



universität  
wien

# DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„A comparative study of video-mediated and audio-only  
listening tests in the Austrian EFL classroom“

verfasst von / submitted by

Julia Zirhan

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of  
Magistra der Philosophie (Mag. phil.)

Wien, 2016 / Vienna, 2016

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

A 190 344 313

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Lehramtsstudium UniStG

UF Englisch UniStG

UF Geschichte, Sozialkunde, Polit. Bildg. UniStG

Betreut von / Supervisor:

Mag. Dr. Helen Heaney, BA

## Table of contents

Acknowledgements.....	iv
Inscription.....	v
Abbreviations and acronyms.....	vi
List of tables.....	vii
List of figures.....	ix
<b>1. Introduction .....</b>	<b>1</b>
1.1. Research questions .....	3
1.2. Hypotheses and null-hypotheses .....	4
<b>2. Theories of listening comprehension .....</b>	<b>5</b>
2.1. The process of listening comprehension .....	5
2.2. A framework for describing language knowledge and listening ability .....	7
2.3. Taxonomies of listening comprehension skills .....	10
2.4. The development of listening comprehension constructs .....	11
<b>3. Connections between non-verbal behaviour, communication and listening comprehension .....</b>	<b>13</b>
3.1. Definition of kinesics and non-verbal and kinesic behaviour .....	14
3.2. Non-verbal behaviour, kinesic behaviour and communication.....	15
3.3. Non-verbal communication and listening comprehension.....	17
3.3.1. <i>Gestures, communication and listening comprehension</i> .....	19
3.3.2. <i>Lip movements and their effects on listening comprehension</i> .....	23
<b>4. Effects of the inclusion of visual input in teaching and testing L2 listening comprehension .....</b>	<b>26</b>
4.1. Positive effects .....	27
4.1.1. <i>Positive effects of visual input on listener performance</i> .....	27
4.1.2. <i>Increase in authenticity</i> .....	29
4.1.3. <i>Access to cues from non-verbal communication</i> .....	31
4.1.4. <i>Contextualisation of listening texts</i> .....	32
4.1.5. <i>Further positive effects of the inclusion of visual input in teaching and testing listening</i> .....	33
4.2. Negative effects .....	35
4.3. Controversial effects relating to construct validity .....	37

<b>5. Visual competencies involved in video-mediated listening tests and theories of listening comprehension.....</b>	<b>42</b>
5.1. Integrating non-verbal communication and visual competencies in theories of language ability .....	42
5.2. Integrating non-verbal communication and visual competencies in theories of listening comprehension.....	45
5.3. A modified B2 listening comprehension construct for video-mediated listening tests.....	48
<b>6. Methodology.....</b>	<b>50</b>
6.1. Study setting and participants.....	50
6.2. Procedure.....	51
6.3. Teacher interview .....	52
6.4. First testing session .....	52
6.4.1. <i>Baseline listening test</i> .....	53
6.4.2. <i>First student questionnaire</i> .....	53
6.5. Second testing session .....	54
6.5.1. <i>Audio-only and audio-visual listening test</i> .....	54
6.5.2. <i>Test specifications operationalised in the audio-only and the audio-visual listening test</i> .....	55
6.5.3. <i>Description and analysis of the visuals used in the audio-visual listening test</i> .....	56
6.5.4. <i>Second set of student questionnaires</i> .....	60
<b>7. Results.....</b>	<b>63</b>
7.1. Test and item analysis .....	63
7.1.1. <i>Test analysis: Baseline listening test</i> .....	63
7.1.2. <i>Item analysis: Baseline listening test</i> .....	66
7.1.3. <i>Test analysis: Audio-only and audio-visual listening test</i> .....	69
7.1.4. <i>Item analysis: Audio-only and audio-visual listening test</i> .....	72
7.2. Effects of visual input on listening comprehension performance .....	74
7.2.1. <i>Effects of visual input on listening comprehension performance of test takers with different listening ability levels</i> .....	76
7.2.2. <i>Effects of visual input on listening comprehension performance of test takers with different learning styles</i> .....	78
7.3. Importance of visual input in everyday communication and language learning...	80
7.4. Attention paid to the visual input in the audio-visual listening test .....	84

7.4.1.	<i>Differences in the attention paid to the visual input in the audio-visual listening test between test takers with different listening ability levels.....</i>	85
7.4.2.	<i>Differences in the attention paid to the visual input in the audio-visual listening test between test takers with different learning styles .....</i>	86
7.5.	Test taker opinions on the inclusion of visual input in a listening test .....	87
7.5.1.	<i>Perception of visual input as an aid to comprehension or a distracting factor .....</i>	87
7.5.2.	<i>Test taker preference of audio-visual listening tests or audio-only listening test .....</i>	91
7.5.3.	<i>Differences in test taker opinion between test takers with different listening ability levels.....</i>	92
7.5.4.	<i>Differences in test taker opinion between students with different learning styles.....</i>	93
7.5.5.	<i>Test taker comments.....</i>	95
7.6.	Results in relation to the hypotheses and null-hypotheses .....	96
<b>8.</b>	<b>Discussion and conclusion.....</b>	<b>99</b>
<b>9.</b>	<b>Bibliography.....</b>	<b>105</b>
<b>Appendix 1:</b>	<b>Abstracts.....</b>	<b>110</b>
	Abstract.....	110
	Zusammenfassung .....	111
<b>Appendix 2:</b>	<b>Taxonomy of listening comprehension skills .....</b>	<b>112</b>
<b>Appendix 3:</b>	<b>Letter to parents .....</b>	<b>114</b>
<b>Appendix 4:</b>	<b>Baseline listening test (including first student questionnaire, correction keys and test specifications).....</b>	<b>116</b>
<b>Appendix 5:</b>	<b>Audio-only listening test (including AO student questionnaire, correction key and test specifications) .....</b>	<b>131</b>
<b>Appendix 6:</b>	<b>Audio-visual listening test (including AV student questionnaire and test specifications).....</b>	<b>147</b>
<b>Appendix 7:</b>	<b>Calculations.....</b>	<b>162</b>
1.	Frequency statistics .....	162
2.	T-tests .....	162
3.	ANOVA.....	165
4.	Correlations .....	167

## **Acknowledgements**

First of all, I would like to express my sincere gratitude to Mag. Dr. Helen Heaney, BA for her excellent supervision of my diploma thesis. She supported and guided me from the early stages of devising an idea onwards to the process of conducting my study and writing the thesis until the day that I finally handed in my diploma thesis. I want to thank her especially for offering me her expertise and advice and encouraging me when I did not know how to continue. Furthermore, I would like to thank my fiancé, Sebastian Stitz, for always being there for me, for emotionally supporting me throughout my studies and for reminding even in the busiest times that there is a life outside the university. Finally, I want to thank my family for supporting me throughout my entire life, for encouraging me in all my decisions and enabling me to study in Vienna and become a teacher in the near future.

*Thank you very much!*

*To my dear fiancé Sebastian*

## Abbreviations and acronyms

abbr.	abbreviation
ACTFL	American Council on the Teaching of Foreign Languages
AO	audio-only
AV	audio-visual
bifie	Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens
CAID	Cronbach's Alpha if item deleted
CEFR	Common European Framework of Reference for Languages: Learning, teaching, assessment
CITC	corrected item-total correlation
EFL	English as a foreign language
ESL	English as a second language
FL	foreign language
H	hypothesis
H <sub>0</sub>	null-hypothesis
L2	second language (including second and foreign language)
max.	maximum
min.	minimum
N	number of test takers
no.	number
NVC	non-verbal communication
<i>p</i> -value	facility value
Qu.	questionnaire
sig.	significance
std. dev.	standard deviation
t	t value
TLU	target language use
TOEFL	Test of English as a Foreign Language

## List of tables

Table 1	Test takers' contact with English.....	51
Table 2	Test A: Descriptive statistics .....	64
Table 3	Test A: Descriptive statistics of test tasks .....	65
Table 4	Test A: Item statistics and item-total statistics .....	67
Table 5	Test A: Reliability statistics.....	67
Table 6	Test A: Revised item-total statistics .....	68
Table 7	Test A: Revised reliability statistics .....	68
Table 8	Test A: Comparison of control and experimental test groups .....	69
Table 9	Test A: Control vs. experimental test groups .....	69
Table 10	AO/AV test: Descriptive statistics.....	70
Table 11	AO/AV test: Descriptive statistics of test tasks.....	71
Table 12	AO/AV test: Item statistics and item-total statistics.....	72
Table 13	AO/AV test: Reliability statistics .....	72
Table 14	AO/AV test: Revised item-total statistics.....	74
Table 15	AO/AV test: Revised reliability statistics.....	74
Table 16	AV/AO test: Comparison of control and experimental test groups.....	75
Table 17	AO/AV test: Control vs. experimental groups .....	75
Table 18	Test A and AO/AV test: Comparison of means of test groups split by listening ability level.....	77
Table 19	Paired samples t-test: Comparison of performance of test groups split by listening ability level on Test A vs. AO/AV test.....	78
Table 20	AO/AV test: Comparison of performance of lower-level ability students on AO vs. AV test.....	78
Table 21	Test A and AV test: Comparison of performance of experimental group split by learning style on test A vs. AV listening test .....	79
Table 22	AV/AO Qu. item 9: "The following activities are very helpful in improving my English listening skills" .....	83
Table 23	AV Qu. item 10: "How often did you watch the video in the test?" (arranged according to ability level) .....	86
Table 24	AV Qu. item 10: "How often did you watch the video?" (arranged according to learning style).....	87
Table 25	AV Qu. items 14-17 (abbr.): Visual input supported comprehension or was distracting and had a negative effect on comprehension (arranged according to ability level).....	93
Table 26	AV Qu. items 14-17: (abbr.): Visual input supported comprehension or was distracting and had a negative effect on comprehension (arranged according learning style).....	94
Table 27	Richards' (1983) taxonomy of listening comprehension skills: Conversational listening (Richards 1983: 228-229).....	112
Table 28	Richards' (1983) taxonomy of listening comprehension skills: Academic listening (Richards 1983: 229-230).....	113
Table 29	AV/AO Qu. item 9: „The following activities are very helpful in improving my English listening skills.” .....	162
Table 30	AV/AO test: Comparison of performance of control vs. experimental groups on individual test tasks .....	162

Table 31	AV/AO test: Comparison of performance of control vs. experimental groups on individual test items .....	163
Table 32	Paired samples t-test: Comparison of performance of experimental test group split by learning style on test A vs. AV listening test .....	164
Table 33	Paired samples t-test: Comparison of performance of test takers who differed in their opinion on AV Qu. item 14 on test A vs. AV listening test.....	164
Table 34	Paired samples t-test: Comparison of performance of test takers who differed in their opinion on AV Qu. item 16 on test A vs. AV listening test.....	164
Table 35	Paired samples t-test: Comparison of performance of test takers who differed in their opinion on AV Qu. item 17 on test A vs. AV listening test.....	165
Table 36	ANOVA 1: Test of homogeneity of variances .....	165
Table 37	ANOVA 1: One-way ANOVA.....	165
Table 38	ANOVA 2: Test of homogeneity of variances .....	166
Table 39	ANOVA 2: One-way ANOVA.....	166
Table 40	ANOVA 3: Test of homogeneity of variances .....	166
Table 41	ANOVA 3: One-way ANOVA.....	166
Table 42	ANOVA 4: Test of homogeneity of variances .....	167
Table 43	ANOVA 4: One-way ANOVA.....	167
Table 44	Correlation: Listening ability and AV Qu. item 10 (“How often did you watch the video in the test?”) .....	167
Table 45	Correlation: Learning style and AV Qu. item 10 (“How often did you watch the video in the test?”) .....	167
Table 46	Correlation: Listening ability and AV Qu. item 14 (abbr.: Visuals supported listening comprehension and helped answer test questions) .....	168
Table 47	Correlation: Listening ability and AV Qu. item 16 (abbr.: Seeing the speakers was distracting and had negative effect on listening comprehension).....	168
Table 48	Correlation: Listening ability and AV Qu. item 17 (abbr.: Seeing the visual context was distracting and had negative effect on listening comprehension).....	168
Table 49	Correlation: Learning style and AV Qu. item 14 (abbr.:Visuals supported listening comprehension and helped answer test questions).....	169
Table 50	Correlation: Learning style and AV Qu. item 16 (abbr.: Seeing the speakers was distracting and had negative effect on listening comprehension).....	169
Table 51	Correlation: Learning style and AV Qu. item 17(abbr.: Seeing the visual context was distracting and had negative effect on listening comprehension).....	169

## List of figures

Figure 1	Framework for describing language knowledge (Bachman & Palmer 1996: 68) ....	8
Figure 2	A framework for describing listening ability (Buck 2001: 104) .....	9
Figure 3	Bifie B2 listening test construct (Universität Innsbruck 2012) .....	12
Figure 4	A modified B2 listening comprehension construct for video-mediated listening tests - based on the Bifie B2listening test construct (Universität Innsbruck 2012)	48
Figure 5	Test results: Baseline listening test.....	65
Figure 6	Test results: AO and AV listening test .....	71
Figure 7	AV/AO Qu. items 6-8: “In face-to-face conversations, I observe my conversation partner’s gestures / lip movements / facial expressions.” .....	81
Figure 8	AV/AO Qu. items 1-4: “I think that it is easier to comprehend spoken English if I am able to see the context of the speaking situation / the speaker’s gestures / lip movements / facial expressions.” .....	82
Figure 9	AV/AO Qu. item 5: “I think that it is easier to understand English on TV rather than English on the radio.” .....	82
Figure 10	AV Qu. item 10: “How often did you watch the video in the test?” .....	84
Figure 11	AV Qu. items 11-13: Attention paid to facial expressions, lip movements, gestures and visual contextual in the AV listening comprehension test. ....	85
Figure 12	AV Qu. 14: “In general, the visuals in this video-listening test supported my comprehension of the spoken text and helped me answer the test questions.” .....	88
Figure 13	AV Qu. item 15 (abbr.): These elements of NVC supported listening comprehension and helped answer test questions. ....	89
Figure 14	AV Q. items 16-17 (abbr.): Seeing the speakers or the visual context of the speaking situations was distracting and had a negative effect on listening comprehension.....	90
Figure 15	AV Qu. item 18: “I prefer a video-listening test over a conventional audio-listening test.” .....	91
Figure 16	AO Qu. item 12: “I would prefer a video-listening test where I could see the speaker(s) and the visual context rather than a conventional audio-listening test.” .....	92

## **1. Introduction**

Videos have been an integral part of language teaching and learning in the English as a foreign language (EFL) classroom for several decades, especially in connection with listening comprehension. However, their effects on comprehension and learning processes are still unclear and investigations into this topic have revealed rather conflicting results as they indicated positive as well as negative effects for visual input on language learners' listening comprehension skills. Despite the fact that video material is commonly used in teaching practice, its use is scarce and barely researched in EFL and English as a second language (ESL) listening tests, which separates testing practice from current teaching practice.

This study aligns EFL classroom pedagogy and testing practice by integrating visual input in a listening test for Austrian EFL students attending an upper secondary school and comparing an audio-visual (AV) listening testing with a traditional audio-only (AO) listening test. The study intends to establish whether the use of visual input, which provides the learners with non-verbal cues that accompany the spoken message, leads to increased test taker performance. This will be achieved by comparing the results of a listening test presented to a control group in audiotape-mediated format and to two experimental groups in video-mediated format. Furthermore, the study investigates the students' interaction with visual input in everyday communication, focusing, in this regard, on the attention they pay to visual input in general and the importance they attach to visual cues in language learning and communication. This will be achieved by analysing the results of a questionnaire. This analysis will then be compared with the test takers' interaction with the visual input provided in the AV listening test, also elicited in the questionnaire. In this regard, the questionnaire provides insights into the extent to which the test takers engaged with the videos in the testing situation, indicated by the time the test takers reported to have spent observing the video material, and the attention they paid to non-verbal cues in the listening test videos. This will be additionally correlated with the test takers' opinions on the usefulness of visual input to support listening comprehension and their preference for traditional AO listening tests or AV listening tests. Finally, the study aims to explore differences in the attention paid to the visual input in the audio-visual listening test and opinions on the inclusion of visual input in listening tests between

students of a lower and higher level of listening ability and between different types of learners, i.e. visual, auditory and kinaesthetic or tactile, who will be identified through the questionnaire answers.

By exploring the above-mentioned aspects, this study aims to contribute to a scarcely researched area and intends to provide new insights into the effects of visual stimuli on test takers and their listening ability by correlating test taker performance, the attention paid to visual stimuli and the opinions of the test takers on the use of videos in listening tests.

This diploma thesis presents the theoretical and empirical background relevant to the conducted “Comparative study of video-mediated and audio-only listening tests in the Austrian EFL classroom” before outlining the method used in this study as well as presenting and discussing the results. Chapter 2 gives an overview of the theories of listening comprehension, focusing on considerations concerning the process of listening comprehension, frameworks for describing language knowledge and listening ability, taxonomies of listening comprehension skills and listening comprehension constructs. Thereby, Chapter 2 expounds the theoretical foundation on which the listening tests used in this diploma thesis study were based. Afterwards, Chapter 3 focuses on non-verbal communicative behaviour, which is ubiquitous in communication and an integral aspect of visual materials used in video-mediated listening tests. This chapter explores the intricate connection between non-verbal behaviour, communication and listening comprehension in general and provides further insights into the link between gestures, communication and listening comprehension as well as lip movements and their effects on listening comprehension. Chapter 4 subsequently presents the state of research on the inclusion of visual materials in teaching and testing ESL and EFL listening comprehension and discusses positive as well as negative effects of these materials on language learners and their listening ability as well as on listening tests themselves. Concerning the latter, the focus is primarily on the controversially discussed effects of visual materials on the construct validity of listening tests. Afterwards, Chapter 5 outlines the sparse endeavours of pioneers in the field to include non-verbal communication and visual competencies in the theory of language ability and listening comprehension and offers a listening comprehension construct for B2-level listening that includes visual competencies.

After laying the theoretical and empirical foundations, the methodology used in this study is described in Chapter 6, covering the setting and participants as well as the study procedure and its individual aspects, e.g. the AV listening test, the AO listening test and the questionnaires designed for this study. Subsequently, the results of the study are presented and described in Chapter 7, which presents first a test and item analysis of the different listening tests before focusing on investigating the answers to the research questions below and relating the results to the developed hypotheses and null-hypotheses. Finally, Chapter 8 discusses the results of this study and arrives at a conclusion.

### **1.1. Research questions**

In this diploma thesis, I intended to answer the following research questions:

1. *Does the use of video material in a listening test affect the test takers' listening comprehension performance?*
  - a) *If so, does the inclusion of visual stimuli increase or decrease test taker performance?*
2. *How important are visual stimuli in everyday communication and language learning for the test takers?*
3. *To what extent do the test takers attend to the visual input in a video-mediated listening test?*
  - a. *Is there a difference in the attention paid to the visuals between higher- and lower-level ability test takers?*
  - b. *Is there a difference in the attention paid to the visuals between test takers with different learning styles?*
4. *To what degree do the test takers perceive the visual stimuli as an aid to comprehension or a distracting factor?*
  - a) *Is there a difference in the perception of visual stimuli as helpful or distracting between higher- and lower-level ability test takers?*
  - b) *Is there a difference in the perception of visual stimuli as helpful or distracting between test takers with different learning styles?*

## 1.2. Hypotheses and null-hypotheses

Based on these research questions, the following hypotheses were formulated:

Hypothesis 1: *The inclusion of visual stimuli leads to improved test taker performance as the test takers utilise the non-verbal cues provided in the visual material to aid comprehension.*

H1<sub>0</sub>: *There is no difference in the test takers' listening comprehension performance in the presence or absence of visual input in a listening comprehension test.*

Hypothesis 2a: *Test takers with a lower level of listening ability attend to visual cues to a higher degree than higher-level ability test takers in order to compensate for a lack of comprehension.*

H2a<sub>0</sub>: *There is no difference in the attention paid to visual cues between higher- and lower-level ability test takers.*

Hypothesis 2b: *Test takers with a lower level of listening ability perceive the visual stimuli as more helpful and less distracting than higher-level ability test takers.*

H2b<sub>0</sub>: *There is no difference in the perception of the visual stimuli as helpful or distracting between higher- and lower-level ability test takers.*

Hypothesis 3a: *Test takers who classify themselves as visual types of learners attend to visual cues to a higher degree than test takers who indicated that they are auditive or kinaesthetic types of learners.*

H3a<sub>0</sub>: *There is no difference in the attention paid to visual cues between test takers with different learning styles.*

Hypothesis 3b: *Test takers who classify themselves as visual types of learners perceive the visual stimuli as more helpful and less distracting than test takers who indicated that they are auditive or kinaesthetic types of learners.*

H3b<sub>0</sub>: *There is no difference in the perception of the visual stimuli as helpful or distracting between test takers with different learning styles.*

## **2. Theories of listening comprehension**

Second language (L2) and foreign language (FL) listening comprehension was a largely neglected area in language testing for a long time due to the lack of a widely-accepted theory of listening comprehension (Buck 1991: 67, cited in Wagner 2002: 1). In the 1990s, Dunkel (1991: 433-434) and Dunkel, Hening and Chaudron (1993: 180) remarked that no consensus has been reached in the field of language research on the components of listening comprehension, how the process of listening works, what it entails and which factors influence the process. Consequently no widely-acknowledged definition of listening comprehension, neither of native nor second or foreign language listening, was available.

The interest in listening comprehension in language testing has increased considerably since then and L2 and FL listening comprehension have developed into important research areas (Wagner 2002: 1). However, theorists continue to define and describe the complex and multidimensional process of listening comprehension in various ways (Buck 2001: 51) and it was argued that still no agreement has ever been reached in the theory of listening comprehension on a commonly-accepted definition of native, L2 or FL listening comprehension (Ockey 2008: 518; Wagner 2002: 1).

This chapter outlines theories of listening comprehension. It focuses first on theoretical considerations concerning the process of listening comprehension and presents afterwards a framework for describing listening comprehension, which is embedded in a framework for describing language knowledge. Finally, the chapter introduces different taxonomies of listening comprehension skills and listening comprehension constructs.

### **2.1. The process of listening comprehension**

It has been argued that what makes defining listening comprehension such an intricate task is the complexity of the listening process (Dunkel 1991: 434) or probably even listening processes themselves, as different processes presumably work together when trying to comprehend a spoken text (Wagner 2002: 1).

One established description of the listening comprehension process can be found in Buck (2001: 247). He describes it as “a complex process in which the listener takes the incoming data, an acoustic signal, and interprets it based on a wide variety of linguistic and non-linguistic knowledge” (Buck 2001: 247), whereby the former involves knowledge of, for example, lexis, syntax and phonology and the latter includes, for instance, topical knowledge and world knowledge (Buck 2001: 1-2). The interpretation of the acoustic signal is then continually modified in the on-going listening process as new acoustic input becomes available (Buck 2001: 247). Consequently, Buck (2001: 31) views listening comprehension as “an active process of constructing meaning”. Buck’s (2001) description of the listening comprehension corresponds with Ockey’s (2007) understanding of the process, which he depicts as one in which the mind simultaneously processes multiple stimuli and continually revises its interpretation in accordance to new incoming information (Ockey 2007: 518)<sup>1</sup>.

Brown and Abeywickrama (2010) go beyond describing the process of listening on a relatively abstract and general level to establishing possible components of this complex process. According to them, the listening comprehension process involves (1) recognising speech sounds and memorising them temporarily in the short-term memory and (2) determining the speech type as, for example, interpersonal or transactional dialogue or monologue as well as processing the context of the speech type and the content of the communicated message (Brown and Abeywickrama 2010: 161). Furthermore, it also involves (3) using bottom-up or top-down processes to interpret the message and comprehend literal and intended meaning and finally (4) the deletion of the exact acoustic message in order to preserve the relevant information in the long-term memory (Brown & Abeywickrama 2010: 161).

The depiction of listening comprehension as a bottom-up or a top-down process is generally common in the field of research (e.g. Brown & Abeywickrama 2010: 161; Buck 1991: 2-3; Dunkel 1991: 448; Rost 1990: 9). These widespread bottom-up and top-down views refer to the sequence in which the types of knowledge involved in listening are applied in the process of comprehension (Buck 2001: 2). According to the bottom-up view, speech comprehension passes through several consecutive levels or stages, starting, for

---

<sup>1</sup> Ockey’s understanding of the listening comprehension process is based on insights gained from Kintsch’s (1998) theory of text comprehension.

example, with the perception of sounds, followed by the interpretation of their meaning (Kelly 1991: 135), whereby the output of one level becomes the basis for the next higher level (Buck 2001: 2). This view was, however, criticised in the last decades as research suggested that there is no fixed order in which different types of knowledge are applied in the listening process (Buck 2001: 2). This led to Buck's (2001: 3) conclusion that listening comprehension is a top-down process, which means that the types of knowledge involved in the comprehension process occur in a random order, even simultaneously, and that these types of knowledge influence and interact with each other<sup>2</sup>.

Apart from this, listening comprehension is also typically perceived as consisting of two stages, one involving lower-level and one involving higher-level comprehension skills, which occur simultaneously and are interrelated (Brindley 1998; Buck 2001; Wagner 2002). The lower-level comprehension processes involve, for instance, the development of an understanding of the literal level of the aural input (Brindley 1998: 172) or the extraction of basic linguistic information, while the higher-level processes involve utilising the received information for communicative purposes (Buck 2001: 51).

