. However, the interviews are not completely secondary to the questionnaires in terms of simply clarifying responses, but

there are also opportunities of unique data collection specifically from the interviews themselves.

Some of the students had been tutored by me and others had not. I thus had to overcome the challenge of being biased and letting my students speak more than other students. Since I had undergone tutor facilitation training, this was not difficult and my actions and stance of a role as tutor in these interviews allowed for these focus groups to be as similar to an actual tutorial as possible. My tutor facilitation training allowed for me to take the role of facilitator in these interviews. This was important because it allowed me to elicit information from all of the students as I would in a normal tutorial rather than influencing students by voicing my own impressions or constantly selecting specific students to provide me with information. It also allowed for students that had me as a tutor to be more comfortable in sharing their opinions and being more forthcoming with contributions than the other students which would have hopefully had positive influences during group discussions. The persona I adopted as the interviewer was a kind, sympathetic and enthusiastic role of a tutor, facilitating discussion in focus groups by ensuring all students in the group had a chance to share their opinion on a specific question before moving on to the next question. The data collection within focus group interviews occurred successfully and all of the student participants showed enthusiasm and engagement towards the process. Group A had a more outspoken dynamic than Group B, which thus required more questioning from me (see Appendix D and E), despite my attempts to place a few students from similar tutorials together in order to ensure a familiar atmosphere to encourage speaking. With the many factors that influence student social interaction, I cannot accurately determine why Group B was not as forthcoming in their responses as Group A.

Section 4.3. Data analysis

As discussed in 4.1., the two types of analysis include a thematic analysis of both student questionnaires and interview transcriptions, as well as a DMDA of video clips selected from the textbook *Menschen A1.* Since these textbooks have been purchased by the German Studies section of the School of Languages and Literatures for the students and the videos are files on a CD, I could analyse digital copies of them. Thus, the videos were obtained and accessed legally.

The annotations were undertaken after the interviews, although my initial impressions of the videos had already been noted down beforehand in order to record my pre-assumptions as an analyst before exposure to student ideas in the interviews, such as for example, recognising the connotation of German women as messy from Video 2. This was to identify my own pre-assumptions to avoid influencing the students with them in the interview, as well as to ensure that the interviews had not completely influenced my own impressions of the videos. The student data was transcribed before the DMDA of videos to determine which modes were salient to students and consequently which modes needed the most focus, thus determining which systems of analysis would be used for a DMDA. The transcriptions were only formally analysed, however, after the analysis of the videos.

Each video that was chosen and annotated with the licensed software "Multimodal Analysis Video" for the purpose of constructing State Transition Machines (STMs) for the discussions of findings in Chapter 5 and later reporting on analysis in Chapter 6. STMs illustrate different states of the videos in the form of circles with accompanying percentages. States are comprised of single or multiple co-occurring visual, linguistic or multimodal systems, such as the occurrence of an Actor Participant within both the visual and linguistic Experiential Metafunctions (Tan et al. 2016:260), therefore a coupling, see 3.3. Transitions are represented by lines and highlight relationships that exist between states and the frequency of states (Tan et al. 2016:260). These are important because they show how various modes relate to one another in the video, which is the answer to research question 2 and part of the answer to research question 3 as to how they relate to student learning.

I now discuss a description of how the systems and modes were selected for analysis. As mentioned above, the videos were analysed using "Multimodal Analysis Video" and the catalogue within the software formed my coding scheme for the videos that was adjusted in order to answer research questions 2, 3 and 4 of this research. The original catalogue has a Composition system that examined the visual compositional structure of video in terms of a framework from film analysis and a system for Auditory Elements that examined sound, music and song, but since the modes of sound and music were not mentioned as helping or hindering learning by the students, they were left out. Similarly, the system of Composition was found to not be useful in terms of examining how the various modes worked together to answer research question 2 and was consequently removed. The Inter-Relations system was only differentiated in the sample catalogue according to Similarity and co-contextualisation

and Difference and re-contextualisation. I found this system to not be detailed enough in terms of illustrating the implication sequences behind these concepts. Therefore, as mentioned in 3.3., Comparative and Temporal Relation systems were created as systems of Inter-Relations for annotation and discussed in terms of co-contextualisation, with re-contextualisation being assumed when these did not occur. The adjusted catalogue used in this DMDA is depicted in Figure 7. Explanations of these systems are found in 3.3.

Figure 7: Catalogue for DMDA used in this research within Multimodal Analysis Video

Videos were annotated according to the systems depicted in Figure 7 for each Participant, represented as tabs, and then once this was completed, the STMs required for analysis were produced by the software and saved. Further adjustments had to be made to deal with a bug in the software, which caused the STMs to not read accurately when multiple speaker tabs were annotated in a single annotation system strip. This resulted in multiple system strips, one per speaker, being created as pictured in Figure 7 (three Experiential Meaning systems in Figure

7, three information value systems in Figure 7, and three Comparative and Temporal Relation systems not pictured above). Speaker tabs were differentiated by colour and in the STMs they would be presented in one of the strips only, thus these repetitions of systems in this catalogue are only present to work around the bug in the software and to ensure accuracy of the STMs. Debut Video Capture software was used to record digital appendices and PRAAT, a free software for phonological speech analysis, was used to separate speech from silences and to identify the speech time of the total amount of words from an audio clip of the video in order to accurately measure the speech rate of the video clips. Speech rate was calculated, as discussed in 3.3., by dividing the number of words by the time of utterance. While the number of words was counted, the total duration of utterances was determined in PRAAT, dividing this number by 60 to convert it from seconds to minutes. Since the average speech rate was something that needed calculating via a formula and could not be calculated via STM, it was not added to the catalogue in "Multimodal Analysis Video", rather the total speech time was calculated from PRAAT as mentioned above and the calculations are reported in Chapter 5.

In terms of the system of connotation, which is described in 3.3., the video clips were examined as a whole to determine what connotative meanings were being portrayed. The system of connotation is particularly dependent on my interpretations within the DMDA and therefore it is necessary to explain how these interpretations were made consistent. Referring to denotative meaning, from which connotative meaning is derived, gender as a social construct was a prominent factor in all three videos. Each video contained one Participant of each sex. Thus, the videos potentially evoke gender connotations and stereotypes that create stimuli for student impression formation (represented in PCSTIF schemas used in the thesis). Due to the fact that race, class and other similar social constructs were not denotatively prominent features in the videos, the connotations of these social constructs consequently were not elaborated upon. The parts of the video were annotated in which these meanings made themselves apparent. Then a fine-grained examination occurred to determine which of the other visual, linguistic and multimodal systems in the DMDA contributed to providing these connotations. These systems were then selected alongside the specific connotation within the STMs to examine how often these systems overlapped/co-occurred with the connotative meanings identified in the video clips.

Likert scale data within the questionnaires were compiled and represented with graphs in Chapter 5 using Microsoft Excel. The rest of the written information in the questionnaires and focus group interview transcriptions was analysed with a thematic analysis, described according to Braun and Clarke (2006:16-23). Furthermore, Vaismoradi et al. discuss the difference between content analysis as a more descriptive approach that focuses on the quantitative whereas they state that a thematic analysis is "purely qualitative, detailed and nuanced" (2013:400). Since my population group was small, and the research questions demanded information that was qualitative in nature, thus not requiring a larger data sample, the data in this study is analysed qualitatively over a quantitative approach. I consequently chose to perform a thematic analysis over a content analysis, which is the opposite of what Abrams (2016:353) did with her larger amount of data in her case study, for a detailed, qualitative approach. I first familiarised myself with both questionnaires and transcriptions, generating initial codes. I then searched for and reviewed themes, which according to Braun and Clarke, "[capture] something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set" (2006:10). I then defined and named themes and describe these in Chapter 5. A list of themes from questionnaires and transcriptions is provided in Appendix F.

I have used the PCSTIF schemas in Chapter 6 to attempt to clarify how the students form their cultural impressions while watching the videos (see description of PCSTIF in 3.1.). The schemas are attempts by me as an analyst to explain how the students reach the impressions they do. They were constructed based on the theory discussed in 3.1. and the bottom node of my schemas (see Figure 36, for example, in 6.1.) reflects the overall cultural impressions students made of a video and they are in green. These are formed by associations, which are the layer above the green node (e.g. Figure 36 in 6.1.), for example, "friendly". Associations are derived based on observations or stimuli which are the red nodes on the topmost layers of the schemas (see Figure 36 in 6.1.). In my adaptations of the PCSTIF schema, I include associations (i.e. the middle-layer of the schemas) from both myself as the analyst and from the students in order to convey a variety of possible associations from which students could have formulated their impressions. I created the schemas in Paint in Windows, as it did not require elaborate design and it was an easy-to-use, freely available software with which I was familiar.

Section 4.4 Conclusion

This chapter discusses the research design of this thesis, in which a DMDA of three German pedagogic video clips and a thematic analysis of questionnaire and transcription data from two semi-structured focus group interviews were undertaken.

Once STMs were derived from the DMDA, screenshots were taken for illustration (see, for example, Figure 10 in 5.1.). The percentages of states would not only illustrate how frequent one particular system is present within the video clips, but also shows how frequent co-occurrences between systems and modes are, thus providing a representation of the relationship between these, which is what research question 2 seeks to achieve. Through examining connotation and the multimodal systems that are employed to provide connotative meanings in the videos, one can then also understand how culture(s) of German-speaking countries are represented in the videos – an aim of question 4. Furthermore, after themes were identified after the thematic analysis of both questionnaires and student transcriptions, student perceptions on the videos were identified, and the learning opportunities and challenges of learning with pedagogic videos allowed for question 1 to be answered. The triangulation of the data sets from two different types of analysis allowed for questions 3 and 5 to be answered, understanding how the relationship between various modes influence student learning and what cultural assumptions and impressions students make based on the cultural representations within the videos. Once potential problem areas and challenges for language learning and cultural knowledge acquisition within the DMDA are identified and confirmed from a thematic analysis of student responses, I discuss and interpret findings individually in Chapter 5. I then interpret these findings in the light of the literature, which is described in Chapter 2, in Chapter 6.

Chapter 5 - Findings

This chapter describes the findings of both the Digital Multimodal Discourse Analysis (DMDA) of the videos and the thematic analysis of the student questionnaires and interview data. 5.1., 5.2. and 5.3. describe the results of the DMDA of Videos 1, 2 and 3 respectively, with each section illustrating relevant State Transition Machines (STMs). As mentioned in Chapter 4, the software "Multimodal Analysis Video" produces STMs, which are diagrams that highlight the different states in the videos. They illustrate how the various systems within the modes interact with one another, with circles as states and the percentage they represent in the video and lines as transitions to show the relationship between states and how many times states occur (see Figure 10 on the next page, for example). PRAAT screenshots clarify the total duration time of utterances used to measure the average speech rate. 5.4. describes the answers from student questionnaires thematically categorised within graphs along with some references made to student focus-group interview comments that relate to questionnaire answers directly. 5.6. discusses both Focus Group A and B's findings in terms of themes. Chapter 6 presents a deeper analysis of these findings, as while this chapter focuses on discussing the results of the DMDA of videos and themes discovered in student interview answers, Chapter 6 seeks to establish a triangulation between student data, video data and context (i.e. literature).

Section 5.1. Video 1

In Video 1, the male character in the video, Oliver, assists a tourist with directions to the Goetheplatz. However, she is confused by the directions, despite the repetitions, to the extent that he has to take her there himself, with repeated directions and arrows being shown for viewers to fully comprehend. At the end of the video, she does not thank him, but instead simply walks away, much to Oliver's annoyance with his final remarks of "Sehr nett. Vielen Dank. Danke!" (Very nice. Thank you very much. Thank you). See Appendix G for a link to the transcription and translation of Video 1.

Figure 10 depicts the language systems of the experiential metafunction (i.e. Processes, Participants and Circumstances) and their relationship with one another, as well as the different speakers in the video.

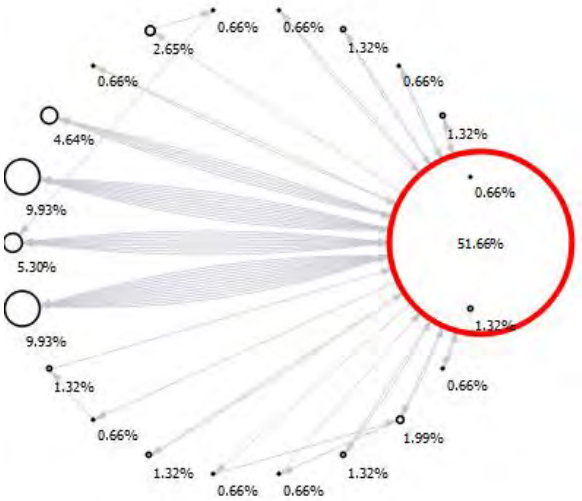

Figure 10: Video 1 Experiential Metafunction Language

Within state machines such as Figure 10, a 'state' (denoted by the circles) represents the system choices that have been utilised for a particular semiotic resource system or a combination of semiotic systems, in terms of total video time (e.g. see Tan et al 2015). There are two 9.93% states in this video: Circumstance of Location and Circumstance of Extent in Material clauses for Oliver. Therefore, in this particular case, it means that Circumstances of Location and Circumstances of Extent each accounted for 9.93% in terms of total video time. These are directions Oliver gives the tourist. For example, "fünfzig Meter geradeaus" (fifty meters straight ahead) has a Circumstance of Extent and "Dann an der Ecke nach rechts" (then turn right at the corner) has a Circumstance of Location. The Actor Participant for this Process in all of these clauses has appeared once and it is *Sie* (you), referring to the tourist and is elided. The states for the tourist include a 5.30% state, which refers to a Material Process with Circumstance of Extent, and the 4.64% state refers to Material Process with Circumstance of Location. The Participant for this Process has never appeared, and it is *ich* (I), referring to the tourist and is elided. For example, the repetitions the tourist makes of Oliver's directions "fünfzig Meter geradeaus" (fifty meters straight ahead) has a Circumstance of Extent and "Dann an der Ecke nach rechts" (then turn right at the corner) has a Circumstance of Location. Therefore, one can argue that the tourist is the most frequent Participant in Video 1. The 2.65% refers to a Material Process with Participant Actor and Circumstance of Extent for Oliver. For example, for the sentence "zuerst gehen Sie fünfzig Meter geradeaus" (first go fifty meters straight ahead) and this reveals a preoccupation in the video on Circumstance of Extent and Location. These system choices are necessary when giving directions for one discusses distance needed to travel, as well as the locations needed

to pass until one reaches the intended destination. Within the video, 51.66% is uncategorisable for the Experiential Metafunction, due to either the language not able to fit into these categories, or no language being present in these moments at all.

Figure 11 show that 7.10% refers to Action Processes with the tourist as Participant Actor.

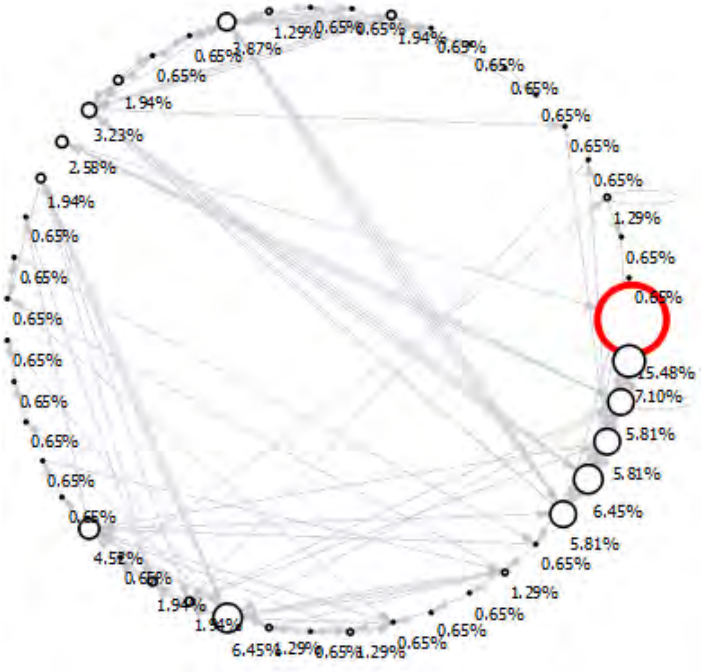

Figure 11: Video 1 Experiential Metafunction Visual

The tourist as Participant Actor is seen visually, for example, when the tourist walks in the beginning of the video. The 6.45% state refers to Oliver as Actor in Action Processes with the tourist as Reactor in Reaction Processes – as these two Processes occur simultaneously. This refers to scenes when Oliver is the speaker, and the tourist is the listener. Another 6.45% state below the aforementioned one in Figure 11 shows an Interaction Process by Oliver, who then shifts between Actor and Reactor Participants (speaker and listener) and the tourist's gestures as Conceptual Processes with a Concept Participant. There are three 5.81% states: (1) an Interaction Process by the tourist with her Participant status shifting as Actor and Reactor (speaker and listener) with Oliver as Actor in Action Processes (e.g. moving his hand to his face or walking), (2) Oliver as Actor in an Action Process (e.g. speaking) and (3) both the tourist and Oliver as Actors in Action Processes (walking together). Within the video, 4.52% includes Oliver's gestures with Conceptual Processes and Concept with the tourist as Actor in an Action Process (speaking). When separating the bright yellow arrows, which appear in the centre shot, into their own State Transition Machine (STM) and thus treating them as separate

Participants, 10.96% of the video includes the arrows as Concept Participants with Conceptual Processes. Action and Interaction Processes are thus dominant, with Conceptual Processes not far behind, indicating the focus to be on directions and movement in Video 1. Within the video, 15.48% is uncategorisable for the visual experiential metafunction.

When examining Video 1 and Mood, 28.28% of the video is comprised of Declaratives made by Oliver and 10.34% of the video is comprised of Declaratives made by the tourist. For example, “An der Ampel nach links” (turn left at the traffic light) which Oliver says and the tourist repeats and 2.76% is made up of Interrogatives made by the tourist. Modality only appears in the 0.69% state with the verb “ n” (to be able to) in an Interrogative by the tourist and as such is not discussed in detail in the thesis, which tells us the levels of obligation are not high in this clip but that also the language is simple for A1 level. Within the video, 53.10% is uncategorisable for Mood. Referring to Themes, 26.53% of Themes are in Oliver’s statements; the tourist is the Theme, even when elided, for example “Sie gehen jetzt fünfzig Meter geradeaus” (you now go fifty meters straight ahead). An 11.56% state refers to the tourist as the Theme for the tourist’s statements in a similar fashion when she repeats what Oliver has said. The three 1.36% states from bottom right in a clockwise direction are: (1) *wo* (where) in utterances of the tourist, (2) Goetheplatz in Oliver’s utterances and (3) *aufpassen* (to pay attention) in Oliver’s utterances. Within the video, 53.06% is uncategorisable for Theme. The tourist is thus the most dominant Theme in Video 1.

Figure 12 presents all three linguistic metafunctions, the experiential, interpersonal and textual, as well as how they relate to one another in Video 1.

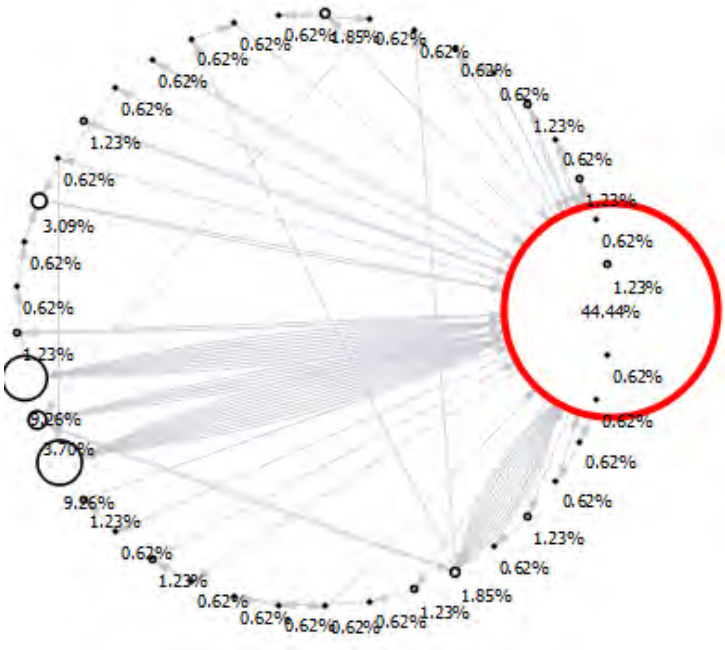

Figure 12: Video 1 Experiential, Interpersonal and Textual Metafunctions for Language

The bottom 9.26% state refers to Material Processes by Oliver with Circumstances of Location in a Declarative Mood and with the Theme of the tourist. An example would be "an der Ampel nach links" (turn left at the traffic light). The other 9.26% state refers to Material Processes by Oliver with Circumstances of Extent in a Declarative Mood and with the Theme of the tourist. Within the video, 3.09% illustrates Material Processes by the tourist with Circumstances of Location in a Declarative Mood and with the Theme of the tourist. The 1.85% state with multiple transitions refers to the utterances made by the tourist in the Declarative Mood. The patterns of co-selection revealed in Figure 12 are referred to as cross-metafunctional couplings in SFL and cross-metafunctional states in DMDA and these are further discussed in 6.1. Within the video, 44.44% does not contain language and so is uncategorisable.

In terms of shots, 61.64% of the video depicts stationary shots with 12.33% dolly shots, stationary shots appear when Oliver and the tourist are conversing, and the dolly shots occur when they are walking. In 6.16% of the video the arrows, zoom in during dolly shots and in 4.11% of the video they zoom in during stationary shots, illustrating the focus on the arrows and directions in the video. A 57.82% state is angled and eye level during the majority of Oliver and the tourist's conversation. A 21.77% state is front and eye level when focusing on a specific person. 3.40% of the video is front and low angle where the focus is on Oliver, for example, and the camera is tilted to look up towards him. This invites viewers into the

conversation with Oliver and the tourist. In 39.46% of the video, Oliver and the tourist both have indirect gaze within a medium shot as they look at one another. In 19.05% of the video, Oliver and the tourist both have indirect gaze within a close shot as they continue to gaze at one another and in 18.37% of the video, both have indirect gaze with a long shot when walking together. The shots indicate a level of formality with the audience and gaze alludes to information being provided to or demanded from the audience depending if it is direct or indirect gaze. In terms of visual information value and saliency, the tourist is standing on the left as Given information, which is during their conversation for the majority of the clip and 56.46% of the video is Oliver (standing on the right) as New information. An 8.84% state reflects Oliver as Given information and the tourist as New information (this happens when they walk together). Within 6.12% of the video Oliver is seen as New information and the tourist as Given information and the arrows illustrating Colour Contrast, Foreground and centred during their conversation. A 3.40% state represents Oliver as Given information and the tourist as New information, and the arrows illustrate Colour Contrast, Foreground and centred when they are walking together. The fact that Oliver is represented as Given information most of the time confirms that he is the dominant speaker of the conversation and the provider of directions, while the tourist's position as New signals that she is the receiver of the information.

Figure 13 depicts Similarity and Contrast between linguistic and visual meanings within these modes.

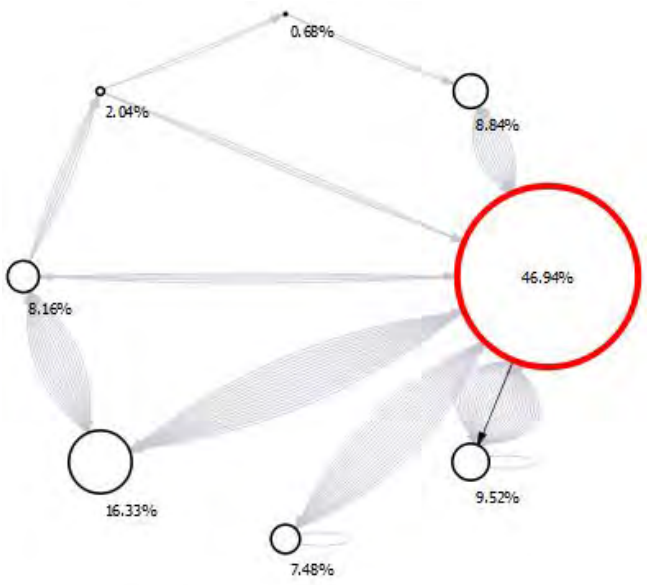

Figure 13: Video 1 Language – Visual Comparative relations

The state 16.33% is Similarity of linguistic and visual meanings for Oliver in terms of speaking and gestures during conversation and often walking. The state 9.52% is Contrast of linguistic and visual meanings for the tourist. The state 7.48% is Contrast for Oliver. In 8.84% of the video Similarity occurs for the tourist. The 8.16% is Similarity for Oliver and for the arrows. 2.04% is Similarity for the arrows only. In terms of Temporal relations, 16.44% of the video is Simultaneity between Oliver speaking with his gestures. In 10.27% of the video, Simultaneity occurs between Oliver, his gestures and the arrows, signifying that the meaning is co-occurring in multiple modes. In the video, 46.94% is uncategorisable, according to Figure 13.

Figure 14 illustrates where connotations and SFL metafunctions overlap.

Figure 14: Video 1 Connotations and all metafunctions for SFL

In 12.21% of the video, the idea that Germans are helpful, friendly and patient is portrayed, the value that people should be helpful, friendly and patient, and potential positive male stereotypes and this is mainly found in Oliver's language actions in helping the tourist by providing her with directions. These ideas, values and connotations about Germans being helpful, friendly and patient are also found in the two 5.23% states at the bottom of the STM in connection with Material Process with Circumstance of Location in the Declarative Mood and the Theme the tourist, as well as Material Process with Circumstance of Extent in the Declarative Mood and the Theme the tourist. The two 3.49% states also represent the above

within the video. In the video, 2.91% of these values, ideas and gender connotations are connected to the tourist's Material Process with Circumstance of Location in the Declarative Mood and the Theme the tourist. A 19.77% state of the video is uncategorisable (see Figure 14).

The total duration of utterances to the second decimal point is measured as 80.96 seconds (see Figure 15). This divided by 60 equates to 1.35, and this was then used in the formula to reach average words per minute. The number of words in Video 1 is 192. Thus 192/1.35 = 142.22 average wpm.

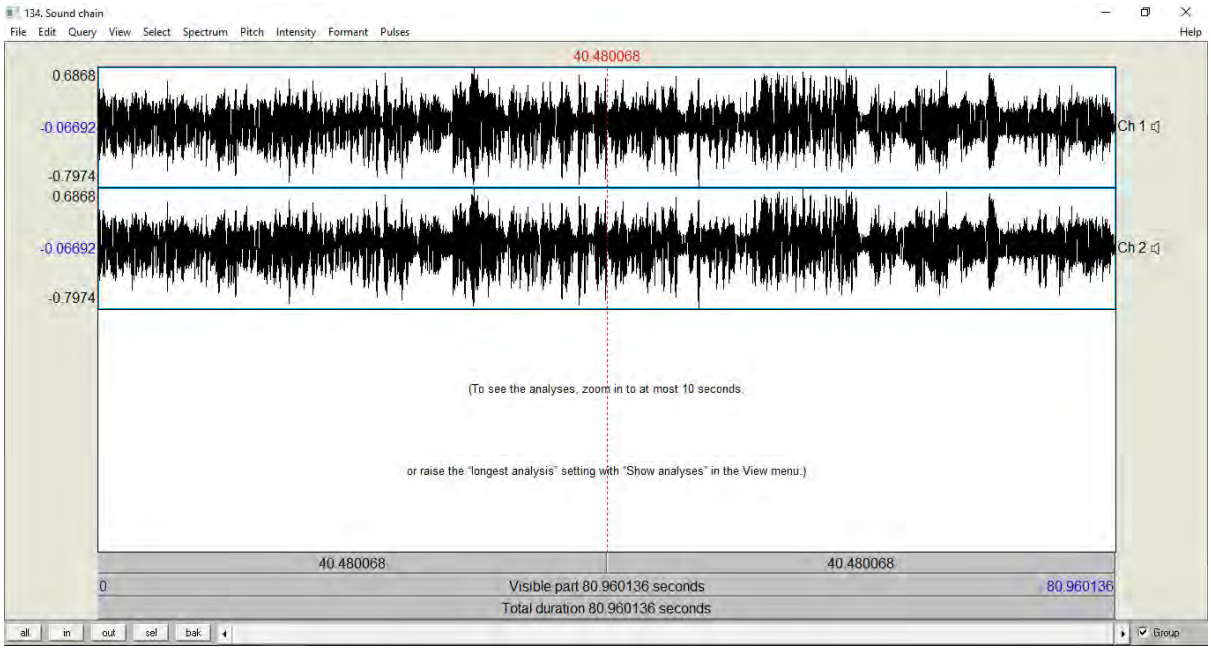

Figure 15: PRAAT sound concatenate of Video 1

Section 5.2. Video 2

The entirety of Video 2 is comprised of camera movements and angles meant to show viewers through the eyes of Mr Waurich as he views Ms emann's *Superwohnung* (super flat), the title of the clip. The odd messy scenes as she shows him each room, however, cause him to eventually run out of the flat after she announces that she wants to show him the bathroom. See Appendix H for a translation of Video 2 and a link to the transcription of the video.

Figure 16 elaborates on the states representing systems of the Experiential Metafunction for language (i.e. linguistic Participants, Processes and Circumstances) in relation to their speakers.

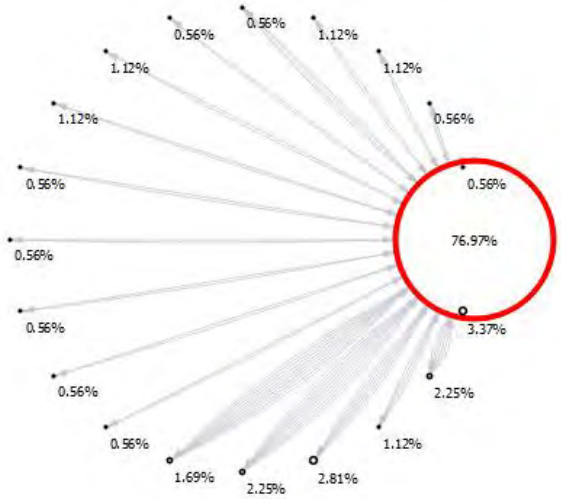

Figure 16: Video 2 Experiential Metafunction Language

All these utterances are by Ms Möllemann, the only speaker. In the video, 2.25% of these utterances are categorised under Material Processes with Actors and no Goal Participants, for example, "Kommen Sie mit" (come with me). Of the total utterances, 3.37% contain identifying Relational Processes, such as "Das hier ist der Flur" (this here is the hallway). Within the video, 2.81% of utterances are attributive Relational Processes and Circumstances of Manner, for example, "Er ist nicht sehr groß" (he/it [the hallway] is not very large). The remaining state of 2.25% nestled before the 3.37% state is comprised of Mental Processes, including Sensor and Phenomenon Participants, for example, "Sie hten die Wohnung ansehen" (you would like to view the flat). According to the STM of Figure 16, only 23,03% of the video is made up of spoken language that has been categorised under SFL categories, with 76,97% of the video being comprised of language uncategorisable (for example, the greeting "Hallo!" (hello), that is neither a Participant, Process or Circumstance), pauses and no utterances. This STM exhibits the predominant Processes and Participant combinations in Video 2.

In this video, 34.44% are visual Action Processes with the Actor being Ms Möllemann, for example, when she is walking around showing Mr Waurich the flat. This illustrates her to be a focus and emphasises her dominant role in the video. The most predominant Mood is Declarative (16.67%), such as "Das ist super!" (that is super!). This is followed by 5%

Imperative, for example, “Kommen Sie” (come), and 1.67% Interrogative, for example “nicht wahr?” (isn’t it?). This shows the presence of giving of information, rather than requests and commands. The most spoken Theme is “das Wohnzimmer” (the living room) with 2.78%. In the video, 2.22% of the video include Themes of Ms Möllemann, and “Kommen” (come), and “Sehen” (look/see), with verbs in the Imperative Mood, e.g. “Kommen Sie” or “Sehen Sie.” Examples of such a theme for Ms emann include “ich mag das Zimmer” (I like the room). In the video, 1.67% of the video’s Themes are “Der Flur” (the hallway), for example “er ist nicht sehr groß” (he/it [the hallway] is not very large).

Figure 17 presents all three linguistic metafunctions, the experiential, interpersonal and textual, as well as how they relate to one another in Video 2.

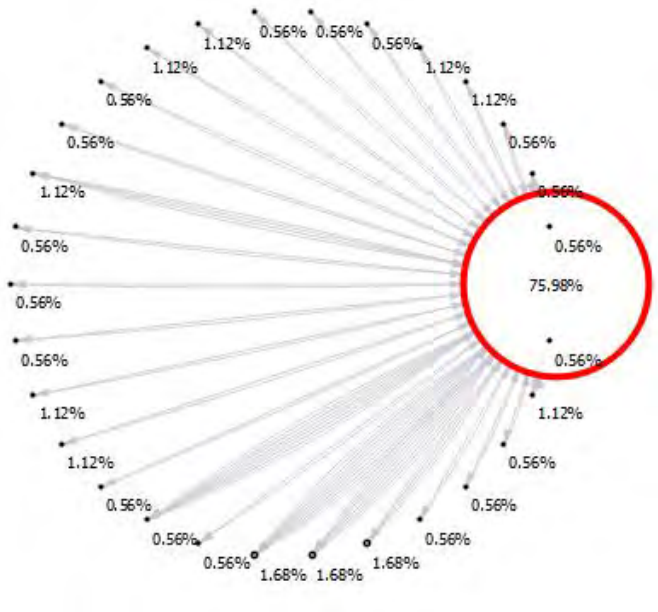

Figure 17: Video 2 Experiential, Interpersonal and Textual Metafunctions for Language

As shown in Figure 17, there are so many possible states that could exist that each of these states is present only for a very small proportion of the video. I thus describe the significant states from Figure 17 for the purpose of examining how frequently the metafunctions for language intersect with one another and thus present more cross-metafunctional couplings/states. The largest 1.68% states starting from the rightmost state in a clockwise direction include the following:

- Attributive Relational Process with Manner, Mood: Declarative Theme: “Der Flur.” For example, “Er ist nicht sehr groß” (he/it [the hallway] is not very large).

- Actor-Material Process, Mood: Imperative, and Theme: “Kommen.” For example, “Kommen Sie” (come).
- Actor-Material Process, Mood: Imperative, and Theme: “Sehen.” For example, “Sehen Sie” (look).

These couplings illustrate how cross-metafunctional states work together to create specific meanings, which are further analysed in 6.2.

Figure 18 highlights the states that concern themselves with zooming in and out, as well as camera movements.

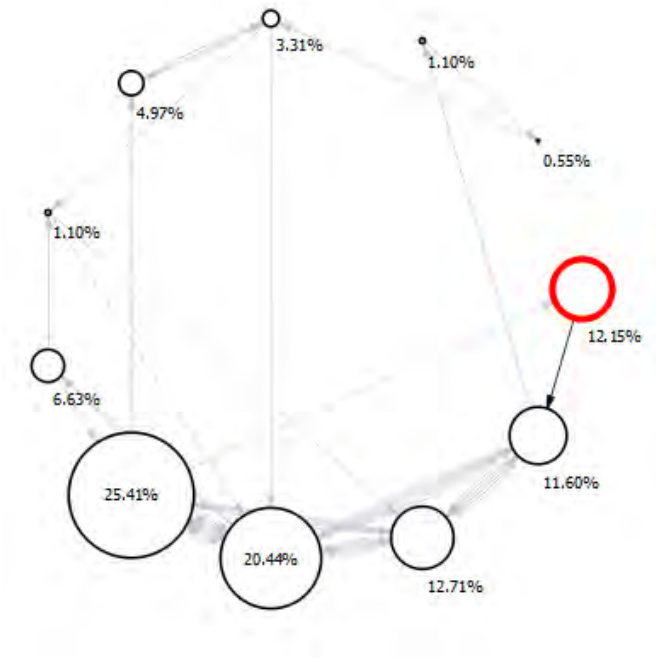

Figure 18: Video 2 Zoom and Camera Movement

Representing the view of Mr Waurich, 12.15% of the video include shots where there is no camera movement as shown in Figure 18, for example as at the beginning of the film, when no shots occur. Of these shots, 25.41% include Dolly or handheld movement – as when Mr Waurich is moving up the stairs. In the video, 20.44% of shots include pan movement and 12.71% are stationary. Pan movement shots occur at times when Mr Waurich is focusing on a specific area of the flat, for example, looking at the messy cupboard, and stationary shots often occur when, for example, the camera is focused on Ms Möllemann. In the video, 11.60% of shots include tilt movements, with 6.63% of shots panning and zooming in simultaneously and 4.97% of shots tilting and zooming in simultaneously. Of these shots, 3.31% include tilting and zooming out simultaneously and the 1.10% states ranging from left

in a clockwise direction include the combinations of zoom out and pan, as well as zoom in and stationary. Within the video, 0.55% of shots are zoom out and stationary, the least prevalent state. These all occur during times Mr Waurich is focused on specific areas of the flat, for example the stain on the bedroom floor, a room which is particularly messy, thus emphasising the messiness.

Figure 19 presents states that show horizontal and viewing perspective and their relationships with one another.

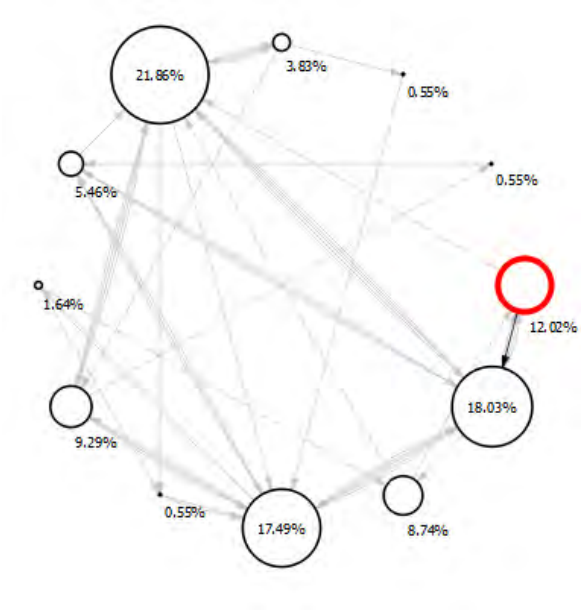

Figure 19: Video 2 Horizontal and Vertical Viewing Perspective

In terms of horizontal and viewing perspectives representing the view of Mr Waurich, Figure 19 illustrates that 12.02% of the video have none of these viewing perspectives. The majority state, 21.86%, includes high angle shots, for example a focus on the stain on the bedroom floor. A 18.03% state represents eye level, with the 17.49% state representing the combination of front angle and eye level when looking at Ms Möllemann. The 9.29% state depicts front and high angle combinations and the 8.74% state low angle, also when looking at Ms Möllemann. Of these shots, 5.46% shows a combination of angled high, and 3.83% angled eye level as the camera shifts up and down when viewing the hallway, for example. In the video, 1.64% of shots are front low angle, for example looking up at Ms Möllemann. The frequency of shots from a high angle indicate disdain from Mr Waurich towards Ms Möllemann and the messy flat, and there is more involvement over detachment between the characters and the viewers due to more front shots being present than angled shots. To describe this, Kress and van Leeuwen state:

> The difference between the oblique and the frontal angle is the difference between detachment and involvement. The horizontal angle encodes whether the image-producer (and hence, willy-nilly, the viewer) is 'involved' with the represented participants or not. The frontal angle says, as it were, 'What you see here is part of our world, something we are involved with.' The oblique angle says, 'What you see here is not part of our world; it is their world, something we are not involved with.' (Kress and van Leeuwen 2006:136)

The amount of involvement over detachment in this video and Video 3 could add to the reason why students' impressions of Germans were friendly and helpful mentioned by students in 5.5. and 5.6.

In the video, 18.23% of shots include direct gaze in combination with long shot when viewing Ms Möllemann, for example in the living room. An 11.60% state illustrates indirect gaze and long shot combinations. This is also visible in the living room when she is looking out of the window. A 3.87% state is direct gaze and a 3.31% state represents long shot. Ms Möllemann does not appear in a shot aside from long shot. A 1.66% state represents indirect gaze. This indicates the formal nature of the relationship between Ms Möllemann and Mr Waurich. In terms of gaze and social distance of Ms Möllemann, 61.33% of the video shows none of these systems.

Information value (or position of Participant in frame), framing and saliency, lighting and contrast occur around Ms Möllemann and are depicted by Figure 20.

Figure 20: Video 2 Information Value, Framing, Saliency (Framing cluster)

The 56.28% represents a state where none of these are present. The 8.20% state represents where Ms Möllemann is positioned centre. Within 6.01% of the video she is Given

information. With the state of 5.46% begins what I have termed “Framing clusters”, states in which saliency, information value and framing intersect. These are syndromes, couplings that are returned to periodically during the course of the video. This state includes lighting, centre and frame lines due to positionality. The frame lines are a door frame within which she stands. The 4.92% state includes Ms Möllemann placed centre and frame lines due to positionality. The 3.83% state includes brighter lighting in addition to all the characteristics mentioned previously. The 3.28% state has Ms Möllemann as New information. There are two 2.73% states from bottom left in a clockwise direction: (1) frame lines due to positionality and (2) another framing cluster of colour contrast and lighting, Given and frame lines due to positionality. The 2.19% state shows Ms Möllemann as Given information and frame lines due to positionality. There are also two 1.64% states from top left in a clockwise direction, (1) colour contrast and lighting, Ms Möllemann centre and frame lines due to positionality, as well as (2) lighting and Ms Möllemann centre. The 0.55% state includes lighting, Ms Möllemann as Given and frame lines due to positionality, hence yet another framing cluster. These provide more emphasis on Ms Möllemann during times she tries to draw Mr Waurich’s attention away from the messiness of the flat.

In terms of Comparative relations of the video with language and visuals, in 23.89% of the video, language contrasts with visuals and in 6.11%, Similarity between language and visuals is shown, while 70.00% are just visuals without language with which to compare. In terms of Temporal relations, there is also 6.11% Simultaneity where Ms Möllemann presents the rooms when referring to them, signifying there is less of a simultaneous relationship between language and visuals than in Video 1, but that there is Simultaneity present in the video. In the video, 27.07% portrays the idea that German women are disorganised with the value connotations that people should be neat and tidy. Thus, these connotations may give rise to stereotyping against women. This mainly occurs during zooms and camera movements that are followed by Ms emann’s interjections to regain Mr Waurich’s attention. In the video, a 6.59% state includes these ideas, values and gender roles simultaneously evoked from zoom in and pan and 6.04% of the video includes these ideas, values and gender roles simultaneously invoked from pan only. In the video, 4.40% includes these ideas, values and gender roles and tilt.

The total duration of utterances to the exact decimal is measured as 36.76 seconds (see Figure 21). This divided by 60 equates to 0.61, which is used in the formula to reach average words

per minute. The number of words in Video 2 is 189. Thus 189/0.61 = 309.84 average wpm, faster than Video 1.

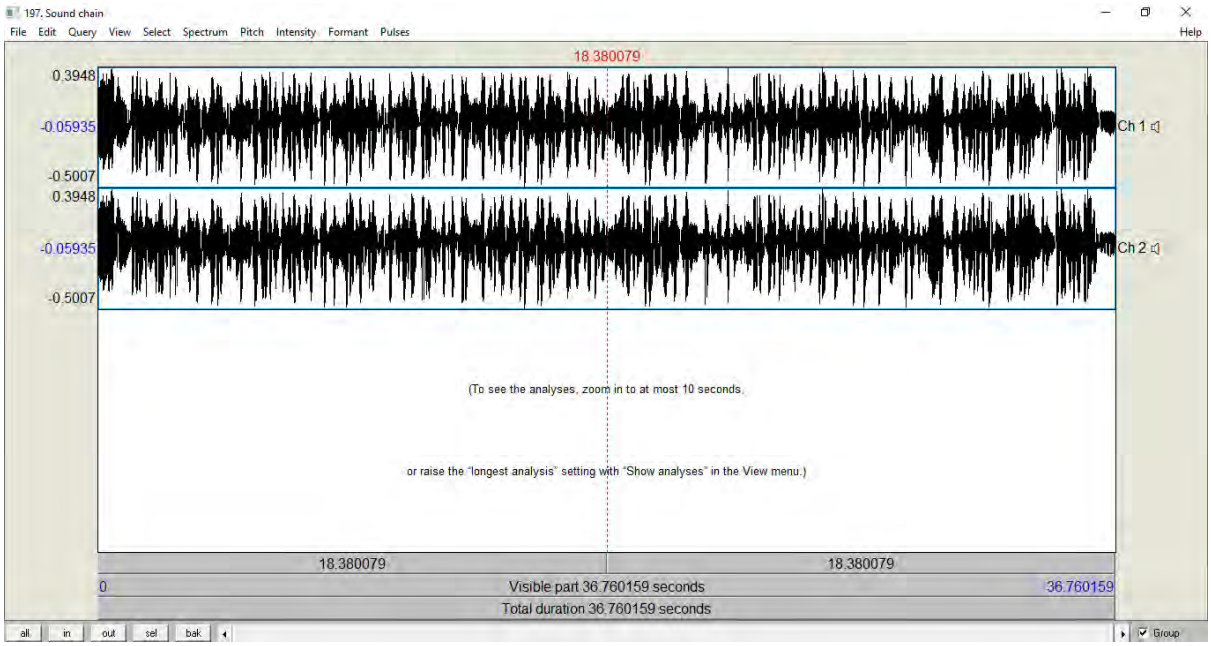

Figure 21: PRAAT sound concatenate of Video 2

Section 5.3. Video 3

In Video 3, Martin Zürchner, a character in the video, shows viewers his city of origin – Bern in Switzerland. He depicts what *Bärndütsch* (Bernese German) sounds like with an example utterance from Salome, the second character in the video, shows viewers "der Zytglogge" (the clock tower) and illustrates that both the Berlin and Bern coats of arms have bears on them, before showing viewers actual bears in bear pits. He ends off his tour in front of the Federal Palace of Switzerland, which he does not name or mention, but rather asks learners what they thought of Bern and says that they should travel to it when in Switzerland. See Appendix I for a translation of Video 3 and a link to the video transcription of the video.

Figure 22 depicts the predominant linguistic Processes and Participant combinations in Video 3.

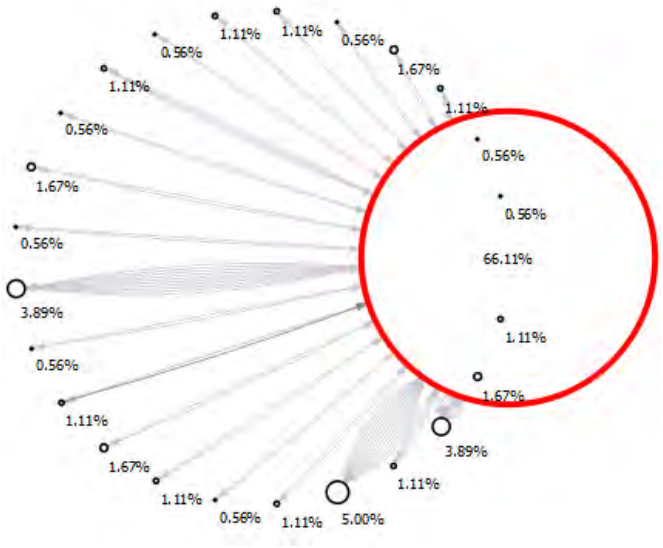

Figure 22: Video 3 Experiential Metafunction Language

According to Figure 22, language that can be analysed using SFL is present for 33.9% of the video. This language is just from the dominant speaker Martin, as Salome's utterance "a lieba Grüeßa a au wiewe Dütschleera" (loving greetings to all German learners) does not classify as a clause and the narrator only has one clause "Das ist meine Stadt" which is an identifying Relational Process (a 1.11% state), in the Declarative Mood (a 1.11% state, not pictured in Figure 22) and has Bern as a Theme (a 1.11% state, also not pictured in Figure 22). The 5.00% state represents attributive Relational Processes with Carrier and Attribute. The 3.89% states beginning from bottom right in clockwise direction include (1) identifying Relational Processes with Identifier and Identified, for example, "Mein Name ist Martin Zürchner" (my name is Martin Zürchner). (2) Mental Processes with Sensor and Phenomenon, for example „Wollt ihr mal einen Satz auf Bärndütsch en?" (do you want to hear a sentence in Bernese German?). In the video, 66.85% of language is uncategorisable or there are pauses or simply no utterances.

Within 40.98% of the video there is solely Action Processes and the Actor is Martin, for example Martin speaking into the camera. There is 2.19% of the video that records moments when both Actors Martin and Salome are completing Action Processes and 0.55% is only when Salome as Actor completes Action Processes. Conceptual Processes with the coat of arms as the Concept Participant form 2.73% of the video and a 1.64% state is Salome participating in Action and Reaction Processes as Actor and Reactor. In the video, 46.99% is

uncategorisable for visual Participants and Processes. Thus, this illustrates that most of the visual Processes are Action with a few Conceptual Processes.

As Salome's utterance does not classify as a clause, both Mood and Themes are also only from Martin. In the video, 25% is Declarative, with 6.67% Interrogative and 1.11% is Imperative and 67.22% is unclassified. An utterance in the Declarative Mood is, for example, "Die Berliner haben einen Bären" (the Berliners have a bear). An example of an utterance in the Interrogative Mood is "Wollt ihr mal einen Bären sehen?" (do you want to see a bear?) and an example of an utterance in the Imperative Mood is "Und seht mal:" (and see). This shows information is provided more than demanded in the video.

Figure 23 depicts the linguistic Themes found in Video 3.

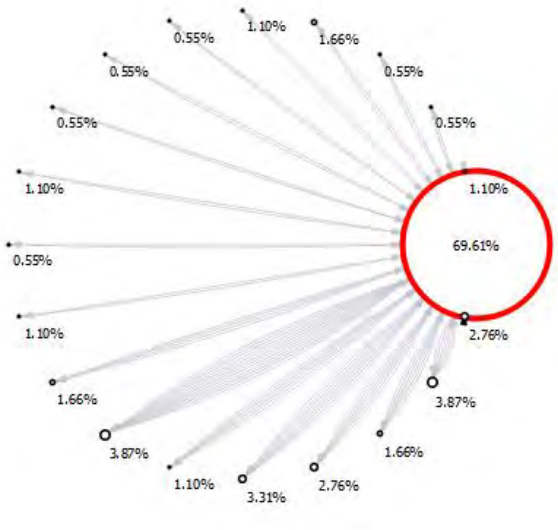

Figure 23: Video 3 Themes

In Figure 23, the 3.87% states from bottom right in a clockwise direction are Bern and "der Zytglogge" (the clock tower), for example, "Bern hat 130 000 Einwohner" (Bern has 130 000 residents) and "Das ist der Zytglogge" (this is the clock tower). The 3.31% state represents the viewers as Theme, termed *Leute* (people) because Martin greeted the viewers as "Hallo Leute" (hello people). The 2.76% states represent Martin as Theme, for example, "und ich lebe hier" (and I live here) and *Bärndütsch*, for example, "Bärndütsch, das heißt Berner Deutsch" (Bärndütsch, that is Bernese German). The 1.66% states from the closest to the uncategorisable state in a clockwise direction represent "kommt" (come), such as in "kommt einfach mit!" (come with!) and "wie" (how), such as in "wie gefällt euch Bern?" (how do you

like Bern?). The 1.10% state with 6 transitions is about Salome, such as “Salome, sag mal was!” (Salome, say something!). The 69.61% state represents the uncategorisable, as these are parts of the video with no spoken language or non-finite/non-clausal utterances such as Salome’s utterance.

Figure 24 presents all three linguistic metafunctions, the experiential, interpersonal and textual, as well as how they relate to one another in Video 3.

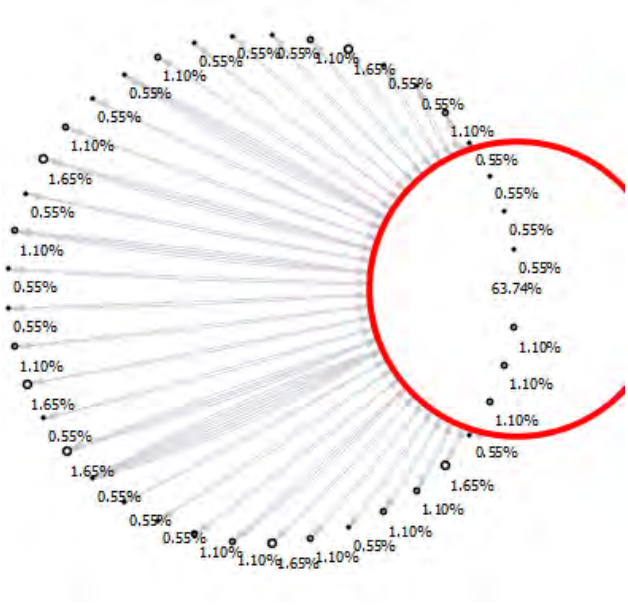

Figure 24: Video 3 Experiential, Interpersonal and Textual Metafunctions for Language

Noteworthy states in Figure 24 include 1.65% of just the Interrogative Mood and another that includes Mental Processes with Sensor and Phenomenon in the Interrogative Mood with “wie” (how) as Theme, for example, “wie gefällt euch Bern?” (how do you like Bern?). There are two 0.55% states. The one with the most transitions is only the Declarative Mood and the other includes Mental Processes with Sensor and Phenomenon as Interrogative and the Theme is “Leute [ihr],” (people[you all]) for example, “Wollt ihr mal einen Satz auf Bärndütsch n?” (do you all want to hear a sentence in Bernese German?). These states illustrate the majority of cross-metafunctional interactions with Mental Processes.

In the video in relation to zoom and camera movement, 79.33% contains stationary camera movement, for example when focused on Martin and Salome. A 6.15% represents stationary and zoom out and a 0.56% represents pan. This shows that there is very little camera movement that occurs in this video. In the video, 13.97% is uncategorisable for zoom and

camera movements, due to the introduction and conclusion of the video showing a black screen, as well as scene transitions and fade outs. In terms of viewing perspective, 53.63% of the video is front and eye level, for example, when focused on Martin explaining *Bärndütsch*. Furthermore, 11.73% of the video's viewing perspective is Angled and Eye Level, for example, when focused on Martin by "der Zytglogge" (clock tower). A 10.61% state represents front and low angle and a 7.26% state represents only low angle shots, such as when focused on Martin at the end of the video. This depicts that the majority of the video illustrates involvement in interaction with Martin, over detachment in terms of front, eye-level shots. Within the video, 13.97% is uncategorisable in terms of horizontal and vertical perspective, due to shot transitions and the black introduction and conclusion screen of the video.

Furthermore, 46.96% of the video has Martin gazing directly in a medium shot. Within the video, 12.71% has Martin gazing directly in a close shot. A 7.18% state represents when both Martin and Salome have direct gazes within a medium shot. A 3.87% state is Martin gazing indirectly in a medium shot. A 1.66% state includes an indirect gaze by Salome while Martin has a direct gaze, both in a medium shot. This illustrates a lack of formality in comparison to, for example, Video 2, where more long shots are present than medium and close shots. Within the video, 25.97% is uncategorisable, which include shot transitions or scenes where actors are not present.

Figure 25 shows information value, i.e. Participants being presented as Given information, New information or centre and consequently given primary focus.

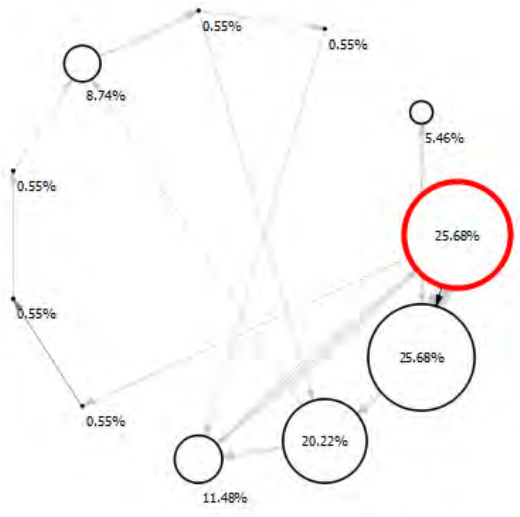

Figure 25: Video 3 Information Value

In Figure 25, a 25.68% state is Martin as Given information. The other 25.68% is uncategorisable. In 20.22% of the video Martin is just centre. The 11.48% state is Martin as New information. In 8.74% of the video, Martin is presented as Given information with Salome as New information. In 5.46% of the video Martin is presented as Given Information with the coats of arms as New information. The 0.55% state is the only time Salome appears by herself without Martin as Given information, before moving to a position as New information. This confirms Martin as the dominant speaker and main provider of information.

For Comparative Relations, 21.11% of the video consists of Martin's utterances indicating Contrast, such as when he's talking in the beginning of the video and there is no Similarity of the content of his utterances to the visuals. and 14.44% of the video indicates Similarity, such as when talking about the clock tower in front of it. For Salome, 1.67% of the video, the only time she speaks, indicates Contrast. A 62.78% of the video is uncategorisable in terms of Comparative relations. With regards to Temporal relations 80% of the video is uncategorisable. A 17.22% state of the video is categorised as Sequentiality when Martin introduces "der Zytglogge" and the bears before the camera shows them. There is also 2.78% Simultaneity with the coat of arms and Martin's speech about them present in the video. This illustrates Sequentiality is present in this video over Simultaneity, and together with the lack of Similarity over Contrast they form a general pattern in which Martin's speech does not tend to refer to what is in the visuals at that moment.

Figure 26 illustrates connotations, which were annotated based on various systems in combination, including language Martin and Salome use and their facial expressions.

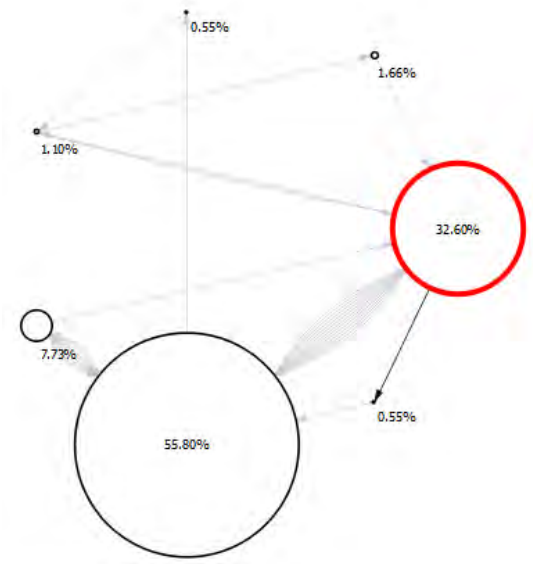

Figure 26: Video 3 Connotations

In Figure 26, 55.80% of the video show potential positive male stereotypes of Martin showing us his city and being depicted as helpful and friendly, thus keeping with the value that Germans and the value of people should be helpful and friendly. Within the video, 7.73% depicts only potential positive male stereotypes without the idea and value connotations as initially and finally pictured where he is smiling at the camera and not necessarily speaking. The 1.66% state, in which Salome speaks, also shows that Germans are helpful and friendly by her saying something in *Bärndütsch* and showing potential positive female stereotypes. Nevertheless, 1.10% depicts only possible excitatory negative female stereotypes present as she listens to Martin and he takes the lead. This emphasises her limited screen time and when she is on screen, she is stereotyped against. Within the video, 32.60% is uncategorisable for connotations.

The total duration of utterances to the second decimal point is elicited as 48.52 seconds (see Figure 27). This divided by 60 equates to 0.80, and this was then used in the formula to reach average words per minute. The number of words in Video 3 is 215. Thus 215/0.80 = 268.75 average wpm. This means that Video 3 is not as fast as Video 2 but faster than Video 1, which illustrates it is easier to comprehend than Video 2 but not as easy as Video 1.

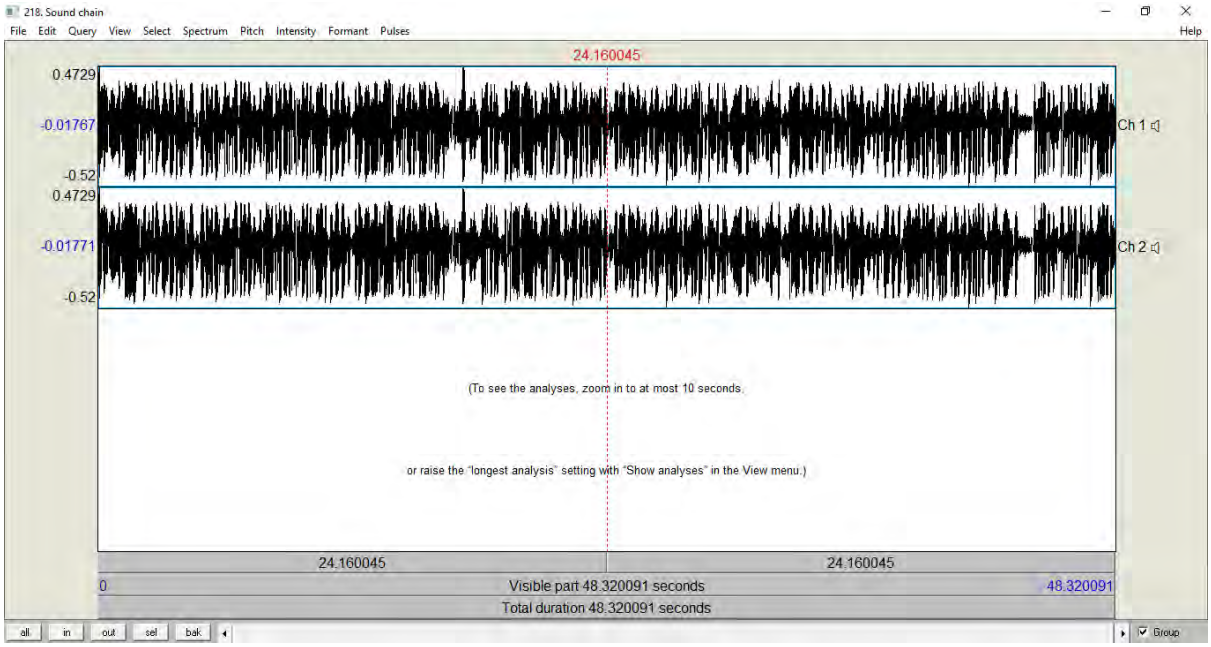

Figure 27: PRAAT sound concatenate of Video 3

5.1, 5.2. and 5.3. illustrate the findings from DMDA on all three videos and one can see that not only is Video 2 the faster in terms of speech, but also makes use of long shots more than the other videos, indicating a more formal setting than the other two videos, as there is greater distance between the Participant and the viewer. Video 1 has more Similarity than Contrast than the rest of the videos and it is also the video that is the slowest in terms of speech rate, which allows for easier comprehensibility. Each video has dominant speakers as emphasised by both the amount of language these individuals speak, as well as visual action surrounding them. An example of this includes centre or Given information positions given to individuals such as Ms Möllemann, Oliver and Martin at various times during the videos and the dominant speakers. Furthermore, the types of prevalent Moods being Declarative are all indicative of the various types of information being provided by the videos from descriptions of flats in Video 2 and a city in Video 3 to directions to a destination in Video 1. The systemic patterns established in this chapter from the various systems of the DMDA for each video is described in more detail in Chapter 6, with 6.1. discussing Video 1, 6.2. discussing Video 2, 6.3. discussing Video 3 and a general analysis that includes general video analysis results and student interview data in 6.4. The next section will discuss the findings of data collected from students in the form of questionnaires and transcriptions from focus group interviews.