## **2.2. A framework for describing language knowledge and listening ability**

Currently the most widely-accepted framework for describing language knowledge is the theoretical framework by Bachman and Palmer (1996: 68), based on Bachman (1990: 87) and illustrated in in Figure 1, in which language knowledge consists of organisational and pragmatic knowledge. Organisational language knowledge itself is divided into grammatical knowledge (Bachman & Palmer 1996: 68) and textual knowledge, which involves understanding the “conventions for joining utterances together to form a text”, which functions by applying the “rules of cohesion and rhetorical organisation” (Bachmann 1990: 88). Pragmatic knowledge comprises illocutionary and sociolinguistic knowledge (Bachman & Palmer 1996: 68).<sup>3</sup>

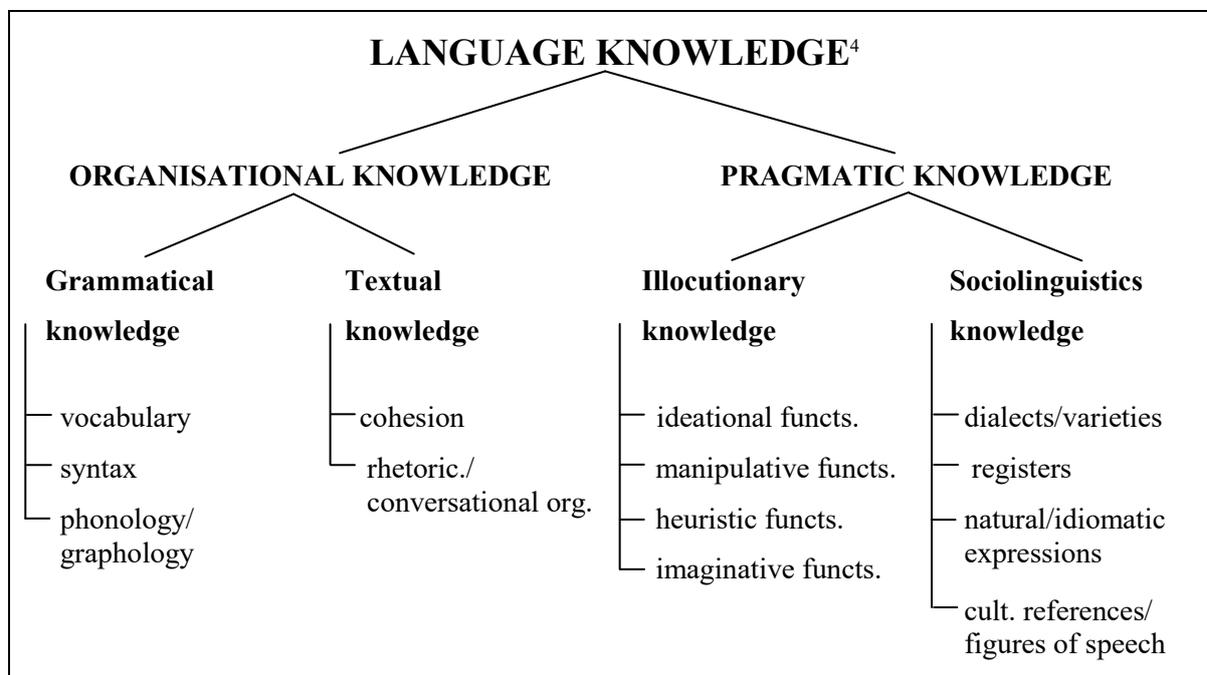
Concerning this framework for describing language knowledge, Bachman stresses that it is not a finite and complete framework (Bachman 1990: 81) and not hierarchically structured

---

<sup>2</sup> As the different types of knowledge involved in listening comprehension interact with each other, the process of listening comprehension was sometimes termed an interactional process (Buck 2001: 3).

<sup>3</sup> For an in-depth explanation of the framework of language knowledge see Bachman and Palmer (1996: 68-70).

as the layout of the model might suggest (Bachman 1990: 86). Language should rather be seen as a dynamic process (Bachman 1990: 83) in which the different components of language knowledge interact with each other (Bachman 1990: 86).



**Figure 1 Framework for describing language knowledge (Bachman & Palmer 1996: 68)**

This framework of the areas of language knowledge also exists within the context of a broader “framework of communicative language ability”, which involves beside language knowledge also topical knowledge, the personal characteristics of the language user, his or her strategic competence and the affective schemata involved in language use (Bachman & Palmer 1996: 63). These individual components of language use and performance interact with each other and with the characteristics of the language use task and the language use situation (Bachman & Palmer 1996: 63).

Based on the above-mentioned models of language knowledge and communicative language ability, Buck (2001: 104) developed the following widely-acknowledged framework for describing listening ability, quoted in Figure 2. It involves, on the one hand, language competence, consisting of grammatical knowledge, discourse knowledge, pragmatic knowledge and socio-linguistic knowledge, and, on the other hand, strategic competence, which consists of cognitive and metacognitive strategies (Buck 2001: 104).

<sup>4</sup> Explanation of abbreviations: rhetoric. means rhetorical, org. means organisation, functs. means functions and cult. means cultural.

**Language competence:** the knowledge about the language that the listener brings to the listening situation. This will include both fully automated procedural knowledge and controlled or conscious declarative knowledge. Language competence consists of:

**Grammatical knowledge:** understanding short utterances on a literal semantic level. This includes phonology, stress, intonation, spoken vocabulary and spoken syntax.

**Discourse knowledge:** understanding longer utterances or interactive discourse between two or more speakers. This includes knowledge of discourse features, such as cohesion, foregrounding, rhetorical schemata, story grammars and knowledge of the structure of unplanned discourse.

**Pragmatic knowledge:** understanding the function or the illocutionary force of an utterance or longer text and interpreting the intended meaning in terms of that. This includes understanding whether utterances are intended to convey ideas, manipulate, learn or are for creative expression as well as understanding indirect speech acts and pragmatic implications.

**Sociolinguistic knowledge:** understanding the language of particular socio-cultural settings, and interpreting utterances in terms of the context of situation. This includes knowledge of appropriate linguistic forms and conventions characteristic of particular sociolinguistic groups, and the implications of their use, or non-use, such as slang, idiomatic expressions, dialects, cultural references, figures of speech, levels of formality and registers.

**Strategic competence:** includes the cognitive and metacognitive strategies, or executive processes, that fulfil the cognitive management function in listening. This is the ability to use language competence, and includes all the compensatory strategies used by second-language listeners. It consists of:

**Cognitive strategies:** those mental activities related to comprehending and storing input in working memory or long-term memory for later retrieval.

**Comprehension processes:** associated with the processing of linguistic and non-linguistic input.

**Storing and memory processes:** associated with the storing of linguistic and non-linguistic input in working memory and long-term memory.

**Using and retrieval processes:** associated with accessing memory, to be readied for output.

**Metacognitive strategies:** those conscious or unconscious mental activities that perform an executive function in the management of cognitive strategies.

**Assessing the situation:** taking stock of conditions surrounding a language task by assessing one's own knowledge, one's available internal and external resources and the constraints of the situation before engaging in the task.

**Monitoring:** determining the effectiveness of one's own or another's performance while engaging in the activity.

**Self-evaluating:** determining the effectiveness of one's own or another's after engaging in the activity.

**Self-testing:** testing oneself to determine the effectiveness of one's own language use or the lack thereof.

Figure 2 A framework for describing listening ability (Buck 2001: 104)

### **2.3. Taxonomies of listening comprehension skills**

The complex and multidimensional process of listening comprehension has been described by many theorists in taxonomies that list the microskills involved in effective listening comprehension (e.g. Brown & Abeywickrama 2010; Buck et al. 1997, Buck & Tatsuoka 1998, Richards 1983). These taxonomies usually originate from insights from classroom practice and literature analysis, which induced Dunkel (1991: 447) and Buck (2001: 51) to state that empirical support for most of these taxonomies is still missing. Consequently, no taxonomy of listening comprehension skills can be considered as a finite list of all skills involved in listening comprehension; nevertheless they give insights into the state of research concerning listening comprehension (Buck 2001: 51).

An early comprehensive enumeration of listening microskills involved in different forms of listening situations was developed by Richards (1983: 228-230). He listed 33 microskills for conversational listening, including the skill to recognise reduced forms of words, cohesive devices and vocabulary items in spoken discourse, and 18 microskills involved in non-participatory, academic listening, i.e. listening to lectures (Richards 1983: 228-229). This form of academic listening comprised skills such as identifying the scope and the purpose of a lecture and recognising crucial vocabulary items related to the subject of the lecture (Richards 1983: 229-230). Richards also integrated microskills that involve the utilisation of non-verbal clues in his taxonomy, which will be explained in greater detail in Section 5.2. Richards' (1983) entire list of listening microskills is given in Appendix 2 (see Tables 27 and 28).

More than a decade later, Buck and Tatsuoka (1998) also developed a detailed list of skills or attributes needed in free response L2 listening tests. Their empirical research led to the identification of 15 cognitive attributes involved in listening, including the skill of scanning fast speech in real time, processing large quantities of information and using background knowledge to process the spoken text (Buck & Tatsuoka 1998: 141-142).

A recent example of a taxonomy of listening comprehension skills is the short and very general definition of four listening subskills by Brown and Abeywickrama (2010: 162), which corresponds to their list of mental processes occurring during listening comprehension, as defined in Section 2.1 on page 6. This list of listening microskills

involves (1) the ability to comprehend phonemes, grammar, words, intonation or other surface level elements of the spoken text, (2) the ability to comprehend the pragmatic context of the auditory input, (3) the ability to determine the meaning of the aural message and finally (4) the ability to develop a global understanding of the aural text, i.e. comprehend the gist (Brown & Abeywickrama 2010: 162).

#### **2.4. The development of listening comprehension constructs**

Based on the theories of listening comprehension, including frameworks for describing language and listening abilities, taxonomies of listening comprehension skills and theoretical considerations concerning the process of listening, test developers have constructed various listening constructs used in diverse listening comprehension tests. As the definition of a listening test construct depends on the purpose of a test as well as the target-language use (TLU) situation envisaged for the test (Buck 2001: 112), no generally applicable listening test construct exists.

Buck (2001: 114) developed quite an inclusive “default listening construct” which he argues is applicable in various testing situations. This construct defines listening as the ability “to process extended samples of realistic spoken language, automatically and in real time”, the ability “to understand the linguistic information that is unequivocally included in the text” and the ability “to make whatever inferences are unambiguously implicated by the content of the passage” (Buck 2001: 114). Despite the generalisability of this construct, Buck (2001: 112) emphasises that the construct can and should be modified according to the purpose of an assessment and the TLU situation involved in a test.

The following construct definition of B2 listening (Universität Innsbruck 2012), which was developed on behalf of bifie, the “*Federal Institute for Educational Research, Innovation and Development of the Austrian School Sector*”<sup>5</sup>, is important in the Austrian context in general and the context of this diploma thesis in particular. This construct definition of listening comprehension is used for English listening exams which are part of the Austrian school leaving examinations, the “Matura”<sup>6</sup>, and based on the “*Common European*

---

<sup>5</sup> The English translation is taken from bifie (2013). The German name for bifie is: „Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens“.

<sup>6</sup> The Austrian school leaving examination is referred to as Matura throughout the thesis.

*Framework of Reference for Languages*”, short CEFR, (Council of Europe 2001), which is generally the basis for most Austrian listening comprehension constructs.

<p><b>General purpose</b></p> <ul style="list-style-type: none"><li>a) listening for gist</li><li>b) listening for specific information</li><li>c) listening for important details</li><li>d) search listening</li><li>e) listening for main ideas and supporting details</li><li>f) listening to infer (propositional) meaning</li><li>g) listening to deduce the meaning of a word</li></ul> <p><b>Specific purpose</b></p> <ul style="list-style-type: none"><li>1. Can understand the main ideas of propositionally and linguistically complex speech on both concrete and abstract topics delivered in a standard dialect, including technical discussions in his/her field of specialisation.</li><li>2. Can follow extended speech and complex lines of argument provided the topic is reasonably familiar and the direction of the talk is sign-posted [sic] by explicit markers.</li><li>3. Can with some effort catch much of what is said around him/her, but may find it difficult to participate effectively in discussion with several native speakers who do not modify their speech in any way.</li><li>4. Can follow the essentials of lectures, talks and reports and other forms of academic / professional presentation which are propositionally and linguistically complex.</li><li>5. Can understand announcements and messages on concrete and abstract topics spoken in standard dialect at normal speed.</li><li>6. Can understand most radio documentaries and most other recorded or broadcast material delivered in standard dialect and can identify the speaker’s mood, tone etc.</li><li>7. Can use a variety of strategies to achieve comprehension including listening for main points and checking comprehension by using contextual clues.</li></ul>
--

**Figure 3 Bifie B2 listening test construct (Universität Innsbruck 2012)**

This chapter revealed that there are different theoretical definitions of listening comprehension, various descriptions of the process or processes involved in listening comprehension, frameworks for describing language and listening ability, a variety of taxonomies listing subskills of listening comprehension and different construct definitions of listening. So far no agreement has been reached concerning these theoretical issues; however, Dunkel, Hening and Chaudron (1993: 182) argued that different listening comprehension constructs and definitions of listening comprehension can exist simultaneously and that there is not necessarily a need for a generally accepted model of listening comprehension as these models depend on the purpose of listening and the purpose of the assessment.

### **3. Connections between non-verbal behaviour, communication and listening comprehension**

Although it has always been indisputable that non-verbal communication (NVC) is ubiquitous in interpersonal communication (Burgoon 1994: 232; Antes 1996: 439), amounting to approximately 69% of the meaning communicated in interpersonal communication, according to Philpott (1983; quoted in Burgoon 1994: 232)<sup>7</sup>, non-verbal communicative behaviour has been overlooked in various disciplines (Kellerman 1992: 239). While NVC tended to be neglected in applied linguistics, communication theory and in the field of L2 teaching in the 1990s, its importance was already widely acknowledged in fields such as psychiatry, psychology and anthropology (Kellerman 1992: 239). This lack of academic interest in NVC in the above-mentioned disciplines was based on the conviction that NVC is “incidental rather than essential to the communication process” and that it lacks the normative and rules-based nature which governs language use (Kellerman 1992: 239). Communication was then predominantly seen as being based only on verbal messages, although some linguists already stressed that the voice is only one means to communicate and only “one component of a larger system of bodily expression” (Kelly *et al.* 1999: 578). It was further pointed out that kinesic behaviour and speech actually interact to co-determine the meaning of utterances and it was requested that “researchers need to take on a broader view of communication and include [...] information from other modalities”, such as kinesic behaviour (Kelly *et al.* 1999: 589).

A decade later, Harris (2003: 180) still noted that non-verbal communication is largely ignored in the field of L2 teaching and stated that “nonverbal communication [...] forms an important part of everyday speech transmission and [...] should [therefore] occupy a more central position in second and foreign language [...] teaching than it currently does”. However, he observed that the opinion on integrating non-verbal communication at least in L2 teaching was changing slightly and that more and more scholars were starting to emphasise the importance of non-verbal communication in the production and reception stage of message transmission (Harris 2003: 180).

---

<sup>7</sup> Philpott conducted a meta-analysis of studies investigating the amount of meaning that is communicated verbally or non-verbally in interpersonal communication.

This chapter first presents definitions of kinesics and kinesic and non-verbal behaviour, before investigating the connection between non-verbal behaviour and communication in general and non-verbal behaviour and listening comprehension in particular. Concerning the latter, the chapter goes into greater detail describing the connection between gestures, communication and listening comprehension and further focuses on lip movements and their effects on listening comprehension.

### **3.1. Definition of kinesics and non-verbal and kinesic behaviour**

Kinesics is described as “the study of gesture and facial expression in a language system” and developed into an important subfield in linguistics, especially in the field of semiotics, in the early 1990s (Antes 1996: 439). While the definition of kinesics is relatively straightforward, there are different and sometimes conflicting definitions of non-verbal and kinesic behaviour, as explained below.

Kellerman (1992: 239) defines non-verbal behaviour as any form of behaviour that does not encompass language production, while Burgoon (1994: 232) also views what she terms vocalics or paralanguage, meaning the use of vocal cues that are not formed by words, such as pitch, velocity and intonation, as an aspect of non-verbal behaviour. Despite this categorisation difference, the two definitions are similar in the way that both depict non-verbal behaviour as a very broad concept which includes kinesic behaviour (Kellerman 1992: 239; Burgoon 1994: 232).

The term kinesic behaviour has also been used by several researchers (Buck 2001: 7; Burgoon 1994: 232; Kellerman 1992: 240; Morley 1991: 89; Rost 1990: 78) as an umbrella term for body language or body movements that are message-related and which include, for example, gestures, gaze, facial expressions, eye contact, posture, interpersonal distance as well as head, trunk and limb movements. Kinesic behaviour, which is for the most part produced and perceived unconsciously (Kellerman 1992: 240), forms an important aspect not only of non-verbal behaviour but also of non-verbal communication, as kinesic behaviour like other forms of NVC, such as physical appearance, haptics, which means the use of touch, and proxemics, which refers to the use of distance in conversations, conveys messages (Burgoon 1994: 232) and plays an crucial role in the process of encoding and decoding information in communication (Kellerman 1992: 239).

Kinesic behaviour is produced co-verbally, which means that it accompanies spoken language and is tied to the semantic content of the spoken message (Kellerman 1992: 240). Furthermore, the production as well as the understanding of kinesic behaviour is described as culturally dependent and language specific (Morley 1991: 89), which can be best seen by comparing the quantity and quality of body movements of a so-called ‘high-gesture culture’, existing typically in Italy or Spain, and a ‘low-gesture culture’, found, for example, in Great Britain (Kellerman 1992: 252).

### **3.2. Non-verbal behaviour, kinesic behaviour and communication**

The relationship between non-verbal behaviour and communication is a very complex one and it is controversially debated to what degree non-verbal behaviour can be seen as communicative. One perspective in research, called the “receiver (decoder) orientation”, suggests that all non-verbal behaviour is potentially communicative as the receiver decides what does or does not count as a communicative message (Burgoon 1994: 229-230). Thereby all human behaviour is turned into potential communicative behaviour, which is problematic according to Burgoon (1994: 230) because the concept of NVC thus becomes too broad and inclusive and consequently loses its meaningfulness. A further position in research describing the connection between non-verbal behaviour and communication is the “source (encoder) orientation“, which argues that only intentional behaviours or actions count as communication (Burgoon 1994: 230). This perspective, which narrows down the spectrum of non-verbal communication considerably, is also perceived as problematic as it is impossible to decide which non-verbal actions are intentional and which are unintentional (Burgoon 1994: 230). As a possible solution, Burgoon (1994: 231) suggests a further perspective, which she terms “message orientation”, in which the focus is on the actions and behaviours themselves. She defines non-verbal communication as consisting of

those behaviors other than words themselves that form a socially shared coding system, [...] are typically sent with intent, typically interpreted as intentional, used with regularity among members of a speech community, and have consensually recognizable interpretations (Burgoon 1994: 231).

According to this perspective, non-verbal behaviour can be habitual, can be performed consciously but also unconsciously, should be encoded frequently with a consistency in meaning and needs to be interpreted as meaningful (Burgoon 1994: 231-232).

Consequently, non-verbal behaviour that seems to be incidental and lacks coherence in meaning is excluded from the definition of NVC (Burgoon 1994: 232).

A common argument in research concerning the relationship between non-verbal communicative behaviour and interpersonal communication is that the non-verbal dimension, also referred to as the non-linguistic or extralinguistic dimension, is one of three different dimensions or channels in which messages are conveyed and perceived in interpersonal communication (Morley 1991; Rost 1990; Capper 2000). Rost (1990: 78) argued that speech acts are not only produced and perceived via linguistic means but “through a convergence of linguistic, paralinguistic and non-linguistic cues”, all of which need to be attended to in order to understand the intended meaning. This interplay of verbal, vocal and physical or visual cues, which corresponds to the linguistic, the paralinguistic and the non-verbal or extralinguistic dimension of message production and perception, is also described in Morley (1991: 88-89) and Capper (2000). They characterise the linguistic channel as comprising elements such as grammar and vocabulary, while the paralinguistic dimension is defined as the way in which the verbal message is constructed vocally by the means of, for example, pitch, stress or prosody, and the non-verbal channel is seen as transmitting information via body language (Morley 1991: 89; Capper 2000).

Kellerman (1992: 240-245), who developed a widely accepted model based on four essential links between the verbal and the non-verbal communicative channel, investigated how the non-verbal message is linked to the linguistic message. Focusing on kinesic behaviour, which forms an integral aspect of non-verbal communication, she argued that kinesic behaviour can be connected to the linguistic message (1) on the phonological level or (2) the semantic level of the spoken message, or (3) it can signal discourse features or (4) have regulatory and controlling functions in interaction (Kellerman 1992: 240). The first of the four links becomes apparent when observing eye or eyebrow movement, which is often linked to phonological features such as stress patterns and intonation, e.g. eyebrow raising at the end of a question corresponding to the rising intonation (Kellerman 1992: 241). Iconic gestures<sup>8</sup> or facial expressions revealing emotions are typically linked to the semantic level of the linguistic messages, while beat gestures usually function to indicate

---

<sup>8</sup> Definitions of the different types of gestures are given in Section 3.2.1.

discourse features, emphasising, for example, important content elements of the linguistic message (Kellerman 1993: 241-243). Beyond this, shifts in posture and position also frequently serve to indicate discourse features, correlating usually with structurally larger units of the spoken text, thereby providing crucial decoding cues for the listener and signalling, for instance, the end of a section in discourse and the beginning of a new one (Kellerman 1992: 243-244). Kinesic behaviour that serves to regulate interaction frequently assists turn-taking and maintaining the flow of speech or functions as important listener feedback for the speaker (Kellerman 1992: 244). This can be achieved, for example, through directing gaze, making eye contact or through bodily gestures and posture (Kellerman 1992: 244).

Non-verbal communication, which unlike earlier assumptions<sup>9</sup> is nowadays argued to be a “full-fledged semiotic system” (Antes 1996: 439) governed by semantic and syntactic rules (Burgoon 1994: 231), can generally serve a variety of functions in communication. NVC can be used to illustrate, repeat, reinforce, augment, accentuate or contradict the accompanying words but it can also substitute for words, like a shoulder shrug (Buck 2001: 47) or gestures to demonstrate actions or indicate objects in the situational context (Kelly *et al.* 1999: 578). Beyond this, NVC can also precede words or even be unrelated to the verbal message (Burgoon 1994: 239).

As most researchers have focused so far on the accompanying function of NVC, non-verbal communicative behaviour received a secondary status and was subordinated to verbal communication, although empirical research suggests that “nonverbal cues not only are essential to verbal message encoding and decoding but are also ‘message bearers’ in their own right” (Burgoon 1994: 239). Research has also indicated that people tend to believe non-verbal cues over verbal cues and rely for more purposes and with a higher frequency on non-verbal information (Burgoon 1994: 236).

### **3.3. Non-verbal communication and listening comprehension**

It is a common perspective in research that speech comprehension is a bi-sensory or bi-model process that comprises the modality of hearing and seeing (Harris 2003: 184;

---

<sup>9</sup> It was still commonly assumed in the 1990s, as stated in the introduction to Chapter 3, that non-verbal communicative behaviour lacks the formal rules and norms that language possesses.

Kellerman 1990: 274). However, it is still debated how the information from the two modalities combines in speech comprehension (Kelly *et al.* 1999: 579). Some researchers (e.g. Antes 1996; Dunkel, Henning & Chaudron 1990) view the visual, non-verbal channel as a secondary and subordinated system that serves to support the information perceived on the audio channel and functions as a back-up system used to compensate for a degraded auditory signal (Kellerman 1990: 274). However, other researchers (e.g. Gruba 1993; Gruba 1995) perceive the visual modality as having its own characteristics and being equally important to the auditory perception (Kellerman 1990: 274).

Despite the yet unclear nature of the relationship of verbal and visual information in the process of speech perception, a consensus is developing in research on the importance of visual information in listening comprehension (Wagner 2007: 67). This importance is nowadays recognised in the field of native as well as second language listening comprehension (Wagner 2007: 67) and several researchers (e.g. Buck 2001; Burgoon 1994; Hardison 2003) set out to explore the connection between NVC and native and L2 listening comprehension. However, concerning FL listening, the role of the visual channel in speech perception is still relatively unexplored.

Investigating the interplay of NVC and listening comprehension, Hardison (2003: 496-497) confirmed that visual cues can be used to compensate for comprehension problems and supported the claim that information derived from different sensory channels is combined in speech perception. Beyond this, Buck (2001: 47) argued that the information derived from kinesic cues can actually modify the interpretation of a verbal message entirely, for example, when sarcasm or disbelief is indicated non-verbally through facial expressions. This led Buck (2001: 48) to conclude that non-verbal communication “has the potential to influence or change the listener’s interpretation of the speaker’s words in a significant way”. Furthermore, Burgoon (1994: 235) established that adult listeners attach more importance to non-verbal cues than to verbal information to gain information about social meanings such as interpersonal relationships in communication, while the verbal channel is more important in understanding factual and persuasive messages. This corresponds to Bostrom’s study (1997: 27, cited in Buck 2001: 47), which suggests that listeners prefer to believe non-verbal cues over verbal cues when making inferences about the speaker’s emotions in situations where the cues derived from the two modalities

contradict each other. Bostrom (1997, cited in Buck 2001: 47) further argued that the decoding of visual signals is done with greater accuracy than the decoding of verbal cues.

### **3.3.1. Gestures, communication and listening comprehension**

Edward Sapir (1999 [1928]: 169) wrote in 1928 that “we respond to gestures with an extreme alertness and, one might almost say, in accordance with an elaborate and secret code that is written nowhere, known to none, and understood by all”. Although more than 80 years old, this statement still holds a certain truth as the majority of people sharing the same culture and language produce the same types of gestures in similar ways with similar meanings without being aware of the underlying elaborate code of gesture production inherent to their language and cultural community. However, an increasing number of researchers have set out to investigate the secret code of gesture production and arrived at definitions and various categorisations of gestures.

Generally gestures are defined as symbolic bodily movements that usually occur in synchronisation with speech and are semantically as well as syntactically related to the verbal message (Kelly *et al.* 1999: 578) and the speaker’s intention of expressing something, which excludes, amongst others, functional body movements such as using the hands to sip a drink, self-regulating gestures such as scratching and symptomatic movements (Gullberg 2006: 104) such as a Parkinsonian tremor. Gestures are produced mostly with the hands, arms or the head and are ubiquitous in interpersonal communication. The use of gestures differs from culture to culture, as different cultures have their own “gestural repertoires” and systems of gesture use (Gullberg 2006: 103) which are governed by cultural norms and conventions and linked to the structure of the individual languages (Gullberg 2006: 107). Despite exhibiting a considerable degree of uniformity within communities sharing the same culture and language, the use of gestures still varies from individual to individual and is also shaped by factors such as the situational context, the mood and status of the interlocutors and the topic of a conversation (Gullberg 2006: 107-108). Unlike language, which presents information in a highly structured way, gestures usually convey messages in a more holistic and pictorial way, serving to visually depict or highlight meaning in communication (Kelly *et al.* 1999: 578). Apart from this, gestures can also fulfil the functions of, for example, regulating turns, marking agreement or disagreement, directing attention, eliciting feedback (Gullberg 2006:

105), replacing speech, clarifying ambiguous utterances and guiding the listener's interpretation of the spoken text (Antes 1996: 439-440).

There are different categorisations of gestures which share certain similarities (Ekman & Friesen 1969; Goldin-Meadow 1999; Gullberg 2006; Harris 2003; Kellerman 1992; Kelly *et al.* 1999; McNeill 1985; Sueyoshi & Hardison 2005). All of them include a category called emblems, which denotes culturally specific gestures (Gullberg 2006: 105) that are independent of speech, have standardised forms and direct verbal translations, can be used as substitutes for speech and are produced and perceived consciously (Goldin-Meadow 1999: 419-420). Typical examples of emblematic gestures are the American thumbs-up sign and the Ok sign (Gullberg 2006: 105). Beats or beat gestures are also common to most classifications of gestures and constitute rhythmic elements that are commonly used by the speaker to control the pace of speech (Sueyoshi & Hardison 2005: 663) and visually emphasise important discourse elements (McNeill 1985: 354). A further common category is that of iconic gestures, also termed illustrators by Harris (2003: 182) and Ekman and Friesen (1969: 68), which carry a meaning that is parallel to the verbally expressed meaning (McNeill 1985: 356) and visually illustrate what is said (Sueyoshi & Hardison 2005: 663). According to McNeill (1985: 356), whose categorisation of gestures is the most widely quoted, metaphoric gestures are a sub-category of iconic gestures. Like iconic gestures, they are visual images of the verbal content, conveying, however, a more abstract idea or concept (McNeill 1985: 356). The last category that is present in most classifications is the category of deictic gestures, which serve to direct the attention of the interlocutor (Kelly *et al.* 1999: 579) as they point, for example, to concrete objects or people in the physical surroundings or in a more abstract sense to points in time or space (Sueyoshi & Hardison 2005: 663).

The relationship between gestures and speech production as well as perception is still unclear; however, it is indisputable that gestures and speech are linked (Gullberg 2006: 107). McNeill (1985) conducted a very influential study investigating this complex relationship and suggested that gestures and speech are linked internally (McNeill 1985: 353) and share a mental computational stage (McNeill 1985: 350). He describes the system of gesture production as being part of the psychological structure of speaking, which is contrary to the common idea that body language and language are always separate systems (McNeill 1985: 350-351). In support of his theory, McNeill (1985: 351-360) argued that

gestures and linguistic units are usually temporally synchronised in communication, with 90% of the gestures occurring simultaneously with the linguistic item of the same semantic and pragmatic content and 10% occurring in speaking pauses, usually with the intention of restarting the conversation. Moreover, McNeill (1985: 351) supported his idea that gesture production and language processes are linked internally with the argument that gestures develop in connection with speech in the ontogenesis of a person, which was also observed by Goldin-Meadow (1999: 423). Beyond this, he undergirded this theory by establishing that aphasia has a similar effect on the production of speech and gestures (McNeill 1985: 351), which was further suggested by Hadar *et al.* (1998: 109).

Goldin-Meadow (1999) studied the relationship of gestures and language production and perception, focusing especially on the functions that gestures fulfil in the communication process. She argued that gestures can support speech comprehension if they convey the same message as communicated verbally but hinder comprehension if the message they communicate conflicts with the verbal message (Goldin-Meadow 1999: 426). A further supporting factor is that “[g]estures also provide listeners with a second representational format, one that allows access to the unspoken thoughts of the speaker and thus enriches communication” (Goldin-Meadow 1999: 428). Concerning speech production, Goldin-Meadow suggested that gestures facilitate the retrieval of words from the mental lexicon, thereby reducing the cognitive load (Goldin-Meadow 1999: 427). This was also suggested by Hadar *et al.* (1998: 107).

Hadar *et al.* (1998) further explored the functions of iconic gestures in the process of language production, while Harris (2003) investigated their functions in language comprehension. Hadar *et al.* (1998) studied the production of iconic gestures accompanying the foreign language production of patients with deficits in retrieving words and patients who had conceptual deficits that affected the early process of constructing a message as well as the gesture production of an unimpaired control group. Their neuropsychological study led to the conclusion that gestures indeed serve to facilitate the retrieval of words (Hadar *et al.* 1998: 107) as the patient group that suffered from word retrieval deficiencies had a significantly higher production rate of gestures compared to the other patients and the control group (Hadar *et al.* 1998: 121). Harris’ (2003) study investigating the connection of iconic gestures and language comprehension revealed that the gestures themselves carry a definite message and support the decoding process of

speech in L2 listening (Harris 2003: 180). He concluded that gestures and other forms of NVC are “important information-carrying elements in communicative acts” and any limitation to perceive them would “impoverish or impair message reception” (Harris 2003: 181).

Iconic gestures, in particular those used to communicate experience, were also examined in a famous study conducted by Riseborough (1981). In her experiments the test subjects first had to guess objects from descriptions that were presented only acoustically, acoustically and visually by a speaker who did not perform any gestures, or acoustically and visually by a speaker who performed iconic gestures (Riseborough 1981: 173-174). In a second and third experiment the subjects had to recall words from lists of words and from a story presented in the aforementioned modes (Riseborough 1981: 175-178). The experiments revealed that the subjects performed significantly better at guessing objects when the audio-visual descriptions were accompanied by iconic gestures, which suggests that gestures facilitated the decoding of the linguistic message (Riseborough 1981: 174). Equally, in the second and third experiment the recall rate increased significantly in the presence of iconic gestures, which further supported the argument that gestures facilitate the retrieval of words from the mental lexicon (Riseborough 1981: 177-180).

Kelly *et al.* (1999) researched the role of gestures in communication, focusing on deictic and iconic gestures. They investigated the role of deictic gestures in pragmatically ambiguous indirect requests and observed that the study subjects were significantly more likely to understand the intention behind indirect requests if the requests were accompanied by deictic gestures (Kelly *et al.* 1999: 579-581). Concerning iconic gestures, they examined the effects of the absence or presence of gestures on the ability to recall the gist of a story and established that study participants who were presented with stories accompanied by iconic gestures significantly outperformed those participants who only heard the stories without the corresponding gestures (Kelly *et al.* 1999: 586-587). Based on their results, Kelly *et al.* (1999) concluded that the presence of gestures can have a positive effect on memory (Kelly *et al.* 1999: 584) and can support the comprehension of the speaker’s intention, as “speech and gesture mutually disambiguate the meanings of one another” (Kelly *et al.* 1999: 577).

### ***3.3.2. Lip movements and their effects on listening comprehension***

A further aspect of non-verbal communication that has been extensively researched is lip movements and their effects on listening comprehension. It has been confirmed by several researchers that lip movements have a considerable impact on speech perception (Ockey 2007: 519).

McGurk and MacDonald (1976) were the first to reveal the extent of this previously unrecognised impact. To counter the then prevalent assumption that speech perception is a purely auditory process, they conducted an experiment in which test subjects watched a film showing a person articulating the syllable [ga] while /ba/ was heard on the audio-track (McGurk & MacDonald 1976: 746)<sup>10</sup>. They found that most native-speaker subjects who perceived the correct sound when they were presented only with the auditory information reported to have heard the syllable /da/ when watching the film, revealing an integration of visual and auditory information (McGurk & MacDonald 1976: 746), which is nowadays called the “McGurk effect” (Hardison 1999: 217). In a second experiment, in which the articulation of [ba-ba], [ga-ga], [pa-pa] or [ka-ka] was shown on a video-tape while the test subjects heard /ga/ instead of [ba], /ba/ instead of [ga], /pa/ instead of [ka] and /ka/ instead of [pa], McGurk and MacDonald (1976: 746) revealed that there is an age difference in the influence of lip movements on speech comprehension. In an audio-only condition, adults outperformed primary school children and pre-school children with 99% correct perceptions of the syllable as compared to 97% and 91% (McGurk & MacDonald 1976: 746). However, in an audio-visual condition, adults had an error rate of 92% and were outperformed by school children and pre-schoolers, who had error rates of 52% and 59% respectively (McGurk & MacDonald 1976: 746). This led to the conclusion that adults are more likely to be influenced by visual information in their listening comprehension process (McGurk & MacDonald 1976: 747), which was later again suggested by Campbell and Dodd (1987: 85) and Burgoon (1994: 235).

A decade later, Campbell and Dodd (1987: 89) investigated the influence of visual information, in the form of lip movements, on the perception of consonants. They conducted an experiment in which test subjects were presented with lists of words in six different conditions, i.e. audio-only, vision-only, audio-visual with visual and auditive

---

<sup>10</sup> Square brackets are used to indicate articulated syllables and slashes to indicate phonemes.

input being presented in temporal synchrony and three audio-visual conditions in which the auditory input was presented with an increasing time delay in relation to the visual input (Campbell & Dodd 1987: 87-88). They found that the native-speaker test subjects performed best in the audio-visual condition in which auditive and visual information were presented in temporal synchrony (Campbell & Dodd 1987: 89). Despite the time delay of the auditive information in the three out-of-synchrony audio-visual conditions, the test subjects still performed significantly better in these conditions than when receiving information only auditorily or visually (Campbell & Dodd 1987: 90). Based on these results, Campbell and Dodd (1987: 89) argued that visual information derived from lip movements can increase listening comprehension performance.

Hardison (1999; 2003; 2005), whose studies supported previous results, followed in the research on the integration of visual and auditory input in speech perception. Her first investigation focused on the sound identification of non-native listeners living in the USA for less than six months (Hardison 1999). She established that also non-native speakers performed better in identifying speech sounds when videos showing the lip movements of the speakers were present (Hardison 1999: 214). This was particularly striking concerning the identification of speech sounds that were considered to be problematic based on the native tongue of the participants (Hardison 1999: 214), such as /f/ and /r/ for Japanese and Korean speakers (Hardison 1999: 228). Especially when the participants had to identify speech sounds pronounced in noisy conditions, the information derived from the visual input became more important and the subjects integrated visual and auditory information in the comprehension process, which confirmed the existence of the ‘McGurk effect’ (Hardison 1999: 253).

In a later study, Hardison (2003) further explored the influence of visual cues attained from lip movements on speech perception, focusing on the effects of perceptual training in which L2 speakers of English received either audio-visual training or audio-only training intended to improve their ability to identify speech sounds. Already a pre-test, conducted before the perceptual training commenced, indicated that the subjects made use of the visual cues derived from lip movements to improve their listening comprehension performance (Hardison 2003: 504), while the post-test clearly revealed that audio-visual perceptual training resulted in “significantly greater improvement in perceptual accuracy” and consequently better performance on the comprehension test (Hardison 2003: 505).

Hardison (2003: 514) emphasised that the study “demonstrated the role of the visual modality as a second channel of input for L2 perceptual learning”. These results were confirmed in a further study conducted by Hardison (2005) focusing on the influence of the same form of perceptual training on word identification by ESL students. This study revealed that the information derived from lip movements facilitated word identification and that ESL learners use audio-visual cues to their advantage (Hardison 2005: 593). The study led Hardison (2005: 579) to argue that visual cues, such as lip movements, provide crucial information about the perceived speech sounds and are particularly important when the auditory signal is degraded or hearing is impaired, as they can serve to compensate for the loss of auditory information.

This chapter explored the complex connection between non-verbal behaviour, kinesic behaviour, communication and listening comprehension. It revealed that while kinesics can be straightforwardly defined as the field of study investigating the use of gestures and facial expressions in a language system (Antes 1996: 439), conflicting definitions of non-verbal behaviour exist which either include or exclude vocalics and paralanguage. Despite this difference, non-verbal behaviour is always defined as including forms of behaviour that do not involve language production. Consequently, kinesic behaviour, which refers to body movements that are message-related, can be seen as a subcategory of the broader concept of non-verbal behaviour. Kinesic behaviour further forms an integral aspect of non-verbal communication, which was shown to be an important aspect of human communication. Although the exact nature of the relationship between non-verbal behaviour and communication in general and kinesic behaviour and listening comprehension in particular is still unclear, research nevertheless established that non-verbal behaviour and verbal communication are inextricably linked and that forms of kinesic behaviour, especially gestures and lip movements, can support listening comprehension of native as well as L2 listeners.

#### **4. Effects of the inclusion of visual input in teaching and testing L2 listening comprehension**

Although video materials were already widely used in second and foreign language teaching in the early 1990s, the impact of visual media on language learning and teaching was barely researched at that time (Gruba 1993: 85). Furthermore, the possible effects of the presence of visual cues on listening comprehension which were addressed in Chapter 3 were generally neglected in the theory of L2 and FL learning and teaching (Kellerman 1990: 272). Finally, despite their widespread use in the practice of language teaching, visual materials were hardly ever used in listening comprehension tests or any other form of language tests, which was largely due to a lack of research on the effects of visual input on language testing (Gruba 1993: 85).

While video media are nowadays even more widespread in language teaching than they were in the 1990s and form an integral part of everyday life, their use is still largely restricted to L2 or FL teaching and avoided in language testing. Nevertheless, a number of studies have been conducted exploring the effects of visual input in listening comprehension tests (e.g. Coniam 2001; Ginther 2001, 2002; Gruba 1993, 1994, 2014; Ockey 2007; Herron *et al.* 1995), yielding conflicting results concerning the impact of the use of visual input on L2 or FL listeners and test takers. Several researchers established that the visual materials used in listening tests have a positive effect on, for example, test taker performance (e.g. Ginther 2002; Herron *et al.* 1995; Jones 2003; Wagner 2010) and test authenticity (e.g. Bacon & Finnemann 1990; Suvorov 2011; Wagner 2002), while others argued that the presence of visuals in listening comprehension tests results in decreased test taker performance (e.g. Coniam 2001; Suvorov 2011) and has a negative impact on test validity (e.g. Buck 2001). Consequently, the impact of visual elements on listening comprehension remains a highly controversial topic in the field of research; however, it is nowadays widely acknowledged that “in the modern world the use of video in the classroom cannot be ignored” (Coniam 2001: 5).

This chapter outlines the positive and negative effects of the inclusion of visual input in teaching and testing listening comprehension, as established empirically, before discussing the controversy over the effects of the presence of visual materials on test validity.

## **4.1. Positive effects**

In the emerging body of research investigating the impact of visual input on listening comprehension, studies revealed so far more positive than negative effects of the use of visuals in listening teaching and testing. They showed, for example, that visuals can have positive effects on listener performance, increase the authenticity of listening texts, provide access to cues from non-verbal communication and contextualise listening texts.

### ***4.1.1. Positive effects of visual input on listener performance***

Already in the 1980s, Mueller (1980: 339-340) investigated the effects of visual input in listening tests and established that visuals, in this case line drawings which were presented to the test takers either before or after a listening passage significantly enhanced the performance of lower proficiency test takers compared to a control group that did not receive any visual input. This result was later confirmed by several studies exploring the use of visuals in listening tests (Ginther 2002; Shin 1998; Wagner 2010). Ginther (2002: 133-134) investigated listening comprehension items in the computer-based “Test of English as a Foreign Language” (TOEFL) that included still images depicting the context of the speaking situation, i.e. context visuals, or providing additional information about the content of the spoken text, i.e. content visuals, which were, for instance, diagrams or tables<sup>11</sup>. It was revealed that the content visuals, which enriched the imagery of the participants, facilitated listening comprehension, while the still images depicting the context of the speaking situation did not have a significant impact on listening comprehension (Ginther 2002: 157).

Moreover, Shin (1998: 74) conducted a study comparing an audio-mediated and a video-mediated listening test which revealed that the students taking the video-mediated listening comprehension test performed significantly better than the control group taking the traditional listening test. This study indicated that the students in the experimental group profited from the contextual cues provided by the video and used those cues to compensate for comprehension difficulties (Shin 1998: 74). This was further argued by Wagner (2010), who also compared the results of a traditional and a video-mediated listening test. Similarly to Shin’s study, the students taking the video-listening test scored significantly

---

<sup>11</sup> An in-depth explanation of the terms context visuals and content visuals can be found in Section 6.5.3.

higher than the control group (Wagner 2010: 439), which led Wagner (2010: 508) to conclude that the visual cues present in the video-mediated test contributed to the increased performance of the test takers.

The fact that cues derived from the visual input present in audio-visual listening tests can facilitate the comprehension of L2 listeners was also established by Sueyoshi and Hardison (2005), who studied the effects of facial expressions and gestures on the listening comprehension of low-intermediate and advanced ESL learners. They revealed that the study participants taking a test that was accompanied by a video showing either the face and the gestures or only the face of the speaker outperformed the group taking an audio-only listening test (Sueyoshi & Hardison 2005: 661). To explain the increase in student performance, Sueyoshi and Hardison (2005: 678) argued that the visual input possibly facilitated the memorising as well as the recalling of information, which was previously suggested by Jones (2003: 41).

Jones (2003) examined the effects of verbal and visual annotations accompanying an aural text which were intended to assist students in comprehending and acquiring vocabulary from spoken texts. Jones' (2003: 41) study revealed that the students were able to memorise and later recall word translations and the gist of the listening comprehension significantly better when presented with verbal and visual annotations accompanying the verbal text than when only hearing the text or being presented with either visual or verbal annotations.

Furthermore, two longitudinal studies (Herron *et al.* 1995; Secules *et al.* 1992) explored the effects of visual materials used in language teaching on the listening comprehension performance of L2 listeners over a longer period of time. Secules *et al.* (1992: 481) conducted the first longitudinal study, in which two control groups of students learning French as a foreign language were taught following a traditional curriculum for one semester, while two experimental groups learned French following the same curriculum but additionally watched an instructional video series called 'French in Action'. Secules *et al.* (1992: 482-483) found that the experimental classes scored significantly higher than the control groups in a post-treatment listening test and received similar results in reading and writing tests, a finding which was later confirmed by Herron *et al.* (1995: 784). However, vocabulary was learned more efficiently in the control groups, which followed traditional

instructional methods such as oral drills (Secules *et al.* 1992: 486). In their study, Herron *et al.* (1995) compared the learning outcome of an EFL language course which took place over the period of a year and was built on video-based language instruction to the language performance after a course that was based on the same curriculum but built on text-based instruction. It was revealed that the students' reading, speaking and grammar skills were very similar in both courses (Herron *et al.* 1995: 784). However, the group receiving video-based language instruction significantly outperformed the other students in listening comprehension at the end of the course and achieved slightly, albeit not significantly higher, results in a writing post-test (Herron *et al.* 1995: 784).

#### ***4.1.2. Increase in authenticity***

According to Buck (2001: 83), the term authentic denotes texts that are taken from TLU situations, which are “the real-world situation[s] in which the language [is] used” (Buck 2001: 83), or have characteristics of TLU texts (Buck 2001: 85). Widdowson (1978: 80), in turn, calls such texts genuine texts, whereas authentic means for him that the usage of a text in a testing situation corresponds to the actual usage of the text in the TLU situation. Irrespective of these differences in the definition of authenticity, various researchers (e.g. Buck 2001; Coniam 2001; Ginther 2001; Kellerman 1990; 1992) have highlighted that the inclusion of visual materials, which can either be seen as authentic or genuine texts, in listening teaching and testing leads to increased authenticity as more aspects of the TLU situation are replicated in the teaching or testing situation (Buck 2001: 111). This corresponds to Bachman and Palmer's (1996: 23) view on test authenticity which they argue is generally achieved when the performance on a test “corresponds to language use in specific domains other than the language test itself” (Bachman & Palmer 1996: 23).

Already in the early 1990s, Kellerman (1990: 272) pointed out that by using audio-tapes in listening comprehensions, teachers “are temporarily inflicting on students a handicap equivalent to the loss of sight”. The students are consequently forced to rely only on hearing, which corresponds to a limited range of listening situations outside the classroom such as listening to the radio, loudspeaker announcements or to an interlocutor on the telephone, which belong to the most difficult listening conditions for ESL and EFL learners (Kellerman 1990: 272). Consequently, Kellerman (1992) called the absence of visual information in listening comprehensions a “divergence from the norm” (Kellerman

1992: 245) and emphasised that audio-mediated listening tests cannot recreate “real-life” or authentic listening, except for the listening involved in the aforementioned TLU situations, in which visual information is naturally absent (Kellerman 1992: 251). Based on this argument, Kellerman (1990: 279) suggested that visual materials could be used for listening comprehensions in order to include “the visual information inherent in all normal speech situations” and because they constitute an “authentic mode of presentation”.

Kellerman’s (1990; 1992) call for authentic listening materials was supported at that time by Bacon and Finneman (1990), who conducted a survey examining the affective and cognitive responses of learners of Spanish as a foreign language to authentic oral and written input. They established that authentic materials can be used for the cognitive reason that they “provide the necessary context for appropriately relating form to meaning in the language acquisition process” and for the affective reason that they serve as motivators and support the overcoming of cultural barriers in language learning (Bacon & Finnemann 1990: 459). Furthermore, their study revealed that authentic oral input had a positive effect on the students’ level of satisfaction and on the development of listening comprehension skills (Bacon & Finnemann 1990: 469), which was later supported by Bachman and Palmer (1996: 32), who stated that the perception of a test as authentic can increase student performance and motivation.

More than a decade later, several researchers (e.g. Buck 2001: 172; Coniam 2001: 4; Ginther 2001: 1; Ockey 2007: 533; Wagner 2008: 220) noticed the superior authenticity of video listening tests in comparison to traditional audio-mediated listening tests. Buck (2001: 123) stated that presenting listening texts in video format is “a more realistic replication of real-world listening”. Suvorov (2011: 2) added that video-mediated listening tests are more authentic as they provide vital information about the context of the speaking situation, the discourse type and the role of the speakers. Furthermore, he argued that they grant access to aspects of the target language culture and to kinesic behaviour, whereby cues provided by gestures, facial expressions and other forms of body language can aid comprehension (Suvorov 2011: 2). Alike Kellerman (1990: 272), Ockey stressed, that in most real-life communicative situations information from the visual channel is present (Ockey 2007: 517) and video-listening comprehensions consequently relate a language test more directly to the TLU domain and make the inferences based on the test results more indicative of the actual performance in TLU situations (Ockey 2007: 533), a claim which

was supported by Ginther (2001: 1). Finally, Weyers (1999: 340) added that the inclusion of video materials in listening comprehensions also increases the students' amount of exposure to authentic target language.

#### ***4.1.3. Access to cues from non-verbal communication***

An important topic that is commonly discussed in connection with video-mediated listening tests is the effect of visual cues such as cues derived from kinesic behaviour, the physical appearance of the speakers and the situational context on ESL or EFL test takers. As discussed in Chapter 3 non-verbal communicative behaviour serves many important functions in communication and the information derived from non-verbal communication has a considerable impact on listening comprehension. Beyond this, several researchers, including Buck (2001), Kellerman (1992), Riseborough (1981), Sueyoshi and Hardison (2005) and Wagner (2002; 2008; 2010), established that cues from non-verbal behaviour not only influence but also assist the listening comprehension process.

One of the earliest studies that indicated a positive impact of visual cues on listening comprehension was the aforementioned study by Riseborough (1981: 178)<sup>12</sup>. It was found that the performance of the test takers who were presented with the auditory text accompanied by a video showing a speaker performing iconic gestures remained constant when noise was added to the audio track, while the scores of the other test takers, who only heard the listening or heard and saw a speaker who did not perform any gestures, dropped considerably (Riseborough 1981: 181). Based on this result, Riseborough (1981: 182) concluded that the listeners utilised the visual information derived from the speaker's gestures to compensate for comprehension difficulties when the auditory signal was degraded or incomplete.

This conclusion was later supported by Buck (2001: 50), Kellerman (1992: 247) and Gullberg (2006: 111). Buck (2001: 50-51) stated that the utilisation of non-verbal cues is an important aspect of the compensation skills of L2 or FL learners, warning, however, that non-verbal signals can be interpreted incorrectly by members of other cultures as they are culturally dependent. Kellerman (1992: 249-250) further suggested that cues from kinesic behaviour also assist in the processing of ambiguous discourse and comprehension

---

<sup>12</sup> For more information see Section 3.3.1.

in general as the redundancy in communication is increased and information can be derived not only from the auditory but also from the visual channel. Kellerman (1992: 249) also highlighted the importance of the presence of visual cues in EFL or ESL listening comprehensions as it allows for more realistic spoken discourse. While traditional listening comprehension texts sometimes need to be overly explicit to counter ambiguities, for example, concerning the addressing of interlocutors, which might not correspond to realistic target language use (Kellerman 1992: 249), the visuals present in video-listening comprehensions automatically disambiguate the spoken text. Moreover, Gullberg (2006: 116) argued for the inclusion of visual materials in listening comprehensions as they provide access to the gestural repertoire of the target language, knowledge of which she views as important for being communicatively competent and “culturally fluent” in a language.

Finally, Wagner (2002; 2008) conducted two studies the effects of the presence of visual cues in video-mediated listening comprehensions. Wagner’s first study (2002: 16-17) provided evidence that the listeners utilise cues from non-verbal communicative behaviour to assist their listening comprehension process, which led him to conclude that it is important to include the non-verbal cues that occur naturally in everyday communication in listening tests. In his second study, Wagner (2008) investigated how test takers utilise non-verbal cues in video-listening tests using a verbal report methodology. He found that the test takers frequently referred to cues derived from the visuals in their verbal reports<sup>13</sup>. Often the participants referred to visual cues, especially cues from gestures, unconsciously, for example, by copying the gestures a speaker produced, which might indicate “the automatic nature of the processing of the nonverbal information” (Wagner 2008: 230). The study revealed that the visual information affected the test takers’ comprehension as well as interpretation process and suggested that the listeners also used the incoming visual information to monitor these processes (Wagner 2008: 235-237).

#### ***4.1.4. Contextualisation of listening texts***

A further positive aspect of the presence of visuals in listening comprehensions noted in research is the contextual information they provide. Videos grant access to an abundance

---

<sup>13</sup> In these verbal reports the participants simultaneously expressed, how they are answering the test questions.

of “contextualised language” (Herron *et al.* 1995: 775), which means that they not only present the discourse that surrounds the sentences or utterances but also give insights into the sociolinguistic situation, which defines the form and functions of the spoken discourse to a considerable extent (Bachman 1990: 81-82).

Moreover, ambiguities which are often present in audio-mediated listening tests as the information, which is usually perceived auditorily and visually, is reduced to the modality of hearing alone are diminished in video-mediated listening comprehensions (Herron *et al.* 1995: 775). This is further supported by Kelly *et al.* (1999: 577) and Sherman (2003: 2), who stated that visual contextual information serves to disambiguate the pragmatic meaning of an utterance and supports its comprehension.

Furthermore, Buck (2001: 247), in his definition of the listening comprehension process, stated that the interpretation of the meaning of a spoken text, which constitutes an important aspect of the listening comprehension process, is influenced to a great extent by the context of the communicative situation. He notes that visual information supplements the verbal information and provides the context that is crucial to the interpretation of the spoken message (Buck 2001: 247).

#### ***4.1.5. Further positive effects of the inclusion of visual input in teaching and testing listening***

In research various other positive effects of the inclusion of video materials in listening testing are discussed. One positive effect, which is well-established in the field, is the alignment of teaching and testing practice through the introduction of video media in listening testing (Buck 2001: 172). Videos and video equipment were increasingly accessible from the 1970s onwards, becoming widespread in ESL and EFL teaching (Gruba 1997: 339). Nowadays visual media play an undisputable role in the communicative language teaching approach (Coniam 2001: 3), in which the main goal is communicative effectiveness in TLU situations and not necessarily absolute grammatical correctness in communication (Buck 2001: 83); however, they are still excluded from testing situations, which artificially dissociates assessment procedures from pedagogic practice.

Furthermore, as communicative tests, which are based on the communicative language teaching approach, are tests that “better simulate the characteristics of target-language use in the real world” (Buck 2001: 92), test developers set out to find genuine texts for language tests. This led to a disproportional use of broadcast materials in listening tests, which are readily accessible, however, unrepresentative of the range of different text types that exists in the target language world (Buck 2001: 87). The integration of video not only in language teaching but also in language testing allows for the presentation of an abundance of different text types, simulating the characteristics of TLU in the world beyond the classroom.