Section 5.4. Questionnaires

This section discusses the findings from student questionnaires, with references to graphs made of Likert scales and a discussion of any significant written elaborations made by students. The results are organised per question of the questionnaire in this section. Bar graphs have categories on the x axis and number of participants on the y axis as can be observed, for example, in Figure 28 below, and the four columns indicate Likert scale levels with the first and last levels being described by the key below the graph. Pie charts show participant numbers as percentages in terms of categories, as can be observed in Figure 32, for example, with Likert scales being represented by the colours described in the key on the right-hand side of the pie charts.

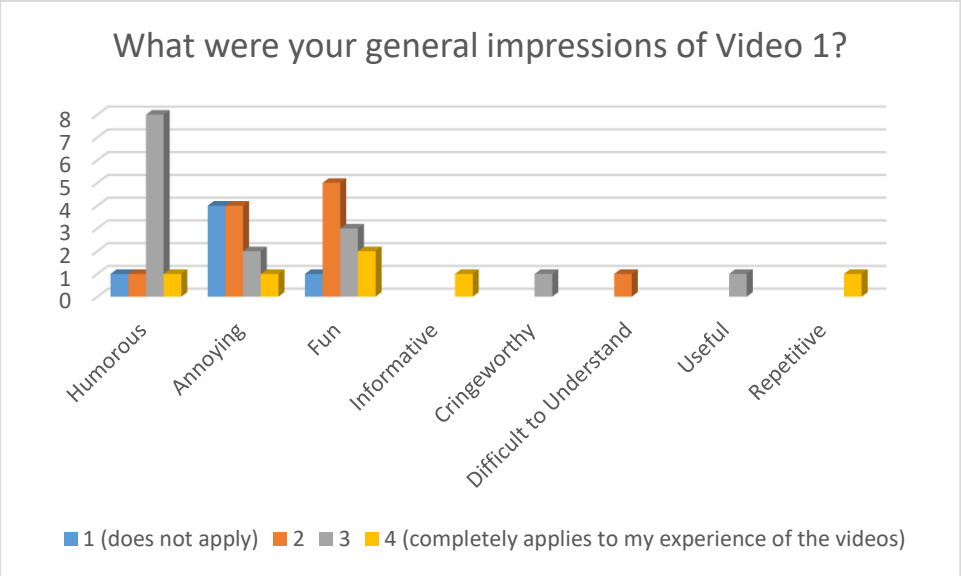

Figure 28: Responses to Question 1 - What were your general impressions of Video 1?

The general impressions of Video 1 were positive (see Figure 28), and impressions were that the video was fairly humorous and not very annoying. The categories that only have one participant are ones that were provided by individual students as further impressions of the videos. The students did not see it as providing much fun and other comments for the video included informative, cringeworthy, not so difficult to understand, useful and repetitive. This was in order to corroborate to the positive or negative nature of the cultural impressions students ascertained, as discussed in 4.1.

The general impressions of Video 2 were similar to Video 1 (see Figure 29), and students found it fairly humorous and not very annoying. Most students found it less fun than Video 1

and other comments for the video included informative, cringeworthy, (not so) difficult to understand, useful and difficult to hear.

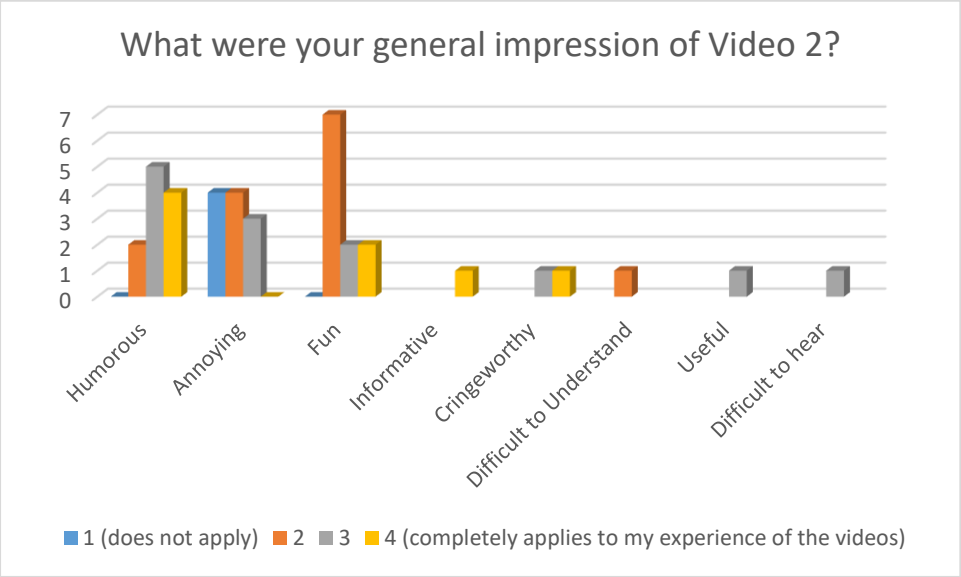

Figure 29: Responses to Question 1 - What were your general impressions of Video 2?

The general impressions of Video 3 were also positive (see Figure 30), although students found this video less humorous than Video 1 and 2. The students found the video was not very annoying and fairly fun. Other comments for the video included informative, difficult to understand, not very useful and uncomfortable.

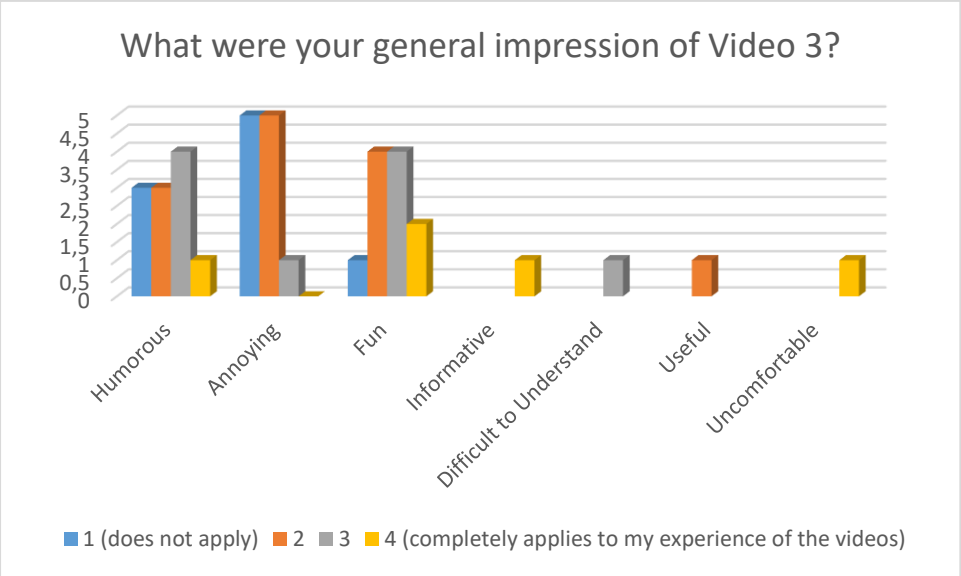

Figure 30: Responses to Question 1 - What were your general impressions of Video 3?

According to Figure 31, students found Video 1 the most helpful. Seven students identified directions as the focus of Video 1 in the written component to this question when prompted for an explanation of their responses. For example, P4 states: "Understanding directions in a foreign place is a very important skill to learn" (Appendix B). The themes from a thematic analysis of the questionnaire responses of Video 1 in Question 2 include listening practice which assists in memory and pronunciation (five participants) and the repetition of the video, while often boring to some students, assists with memory and easy comprehension (three participants). For example, P6 states: "Even though the repetition is boring it is memorable and this helps with vocabulary learning" (Appendix B). Furthermore, P1 elaborates: "This is a challenging chapter of Menschen and the directions are repeated 3 times to ensure ample listening practice and to display the correct usage of these structures/expressions" (Appendix B). Two participants found the visual aids helpful and that the videos helped with vocabulary (such as mentioned by P6).

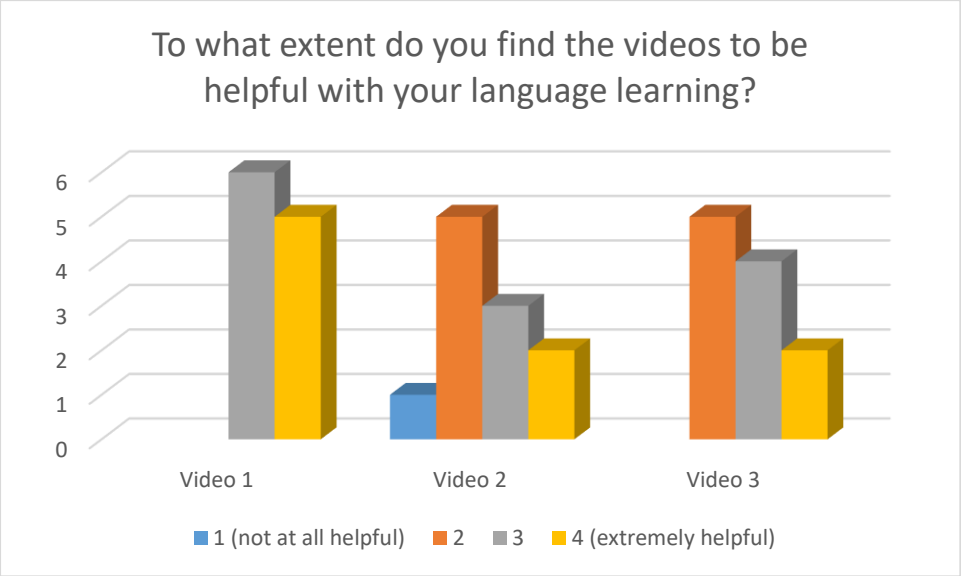

Figure 31: Responses to Question 2 - To what extent do you find the videos to be helpful with your language learning?

For Video 2, four students were distracted by visuals, with two participants referring to the dirtiness or filthiness of the flat. For example, P7 states: "Was nice to listen to however got distracted by the filthy house the woman lived in. So, ended up not focussing on what was being said" (Appendix B). They identified the video could be helpful in terms of descriptive words for a flat (three participants) and for listening practice (three). Another example is a statement made by P14: "great for helping with dictation" (Appendix B). There was much

more criticism of this video clip, for example P9 said that the language used was too simple: "Very simple language used, it could have used better or more descriptive words" (Appendix B). The visuals were found to be more important in real life than linguistic descriptions of a flat, according to students, thus illustrating that the video's task is not aligned to the requirements of a real life situation in Germany as perceived by students in South Africa. For example, P4 states: "As a foreigner it's not that important to understand descriptions of an apartment, sight is the most important" (Appendix B). Three students identified Swiss-German as a major topic in Video 3. A good example of this is found in a statement made by P11: "Exposure to a new form of the language entirely, words we've learned tied together with visuals, good pace and new words" (Appendix B). Three participants, including P11 enjoyed the video and three claimed to have learned new vocabulary through it, illustrating that perhaps the videos are not simply tools for revision, which one could assume because the clips are based on the content of a previous chapter. The videos may be designed to be used as revision tools, although this is not explicitly stated by Hueber. Two participants found Video 3 easy to comprehend and that it provided helpful visual aids, but one participant also found the videos confusing and another two thought there was no point to the video. For example, P8 stated: "I didn't really see what they were trying to teach us in that video" (Appendix B).

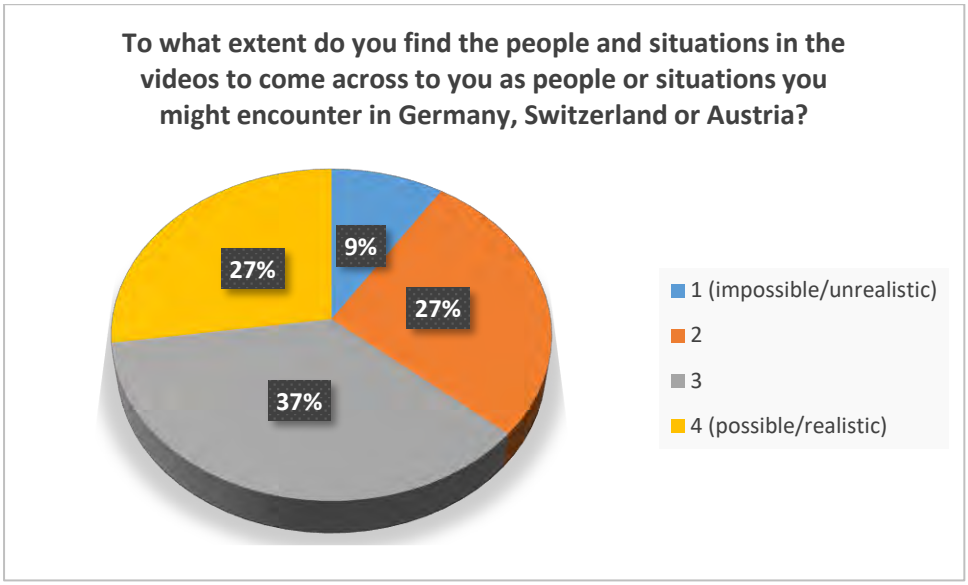

Figure 32: Responses to Question 3 - To what extent do you find the people and situations in the videos to come across to you as people or situations you might encounter in Germany, Switzerland and Austria?

As illustrated in Figure 32, 37% of participants (four) in the study thought that the portrayals of Germans were fairly realistic, with 27% (three) thinking they were very realistic and another 27% (three) thought they were not very realistic. Seven students found situations in Germany and Switzerland portrayed in the videos realistic, whereas three students said the people were unrealistic and/or there is a terrible quality of acting in the videos. For example, P3 states: "Tourists ask for directions anywhere, people can be unpredictable" (Appendix B). This supports the view that the videos are experienced by students as realistic.

For question 4a (What do you think the videos are trying to portray about German speakers in German-speaking countries?), the students found that the videos portray Germans as friendly (three students), amusing (two students), diverse (two students) and human, i.e. not displaying static behaviour (two students). One student also found the videos portray Germans as helpful, another as polite, a third as patient, and a fourth as welcoming. A student also said Germans enjoy teaching about German culture (see Appendix B for these responses).

Figure 33 illustrates that 60% of students who participated in this study, six, were in fair agreement of the portrayals shown in the videos, whereas 30% of students that participated, three, were in fair disagreement.

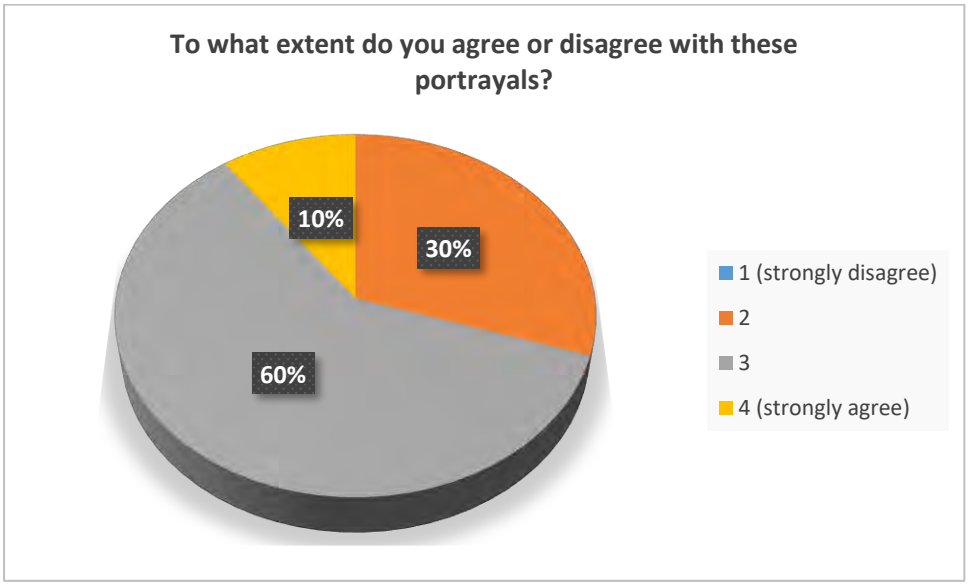

Figure 33: Responses to Question 4b - to what extent do you agree or disagree with these portrayals?

Question 5 of the questionnaire contained the question: “What did you find difficult to understand in these videos regarding: a. language in the videos?” and “b. the characters’ actions or behaviours in the videos?” For 5a, six students answered that the speech rate is too fast. For example, P11 stated: “In some parts of the videos the dialogue was very rapid, and I couldn’t quite make some of it out” (Appendix B). No comprehension difficulties were found in two responses. Three students claimed the Swiss German spoken by the character Salome in Video 3 was incomprehensible, and two claimed there was pronunciation variation between one dialect (*Bärndütsch*) and *Hochdeutsch* or Standard German. P3 claimed that “the last video was not proper German to my knowledge” (Appendix B). For 5b, six students found the actions and people to be unrealistic, with comments on the bad acting of video actors such as one made by P6, “Inexplicable, but what can you expect from failed actors, badly paid, who have to portray unrealistic scenes. Their behaviour is not the focus” (Appendix B). Four students had no difficulties, two students found videos overdramatic, especially Video 2 as P4 describes: “The characters behaviour were actually easy to understand. However, Frau emann was peculiar” (Appendix B). P10 commented that video actor reactions were unrealistic (see Appendix B) and P11 was unsure if realistic/unrealistic actions were portrayed.

According to Figure 34, 55% of participants (six) in the study would recommend the video as a study aid, whereas 36% of participants (four) in the study would not often recommend it as a study aid. In the transcription for Focus Group A, two participants said one should manage their expectations when gauging what they could receive from the videos designed at elementary level in terms of language and German culture (i.e. not expect too much from the videos). Other reasons for low ratings students listed included that they felt uncomfortable with what they saw as the cringeworthy nature of the videos and one student felt intermediate knowledge is needed, i.e. the videos were not designed for students with no prior vocabulary knowledge. Reasons for positive ratings students listed included that the videos helped with memory and German language exposure. In the transcription for Focus Group B, two participants thought that audio exercises would be more helpful than videos, as mentioned, that the videos would be a distracting study aid due to the overdramatisation of acting experienced by students and thus would not be helpful. Different learning styles were mentioned, stating the videos could accommodate those who preferred visual aids in learning. Others mentioned the videos could allow for mapping visuals to sound to activate one’s

memory and provide cultural information. Both groups thus found that the videos had positive aspects for memory in terms of language learning. While there was a variation of remarks on the videos as a study aid, the students' experience of the overdramatised acting within the video caused some students to feel uncomfortable and distracted, which resulted in them not finding the videos an effective study tool.

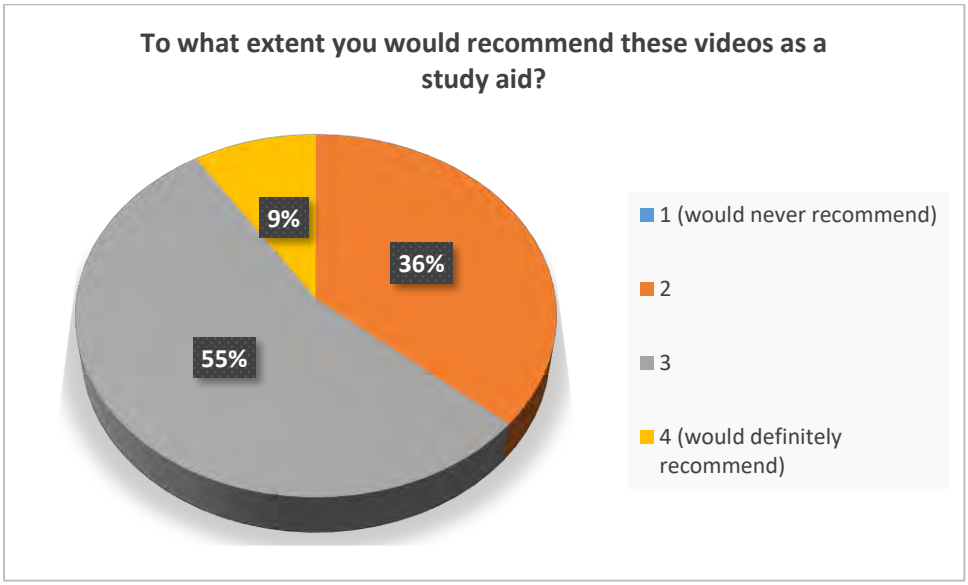

Figure 34: Responses to Question 6 - To what extent would you recommend these videos as a study aid?

As illustrated by Figure 35, 46% of student participants (five), think the videos are not being used quite as effectively as possible in the German Studies 1 course, whereas 36% of them think they are used fairly effectively (four participants). In the transcription for Group A, two students suggested better task design set to a longer video and more opportunities needed for discussions in German. This goes beyond recommendations for the use of these videos; the students also recommend that the videos be modified. The tasks, like the videos, are designed by Hueber and are not the responsibility of the German Studies Section at Rhodes University, although teachers/ lecturers/ facilitators could design their own tasks for use with the videos. Two students also said one video should have three tasks, rather than three clips with a task each. A single longer video with tasks set to it, students felt, would be more beneficial than three shorter clips on different topics with less time spent on each. A participant also said the task for Video 2 should allow for more descriptions of the flat. In the transcription for Group B, students stated that the videos should be used when first learning the section once the basic vocabulary has been acquired and teaching the vocabulary alongside the video is helpful.

They further asked for repeated viewings, as it is better to do the task after the viewing or during a repeated viewing.

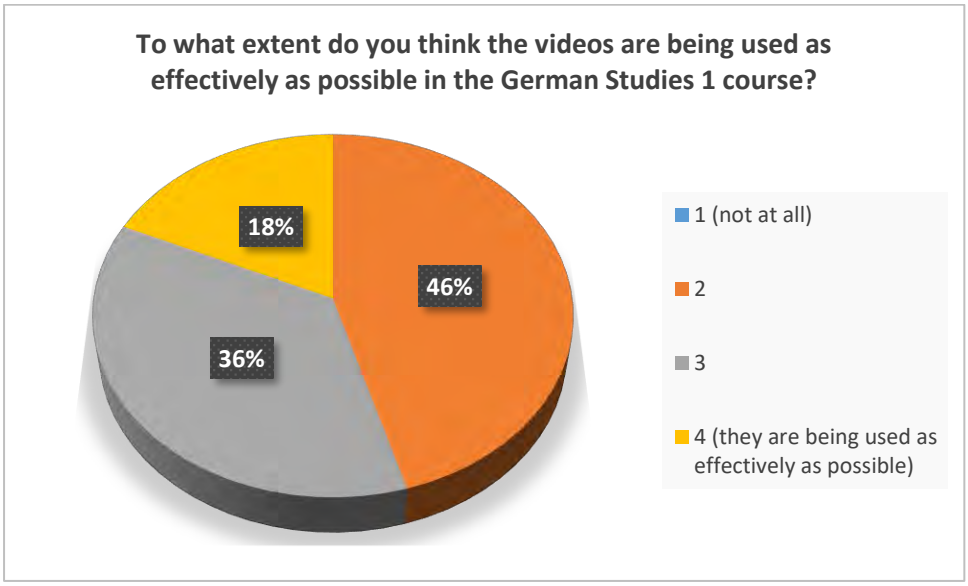

Figure 35: Responses to Question 7 - to what extent do you think the videos are being used as effectively as possible in the German studies course?

Section 5.5. Responses to Interview Questions in Focus Groups

In this section, student responses are reported on in a question-by-question format in a similar manner to 5.4. This is to present trends of opportunities and challenges that students found in the videos in an easier manner, which can later be compared to specific videos in sections 6.1., 6.2., 6.3, as well as within a general analysis in 6.4. Appendix D and E are referenced after each student participant answer throughout this section in order to ensure responses can be easily examined in the context of their relevant appendix.

1. *When watching the videos and filling in the questionnaires, did any of you notice anything interesting about them that you want to discuss?*

In Group A, five people discussed how the visual aids such as the arrows in Video 1 were helpful in the videos, for example P3 (see Appendix D). Three participants mentioned they were helpful in Video 1, P4 mentioned they were helpful in Video 2 (seeing the flat): “Like the the apartment scene, like the showing of like gesturing to the dishes or the view or just the visual like aids as in like the everyday items” and P7 mentioned they were helpful in Video 3 (seeing the clock tower, which was content in the discussion): “Thank you for showing me

the clock ‘cause I well you know. I wouldn’t have known what you were talking about” (see Appendix D). Two participants also mentioned the repetition was useful for learning, for example, P1: “Well, for example, we’re not watching videos related to any like everything we’ve done, they always have a group of target language, for example, directions and we know okay in this video they’re gonna be talking about directions. We have a fixed number of words and expressions we can learn to help us. And then we get like a great demonstration, especially in that video like three times on how to use them effectively” (Appendix D).

Two further participants mentioned the speech rate was too fast in Video 2, for example P7 states: “I don’t find um its specific videos depending ‘cause the first and second one were fine, but where they were in conversation, I was like what is this lady saying, ‘cause you speaking hella fast. ‘Cause I’m still trying to translate what you saying into English to actually understand what you’re saying so ja I agree with that” (Appendix D). Another two participants mentioned that either the Swiss German accent or dialect was unclear. For example, P6 states: “Although I did find when they do other accents, um, that’s a little over my level. Like the Swiss accent. Is it really necessary right now, maybe later. Try to figure that out, but right now can I just try and understand high the high German” (Appendix D).

Three participants thought Video 1, and the directions it conveyed, assisted with pronunciation, and another three participants thought it assisted with memory:

> P1: Well essentially these are listening activities so like the primary well it depends on the task given as well I mean we could be listening for just or as you know for like specific details so it it depends on that first of all, so task-related, and ja I don’t know, listening is essentially to help you with your pronunciation and your comprehension and...
>
> P6: Memory and repeating it three times and with the visual aid. Especially if you’re a visual learner it’s definitely gonna help determine with the vocabulary. (Appendix D)

Two participants stated the overdramatisation functioned as a learning aid through the technique of using repetition and that this was memorable. For example, P4 states: “I think maybe the like extra dramatisation is to like help to aid us in learning ‘cause we don’t ‘cause we can’t function in like a normal day to day conversation we don’t know what’s going on so the gestures are like to aid us” (Appendix D)

Two students mentioned the videos provided speaking opportunities in class, and two students thought the videos assisted with various learning styles such as visual learners in learning. For example, P3 states: “It shows uh, more of a conversational setting. So we’re not

just learning sentences at a time. Helps to show how you can like string those sentences together and hold a conversation with people" (Appendix D) and P7 states: "Ja I'm a visual learner so when I see like a picture I'm like kay cool and hearing it also helps with pronouncing and ...yeah definitely" (Appendix D). In Group B, three further participants mentioned the speech rate was too fast in Video 2, for example in a statement made by P11:

> "So like the first one was like really slow and like repeated a lot of stuff and then the second one was like, it had like really slow parts and then it had like that one really fast part where no one could actually hear what she was saying and then the third one was quite like was slow enough but it was a lot faster than the other ones" (Appendix E).

2. *To what extent would you say the interactions and the depictions of the German-speaking countries in these videos match your idea of what Germany, Austria and Switzerland are like?*

From Group A, three participants agreed that asking for directions was realistic in terms of their experiences, but overall participants came to the conclusion that not saying thank you/ impoliteness is unrealistic, that escorting someone to their destination is unrealistic and that the actions and people were at times overdramatic. P3 states here:

> Again with the overdramatisation. I don't think that people just walk around discussing everything around them, 'cause if you go to visit a country like that, they're not gonna randomly take you aside and be like hey look at this and this and this. You seeee? This is this how many years old and this is the explanation of that. They'll just be like hey get a tour guide (Appendix D).

P1 attempted to translate the situations from his "mother country" (South Africa), clearly referring to his own culture, stating all situations are realistic based on his comparison of his own experiences:

> I don't know I just consider each of these situations and how I'm bound to use them in my mother-tongue in my mother country every now and then when I'm looking for an apartment or moving or if I need directions, which I very often do, or if I'm planning to go to visit some place I'm reading up about that place and there's bound to be some information about you know why that place is great or not great. So ja all of these are definitely used in my English so I don't see how it wouldn't be relevant in Germany (Appendix D).

P3 also said the portrayals are that Germans are friendly in all three videos and P4 that Germans are helpful in the videos, providing an example of Oliver escorting the tourist, and that this was realistic (Appendix D).

In Group B, three students mentioned the videos made them uncomfortable due to the overdramatisation, others mentioned distracting and cringeworthy as feelings towards the

unrealistic nature of some of the videos, for example P8 states: “So like in the videos they were always trying to like crack jokes and like smile into the camera and laugh and that was uncomfortable” (Appendix E).

Six participants mentioned a previous *Menschen A1* video clip they had seen, in which the salesperson and the client burst into song as another example of an unrealistic situation they had experienced with the *Menschen A1* videos. Nevertheless, they felt the nature of the video helped them remember it and the phrase “Ich brauche eine Rechnung” (I need a receipt), which is a line in the song in the video (Appendix E). In addition, they felt that the portrayals of Germans and German culture in the video included that Germans were friendly, funny and happy people, yet P9 stated that this portrayal is not accurate as Germans are just human (their behaviour is not static) and they are like everyone else:

> P9: Well I’ve been to Germany and I’m not saying the society is the same as here but people are pretty much the same. Like there isn’t any like vast difference like they’re not completely like in the zone like everyone but they’re not like crazily out of their way friendly they’re just like people
> I: Mhm So okay so they’re just like people, they don’t go out of their way or?
> P9: No not all of them are like. So the way like for me the videos seem to portray the Germans they go out of their way to be like oh we’re friendly and like you know smiley and super happy and I’m sure some people in Germany are but like some of them not all of them. I’m sure there are some who are like people go (Appendix E)

This could mean that humans tend to have a similar nature around the world regardless of cultural differences or that she covertly orientates to a norm in which cultures from the Global North are considered normal.

3. *From your personal experience, what are the advantages of language learning from these videos, especially in comparison to learning from books?*

Due to time limits, this question was only asked in Group B, as Group A tended to provide more elaborate answers to questions, which resulted in less questions being asked in this group than in Group B. In Group B, two participants mentioned the videos provide the opportunity to hear words actually spoken and how they are pronounced, for example P9 states: “Well, I mean obviously you get to hear like the words actually being spoken, when you’re reading a book you only have like you know your head voice” (Appendix E). P7 mentioned this is effective for remembering the phrases in the tasks (Appendix E).