Motivation is a further positive effect of the integration of visual input in listening teaching as well as testing. Sherman (2003: 2) noted that the use of videos in the language classroom can increase the students’ motivation as videos are eye-catching, modern and authentic. Several studies (Progosh 1996; Sueyoshi & Hardison 2005; Wagner 2010) further established that videos can have a positive effect on the level of motivation. The test subjects in Sueyoshi and Hardison’s (2005: 661) study, which investigated the complex effects of visuals on listening comprehension, indicated a positive attitude towards the inclusion of visual input in listening comprehension tests. The participants in Progosh’s (1996) study, which explored the participants’ opinions on video-mediated tests, also clearly stated that they prefer audio-visual over audio-only listening tests (Progosh 1996: 34), with 91.9% of the test takers reporting to prefer video-mediated listening tests (Progosh 1996: 40). This preference for video-mediated listening comprehensions was later confirmed by Wagner (2010: 494).

Moreover, it was argued that visual materials included in teaching and testing English provide access to vital information about the target language culture. Sherman (2003: 2) argued that videos do not only serve as language models and linguistic resources but they function as a “window on English-language culture”. Kellerman (1992: 253) added that videos also offer insights into the kinesic behaviour of the target culture and can consequently increase the language learners’ familiarity with the non-verbal communication system of the target culture. In connection to this positive effect of visual materials, it was noted by Kellerman (1992: 253) and Antes (1996: 443), focusing however on L2 learners, that the knowledge of the kinesic behaviour of the target culture can be seen as a part of a learner’s communicative competence. Based on this argument they

emphasised the importance of including videos providing access to the kinesic behaviour of the target culture in language teaching (Antes 1996: 443; Kellerman 1992: 253).

Furthermore, it was argued that seeing the speaker and the speaking situation in video listening tests supports the activation of relevant background schemata and background knowledge (Buck 2001: 172; Suvorov 2011: 2). Finally, Progosh (1996: 38) also stated that these tests are likely to have a positive washback effect as they are both authentic and direct, which are two important aspects that can lead to beneficial washback (Messick 1996: 2).

#### **4.2. Negative effects**

Contrary to the above-mentioned studies, which revealed the positive effects of the inclusion of visual input in listening teaching and testing, some researchers have also raised doubts about the inclusion of visual input when testing listening (e.g. Buck 2001; Gruba 1997) and a few studies (e.g. Coniam 2001; Gruba 1994; Suvorov 2011) also revealed negative effects of this inclusion. However, these conflicting study results could be affected by a number of factors such as differences in the nature of the aural and the visual text (Ockey 2007: 520) and the text genres used as well as different question formats and varying proficiency levels of the study participants (Wagner 2008: 220).

Alongside his emphasis on the positive aspects of using visual input when teaching listening, Buck (2001) warns that the effects of videos on listening comprehension teaching and testing are yet unclear and barely researched (Buck 2001: 253). It is still doubtful whether the access to visual cues leads to increased comprehension (Buck 2001: 47) as studies revealed positive as well as negative effects on student performance. Furthermore, both Buck (2001: 173) and Wagner (2008: 219) remarked that there are still practical concerns when using visuals in listening tests because video-mediated listening tests require more resources in construction as well as administration. Shin (1998: 58) further noted that there is also a lack of practical guidelines for developing video-listening comprehension tests.

Moreover, Buck (2001) raised the concern that by integrating visual information in listening comprehension tests, which requires the test takers to attend to a video while

listening to the audio track, reading the test questions and writing responses (Buck 2001: 254), the cognitive load put on the test takers is increased considerably, which might affect the test results (Buck 2001: 123). This was supported by Gruba (1994) and Coniam (2001), whose studies indicated that the test takers were often overwhelmed by listening, writing and watching at the same time, which caused them to be distracted (Coniam 2001: 1) or abandon the visual input and focus mainly on listening (Gruba 1994: 34).

Gruba (1997: 337) argued that further problems with video-mediated listening tests are that test takers differ in their ability to interpret non-verbal cues and might interpret visual cues differently based on their cultural background. Furthermore, criteria for assessing visual competencies, such as the ability to interpret non-verbal signals, are so far non-existent, which still applies to the present day (Gruba 1997: 337). Based on these arguments, Gruba (1997: 341) noted that video-mediated listening tests face a reliability and validity problem because the inclusion of visuals in listening tests adds further relatively unpredictable variables to the test, which might affect its reliability and complicates its validation.

Suvorov's (2011) study further suggested a negative impact of visual input on test taker performance. Suvorov (2011: 6-7) compared an audio-mediated listening test to listening comprehensions accompanied by videos or still images and found that test takers performed equally well in traditional audio-only listening tests and listening comprehensions accompanied by photographs but significantly better than in video-mediated listening tests. This led Suvorov (2011: 7) to conclude that the use of video as a presentation mode had a negative impact on the participants' performance. To explain this negative impact, it was suggested that the visuals might have confused and distracted the test takers and impaired their listening comprehension process (Suvorov 2011: 2).

This was further established by Coniam (2001), who additionally investigated the participants' opinions on the different types of tests. Coniam's study did not reveal any significant difference between the performance of the participants taking the audio-only listening test and those taking the audio-visual test; however the majority of test subjects taking the audio-visual test reported that they did not gain any advantage from the visuals provided (Coniam 2001: 1). The test takers reported that they were distracted by having to focus on both the test booklet and the screen showing the video and had the impression that they would have performed better in an audio-only condition (Coniam 2001: 9). This

contrasts with the subjects taking the audio-only test, who predominately thought that they would have performed better in an audio-visual condition (Coniam 2001: 9). Concerning the test takers' preference, the study revealed that the majority of the audio-visual test group preferred having a traditional audio-only test (Coniam 2001: 9-10) which, in combination with the aforementioned results, induced Coniam (2001: 1) to conclude that the traditional audio-only mode of presentation should be preferred when testing listening.

### **4.3. Controversial effects relating to construct validity**

The most widely debated topic concerning the integration of visual input in listening tests is their effect on the validity of the test construct. Validity was defined by Messick (1996: 6) as “an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on the test scores or other modes of assessment”. Validity is further not seen as a property of the test or assessment instrument itself but rather of the test results and the inferences that can be made from these results (Messick 1996: 6). The test results themselves can have a certain degree of validity, as validity is a matter of degree and not absoluteness (Buck 2001: 195). For an assessment to be valid it is crucial that none of the important aspects of the focal construct is left out, ensuring the “construct validity” of the test (Messick 1996: 3), which is defined as “the extent to which the test measures the right construct” (Buck 2001: 1). There are two inadequate forms of construct operationalisation (Buck 2001: 94). On the one hand, there is the problem of “construct underrepresentation”, which entails an incomplete operationalisation of the test construct, meaning that something that is present in the construct is not measured in the test (Buck 2001: 94). On the other hand, there is the introduction of “construct-irrelevant variance”, meaning that the test accidentally measures abilities which are not included in the construct (Buck 2001: 94). Although a certain degree of construct underrepresentation as well as construct-irrelevant variance usually cannot be avoided in testing, a valid test should achieve the standard of “minimal construct underrepresentation” and focus on “construct-relevant variance” (Messick 1996: 4-5), whereby variance denotes the “variability between test takers” or “the extent to which the test takers are spread out from the average” (Buck 2001: 100).

So far the impact of the inclusion of visual input on the test validity of listening tests has been barely researched (Ockey 2007: 518). Evidence that video-listening tests are more or at least equally valid as traditional listening tests as they might allow to draw more accurate and generalisable conclusions about the listening skills of language learners is still non-existent (Wagner 2010: 495). Several researchers (Buck 2001: 172; Coniam 2001: 12; Wagner 2008: 219; Wagner 2010: 493) have suggested that video listening tests actually assess something beyond listening skills and hence abilities which are not included in the focal construct of the listening tests. This caused Wagner (2008: 219) to state that the construct validity of such types of listening tests is still a major concern in test development and Coniam (2001: 5) to emphasise that the inclusion of visuals in listening tests should be carefully considered, especially in high-stakes tests, as the validity of the test might be compromised. Buck (2001: 172) further assumed that language learners differ in their ability to attend to and utilise non-verbal cues to aid their listening comprehension process, which was later confirmed by the results of a study conducted by Wagner (2008: 218). Buck (2001: 172) argued that this negatively affects listening tests as any type of listening assessment should focus on the processing and utilisation of verbal and not visual information and stressed that test developers should be more “interested in the test takers’ language ability, rather than the ability to understand subtle visual information”.

However, Wagner (2007: 68-69) argued that Buck’s (2001) perspective is problematic when it comes to assessing communicative language ability as visual information is crucial in almost all communicative situations and non-verbal behaviour is an essential aspect of interpersonal communication. As communicative language tests are primarily aimed at allowing teachers or test administrators to draw conclusions about a learner’s ability to use language communicatively in a TLU situation (Bachman & Palmer 1996: 44), the characteristics of this situation should inform the development of a language test construct, which was also highlighted by Buck (2001: 112). If a language test has the purpose of measuring a person’s ability to listen to a telephone conversation or a radio programme, the inclusion of visual input in the listening test as well as visual competencies in the listening test construct would be inappropriate for the purpose of the test and would not replicate the TLU domain (Wagner 2010: 508). However, if a test should enable testers to make inferences, for example, about the candidates’ skills to listen to a face-to-face conversation, a lecture or a TV programme, in which visual information is naturally present, the exclusion of the visual channel from the listening test and its construct would

be inappropriate (Wagner 2010: 508) and “would introduce a source of invalidity into the assessment, [sic] because of construct underrepresentation” (Wagner 2002: 17). This insight led Wagner (2007: 68-69) to stress the importance of including visual skills in construct definitions of L2 listening comprehension. Furthermore, he argued that “the ability to understand the subtle visual information communicated by a speaker might in fact be a very important part of a person’s communicative language ability”, as it constitutes an integral aspect of most TLU situations and TLU tasks (Wagner 2010: 509). Based on this argument, Wagner (2010: 509) stated that the inclusion of visual input in a listening test might lead to construct-relevant variance, provided that the test is intended to measure the listening ability in a TLU situation in which visual information is usually present and that the abilities to interpret cues derived from the visual input are integrated in the test’s construct.

The argument that the general exclusion of visual information in listening comprehension tests might lead to construct underrepresentation is supported by several researchers (e.g. Burgoon 1994; Gruba 1997; Progoosh 1996; Shin: 1998). While Progoosh (1996: 35) stressed that in most communicative situations the listener is able to see the speaker, Burgoon (1985: 347) stated that the verbal and the visual channel “are inextricably intertwined in the communication of the total meaning of an interpersonal exchange” as well as in the comprehension of the communicated meaning. Gruba (1997: 338) further added that non-verbal elements of communication are a “key [to] comprehending the world”, while Shin (1998: 74) established that the test takers profited from visual contextual cues present in his video-mediated listening test. As the test subjects would deploy their ability to utilise non-verbal cues to aid their listening comprehension outside the testing situation in the TLU domain, Shin (1998: 74) argued that these abilities are construct relevant.

Based on Kellerman’s (1992: 249) finding that the language in audio-listening comprehensions sometimes has to be overly explicit to compensate for the lack of visual cues, Wagner (2002: 9) further argued that this “compensation for the lack of visual cues could be seen as an introduction of sources of invalidity into the assessment”. Moreover, there are hardly ever more than two speakers in traditional audio-mediated listening tests because of the difficulty for listeners to distinguish between the speakers, which poses a further threat to the validity of the assessment, as it considerably limits the selection of

listening text samples from TLU domains that can be used in listening comprehension tests (Wagner 2008: 237). Wagner (2008: 238) suggested that this problem could be overcome by including visual input in listening comprehensions, making them more representative of the broad variety of 'real life' listening situations.

Finally, while Buck (2001: 172) argued that including the visual channel in listening comprehension tests would favour learners who are better at utilising non-verbal cues, Wagner (2008: 238) countered that the exclusion of visual input from listening tests would create a disadvantage for these learners as it would deprive them of the visual cues which are usually present in almost all TLU listening situations. Consequently, the listening ability of the language learners who are adapted to processing and utilising non-verbal cues to aid listening comprehension would be underestimated and the inferences based on the audio-listening tests would not be generalisable and indicative of their listening comprehension skills in TLU situations (Wagner 2008: 238).

This chapter showed that the effects of visual input on teaching and testing listening in general and L2 listening comprehension in particular are still unclear. On the one hand, several studies revealed positive effects of the use of visual materials. Research showed that the inclusion of visual stimuli in a listening test can lead to an increase in test taker performance as the test takers make use of visual cues to assist listening comprehension. Moreover, longitudinal studies indicated that students profited from video-based language instruction and significantly improved their listening comprehension skills in comparison to students who are taught following a traditional curriculum. Furthermore, the chapter revealed that researchers agree to a large extent that the use of visual materials increases the authenticity of listening comprehensions and contextualises listening comprehension texts. On the other hand, the results of several studies suggested that the presence of visual input in a listening test can have a negative effect on the test takers' listening skills, largely because they feel distracted and overwhelmed by the difficult task of focusing not only on listening and answering test questions but also on watching test videos. Finally, the chapter revealed that the effects of visual input on the construct validity of listening tests are controversially debated. It was argued that video-mediated listening tests assess something beyond listening and hence skills that are not included in the test construct. Consequently, the inclusion of visual input in a listening test threatens its construct validity. Contrary to this point of view, it was argued that the inclusion of visuals in testing situations leads to

construct-relevant variance as visual information is naturally present in most listening situations outside the classroom and the ability to use cues derived from the visual input to assist listening comprehension constitutes an important and construct-relevant ability in listening comprehension.

## **5. Visual competencies involved in video-mediated listening tests and theories of listening comprehension**

Since high-quality video media became generally available in the 1980s, a trend to integrate videos in language courses, especially L2 courses, has started and language teachers have been increasingly urged by text book designers to incorporate visual materials in their teaching practice (Wagner 2007: 67; Herron *et al.* 1995: 775). This trend has continued until the present day and Wagner (2008: 218) predicted that the use of video in language teaching will increase in future years due to the “pervading influence of video technology and the Internet in daily life”.

Buck (2001: 172) argued that the integration of visual input in listening comprehension tests, which would conform with this ongoing trend, would align current teaching and testing practice. However, the use of videos is still refrained from in testing listening today, as the effects of visual input on listening comprehension are still unclear, as discussed in Chapter 4, and because the visual competencies that are involved in listening comprehension in TLU situations which provide the listener not only with auditive but also with visual information are largely excluded from theories of listening comprehension, which forms the basis for listening test development.

This chapter outlines the work of pioneers who took the first steps in the direction of including non-verbal communication and visual competencies in the theory of language ability and listening comprehension and provides a suggestion for a listening comprehension construct for B2 listening that combines traditional listening comprehension skills and subskills with visual competencies.

### **5.1. Integrating non-verbal communication and visual competencies in theories of language ability**

As mentioned in Section 4.3., it is commonly argued that video-mediated listening tests assess something beyond listening (Buck 2001: 172; Coniam 2001: 12; Wagner 2008: 219; Wagner 2010: 493) which threatens the validity of an assessment that is based on a listening comprehension construct excluding skills above and beyond listening. Based on this argument, a small number of researchers (Canale & Swain 1980; Wagner 2007; 2008)

and organisations (ACTFL 2015; Council of Europe 2001) set out to describe these particular skills and integrate them in the theory of language ability. These skills, which had not yet been formally defined or labelled, were described, for example, as “the ability to understand subtle visual information” (Buck 2001: 172) and the ability to utilise information derived from non-verbal cues to aid comprehension (Wagner 2010: 509). These skills involved in video-mediated listening tests, which will be referred to as visual competencies in the context of this diploma thesis, can be seen as an aspect of visual literacy, which belongs to the so-called “Digital-Age Skills” (Dallow 2010: 98). Visual literacy has been defined as “the ability to interpret, use, appreciate, and create image and video using both conventional and 21<sup>st</sup> century media in ways that advance thinking, decision making, communication, and learning” (Dallow 2010: 98).

Already in the 1980s, Canale and Swain (1980) integrated such visual competencies in a description of language competence. They developed a “theoretical framework for communicative competence” in which communicative competence consists of grammatical, sociolinguistic and strategic competence (Canale & Swain 1980: 28). Besides describing communication as consisting of verbal as well as non-verbal signals (Canale & Swain 1980: 29), they defined strategic competence as being “made up of verbal and non-verbal communication strategies that may be called into action to compensate for breakdowns in communication due to performance variables or [...] insufficient competence” (Canale & Swain 1980: 30).

More than two decades later further important steps in the direction of including non-verbal communication and visual competencies in the theory of language ability were taken by the Council of Europe and almost two decades later again by the American Council on the Teaching of Foreign Languages (ACTFL). The Council of Europe (2001), which developed the influential CEFR, acknowledged the importance of non-verbal input in communication and integrated visual competencies in their descriptions of language competence. The council views non-verbal communication as being a crucial aspect of ‘communicative language activities and strategies’, which are important to carry out communicative language tasks (Council of Europe 2001: 57). Furthermore, the interplay between language skills and visual competencies was highlighted by stating that C2-level interaction requires interacting “with ease and skill, picking up and using non-verbal and intonational cues apparently effortlessly” (Council of Europe 2001: 28).

The ACTFL, which also developed performance descriptors for language learners that are especially important in the American educational context, took the integration of visual competencies in the theory of language ability a step further. They incorporated the utilisation of non-verbal cues in several different categories of language competence, as, for example, in the aspect of “Language Control” in interpersonal communication for novice language learners, which was described as the student “[c]an usually comprehend highly practiced and basic messages when supported by visual or contextual clues, redundancy or restatement, and when the message contains familiar structures” (ACTFL 2015: 15). Furthermore, the council views the utilisation of kinesic behaviour as an important aspect of the communication strategies used in interpersonal communication and states that novice learners may use gestures and facial expressions to maintain communication (ACTFL 2015: 15). Moreover, the production as well as the comprehension of non-verbal behaviour are vital aspects of “Cultural Awareness”, which constitutes a further aspect of interpersonal communication (ACTFL 2015: 15). While novice learners are simply required to use highly practised “culturally appropriate gestures”, intermediate learners should further recognise gestures in daily interactions and advanced language learners should use “cultural knowledge to conform linguistically and behaviorally in many social and work-related interactions” (ACTFL 2015: 15). The ACTFL (2015: 16) further emphasises the importance of visual competencies in interpreting language and states that novice learners should understand “words, phrases, and formulaic language that have been practiced and memorized [...] with strong visual support”. Finally, novice as well as intermediate and advanced learners should also use visual support to comprehend spoken texts, which is seen as further important strategy in interpreting language (ACTFL 2015: 17).

In addition to this fundamental pioneering work, Wagner (2010) took existing theories of language ability, especially Bachman and Palmer’s (1996) influential framework for describing language knowledge, as a basis for considerations on how to integrate aspects of non-verbal communication and visual competencies. Wagner (2010: 509) argued that the ability to utilise non-verbal cues can be interpreted as a part of pragmatic knowledge, defined by Bachman and Palmer (1996: 69) as an aspect of language knowledge that allows the language learner to “create or interpret discourse by relating utterances or sentences and texts to their meanings, to the intentions of language users, and to relevant

characteristics of the language use setting”<sup>14</sup>. In the context of this definition, he argued that listeners not only use their pragmatic knowledge to relate linguistic messages to their meanings and to the intentions of the speaker but also interpret non-verbal communicative behaviour by using pragmatic knowledge (Wagner 2010: 509). Furthermore, Wagner (2008: 239) argued that the utilisation of non-verbal cues might support the listener in relating spoken discourse to the intentions of a speaker and thereby promote the pragmatic competence of the listener. This was supported by Gullberg (2006) and Kelly *et al.* (1999), who established that gestures are inextricably linked to the speaker’s intention of expressing something (Gullberg 2006: 104) and that the utilisation of non-verbal cues can support the comprehension of the speaker’s intention (Kelly *et al.* 1999: 577)<sup>15</sup>.

## **5.2. Integrating non-verbal communication and visual competencies in theories of listening comprehension**

Alongside the sparse endeavours to integrate non-verbal communication and visual competencies in theories of language ability, an even smaller number of researchers has so far attempted to incorporate non-verbal communication and visual competencies in theories of listening comprehension. While Richards (1983) and Buck (2001) included non-verbal input and visual competencies in the general theory of listening comprehension, Gruba (1997), Ockey (2007), Suvorov (2011) and Wagner (2002) focused on the integration of visual competencies in construct definitions for listening comprehension tests.

The first step towards including visual competencies in the theory of listening comprehension was taken by Richards (1983), who included non-verbal communication skills in his taxonomy of listening comprehension skills<sup>16</sup>, which was discussed in section 2.3. Richards (1983: 229) included the “ability to make use of facial, paralinguistic, and other clues to work out meanings” and the “ability to signal comprehension or lack of comprehension, verbally and non-verbally” in his list of the micro-skills involved in conversational listening. Concerning microskills involved in academic listening, Richards (1983: 230) argued that listeners in academic listening situations need to have the “ability

---

<sup>14</sup> For further information concerning Bachman and Palmer’s (1996) framework for describing language knowledge see Section 2.2.

<sup>15</sup> For further information see Section 3.3.1.

<sup>16</sup> Richards’ (1983) entire taxonomy of listening comprehension skills can be found in Appendix 2.

to follow different modes of lecturing” such as “spoken, audio, audio-visual” modes and need to be able to recognise the functions of “non-verbal cues as markers of emphasis and attitude”. Furthermore, it can be suggested that the ability to process and utilise non-verbal cues, which Bostrom (1997: 27, cited in Buck 2001: 47) argued to provide more accurate information about the speaker’s emotions and attitudes than verbal cues, forms an important aspect of detecting the attitude of a speaker towards a topic, which also belongs to Richards’ (1983: 229) list of academic listening micro-skills.

Despite Richards’ (1983) pioneering work in integrating visual competencies in a taxonomy of listening comprehension skills, it was not until the mid-1990s, when the number of studies investigating video-listening tests increased, that the integration of visual competencies in theories of listening comprehensions slowly developed into an area of interest and a matter of debate in the field of research. Gruba (1997: 337) was the first to contemplate the lack of acknowledgement of the utilisation of non-verbal cues in definitions of listening comprehension and listening comprehension constructs.

Only at the beginning of the new millennium did a definition of listening comprehension emerge that stressed the important role of non-linguistic input in listening comprehension. This definition was Buck’s (2001: 104) framework for describing listening ability, presented in Section 2.2. In this widely accepted framework Buck (2001: 104) highlighted that the comprehension processes as well as the storing and memory processes, which form a vital part of the strategic competence of a language user, not only involve the processing and storing of linguistic input but also of non-linguistic input. Although this framework for describing listening ability can be seen as a crucial step towards the integration of the visual modality in definitions of listening comprehension, it focused only on the role of non-linguistic input in listening comprehension and did not yet include visual competencies that might be involved.

Generally visual competencies continued to be overlooked and neglected in theories of listening comprehension, which was further noted by Wagner (2007: 67) and Suvorov (2011: 2). This general neglect is reflected in the limited use of visual input in testing situations, which is partly due to the still unclear role of visual competencies in listening comprehension in video-mediated listening tests (Suvorov 2011: 2). This avoidance of visual input in listening tests is further influenced by the fear of test developers and

researchers that the inclusion of visuals will lead to construct-irrelevant variance, as something beyond the skill of listening comprehension is measured (Suvorov 2011: 2), as discussed in Section 4.3. To counter this potential negative effect on the construct validity of video-mediated listening tests, Wagner (2002: 8) argued that the visual clues provided by the videos have to be taken account of in the listening comprehension constructs deployed in video listening tests. This was further supported by Gruba (2014: 1004), who suggested that “language assessments that make use of new media and technologies may require both a reworking of established constructs and an identification of emerging ones to better acknowledge candidate performance of new literacies”, and Ockey (2007: 532), who suggested an expansion of current construct definitions of listening comprehension.

Although the absence of visual competencies in the construct definitions used in video-mediated listening tests was discussed at different points in time, only Gruba (1997) and Ockey (2007) have provided concrete suggestions so far on how to integrate these complex skills in listening constructs. Gruba (1997: 337) argued that if a listening test is presented in a visual mode of presentation, Rubin’s (1995) definition of listening comprehension can serve as a basis for the listening test construct. This definition describes listening as “an active process in which listeners select and interpret information which comes from auditory and visual cues in order to define what is going on and what the speakers are trying to express” (Rubin 1995: 7). Ockey (2007: 517), who conducted a study that compared the effects of still images and video input on non-native listeners and investigated how the test takers engaged with the different modes of presentation, took the integration of visual competencies in listening constructs a step further. He found that the test takers engaged only to a very limited extent and in similar ways with still images but varied widely in the ways in which and the extent to which they engaged with the video input (Ockey 2007: 517). Based on these results, Ockey (2007: 532) argued that the constructs of listening comprehension tests which are accompanied by still images do not necessarily need to be adapted, while the inclusion of videos in listening tests requires a rethinking of existing listening constructs as it results in “measuring test takers’ abilities to use visual cues to aid comprehension” (Ockey 2007: 532). To avoid construct-irrelevant variance resulting from the test takers’ varying visual competencies, Ockey (2007: 532) advocates an “enlarged academic listening construct” that should account for the visual information “present in the ‘cognitive environment’” of the test takers. This enlarged construct should include the test takers’ ability “to gain information from lip movement,

facial gestures, and other body language such as hand movements” and might additionally include the “ability to simultaneously use visual cues, process audio input, and take notes” (Ockey 2007: 532). Ockey (2007: 532) finally concludes that such a new construct definition of listening comprehension would correspond to the current state of knowledge concerning the process of listening comprehension, as it is currently assumed that this process depends on everything in the cognitive environment of the listener that might influence the perception of the message, which also includes visual information in almost all listening situations.

### **5.3. A modified B2 listening comprehension construct for video-mediated listening tests**

Based on insights gained from the above-mentioned approaches and ideas on how to integrate visual competencies in the theory of language ability and listening comprehension which are still in their initial stages a modified construct definition of B2 listening comprehension which combines listening skills and visual competencies was designed in the context of this diploma thesis.

This modified listening construct, which is based on the construct definition of B2 listening (Universität Innsbruck 2012) developed on behalf of the bifie (2016), was further used as a basis of the video-mediated listening comprehension test that was compared with an audio-only listening comprehension test in this diploma thesis study<sup>17</sup>. Furthermore, this construct definition of listening comprehension (Figure 4) can serve as a basis for other video-mediated listening tests and can be modified to fit the purpose of different audio-visual listening tests and the TLU situations reproduced in these tests.

#### **General purpose**

- a) listening for gist
- b) listening for specific information
- c) listening for important details
- d) search listening
- e) listening for main ideas and supporting details
- f) listening to infer (propositional) meaning
- g) listening to deduce the meaning of a word

<sup>17</sup> This is explained in greater detail in Chapter 6. For the results of this comparison see Chapter 7.