4. *What do you think are the disadvantages of language learning from these videos from your personal experiences, especially in comparison to learning from books?*

Four participants expressed their concerns that they are unsure if the situations and actions of the people are portrayed realistically or not as many of the participants had not been to Germany before. For example, P12 states: "Because now okay you're in German class, you're watching those videos then you happen to go to Germany and then how will you not expect that to happen? Like you will expect that to happen. But maybe you won't find it like that, so that makes me sceptical" (Appendix E).

Two participants mentioned the videos were overdramatic and this could be misleading as to how Germans are in real life. They made it clear, however, that the repetition and overdramatisation, as well as visual aids, were for their comprehension, and they thought asking for directions and viewing an apartment for rent were realistic situations that could happen (see Appendix E).

5. *What do you think these videos are useful for when learning for German Studies?*

Both groups were asked this question, but discussions were short as this had already been discussed in questionnaires. Students mentioned useful features of learning with video including mapping visuals to sound (P11 in Appendix E), and that the videos were good for memory and language used in context, particularly due to the overdramatisation, as noted from a discussion with Group A from Appendix D:

P4: It like provokes, like conversation and a debate because it's memorable and overdramatised, it like it intrigues you. So you almost want to learn what's going on. Like a soapie, it's like a soapie
P1: A bad idols audition
P7: It's a terrible soapie

The videos were said to illustrate a diversity of accents and dialects, as well as potentially cultures. For example, P12 states:

> They also help in understanding that um Germans like they're different like the Swiss and the German so they're different not like the same, so you'll be able to maybe if you understand the language you'll be able to distinguish between the Swiss and the German (Appendix E)

6. *Regarding language in the videos, what did you find difficult to understand?*

Three participants mentioned the speech rate in Video 2 in particular was too fast, although a participant mentioned that Video 3 was faster than Video 1, but not as fast as Video 2. For example, P12:

> P12: Well since P9 said in some others they were fast so I was not able to pick up like the the words and the other phrases to translate quickly as the video was going like you're stuck on one word that was spoken and then you're busy trying to translate it and then when you come back to the video it's just way over that it's just its moving on to others things. So ja.
> I: So you found it very fast. Was it just Video 2 that was the fastest or?
> P12: Yeah, it was Video 2 (Appendix E).

This confirms my speech rate calculations in that the speech rate in Video 2 is faster than the other two videos and this is discussed further in 6.2. They felt that they miss out when doing the task simultaneously with the videos during a tutorial as they're only focusing on completing the task and finding information on it rather than viewing the video completely without interruption. For example, P11 states:

> Yeah, I agree with that thing about like when there's a question about it you only focus on finding the answer and miss out everything else and like sometimes you even miss out on the answer because I don't know, it's just again the speed is weird because sometimes with the answer they speed through that and then take their time with random information just feel like its weirdly balanced like having a written task while you're watching the video is very jarring whereas if you had to watch it and then someone be like okay what did you notice about such and such, you could do it like orally and then write it down like afterwards (Appendix E).

P8 also found the background noises distracting (Appendix E).

7. *End remarks - Opinions of the flat and writing in Video 2*

From Group A, two students made the remark that the flat in Video 2 was really tiny and one student, who had been to Germany, commented that the flat is huge if it was situated in Berlin. Writing visible in Video 2 in the form of a written advertisement held up to the camera was not memorable until it was mentioned and all participants claim they only saw "Superwohnung" (see Appendix D). In Group B, four participants discussed the physical representation of the flat, identifying that as a focus of the video. One person said it was small, another said it was cheap, but was then questioned by another participant as to how they know it is cheap; it could be expensive. Writing was not memorable until I mentioned it and all participants claim they only saw "Superwohnung" (see Appendix E). This links to insights gained from a cognitive theory of multimedia learning, as shown in 6.2.

Section 5.6. Conclusion

In this chapter, students did not find the videos too challenging for task completion, and opportunities for both language learning, such as for pronunciation and memory, and cultural representations of friendliness of Germans were the majority identified. These are in line with findings also reported from a DMDA of all three videos. According to many student participants, all three videos contained bad acting from actors that made students cringe and uncomfortable, but this did not hamper learning or task completion too greatly. Of the three videos, Video 2 was found to be the most challenging with the fastest speech rate of the three videos and visuals of the flat that were distracting and this is in line with DMDA findings in 5.2. The presence of Swiss German in Video 3 was also found as a distraction to students, with many not understanding the point of the video. In line with the highest rating of Similarity and Simultaneity between language and visuals in Video 1 found in the DMDA, students found the visual aids and repetition in Video 1 the most helpful in terms of language learning. All of this, as well as further findings mentioned in this chapter are analysed and compared with one another in greater detail in Chapter 6.

Chapter 6 - Analysis

This chapter analyses the findings from Chapter 5 in deeper detail and is organised in a similar fashion to Chapter 5. 6.1., 6.2. and 6.3. examine Videos 1, 2 and 3 specifically. The data from both DMDA and student interviews is triangulated with literature and theory analysed from Chapters 2 and 3, and this chapter thus differs from the more separated reports in Chapter 5. Similar triangulation occurs with general data that came from student interviews in 6.4.

Section 6.1. Video 1

The states referred to in this section were determined in 5.1. and the student data from 5.3., 5.4 and 5.5. Material Processes in the linguistic mode and Action and Interaction Processes in the visual mode coincide with one another, illustrating that there is a similarity in Process between modes, as I observed Action Processes in particular as analogous to Material Processes in the linguistic mode. This is a syndrome as defined in 3.3. of Chapter 3 and it illustrates a relationship between various modes in this video, thus providing an answer to research question 2. While Oliver has the majority of utterances with Material Processes (two 9.93% states), the tourist is not only the most frequent Participant and Theme in most of the utterances, but she is also the one with the most visual Action Processes (7.10%). Both Participants have turns as Actors and Reactors and this equalises both Participants as Interactors within the visual Interaction Process. Furthermore, this Similarity between Material Processes in the linguistic mode and Action Processes in the visual mode is further demonstrated by the high amount of Similarity (36.01% in total, including a 0.68% state) in the video in terms of Comparative relations, in comparison to the lower amount of Contrast in the video (17% in total). This epitomises the mark of a successful video for pedagogic purposes when meanings within modes in the video match one another (see principles of spatial and temporal contiguity discussed in terms of the cognitive theory of multimedia learning in Chapter 3).

The focus of the video, however, is not just on the Action Processes and Material Processes, but also on the Circumstances and Conceptual Processes that accompany them, as the Process and Participant “Sie gehen” (you go) is elided in the majority of clauses. Within the 9.93% states, the majority Circumstances are identified as being Extent and Location, which describe the directions that Oliver gives the tourist, and which she then repeats and Oliver

also repeats. This is further emphasised as information regarding directions which is evident by the large number of Declaratives (28.28% for Oliver and 10.34% for the tourist). Approximately 6% of the video shows that during an interaction, the tourist depicts gestures that count as Conceptual Processes with a Concept Participant and 4.52% of the video refers to Conceptual Processes and Concept Participants made by Oliver. Exactly 10.96% of the video in my analysis is further categorised as Conceptual Processes, with the Concept Participants being the large, yellow arrows that appear foregrounded in colour contrast and centred alongside Oliver's speech and gestures and both Oliver and the tourist's actions within the video. Despite the arrows appearing sequentially and the gestures and Circumstances appearing simultaneously, one can see this as a pattern of three-way couplings (states) and thus representing a syndrome.

This repetition and Simultaneity are also observed within Temporal relations, with 16.44% of the video realizing Simultaneity with Oliver and his gestures and 10.27% Simultaneity between Oliver's speech, gestures or walking and the arrows. This does not include the sequential parts of the video in which Oliver and the tourist also walk through the directions (represented by the 12.33% dolly shots), or in which the tourist repeated the directions, which is a further repetition of Sequentiality. The linguistic task for this video calls for the ordering of directions to the Goetheplatz (Evans et al. 2012:88) and this clarifies that the directions that take so much of the video's speech and that occur with such repetition and Simultaneity are the focus of the video.

Within the video, 61.64% consists of stationary shots and should illustrate minimal involvement of the audience in the action of the video, as only 12.33% of the video contains dolly shots and 10.27% of the video consists of dolly shots with arrows zooming in and out in a Simultaneity between language and visuals. However, 57.82% of the video is angled and eye level and the audience is thus an accepted participant in the video for the majority of its duration and through DMDA students may be made to feel an equal part of this conversation. Information is provided to them exactly as it is to the tourist. Furthermore, Oliver acknowledges the audience at the end of the clip when he speaks to them, commenting on the tourist's rude behaviour by stating "Sehr nett. Vielen Dank. Danke." (very nice. Thank you very much. Thank you) sarcastically, and 21.77% of the video depicts the frequency of front and eye level shots in the video. This is further emphasised by the majority of the video

(39.46% and 19.05% states) consisting of medium to close shots and the presence of indirect gaze which typically alludes to information being provided, rather than demanded.

Oliver is the dominant speaker and presented visually as New information for most of the video, while the tourist is presented mostly as Given information. This illustrates that he is providing the answers of directions to the tourist's question of where the Goetheplatz is located. As Oliver is the dominant speaker assisting the tourist with directions, it is no surprise then that 12.21% of the video exhibits positive male connotations and that through his actions, Oliver is portraying Germans as helpful, friendly and patient, with the value that people should be helpful, friendly and patient. Around 10% of the video connects these connotations with Material Processes and Circumstances of Location and Extent in the Declarative Mood (a syndrome). The tourist is the Theme for these utterances. While these are the value and gender connotations portrayed, also present, though less frequently, is the idea connotation that Germans are unintelligent in the way that the tourist is lost and keeps getting confused. The idea connotation that Germans keep getting lost because they are adventurous and positive female connotations do not even appear in terms of significant numeric states, and thus cannot be counted as a significant connotation grouping in this clip.

From the analysis, two points to be triangulated with the student data can be established:

1. There is a high amount of Similarity between visuals and language in Video 1 and this repetition occurs with a focus on directions as information provided to the audience that the Theme the tourist should take. This is further confirmed by the focus on the ordering of directions in the linguistic task for students.
2. There exist more positive connotations of men than women in Video 1 with Oliver illustrating how Germans can be friendly, helpful and patient.

Discussing student impressions of Video 1, more students found Video 1 humorous than the other two videos. When the video was played to the students, the students' laughed when the tourist left Oliver without thanking him at the end of the video, a situation which Group A thought was unrealistic. Laughter could be a reaction to something 'funny', as stated by students themselves of the video, or a reaction of embarrassment, shock and discomfort (Baumeister 2001:213). This is confirmed by theories of humour, including superiority or relief theories, in which laughter is often a physical response from humour that establishes superiority or acts as a defence mechanism against painful situations (Holoch 2012:20).

Furthermore, ‘cringeworthy’ was a term described by P3 for Video 1, and this was later linked to the overdramatisation of the acting in the interviews. Thus the humorous impression in the video could have been formed by the impoliteness displayed by the tourist, as well as other overdramatisations present within the video itself. Nevertheless, despite viewing the impoliteness portrayed by the tourist, the students only formulated positive impressions of Germans from the videos in general, with both groups stating that Germans are friendly, and other words were used including ‘helpful,’ ‘funny’ and ‘happy.’ Only one student mentioned ‘patient’ in reference to this particular video. All the impressions of German-speaking cultures by students were based on behaviour of actors, for example, with Oliver helping the tourist by repeating directions and taking her to Goetheplatz, his behaviour is recognised as ‘helpful’ and many student participants felt this is what the video was trying to say about Germans in general. The video points to polarised friendly and helpful or impolite covert cultural values and the little “c” of culture. Thus, according to PCSTIF, the following schema (Figure 36) can be created for impression formation for Video 1:

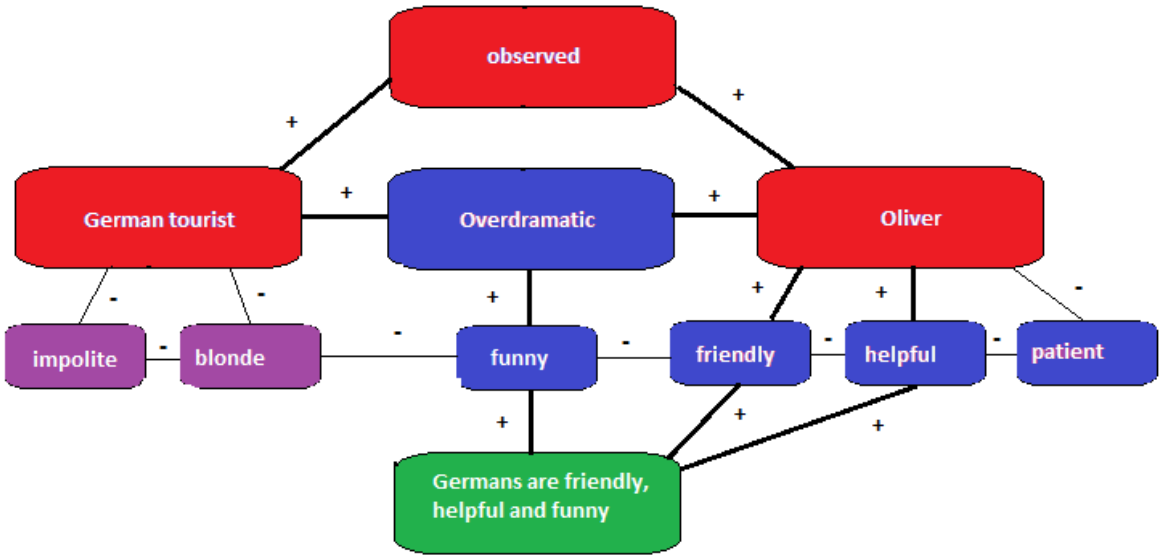

Figure 36: PCSTIF schema of Video 1

In Figure 36, the red indicates what was observed, the purple indicates my impressions, the blue indicates impressions by students or by students and me, and the green indicates the ultimate student impression based upon a collection of the stimuli and their associations made from observed behaviour above the node. I use the term “overdramatic” in this schema to refer to all the instances where students referred to the over exaggerated nature of the actors’ portrayals. One can see in Figure 36 that far more excitatory links are present for Oliver than

for the tourist and thus only the positive connotations for Germans as friendly and helpful were acknowledged by student participants. Despite the impoliteness of the tourist being acknowledged, most likely due to the fewer excitatory links present above and her not being the dominant speaker, it was not the first impression that was made or expressed consciously by students and only by me who had viewed it countless times in a DMDA, thus I had more opportunities to formulate more impressions. The fact that Group A stated that the escorting to the destination and the tourist's impoliteness both were unrealistic actions also could have resulted in that association to be inhibitory, i.e. it was not a predominant impression for students.

Thus the students correctly identified the focus of the video to be on directions, which the analysis of the video confirms, based upon the large amount of linguistic Processes and the presence of the salient arrows and Action Processes (a syndrome). Three participants found the visual aids of the gestures, the escorting to the destination and the arrows helpful to map onto the language to memorise. This and the repetition of the simultaneous language and visuals that the analysis of the video also uncovered was further identified as helpful by students for memory. While a couple of students mentioned in the interviews that the repetition could be boring, the majority found this a helpful tool for language learning. They also found the presence of spoken language and its repetition assisted in the learning of the correct pronunciation. This corresponds to Mayer's (2008:765) principles of spatial and temporal contiguity: when the corresponding visuals and words are presented closer to one another or simultaneously, better learning is achieved.

Overall, the high Similarity and Sequentiality ratings contribute to language learning in terms of the spatial and temporal contiguity principles, and the repetition and visual aids assist student memory and pronunciation. The Themes, information value and dominant speaker causing Oliver to be the focus all contribute to providing information on directions, which corresponds well to the ordering of said directions in the linguistic task that students had no trouble completing. In terms of cultural representations, men were seen in a more positive light than women and while I in DMDA identified connotations of German (blonde women) being unintelligent, it was not expressed as salient by students, rather Oliver as a friendly, helpful and patient representative of Germans permeated into overall positive connotations of culture(s) of German-speaking countries.

Section 6.2. Video 2

The states referred to in this section were established in 5.2. and the student data from 5.3, 5.4 and 5.5. Despite 39.44% of the clip containing visual Action Processes, the dominant utterances contain identifying Relational Processes with Identifier and Identified Participants. This shows that the linguistic focus is indeed on describing the flat, for example, a Relational Process clause includes the utterance "Das hier ist der Flur" (this here is the hallway). There are more that mention the kitchen, living room, bedroom and bathroom. The second highest number of clauses contains an attributive Relational Process with Carrier and Attribute Participants and Circumstances of Manner (2.81%). This illustrates that descriptions of Identifiers and Carriers are also prevalent in the video, further emphasised by the predominant Declarative Mood of utterances (16.67%), the most spoken theme of "das Wohnzimmer" (2.78%) and the relational clustering of "der Flur".

The visual Action Processes are of Ms Möllemann speaking and moving from room to room to visually depict the rooms she discusses. While the percentages of actual linguistic content are not high in contrast to the visual Action Processes, the linguistic content follows a pattern that is conducive to the completion of the linguistic task. The task that accompanies this video is a comprehension task that mainly seeks adjectival answers from participants, for example, 1a asks "Wie ist der Flur?" (what is the hallway like?) and the answers are provided as "nicht sehr groß, praktisch" (not very big, practical). It is an exercise relating to a chapter in the textbook which teaches the vocabulary relating to the interior of flats (see 5.2.).

"Kommen" and "sehen" (come and see) (2.22%) are the second most spoken Themes and this coincides with the Imperatives which is the second most frequent Mood (5.5%) (see 5.2.). The Imperatives "kommen Sie" and "sehen Sie" always occur during moments when Mr Waurich pauses to inspect messy areas, and Ms Möllemann uses these Imperatives to rush Mr Waurich through the flat and cause him to stop focusing on these areas. However, this is unsuccessful as shown by the amount of zooming and panning in the video. Her embarrassment is further made clear by frequent nervous chuckles between moments in which she explains that she has not yet made the bed or washed the dishes, as well as "ähm" and "mhm" utterances that usually precede such explanations. Ms lemann appears as Theme more than Mr Waurich, highlighting the emphasis to be more on her than Mr Waurich, even though she states his name repeatedly to shift his attention from the messy

areas, thus also attempting to shift the focus away from herself as the messiness is due to her actions.

The meanings conveyed by the visual mode relay Ms emann's embarrassment and emphasise her messy flat. The high amount of dolly or handheld movements (25.41%, see 5.2.) shows how the viewers are viewing the entire flat from the perspective of Mr Waurich, and also depict moments where he views the various rooms described by Ms Möllemann. The messiness of the flat is emphasised by the preponderance of pan (20.44%) and tilt (11.60%) movements, as well as parts of the stationary shots (12.71%) and the remaining combinations of zooming in/out and camera movements (16.56%). This stands in contrast and irony to the title being a macroTheme of the clip, "die Superwohnung" (the fantastic flat).

The high number of high angle shots (21.86% see 5.2.) emphasises Mr Waurich's disdain towards Ms emann's mess, and thus by extension towards Ms emann herself. The second most frequent shot is eye-level, and Ms Möllemann is depicted more from the front than at an angle, implying involvement over detachment (Kress and van Leeuwen 2006:134), which is further confirmed with the high amount of direct gaze (18.23% and 3.87%) within the video from her. The audience is given the view from Mr Waurich's perspective, so they are involved in the entire scene. There are rare moments when low angle is used, and that is either just to examine items piled up higher than eye-level or when he looked up at the apartment block when he arrived at the flat. The formal relationship between both Ms Möllemann and Mr Waurich is emphasised by the fact she only appears in long shots and is kept socially distant from Mr Waurich. This further implies a level of formality in German cultural norms regarding interpersonal interaction, but this was not identified by the students. This also shows that Mr Waurich does not feel affinity towards her, and this, in combination with the frequent high angle shots, portrays Ms Möllemann in a negative light with regards to her untidiness.

The emphasis on Ms Möllemann as important information is not only emphasised by topical Theme, but also by information value, as she is found more in a centre position than as Given information (8.20% in comparison to 6.01%). She is also focused on within framing clusters (syndromes) that appear, often when she uses Imperatives "kommen Sie" and "sehen Sie" (come and see). An example of a framing cluster (see 5.2.) within the annotations is pictured below (Figure 37).

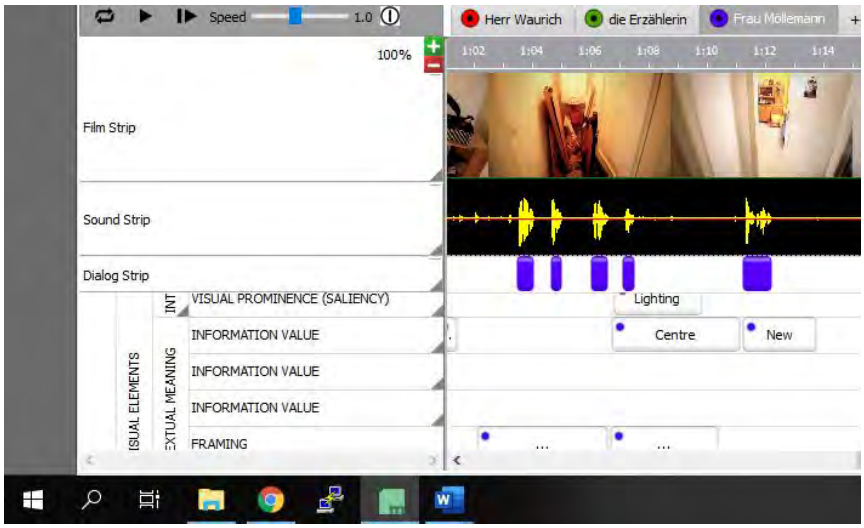

Figure 37: An example of a framing cluster in annotation where the annotation with ellipsis is "framing due to positionality"

The data establishes that these clusters occur frequently, for example in the states 5.46% lighting, centre and frame lines, 2.73% colour contrast and lighting, Given and frame lines, 1.64% colour contrast and lighting, centre, frame lines, as well as 0.55% lighting, Given and frame lines. There are also times when framing and centre co-occur without the lighting (4.92%). These framing clusters/framing syndromes occur when Ms Möllemann appears in doorways. An example shot of this is Figure 38, when she initially greets Mr Waurich: she appears centre, framed by the door and the door frame and the lighting behind her illuminates her more than the foreground. Framing clusters such as these are used to further place a focus on Ms Möllemann, and when paired with her Imperatives, it places emphasis on her shifting focus at the messiness of her flat.

Figure 38: Example shot of a framing cluster

Due to the framing clusters, Imperatives, high frequency of Ms Möllemann-related Themes, and the camera movement's focus on the messiness of the flat, connotations of German women as disorganised are found in the video. As illustrated in 5.2., 27.07% of the video portrays this with the value connotation that people should be neat and tidy and should be embarrassed and ashamed if this is not the case. The zooming in and panning of the mess with high percentages emphasises this, as well as constant Imperatives to rush Mr Waurich away from the mess and focus on Ms Möllemann thematically that are mentioned earlier in this chapter.

Despite the fact that patterns of language conform to the task and context of it within the chapter of the textbook, only 6.11% of the language is similar in content to visual and displays Simultaneity, with a 23.89% Contrast and the rest uncategorisable. This could be a reason why students found this video of the three the most challenging to follow, as most of the language content was dissimilar and did not correlate with the visuals that they had to process. The descriptions of the room given by Ms Möllemann do not always correlate with the visuals on the screen. Furthermore, while Relational Processes were the majority of Processes found in the clauses of the video, Action Processes were the majority of Processes found in visuals, rather than State or Conceptual Processes, to allow for co-contextualisation to occur. Thus, there is too frequent re-contextualisation and opposition between visuals and language that may present a challenge in cognitive processing for students. Due to the high frequency of relational Processes and names of rooms used as Themes, the adjectival content of the videos is provided in terms of the linguistic mode on its own. There are not, however, frequent instances of language that is categorisable in such a clip, and it is dominated by visuals that contrast in opposing Processes. Furthermore, the clip carries possible negative connotations towards German women and that being disorganised and messy is something for which one should be embarrassed.

Despite the connotations found in analysis, these were not discussed by the students, and therefore, whether or not they were identified by them remains unknown. P6 in Focus Group A claimed that identifying portrayals of German culture were not the point of the videos' creation. Thus, one could argue according to PCSTIF, that these potential connotations were inhibited amongst students. As the PCSTIF Figure 39 depicts, this could be due to Mr Waurich's dislike of mess becoming apparent on his sudden exit of the flat, thus Mr Waurich could be associated with organisation and cleanliness, which then nullifies Ms lemann's

negative associations and perhaps prevents negative associations from becoming negative impressions of Germans in general. If students thought this video further enforced connotations of friendliness and helpfulness amongst Germans, as they felt in terms of Video 1, they did not specify this.

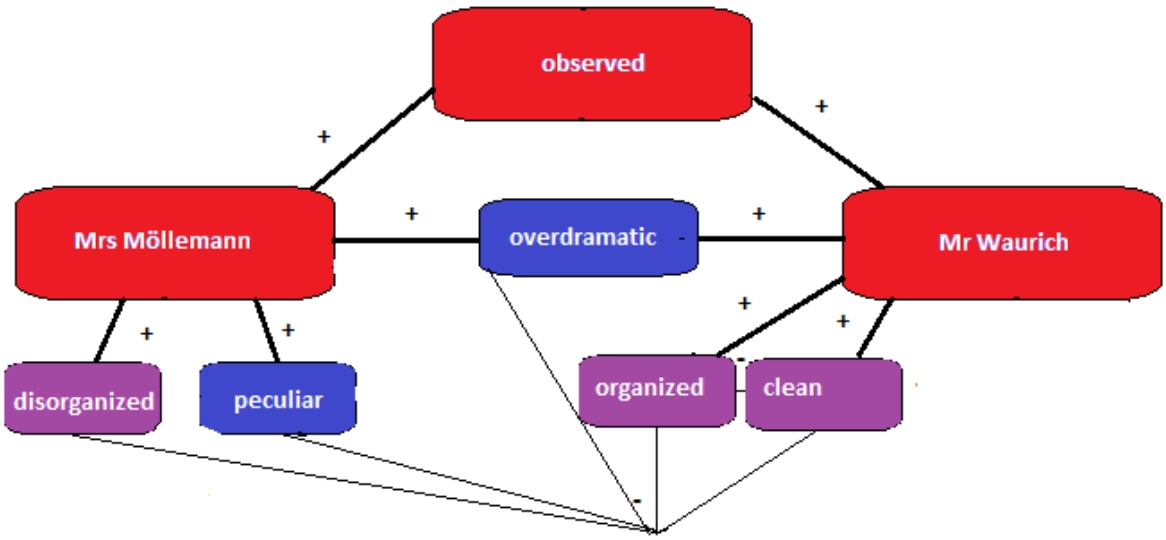

Figure 39: PCSTIF schema of Video 2

In Figure 39, the red indicates what was observed, the purple indicates my impressions and the blue indicates impressions by students or by students and me. I use the term "overdramatic" in this schema to refer to all the instances where students referred to the exaggerated nature of the actors' portrayals. While Figure 39 clearly highlights negative associations identified by students, no impressions were formulated, which differs from Figure 36, the PCSTIF schema for Video 1, in which more positive associations resulted in an overall positive impression of Germans. While "friendly" and "helpful" were general attributes observed by students, P7 elaborates on the helpfulness, explaining: "I found it like it seemed as though if you are a foreigner anyone would be willing to help you figure out what's happening and stuff like that I was like okay cool I'd go to Germany people would help me. Ja." (Appendix D). Thus based on the explanation that the helpful nature was attributed to the actions around a foreign presence, such as the tourist or the viewer in Video 3, I found friendliness and helpfulness to be attributes with stronger visual and linguistic

evidence within Videos 1 and 3 than Video 2 and this also explains the lack of these associations mentioned in Figure 39.

From the above analysis, two points can be corroborated with the student data, namely, the visuals illustrating the messiness of the flat, with students identifying the messiness of the flat (see 5.5.) rather than attributing the disorganisation to Ms Möllemann, as well as identification of the physical and descriptive representation of the flat as the focus of the video. Not only were adjectives provided to students as vocabulary, for example “groß” (big), but students identified this video as assisting with listening practice. Nevertheless, students did not see the point of the video in terms of it resembling real life situations, stating that one would not need adjectives as vocabulary when going to view a flat that they were interested in renting. Overall, students had neutral and negative impressions of this video, with Video 2 being rated as less fun than of the other two video clips.

Many language learning challenges were stated for this video including that the language used was too simple, and four students mentioned in questionnaires that they were distracted by the visuals, with only one student claiming they were helpful for language learning in the video. One participant in their questionnaire particularly pointed out that they thought Ms Möllemann and her actions were “peculiar” in the video with follow-up interviews discussing the bad acting, clearly illustrating that overdramatisation exhibited in the videos by actors could also be applied to Video 2. While this was mentioned as assisting students in Video 1, the use of the word “peculiar” and its negative associations provided by a student in the questionnaire and interview discussions around this video indicate that the overdramatisation was viewed in a negative light by students regarding Video 2. This could be due to the lack of overdramatisation in the form of repetition, which was seen as helpful in Video 1 by students.

As depicted in Chapter 5, Video 2 has the fastest speech rate with 309.84 wpm, which was a common complaint of the student participants in the study, as six students mentioned in the questionnaires that the speech rate was too fast and this also was a topic in the focus group interviews (see Chapter 5). The complaint was only made against Video 2, and not against Video 3 (268.75 wpm) and Video 1 (142.22 wpm), illustrating that their average word rates per minute were adequate for students to comprehend. This indicates a potential challenge of the videos in that a high speech rate could impact negatively on learning and it is recommended that facilitators of these videos pay attention to speech rate in determining

whether another playback of these videos is necessary. It also suggests that a range of speech rate for A1 GFL students in South Africa should not exceed 270 wpm and should not reach 300wpm or above for excellent comprehension.

Writing being present in the video was not memorable by students until I mentioned it and all student participants claim they only saw "Superwohnung" on the piece of paper Mr Waurich held up in the beginning of the video, missing the finer details written on the written advertisement. This corresponds to Mayer's (2008:764) redundancy principle, which states that people learn better through video when on-screen text is not present along with visuals and narration, even if the narration and on-screen text match in content. The writing was up for a short time as well, and this indicates that the lack of frequency of writing within the video may play a role in a lack of salience – thus suggesting that facilitators of these videos should be aware that students will pay less attention to writing in videos. If the writing is necessary for comprehension, facilitators should be sure to either, pause the video during the task or forewarn students to pay attention to it before the viewing of the video.

Section 6.3. Video 3

The states referred to in this section are taken from 5.3. and the student data from 5.3, 5.4 and 5.5. The language used in the video is conducive to the task at hand, as the majority of clauses contain Relational Processes, even the narrator's clause with the title of the clip and macroTheme "Das ist meine Stadt" (this is my city) (see 5.3.) and 25% of utterances are in the Declarative, which provide information. The task for Video 3 was a basic comprehension task allowing for students to cross out the correct answer, with answers containing two numeral figures and two nouns, example questions from this task are "Bern hat ____ Einwohner" (Bern has ____ residents) and students must choose between "150 000", "130 000" and "120 000". The most frequent Participants, however, were Carrier and Attribute, followed by Identified and Identifier, thus illustrating that the discussion content of the video led more to characteristics of objects being described than identified. Material Processes only scored 1.66% of the video content with Relational processes at 15.64% and the most prevalent linguistic Process, yet the majority of the video was comprised of Action Processes with Martin as Actor (40.98%, see 5.3), and Conceptual Processes only scored 2.73% of Processes. This shows a mismatch of visuals and speech that may lead to learning difficulties according to the cognitive theory of multimedia learning (see 3.3.). This is further

emphasised by the 21.11% Contrast state in Comparative relations, which is higher for Martin than Similarity and the low Simultaneity rate (2.78%) in the video (see 5.3).

Moving shots show more involvement from the audience. While there is less involvement from the audience in Action Processes with 79.33% of shots being stationary, the language, viewing perspective, gaze and social distance illustrate a high percentage of viewer participation elicited in the video overall. This seems, like in Video 2, to be yet another instance of mismatch between modes. The second highest Theme in the video is "Leute" (people), which Martin refers to the viewers in greeting "Hallo Leute" (hello people), followed by "kommt" (come) and "wie" (how). Furthermore, while the other Mood percentages were not as high, Interrogative and Imperative Moods do make an appearance in the video and the majority are directed at the audience, for example, "wie gefallen euch die Uhr und der Turm?" (how do you like the clock and the tower?). In addition, the highest state for viewing perspective combinations is front and eye level (53.63% see 5.3.); with another two 10% states being angled and eye level and front and low angle (see 5.3.). This majority viewing perspective of front and eye level depicts the viewer as a recognised participant that is involved in the discussions. This is also complemented by gaze and social distance in the videos, with the most frequent gazes being direct, and shots being medium and close shots (see 5.3.). Thus, the viewer becomes a noteworthy participant in the video, and this is shown by both the linguistic and visual modes.

Bern is positioned as a city that is an interesting place to visit, and this is already illustrated by the top Themes of the video being Bern and the "Zytglogge." Choices of what has been shown of Bern include the coat of arms, the "Zytglogge", *Bärndütsch* and the bear pits. The bear pits are an interesting choice as people could harbour negative feelings towards animal captivity. There is a Sequentiality of language first and visuals second that also emphasises the Theme and Relational Processes (17.22% of the video) in which Martin describes the "Zytglogge" and discusses bears before the camera depicts shots of them (see 5.3.). Martin is also in front of the Federal Palace of Switzerland but never mentions it or introduces it to his audience, so unless one already knows this building, one will not understand its semiotic significance. If signs of cultural significance are present in videos, this presents a learning opportunity for students that should be utilised, rather than simply ignored. Signs that carry cultural provenance include the Federal Palace of Switzerland, the "Zytglogge", *Bärndütsch* as a dialect, as well as the coats of arms for Berlin and Bern shown in the video. These signs

are icons of Bern, and by extension, Switzerland as a German-speaking country, thus they play a role in building students' knowledge of German-speaking cultures.

Martin is the dominant speaker in the video. This is not only illustrated in the number of utterances he has (54), but also in the fact that he is shown the majority of the time as Given information (25.68%) or centre (20.22%) (see 5.3.). In comparison, Salome is only Given information for 0.55% of the video, and the rest of her duration is positioned as New information, which is not frequent within the video's duration (see 5.3.). This highlights that Salome is not given a lot of screen time; she is simply there to provide an example sentence in *Bärndütsch*, and even Martin explains that sentence and not her. Furthermore, Salome's utterance does not classify as a clause, and consequently only Martin's language is analysed in terms of the metafunctions (see 5.3.). Martin is also a more prominent Theme than Salome (see 5.3.).

Despite the connotation within the video mainly being that Germans are friendly and helpful, there are differences in gender connotations. The majority of connotations are positive male associated with masculinity (55.80%), in which Martin shows viewers the city, and he is portrayed as helpful and friendly. By contrast, due to Salome's short screen time and limited speech, only 1.66% of the video shows positive women connotations associated with German women being helpful and friendly, when she says something in *Bärndütsch*. I explain at the end of the chapter why the positive connotations are associated with Germans rather than Swiss-Germans. Martin thus has the most power in the video, and Salome is recognised only for an example she provides and nothing else.

Therefore, from the analysis four points to be corroborated with the student data can be established:

1. There is a low amount of Similarity between visuals and language in Video 3, so similar to Video 2 and less similar to Video 1, but what does occur is that objects are presented after they are spoken, rather than simultaneously, and this is still good for learning (see the spatial contiguity principle of the cognitive theory of multimedia learning in Chapter 3).
2. The top Themes of the video are Bern, "Bärndütsch" and the "Zytglogge."
3. "Bärndutsch" may not be a prominent Theme, but it holds cultural provenance alongside other objects in the video and this along with Salome's sole purpose in the

video being to provide an example of it could have provided it with a specific salience.

4. Gender connotations may differ in terms of positivity and negativity but there are general positive connotations of speakers of German in general being friendly and helpful in Video 3.

Three participants stated specifically in the questionnaires that they enjoyed Video 3, despite it being rated less humorous than the other videos and rather average in terms of the categories presented to students, which were annoying and fun. Two participants found the visual aids helpful, with one student specifying in the Group A interview that it was helpful that the tower ("Zytglogge") was shown while being described. This suggests that students prefer simultaneous presentation, and even though this was lacking in Video 3 in comparison to Video 1, the Sequentiality of temporal visual-linguistic relations allowed for learning to still be effective in terms of the interplay of language and visual aids. Three students claimed to have learned new vocabulary as I had to answer, for example, the question what "Wappen" (coat of arms) was before the video was shown. The videos are presented as revision tools, therefore students learning new vocabulary from the videos does not mean that the videos can be used as learning tools for vocabulary introduction.

Three students identified the Swiss-German dialect as a focus and it was seen rather negatively, with one student finding it confusing and another two finding no point to the video due to the dialect not being a form of German they are learning. P6 in Group A stated it is a good video for cultural studies, but at the current language level, she does not see the point of Swiss-German being shown to them. This connects to the coherence principle in Mayer (2008:764), which states that extraneous material can distract students from more important material.

Gender connotations were not remarked on by students, but while connotations of Swiss people being friendly and helpful were not explicitly connected to Video 3 by students, the videos were mentioned as a salient portrayal from the three videos in general and thus this indicates that they formulated the same impressions that I did when viewing the video. Martin showing viewers the city emphasised the connotation, further illustrated by Salome also providing an example of *Bärndütsch* for viewers. According to PCSTIF, the following

schema (Figure 40) can thus represent impressions formulated by both students and me of Video 3:

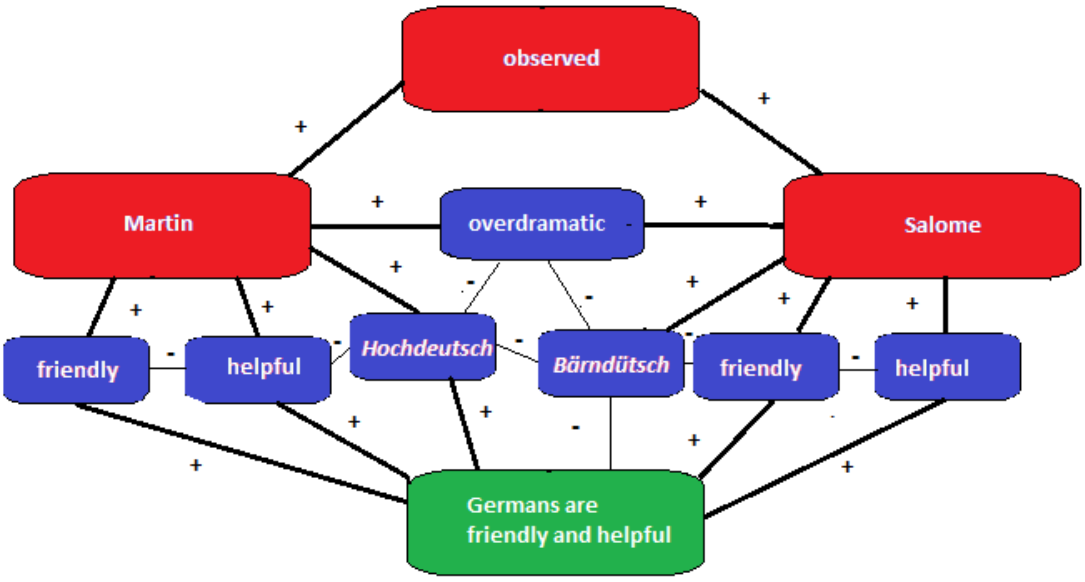

Figure 40: PCSTIF schema of Video 3

In Figure 40, the red indicates what was observed, the blue indicates impressions by students or by students and me and the green indicates the ultimate student impression based upon a collection of the associations made from observed behaviour above the node. I use the terms *Hochdeutsch* and *Bärndütsch* to identify how the students did observe a difference in the German spoken by each character. Even though Salome does not speak *Hochdeutsch*, but rather *Bärndütsch*, it would appear she is generalised as "German" and her presence further emphasises Martin's connotations of Germans being friendly and helpful, as well as those emphasised by Oliver in Video 1, due to the fact that students overall impressions were that Germans are friendly, helpful, funny and patient in the videos. There was no distinction made in this regard between Germans and Swiss-Germans.

Section 6.4. General analysis

In this section, I analyse general findings that relate to the videos overall from student interviews in terms of a thematic analysis and that are not related to specific videos. I triangulate these general findings with the literature discussed in Chapter 2. This will assist in answering research questions 1 and 5.

Overall, students commented that the videos assisted them regarding pronunciation, vocabulary (especially Video 2 and 3), general comprehension skills (especially Video 1) and listening comprehension, particularly in terms of visual aids (see Chapter 5). This corroborates studies reported on in Chapter 2. Pronunciation improvement is particularly corroborated by White *et al.* (2000:167) and their study on Spanish and Diab *et al.'s* (2016:11) South African study on isiZulu in 2.2. Listening comprehension and speaking are major themes in the literature in 2.2 in terms of the advantages of learning foreign languages, including German, with video, and this combined with visual aids that students mention identifies *Hör-Sehverstehen* (auditory and visual comprehension) in 2.2. as a benefit of video. Vocabulary as a learning opportunity from videos is also corroborated by many authors in 2.2., but students do not frequently mention interest or motivation generated by the videos, nor speaking as a skill obtained by these videos, although speaking was mentioned by one or two throughout the questionnaire and interview data as a potential advantage. This would be a skill that would need to be encouraged by teachers or facilitators of the video, so that as Dal (2010) warns, video watching does not become a passive activity, in which no speaking opportunities occur from them in the classroom. Students showed confusion in Video 2 and 3 that could be connected to a mismatch of visuals with language, a disadvantage of which Gruba (1999) warns. Jewitt (2006) furthermore voiced the concern that images can distract students, but rather than the visuals, students seemed to find the overdramatisation of pedagogic videos distracting, despite this feature of the videos also being found as providing good mnemonic techniques for vocabulary and pronunciation, especially when combined with repetition. Other difficulties mentioned that the literature did not predict include rapid speech rate and pronunciation variation as potential learning challenges.

"The pedagogic challenge par excellence" described by Kramsch and Andersen (1999:32) in 3.1. (a gap between text or representation and context or the lived) is apparent in the differences in the data between students and me. For example, the negative connotations that I found, such as that Germans are unintelligent in Video 1 or that German women are disorganised in Video 2 are not noticed by the students, rather connotations of videos overall are mostly positive features: friendly, funny/amusing, diverse and human (see 5.4 and 5.5). These traits are examples of Hall's (1959) model of covert/ invisible culture in terms of attitudes that manifest in overt/ visible actions such as the humour in the videos or Germans' actions within the videos of being hospitable by either giving directions and walking the

person to the destination as in Video 1. Furthermore, the traits identified by learners that are similar and different from my interpretations depict Neuner's "interim world" (2003:50) and the filtering of culture from teacher and authorities to learner perceptions, which is mentioned in 2.2. This is also manifested as shown by my interpretation of semiosis as discussed by Kramsch and Andersen (1999:32) in 3.1.

However, this gap between "the imitation and the lived" is identified by many students, mainly due to the overdramatisation identification in Chapter 5 and discussed in terms of the videos separately earlier in this chapter. In Group B, many students discussed how they were unsure of whether this was realistic or not as they had never been to Germany, and thus did not have any confidence in the veracity of their predetermined perceptions, as Neuner (2003:18) states, through which to filter the videos. This was shown by the exercise of judging the flat after the viewing of Video 2 and how Group A thought the flat was tiny (in comparison to the space in South Africa) versus Group B's discussion of whether the flat was cheap versus expensive to rent. Furthermore, the majority seem to agree with this distorted and mainly positive "interim world" (Neuner 2003:50) provided by the videos. There were 30% of students that were in fair disagreement, with one student (P1 in questionnaires Appendix B and Group A transcription in Appendix D) even refusing to answer the question 4b in the questionnaire stating the portrayals were not applicable, as representing culture was determined by the student as not being an important purpose of the videos. P1 and P6 discussed in the Group A interview (Appendix D) how the cultural portrayals of these videos were not the point of the video, rather the language was, illustrating that students were aware of the difference between representation and reality. However, they also, as determined by the history of cultural representation in textbooks in 2.1., seem to ignore the cultural components of the video and identify them as mainly assisting in language learning, consequently not seeing them as a unified whole. Furthermore, P1 claimed in an interview that the situations were seemingly realistic because he could translate them from his home culture. While students perceive other cultures around their own, it is important not to allow attitudes where cultures are viewed as more realistic to be perpetuated if they are similar to the video viewer's culture. As just because an aspect cannot be identified in terms of an L1 culture does not mean it is unrealistic. Thus, it is important to have focus group discussions with students in class, in order for condensed, idealised stereotypes to not be taken as truth, and that students feel Germans/ Swiss people are more "human" (see 5.4. and 5.5.) and

individuals over a national representation inform others of this too, allowing for a balancing of impression formation by students.

Despite the attempt at DACH (Germany, Austria and Switzerland) and cultural integration in these videos, students were often further confused by the presence of German-speaking cultural diversity between videos and the Swiss German dialect. One student (P3) claimed the Swiss German was, to her knowledge, not "proper" German. While this opinion may have originated through lived experience, media consumption or the textbook, it does show that the textbook and these resources, while attempting to treat all German-speaking countries equally, do not always manage to convey this to students. This is supported by the findings in the literature mentioned in 2.1. Overall students had difficulties with cultural portrayals that include what they termed unrealistic situations and overdramatisation, which they also admitted was an effective learning aid at times, as well as of the Swiss-German dialect within the video, but these did not interfere with their experience of success with learning language through them.

Furthermore, different learning styles were mentioned in both group transcriptions, with students stating that the videos could accommodate those who preferred visual aids in learning. As illustrated in Chapter 5, however, it is not just the presence of the visuals alone that assists learning, but how the modes interact with one another facilitates learning to take place, and learning styles as a theory of learning has become more of a widespread belief than grounded in actual practice, as discussed in Chapter 2. As shown in 5.4., most students see the videos as a study aid (55% of participants in the study, or 6 participants) and would recommend the videos, stating they helped with remembering of vocabulary and grammatical structures and exposure to German. In the transcriptions, a *Menschen A1* clip was mentioned that they had remembered, even though it did not feature as one of the videos shown to them, due to the overdramatisation present in this video. The overdramatisation, despite the complaints of students being distracted and uncomfortable because of this, seemed to be recognised as a tool to benefit their learning, rather than a challenge for most people.

Based on the findings in 5.4. and 5.5. regarding the high number of students participants who do not think the videos are being used quite as effectively as possible, more facilitator group discussions on the videos are necessary, with perhaps extra tasks for the videos that allow students more speaking opportunities and repeated viewings for certain videos such as Video

2. The vocabulary of that section should also be taught alongside the video by the teacher in terms of spoken language and written text to reinforce what is being seen and heard from the videos.

Video 1 has been shown to be the most effective video in terms of language learning, due to its adherence to the temporal and spatial contiguity principles and presence of intermodal syndromes, repetition, use of visual aids in terms of the arrows and speech rate, on which students have commented. Video 2 and 3 have displayed more Sequentiality over Simultaneity, which still adheres to the temporal contiguity principle, but have provided their own difficulties, including fast speech rate in Video 2 and the introduction of a Swiss German dialect in Video 3. In general, students commented that the videos assisted them in terms of language with pronunciation, vocabulary (especially Video 2 and 3), general comprehension skills (especially Video 1) and listening practice, particularly in terms of visual aids (see Chapter 5), thus illustrating the presence of *Hör-Sehverstehen* (auditory and visual comprehension) developed by the videos. Connotations of the videos are overall positive with Germans being seen as friendly, helpful and amusing and negative connotations not being consciously expressed by students. This has been shown to often be due to inhibitory connections in terms of PCSTIF. “The pedagogic challenge par excellence” described by Kramsch and Andersen (1999:32) is identified by many students and it is thus important to have discussions to which the facilitator and other students can contribute in ensuring idealised positive stereotypes are not taken as truth and a balanced impression of culture(s) of German-speaking countries is formed.

Chapter 7 - Conclusion

This chapter concludes this thesis by providing the key findings in 7.1., organised as responses to research questions. This chapter then discusses the limitations and opportunities for future research from this thesis (7.2.). It concludes in 7.3. with final remarks and reflections of this research.

7.1. Key findings

As mentioned above, this section organises key findings of this research by responses to research questions discussed in 1.4., in order to highlight how the research goals of this thesis have been reached.

Response to Question 1: What learning opportunities and challenges do RU students from diverse backgrounds experience in using pedagogic videos as a multimodal learning resource in German Studies 1?

Overall, improving general comprehension skills through the exposure to spoken German was also found to be a learning opportunity through using these videos, illustrating videos as helpful in building skills of *Hör-Sehverstehen* (auditory and visual comprehension). Interest, motivation were not mentioned by students, and speaking very little, the latter of which is a skill that may consequently not be fully realised by these videos. The overdramatisation present within the videos may have been a distraction to some students, but especially in Video 1, it was perceived to be a useful learning aid regarding vocabulary acquisition.

Students expressed that Video 1 is good for learning vocabulary for directions in German and its repetition is good for pronunciation and the spoken repetition and visual aids are also good for memorisation. A mismatch between video and reality was found to be a potential challenge in Video 1 in terms of an unrealistic manner of giving directions (see question 2 in 5.5.). This mismatch between the video and reality was also found in Video 2 by students, where they failed to find the video helpful in terms of learning how to deal with a real situation in Germany of viewing a potential flat for rent (see an example from P4 in 5.4.). Video 2 was found to be beneficial for learning, physical descriptions of a flat and to help with listening practice, but in general students had fairly negative impressions of the video. Overall, students felt that Video 2 contained too simple vocabulary, that the speech rate was

too fast and that the visuals were distracting, which could have been caused by overdramatisation and/ or the high amount of re-contextualisation and opposition between visuals and language discussed in the responses of Question 2 and 3 later in this section. The students recognised that Video 3 taught them new vocabulary and this is in contrast to the supplementary nature of the pedagogic videos within the textbook, proving that students may not be aware as to the purpose of these videos, and this may be considered as a challenge that educators would need to address. Furthermore, this also highlights that the other areas of the textbook from which they should have learned this vocabulary was either not examined by students previously or if it was, it was not memorable enough for students or repeated enough to ensure it left the area of their short-term memory. The pedagogic videos, however, consequently, allow for these words to be repeated in perhaps a way that is more memorable than their appearance within a static text, which provides a language learning opportunity. The manifestation of how these benefits and challenges are potentially realised within the videos themselves and the relationships between student perceptions of benefits and challenges are discussed in greater detail in the responses of research questions later in this section.

Response to Question 2: What is the relationship between the various modes within these videos?

With regard to Video 1, two inter-modal syndromes were discovered in this research (the Material-Action Process linguistic and visual combination and two types of visual Conceptual Processes combined with linguistic Circumstances of Extent and Location). These syndromes contribute to a high amount of Similarity in the Video and a low amount of Contrast. Video 1 has the highest amount of Similarity in comparison to the other two videos, which lack as many inter-modal syndromes. This shows the greater the number of inter-modal syndromes present within the video, the higher amount of Similarity that is present in the videos in terms of Comparative relations. In contrast, inter-modal syndromes were not observed in Video 2, rather visual Action processes were used to emphasise the names of the rooms through which the character of Ms Möllemann moves to be emphasised when discussed linguistically. Despite this, her descriptions of the rooms did not always match the visuals, resulting in a low Similarity rate in terms of Comparative relations. Like Video 2, Video 3 has a mismatch of linguistic and visual Processes (Action versus attributive Relational). The second most common Theme of “Leute”, common states of front and eye level angles, direct gaze and medium and close social distance illustrate the viewer as

participant in the video, as with Video 1 and 2. Due to the frequent instances of "Bern" as Theme and the visual semiotic signs of cultural provenance discussed in 6.3, Switzerland and Bern are positioned as interesting places to visit. Visuals with cultural significance used without linguistic introduction may result in an imbalance in the use of modes and no provisions made for learning about these cultural significant objects, as shown in Video 3 with the Federal Palace of Switzerland at the end of the video.

The DMDA of these videos depicts how the visual and auditory modes do not always combine in a way that presents a Similarity in meaning or Sequentiality, but Video 1, for example, exhibits a high amount of Similarity, while Video 3 has more instances of Sequentiality than any of the other three videos.

Response to Question 3: How do the relationships between the various modes in these pedagogic videos influence student learning?

Overall, Video 1 is an example of a helpful pedagogic video for the context of South African Higher Education. The high amount of Similarity in Video 1 adheres to the spatial and temporal contiguity principles in terms of Mayer's (2005) cognitive theory of multimedia learning (see 3.2. for details on principles). Close shots, indirect gaze, Declaratives, gestures and arrows emphasise information on directions being provided in the video, with arrows as visual aids assisting students. Syndromes and Sequentiality were also found helpful to students and the content repeated is important for the accompanying linguistic task on directions (see 6.1. for a description of this task). In terms of Video 2, the fast speech rate and the lack of Similarity between visuals and language negatively impact students' perception of their learning. The lack of Similarity between visuals and language was also a potential challenge in Video 3, but the rate of Sequentiality was greater in Video 3 than in Video 2 and this aided learning in terms of the temporal contiguity principle (see 3.2.). Sequentiality in combination with the high number of Relational Processes in the video were conducive for learning, as objects such as "der Zytglogge" were described and then shown.

Response to Question 4: How are the German-speaking characters represented in these videos with regard to culture(s) of German-speaking countries?

The dominant speakers Oliver in Video 1 and Martin in Video 3 evoke connotations that German men are friendly and helpful. Thus, culture(s) of German-speaking countries, including Swiss-Germans, are presented as being cultures that are friendly and helpful. In Video 1 and 3 these connotations are only extended to the male dominant speakers and not to

the women in the videos. In Video 2, Ms Möllemann is presented as disorganised due to framing cluster syndromes, camera movements, high frequency of Themes related to Ms. Möllemann and Imperatives and this stereotype may be transferred to other German women. Students, however, did not express a stereotype or impression of Germans as disorganised or German women as disorganised based upon this portrayal, discussed in response to question 5 below. Nevertheless, the portrayal of German culture in Video 2 is more negative than representations of culture(s) of German-speaking countries in Video 1 and 3, due to the presence of Ms. emann's disorganisation based upon the untidy visual presentation of her flat.

Response to Question 5: What cultural assumptions and impressions do students make based on the above representations?

The portrayal of Germans as impolite, like the tourist was to Oliver, and his escorting of her to her destination, is seen as unrealistic by students. Perhaps this was also a reason why the negative connotation of German women as disorganised from the depiction of Ms Möllemann was not salient to students. Student laughter during video viewings could be due to students finding videos humorous (some student stated they found the videos humorous and Germans portrayed as humorous but they did not explain specifically that their laughter during viewing was due to this). However, this could not only be due to videos being funny but also shocking, thus laughter could be due to discomfort, as some students mentioned Video 3 made them feel uncomfortable and that Videos 1 and 2 were cringeworthy (this in relation to theories of humour is discussed in 6.1. and 6.4.). This could be connected to the humorous representation of Germans, which was observed by students, or due to the identified overdramatisation present in all three videos. Despite negative impressions and impoliteness present in videos, only positive cultural assumptions and impressions were formulated by students. In general, our students that took part in this study find Germans portrayed as friendly and helpful in the videos. The majority of students agreed with this distorted presentation of culture(s) of German-speaking countries with only 30% that were in disagreement. *Bärndütsch* in Video 3 is not seen as helpful or recognised as "proper German" (i.e. recognised as different from the standard variety) and it proves distracting to students. This correlates to the coherence principle discussed in 3.2. Differences in culture in terms of flat size and pricing, as well as lack of experience in Germany, made it difficult for some students to accurately determine German cultural norms.

Response to Question 6: How can the videos be used in a manner that maximises the benefits and minimises the challenges outlined in Question 1?

This research shows the opportunities and challenges that these videos offer students in South African HE, which were not obvious as *Menschen* was not designed for a South African context. Video 1 has been highlighted as a type of pedagogic video that students find helpful to learning and recommendations for teachers / lecturers / facilitators have been generated in order to ensure that the opportunities are maximised and the challenges are minimised. The use of these videos is thus recommended in the classroom provided these recommendations answering research question 6 are observed.

Recommendations have been formulated based upon the above findings. Firstly, more videos should be used that illustrate a high Similarity rate between visuals and language. Teachers/ lecturers/ facilitators should ensure that the visuals adequately correspond (even if only sequentially) to language to ensure more effective learning takes place. If they do not, teachers/ lecturers/ facilitators could compensate by allowing for a second viewing of the video for students to hear or view information they may have missed upon the first viewing. Secondly, teachers/ lecturers/ facilitators should consider the speech rate of the videos. In a South African context, it is thus recommended that videos with a speech rate from 270 wpm should be given repeated viewings at A1 level, particularly if students complain of a rapid speech rate. Thirdly, teachers/ lecturers/ facilitators should allow for focus group discussions after watching the videos on how culture is presented within the videos in order to balance stereotype formation, as teachers/ lecturers/ facilitators' explanations of potential stereotypes as well as students' experience with Germans and in Germany can be used to counter damaging stereotypes. In addition, teachers/ lecturers/ facilitators need to be aware that due to the redundancy principle and often lack of frequency, writing within the videos may not be salient enough for students and thus if the writing present in the videos is important for comprehension, teachers/ lecturers/ facilitators need to ensure students have noticed it in videos.

Furthermore, teachers/ lecturers/ facilitators should point out important semiotic signs of cultural significance, in order for the cultural provenance to not be lost to students. This is to further ensure that students are aware that these videos are designed to impart some cultural knowledge and that they are not just language learning resources. Teachers/ lecturers/

facilitators should also teach relevant vocabulary before the video using both spoken language and written text to enforce what has been seen and heard in the videos.

Penultimately, teachers/ lecturers/ facilitators should foster interest and motivation for engagement with the videos, as well as ensure that enough speaking at A1 level occurs during activities with the videos, in order for speaking to not to be neglected. Finally, it is important that attitudes, in which cultures are viewed as not realistic if they differ to the video viewer's culture, are not given enough room to foster and teachers/ lecturers/ facilitators should ensure that discussions on cultural variation occur enough in the classroom in order for this to not occur.

7.2. Limitations of Study and Opportunities for Future Research

This research examines three videos from the *Menschen A1* course book. Further studies could examine more than three videos to ensure that the findings in this video encompass a greater data set, perhaps illustrating further examples of how the visual mode and linguistic mode co-occur in a manner that provides language learning opportunities to students. At the time of study, a multimodal corpus was not a feasible option, but an advancement in this area of research could yield studies of pedagogic videos that could allow for a fine-grained analysis of the videos across a larger data set. The parameters and scope of this study only required three videos, in order to conduct a fine-grained analysis of videos. More analysts would have been required in this case to analyse videos across a larger data set in a short space of time, and, while this research only focused on *Menschen A1* pedagogic videos specifically, and consequently examined three videos, the examination of more videos could be an area of future research to investigate if the findings in this thesis are the case for a wide variety of pedagogic videos (different GFL videos at different levels aside from the *Menschen* series) as well as in other contexts outside of South Africa. In this way, the recommendations that apply to this research and for the *Menschen A1* videos can further be used or adapted with additional recommendations with other pedagogic videos. Videos designed for higher levels of the CEFR would convey more complex language structures and tasks, thus possibly providing different kinds of challenges for students that may not be found at beginner level.

This study focused solely on student impressions and their learning experiences. A study that could examine the impressions of facilitators that use the videos in the GFL classroom would provide a holistic perspective to this research. A comparison of pedagogic videos to videos designed for a mother-tongue audience ('authentic' videos) could indicate the advantages and disadvantages of both of these types of video in the GFL classroom. It would provide an understanding of the similarities and differences between these types of videos, as well as the similarities and differences in their implementation within the GFL classroom. This research could also be undertaken in other FLs, L2s and L1s to examine whether the findings of this research also occur for other FLs or South African languages in a South African context and whether the recommendations generated in this thesis could apply to them as well, or whether linguistic differences provide additional or fewer recommendations.

This study also opens up opportunities for research that builds on the DMDA framework that was used – other semiotic resources of the modes examined in this research, such as incongruities between language and visuals could be examined in terms of student language learning opportunities and challenges and a set framework for pedagogic videos could be established. A set framework could be used to analyse pedagogic videos in languages other than German to show where the videos are useful to students and where they are confusing to students.

Further development on the education theories mentioned in this research such as Mayer's cognitive theory of multimedia learning and PCSTIF can be undertaken, as while this study has already attempted to develop PCSTIF, further research of this theory and its schemas is needed to examine how interpretations can differ amongst different people of different cultures and backgrounds.

Since technology is changing so rapidly, adaptations to Mayer's (2005) cognitive theory to examine the different aspects of how different semiotic resources within a mode are cognitively processed when viewing videos, for example, background noises and music in comparison to voices within the auditory channel from the auditory mode, could provide a complete picture of the complexity of how semiotic resources within modes are processed within pedagogic videos, but also within audio-visual media. Furthermore, Bandura's (1989) theory could be further examined in this kind of context by directing students to further reflect on what aspects of their lives outside of the classroom have allowed them to

experience the pedagogic videos in the ways that they have, as this was outside the scope of this research and possibly influenced the formation of student impressions when watching the videos. This would inform us exactly how these external factors contribute to shaping a student learning experience when learning with pedagogic videos.

Interest and motivation were not examined in this research, and consequently, studies on these learning aspects in relation to pedagogic videos both for GFL and other FLs, L2s and L1s, could be undertaken to examine exactly how these aspects relate to learning with this type of learning resource. Nevertheless much was gained from this research in particular; specifically the recommendations that have the potential to shape future GFL teaching and learning practices in South Africa for these kinds of videos at A1 level.

7.3. Final Remarks

As 7.1. has illustrated within the response to research question 6, recommendations from this research that can shape teaching and learning practices have been established. They allow clarification for GFL instructors on how to use this type of video at A1 level in a South African context in order to maximise the opportunities and minimise the challenges found in this research. As an online GFL teacher now in 2020, this research has provided me with knowledge on teaching with technological tools such as pedagogic videos to improve my students' learning experiences.