### **Specific purpose**

1. Can understand the main ideas of propositionally and linguistically complex speech on both concrete and abstract topics delivered in a standard dialect, including technical discussions in his/her field of specialisation.
2. Can follow extended speech and complex lines of argument provided the topic is reasonably familiar and the direction of the talk is sign posted by explicit markers.
3. Can with some effort catch much of what is said around him/her, but may find it difficult to participate effectively in discussion with several native speakers who do not modify their speech in any way.
4. Can follow the essentials of lectures, talks and reports and other forms of academic / professional presentation which are propositionally and linguistically complex.
5. Can understand announcements and messages on concrete and abstract topics spoken in standard dialect at normal speed.
6. Can understand most radio documentaries and most other recorded or broadcast material delivered in standard dialect and can identify the speaker's mood, tone etc.
7. Can use a variety of strategies to achieve comprehension including listening for main points and checking comprehension by using contextual clues.

### **Visual competencies involved in B2-level listening comprehension in audio-visual listening conditions**

Language learners who have reached B2 level in listening comprehension have acquired the following visual competencies:

- I. Can utilise information derived from the gestures of a speaker to support their comprehension of the spoken message.
- II. Can utilise information derived from the facial expressions of a speaker to support their comprehension of the spoken message.
- III. Can utilise information derived from the lip movements of a speaker to support their comprehension of the spoken message.
- IV. Can utilise information derived from the visual context of the speaking situation to support their comprehension of the spoken message.
- V. Can gain information about the intentions, emotions and attitudes of a speaker from his or her gestures and facial expressions.
- VI. Can utilise information derived from lip movements to assist speech perception.
- VII. Can gain information about the roles of different speakers in a conversation, their interpersonal relationship and the discourse type of the spoken text from the visual context of the speaking situation.
- VIII. Can simultaneously process and utilise auditory and visual information and answer test questions.

**Figure 4 A modified B2 listening comprehension construct for video-mediated listening tests  
- based on the bific B2 listening test construct (Universität Innsbruck 2012)**

## **6. Methodology**

After describing the setting and the participants, this chapter outlines individual aspects of the procedure involved in this study. The study included a teacher interview, a first testing session, consisting of a baseline listening test and a student questionnaire, and a second testing session, consisting either of an audio-visual (AV) listening test and a student questionnaire or an audio-only (AO) test and a different student questionnaire. Concerning the second testing session, the AO and AV tests and the test specifications on which these tests were based are explained in greater detail and the visual materials accompanying the AV test are described.

### **6.1. Study setting and participants**

The study was conducted in an upper secondary school in Vöcklabruck, in Upper Austria, in two 7<sup>th</sup> grade classes consisting of 50 students, 36 female and 14 male, who were split into three groups and taught by three different teachers. Group A consisted of 18 students, group B of 17 and group C of 15 students. All of the participants had German as their mother tongue, were 16 or 17 years old and learned English as a foreign language. Consequently, they represented a rather homogenous test group.

According to the Austrian curriculum for modern foreign languages, the listening skills of students attending the 7<sup>th</sup> grade of an upper secondary school are at B2 level (Bundesministerium für Bildung und Frauen 2015: 6). This means that, according to the CEFR (Council of Europe 2001: 27), the students are able to understand “extended speech and lectures and follow even complex lines of argument provided that the topic is reasonably familiar”, “understand most TV news and current affairs programmes” and “the majority of films in standard dialect”.

The first student questionnaire revealed that the students mainly come into contact with English in the same situations, which further underlined the homogeneity of the test group. The test subjects are mainly in contact with English in school (100%), through internet surfing (95.9%), watching TV (73.5%) and to a lesser degree also through reading English books (53.1%) and reading or writing, for example, e-mail, letters and blogs (40.8%), as illustrated in Table 1 below.

**Table 1 Test takers' contact with English**

<b>Qu. 1 item 3: "In which situations are you in contact with English?"<sup>18</sup></b>	<b>'yes' in %<sup>a</sup></b>	<b>'no' in %<sup>b</sup></b>
in school	100	0
with friends	26.5	73.5
with family	4.1	95.9
internet surfing	95.9	4.1
reading or writing e-mails, letters, blogs ...	40.8	59.2
reading books	53.1	46.9
reading the newspaper	8.2	91.8
listening to the radio	36.7	63.3
watching TV	73.5	26.5

<sup>a</sup> Percentage of students who reported to be in contact with English in the following situations.

<sup>b</sup> Percentage of students who did not report to be in contact with English in the following situations.

## 6.2. Procedure

Prior to the study, the headmaster and the classes' English teachers were asked for permission to conduct the diploma thesis study. The students and their parents were informed about the purpose and the procedure of the study and received some introductory information concerning the research on video-mediated listening tests in the form of a letter to the parents, which can be found in Appendix 3, and agreed on the students' participation in the study.

First a short teacher interview was conducted, which was intended to establish the teachers' and students' prior experience with videos in English language teaching and learning as well as with videos in listening tests. Afterwards the two main parts of the study, namely the first and the second testing session, were conducted in altogether six English lessons, one lesson per group in the first and second session.

As part of the first testing session, the baseline listening test established whether there was a difference in the level of listening ability of the participating groups. It was followed by the first student questionnaire, which asked the participants if they were already familiar

---

<sup>18</sup> Qu. stands for questionnaire. Qu. 1 refers to the questionnaire conducted after the baseline listening test in the first testing session.

with the selected listening tasks, requested them to provide personal background information and established their learning style.

The second testing session in this study comprised two tests, an AV and an AO listening test that involved the same listening tasks and test items as the AV test, however, without the provision of visual input. These tests provided the basis for comparisons of student performance in the presence or absence of visual cues. Following the AV and the AO listening test, the students received further student questionnaires. These questionnaires investigated the importance the students attach to visual clues in everyday conversation, the attention paid to the videos in the AV listening test and their opinion on the inclusion of visual input in listening tests and established whether they had taken any of the selected listening comprehensions before.

### **6.3. Teacher interview**

The teacher interview was conducted shortly before the baseline tests and the first student questionnaire. The three English teachers were asked the following questions:

- *How often do you use video materials in teaching English?*
- *Which skills do you usually target when using video materials?*
- *Did you ever administer a video-listening test in your group?*
- *If yes, how often did you administer video-listening tests in your group?*

The interview established that groups A and C had a rather limited experience with video materials in English lessons, while the teacher of group B reported that she uses videos quite often, usually as a means to introduce a topic. When using videos materials in English language teaching, the teachers reported that they focus mainly on content aspects of, for example, films, trailers or music videos and never specifically focused on training listening skills with the help of visual input. Concerning the administration of video-mediated listening tests, all teachers stated that they had never administered such a test before.

### **6.4. First testing session**

The first testing session took place after the teacher interview had been conducted. For this testing session, the participants received a test booklet containing the baseline listening

tests and the first student questionnaire, which can be found in Appendix 4, followed by the correction key and the test specifications, including the listening comprehension construct on which the test was based. Of the 50 students altogether in the three groups, 48 participated in this testing session.

#### **6.4.1. Baseline listening test**

The baseline listening test was a shortened version of the standardised and pretested English Matura from the second main test date on 10<sup>th</sup> May 2011, provided by bifie under the link: <https://www.bifie.at/node/800> (bifie 2011). Because of the time constraints of an English lesson, only the first three of the originally four tasks were presented in order to allow time to conduct the first student questionnaire.

All of the listening tasks were played twice<sup>19</sup>. The students had 45 seconds' time before hearing the first recording to study the listening tasks, which were a multiple-choice, a matching and a limited production task. Afterwards all three listening tasks were played twice with 20 seconds between the texts to complete the preceding task and 45 seconds after the last recording was played the second time to complete the tasks and check the answers. Altogether the listening test took 32 minutes and contained 23 test items.

#### **6.4.2. First student questionnaire**

The first student questionnaire contained 24 items. Items 1 to 3 asked for personal background information, i.e. gender, native language or languages, in case the participants have several mother tongues, and the situations in which the students are in contact with

---

<sup>19</sup> This was decided after considering the advantages and disadvantages of playing a task once or twice in a testing situation. Buck (2001: 171) argues that there are equally convincing arguments for playing a listening tasks only once or twice, but stresses that the ultimate decision depends on the purpose of the test and the testing situation. A striking argument for playing the task only once is that it has a higher degree of interactional and situational authenticity (Buck 2001: 171), with the former denoting that the knowledge and the competences needed for the test tasks correspond to those specified in the construct (Buck 2001: 126) and the latter referring to the correspondence of the test task to 'real-world' listening in TLU settings (Buck 2001: 171-172), where the listeners usually also hear a text only once with no chance of reviewing. Consequently, listening tests which involve playing the tape only once favour the automatic and fast processing of spoken language which is necessary in real-world listening. However, playing the task twice reduces the difficulty of test (Buck 2001: 170) and corresponds to the standard listening test procedures in the Austrian school framework, for example, the Matura listening exam procedures, and consequently constitutes what the students are used to from other testing situations. For these reasons and because of the potential threat of introducing construct-irrelevant variance by using a testing method with which the students are rather unfamiliar, it was decided to play the recording twice.

English. The next item asked the student whether they were already familiar with this listening test, which is freely available on the bifie homepage. Three students reported that they are familiar and one that he is partly familiar with the listening comprehension test. However, since these students did not perform better than the other test takers, which indicates that they remembered the listening test at most vaguely, they were left in the calculations presented in Chapter 7.

The next 20 items were intended to establish which types of learners the participants are, i.e. visual, auditive or kinaesthetic. These questionnaire items were taken from the website “Education Planner.org” (2011) and were slightly adapted to fit the age group and the estimated vocabulary knowledge of the test subjects.

## **6.5. Second testing session**

In this testing session, one of the three student groups, group A, served as the control group and took the AO listening test, while groups B and C participated in the AV listening test. Of the 50 study participants, 47 were present in this testing session; 18 of them took the AO test, while 29 students took the AV test.

Similar to the procedure in the first testing session, the test takers in both groups received test booklets at the beginning of the experiment. The test takers in the control group received a test booklet which included the AO listening test and a student questionnaire adapted to the audio-only condition of the listening comprehension. The two experimental test groups, on the other hand, received a test booklet which contained the AV listening test and a questionnaire that corresponded to the audio-visual mode of presentation. The test booklets, the correction key for the listening tasks as well as the test specifications operationalised in the tests can be found in Appendix 5 and 6.

### **6.5.1. *Audio-only and audio-visual listening test***

The AV and the AO test, which were the same listening test with and without the visual input, consisted of standardised and pretested Matura listening comprehension tasks taken from the bifie (2016) homepage like the baseline listening test. However, these Matura listening comprehension tasks were originally based on videos of which only the sound track was used as input in the Matura listening comprehensions. For the purpose of this

diploma thesis, the original videos were retrieved and used as input in the AV test, while the sound track of the videos was used for the AO test.

The AO and the AV listening test contained three listening tasks which consisted of 22 test items altogether and took 30 minutes to conduct. Similar to the procedure in the baseline listening test and in Austrian Matura listening exams, the recordings were played twice with 45 seconds' time before and after the listening texts were played to study or complete the tasks and 20 seconds between the tasks.

The first listening task was a limited production task that can be found under the following link: <https://www.bifie.at/node/2800> (bifie 2015a). The second and the third listening task were multiple choice tasks and can be accessed via the following links: <https://www.bifie.at/node/2837> (bifie 2015b) concerning the second listening task and <https://www.bifie.at/node/1830> (bifie 2012) concerning the third task. The corresponding videos that were presented in the AV listening test are available under the links <https://www.youtube.com/watch?v=uSJW-MC7jDM> (YouTube 2012), [https://www.ted.com/talks/pete\\_alcorn\\_s\\_vision\\_of\\_a\\_better\\_world#t-209334](https://www.ted.com/talks/pete_alcorn_s_vision_of_a_better_world#t-209334) (Alcorn 2009) and <https://www.youtube.com/watch?v=sr07uR75Qk0> (Jobs 2011).

The items were the same for the control and the experimental groups and could be answered without seeing the visuals. In order to determine to what extent the visual input could have helped in answering the listening test items, the videos were analysed for non-verbal cues that could have a positive effect on the students' listening comprehension. This will be described in further detail in Section 6.5.3.

### ***6.5.2. Test specifications operationalised in the audio-only and the audio-visual listening test***

Two listening comprehension constructs, which corresponded to the proficiency level of the students, were defined and integrated in the test specifications for both tests. As the listening tasks were the same for both tests, apart from the provision of visual input in the AV test, the test specifications and the listening test constructs differed only in one aspect, namely the inclusion of visual competencies in the listening test construct.

The listening comprehension constructs used in these tests were based on the construct definition of B2 listening (Universität Innsbruck 2012) developed for bifie as this is the listening comprehension construct operationalised in Matura listening tests which were used in the AO and the AV listening test. The bifie listening construct was, however, modified to fit the purpose of the listening tests, the TLU situation and the audio-visual listening condition in the AV test.

The test specifications for the AO and the AV test, which contain the two different construct definitions of listening comprehension and “make the theoretical framework which underlies the test[s] explicit” (Alderson, Clapham & Wall 1995: 17), were based on the test specifications operationalised in the Matura listening tests (Universität Innsbruck 2012) as well as Alderson, Clapham and Wall’s (1995: 11-15) and Bachman and Palmer’s (1996: 48-49) guidelines to draw up test specifications.

### ***6.5.3. Description and analysis of the visuals used in the audio-visual listening test***

In the description of visual materials used in listening tests, a general distinction is made between context visuals and content visuals (Ginther 2001: 2; Ockey 2007: 520; Suvorov 2011: 3). Context visuals provide “information about the context in which the verbal exchanges occur” (Ginther 2001: 2) and “set the scene for the communication” (Ockey 2007: 520). They complement the information perceived on the verbal channel as they provide information about, for instance, the setting, the appearance of the participants, their facial expressions, lip movements and gestures, and further indicate a change in speakers (Ginther 2001: 2). A typical example of a video that belongs in this category would be one depicting a conversation between two people which shows only the two speakers and their physical surroundings.

Content visuals, in contrast, complement the aural information (Ockey 2007: 520) and provide information that is related to the content of the spoken message (Suvorov 2011: 3). They illustrate what it said, for example, by showing images, graphs and tables (Ockey 2007: 520). An example of this type of visual would be a video of a lecture focusing not on the lecturer and the audience but displaying the PowerPoint slides that depict important content elements of the lecture. However, most visual materials, especially authentic videos, do not simply fall into one category as they present a mixture of contextual as well

as content information. An example of this mixed type of visual material would be a TV news report showing not only the news reporter and his or her physical surroundings but also images, videos or charts related to the content of the news report.

For the AV test in this diploma thesis study, three different videos were used. The first video was a TV report called “Comic Books Go Digital” (YouTube 2012), which dealt with the introduction of digital comic books in the comic book business. These digital comic books as well as their effects on the comic books business and on the readership were described by several speakers, who were either developers of comic books, owners of comic book stores or comic book readers. This video can be seen as a mixture of context and content visuals as the video sometimes focused only on the speakers and the setting, which is mostly a comic book store, and sometimes visually illustrated the content of what is said, for example, by showing traditional comic books or digital comic books. An analysis of the video, which was intended to establish whether visual cues provided by this video could have been useful for the test takers, revealed that the presence of visual cues could have helped answer several test questions.

The first test item, which asked what was recently introduced in a comic book store, might have been easier in the presence of the video as the video showed a speaker clearly pronouncing the answer, whereby the lip movements of the speaker might have supported speech perception. Furthermore, after the speaker revealed the answer to the test item, which was “digital comic books” (YouTube 2012: 0:24 min.), the video showed such new comic books. Apart from this, also the facial expressions of the speaker, especially his eye and eyebrow movements, might have assisted the students’ comprehension as they indicated which aspects of the spoken message were important.

While providing no assistance in answering the second, sixth and seventh test item, as the answers to these items were given by a speaker in voice-over, while the accompanying video footage did not directly relate to the test item answers, the video might have facilitated answering items three, four and five. Concerning the third test item, a commentator said that Mark Waid’s comic books “fit the digital screen” (YouTube 2012: 1:15 min.), which is the answer to the test item, while such a digital screen was shown. Afterwards, Mark Waid described the format of his digital comic books, using hand gestures to demonstrate the format of a typical digital screen. What could have been of

special help concerning this test item is that the name of the new speaker, Mark Waid, which also appears in the test question, was shown at the bottom of the video to introduce the speaker. This could have indicated for the test takers to move forward and focus on test item three.

Regarding item four, the commentator said in voice-over that readers can get extra historical information from digital comic books, thereby providing the answer to the test item, while a hand was shown that zooms in on historical people in a digital comic book. This is similar to item five, as the answer to the item was also given by an invisible commentator, who stated that readers are nowadays looking for “something other than violence” (Youtube 2012: 2:07 min.), while the video showed covers of comic books that depict violence. Further visual information was provided afterwards by an owner of a comic book store talked about previously popular violent comic books while pointing at the covers of such comic books and visually portraying violence by gesturing a punch.

This video can be said to offer an abundance of visual cues that could have assisted the students not only in answering the test items but in generally comprehending the spoken text. Especially the clearly visible lip movements of the speakers and the visible indication of a change in speakers might have helped the general comprehension of the TV report. However, the abundance of visual input, the quickly changing scenes and the constant shift between content and context visuals could have increased the cognitive load of the listening task and caused the test takers to feel distracted.

The second video used in the AV listening test was a video of a talk by Pete Alcorn (2009) in front of a live audience about his vision of the world in 2200, which was accompanied by PowerPoint slides. In his talk, Pete Alcorn started by deconstructing Robert Malthus’ pessimistic model of the world’s future, continued by analysing past historical events to predict future developments and ended with a positive outlook on the world’s future. Like the previous video, this video combined visual contextual and content information as it focused sometimes on the context of the speaking situation, i.e. the speaker and his physical surroundings, and sometimes on the PowerPoint slides, which showed images and graphs related to the content of the talk.

In this video, especially the images depicting the content of the spoken text might have provided valuable cues for answering the test questions. This might have been the case for answering test items eight, nine, ten, eleven and twelve. Concerning item eight, a picture of a house that is for sale with a sign saying “reduced price” in front of it was shown on the PowerPoint, followed by Pete Alcorn stating that “the price of the land drops” (Alcorn 2009: 1:25 min.), which provided the answer to the test question. The clearly visible lip movements of the speaker might have also aided speech perception. A further rather direct clue might have been provided by the PowerPoint for answering item twelve, as a slide showed a painting that depicted a violent scene probably of a revolution or a riot, while the speaker said that “transitions are dangerous times” (Alcorn 2009: 2:59 min.), which is the answer to the test item.

The link between the PowerPoint slides and the answers to items nine and ten were less direct but still potentially helpful. Regarding item nine, the PowerPoint showed an image of female workers demanding higher wages while the speaker explained that people will earn more money if the population declines. Concerning item ten, the PowerPoint showed a painting depicting middle class men while Alcorn explained that the decline in population following the great plague led to the “birth of the middle class” in Europe (Alcorn 2009: 2:01 min.), thereby indicating the answer to the test item.

Finally, the gestures of the speaker might have assisted the pupils in answering item eleven as the speaker explained that people believed that the past was always better than the future and that human history is seen as a down-hill slide (Alcorn 2009: 2:30 min.), which should indicate the test answer while he gestures a down-hill slide.

The third and last video used in the AV listening test showed a commencement speech by Steve Jobs (2011) at the University of Stanford in which he talked about his early life, the decision to drop out of college and stay on the university campus to take up courses that interested him and how this decision influenced his life. As this video focused predominantly on the speaker and sometimes showed the audience of the commencement speech, it can be assigned to the category of context visuals.

The video showed, for the most time, the upper body and the face of the speaker. The lip movements of the speaker were clearly visible in this video and might have supported the

test takers' speech perception. By observing the face of the speaker, the test takers could have derived further cues that might have assisted them in comprehending the spoken text and helped them in answering the test questions as the speaker stressed important content elements, which often corresponded to answers of the test items, by making and maintain eye contact with the audience. As Jobs did not use visual support for his speech, for example, in form of a PowerPoint, and did not gesture at all, only the cues derived from the lip movements and facial expressions of the speaker might have helped the students in answering the test questions.

#### ***6.5.4. Second set of student questionnaires***

After the AO and the AV listening test, the students filled in a further student questionnaire that matched either the audio-only or the audio-visual listening condition of the preceding test. The questionnaire for the student group taking the AO test<sup>20</sup> consisted of 14 items, while the questionnaire for the experimental groups taking the AV test<sup>21</sup> contained 20 items.

These questionnaires were intended to provide necessary information for the investigation of the test takers' interaction with the video material, their use of visual cues to aid listening comprehension and their opinion on the inclusion of visual materials in listening comprehension tests. The results of these questionnaires were crucial in answering the research questions and corroborating or disproving the hypotheses. Furthermore, the questionnaires established whether the students were already familiar with the listening comprehension tasks.

The first section of these questionnaires, which is called "Visual cues in everyday communication", was the same for the AV and the AO group. This part was intended to provide information about the attention the students pay to visual input in everyday communication and language learning. These questions were adapted from Sueyoshi and Hardison (2005: 697-698).

---

<sup>20</sup> This questionnaire is referred to as the "AO questionnaire" or "AO Qu." in Chapter 7.

<sup>21</sup> This questionnaire is referred to as the "AV questionnaire" or "AV Qu." in Chapter 7.

The next part of the questionnaire, called “Attention paid to visual cues in the video-listening test”, was only part of the questionnaire for the student groups taking that test. This section of the questionnaire served to establish how much attention the students paid to the videos in the testing situation and to the non-verbal cues provided by the videos, i.e. cues derived from facial expressions, lip movements, gestures and the context of the speaking situation, the latter including also information gained from content visuals, such as the PowerPoint slides in the second video. As this part of the questionnaire asked the students to indicate the attention they paid to these non-verbal cues separately for each of the three video listening texts, whereby visual cues derived from gestures were left out for the third video as the speaker did not use any gestures, it also served to establish whether there was a difference in the level of attention paid to the different videos.

The following section, which is named “Your opinion on the inclusion of visual input in a listening test”, was part of both questionnaires, but differed for the experimental groups and the control group. In this section, the test takers taking the AV test were asked about their opinion on the visuals included in the video listening test and indicated whether the visuals in general and which aspects of non-verbal communication in particular facilitated their comprehension or if they had a negative and distracting effect. Furthermore, they were asked about their preference for video-mediated or audio-only listening tests. Concerning the AO questionnaire, the students were asked whether they thought that the inclusion of visual input in the listening test would have helped their understanding of the spoken text and whether they would prefer video-listening tests over conventional audio-listening tests.

Finally, the last two items, which were almost identical in the AV and AO questionnaire, established whether the test takers were already familiar with some or all of the listening texts, which might compromise the test results, and invited the students to make additional comments concerning the tests. The students were invited, for example, to comment on the video quality and the inclusion of videos in listening tests, regarding the AV test, as well as the sound quality, the difficulty of the test and time to answer the test questions. Similarly to the baseline listening comprehension, only a small number of students were familiar with the selected listening comprehension tasks. One test taker was familiar with the first task and two test takers had taken the third listening comprehension “Connecting the Dots” before. However, as these test takers, who were not the ones familiar with the listening

tasks in the baseline listening test, did not perform better than the average test taker, which indicates that they remember the listening tasks at most vaguely, they were left in the calculations.

## 7. Results

In this chapter, the results of the conducted “Comparative study of video-mediated and audio-only listening tests in the Austrian EFL classroom” are presented. For the interpretation of the data, the statistical program IBM SPSS Statistics 21 (IBM Corporation 2012) was used. First, a test and item analysis is presented, before revealing the effects of visual stimuli on listening comprehension performance. Afterwards, the importance the study participants attach to visual cues in language learning and everyday communication is discussed before investigating the attention the test takers paid to the visual input in the testing situation. Finally, the test takers’ opinions on the inclusion of visual material in a listening test is presented and the results are related to the hypotheses and null-hypotheses.

### 7.1. Test and item analysis

In this section, a test analysis and an item analysis of the baseline listening test as well as the AO and AV listening tests is presented. The test analysis provides information about the students’ performance on the tests and the difficulty or ease of the different tests, while the item analysis establishes how well the individual test items performed in the tests. The item analysis further determines the reliability of the tests and reveals items which did not successfully discriminate between higher- and lower-level ability test takers. Based on the item analysis, several items were dropped in order to produce reliable test results, which were then used for further calculations intended to answer the research questions and corroborate or disprove the hypotheses, introduced in Sections 1.1 and 1.2.

#### 7.1.1. Test analysis: Baseline listening test

The analysis of the results of the baseline listening test, referred to as “Test A” in the statistics, revealed that the test was altogether moderately difficult as the mean of the test scores was 70.56%, shown in Table 2 below. The mean was only minimally influenced by outliers as the median and mode with 69.57% are very similar to the mean. The relatively large standard deviation of 13.22%, the range of 56.52%, the minimum of 43.48% and the maximum of 100% reveal that the 48 test takers varied considerably in their achievement on this test. This is further supported by the high variance (175.15%) of the test takers.

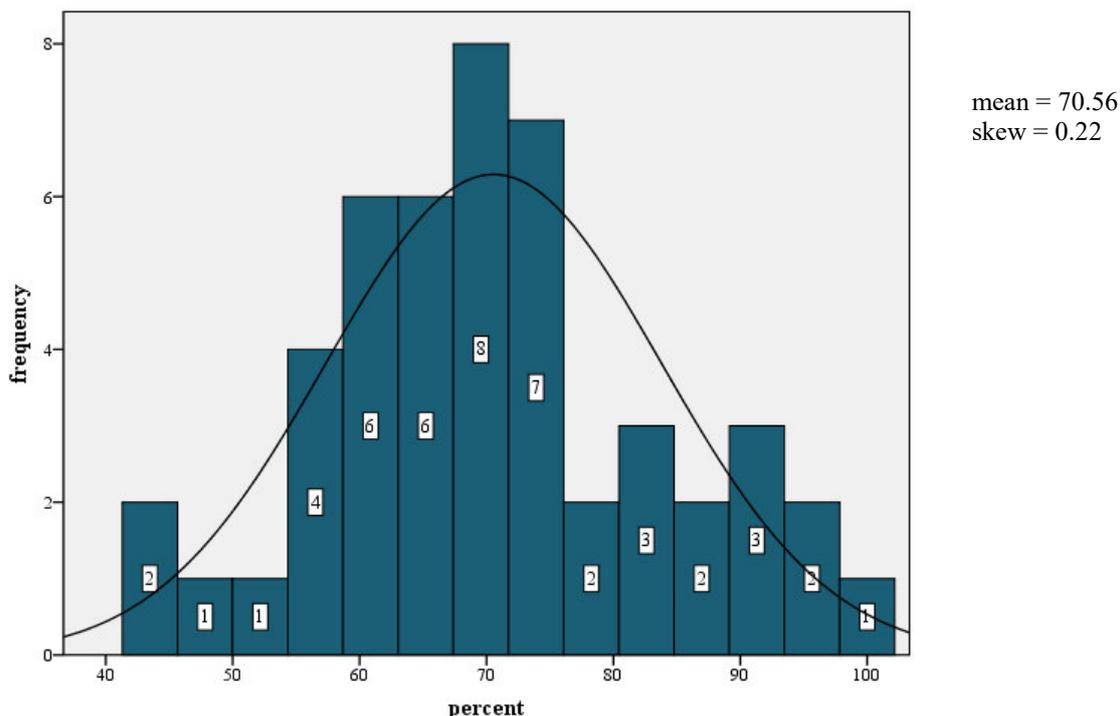
Generally, it can be concluded that the population of test takers in this study is relatively heterogeneous concerning their listening comprehension skills.

**Table 2 Test A: Descriptive statistics**

scores: Test A <sup>a</sup>	
no. of test takers	48
no. of test items	23
mean	70.56 <sup>a</sup>
median	69.57
mode	69.57
min.	43.48
max.	100.00
range	56.52
std. dev.	13.22
variance	175.15

<sup>a</sup> values below mean percentages

The following histogram, depicted in Figure 5, shows that most students achieved approximately 55% to 75% on the baseline listening test, while only four students achieved less than 55% and 13 of the 48 test takers reached more than 75%. This further confirms that the test was moderately difficult. The normal curve indicates that test results are relatively normally distributed. Calculating the skew of the results (0.22) shows that the distribution of the test's results is slightly positively skewed, indicating that there are more students at the lower end of the distribution (Green 2013: 44). The histogram further illustrates the previous suggestion that the test takers represent a relatively heterogeneous group as they are widely spread above and below the mean of 70.56%.



**Figure 5 Test results: Baseline listening test**

An analysis of the individual tasks, which is illustrated in Table 3 below, revealed that the test was moderately difficult because test task 2 was very easy, indicated by the high mean of 94.01% as well as the mode and median of 100%, whereas task 1 seemed to be very difficult. The difficulty of task 1 is indicated by a very low mean (46.88%), mode (37.50%) and median (43.75%). Task 3 was fairly easy, which can be seen by the high mode of 100% and median of 71.43%. The mean of this task is lower with 70.54%, resulting from the influence of low outliers. The range of 100% in task 1 and 78.57% in task 3 further indicated the heterogeneity of the test population. The small range of 37.50% in task 2 revealed that there was little variation in the test takers' achievement on this task. This result, in combination with the high mean, mode and median, suggests that task 2 might have been too easy for the test takers, which contrasts with the results of task 1, which might have been too difficult.

**Table 3 Test A: Descriptive statistics of test tasks**

test tasks	mean	mode	median	max.	min.	range
task 1	46.88 <sup>a</sup>	37.50	43.75	100	0	100
task 2	94.01	100	100	100	62.50	37.50
task 3	70.54	100	71.43	100	21.43	78.57

<sup>a</sup> all values mean percentages

### 7.1.2. *Item analysis: Baseline listening test*

The conclusions drawn in the test analysis were further examined by calculating the facility value, also named *p*-value, of the individual test items, which is illustrated in Table 4. To establish whether the items successfully discriminate between higher- and lower-level ability test takers, the corrected item-total correlation (CITC), or discrimination index, of the test items was calculated. Table 4 further shows Cronbach's Alpha if item deleted (CAID), which provides information about the individual test item's reliability (Green 2013: 38). The CAID needs to be seen in relation to the test's overall Cronbach Alpha value, shown in Table 5, which should be higher than 0.7 to constitute an acceptable level of test reliability (Pallant 2001: 85).

The *p*-value of the test items confirms that task 2, which consisted of items 9 to 16, was relatively easy for the test takers. Most of the test takers answered these test items correctly, which can be seen by the items' *p*-values, which range from 88% in item 14 to 100% in item 15 and are consequently far removed from the desired range of 33% to 66%, which offers the most valuable information about a test taker's proficiency (Alderson, Clapham & Wall 1995: 84)<sup>22</sup>. Moreover, also the facility value of items 18, 19, 22 and 23, which were part of task 3, indicates that these items, which were answer correctly by 81% to 98% of the test population, had a very low level of difficulty.

Furthermore, the item analysis of the baseline test revealed that several items did not successfully discriminate between stronger and weaker test takers. Following Henning (1987: 53), who argues that +0.25 constitutes an acceptable CITC threshold value<sup>23</sup>, items with CITC values of +0.25 and higher were seen as discriminating successfully between higher- and lower-level ability test takers in this study. The discrimination index of the test

---

<sup>22</sup> Alderson, Clapham and Wall (1995: 84) argue that items with *p*-values ranging from 33% to 66% offer the most valuable information about a test taker's level of proficiency as such items can have maximum discrimination indices of 1.00. Green (2013: 26) further suggests that *p*-values between 30% and 70% are the next best indications of a test taker's ability (Green 2013: 26). Facility values in-between 20% and 80% might still be indicative of the test taker's proficiency if they discriminate between higher- and lower-level ability test takers, which is indicated by the CITC value, and "contribute to the test's internal consistency" (Green 2013: 26), revealed by the CAID value. *P*-values higher than 80% or lower than 20% indicate that the majority of test takers answered the test item either correctly or incorrectly (Green 2013: 27). Such items are relatively unlikely to discriminate successfully between higher- and lower-level ability test takers (Green 2013: 27), which is, however, established by the CITC value. Alderson, Clapham and Wall (1995: 84) further add that especially items which are too easy are problematic because not all test takers who answered them correctly can be confined to the group of higher-level ability test takers. Consequently, items with a *p*-value of 80% and higher cannot have a CITC value above 0.6 (Alderson, Clapham & Wall 1995: 84).

<sup>23</sup> Henning (1987: 52-53) uses the term "point biserial correlation" instead of CITC.

items shows that item 1, items 5 to 11, item 13 to 15, items 18, 20, 22 and 23 did not successfully distinguish between weaker and stronger students.

As the CITC value of items 9 and 22 was negative and 0 for item 15, these items were removed from the calculations in order to increase the discrimination index of the remaining items. This decision was also influenced by the CAID index of the items, which was higher than the test's overall Cronbach Alpha value of 0.602, consequently indicating that a deletion of these items would increase the reliability of the test.

**Table 4 Test A: Item statistics and item-total statistics**

	<i>p</i> -value	corrected item-total correlation	Cronbach's Alpha if item deleted
Item1	0.38	0.240	0.585
Item2	0.35	0.255	0.582
Item3	0.65	0.348	0.568
Item4	0.25	0.268	0.581
Item5	0.54	0.213	0.589
Item6	0.63	0.169	0.596
Item7	0.71	0.204	0.590
Item8	0.25	0.162	0.596
Item9	0.98	-0.076	0.608
Item10	0.98	0.124	0.600
Item11	0.90	0.043	0.606
Item12	0.92	0.270	0.586
Item13	0.94	0.076	0.602
Item14	0.88	0.076	0.604
Item15	1.00	0.000	0.603
Item16	0.94	0.382	0.579
Item17	0.44	0.332	0.570
Item18	0.81	0.095	0.609
Item19	0.85	0.343	0.575
Item20	0.65	0.117	0.607
Item21	0.56	0.342	0.567
Item22	0.98	-0.026	0.606
Item23	0.90	0.226	0.587

**Table 5 Test A: Reliability statistics**

Cronbach's Alpha	no. of items
0.602	23

Afterwards, the procedure of removing items from the calculations was repeated until the remaining items, except for one item, successfully discriminated between test takers with different ability levels and the reliability of the test could not be increased any further. Altogether 13 items had to be removed from the test. Table 6 depicts the discrimination index of the remaining ten test items, which ranges from 0.233 in item 7, the removal of which would have negatively affected the test's reliability, to 0.412 in item 1, and the items' CAID values, which are all lower than the test's overall Cronbach's Alpha of 0.641, shown in Table 7. The small number of test items, which can decrease the test reliability, could explain why the test's Cronbach Alpha value is lower than Pallant's (2001: 85) desired minimum level of 0.7. This can be further explained by the fact that all of the test takers have roughly the same level of ability, which is B2 as defined in the CEFR (Council of Europe 2001: 27), and have been taught following the same curriculum, which might cause the variance to be lower than in more diverse test groups; hence the discrimination between the test takers might be weaker (Green 2013: 39). Generally, the analysis revealed that the test has a limited degree of reliability and comparisons of this test with the AV or AO test need to be treated with caution.

**Table 6 Test A: Revised item-total statistics**

	<b>corrected item – total correlation</b>	<b>Cronbach's Alpha if item deleted</b>
Item1	0.412	0.592
Item2	0.296	0.619
Item3	0.359	0.604
Item4	0.283	0.621
Item5	0.332	0.611
Item7	0.233	0.632
Item16	0.250	0.630
Item17	0.343	0.608
Item19	0.335	0.613
Item21	0.256	0.630

**Table 7 Test A: Revised reliability statistics**

<b>Cronbach's Alpha</b>	<b>no. of items</b>
0.641	10

As the baseline listening test was designed to establish whether there was a difference in the level of listening ability of the participating groups, the mean, mode and median of the

revised test with the remaining 10 items was calculated separately for group A, serving as the control group in the second testing session, and the two experimental groups, groups B and C. This is illustrated in Table 8. This table reveals that the experimental groups performed better than the control group on the baseline listening test, which can be seen by the mean, median, and mode, which are 60% for the experimental groups compared to 50.59, 50 and 40% for the control group. This might suggest that the experimental groups have a slightly higher level of listening ability than the control group.

**Table 8 Test A: Comparison of control and experimental test groups**

testing group	N	mean	mode	median
control group	17	50.59 <sup>a</sup>	50	40
experimental groups	31	60	60	60

<sup>a</sup> values mean percentages

An independent samples t-test, shown in Table 9, was run to investigate whether the difference in listening performance between the control and experimental groups was significant. The test showed that the two groups did not significantly differ in their performance as the t-test's two-tailed significance value is 0.163 and consequently above the threshold of 0.05. This similarity in listening ability between the groups can be seen as an ideal basis for later comparisons of the control and experimental groups' performances on the AO or AV listening test.

**Table 9 Test A: Control vs. experimental test groups**

Independent samples t-test	Levene's test for equality of variances		t-test for equality of means	
	F	sig.	t	sig. (2-tailed)
equal variances assumed	0.747	0.392	-1.147	0.163

### **7.1.3. Test analysis: Audio-only and audio-visual listening test**

Following the same procedure as in the analysis of the baseline listening test, the descriptive statistics of the AO and the AV tests and the individual test tasks were analysed. Although the AO and AV listening test consisted of the same texts and items, a separate analysis was conducted for the control group taking the AO test and the

experimental groups taking the AV test, which were calculated as one group, as the presence of the visuals could have affected the performance on the test.

Table 10 depicts the descriptive statistics for both tests and reveals that the mean (61.99%) and median (61.36%) of the test scores were higher for the AV test than for the AO test, which had a mean of 58.71% and a median of 56.82%. However, the mode of the AV test (59.09%) was lower than the mode of the AO listening test (81.82%). Generally, the experimental groups performed better on the test, which has to be seen, however, in relation to the results of the baseline listening test, which already established a certain, albeit not significant, difference in listening ability between the experimental and control group. The table generally reveals that the AO and AV listening test were more difficult than the baseline listening test, which had a mean of 70.56% and a mode and median of 69.57%. It further shows that the test takers taking the AO test varied considerably in their performance on the test, which can be seen by the high range of 81.82% and the extremely high variance of 440.11%, whereas the experimental test groups varied to a far lesser extent.

**Table 10 AO/AV test: Descriptive statistics**

scores <sup>a</sup>	AO test	AV test
no. of test takers	18	29
no. of test items	22	22
mean	58.71 <sup>a</sup>	61.99 <sup>a</sup>
median	56.82	61.36
mode	81.82	59.09
min.	13.64	27.27
max.	95.45	90.91
range	81.82	63.64
std. dev.	20.08	16.16
variance	440.11	261.18

<sup>a</sup> values below mean percentages

The histogram depicted in Figure 6 reveals that the majority of test takers achieved between approximately 42% and 88% on the AO and AV listening test, while only six test taker scored below 42% and two test takers achieved more than 88%. The histogram shows that the AO and AV listening were moderately difficult and relatively normally distributed, which is indicated by the normal curve. The skew of the results (-0.24) as well as the

normal curve reveal that the distribution of the results of the two tests is slightly negatively skewed, which indicates that more test takers are distributed above the combined mean of the two tests (60.74%).

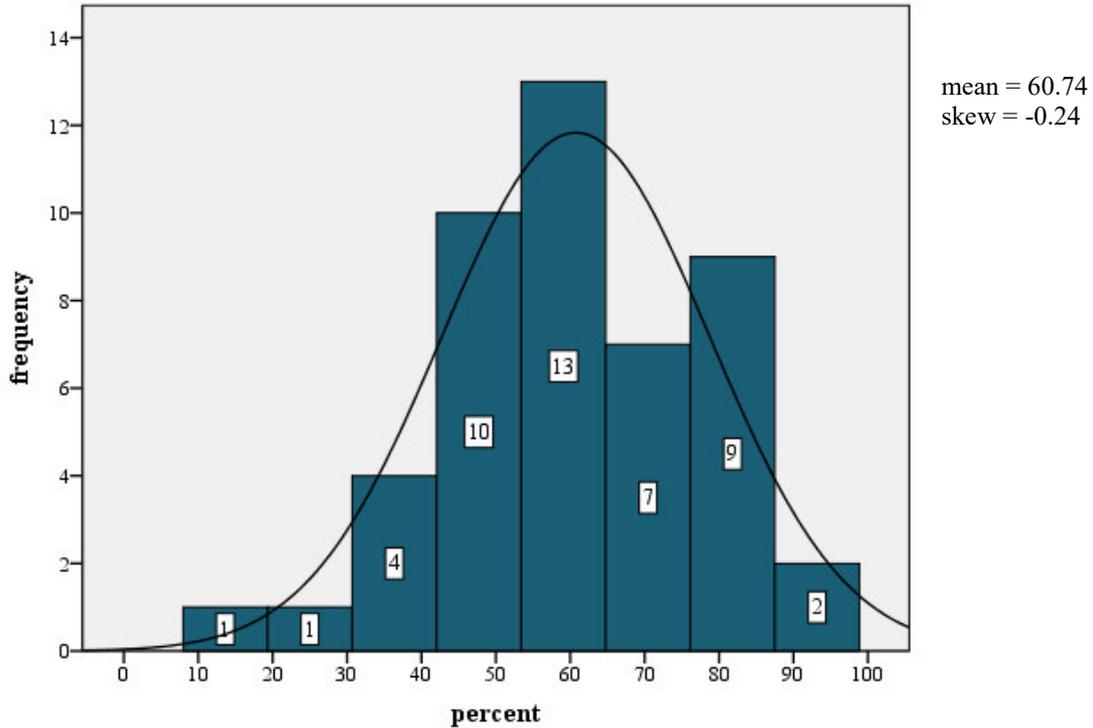


Figure 6 Test results: AO and AV listening test

Table 11 shows the test takers’ performance on the individual tasks for the AO as well as the AV test. It can be seen that the AV test group performed far better than the AO group on the first task, slightly better on the second task and minimally better on the third task, which can be seen by the higher mean, mode and median values of the AV group. Only the median for task three is equal for both test groups.

Table 11 AO/AV test: Descriptive statistics of test tasks

	task 1		task 2		task 3	
	AO	AV	AO	AV	AO	AV
mean	52.78 <sup>a</sup>	60.34	57.41	60.34	64.20	64.37
median	50.00	71.43	58.33	66.67	66.67	66.67
mode	42.86/ 71.43 <sup>b</sup>	71.43/ 85.71 <sup>b</sup>	33.33	66.67	44.44 <sup>24</sup>	55.56/ 77.78 <sup>b</sup>

<sup>a</sup> all values mean percentages 3, <sup>b</sup> two mode values available

<sup>24</sup> Four other mode values are available (55.56, 66.67, 77.78 and 100).

#### 7.1.4. Item analysis: Audio-only and audio-visual listening test

Afterwards, the  $p$ -values as well as the CITC and CAID values were calculated for the AO and AV listening tests, as depicted in Table 12 below. An analysis of the items' facility values confirmed the above-mentioned difference in performance of the two test groups. Several items (e.g. items 4, 5, 7, 11, 16 and 20) appeared to be easier for the AV group compared to the AO test group, recognisable by the higher  $p$ -values, while other items were more difficult for the experimental test groups (e.g. items 3, 6, 12, 17, 19 and 22).

**Table 12 AO/AV test: Item statistics and item-total statistics**

	$p$ -value		corrected item – total correlation		Cronbach's Alpha if item deleted	
	AO	AV	AO	AV	AO	AV
Item1	0.61	0.62	0.630	0.301	0.785	0.657
Item2	0.61	0.66	0.483	0.484	0.793	0.637
Item3	0.67	0.59	0.191	0.093	0.809	0.680
Item4	0.78	0.93	0.266	0.437	0.806	0.655
Item5	0.44	0.86	0.386	0.220	0.799	0.666
Item6	0.11	0.03	0.253	0.068	0.805	0.675
Item7	0.56	0.69	0.223	0.018	0.808	0.690
Item8	0.56	0.52	0.413	0.248	0.797	0.663
Item9	0.11	0.28	0.379	0.158	0.800	0.672
Item10	0.72	0.72	0.271	0.104	0.805	0.677
Item11	0.78	0.90	0.213	0.293	0.807	0.661
Item12	0.56	0.45	0.276	-0.010	0.805	0.691
Item13	0.72	0.76	0.242	0.453	0.806	0.643
Item14	0.28	0.34	0.546	0.097	0.791	0.679
Item15	0.67	0.72	0.219	0.392	0.808	0.648
Item16	0.61	0.76	0.660	0.273	0.783	0.661
Item17	0.44	0.28	0.642	0.425	0.784	0.645
Item18	0.78	0.83	0.119	0.471	0.811	0.644
Item19	0.78	0.62	0.057	0.278	0.814	0.660
Item20	0.61	0.76	0.342	0.273	0.801	0.661
Item21	0.67	0.69	0.691	0.200	0.782	0.668
Item22	0.94	0.79	0.509	0.255	0.799	0.663

**Table 13 AO/AV test: Reliability statistics**

Cronbach's Alpha		no. of items
AO	AV	
0.808	0.674	22

The analysis of the test items revealed, furthermore, that the CITC value of the individual items varied to a great extent between the control and the experimental test groups. The first calculation of the items' CITC values showed that several items (e.g. items 5, 6, 9, 10, 14, 21) discriminated successfully between weaker and stronger students in the control group, while they were unable to achieve this in the AV test group. The opposite was true for a number of other items (e.g. items 11, 13, 15, 18, 19), which had a CITC value higher than +0.25 for the AV group and a lower value for the AO test group. Item 12 had a negative CITC value of -0.018 for the AV and a rather low, however, still acceptable value for the AO group. This item was removed from the calculations, together with items 3 and 7, which had CITC values below the threshold level for both test groups. The CAID index of items 7 and 3 also revealed that a removal of these items would lead to an increased or at least equal test reliability level for both tests, which can be seen by comparing the items' CAID values with the tests' overall Cronbach's Alpha values shown in Table 13, which are 0.808 regarding the AO and 0.674 regarding the AV test. These Cronbach's Alpha values generally suggest that the AO test can be seen as a very reliable test, while the AV test did not reach the desired reliability level of 0.7.

The removal of items 3, 7 and 12 changed the CITC values of the remaining test items, shown in Table 14, and led to an increase in the reliability of the AV listening test, which can be seen by the test's overall Cronbach Alpha of 0.724, depicted in Table 15, which surpasses the test reliability threshold level. Although the CITC index of several items (i.e. items 10, 11, 13, 15, 18, 19 in the AO test and items 1, 5, 6, 9, 14 and 21 in the AV test) is below the threshold of +0.25, they were left in the calculation as their CITC value was only weak in one test group and strong in the other test group. The items' CAID index shows that a removal of these items would have led to an increased or equal CAID value for one test, however, a decreased CAID value for the other test and consequently a reduced test reliability.

**Table 14 AO/AV test: Revised item-total statistics**

	corrected item – total correlation		Cronbach’s Alpha if item deleted	
	AO	AV	AO	AV
Item1	0.594	0.244	0.784	0.718
Item2	0.496	0.399	0.790	0.702
Item4	0.323	0.348	0.802	0.712
Item5	0.346	0.188	0.800	0.724
Item6	0.261	0.112	0.804	0.724
Item8	0.394	0.255	0.797	0.717
Item9	0.354	0.183	0.800	0.723
Item10	0.240	0.286	0.806	0.714
Item11	0.198	0.299	0.808	0.714
Item13	0.240	0.499	0.806	0.694
Item14	0.539	0.084	0.788	0.733
Item15	0.192	0.414	0.809	0.701
Item16	0.661	0.309	0.779	0.711
Item17	0.632	0.438	0.781	0.699
Item18	0.128	0.472	0.811	0.699
Item19	0.026	0.338	0.816	0.709
Item20	0.399	0.283	0.797	0.714
Item21	0.756	0.219	0.774	0.720
Item22	0.517	0.340	0.797	0.709

**Table 15 AO/AV test: Revised reliability statistics**

Cronbach’s Alpha		no. of items
AO	AV	19
0.806	0.724	

## **7.2. Effects of visual input on listening comprehension performance**

In this section the effects of the visual materials used in the AV listening test on the test takers’ listening comprehension performance are investigated. To establish if there is a difference in the test takers’ performance in the presence or absence of visual cues, the results of the AO test are compared with the results of the AV test. This chapter further intends to provide an answer to the first research question, which reads as follows: “*Does the use of video material in a listening test affect the test takers’ listening comprehension performance?*” and “*If so, does the inclusion of visual stimuli increase or decrease test taker performance?*”

First the means achieved in the AO or AV listening test were compared. Table 16, which is based on the 19 items of the revised AO and AV listening test, reveals that the experimental groups scored higher on the test than the control group with an average of 63.52 % compared to 59.06 %. However, this has to be seen in relation to the results of the revised baseline listening test, which already established a slightly, though not significantly, higher listening ability in the experimental test groups, as mentioned in Section 7.1.2.

**Table 16 AV/AO test: Comparison of control and experimental test groups**

test group	mean in %	no. of test takers
control group	59.06	18
experimental groups	63.52	29

To ascertain whether the difference in listening performance on the AO and AV listening test is significant, an independent samples t-test, shown in Table 17, was conducted. This test revealed that the difference between the control group and the experimental test groups was insignificant (sig. = 0.448). Consequently, the first research question can be answered by stating that the presence of videos in a listening test does not affect the test takers' listening comprehension ability to a significant degree.

**Table 17 AO/AV test: Control vs. experimental groups**

Independent samples t-test	Levene's test for equality of variances		t-test for equality of means	
	F	sig.	t	sig. (2-tailed)
equal variances assumed	0.468	0.497	-0.766	0.448

Further comparisons of the test takers' performance on the individual tasks of the AO or AV test did not reveal any statistically significant difference between the control group and the experimental groups, which indicated that none of the three videos used as input in the AV test significantly increased nor decreased the test takers listening comprehension performance. These calculations can be found in Table 30, included in Appendix 7 page 162.

To establish whether the presence of visual cues might have facilitated answering individual test items, which was suggested in Section 6.5.3 in relation to several test items<sup>25</sup>, independent samples t-tests were conducted to reveal any significant difference in the performance of the control group and the experimental groups on these particular items. These t-tests, which are summarised in Table 31 in Appendix 7 page 163, revealed that the performance of the test takers taking the AV listening test on these test items did not significantly differ from the performance of the control group. The only exception was test item 5, which was significantly easier for the students taking the AV test, indicated by a *p*-value of 0.86 compared to 0.44 for the control group and a significance value of 0.010. In this case the video in the AV test, which showed covers of violent comic book that used to be popular and a person visually portraying violence while a commentator states that readers are nowadays looking for “something other than violence” (Youtube 2012: 2:07 min.), could have indeed helped the students in answering the test item; however, also unknown factors might have influenced this result.

### ***7.2.1. Effects of visual input on listening comprehension performance of test takers with different listening ability levels***

To establish whether the presence of visual input had an effect on test takers with different listening ability levels the mean test scores of these groups on the revised baseline listening test (10 items) and on the revised AO or AV test (19 items) were compared and paired-samples t-tests were conducted<sup>26</sup>. Based on the test scores achieved in the baseline listening test, the test takers were first split into three ability levels: lower ability level, average ability level and higher ability level<sup>27</sup>.

The comparison of the percentual mean test scores achieved by these ability level groups on the baseline and the AO or AV listening tests is depicted in Table 18 below, which sorts the students with different ability levels according to their testing group (control vs.

---

<sup>25</sup> These test items are items 1, 3, 4, 5, 8, 9, 10, 11 and 12, of which item 3 and 12 were deleted as a result of the item analysis, as well as items 14 to 22, which were part of the third listening task. Items 14 to 22 were included as the clearly visible lip movements of the speaker in the third video might have assisted the speech perception of the test takers taking the AV listening test.

<sup>26</sup> These paired-samples t-test were intended to reveal any significant differences between the performances of the ability level groups on the different tests.

<sup>27</sup> The lower ability level group consisted of 17 students achieving 1 to 4 test points on the baseline listening test, while the average ability level group comprised 15 students who achieved 5 to 6 points and the higher ability level group 16 students achieving 7 to 10 points.

experimental). The table reveals that the performance of the students classified as higher level ability listeners slightly decreased from 85% (baseline test) to 81.58% (AO test) concerning the control group and 80.83% to 75.12% concerning the experimental groups. While the performance of the students with an average listening ability taking the AV test also decreased, the average-level ability test takers in the control group performed slightly better on the AO test compared to the baseline listening test. The test takers in the experimental group who achieved the least percentage points in the baseline listening test drastically improved their performance in the AV listening test and achieved an average of 61.40 percent, thus surpassing the students who were classified as having an average listening ability level, who obtained 51.46 percent. However, also the lower-level ability test takers in the control group improved considerably and achieved 46.71% in the AO test compared to 30% in the baseline listening test.

**Table 18 Test A and AO/AV test: Comparison of means of test groups split by listening ability level**

test group	listening ability level	mean in %		no. of test takers	
		test A	AO/AV test	test A	AO/AV test
control group (AO test)	lower ability level	30.00	46.71	8	8
	average ability level	56.00	60.53	5	6
	higher ability level	85.00	81.58	4	4
experimental groups (AV test)	lower ability level	35.56	61.40	9	9
	average ability level	57.00	51.46	10	9
	higher ability level	80.83	75.12	12	11

Furthermore, paired samples t-tests comparing the performances of test takers with different ability levels on the baseline and the AO or AV test were conducted. These calculations, which are shown in Table 19, revealed that the performance of the test takers with an average or higher ability level did not differ significantly from one to the next testing session. However, they showed that there was a significant difference in the performance of the lower-level ability test takers on the baseline listening test and the AO and AV test. This is indicated by the significance value of 0.039 for the control group and 0.011 for the experimental group. The mean values further illustrate the improvement in performance of the lower-level ability test takers. These values show that the lower ability students in the control group improved by 16.71%, while the lower ability students in the experimental group increased their average performance by 25.85%.