This thesis was the product of research around the domain of multimodality since 2016 and it would not have been undertaken in this manner if it had not been for the conversations I had shared with the academics mentioned in the Acknowledgements, both inside and outside of the university space. I have progressed from my Honours research and a multimodal social semiotic analysis undertaken by myself to a more fine-grained DMDA in combination with student interviews. This has resulted in a reflection on the strength of the methods undertaken in this research in comparison to a simple multimodal social semiotic analysis. This research has shown that interviews provide further support to a DMDA, as often an analyst identifies aspects of a multimodal text that are not salient to other viewers, but also other viewers can confirm aspects of the text that an analyst has seen to illustrate that they are indeed to be found within the text. The arrival of the Covid-19 pandemic has emphasised the need for more research such as mine, as digital learning resources such as pedagogic videos have

become a necessary tool to further education in a time where face-to-face social interaction is not possible between student and teacher and at the time of writing, the pandemic has not yet ended. Furthermore, Covid-19 has illustrated countries need to be equipped in terms of online and digital teaching and learning resources in preparation for future pandemic situations, and more research is needed on providing knowledge as to their assisted mediation, for the literature in Chapter 2 has shown that teachers as mediators are vital for the successful implementation of technology in the classroom. Multimodal research has never been more of a necessity and my hope is that this type of research can allow us to better understand language learning and cultural knowledge acquisition via technological resources in a language-rich country like South Africa.

Bibliography

Abrams, Z.I. (2016) Possibilities and challenges of learning German in a multimodal environment: A case study. *ReCALL*, 28 (3): 343–363.

Allen, H.W. (2008) "Textbook Materials and Foreign Language Teaching: Perspectives from the Classroom." In *NECTFL Review: A Journal for K-16+ Foreign Language Educators*. Vermont, 2008. pp. 5–28. Available at: www.bennington.edu.

Altay, İ.F. and Ünal, D.Ç. (2017) An Investigation of Foreign Language Courses with and without Video Technology. *International Journal of Curriculum and Instruction*, 9 (2): 49–69.

Amundrud, T.M. (2017) *Analyzing classroom teacher-student consultations: A systemic multimodal perspective*. Available at: https://www.academia.edu/36343207/Analyzing_classroom_teacher_student_consultations_A_systemic_multimodal_perspective?email_work_card=view-paper (Accessed: 27 May 2020).

Annas, R. (2016) Deutsch an Universitäten im südlichen Afrika: Zur Entwicklung des Fachs seit 2003. *Acta Germanica,* 44(1): 105-118.

Archer, A. (1997) *The function of the visual in teaching English as an additional language: The case of video.* Available at: http://hdl.handle.net/10539/21355 (Accessed: 27 May 2020).

Archer, A. (2004) *Access to academic practices in an engineering curriculum: drawing on students' representational resources through a multimodal pedagogy*. Available at: http://hdl.handle.net/11427/23682 (Accessed: 27 May 2020).

Archer, A. (2017) Using multimodal pedagogies in writing centres to improve student writing. *Stellenbosch Papers in Linguistics Plus*, 53 (1): 1–12.

Archer, A. and Newfield, D. (2014) "Challenges and Opportunities of Multimodal Approaches to Education in South Africa." In Archer, A. and Newfield, D. (eds.) *Multimodal Approaches to Research and Pedagogy: Recognition, Resources, and Access*. Oxon, New York: Routledge.

Ashby, W. (2003) *Authoring the German "other": A semiotic, narrative discourse analysis of the culture box in beginning L2 German textbooks Item Type text; Dissertation-Reproduction*. Available at: http://hdl.handle.net/10150/280250 (Accessed: 27 May 2020).

Banda, F. and Oketch, O. (2011) Localizing HIV/AIDS discourse in a rural Kenyan community. *Journal of Asian and African studies*, 46 (1): 19–37.

Bandura, A. (1971) *Social Learning Theory*. New York: General Learning Press.

Bandura, A. (1989) "Social cognitive theory." In Vasta, R. (ed.) *Annals of child development. Vol. 6. Six theories of child development*. Greenwich, CT: JAI Press. pp. 1–60.

Bateman, J., Wildfeuer, J., and Hiipala, T. (2017) *Multimodality: Foundations, Research and Analysis – A Problem-Orientated Introduction*. Berlin/Boston: Walter Gruyer.

Bauer, G.U. (2016) 'Caught' between foreign cultural and development policies: perspectives of German Studies and German as a foreign language in Africa. *International Journal of Cultural Policy*, 22 (4): 611–630.

Baumeister, R. (2001) *Evil: Inside Human Violence and Cruelty*. New York: Holt Paperbacks.

Berk, R.A. (2009) Multimedia teaching with video clips: TV, movies, YouTube, and mtvU in the college classroom. *International Journal of Technology in Teaching and Learning*, 5 (1): 1–21. Available at: www.ronberk.com (Accessed: 27 May 2020).

Bezemer, J. and Jewitt, C. (2010) "Multimodal Analysis: Key issues." In Litosseliti, L. (ed.) *Research Methods in Linguistics*. London: Continuum. pp. 180–197.

Biechele, B. (2000) Einblicke. Ein deutscher Sprachkurs. 4 Videocassetten, 6 Begleithefte für Selbstlerner,Handbuch für den Unterricht, 4 Tonkassetten für Selbstlerner, Dialogbuch. *Informationen Deutsch als Fremdsprache*, 27 (2–3): 166–168. Available at: https://www.degruyter.com/view/journals/infodaf/27/2-3/article-p166.xml (Accessed: 28 May 2020).

Biechele, B. (2006) *Film / Video / DVD auch auf Niveaustufe A1?* Available at: https://www.iik.de/node/210 (Accessed: 28 May 2020).

Biechele, B. (2007) "„Ich sehe was, was du nicht siehst" – Reflexionen zum Lernen mit Spielfilmen im Unter- richt Deutsch als Fremd- oder Zweitsprache." In Esser, R. and Krumm, H.-J. (eds.) *Bausteine für Ba- bylon: Sprache, Kultur, Unterricht. Festschrift zum 60. Geburtstag von Hans Barkowski*. München: Iudicium. pp. 194–205.

Biechele, B. (2017) "Wege zu einer 'film literacy' im Unterricht Deutsch als Fremd- und Zweitsprache: Potenziale und methodische Ansätze." In Brede, J.R. and Helmes, G. (eds.) *Vielfalt und Diversität in Film und Fernsehen: Behinderung und Migration im Fokus*. Münster – New York: Waxmann-Verlag. pp. 193–211.

Biechele, B., Böttcher, D. and Kittner, R. (2001) Ausgewählte Ergebnisse einer empirischen Untersuchung zu subjektiven Theorien. Erfahrungen von DaF-Studierenden in Bezug

auf (neue) Medien und Lernen. *Informationen Deutsch als Fremdsprache*, 28 (4): 343–368. Available at: https://www.degruyter.com/view/journals/infodaf/28/4/article-p343.xml (Accessed: 28 May 2020).

Bloor, T. and Bloor, M. (2004) *The Functional Analysis of English*. 2nd ed. New York: Oxford University Press.

Bock, K. (2015) *Landesbilder und Stereotype in Lehrwerken für den DaF-Unterricht*.

Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3 (2): 77–101.

Breen, R.L. (2006) A Practical Guide to Focus-Group Research. *Journal of Geography in Higher Education*, 30 (3): 463–475.

Brinitzer, M., Hantschel, H.-J., Kroemer, S., et al. (2013) *DaF unterrichten Basiswissen Didaktik Deutsch als Fremd- und Zweitssprache*. Stuttgart: Ernst Klett Sprachen GmBH.

Brunsing, T. (2016) Landeskundliche Abbildungen in Lehrbüchern für Deutsch als Fremdsprache. Die Entwicklung des Bildeinsatzes in den Jahren 2000 bis 2010. *Informationen Deutsch als Fremdsprache*, 43 (5): 494–515. Available at: https://www.degruyter.com/view/journals/infodaf/43/5/article-p494.xml?language=de (Accessed: 27 May 2020).

Byram, M. (2000) *Routledge Encyclopedia of Language Teaching and Learning*. London, New York: Routledge.

Byram, M. and Hu, A. (2013) *Routledge Encylopedia of Language Teaching and Learning*. Oxon, New York: Routledge.

Canale, M. and Swain, M. (1980) Theoretical Bases of Communicative Approaches To Second Language Teaching and Testing. *Applied Linguistics*, 1 (1): 1–47.

Chudak, S. (2012) Lernstrategien im Umgang mit Hör-Seh-Texten (Versuch einer Übersicht) Language learning strategies for dealing with audiovisual texts (An Attempt at an overview). *Glottodidactica. An International Journal of Applied Linguistics*, 39 (2): 87.

Çiftçi, Y. (2016) A Review of Research on Intercultural Learning through Computer-Based Digital Technologies. *Journal of Educational Technology & Society*, 19 (2): 313–327.

Collins, M.G. (2017) *An action research approach: developing intercultural competence in German Studies at Rhodes University*. Available at: http://hdl.handle.net/10962/63910 (Accessed: 27 May 2020).

Condray, K. (2012) *III. Elementary and Intermediate German Textbooks (Adults, A1–A2)*. Cherry Hill: American Association of Teachers of German.

Council of Europe (2019) *Introduction*. Available at: https://www.coe.int/en/web/common-european-framework-reference-languages/introduction-and-context (Accessed: 9 June 2019).

Dal, M. (2012) "Digital Video Production and Foreign Language Learning." In *Conference: World Conference on Educational Multimedia, Hypermedia and Telecommunications*. 2012. pp. 1–13. Available at: https://www.researchgate.net/publication/277984755_Digital_Video_Production_and_Foreign_Language_Learning (Accessed: 27 May 2020).

Datenerhebung (2020). "Deutsch als Fremdspache weltweit". Available at: https://www.auswaertiges-amt.de/blob/2344738/b2a4e47fdb9e8e2739bab2565f8fe7c2/deutsch-als-fremdsprache-data.pdf (Accessed: 28 February 2021).

Department of Basic Education (2012) *Curriculum and Assessment Policy Statement Grades 7 - 9: Languages German Second Additional Language*. Available at: https://www.education.gov.za/LinkClick.aspx?fileticket=yB_flwk7ycQ%3D&tabid=573&portalid=0&mid=1571&forcedownload=true (Accessed: 27 May 2020).

Diab, P., Matthews, M. and Gokool, R. (2016) Medical students' views on the use of video technology in the teaching of isiZulu communication, language skills and cultural competence. *African Journal of Health Professions Education*, 8 (1): 11.

Divjak, D. (2019) *Frequency in Language: Memory, Attention and Learning*. Singapore: Cambridge University Press.

Dixon, Q.L., Zhao, J., Shin, J-Y, Wu, S., Su, J-H, Burgess-Brigham, R, Gezer, M.U., Snow, C. (2012) What we know about Second Language Acquisition: A Synthesis From Four Perspectives. *Review of Educational Research*, 82 (1): 5-60.

Drewelow, I. (2012) Learners' perceptions of culture in a first-semester foreign language course. *L2 Journal*, 4 (2): 283–302. Available at: https://escholarship.org/uc/item/73m1p5dx (Accessed: 27 May 2020).

Evans, S., Pude, A. and Specht, F. (2012) *Menschen A1 Kursbuch*. Ismaning: Hueber Verlag.

Farias, M.A., Obilinovic, K. and Orrego, R. (2011) Engaging Multimodal Learning and Second/Foreign Language Education in Dialogue. *Trabalhos em Lingüística Aplicada*, 50 (1): 133–151.

Ferreira-Meyers, K.A.F. and Horne, F. (2017) Multilingualism and the language curriculum in South Africa: contextualising French within the local language ecology. *Stellenbosch Papers in Linguistics Plus*, 51: 23–40.

Ferris, F.S. and Banda, F. (2015) "Poof! a'm heppily saving the Lord…": multimodality and evaluative discourses in male toilet graffiti at the University of the Western Cape. *African Identities*, 13 (4): 243–261.

Gachago, D., Cronje, F., Ivala, E., et al. (2014) Using Digital Counterstories as Multimodal Pedagogy among South African Pre-service Student Educators to produce Stories of Resistance. *Electronic Journal of e-Learning*, 12 (1): 29–42.

Garrett, N. (1991) Technology in the service of language learning: Trends and issues. *The Modern Language Journal*, 75: 74–101.

Garza, T.J. (1996) The Message Is the Medium: Using Video Materials To Facilitate Foreign Language Performance. *Texas Papers in Foreign Language Education*, 2 (2): 3–21.

Gottschalk, G.H. (1965) Closed-Circuit Television in Second Semester College German. *The Modern Language Journal*, 49 (2): 86–91.

Gruba, P.A. (1999) *The role of digital video media in second language listening comprehension*.

Guest, G., Bunce, A. and Johnson, L. (2006) How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods*, 18 (1): 59–82.

Hall, E.T. (1959) *The Silent Language*. 1st ed. New York: Doubleday & Company Inc.

Halliday, M.A.K. (1976) *System and Function in Language*. Kress, G.R. (ed.). London: Oxford University Press.

Halliday, M.A.K. and Matthiessen, C.M.I.M. (2014). *Halliday's Introduction to Functional Grammar*. Oxon, New York: Taylor and Francis.

Herman, C. and Sehoole, C. (2014) *Research and PhD Capacities in Sub-Saharan Africa: South Africa Report*.

Herron, C., Cole, S.P., Corrie, C., et al. (1999) The Effectiveness of a Video-Based Curriculum in Teaching Culture. *The Modern Language Journal*, 83 (4): 518–533.

Herron, C., Dubreil, B., Corrie, C., et al. (2002) A Classroom Investigation: Can Video Improve Intermediate-Level French Language Students? Ability to Learn about a Foreign Culture? *The Modern Language Journal*, 86 (1): 36–53.

Hodge, R. and Kress, G. (1988) *Social Semiotics*. Cornwall: Cornell University Press.

Holoch, A.M. (2012) *The serious work of humor in postcolonial literature*. Available at: https://ir.uiowa.edu/cgi/viewcontent.cgi?article=5684&context=etd (Accessed: 27 May 2020).

Hueber.de (2019) *Lehrwerk entdecken: Entdecken Sie die Besonderheiten von Menschen*. Available at: https://www.hueber.de/menschen/entdecken?tabid452019=6 (Accessed: 25 February 2019).

Iedema, R. (2001) "Analysing film and television: a social semiotic account of Hospital: an Unhealthy Business." In van Leeuwen, T. and Jewitt, C. (eds.) *Handbook of Visual Analysis*. London, California, New Delhi: SAGE Publications Ltd. pp. 183–206.

Jewitt, C. (2006) *Technology, Literacy, Learning*. Oxon, New York: Routledge.

Jewitt, C. (2009) "Different Approaches to Multimodality." In *The Routledge Handbook of Multimodal Analysis*. London: Routledge. pp. 28–39.

Jewitt, C., Bezemer, J. and O'Halloran, K. (2016) *Introducing Multimodality.* Oxon: Routledge.

Kalender, S. and Pude, A. (2013) *Menschen A1: Deutsch als Fremdsprache Lehrerhandbuch*. Ismaning: Hueber.

Kaltenbacher, M. (2004) "Multimodality in language teaching CD-ROMs." In Ventola, E., Charles, C. and Kaltenbacher, M. (eds.) *Perspectives on Multimodality*. Amsterdam, Philadelphia: John Benjamins Publishing Company. pp. 119–136.

Karpinnen, S. (2016) *Kultur der Deutschsprachigen Länder in Finnischen Lehrbüchern: Gemeinsamkeiten und Unterschiede zwischen zwei DaF-Lehrbüchern des gymnasialen Niveaus B2*. Available at: https://jyx.jyu.fi/handle/123456789/49547 (Accessed: 27 May 2020).

Kellerman, S. (1992) 'I See What You Mean': The Role of Kinesic Behaviour in Listening and Implications for Foreign and Second Language Learning. *Applied Linguistics*, 13 (3): 239–258. Available at: https://academic.oup.com/applij/article/13/3/239/235714 (Accessed: 27 May 2020).

Kelly, S.D., McDevitt, T. and Esch, M. (2009) Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes*, 24 (2): 313–334.

Knox, J.S., Patpong, P. and Piriyasilpa, Y. (2011) "**ขาวหน าหน** ่ ้ ็ ่**ง** (Khao naa nung): A Multimodal Analysis of Thai-language Newspaper Front Pages." In Martin, J.R. and Bednarek, M. (eds.) *New Discourse on Language: Functional Perspectives on*

Multimodality, Identity, and Affiliation. London & New York: Continuum International Publishing Group. pp. 80–110.

Königs, F. (2003) Teaching and learning foreign languages in Germany: A personal overview of developments in research. *Language Teaching*, 36: 235–251.

Kramsch, C. (1993) Language Study as Border Study: Experiencing Difference. *European Journal of Education*, 28 (3): 349.

Kramsch, C. (1995) The cultural component of language teaching. *Language, Culture and Curriculum*, 8 (2): 83–92.

Kramsch, C. (2008) "The Intercultural Yesterday and Today: Political Perspectives." In Schultz, R.A. and Tschirner, E. (eds.) *Communicating across Borders*. Munich: Iudicium. pp. 1–361.

Kramsch, C. (2013) Culture in foreign language teaching. *Iranian Journal of Language Teaching Research*, 1 (1): 57–78.

Kramsch, C. and Andersen, R.W. (1999) Teaching Text and Context Through Multimedia. *Language Learning & Technology*, 2 (2): 31–42.

Kress, G. (2015) Semiotic work: Applied Linguistics and a social semiotic account of Multimodality. *AILA Review*, 28: 49–71.

Kress, G. and van Leeuwen, T. (1996) *Reading images: The grammar of visual design*. London: Routledge.

Kress, G. and van Leeuwen, T. (2001) *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold Publishers.

Kress, G. and van Leeuwen, T. (2006) *Reading images: The grammar of visual design*. 2nd Edition. Oxon: Routledge.

Kruger, J.L. and Doherty, S. (2016) Measuring cognitive load in the presence of educational video: Towards a multimodal methodology. *Australasian Journal of Educational Technology*, 32 (6): 19–31.

Kunda, Z. and Thagard, P. (1996) Forming Impressions From Stereotypes, Traits and Behaviors: A Parallel-Constraint-Satisfaction Theory. *Psychological Review*, 103 (2): 284–308.

Lange, L. (2017) 20 Years of higher education curriculum policy in South Africa. *Journal of Education*, (68): 31–58.

Lay, T. (2009) Film und Video im Fremdsprachenunterricht. Eine empirisch quantitative Erhebung zur didaktisch-methodischen Implementierung filmspezifischer Arbeit im universitären Deutschstudium Taiwans. *Zeitschrift für Interkulturellen*

Fremdsprachenunterricht, 14 (1). Available at: https://tujournals.ulb.tu-darmstadt.de/index.php/zif/article/view/210 (Accessed: 28 May 2020).

van Leeuwen, T. (2005) *Introducing Social Semiotics*. Oxon, New York: Routledge.

Leitner, N. (2015) *Repräsentanz von Vielfalt in DaF- Lehrwerken*. Available at: https://othes.univie.ac.at/38487/ (Accessed: 27 May 2020).

Levno, A.W. and Pfister, G.G. (1980) An Analysis of Surface Culture and Its Manner of Presentation in First-Year College French Textbooks from 1972 to 1978. *Foreign Language Annals*, 13 (1): 47–52.

Lim, F.V. (2004) "Developing an integrative multi-semiotic model." In O'Halloran, K. (ed.) *Multimodal discourse analysis: Systemic functional perspective*. London: Continuum. pp. 220–246.

Lim F.V. (2011) *A Systemic Functional Multimodal Discourse Analysis Approach to Pedagogic Discourse*.

de Linde, Z. and Kay, N. (2016) *The Semiotics of Subtitling*. Oxon, New York: Routledge.

Liu, Y. and O'Halloran, K. (2009) Intersemiotic Texture: analyzing cohesive devices between language and images. *Social Semiotics*, 19 (4): 367–388.

Löschmann, M. (1989) Fernsehen und Video als Lehr- und Lernmittel im Fremdsprachenunterricht. *Deutsch als Fremdsprache - Zeitschrift zur Theorie und Praxis des Deutschunterrichts für Ausländer*, 26 (1): 24–28.

Macedonia, M. and von Kriegstein, K. (2012) Gestures Enhance Foreign Language Learning. *Biolinguistics*, 6 (3–4): 393–416.

MacWilliam, I. (1986) Video and language comprehension. *ELT Journal*, 40 (2): 131–135.

Mafofo, L. and Banda, F. (2014) Accentuating institutional brands: A multimodal analysis of the homepages of selected South African universities. *Southern African Linguistics and Applied Language Studies*, 32 (4): 417–432.

Maijala, M. (2007) Was ein Lehrwerk können muss – Thesen und Empfehlungen zu Potenzialen und Grenzen des Lehrwerks im Unterricht Deutsch als Fremdsprache. *Informationen Deutsch als Fremdsprache*, 34 (6): 47–52. Available at: https://www.degruyter.com/view/journals/infodaf/34/6/article-p543.xml (Accessed: 27 May 2020).

von Maltzan, C. (2009) Sprachenpolitik und die Rolle der Fremdsprachen (Deutsch) in Südafrika. *Stellenbosch Papers in Linguistics Plus*, 38: 205–214.

Marefat, F. (2002) Communicating Culture through Video. *Intercultural Communication Studies*, 11 (4): 141–148. Available at:

https://pdfs.semanticscholar.org/d76a/e5850942f26c07d24364ea0079b25b966589.pdf (Accessed: 27 May 2020).

Martin, J. R. and Rose, D. (2007) *Working with Discourse: Meaning beyond Clause*. London, New York: Continuum.

Mayer, R.E. (2005) "Cognitive Theory of Multimedia Learning." In Mayer, R.E. (ed.) *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press. pp. 31–48.

Mayer, R.E. (2008) Applying the Science of Learning: Evidence-Based Principles for the Design of Multimedia Instruction. *American Psychologist*, 63 (8): 760–769.

McConachy, T. (2018) Critically engaging with cultural representations in foreign language textbooks. *Intercultural Education*, 29 (1): 77–88.

Mekheimer, M.A.A.G. (2011) The Impact of Using Videos on Whole Language Learning in EFL Context. *Arab World English Journal*, 2 (2): 5–39. Available at: https://awej.org/images/AllIssues/Volume2/Volume2Number2April2011/1.pdf (Accessed: 27 May 2020).

Ministry of Education (2002) *Language Policy for Higher Education*. Available at: http://www.dhet.gov.za/Management Support/Language Policy for Higher Education.pdf (Accessed: 27 May 2020).

Ministry of Education (2017) *Draft Language Policy for Higher Education*. Available at: http://www.dhet.gov.za/Policy and Development Support/Government Notice Revised Language Policy for Higher Education.pdf (Accessed: 27 May 2020).

Mitra, B., Lewin-Jones, J., Barrett, H., et al. (2010) The use of video to enable deep learning. *Research in Post-Compulsory Education*, 15 (4): 405–414.

Neuner, G. (2003) "Socio-Cultural Interim Worlds in Foreign Language Teaching and Learning." In Byram, M. (ed.) *Intercultural Competence*. Strasbourg: Council of Europe. pp. 17–62. Available at: https://rm.coe.int/16806ad2dd (Accessed: 27 May 2020).

Nielsen, M. (2016) Die heile Welt in der Werbung - Stereotype als Bestandteil von Werbstil. *Linguistik online*, 79 (5): 231–252.

Nova, J.C., Chavarro, I.O. and Córdoba, A.T.Z. (2017) Educational videos: a didactic tool for strengthening English vocabulary through the development of affective learning in kids. *Gist Education and Learning Research Journal*, (14): 68–87.

O'Halloran, K. (2005) *Mathematical discourse: Language, symbolism and visual images*. London: Continuum.

O'Halloran, K. (2008) Systemic functional-multimodal discourse analysis (SF-MDA): constructing ideational meaning using language and visual imagery. *Visual Communication*, 7 (4): 443–475.

O'Halloran, K. and Lim Fei, V. (2014) "Systemic functional multimodal discourse analysis." In Norris, S. and Maier, C. (eds.) *Texts, Images and Interactions: A Reader in Multimodality*. Berlin: Mouton de Gruyter. pp. 137–153.

O'Halloran, K., Tan, S. and Marissa, K.L.E. (2015) Multimodal analysis for critical thinking. *Learning, Media and Technology*, pp. 1–24.

O'Halloran, K., Tan, S., and Smith, B.A. (2011) Multimodal analysis within an interactive software environment: critical discourse perspectives. *Critical Discourse Studies*, 8 (2): 109–125.

O'Halloran, K., Tan, S. and Wignell, P. (2018) "SFL and Multimodal Discourse Analysis." In Thompson, G., Bowcher, W.L., Fontaine, L., et al. (eds.) *The Cambridge Handbook of Systemic Functional Linguistics*. Cambridge: Cambridge University Press.

O'Toole, M. (1994) *The Language of Displayed Art*. London: Leicester University Press.

Petersen, U.H. (2012) *Erster Einstieg in die Systemisch Funktionale Grammatik des Deutschen*. In 2012. pp. 1–14.

Pisarenko, V. (2017) Teaching a foreign language using videos. *Social Sciences*, 6 (4).

Plass, J.L., Chun, D.M., Mayer, R.E., et al. (1998) Supporting Visual and Verbal Learning Preferences in a Second-Language Multimedia Learning Environment. *Journal of Educational Psychology*, 90 (1): 25–36.

Pretorius, E.J. and Mokhwesana, M.M. (2009) Putting reading in northern sotho on track in the early years: Changing resources, expectations and practices in a high poverty school. *South African Journal of African Languages*, 29 (1): 54–73.

Ramirez, A.G. and Hall, J.K. (1990) Language and Culture in Secondary Level Spanish Textbooks. *The Modern Language Journal*, 74 (1): 48–65.

Rapp, B. (2002) "Uncovering the cognitive architecture of spelling." In Hillis, A. (ed.) *The handbook of adult language disorders*. New York: Psychology Press. pp. 47–62.

Rhodes University Language Policy (2019) *Rhodes University Language Policy*. Available at: https://www.ru.ac.za/media/rhodesuniversity/content/institutionalplanning/documents/Language_Policy_2019.pdf (Accessed: 27 May 2020).

Risager, K. (2018) *Representations of the World in Language Textbooks*. Bristol: Multilingual Matters.

Roche, J.M. (2010) "Audiovisuelle Medien." In Wiegand, H.E. (ed.) *Handbücher zur Sprach- und Kommunikationswissenschaft*. Band 2. Berlin/New York: Walter de Gruyter. pp. 1243–1251.

R , D. (2010) "Die Funktion von Medien im Deutsch als Fremd- und Deutsch als Zweitsprache-Unterricht." In Wiegand, H.E. (ed.) *Handbücher zur Sprach- und Kommunikationswissenschaft*. Band 2. Berlin/New York: Walter de Gruyter. pp. 1199–1214.

Rösler, D. (2012) *Deutsch als Fremdsprache: Eine Einführung*. Stuttgart: J.B. Metzler Verlag.

R sler, D. (1994) *Deutsch als Fremdsprache*. Stuttgart: Verlag J.B. Metzler.

Ruthner, A.R. (2013) Der Einsatz authentischer Texte zur Vermittlung interkultureller Kompetenz im DaF-Unterricht. *Temeswarer Beiträge zur Germanistik*, 10: 209–220. Available at: http://www.diacronia.ro/en/indexing/details/A13062 (Accessed: 27 May 2020).

Salaberry, M.R. (2001) The use of technology for second language learning and teaching: A retrospective. *Modern Language Journal*, 85 (1): 39–56.

Sass, A.((2007) Filme im Unterricht – Sehen(d) lernen. *Fremdsprache Deutsch*, 36: 5–13.

de Saussure, F. (1974) *Course in General Linguistics*. Culler, J. (ed.). London: Fontana.

Scherpinski, A. (2014) Potenzial audiovisueller Texte für das Fremdsprachenlernen - eine vergleichende quantitative Untersuchung mit koreanischen Germanistikstudierenden zu Verstehensleistungen bei auditiver und audiovisueller Textrezeption. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 19 (2). Available at: https://tujournals.ulb.tu-darmstadt.de/index.php/zif/article/view/38 (Accessed: 28 May 2020).

Schwerdtfeger, I.C. (1980) Perception-Oriented Foreign Language Learning and Teaching: Theory into Practice-Notional Syllabuses. *System*, 8: 23–34.

Schwerdtfeger, I.C. (1989) *Sehen und Verstehen. Arbeit mit Filmen im Unterricht Deutsch als Fremdsprache*. Berlin: Langenscheidt.

Scott, C. (2010) The enduring appeal of 'learning styles.' *Australian Journal of Education*, 54 (1): 5–17.

Shaughnessy, M. (2003) CALL, commercialism and culture: Inherent software design conflicts and their results. *ReCALL*, 15 (2): 251–268.

Sigurðardóttir, S.Þ. (2016) *Warum sind Filme ein gutes Lehrmittel im Fremdsprachenunterricht?* Available at: http://skemman.is/stream/get/1946/25938/58956/2/MA_Svanborg_Sigur%C3%B0ardo$0301ttir_4._sept_2016.pdf (Accessed: 27 May 2020).

Spitzberg, B.H. and Changnon, G. (2009) "Conceptualizing intercultural competence." In *Conceptualizing intercultural competence*. California: SAGE Publications Ltd. pp. 2–52.

Stein, P. and Newfield, D. (2006) Multiliteracies and multimodality in English in Education in Africa: Mapping the terrain. *English Studies in Africa*, 49 (1): 1–22.

Stempleski, S. (2002) "Video in the ELT Classroom: The Role of the Teacher." In *Methodology in Language Teaching*. Cambridge University Press. pp. 364–367.

Stickler, U. and Shi, L. (2016) TELL us about CALL: An introduction to the Virtual Special Issue (VSI) on the development of technology enhanced and computer assisted language learning published in the System Journal. *System*, 56: 119–126.

S 2004) "In between Modes: Language and Image in Printed Media." In Ventola, E., Charles, C. and Kaltenbacher, M. (eds.) *Perspectives on Multimodality*. Philadelphia & Amsterdam: Benjamins. pp. 119–136.

Talaván, N. (2007) Learning Vocabulary through Authentic Video and Subtitles. *TESOL-SPAIN Newsletter*, 31: 5–8. Available at: https://www.researchgate.net/publication/235557228_Learning_Vocabulary_through_Authentic_Video_and_Subtitles (Accessed: 27 May 2020).

Tan, S., O'Halloran, K. and Wignell, P. (2016) Multimodal research: Addressing the complexity of multimodal environments and the challenges for CALL. *ReCALL*, 28 (3): 253–273.

Tan, S., O'Halloran, K. and Wignell, P. (2018) "Multimodality." In de Fina, A. and Georgekopoulou, A. (eds.) *Handbook of Discourse Studies*. Cambridge: Cambridge University Press. pp. 1–29.

Tan, S., Smith, B.A. and O'Halloran, K. (2015) Online leadership discourse in higher education: A digital multimodal discourse perspective. *Discourse & Communication*, 9 (5): 559–584.

The Constitution of South Africa (1996) *The Constitution of South Africa*. Available at: https://www.justice.gov.za/legislation/constitution/SAConstitution-web-eng.pdf (Accessed: 27 May 2020).

Thesen, L. (2001) Modes, literacies and power: A university case study. *Language and Education*, 15 (2–3): 132–145.