**Table 19 Paired samples t-test: Comparison of performance of test groups split by listening ability level on Test A vs. AO/AV test**

testing group	listening ability		mean <sup>a</sup>	t	sig. (2-tailed)
control group	lower ability level	pair: results test A in % -	-16.71	-2.536	0.039
	average ability level		-1.89	-0.168	0.875
	higher ability level		3.42	1.025	0.381
experimental groups	lower ability level	results AO/AV test in %	-25.85	-3.271	0.011
	average ability level		7.57	1.684	0.136
	higher ability level		6.70	1.492	0.167

<sup>a</sup> The mean value is calculated by subtracting the percentual mean of the AO/AV test from the percentual mean of the baseline listening test (Green 2013: 99)

To reveal whether the performance of the lower-level ability test takers in the experimental group, who showed the greatest improvement, was significantly different from the performance of the students with a lower ability level in the control group, which might have been an indication that they profited from the visual cues available in the AV test, an independent samples t-test was run. This test, shown in Table 20, revealed that the performances of these groups of test takers did not significantly differ from each other, indicated by the significance value of 0.095.

**Table 20 AO/AV test: Comparison of performance of lower-level ability students on AO vs. AV test**

Independent samples t-test	Levene's test for equality of variances		t-test for equality of means	
	F	sig.	t	sig. (2-tailed)
equal variances not assumed	6.703	0.021	-1.811	0.095

### ***7.2.2. Effects of visual input on listening comprehension performance of test takers with different learning styles***

This section investigates whether the presence of visual input affected the performance of test takers with different learning styles in different ways, causing, for example, test takers who are predominantly visual learners or bimodal auditory and visual learners to perform better in the AV listening test than in the baseline listening test. To achieve this, the average performances on the baseline test of the test takers in the experimental group, grouped according to their learning styles, were compared with their average performances

on the AV listening test. Prior to this, the test takers were classified according to their learning style into different learning style categories<sup>28</sup>.

Table 21 below shows that the predominantly tactile learners achieved best on the baseline listening test, which can be seen by the high mean of 83.33 percent, followed by the bimodal auditory and tactile learners with a percentual mean of 80. These test takers achieved similarly well on the AV listening test. The test takers with a predominantly visual learning style as well as the bimodal visual and auditory learners, who might be expected to perform better in the presence of visual input due to their learning style, performed actually worse in the AV listening test. Comparing the means of these two groups of test takers with the total means of the tests, it can be seen that these students had an average or even above average performance in test A, while their performance was lower than the average of 63.16 percent in the AV listening test.

**Table 21 Test A and AV test: Comparison of performance of experimental group split by learning style on test A vs. AV listening test**

Learning style	mean in %		N	
	test A	AV test	test A	AV test
mixed learning style	55.56	65.79	18	16
bimodal visual/auditive	60.00	51.32	4	4
bimodal auditory/tactile	80.00	89.47	1	1
predominantly visual	70.00	56.14	3	3
predominantly auditory	40.00	47.37	2	2
predominantly tactile	83.33	78.95	3	2
total	60.00	63.16	31	28

An ANOVA analysis, illustrated in Tables 36 and 37 in Appendix 7 page 165, revealed that the differences found in the performance of the test takers with different learning styles within the baseline (sig. = 0.128) and AV listening test (sig. = 0.165) were statistically insignificant. Furthermore, paired samples t-tests were conducted to compare

---

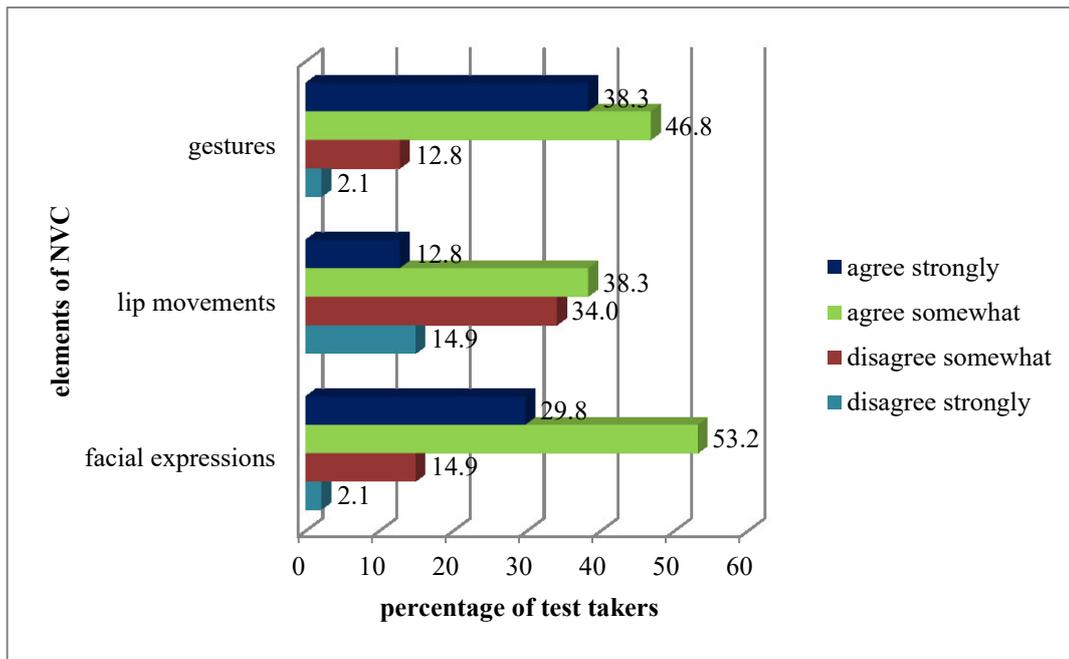
<sup>28</sup> These learning style categories are: mixed learning style, bimodal visual/auditive, bimodal auditory/tactile, predominantly visual, predominantly auditory and predominantly tactile. The test takers were assigned to the learning style categories according to the percentage points they achieved for the individual learning styles (i.e. visual, auditory and kinaesthetic) in the learning styles questionnaire (see Appendix 4). Test takers who obtained 25 to 40 percent for every learning type were classified as having a mixed learning style. If the test takers achieved 35 percent and more for two learning styles and less than 20 percent for the remaining learning style, they were classified as bimodal learners. Finally, test takers who had more than 50 percent for one learning style were assigned to the categories predominantly visual, predominantly auditory or predominantly tactile.

the performance of the different learning style groups on the baseline listening test with their performance on the AV listening test. These tests are summarised in Table 32, which is presented in Appendix 7 page 164, and revealed that there was also no significant difference between the performances of the test takers grouped by learning style on the different listening tests.

### **7.3. Importance of visual input in everyday communication and language learning**

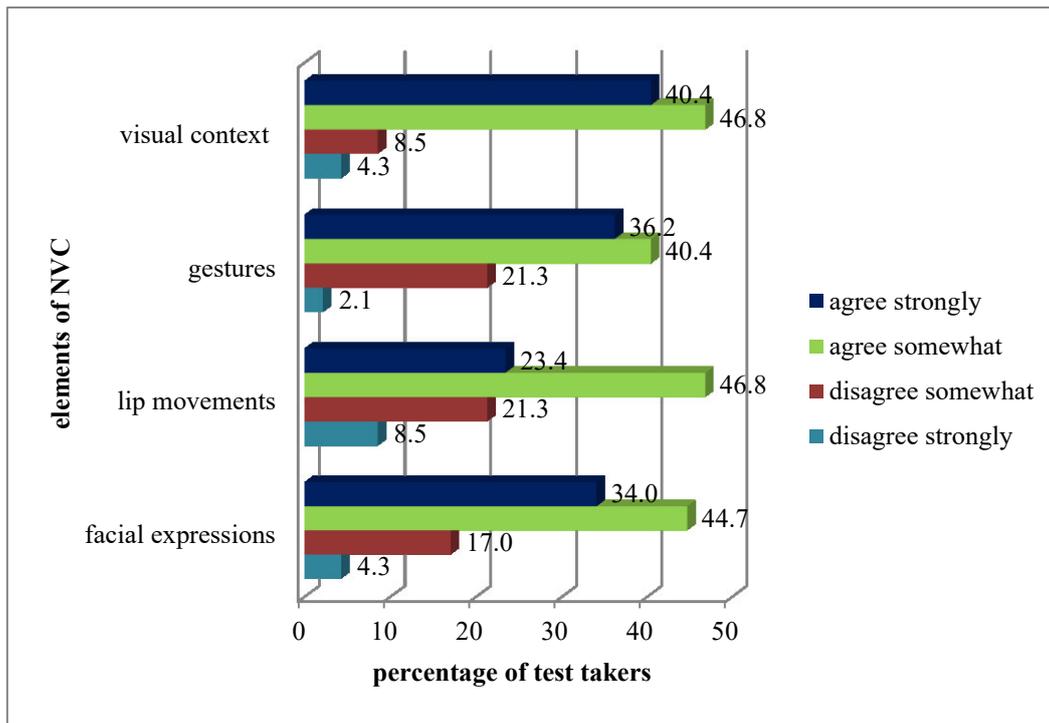
This section explores the importance the test takers attach to visual cues in language learning and communication, thereby providing an answer to the second research question, which asks “*How important are visual stimuli in everyday communication and language learning for the test takers?*”.

To answer this research question, the test takers’ answers to the questionnaire section “Visual cues in everyday communication” included in both the AV and AO questionnaire were analysed. Figure 7 illustrates the attention the students in the experimental as well as the control group report to pay to non-verbal elements of communication, in this case gestures, lip movements and facial expressions, in face-to-face conversations. The diagram shows that gestures and facial expressions play a more important role in face-to-face conversations than lip movements. This can be seen by the large percentage of test takers who either agree strongly to paying attention to gestures (38.3%) and, to a lesser degree, to facial expressions (29.8%) or agree to certain extent to focusing on gestures (46.8%) or facial expressions (53.2%). In comparison, only 12.8% of the test takers are convinced that they pay attention to the lip movements of the interlocutor in face-to-face conversations, while 14.9% do not attach any importance to lip movements and further 34% indicate that they pay only little attention to lip movements.



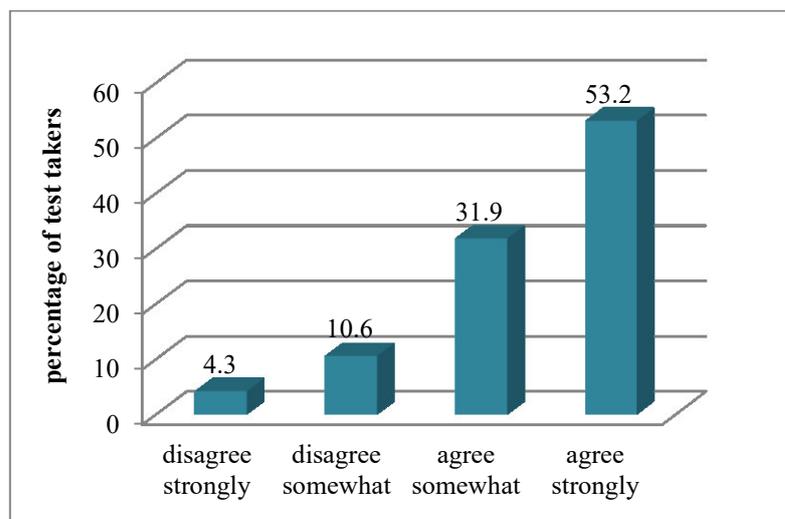
**Figure 7 AV/AO Qu. items 6-8: “In face-to-face conversations, I observe my conversation partner’s gestures / lip movements / facial expressions.”**

The vital role gestures and facial expressions play in everyday communication and language learning is further illustrated in Figure 8 where 36.2% of the students agreed strongly that it is easier to understand English if they can see the gestures of a speaker and 34% agreed strongly that cues derived from the facial expressions of a speaker helps them in comprehending English. Moreover, a large number of students agreed to a certain extent that gestures (40.4%) and facial expression (44.7%) support listening comprehension. While the information derived from the lip movements of speaker is less supportive in comprehending English, seeing the visual context of the speaking situation appears to be even more important than gestures and facial expressions as 40.4% agreed strongly and 46.8% to some extent that receiving visual contextual information aids comprehension.



**Figure 8 AV/AO Qu. items 1-4: “I think that it is easier to comprehend spoken English if I am able to see the context of the speaking situation / the speaker’s gestures / lip movements / facial expressions.”**

The importance of visual input in EFL listening comprehension is further supported by the fact that 53.2% of the test takers reported that it is easier to understand English on TV, in the presence of visual cues, rather than on the radio, which requires the test takers to rely only on hearing. This is illustrated in Figure 9 below. Moreover, 31.9% of the test takers agree somewhat with the statement that English on TV is easier to understand than on the radio, while only 10.6% disagreed somewhat and 4.3% disagreed strongly.



**Figure 9 AV/AO Qu. item 5: “I think that it is easier to understand English on TV rather than English on the radio.”**

Finally, Table 22 reveals that the test takers prefer activities in which visual cues are present to improve their English listening skills over language learning activities in which information comes only from the auditory channel<sup>29</sup>. While listening to the radio and to music is not perceived as particularly helpful in improving EFL listening skills, all activities in which the students receive information from the auditory as well as the visual channel are perceived as rather useful language learning activities. In this respect, especially having conversations in English and watching videos on the internet supports the students in advancing their listening comprehension skills.

**Table 22 AV/AO Qu. item 9: “The following activities are very helpful in improving my English listening skills”**

<b>Language learning activities</b>	<b>mean<sup>a</sup></b>
watching TV	3.26
watching videos on the internet	3.38
listening to the radio	2.55
listening to music	2.66
having conversations in English	3.47
attending to your English lessons	3.00

<sup>a</sup> 1 = disagree strongly, 2 = disagree somewhat, 3 = agree somewhat, 4 = agree strongly

Generally, the analysis conducted in this chapter revealed that visual cues play a crucial role for EFL learners in everyday communication and in learning English, as the students perceive cues derived from elements of non-verbal communication as facilitative in listening comprehension and engage mostly in activities which involve auditory as well as visual information to improve their listening skills. Consequently, it can be argued that AV listening tests, which provide visual alongside auditory input, closely simulate listening situations in which the students naturally engage in and thus constitute a more authentic form of listening test.

<sup>29</sup> A table showing not only the average opinion of the students but the exact test taker responses can be found in Appendix 7 on page 162 (see Table 29).

#### 7.4. Attention paid to the visual input in the audio-visual listening test

This section investigates the answer to the third research question, “*To what extent do the test takers attend to the visual input in a video-mediated listening test?*”, and examines whether there are differences in the attention paid to the visual input between test takers of different ability levels and test takers with different learning styles.

Figure 10 shows that a small majority of test takers, 58.6% altogether, paid close attention to the visual input, either watching the videos often (48.3%) or very often (10.3%), while a smaller proportion of test takers reported that they watched the videos hardly ever (24.1%) or never (17.2%).

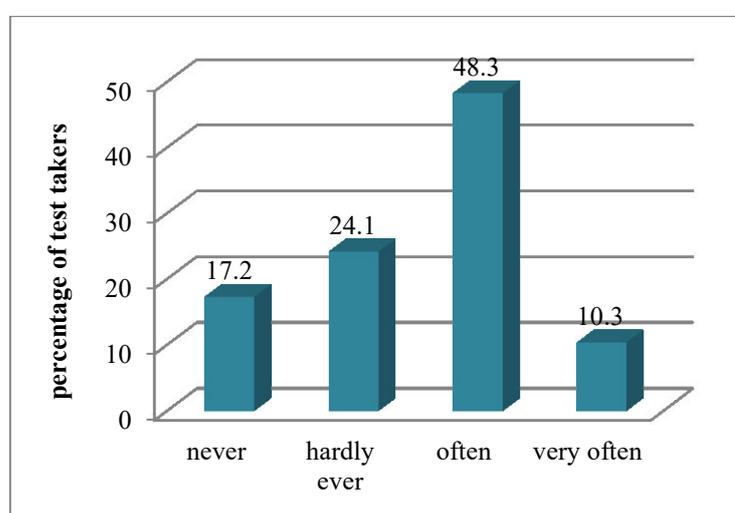


Figure 10 AV Qu. item 10: “How often did you watch the video in the test?”

To determine whether there was a difference in the attention paid to the different videos shown in the AV test and the individual aspects of non-verbal communication to which the videos provided access, the answers to AV questionnaire items 11 to 13 were analysed. A graphic depiction of this analysis can be seen in Figure 11. It was revealed that the test takers generally paid little attention to the third video, indicated by the low means of 2.55 concerning facial expressions, 2.34 concerning lip movements and 2.79 concerning the visual context of the speaking situation. This might suggest that the test takers could not gain much from the visual information in this video, which showed mostly the face and upper body of speaker who did not use any gestures or a visual support for his speech, to aid comprehension. By contrast, the students paid far more attention to the first and second video. As established in the preceding Section 7.3, the lip movements of the speakers were

of little importance to the test takers in all of the three test videos. Contrary to the high importance of facial expressions in everyday communication, the test takers reported that they paid only little attention to the facial expressions of the speakers in the test videos. The test takers paid close attention to the gestures of the speaker in the second video, and the visual context of the speaking situation in the first video, which was only moderately important in the other video-listening comprehension tasks.

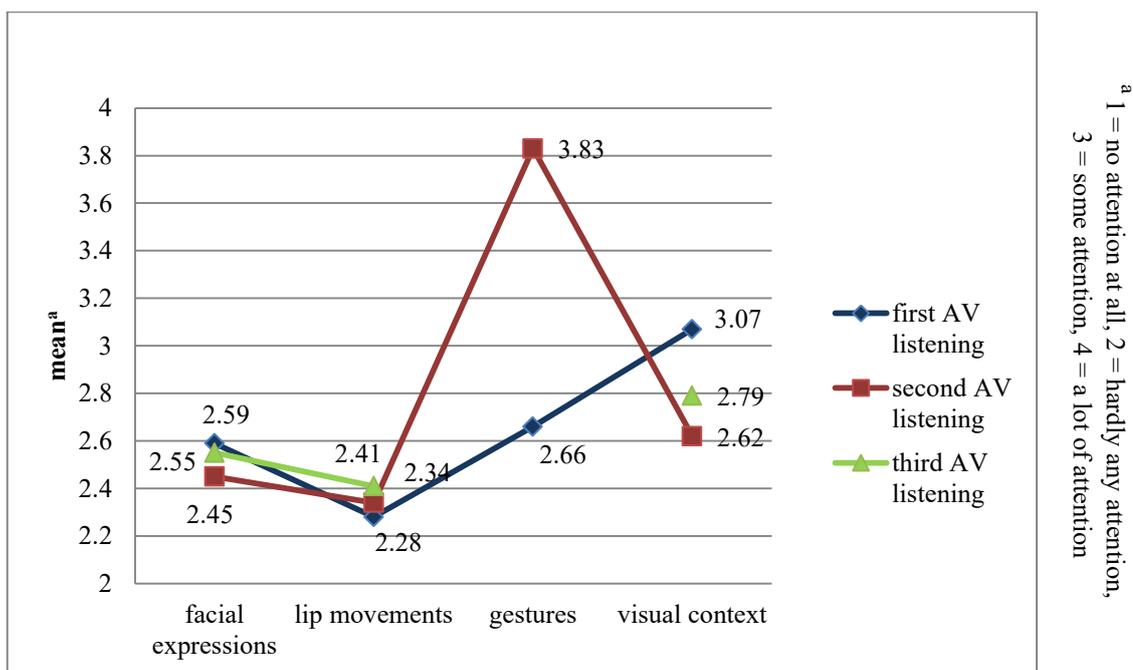


Figure 11 AV Qu. items 11-13: Attention paid to facial expressions, lip movements, gestures and visual context in the AV listening comprehension test

#### 7.4.1. Differences in the attention paid to the visual input in the audio-visual listening test between test takers with different listening ability levels

In order to detect differences in the attention paid to the videos between test takers with different listening ability levels and thus answer the first sub-question of the third research question, which asks “*Is there a difference in the attention paid to the visuals between higher- and lower-level ability test takers?*”, the average level of attention paid to the videos was calculated for every ability level group. Table 23 shows that the test takers with a lower level of listening ability attended more closely to the videos than students with average- or higher-level listening skills, who were very similar in their level of attention paid to the test videos.

**Table 23 AV Qu. item 10: “How often did you watch the video in the test?” (arranged according to ability level)**

<b>ability level</b>	<b>mean<sup>a</sup></b>	<b>N</b>
lower ability level	2.67	9
average ability level	2.44	9
higher ability level	2.45	11

<sup>a</sup> 1 = never, 2 = hardly ever, 3 = often, 4 = very often

In order to see if the difference between the test takers with a lower ability level and the other ability level groups is significant, a Spearman correlation test was conducted<sup>30</sup>. This test, which can be found in Table 44 in Appendix 7 on page 167, revealed that there is no significant correlation between the test takers’ listening ability level and the attention they paid to the visual input in the AV listening test (sig. = 0.589).

**7.4.2. Differences in the attention paid to the visual input in the audio-visual listening test between test takers with different learning styles**

This section investigates the second sub-question of the third research question, which reads as follows: “*Is there a difference in the attention paid to the visuals between test takers with different learning styles?*”. To answer this question, the average attention the test takers paid to the visual input was calculated, which is depicted in Table 24 below, which arranges the test takers according to their learning style. This table shows that the test takers who have a bimodal visual and auditive learning style watched the video often, while the students with other learning styles watched the videos on average hardly ever or in-between hardly ever and often.

---

<sup>30</sup> The Spearman procedure was used because the data investigated were ordinal data which did not allow the use of Pearson correlation test (Green 2013: 82).

**Table 24 AV Qu. item 10: “How often did you watch the video?” (arranged according to learning style)**

Learning style	mean <sup>a</sup>	N
mixed	2.50	16
bimodal - visual/auditive	3.00	4
bimodal - auditive/tactile	2.00	1
predominantly visual	2.33	3
predominantly auditive	2.00	2
predominantly tactile	2.50	2

<sup>a</sup> 1 = never, 2 = hardly ever, 3 = often, 4 = very often

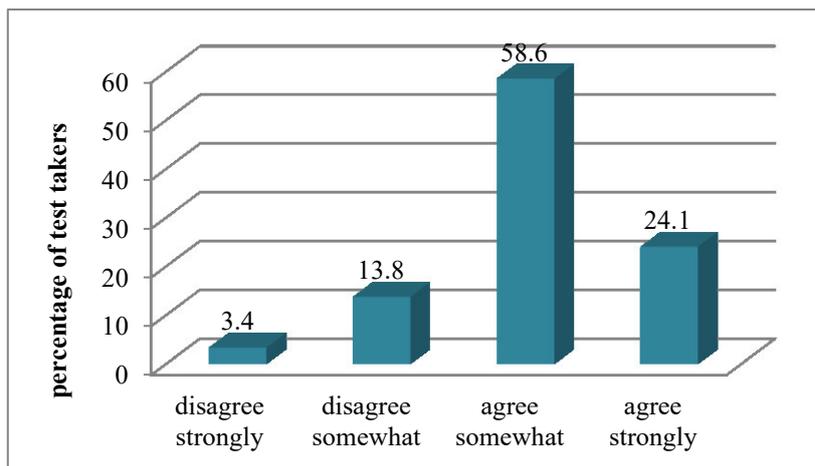
Following this comparisons of means, a Spearman correlation test was run. This test, which is depicted in Table 45 on page 167 in Appendix 7, revealed that there was no significant correlation between the learning style of the test takers and the attention they paid to the videos in the AV listening test. This can be seen by the significance value of 0.691.

## **7.5. Test taker opinions on the inclusion of visual input in a listening test**

This section investigates the test takers’ opinions on the use of videos in listening tests, focusing on their perception of the visual input as helpful or distracting and their preference for video-listening tests or traditional AO listening tests, and establishes whether test takers with different ability levels and learning styles differ in their opinion. Thus the section provides an answer to the third research question, “*To what degree do the test takers perceive the visual stimuli as an aid to comprehension or a distracting factor?*”. This is achieved by analysing the answers to the AV as well as the AO questionnaire. The results of this analysis is further correlated with the test takers’ performances.

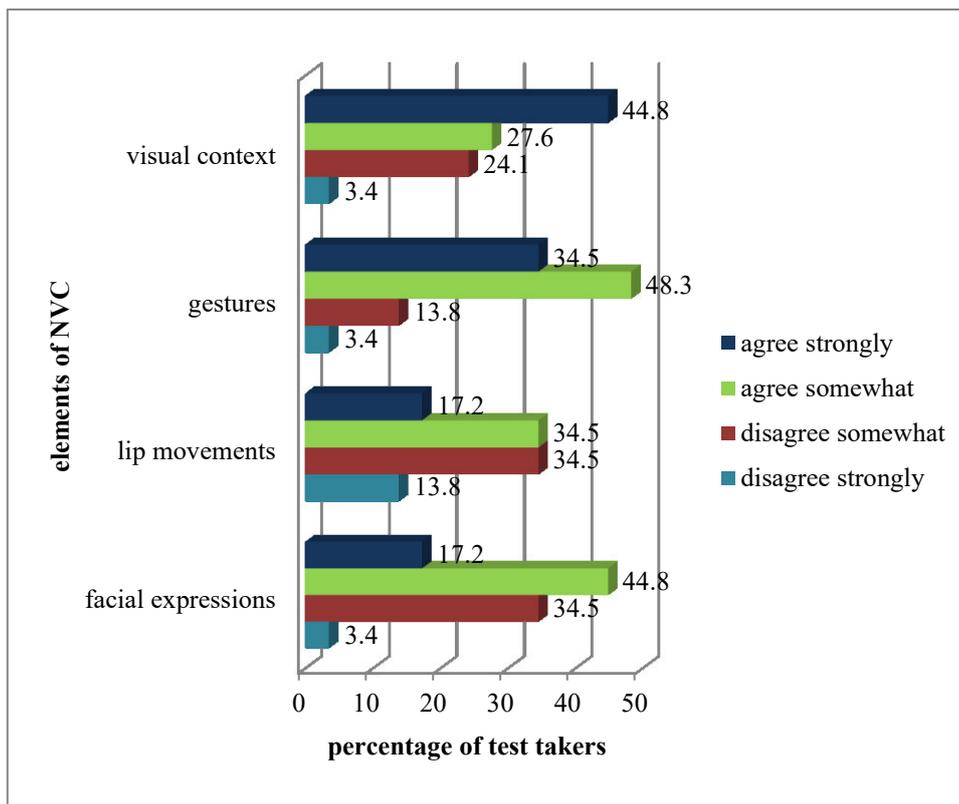
### **7.5.1. Perception of visual input as an aid to comprehension or a distracting factor**

Figure 12 illustrates that 58.6% of the 29 test takers who took the AV listening test felt that the videos supported them somewhat in comprehending the spoken text and helped them answer the test questions. Furthermore, 24.1% of the test takers strongly agreed with the statement that the visual input facilitated listening comprehension, while only 13.8 and 3.4% disagreed to a certain extent or disagreed strongly.