Tschirner, E. (2001) Language acquisition in the classroom: The role of digital video. *Computer Assisted Language Learning*, 21 (1): 305–319.

Uhlen, R. (2018) "Videos im DaF-Unterricht." In *Workshop österreichische und deutsche Landeskunde Videos im DaF-Unterricht*. Odessa Oblast, 2018. pp. 1–3.

Ulrich, N. (2004) Towards a methodology for analysing the cultural content of Modern Foreign Language textbooks. *Internationale Schulbuchforschung*, 26 (2): 165–179.

University of Pretoria (2020) *German*. Available at: https://www.up.ac.za/modern-european-languages/article/39497/about-german (Accessed: 9 June 2020).

University of Witwatersrand Johannesburg (2003) *University of Witwatersrand Johannesburg Language Policy*.

Vaismoradi, M., Turunen, H. and Bondas, T. (2013) Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing and Health Sciences*, 15 (3): 398–405.

Vanderplank, R. (2009) Déj vu A decade of research on language laboratories, television and video in language learning. *Language Teaching*, 43 (1): 1–37.

Wahba, S. and Zappen-Thomson, M. (1994) Video im Fremdsprachunterricht - ein praktisches Beispiel des Einsatzes an der Universität Namibia. *Deutschunterricht im südlichen Afrika*, 25 (1): 29–47.

Weber, A. (2018) Teaching Grammar through Literature in Foreign Language Learning with German as an Example. *Per Linguam*, 34 (1): 78–89.

Weber, U.S. and Domingo, R.S.C. (2011) "Adding another colour to the rainbow: An attempt at importing German cultural competence in a South African university context." In Harden, T. and Witte, A. (eds.) *Intercultural Competence. Concepts, Challenges, Evaluations*. Oxford: Peter Lang Verlag. pp. 117–191.

Wehage, F.-J. (1997) Einsatz von Videos im DaF-Unterricht aus US-amerikanischer. *Informationen Deutsch als Fremdsprache*, (4): 523–529.

Weninger, C. and Kiss, T. (2015) "Analyzing culture in foreign/second language textbooks: Methodological and conceptual issues." In Curdt-Christiansen, X. and Weninger, C. (eds.) *Language, ideology and education: The politics of textbooks in language education*. Abingdon: Routledge. pp. 50–66. Available at: https://www.taylorfrancis.com/books/e/9781315814223/chapters/10.4324/9781315814223-11 (Accessed: 27 May 2020).

White, C., Easton, P. and Anderson, C. (2000) Students' Perceived Value of Video in a Multimedia Language Course. *Educational Media International*, 37 (3): 167–175.

Wilcox, K.C. (2010) The impact of student beliefs on the effectiveness of video in developing cross-cultural competence. *CALICO Journal*, 27 (1): 91–100.

Willingham, D.T., Huges, E.M. and Doboloyi, D.G. (2015) The Scientific Status of Learning Styles Theories. *Teaching of Psychology*, 42 (3): 266–271.

Willis, D. (1983) "The Potential and Limitations of Video." In McGovern, J. (ed.) *Video Applications in English Language Teaching*. Oxford, New York, Toronto, Sydney, Paris, Frankfurt: Pergamon Press. pp. 17–28.

Zappavigna, M. (2011) Ambient affiliation: A linguistic perspective on Twitter. *New Media & Society*, 13 (5): 788–806.

Zappavigna, M., Dwyer, P. and Martin, J.R. (2008) "Syndromes of meaning: exploring patterned coupling in a NSW Youth Justice Conference." In Mahboob, A. and Knight, N. (eds.) *Questioning Linguistics*. Cambridge: Cambridge Scholars Publishing. pp. 164–184.

Appendix A – Sample Questionnaire

Participant Number: _____

1. What were your general impressions of the videos?

   On a scale of 1 to 4, please rate with a ✓, with (1 = this does not apply at all, 4 = this completely applies to my experience of the videos). If there are any options you wish to add, please do so beside "Other:" below.

   I found Video 1:

Humorous	1	2	3	4
Annoying	1	2	3	4
Fun	1	2	3	4
Other:_____	1	2	3	4

   I found Video 2:

Humorous	1	2	3	4

Annoying	1	2	3	4
Fun	1	2	3	4
Other:_____	1	2	3	4

I found Video 3:

Humorous	1	2	3	4
Annoying	1	2	3	4
Fun	1	2	3	4
Other:_____	1	2	3	4

2. a. To what extent do you find the videos to be helpful with your language learning? On a scale of 1 to 4, please rate with a ✓ (with 1 = not at all helpful, 4 = extremely helpful).
   b. Please explain the reasons for your answers above.

Video 1

1	2	3	4

_____

_____

_____

Video 2

1	2	3	4

_____

_____

_____

Video 3

1	2	3	4

_____

_____

_____

3. a. To what extent do you find the people and situations in the videos to come across to you as people or situations you might encounter in Germany, Switzerland and Austria? On a scale of 1 to 4, please rate with a ✓ (with 1 = impossible/unrealistic, 4 = possible/realistic)?

1	2	3	4

b. Please explain **in one sentence** the reason for your answer above.

_____

4. a. What do you think the videos are trying to portray about German speakers in German-speaking countries?

_____

_____

_____

b. On a scale of 1 to 4, please rate with a ✓ to what extent do you agree or disagree with these portrayals (with 1 = strongly disagree, 4 = strongly agree)?

1	2	3	4

5. What did you find difficult to understand in these videos regarding:

a. language in the videos?

_____

_____

_____

b. the characters' actions or behaviours in the videos?

_____

_____

_____

6. On a scale of 1 to 4, please rate with a ✓ to what extent you would recommend these videos as a study aid? Please place ✓ in one of the boxes below (with 1 = would never recommend, 4 = would definitely recommend).

1	2	3	4

7. To what extent do you think the videos are being used as effectively as possible in the German Studies 1 course? On a scale of 1 to 4, please rate with a ✓ (with 1 = not at all, 4 = they are being used as effectively as possible)

1	2	3	4

THANK YOU FOR PARTICIPATING!

Appendix B – Questionnaire Answers by Students

1. What were your general impressions of the videos?

   On a scale of 1 to 4, please rate with a ✓, with (1 = this does not apply at all, 4 = this completely applies to my experience of the videos). If there are any options you wish to add, please do so beside "Other:" below.

   I found Video 1:

	P1	P3	P4	P6	P7	P8	P9	P10	P11	P12	P14
Humorous	3	2	3	3	3	4	3	3	1	3	3
Annoying	1	3	2	2	1	4	2	1	3	1	2
Fun	4	2	3	2	2	1	2	3	2	4	3
Other: Informative	4										
Other: Cringeworthy		3									
Other: Difficult to understand			2								

Other: useful				3							
Other: repetitive						4					

I found Video 2:

	P1	P3	P4	P6	P7	P8	P9	P10	P11	P12	P14
Humorous	4	3	4	3	2	2	4	3	3	4	3
Annoying	1	3	1	2	3	2	3	1	2	2	1
Fun	4	2	4	2	2	2	2	3	2	3	2
Other: Informative	4										
Other: Cringeworthy		3				4					
Other: Difficult to understand			2								
Other: useful				3							
Other: Difficult to hear					3						

I found Video 3:

	P1	P3	P4	P6	P7	P8	P9	P10	P11	P12	P14
Humorous	1	3	1	2	2	3	4	3	2	3	1
Annoying	1	2	1	2	2	3	2	1	1	2	1
Fun	4	3	3	2	3	1	3	2	2	4	2
Other: Informative	4										
Other: Difficult to understand			3								
Other: useful				2							
Other: uncomfortable						4					

2. a. To what extent do you find the videos to be helpful with your language learning? On a scale of 1 to 4, please rate with a ✓ (with 1 = not at all helpful, 4 = extremely helpful).

	P1	P3	P4	P6	P7	P8	P9	P10	P11	P12	P14
Video 1	4	3	4	4	3	3	3	4	3	4	3
Video 2	4	2	2	3	2	1	2	4	2	3	3
Video 3	4	2	3	2	3	2	3	2	3	4	3

a. Please explain the reasons for your answers above.

Video 1:

P1:

This is a challenging chapter of Menschen and the directions are repeated 3 times to ensure ample listening practice and to display the correct usage of these structures/expressions

P3:

It provide vocabulary for giving directions

P4:

Understanding directions in a foreign place is a very important skill to learn.

P6:

Even though the repetition is boring it is memorable and this helps with vocabulary learning.

P7:

Helpful as not only did it aid in listening but the directions were also physically shown in the end.

P8:

I was able to hear how to give someone directions.

P9:

The sentences were repeated many times, so it was easy to understand.

P10:

If I hear the people talking in the videos, I feel like I will be able to remember it more.

P11:

Good revision of directions and instructions.

P12:

I always have a problem with directions but after watching this video. I think it did help me understand directions in German.

P14:

Seeing and hearing them speak helps with pronunciation.

Video 2:

P1:

We are presented with a content in which to be critical and are thus encouraged to personalise the target language through opinion formation.

P3:

The visuals were not so good. Though they did use descriptive words that could help with vocabulary.

P4:

As a foreigner it's not that important to understand descriptions of an apartment, sight is the most important.

P6:

The video is helpful as it reinforces the adjectives learned in this unit, but the questions could have been designed better - to ask what the room was really like, for example.

P7:

Was nice to listen to however got distracted by the filthy house the woman lived in. So ended up not focussing on what was being said.

P8:

I was too distracted by how dirty the living place was.

P9:

Very simple language used, it could have used better or more descriptive words.

P10:

I did hear a lot phrases and that phrases that I heard before but forget about it, so it did revise some language.

P11:

The visuals were a bit too distracting and took away from the language element.

P12:

Some of it was confusing because I have a hard time picking up a new language but it was good.

P14:

great for helping with dictation.

Video 3:

P1:

Another great demonstration of the target language. I also enjoyed the introduction to Swiss-German. It is very different to the standard German we have been learning.

P3:

It was okay. Provided new vocabulary.

P4:

It is a useful skill to understand discriptions of a city you are wanting to live/visit in.

P6:

At this level, I find the introduction of different accents and dialects more confusing than anything else.

P7:

Was a fun video, was easy to follow as he spoke clearly and pictures being shown also aided in understanding

P8:

I didn't really see what they were trying to teach us in that video.

P9:

Didn't think there was anything to learn while watching the video clip.

P10:

I can't really remember anything from Video 3. I just looked at the answers I had than I can remember a bit.

P11:

Exposure to a new form of the language entirely, words we've learned tied together with visuals, good pace and new words.

P12:

Helpful. I got to learn how to pronounce some words and learn new vocabulary.

P14:

good to get an idea of what different German speaking countries are like

3. a. To what extent do you find the people and situations in the videos to come across to you as people or situations you might encounter in Germany, Switzerland and Austria? On a scale of 1 to 4, please rate with a ✓ (with 1 = impossible/unrealistic, 4 = possible/realistic)?

P1	P3	P4	P6	P7	P8	P9	P10	P11	P12	P14
4	3	1	4	3	4	3	3	2	3	4

b. Please explain **in one sentence** the reason for your answer above.

P1: **All three target language groups are used/beneficial in my mothertongue/motherland.**

P3: **Tourists ask for directions anywhere, people can be unpredictable.**

P4: **I think that their quality of acting was so terrible that it is hard to believe that they are even real people.**

P6: **Even though it is clearly staged, I feel like it still helps me to hear the language in use.**

P7: **Sometimes the situation is believable but mostly absurd.**

P8: **The first two videos I feel as though they are more likely to happen because directions are something you'll always need in a foreing country and renting accomodation is also pretty necessary. The third video 1 found unlikely to happen.**

P9: **They are realistic scenarios, but they are portrayed unrealistically.**

P10: **If you gonna go to these countries, you will always ask for direction.**

P11: **The situations are realistic but the interactions are very staged, as are the personalities.**

P12: **I am not so sure if I will find them like that.**

P14: **relevant everyday things which a person is likely to experience when living or visiting these countries.**

4. a. What do you think the videos are trying to portray about German speakers in German-speaking countries?

P1: **I was oblivious to any 'portrayals.' I felt the videos engaged with amusing contexts to encourage observer interest.**

P3: **That they are all super friendly and polite**

P4: **The videos are trying to portray that Germans are very friendly people.**

P6: **These are not particularly flattering but I feel like they are trying to give a sense at their variety and humanness.**

P7: **That they enjoy teaching those who are not educated about their language and country.**

P8: **That they are "funny"**

P9: **I have no idea. That they are very friendly?**

P10: **I think, how they talk to each other.**

P11: **That they're helpful, patient and welcoming/hospitable.**

P12: **They are trying to portray that they are very different in the way they speak the language.**

P14: **That they're just like everyone else.**

b. On a scale of 1 to 4, please rate with a ✓ to what extent do you agree or disagree with these portrayals (with 1 = strongly disagree, 4 = strongly agree)?

P1	P3	P4	P6	P7	P8	P9	P10	P11	P12	P14
N/A	2	4	2	3	3	3	3	2	3	3

5. What did you find difficult to understand in these videos regarding:

a. language in the videos?

P1: **All language usage seemed aptly graded. Only the Swiss-German was incomprehensible, but the writer was aware of this and swiftly clarified.**

P3: **The last video was not proper German to my knowledge**

P4: **I find that alot of German speakers speak very quickly and they pronounce words**

differently depending on their location.

P6: **The accents in the 3rd video made it unclear. Otherwise it was appropriate to an A1 level.**

P7: **Struggled mostly to hear what was being said. The vocabulary being used was not**

difficult.

P8: **The Swiss German**

P9: **A few words that I did not understand, otherwise the language is very understandable.**

P10: **How fast they talk in the videos and how they pronounce some words.**

P11: **In some parts of the videos the dialogue was very rapid and I couldn't quite make**

some of it out.

P12: **I was not able to pick up and translate the other words and other phrases.**

P14: **Some of the characters spoke a bit fast.**

b. the characters' actions or behaviours in the videos?

P1: **The only questionable actions were used for dramatic effect and did not impede comprehension.**

P3: **They were over played and the actors tried too hard to portray their characters**

P4: **The characters behaviour were actually easy to understand. However, Frau Möllemann was peculiar.**

P6: **Inexplicable, but what can you expect from failed actors, badly paid, who have to portray unrealistic scenes. Their behaviour is not the focus.**

P7: **Their actions were strange and not how one would normally act.**

P8: **Smiling into the camera**

P9: **Many of them seem very fake or exaggeraated, but understandable.**

P10: **How they react to certain situations.**

P11: **I wonder if the interactions would ever actually happen - is this really how average**

people talk to and behave around each other?

P12: **Their actions were good as I was able to understand them better.**

P14: **N/A**

6. On a scale of 1 to 4, please rate with a ✓ to what extent you would recommend these videos as a study aid? Please place ✓ in one of the boxes below (with 1 = would never recommend, 4 = would definitely recommend).

P1	P3	P4	P6	P7	P8	P9	P10	P11	P12	P14
4	2	3	3	3	2	2	2	3	3	3

7. To what extent do you think the videos are being used as effectively as possible in the German Studies 1 course? On a scale of 1 to 4, please rate with a ✓ (with 1 = not at all, 4 = they are being used as effectively as possible)

P1	P3	P4	P6	P7	P8	P9	P10	P11	P12	P14
2	3	2	2	3	4	4	3	2	3	2

Appendix C – Interview Schedule

Interview Questions for discussion

1. When watching the videos and filling in the questionnaires, did any of you notice anything interesting about them that you want to discuss?
2. What were your overall general impressions of the videos?
3. Are there any videos that you really enjoyed and if so, what did you like about them?
4. Are there any videos that you really did not enjoy and if so, what did you not like about them?
5. What did you learn from these videos (starting from Video 1 up until and including Video 3?)
6. From your personal experience, what are the advantages of language learning from these videos, especially in comparison to learning from books?
7. What are the disadvantages of language learning from these videos from your personal experience, especially in comparison to learning from books?
8. What do you think these videos are useful for when learning for German Studies?
9. You all rated on a scale of 1 to 4 whether you agreed/disagreed with the portrayals of German speakers and German-speaking countries in these videos. Why did you choose the rating you did?
10. You all rated on a scale of 1 to 4 whether you would/would not recommend these videos as a study aid. Why did you choose the rating you did?
11. To what extent would you say the interactions and the depictions of the German-speaking countries in these videos match your idea of what Germany, Austria and Switzerland are like?
12. Regarding language in the videos, what did you find difficult to understand?
13. Regarding the characters' actions or behaviours in the videos, what did you find difficult to understand?
14. Do you have any ideas on how we can better use these videos in a lecture/tutorial setting?

Appendix D - Transcription Group A

I: Alrighty guys, so, we have just filled in our forms and what I want to ask you was when watching the videos and filling in the questionnaires, did any of you notice anything interesting about the videos or questionnaires that you want to discuss before we like ask specific questions?

P4: Can I ask an opinion an opinion of yours? Do you think videos are being used effectively?
I: Well this is what this study is trying to find out so you tell me P4, do you think these videos are being used effectively?
P4: I don't think so, I think we need more basic level German videos for like for GS1
I: Ok
P4: I find like for myself the level of these videos are like too advanced for me personally
I: Why? What do you think about the videos makes it too advanced?
P4: Like the speed at which they talk, it's quite rapid.
I: How many of you agree with this? I see P7 is nodding sorry and I'm saying this out loud because
P7: I don't find um its specific videos depending 'cause the first and second one were fine but where they were in conversation, I was like what is this lady saying, 'cause you speaking hella fast. 'Cause I'm still trying to translate what you saying into English to actually understand what you're saying so ja I agree with that
I: Right so are there any other different opinions were you guys okay with the speed or
P6: I disagree
I: Okay, P6
P6: Although I did find when they do other accents, um, that's a little over my level. Like the Swiss accent. Is it really necessary right now, maybe later. Try to figure that out but right now can I just try and understand high the high German.
I: No that makes sense. Ja ja. And what do you think P1?
P1: Um well with these videos it's always a real toss up because they want to try and give you the experience of of like German but also they they need to heavily grade their language and that balance is often very hard to find and I think they've done pretty well I mean they've provided us with lots of visual contextual support and ja
I: Can you give me examples with you think the visual and contextual support?
P1: Well, for example, we're not watching videos related to any like everything we've done, they always have a group of target language, for example, directions and we know okay in this video they're gonna be talking about directions. We have a fixed number of words and expressions we can learn to help us. And then we get like a great demonstration, especially in that video like three times on how to use them effectively
P3: Arrows

laughter

I: With the arrows? Okay ja. Ok so like when it comes to the like you said the the directions for example lets let's take that. So how do you think that kind of thing helps your language learning? By like you say, putting it visually and contextually like that?
P1: Well essentially these are listening activities so like the primary well it depends on the task given as well I mean we could be listening for just or as you know for like specific details so it it depends on that first of all, so task-related, and ja I don't know,

listening is essentially to help you with your pronunciation and your comprehension and...

P6: Memory and repeating it three times and with the visual aid. Especially if you're a visual learner it's definitely gonna help determine with the vocabulary.

P1: It's also nice to take a break from the lecturers

laughter

P6: Hearing from Ja ja someone else

I: Ja ja, hear other voices. So that's really interesting. Um cool guys. Are there any. Do you guys agree? P3 also do you guys agree with what P1 said? Disagree? What do you guys think?

P3: I kinda disagree with everything because I find them very hard to pay attention to especially when they try to push what they're talking about in a really exaggerated way which just makes it feel cringeworthy and then eventually I get put off to the point where my body physically feels ill

I: *groans* okay I'm sorry like you can leave if you're uncomfortable right? You know this.

P3: I know but I love judging stuff, its fine

I: Okay great. Okay great. So, you're you're enjoying the uncomfortable? I remember you were saying that just the sound of the Menschen videos…

P3: It gives like 'cause there's been a couple of times in class when we watching a video and just watching it made me physically ill

I: Oh my gosh that's I don't know how I feel about this like *groans* okay well as long as you're okay to be here, like mentally, its fine. Also, for um but ja what I wanted to say was why why do you think you get that feeling?

P3: I don't know because the videos make it feel very unnatural

I: Okay. So, give me an example from one of the videos. What what what is unnatural for you? Just something anything from the three.

P7 had hand up

I: P7 do you wanna?

P7 shakes head

I: Okay

P3: The third video when they went to the lady's apartment

I: Mhm

P3: and she's stressing on the fact that her place is gorgeous, absolutely gorgeous and standing in front of the camera like she had to like show that she had pride of how gorgeous her place was and clearly the person in the video was not agreeing, that just made me feel really uncomfortable.

I: Mhm so like was it because there was like dissonance or disagreement between the two?

P3: No just the over exaggeration of like the way the characters are and the over exaggeration how they're portraying them

I: Mhm so you think it's the actors really that are making you feel uncomfortable?

P3: Yes, and a little bit of the content

I: Oh really? Like? Can you give me an example?

P3: Um I can't one of the videos we watched in class

I: Mhm

P3: Where they're on an elevator and like all of a sudden, a guy just starts flirting with the girl

I: Wow

P3: While they're stuck on the elevator

I: Was that one of the Menschen videos? Oh, wow okay I'll find out which one it was

P6: The problems at the hotel *Lektion* (Chapter)
I: Ohhh. And so, like this this guy just starts randomly flirting?
P3: And their reactions are like oooohhh okaaaaay *laughs* and it was just really uncomfortable *laughs*
I: Because of the over exaggeration?
P3: Yes
I: Ok wow. OK Do you guys feel the same do you guys get a little level of discomfort?
P7: I think personally if I watched someone that sort of being realistic about it as if I'm also in the situation I'd be like okay I'm gonna listen to you but if you gonna be weird and I'm not gonna start going to flirt with a random person standing next to me that's just but ja but that's my opinion
I: Do you think this is something that would happen in real life?
P6 & P3: No
I: No?
P6: But this doesn't have to be. We're not at a level when we can really operate in German in real life yet
P4: Mhm
I: Interesting
P6: So It can't really give us real life situations
P3: Also when you're stuck in an elevator you don't think about flirting with the person you're stuck with you think about getting out
P6: Jaaa but
P7: Ja I agree with that as well. There's not enough content or
P6: Yeah
P7: To actually
P6: It's got to be appropriate to the level
I: P4 what do you think? You're very quiet
P4: I think I agree with what they're saying I think maybe the like extra dramatisation is to like help to aid us in learning 'cause we don't 'cause we can't function in like a normal day to day conversation we don't know what's going on so the gestures are like to aid us
I: Okay, like can you give me example, like of a of something that aids?
P4: Like the the apartment scene, like the showing of like gesturing to the dishes or the view or just the visual like aids as in like the everyday items
P7: Ja and the last video where he was talking about like the clock thing.
P6: Yeah
P7: Thank you for showing me the clock 'cause I well you know. I wouldn't have known what you were talking about
P3: See stuff like that is helpful
All: Mhm
I: Okay so clearly that helps, the visual aids you say, the gestures you say the. What else?
P1: Another beneficial thing I think from this overdramatisation is that you form an opinion about it which really helps you personalise whatever target language there is and regardless of whether you like it or hate it or are okay with it at least you are forming these personal opinions and engaging with the language
P3: And you will remember it.
P1: Yeah yeah
P3: I mean you don't like it is bad but you're remembering it nonetheless
You will still remember it
P7: That's actually kind of clever

P6: Yeah

I: And ja exactly. Do you guys then think that like these are what do you think these videos are useful for? We've discussed some things so like P1 was talking about the language bits the like so what do you it's doing with this overdramatisation, the gestures? I think P4 was more like reinforcing? Language? Kay, P4's nodding. Sorry, guys I have to do I have to vocalise your gestures.

laughter

P4: It like provokes like conversation and a debate because its memorable and overdramatised, it like it intrigues you. So you almost want to learn what's going on. Like a soapie, it's like a soapie

P1: A bad idols audition

P7: It's a terrible soapie

I: Ja ja So probably like those corny ones right? Okay cool. What would you say the advantages of are language learning from these videos are, especially in comparison to learning from books?

P3: It shows uh more of a conversational setting. So we're not just learning sentences at a time. Helps to show how you can like string those sentences together and hold a conversation with people

I: Cool. P6?

P6: It helps with different types of learners as well. Some people learn by reading, some by seeing so that's and some by hearing and the video is all of those things which is really great

P7: Ja I'm a visual learner so when I see like a picture I'm like kay cool and hearing it also helps with pronouncing and ...yeah definitely

P6: It's good for pronunciation apart from when they speak Swiss German

laughter

I: So would the Swiss German be a disadvantage?

P3: Mhm

P3: It could be confusing

P6: It's just for this level it's not appropriate

P7: It's just adding an extra level of spice I don't need right now

laughter

P6: And it's like if you want to do Cultural Studies I suppose it's fine

P1: It's cool

P6: But don't try learn language from it

I: So any disadvantage of your language learning in comparison to learning from books? From the videos?

P6: I don't think you could have one without the other

P4: Mhm

I: So you say then not

P6: They're not separable

I: Separable?

P4: Mhm

I: Kay. So in other words... disadvantages, but like from the language learning of these videos in terms of books? Are there any?

P6: We don't have a record of it to learn from

I: You don't have a record?

P6: I suppose you have everything written down, your grammar, your vocab. The video you just watch once.

I: Which is interesting because just FYI there are actual transcriptions of these videos available

P6: Ja
P4: Mhm?
I: Mhm
P3: I think the two when they put the books and videos together it's really helpful because like with the exercise we just did, we weren't necessarily like watching the video, we were looking at our exercise and listening
P7: And hearing where's the question coming in so I can give my answer
I: Ja that makes sense. Okay guys some few questions I really need to ask you before we wrap up. You've all rated on a scale of 1 to 4 that's why I've given these back to you whether you would or would not recommend these videos as a study aid. Why did you choose the rating you did? So what are they and why did you choose them? Do you need yours back P3?
P3: I believe mine was two or three
I: Ja two two I believe
P3: Yeah so I personally I wouldn't recommend them
P6: Ja I rated 1 more
P3: Personally 'cause of bias it makes me feel uncomfortable so I don't know how it's gonna make them feel especially if they have similar feelings to me
I: Okay would anybody else? Does anyone have different or similar feelings to P3?
P6: I gave a 3 though I mean I would recommend the video like that you watch the videos I'd always recommend go for as much German exposure as you can I mean,
I: Mhm
P6: they're given to you watch them use them why not?
I: Cool okay.
P1: Just don't like come in with some weird expectations about how you're going to learn oh about German culture uh anything else really from a video targeted at an elementary level student you know it's we're dealing with a very closed set of things and its
P6: It's got to be artificial
P1: Yeah
I: Ja that makes sense
P6: Manage your expectations
I: Mhm that makes sense. So P1 what did you rate?
P1: I'm happy to recommend these as a learning aid to anyone. No one said like it can't be your only German only aid but as an aid
P6: Please don't only watch videos for exams
P1: Absolutely yeah
I: Ja okay cool. And P7?
P7: Ja I said 3 as well I think it just aids you in slowly building up what German is actually all about
I: Okay
P6: And also help you remember
P4: I also said a 3 but I would recommend it as a secondary
P6: Yeah
P4: Form of study. I think you have to know a lot more of the basics to actually
P7: Ja
P4: Interpret what's going on
I: Makes sense. Ja that makes sense. And guys also you rated to what extent would you say the interactions and the depictions of the German-speaking countries in these videos match your idea of what Germany, Austria and Switzerland are like?

P3: Again with the Overdramatisation. I don't think that people just walk around discussing everything around them, 'cause if you go to visit a country like that, they're not gonna randomly take you aside and be like hey look at this and this and this. You seeee? *raised tone* This is this how many years old and this is the explanation of that. They'll just be like hey get a tour guide.
I: Mhm
P3: Like that's what you came here for
P6: You might ask someone for directions
P3: Yeah like directions happen anywhere
P6: It's not unrealistic. Definitely
P3: The only unrealistic thing is that lady did not say thank you
P6: Yeah
P4: It's true
P6: And that the guy escorted her.
I: Ja oh really so what do you think?
P6: That may or may not happen. It depends I guess
I: Mhm. So like do you guys think, do you think it's realistic or unrealistic?
P4: Unrealistic
I: Or in the middle?
P1: I don't know I just consider each of these situations and how I'm bound to use them in my mother-tongue in my mother country every now and then when I'm looking for an apartment or moving or if I need directions, which I very often do, or if I'm planning to go to visit some place I'm reading up about that place and there's bound to be some information about you know why that place is great or not great. So ja all of these are definitely used in my English so I don't see how it wouldn't be relevant in Germany
I: In Germany? Okay cool. Um anything else to contribute on that bit before we...I've got two more questions for you guys and then I'll leave you alone. Um what I wanted to ask. Do you have any ideas on how we can better use these videos in your lectures or tutorials?
P6: Yes
I: P6!

laughter

I: Go for it. The floor is yours
P6: Um I do feel like they are underused. Not uhm not that we don't do them often enough I mean in the books the way that it's designed, you watch a video and you have one task. I would rather have one video with three tasks and really get into the language. So I'd rather just have the first one with the directions maybe okay slightly well it was fine as it was but maybe with slightly fewer cues like the arrows and things they were a little bit like obvious and or possibly play it the first time just the first part when he gives the directions without visual cues and then do a task, like draw the route or something. And then watch it again and fill in the missing words like more tasks with one video so that you're not just looking at the language one to five you're really getting into it
P7: Yeah I think you're gonna get more out of it when you do it like that
P6: Ja so much more out of it
I: Mhm that's interesting
P6: So each video you should have a task for just and a task for close listening
I: Kay cool

P1: Also it would be nice to like set them up with some contextualisation like for example before we watch the direction video you know like have you ever been lost? Discuss it with your partner I think like the communicative approach is brilliant
P3: Are you bad at directions?
P1: Yeah yeah exactly it's really effective at
P6: Or like hey what happens or something?
I: Ja ja okay cool. Also guys going back to your tasks actually you reminded me of that um I actually want to steal one of these quickly just to see um. There's a question here möchte Herr Waurich die Wohnung mieten? Und Sie? Wie finden Sie die Wohnung? Do you guys understand what that question is?
P6: Yeah
I: Okay so would Mr Waurich rent the flat?
P3: Oh God no
P4: Nein
I: Why not?
P3: Because of the physical representation of the flat
I: Describe to me the visual representation of the flat
P3: Kitchen was messy, one cupboard was left open and things were just thrown into it, bed wasn't made there was stuff on the floor there was a weird stain on the floor that even the person with the camera was focusing on
P7: Weird ja
P3: I can't remember the other room. I think the lounge looked okay for a lounge
I: Okay so how do you guys then feel about also the description now P3's also given and what you saw of the flat? Would you how d- Do you like it? Don't you like it?
P1: No
P4: Personally wouldn't
P6: No of course not
P6: But theoretically she's gonna move out and take all of her dirty stuff with her
P1: The kitchen looks so old
P3: And hopefully everything will be clean
P6: Yeah
P7: Ja
P1: The kitchen was sucky
P6: Then you're gonna get a cleaning service and then you're yeah
P3: Then again the place looks really tiny
P1: It's really old
P6: Maybe she doesn't have much money

laughter

P1: I mean if that's like an apartment in Berlin I mean I'm sure it's pretty huge
P6: I think the question could be used better 'cause then you could have this question and say how would you describe the flat and you the student needs to provide different adjectives like um Der Flur ist nicht praktisch, es ist er ist what can I say klein und schmutzig *laughter*
I: Ja ja Also P7 and P4 'cause you've been quiet from this side here. What do you guys think about the interactions of like speaking the speaking and what's going on matching your ideas of Germany, Switzerland and Austria?
P4: I just I personally think that the production of the videos so like the actors they just they seemed unrealistic. So like they could literally be from anywhere, they were just actors to me they weren't Germans or Swiss people they were just they came across as just actors
I: Okay P7?

P7: I found it like it seemed as though if you are a foreigner anyone would be willing to help you figure out what's happening and stuff like that I was like okay cool I'd go to Germany people would help me. Ja
P3: People are so friendly there
P7: Ja
P1: Like flirting in elevators

laughter

I: Ok cool and also for you two 'cause you were quiet there as well do you have any ideas on how we can better use the videos in lectures or tutorials?
P7: I like what P6 said. P6 said it better than I would have said anyway so
I: And P4?
P4: If possible like I know that tutors have a strict curriculum to follow but I would like more discussion of the videos
I: Ok can you give me an example like is it ja give me an example
P4: Maybe just more explanation but because personally I don't know a lot of language and sometimes I miss out like the context, I don't actually understand what's happening so if we could like talk through the video, maybe discussions like a class discussion like this
I: Like this?
P4: Yes
I: So did you find this type of thing helpful?
All: Mhm
I: Okay that's interesting
P3: As much as I dislike the videos it was actually very helpful
I: Oh I'm so glad I'm so glad that makes me happy
P1: Just calming to talk about your dislikes of the videos
I: Venting *laughter* your venting session. Also my last question before we wrap up did any of you notice the writing in the second video? Did any of you manage to read it?
P1: What writing?
P6: Writing?
P3: What writing?
I: Kay that's all I wanted to ask.

laughter

I: Thank you very much guys.
P7: There was writing?

laughter

P1: What like all over the walls or
I: No no there was like a thing in the beginning
P3: Oh yes where he showed the
P4: Oohhh
P3: The little advert thing
P7: Where they
I: Did you guys
P1: I just saw the Superwohnung, that was all I saw

noises of agreement

P3: Why did he put it next to the intercom?
P4: Oh ja he looked on the intercom
P7: He was just like showing it there
P6: Ja it was weird
I: Thank you guys

Appendix E – Transcription Group B

I: So I'm going to set every on and we've all filled in our forms so tell me, guys, before we really dig into the questions. Is there anything interesting that you found in the videos um or in filling in the questionnaires that you want to discuss before start?
P11: I found it normally in the videos they speak quite slowly and then some of the videos they did, but then in other videos they spoke like really fast. It was like a weird
I: Which ones? so you talking about
P11: So like the first one was like really slow and like repeated a lot of stuff and then the second one was like, it had like really slow parts and then it had like that one really fast part where no one could actually hear what she was saying and then the third one was quite like was slow enough but it was a lot faster than the other ones.
I: Okay, cool. How do you guys feel about it, do you also agree with speed being...was it an issue? The speed?
P11: Not not an issue but like it was just interesting that they were so like varied
I: Okay, that's interesting. Do you guys also find a difference in speed?
P14: Yeah
P9: Yeah, but I also don't think it's an issue, it's just a change
I: And P12 do you find speed like what was the speed like? Was it okay for you?
P12: Okay the one where the speed was very fast I wasn't able to understand like some of the words spoken or phrases
I: Which video was it?
P12: Oh the second one
I: The second one okay cool. Ja like do you guys agree that the second one was the faster one? I think P11 said so?
P14: Ja the faster one
P11: Ja that one just that one part with the one question like none of us could answer
I: Ja that prima einkaufen bit? Ja that was quite ja and like I said that "Alle Läden gleich um die Ecke". Ja was quite challenging. Okay cool and then tell me guys, um from your personal experience, what are the advantages of language learning from these videos especially in comparison to learning from books?
P9: Well, I mean, obviously you get to hear like the words actually being spoken, when you're reading a book you only have like you know your head voice
I: Mhm ja *laughter* your head voice
P9: I mean I don't know how else to say it
P8: You pronounce it the way
P9: Ja you pronounce how like you think its
P8: Its pronounced
P9: Sounds but here you actually get to like
P8: Get to hear German speakers
P9: Hear how it's spoken and the pronunciation and ja
I: Okay, so ja so you get to hear your German speakers like P8 said you get like the voices and the pronunciation. P7 is there anything you want to add? 'Cause I heard you like and also from the two of them like pronunciation. Did you also agree with and?
P7: Yeah and it also helps when I think the lecturers or someone even in the exam when they talk about the word then they can always refer back to it back to what they were saying on the video 'cause I think it's more easy to remember things by hearing it
I: Okay, cool. So you also you like audio it helps you remember things?
P7: Yeah

I: Okay, cool awesome. And then do you think what are the disadvantages of language learning from these videos from your personal experiences, especially in comparison to learning from books?
P11: I don't know how realistic these interactions are like is this something that you would literally like see
P8: Yeah
P11: Hap- Like it's just its very like staged which obviously but i feel like would this ever actually happen? A lot of it.
I: Why do you think its staged like this?
P11: Well, obviously 'cause we're beginners, so we kind of have to has to get like laid out for us and
I: Mhm
P11: You know in a very simple way
P8: That's what I also don't like
P11: Yeah
P8: It just feels really like dumb-downed in a sense and it's like if you were to go to Germany that's not the interactions you would have. It's just like misleading
I: Mhm
P9: I feel like the scenarios are actually quite realistic they're just very exaggerated
I: Mhm and why do you think they're exaggerated?
P9: No like with the directions I think that's very a realistic scenario but the fact that they repeated each sentence at least three times is pretty exaggerated
I: And why do you think they're doing that though?
P9: It's obviously for us so we can like actually hear
I: Mhm
P9: And like get to understand the directions like they also said and also showed where to go with the words
P8: And with the arrows
P9: And for us to understand. So I guess it does need to be exaggerated so we fully understand the interaction happening
I: Mhm
P9: So we can link like you know the phrases with the actions
I: Mhm that makes sense and also P8 you said did you say arrows?
P8: And also the arrows in the videos as well.
P9: Yeah yeah
P8: Like the situations are realistic, some of them, it's just the way the videos are
I: So can you give me some examples of the real situations in here?
P8: Asking for directions, that's a real situation,
P14: Going to view an apartment for rent
I: Yeah, yeah for sure. P12, what do you think?
P12: It is like some kind of you know not so like she said not so realistic or what but I'm wondering like if people in Germany really act like that or so it's confusing yes
I: So like what do you think this is telling you guys about like, to what extent would you say the interactions and depictions of the German-speaking countries in these videos then match your idea of what Germany is like? Do you guys ja? Does what you saw match your idea of Germany and Germans? P9?
P9: Well I've been to Germany and I'm not saying the society is the same as here but people are pretty much the same. Like there isn't any like vast difference like they're not completely like in the zone like everyone, but they're not like crazily out of their way friendly they're just like people
I: Mhm So okay so they're just like people, they don't go out of their way or?

P9: No not all of them are like. So the way like for me the videos seem to portray the Germans they go out of their way to be like oh we're friendly and like you know smiley and super happy and I'm sure some people in Germany are but like some of them not all of them. I'm sure there are some who are like people go
I: You think so?
P8: Well I it felt like they were trying to portray Germans as really funny people
I: OOohhh
P8: So like in the videos they were always trying to like crack jokes and like smile into the camera and laugh and that was uncomfortable
I: Uncomfortable?

laughter

I: Did any of you guys feel like discomfort like that? From watching?
P9: Ja
P11: Ja I really think that the exaggerated nature of the videos takes away from the message.
P8: Ja
P11: Like I feel like I watched them and I can't even concentrate on what's going on
P8: Mhm
P11: 'Cause this is just so cringey
P9: Ja
P11: Like if they made its just so like I agree that the situations are realistic and that the interactions were probably happened they're just very exaggerated like P9 said but I just feel like
I: Mhm
P9: Thank you
P11: They were just too exaggerated you know like maybe if they were more natural with like subtitles or something?
I: *laughter*
P11: It would make more sense but it's just very like like they're helpful but also I just can't help watching them and being like bleurgh.
P8: Yeah
P11: I just want to turn away
P9: I think we just need to go on a trip to in third year just video some Germans and use that in just video
P9: In third year just send us off to Germany
I: *laughter* replace it with authentic stuff. So tell me P11, you said that the exaggeration was helpful? How?
P11: No the videos were helpful
I: The videos were helpful
P11: The exaggeration is what takes away from it
I: Okay so you think that takes away from it. P9?
P9: I think to a certain degree 'cause as P11 said, we are still learning, so I feel like if they didn't exaggerate a little bit we'd be a bit lost with you know what it means so I feel like the exaggeration in just helps us understand what they're doing, but I agree in some cases it does feel a little too exaggerated and like a bit forced
I: P12 you wanted to say?
P12: Well it makes me sceptical in some way
I: Yeah?
P12: Because now okay you're in German class, you're watching those videos then you happen to go to Germany and then how will you not expect that to happen? Like you

will expect that to happen. But maybe you won't find it like that, so that makes me sceptical.

I: It's really interesting so that it brings up sceptical and like do you think you can accurately sort of judge like P11 also hasn't been to Germany, so the both of you? How many of you? I know P7 has been. P8 no? P14 no. Do you guys think that you can accurately judge in this video what could happen in Germany and what isn't a thing?

P8: No

P9: No

P8: I don't think anyone will act like that

P14: Ja you know as well but I mean that's not what we're gonna get when we go like realistically that doesn't happen

I: So it's still interesting that you guys can pick up what's realistic and what's not even if you haven't been, but ja some of it can be kind of confusing and misleading like P12 says. Mhm interesting guys you're giving me some thought here. Um I want to also ask you guys what do you think these videos are useful for when learning for German?

P11: Think like mapping knowledge of something to what it actually looks like so like we always learn these things in like theory and seeing it like they're walking around and this is this thing when you think of the word, this is what you should think of kinda thing

I: Okay

P11: Like I think that helps remember what it is and how to use it

I: Mhm. P12?

P12: They also help in understanding that um Germans like they're different like the Swiss and the German so they're different not like the same, so you'll be able to maybe if you understand the language you'll be able to distinguish between the Swiss and the German

I: Mhm. Good no that's great so that's interesting so we even have some Cultural Studies. P7 you have been very quiet. I want to ask you, you've been to Germany, so how does this sort of match to your experiences?

P7: Um so the videos we've watched, the one with the directions I have encountered that even three times

I: *laughter*

P7: But the ones that we watched with Natasha then in class, not really.

I: Can you give me an example like of which video that you watched with Natasha?

P7: Oh my gosh that's

P9: The one in the guitar store.

P11: Yeah

laughter

P9: That was the worst one we watched

I: Oh she played that for you?

P11: Oh no that was the worst thing

P7: That's a good example

P12: It was like saying ich brauche eine Rechnung

I: Ja ich brauche eine Rechnung

P7: *laughter*

I: And ja P7 like what were you saying what did you did you think it was horrible or?

P7: Ja oh my gosh *laughter*

P9: 'Cause of the singing

P11: *laughing*

P14: Oh my gosh ja.
I: Okay that's really interesting that's really ja
P14: Repressed memories *laughter*
P11: Like if I had to go to a shop and actually need like one of those things I would not be able to walk in there and do it because all I would think of is the song and they're not going to respond to that
I: So I really like you I hear though that you guys
P9: Imagine walking down to Pick 'n Pay singing that
I: *laughs* you guys remembered it though?
P8: Yeah
P9: We remember the really cringey ones.
P7: Yeah
P9: I don't remember the exact words
P11: I can remember they couldn't really sing they were just like saying words rhythmically. Saying words rhythmically?
I: Oh wow okay cool. Okay wow important ones guys 'cause you all rated on your questionnaires on a scale of 1 to 4 this is question 6 whether you would or would not recommend these videos as a study aid. Why did you choose the ratings that you guys did?
P9: Well I guess different things work for different people but for me I gave it a 2 because for me studying is the best technique of studying is going like writing over words and saying it myself and these videos would just be distracting
P8: Mhm
I: How would they distract you?
P9: In terms because it's so exaggerated that I'd be so focused on what's happening and not what I'm hearing so I wouldn't really take anything in
P8: Mhm I also think so too. Yeah the situations I don't think they're really necessary I mean you can just listen to the audios and that would be much more helpful I think
P7: I gave it a three I know with me its more if I can hear then I will I will be able to remember it more
I: Okay ja you said you prefer audio kay ja cool
P14: I also gave it a 3 while some people might find it better to write and hear the audio others don't people learn differently so I guess it's also just to accommodate those who prefer visual aids or learn better with actually seeing it happen
I: Mhm
P7: Yeah and we see the stuff visually in the video so ja
I: Cool P12?
P12: I gave it a 3. They're not bad, still they're not good.
I: Ja
P12: It's more like um how a person learns. Some people learn by watching
I: Mhm
P12: And hearing what they're watching so some people maybe go and do that and some not
I: So you say that the reason that its good and that its bad is because it depends on the person learning from it or?
P12: Yeah like they may have different learning styles
I: Mhm
P12: So auditory, like, okay and others ja
I: Ja ja. No that makes perfect sense. Kay cool. P11, what did you rate it?
P11: I said 3 as well. Um it I I really like they're not they're not bad they're just, they're not the best medium for learning but I feel like they really do like I said earlier a map

to like visuals and I feel like there's a lot of like small cultural things as well like how people like you can tell with accents for example, like who's a native speaker and who's not.

I: Oooh can you give me an example of?

P11: I can't remember which video this was but there was one we video we watched in class like the person who was doing the videos accents sounded kinda like odd compared to the others and I think Natasha was like you guys can tell that this is not a

P8: A native speaker

P9: Like they were trying to sound Spanish or something

P11: And then she was like you can tell that this isn't a native speaker and I was like wow actually

I: Mhm

P11: And you can like see we watched this one like dancing dance video and it wasn't like was that a Menschen video? No it wasn't, never mind

I: Ja. P9?

P9: I actually I think the videos were pretty good like when you first learned it like when you do it in class and you first learn a section and you see the video? I feel like that's pretty helpful then just to associate the words with the object but when you're studying later on I feel like it's not helping not helpful.

I: Kay. Cool. Thank you guys. So what did you guys find difficult to understand, I know you guys talked a little about Video 2, so regarding language in the videos, what did you find difficult to understand?

P9: Well for me I thought the language and vocab use was pretty like easy to understand. I feel like only time, well I can't speak for everyone, but the only time people get confused is when they speak a bit fast and they can't hear the words but I feel like all the language use is very simple and in some videos I feel like it could actually be a little more um not complicated a little bit more what's the word um challenging

P11: Yeah

I: Yeah cool P12 what do you think?

P12: Well since P9 said in some others they were fast so I was not able to pick up like the the words and the other phrases to translate quickly as the video was going like you're stuck on one word that was spoken and then you're busy trying to translate it and then when you come back to the video it's just way over that it's just its moving on to others things. So ja.

I: So, you found it very fast. Was it just Video 2 that was the fastest or?

P12: Yeah it was Video 2.

I: Okay cool. P14?

P14: Ja especially when you're like listening out to or listening to answer questions as well in Video 2 when you're sort of looking and trying to figure out certain words and focusing on that then you miss it.

I: Ja

P14: Ja 'cause it's something completely different you know?

I: So, do you miss out? Do you feel on what's happening or? The words or the whole phrase 'cause you're concentrating or?

P14: Ja

I: Okay. Mhm.

P8: And also, sometimes I can't hear them because of the background noises like when finding the answer

P7: *laughter*

I: That's interesting. Oh my word. P7, what do you? What do you think? Was there anything difficult for you?

P7: For me it was also the fast talking like that sentence there and like how you were saying um so the people are talking fast and you're still trying to hear but you have to listen to when they're asking a question and then you're trying to listen they always just go so fast and then there's just another one and then you're just getting lost so ja

I: Mhm thank you guys. Do you have any ideas on how we can better use these videos in a lecture or tutorial setting?

laughter after silence

I: Ja like crickets guys. Any ideas on how we can use these videos? P12 anything?

P12: I have no idea

I: No idea?

P9: I think they're being used pretty well at the moment like I said its best when having them when you first learn a section that's when we're seeing them so well most of them so yeah

P8: Mhm

I: P8 you wanted to say something?

P8: Same

I: Same. P8. Also, you said. Do you prefer it before the sections starts rather than after the section starts? Why do you think?

P8: I feel like after 'cause I mean if we go into the video without knowing any of the vocab I mean we can put together some action and the word but I feel like it's better after we have the base vocab and the *Bildlexicon* (visual encyclopaedia) and

I: Mhm That makes sense. Do you guys agree? disagree?

P14: To what? *laughter* I don't know for me watching it first would be better.

P7: Yeah

P9: Before learning any of the vocab?

P14: Ja before you learn the vocab and then it gives you an idea of what it is

P7: Natasha used to teach the vocab with us sometimes and then she then when it was the video, I found it really helpful

I: Ah okay so that's great so then like you feel like maybe before the video starts even if you do do it in the beginning, you want to go through the vocab and the *Bildlexicon* (visual encyclopaedia) and then do it?

P7: Yeah so you can see the stuff what's going on ja

I: So also, like are you okay with how you know the interaction with the video is? So, like you know the tasks and?

P9: Yeah.

I: It's all good?

P14: Yeah

P11: Yeah I agree with that thing about like when there's a question about it you only focus on finding the answer and miss out everything else and like sometimes you even miss out on the answer because I don't know, it's just again the speed is weird because sometimes with the answer they speed through that and then take their time with random information just feel like its weirdly balanced like having a written task while you're watching the video is very jarring whereas if you had to watch it and then someone be like okay what did you notice about such and such, you could do it like orally and then write it down like afterwards

I: So, you don't like doing the tasks the written tasks while watching?

P11: Ja it kind of defeats the point

P9: Unless you like watch it once and then no writing and then watch it again but then that takes time obviously so

P8: Ja
I: Some good suggestions hey guys. Is there anything else? I have one more thing to talk to you about. Is there anything else you want to add to this question? Do you have any ideas for the lectures and tuts? No? okay. My last thing, back to your tasks, there's that question. Oh, I actually have two last things.
P14: *laughter*
I: Möchte Herr Waurich die Wohnung mieten? Do you guys know what that means?
P9: Something rent?
P11: Ja
P12: Yes
P14: Rent ja
I: Would he?
P9: Oh, does he want...no he does not want
I: Why not?
P9: Because he ran out of there
P8: *laughter*
P9: He ran out of there and was like phew
P8: Yeah
I: Why does he not want to rent it?
P8: It's really dirty
P7: And untidy
I: How do you know it's really dirty?
P14: We saw
P7: We saw the dishes
P8: Yeah we saw
P7: In the first picture there was standing and then
P14: He walked in
P7: And was like *laughter*
P11: Yeah she was trying to like distract him like don't look at the huge mess that I made and it was clean before you got here
I: Okay cool und wie finden Sie die Wohnung?
P12: Yoh
P14: *laughter* Yoh
P9: I mean it's almost clean
P9: It's pretty decent
P11: I don't know it's kind of small
P8: I don't know ja
P9: And you like have like a low paying job you need somewhere to sleep
P11: Ja well I mean you don't know what the rent is
P8: The furniture doesn't even look that good
P14: *laughter* is he going to be staying with her, is he renting it out like for himself or?
P11: Does it come furnished? Like
P9: Well hopefully not
P8: Well I mean like the sinks and what not in
P12: The sink was not okay
P9: Well obviously the sinks
P8: Ja that's built in
P11: But like the bed like I can't believe we're paying
P9: Like I feel if it was clean
P7: Ja me too

P9: If it was clean it could be a pretty decent I'm not saying it's the best for one person who didn't have money to burn
P14: potential
P7: Its good for one person
P8: Who doesn't care?
P11: But how do you know that its cheap though?
P8: *laughter*
P9: I'm assuming 'cause it's like
P11: But you don't know that. It could be expensive. You don't know where the location is.
P8: And did you see that Blick?
I: *laughter*
P9: Please. Is res any better?
P11: Haha
P14: lol
P8: Ja. Honestly
I: Okay guys last thing did any of you notice the writing in the second video
All: Writing?
P12: Which writing?
P14: When he was going up to the?
P11: Oh
P9: Oh that piece of paper
P12: Ooohhh
P14: Ja
P9: It said super flat
P14: Superwohnung
P9: Like does that mean like super cool or super expensive or super big
I: Just a super flat
P9: Like I couldn't read the smaller writing, I could just see Superwohnung 'cause it was bigger the rest of the information was pretty small
P8: And anyone could just walk in, like security none
P9: You still had to press the button like
I: So what does that tell you about Germany P8 do you think? 'Cause you're talking about no security
P8: No
P11: No
P9: Maybe certain areas. We don't know.
P12: Maybe they're advertising. Like you advertise your apartment
P14: You expect
P12: Super apartment. Then so like it's more like more people would come and then it's just like that so not super like you've mentioned on the advert. Yeah it's more like selling it selling your apartment to the people and ja
I: Okay guys. Thank you very much. I think we are done.

Appendix F – Themes from Thematic Analysis

The numbers next to the listed themes indicate the number of student responses in which these themes were found.

Questionnaires

Question 2

Video 1

Directions as focus 7
Listening practice, which can assist in memory and pronunciation 5
Repetition, which is often boring but memorable 3
Vocabulary 2
Visual aids 2
Easy comprehension 2
Authenticity of language 1
Revision 1
Pronunciation 1

Video 2

Distracted by visuals 4
Vocabulary of description of flat 3
Listening 3
Descriptions of flat not important 1
Visuals are more important 1
Task could have been designed better 1
Personalised engagement with language through opinion formation 1
Too simple language used 1
More descriptive words needed 1
Revision 1
Difficulty comprehending 1

Video 3

Swiss-German introduction as theme 3
Enjoyment 3
New vocabulary 3
Visual aids 2
Confusing 1
No point to the video 2
Easy comprehension 2
Authenticity of language 1
Description of city 1
Revision 1
Good speech rate 1

Pronunciation 1

Question 3

Possible situations 7
Unrelatable people/terrible quality of acting 3
Beneficial 1 (helpful in terms of language used in context)
Mostly non relatable situations but sometimes possible/relatable 1
Video 1 and 2 relatable/possible 1
Video 3 unrelatable 1

Question 4

Very friendly 3
Amusing 2
Diversity 2
Humanness 2
Enjoy teaching about German culture/welcoming 2
Helpful 1
Polite 1
Patient 1
How they interact 1

Question 5a

Speech rate too fast 6
No difficulties 2
Pronunciation variation 2
Swiss-German was incomprehensible 3
Swiss-German was incomprehensible but clarified 1
Vocabulary 1

Question 5b

unrealistic actions/people often with terrible quality of acting 6
No difficulties 4
Overdramatic, especially Video 2 (Frau Möllemann) 2
Reactions 1
Unsure if realistic 1

Transcription A

Do you think these videos are being used effectively?

Too advanced 1 - why? Speech rate too fast 2 (Video 2)
Speech rate okay 1
Swiss German accent/dialect unclear 2
no point to video 3 1
Visual aids 5
Directions as theme 1

Simple language 1
Repetition 2
Listening activities 1

Ok so like when it comes to the like you said the the directions for example lets let's take that. So how do you think that kind of thing helps your language learning?

Videos are task dependent 1
Pronunciation 3
Comprehension 1
Memory 3
Learning styles (visual learner) 2
Vocabulary 1
Voices other than lecturers 1
Distracting 1
Dramatic 1
Cringeworthy 1
Uncomfortable 1
Unrealistic 1
Understandably unrealistic due to grade of German 1
Level appropriate 1
Dramatisation as learning aid 2
Gestural aid 1
Personalised engagement with language through opinion formation 1

Do you guys then think that like these are what do you think these videos are useful for?

Speech opportunity 2
Intriguing 1
Conversation in context 1

So, any disadvantage of your language learning in comparison to learning from books? From the videos?

Videos aren't written down 1
Listening 1

You've all rated on a scale of 1 to 4 that's why I've given these back to you whether you would or would not recommend these videos as a study aid. Why did you choose the rating you did?

Uncomfortable 1
German exposure 1
Manage your expectations 2
Building knowledge about German and Germany 1
Memory 1
Intermediate knowledge needed 1

Also, you rated to what extent would you say the interactions and the depictions of the German-speaking countries in these videos match your idea of what Germany, Austria and Switzerland are like?

Overdramatic 1
Asking for directions are realistic 3
Not saying thank you/impoliteness unrealistic 1
Escorting someone to their destination unrealistic 1
All situations are realistic/possible 1 (translated it from mother culture)

Do you have any ideas on how we can better use these videos in your lectures or tutorials?

Underuse books 1
1 video with three tasks 2
Better task design with a longer video 2
Opportunities for discussions in German 2
More discussion of the videos 1

Okay so would Mr Waurich rent the flat?

physical representation of the flat 1
Flat is really tiny 2
Flat is really huge if in Berlin 1
Task for video 2 could allow your own descriptions 1

What do you guys think about the interactions of like speaking the speaking and what's going on matching your ideas of Germany, Switzerland and Austria?

Overdramatic 1
Germans are helpful 1
Germans are friendly 1

Transcription B

Is there anything interesting that you found in the videos um or in filling in the questionnaires that you want to discuss before start?

Speech rate too fast (Video 2) 3

what are the advantages of language learning from these videos especially in comparison to learning from books?

Hear words as actually spoken 2
Pronunciation 2
Memory from hearing it 1

And then do you think what are the disadvantages of language learning from these videos from your personal experiences, especially in comparison to learning from books?

Unsure if realistic 4
Overdramatic 2
Understandably unrealistic due to grade of German 1
Misleading 2
Realistic 1

Repetition and overdramatisation for comprehension 1
Visual aids 2
Asking for directions is realistic 1
To view an apartment for rent is realistic 1

to what extent would you say the interactions and depictions of the German-speaking countries in these videos then match your idea of what Germany is like?

Humanness 1
Germans are friendly 1
Germans are a happy people 1
Germans are funny 1
Uncomfortable 3
Distracting 1
Cringey 1
Some videos more realistic than others (guitar store example given) 1
Overdramatic 1
Actions unrealistic 2

what do you think these videos are useful for when learning for German?

Mapping visuals to sound 1
Memory 1
Language used in context 1
Diversity 1

whether you would or would not recommend these videos as a study aid.

Videos would be distracting as study aid 1
Overdramatic 1
Could just listen to the audios and that would be more helpful 2
to accommodate those who prefer visual aids 1
Some people may have different learning styles 1
Cultural things 1
Mapping visuals to sound 1
Not helpful 1

so, regarding language in the videos, what did you find difficult to understand?

Easy to understand 1
Speech rate too fast 3
Speech could be more challenging 1
When doing the task while watching the videos you miss out 2
Background noises distracting 1

Do you have any ideas on how we can better use these videos in a lecture or tutorial setting?

Best when first learning a section 1
Must learn base vocabulary first 1
Teach vocab with video is helpful 1
When doing the task while watching the videos you miss out 1

Speech rate too fast 1
Better to watch then answer oral questions and write them down 1
Repeated viewings 1

Okay so would Mr Waurich rent the flat?

physical representation of the flat 4
1 person said it was small
1 person assumed it was cheap

Appendix G – Transcription and Translation of Video 1

The followingis a link to the transcription of Video 1 (Clip 13):

https://www.hueber.de/menschen/lehren/download?kategorie=transkription&kategorie_1=&band=a1&band_1

Translation of Video 1:

Clip 13: How do I get to the Goetheplatz

Touristin:	Ähh. Excuse me?
Oliver:	Ja?
Touristin:	Can you help me? Where is the Goetheplatz?
Oliver:	The Goetheplatz?
Touristin:	Ja. Ich would like to go to the Goetheplatz. Is it here in the area?
Oliver:	Goetheplatz. It is not far from here at all. So, pay attention: you now go 50m straight
Touristin:	50m straight
Oliver:	Then right at the corner
Touristin:	Right at the corner.
Oliver:	Then left at the next street and immediately right again.
Touristin:	left at the next street and immediately right again.
Oliver:	then 200m straight
Touristin:	200m straight
Oliver:	Left at the robot
Touristin:	Left at the robot
Oliver:	and now again 400m straight
Touristin:	Now, I go, um. Okay first left
Oliver:	No, first go 50m straight
Touristin:	Ah ja, and then left, ah no, right. Oh dear! How was it?
Oliver:	What's the use, come with me!
Touristin:	Very nice!
Oliver:	50m Straight. Right at the corner. Left at the next street and immediately right again. 200m straight. And here left at the robots. And now another 400m straight. And there is the Goetheplatz.
Touristin:	Aah!
Oliver:	very nice! Thank you very much. Thank you.

Appendix H – Transcription and Translation of Video 2

The following is a link to the transcription of Video 2 (Clip 14):

https://www.hueber.de/menschen/lehren/download?kategorie=transkription&kategorie_1=&band=a1&band_1

Translation of Video 2:

Clip 14: The super flat

Mrs Möllemann:	Ah, hallo! Are you Mr Waurich? You would like to view the flat, not so? Well, then come in! We will start here now: This here is the hallway. It is not very big, but completely practical! <chuckle> Come with me! Mr Waurich? Hello? Well, look! Come! This is the kitchen. Yes, the dishes, I have not yet washed the dishes today. But look, the view is very beautiful, not so? Come with me! Mr Waurich? Come no! Look! <chuckle> Here is the living room. I like the room. It is really comfortable, I find. <chuckle> <chuckle> Yes yes, look outside! Here you have all the shops right around the corner. There you can go shopping fantastically! That is super! Well, and now look! Mr Waurich? Mr Waurich, look! This is the bedroom. Yes, excuse me, I have not yet made the bed today. But I sleep very well here. Well! And now I am showing you, um, Mr Waurich? Hello? Come! Come! The bathroom! Um, Mr Waurich? Mr Waurich?
Mr Waurich:	Phew!

Appendix I – Transcription and Translation of Video 3

The following is a link to the transcription of Video 3 (Clip 15):

https://www.hueber.de/menschen/lehren/download?kategorie=transkription&kategorie_1=&band=a1&band_1

Translation of Video 3:

Clip 15: This is my city.

Martin:	Hello people! Greetings and welcome here in Bern. My name is Martin Zürcher and I live here. Bern has 130 000 residents and is the capital city of Switzerland. I would like to show you guys my city today! Simply come with me! Here one speaks Bärndütsch. Bärndütsch, this means: Bernese German. This is German, but not Hochdeutsch. Do you guys want to hear a sentence in Bärndütsch? Salome, say something!
Salome:	(in Bärndütsch) a loving greeting to all German learners!
	...
Martin:	Did you guys understand anything? No? Salome said: (in Hochdeutsch) a loving greeting to all German learners! This is the Zytglogge, the clock tower. The tower is very old – 800 years. The clock is 500 years old. Well, how do you guys like the clock and the tower? Bern and Berlin have something in common. Do you guys know what? Correct, the city names both start with B. And look: This is how the coats of arms look. The Berlin residents have a bear, and the Bern residents have a bear. Would you guys like to see a bear? Then we are in the right place here at the bear pit. Nice, not so? Well, now you guys have already seen some things of my city. Well, how do you guys like Bern? Maybe you guys are coming sometime to Switzerland? Then come also here, to Bern, and see everything for yourselves. Until then! Goodbye! Cheers/Bye!