**Figure 12 AV Qu. 14: “In general, the visuals in this video-listening test supported my comprehension of the spoken text and helped me answer the test questions.”**

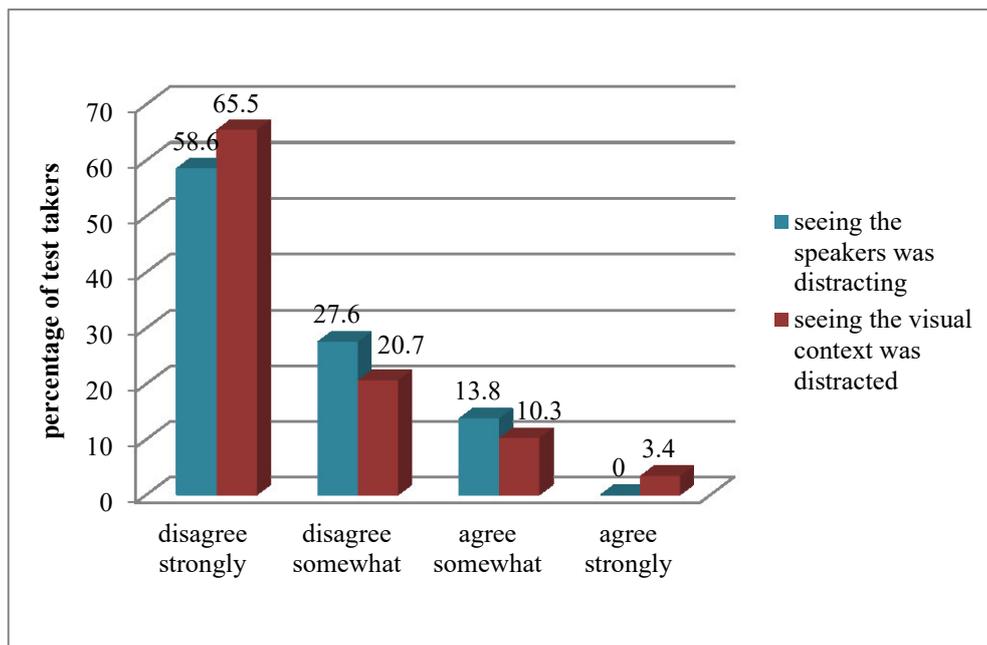
Figure 13 further reveals which aspects of the visual input (i.e. visual context of the speaking situation or gestures, lip movements or facial expressions of the speakers) supported listening comprehension and were perceived as helpful when answering the test questions. The figure shows that seeing the visual context of the speaking situation and the gestures of the speakers supported listening comprehension to a greater extent than seeing the speakers’ facial expressions and lip movements. The majority of test takers agreed strongly (44.8%) or somewhat (27.6%) with the statement that visual contextual information supported listening comprehension and helped them answer the test items. This is similar to the perception of the speakers’ gestures, which were perceived as very helpful (34.5%) or somewhat helpful (48.3%). However, only 17.2% of the test takers felt strongly that seeing lip movements and facial expressions supported comprehension, while 34.5% disagreed to some extent with the usefulness of these forms of NVC in supporting listening comprehension, 13.8% disagreed strongly concerning lip movements and 3.4% concerning facial expressions.



**Figure 13 AV Qu. item 15 (abbr.): These elements of NVC supported listening comprehension and helped answer test questions.<sup>31</sup>**

The AV questionnaire further asked the test takers whether seeing the videos, in particular the speakers and the visual context of the speaking situations, distracted them from listening and consequently had a negative effect on their listening comprehension skills. While only 3.4% reported a negative effect of seeing the visual context, which is illustrated in Figure 14 below, the majority of the test takers did not feel at all distracted by seeing the visual context of the speaking situation (65.5%) or the speakers (58.6%). A large number of test takers further disagreed to some extent with the statements that seeing the speakers (27.6%) or the visual context (20.7%) had a negative effect on their listening comprehension skills.

<sup>31</sup>The exact wording of the questionnaire item can be found in Appendix 6.



**Figure 14 AV Qu. items 16-17 (abbr.): Seeing the speakers or the visual context of the speaking situations was distracting and had a negative effect on listening comprehension.<sup>32</sup>**

To see whether there was a significant difference in the performance of the test takers who perceived the visual input in the AV listening test either as a supportive or a distracting factor, three one-way ANOVA tests were conducted. These ANOVA tests, which can be found in Appendix 7 on pages 166 to 167 (see Tables 38 to 43), compared the performances of the test takers differing in their opinion on AV questionnaire items 14, 16 and 17 on the AV listening test. The tests revealed that there was no significant difference in the performance between the students grouped according to their perception of the usefulness of the visual input in supporting comprehension. To reveal whether students who felt that the videos supported listening comprehension performed better in the AV listening test in comparison to the baseline test, paired samples t-test were conducted. These tests, which can be found in Appendix 7 on pages 164 to 165 (see Tables 33 to 35), established that there was no significant difference in the performance of the test takers grouped according to their perception of the visual input as helpful or distracting on the baseline listening test compared to the AV test. Consequently, it can be concluded that the test takers' perception of the visual input as facilitating or hindering comprehension does not correspond to their performance and that neither test takers who felt that the videos aided their comprehension nor the test takers who perceived them as a distracting factor performed significantly differently than the other test takers.

<sup>32</sup> The exact wording of the questionnaire item can be found in Appendix 6.

### 7.5.2. Test taker preference of audio-visual listening tests or audio-only listening test

Both questionnaires asked the test takers about their preference for AV or AO listening tests. The test takers taking the AV listening test clearly preferred audio-visual listening tests over traditional audio-only listening tests, which is shown in Figure 15 below. While 55.2% of the students agreed strongly with the statement “I prefer a video-listening test over a conventional audio-listening test.” and 27.6% to some extent, only 6.9% disagreed strongly or somewhat with this statement. A further 3.4% did not show a clear preference.

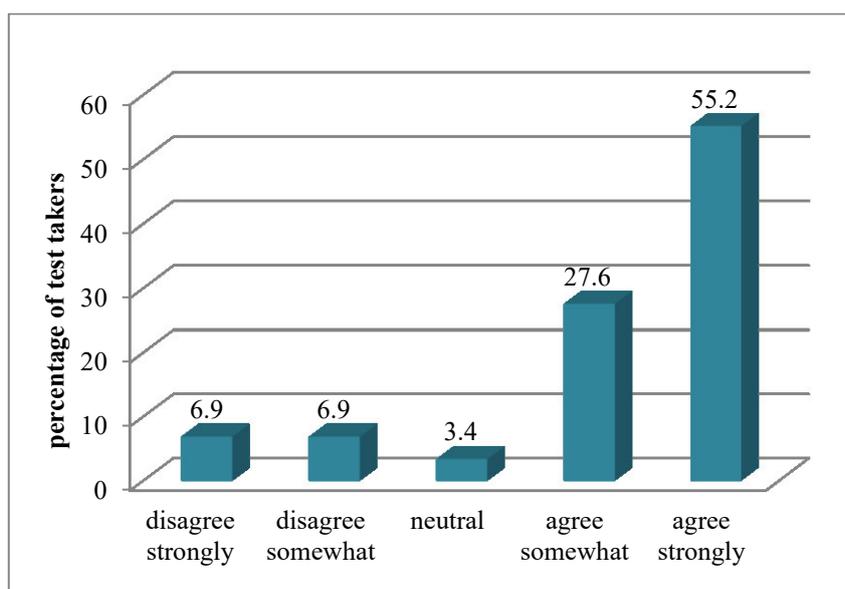


Figure 15 AV Qu. item 18: “I prefer a video-listening test over a conventional audio-listening test.”

A strong preference for AV listening tests over conventional AO listening tests was also expressed by the test takers who took the AO listening test, which is illustrated in Figure 16. Half of these test takers reported that they would definitely preferred a video-listening test where they could see the speakers and the visual context of the speaking situation, while only 5.6% strongly preferred a traditional AO listening test.

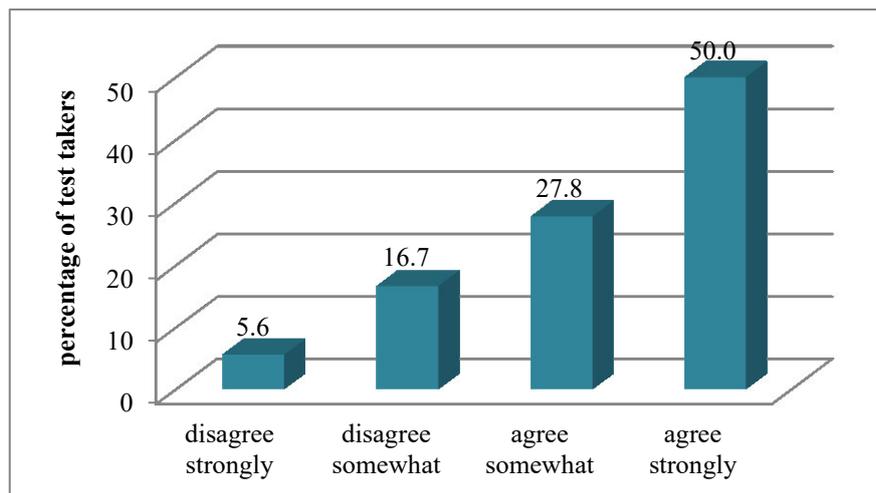


Figure 16 AO Qu. item 12: “I would prefer a video-listening test where I could see the speaker(s) and the visual context rather than a conventional audio-listening test.”

### 7.5.3. Differences in test taker opinion between test takers with different listening ability levels

To answer the first sub-question of the third research question, which is “*Is there a difference in the perception of visual stimuli as helpful or distracting between higher- and lower-level ability test takers?*”, the AV questionnaire answers for items 14, 16 and 17 were analysing according to the test takers’ ability level.

Table 25 depicts the means of the test takers’ answers to these questionnaire items and reveals that the test takers with a lower ability level, who performed significantly better in the AV test than in the baseline listening test, felt that the visual input supported listening comprehension and helped them in answering the test questions to a greater extent than the test takers with an average or higher ability level. The high mean of 3.22 indicates that most students with a lower ability level agreed strongly or somewhat that the videos supported comprehension, while fewer higher-level ability students perceived the visual input as helpful. Furthermore, more test takers with a higher ability level felt that the videos distracted them and had a negative effect on comprehension. This can be seen by the means of the higher-level ability group for items 16 (1.64) and 17 (1.82), which are slightly higher than the means for the average- and lower-level ability group. In general, the table shows that most test takers, regardless of their ability level, perceived the visual input as relatively helpful and did not feel distracted but reveals also a tendency that the

lower-level ability students had an even more positive opinion on the inclusion of visual input in a listening test than the average- and especially the higher-level ability students.

**Table 25 AV Qu. items 14-17 (abbr.): Visual input supported comprehension or was distracting and had a negative effect on comprehension (arranged according to ability level).**

questionnaire item	mean <sup>a</sup>		
	lower ability level	average ability level	higher ability level
AV Qu. item 14 (abbr.): <i>Visual input supported comprehension and helped answer test questions.</i>	3.22	3.11	2.82
AV Qu. item 16 (abbr.): <i>Seeing the speakers was distracting and had a negative effect on listening comprehension.</i>	1.44	1.56	1.64
AV Qu. item 17 (abbr.): <i>Seeing the visual context was distracting and had a negative effect on listening comprehension.</i>	1.22	1.44	1.82
N	9	9	11

<sup>a</sup> 4 = agree strongly, 3 = agree somewhat, 2 = disagree somewhat, 1 = disagree strongly

To see whether this observed tendency is statistically significant, the test takers' ability levels were correlated with the answers to questionnaire items 14, 16 and 17 using Spearman's correlation. The exact results of these correlation tests can be found in Appendix 7 on pages 168 to 168 (see Tables 46 to 48). The correlation coefficient revealed that there was no statistically significant correlation between the test takers' ability levels and their opinions on the inclusion of visual input in the listening test, as the significance values were 0.092 for questionnaire item 14, 0.582 for item 16 and 0.109 for item 17. As a result the question "*Is there a difference in the perception of visual stimuli as helpful or distracting between higher- and lower-level ability test takers?*" can be answered by stating that there is no significant difference in the perception of the visual stimuli between test takers with a lower or higher ability level.

#### **7.5.4. Differences in test taker opinion between students with different learning styles**

To answer the second sub-question of the third research question, "*Is there a difference in the perception of visual stimuli as helpful or distracting between test takers with different*

*learning styles?*”, the answers to items 14, 16 and 17 of the AV questionnaire were correlated with the test takers’ learning styles.

Table 26 below shows the average opinions of the test takers taking the AV listening test arranged according to their learning style on the inclusion of visual input in a listening test. The comparison of means revealed that bimodal visual and auditory learners perceived the visual input as most helpful (mean = 3.25) and felt the least distracted by seeing the speakers (mean = 1) and the visual context (mean = 1), while the bimodal auditory and tactile test taker did not perceive the visuals as a help and felt quite distracted by seeing the speakers. The predominantly tactile test takers felt distracted by the visual contextual information. This analysis showed a tendency of bimodal visual and auditory test takers to perceive visual input in a listening test as a facilitating factor; however, because of the small number of test takers in every category, this analysis has to be treated with caution. Relating this result to the performance of the test takers grouped according to their learning style (see Section 7.2.2), it can be seen that the bimodal visual and auditory learners performed worse in the AV listening test compared to the baseline listening test, although they perceived the visual input as a supportive factor. This suggests that these test takers could not use the visual information to their advantage in the testing situation, which will be further discussed in Chapter 8.

**Table 26 AV Qu. items 14-17: (abbr.): Visual input supported comprehension or was distracting and had a negative effect on comprehension (arranged according learning style).**

Learning style		<i>AV Qu. item 14:</i> Visual input supported	<i>AV Qu. item 16:</i> Seeing speakers distracted	<i>AV Qu. item 17:</i> Seeing visual context distracted
mixed learning style	N	16	16	16
	mean <sup>a</sup>	3.13	1.50	1.31
bimodal - visual/auditory	N	4	4	4
	mean	3.25	1.00	1.00
bimodal - auditory/tactile	N	1	1	1
	mean	2.00	3.00	2.00
predominantly visual	N	3	3	3
	mean	3.00	2.00	2.33
predominantly auditory	N	2	2	2
	mean	2.50	2.00	1.50
predominantly tactile	N	2	2	2
	mean	3.00	1.50	3.00

<sup>a</sup> 4 = agree strongly, 3 = agree somewhat, 2 = disagree somewhat, 1 = disagree strongly

Spearman correlation tests, which can be found in Appendix 7 on page 169 (see Tables 49 to 51), were conducted to see whether there was a significant correlation between the test takers' learning style and their perception of the visual input as a supporting or distracting factor. No significant correlation was found between the test takers' learning style and the answer to questionnaire items 14 and 16. However, the tests revealed that there is a statistically significant correlation, with a significance value of 0.031, between learning style and the perception of the visual context of the speaking situation as a distraction (see Table 51). Consequently, the sub-question investigated in this chapter can be answered by stating that the test takers with different learning styles partly differ in their perception of the visual stimuli.

#### **7.5.5. Test taker comments**

Important insights into the test takers' opinions on the inclusion of visual input in a listening test were also given by their comments on the AV listening test. On the one hand, three test takers reported that they felt distracted by the videos or overwhelmed by the task of attending to the videos, listening to the spoken text and reading and answering the test questions. One of these test takers stated that the AV listening test "was more difficult because you had to read and watch the screen at the same time", which was supported by another test taker who said that she did not find the time to watch the video because she had to focus on the test questions. The third test taker felt distracted only by the first video, which might have been caused by the rapid changes of speakers and scenes, while he reported that the second and third video supported comprehension.

On the other hand, several test takers highlighted positive effects of audio-visual listening tests and stated a clear preference for such tests over traditional AO listening tests. These test takers stated, for example, that seeing what happens in the videos makes it easier to understand the spoken text. Furthermore, a test taker said that AV listening tests are easier to comprehend because she can "identify with the situation". Two test takers referred to the authenticity of AV listening tests stating that these tests are far more "real" and do not appear as "fictional and unrealistic" as conventional listening tests. They further added that they are also able in "normal life" to "see the gestures and the context" and that it is important to connect the gestures of a speaker to the spoken message. Another student stated that "it would be great to have video-listenings in our exams because it can help

some people and those who don't [sic] see positive effects don't [sic] have to watch them because they can look on the [test paper].”

## 7.6. Results in relation to the hypotheses and null-hypotheses

To conclude Chapter 7, the findings of this study are related to the hypotheses and null-hypothesis formulated in the introduction (see Section 1.2).

Hypothesis 1: *The inclusion of visual stimuli leads to improved test taker performance as the test takers utilise the non-verbal cues provided in the visual material to aid comprehension.*

*H1<sub>0</sub>: There is no difference in the test takers' listening comprehension performance in the presence or absence of visual input in a listening comprehension test.*

The comparison of the performance of the control and experimental test groups on the AO or AV test respectively revealed that the inclusion of visual stimuli did not lead to an increase or a decrease in the test takers' listening performance, neither on the test as a whole nor on the individual tasks. The performance of the test groups differed only significantly on one test item, item 5, which could have been caused by the presence of visual cues. The presence of visual input in a listening test did not have a significant effect on test takers with different learning styles either; however, the test takers who were classified as lower ability level test takers based on their performance on the baseline listening test significantly improved their performance in the AV listening test. A significant, though not as drastic increase, was, also found for the lower ability test takers in the control group. Possible reasons for these findings will be discussed in Chapter 8. Taking everything into account, the null-hypothesis H1<sub>0</sub>, stated above, could not be rejected and hypothesis 1 was disproved.

Hypothesis 2a: *Test takers with a lower level of listening ability attend to visual cues to a higher degree than higher-level ability test takers in order to compensate for a lack of comprehension.*

*H2a<sub>0</sub>: There is no difference in the attention paid to visual cues between higher- and lower-level ability test takers.*

Hypothesis 2b: *Test takers with a lower level of listening ability perceive the visual stimuli as more helpful and less distracting than higher-level ability test takers.*

H2b<sub>0</sub>: *There is no difference in the perception of the visual stimuli as helpful or distracting between higher- and lower-level ability test takers.*

In connection with hypothesis 2a, it was found that test takers with a lower level of listening ability attended to the visual input to a slightly, though not significantly, higher degree than higher-level ability test takers. Consequently, null-hypothesis H2a<sub>0</sub> was accepted and hypothesis 2a was disproved. With regard to hypothesis 2b, it was established that the lower-level ability test takers felt least distracted by seeing the speakers and the visual context of the speaking situation and felt more strongly than the students with a higher listening ability level that the visual information supported listening comprehension and assisted them in answering the test questions. However, null-hypothesis H2b<sub>0</sub> needed to be accepted as no significant correlation was found between the test takers' ability levels and the perception of the visual input as helpful or distracting. As a result, hypothesis 2b was not corroborated by the results of this study.

Hypothesis 3a: *Test takers who classify themselves as visual types of learners attend to visual cues to a higher degree than test takers who indicated that they are auditive or kinaesthetic types of learners.*

H3a<sub>0</sub>: *There is no difference in the attention paid to visual cues between test takers with different learning styles.*

Hypothesis 3b: *Test takers who classify themselves as visual types of learners perceive the visual stimuli as more helpful and less distracting than test takers who indicated that they are auditive or kinaesthetic types of learners.*

H3b<sub>0</sub>: *There is no difference in the perception of the visual stimuli as helpful or distracting between test takers with different learning styles.*

With regard to hypothesis 3a, the analysis of the questionnaire data revealed that there was no significant difference in the attention paid to the visual input between test takers with different learning styles. Hence, null-hypothesis H3a<sub>0</sub> was accepted and hypothesis 3a disproved. In connection with hypothesis 3b, no significant difference was found in the

perception of the visual input as helpful between test takers with different learning styles. The analysis did not show a significant correlation between the test takers' learning style and their feeling that seeing the speakers was distracting and had a negative impact on comprehension. However, the test takers grouped according to their learning styles differed significantly in their opinion on questionnaire item 17 ("I think that seeing the visual contexts of the speaking situations distracted me and had a negative effect on my comprehension of the spoken text."). It was revealed that bimodal visual and auditive learners did not feel distracted by seeing the visual context of the speaking situations while predominantly tactile learners perceived seeing the visual context as a distracting and negative factor. Based on these findings, null-hypothesis H3b<sub>0</sub> can be rejected; however, the results of this analysis have to be treated with care because of the small number of test takers in every learning style category, which is discussed in greater detail in Chapter 8.

As the test takers who were classified as predominantly visual learners did not perceive the visual input as more helpful and felt more distracted by seeing the visual context and the speakers than most test takers with other learning styles, hypothesis 3b could not be corroborated by the study's results.

## 8. Discussion and conclusion

This study revealed neither a significant increase nor a decrease in the test takers' listening comprehension performance in the presence of visual input in a video-mediated listening test. Consequently, this study is in line with Coniam's (2001: 1) comparative study of audiotape-mediated and video-mediated listening tests, in which the performance of the test takers taking an audio-visual test did not significantly differ from the performance of the test takers taking an audio-only listening test either. In contrast to the participants in Coniam's (2001) study, who had a negative attitude towards the inclusion of visual input in listening tests and strongly preferred AO tests over AV listening tests (Coniam 2001: 9-10), the subjects in this study showed a clear preference for AV listening tests and indicated a positive opinion on the integration of visual materials in listening tests.

The fact that the presence of visual cues did not lead to a significant increase in test taker performance indicates that the students taking the AV listening tests could not use the visual cues available to aid listening comprehension, which is contrary to the results of several studies (e.g. Ginther 2002; Herron *et al.* 1995; Mueller 1980; Riseborough 1981; Secules *et al.* 1992; Shin 1998; Sueyoshi and Hardison 2005; Wagner 2010) investigating the effects of visual input on test taker performance (see Section 4.1.1). Following Ockey's (2007: 520) and Wagner's (2008: 220) argument that such conflicting study results can be caused by various factors affecting the listening comprehension process, it can be suggested the results of this study might have been affected by factors such as the setting of the study, the number of test items and the sample size.

Most other studies exploring the effects of visual input on test taker performance (i.e. Ginther 2002; Shin 1998; Sueyoshi and Hardison 2005; Wagner 2010) were conducted in an ESL setting in which the study participants are surrounded by English and the non-verbal behaviour of the English language community in their daily life. These participants are consequently more used to reading non-verbal cues and utilising them to aid listening comprehension of English than EFL learners who are mainly in contact with native English on holidays, when watching films in English or through internet surfing and are usually surrounded by a language other than English. Based on this argument, it can be suggested that the participants in this study, which was conducted in an EFL setting, could not gain

an advantage from the presence of visual input in a listening test because they are not used to utilising visual cues to assist listening comprehension in a foreign language. Although the students probably use non-verbal cues extensively to aid listening comprehension in their mother tongue, which is also suggested by the importance they attach to visual cues in everyday communication (see Section 7.3), this might apply only to a limited degree to listening comprehension in a foreign language as the students are usually not in contact with the kinesic behaviour of the target language community. This is particularly apparent concerning lip movements. The participants, for example, in Hardison's (1999, 2003, 2005) studies, who learned English as an L2 in the USA, used cues derived from the lip movements of different speakers to compensate for comprehension difficulties (Hardison 2005: 579) and to support speech perception (Hardison 1999: 214; 2003: 505) and word identification (Hardison 2005: 593). Consequently, they performed significantly better in listening comprehension tests than control groups who did not receive visual information (Hardison 1999: 214, 2003: 504-505). In contrast to these results, the EFL students in this study reported that they paid only little attention to lip movements in everyday communication as well as in the listening test and indicated that the information derived from the lip movements of the speakers supported them only to a very limited and far lesser degree than the information obtained from other elements of NVC, such as the visual context, gestures and facial expressions (See Figure 13, page 89). This suggests that EFL students might not be as used to reading lip movements in relation to English as ESL students and cannot profit as much from cues derived from this element of NVC. A longitudinal study in which the students receive either perceptual training or generally video-based language instruction would be needed in the Austrian school context to investigate whether such training or instruction can have positive effects on the listening comprehension skills of EFL learners and can lead to increased listening comprehension performance, as suggested by Herron *et al.* (1995: 784) and Secules *et al.* (1992: 482-483).

Furthermore, the sample size and the small number of test items might have influenced the results of this study. The generally small number of participants and the particularly small number in the ability levels or learning style groups could have affected the significance of the study's results. This applies also to the small number of test items, especially in the baseline listening test after the removal of various items to improve its reliability. Because of the limited number of items which successfully discriminated between higher- and lower-level ability students, comparisons of the baseline listening test with the AO or AV

listening test need to be treated with caution. A pilot study would have been needed to reveal that task 2 of the baseline test was too easy. Based on the results of a pilot study, this task would have been replaced with another listening task and the reliability of the baseline listening test could have been improved. Generally, these observations imply that a large-scale study, with more participants and a larger number of test items, would be needed to further investigate the effects of visual input on the listening comprehension skills of Austrian EFL students.

By exploring the effects of the presence of visual input on test takers with different ability levels it was revealed that the lower-level ability students performed significantly better on the AV listening test than on the baseline listening test. They improved their performance by 25.85%, while the performance of the test takers in the experimental group who were classified as having an average- or higher-level listening ability remained more or less constant. However, as mentioned in Sections 7.2.1 and 7.6, also the lower-level ability test takers in the control group increased their performance by 16.71% and the analysis of the test data revealed no significant difference between these test takers and the test takers in the experimental groups with a lower ability level. This general increase in the performance of the lower-level ability students could be due to several reasons. Firstly, the classification of the test takers into ability levels groups based on the results of the baseline listening test might not be entirely reliable because of the limited reliability of the baseline listening test and because the performances of students vary considerably from one test to another due to student-related unreliability factors, such as stress, fatigue or anxiety (Brown & Abeywickrama 2010: 28). Consequently, the test takers' abilities might have been over- or underestimated based on the results of the baseline listening test. Secondly, maybe an unknown factor that makes listening difficult for lower-level ability students was not as much present in the AO and AV test as it was in the baseline listening test.

The relatively drastic increase in the listening comprehension performance of the lower-level ability test takers in the experimental group, who outperformed the average-level ability students taking the same test and achieved more points than the lower-level ability students taking the AO test, could have been affected by the presence of visual input and might indicate that they utilised visual cues to assist listening comprehension at least to a certain extent. Although the difference between the lower-level ability students taking the AV and the AO test was insignificant, the study's results might point towards a tendency

for lower-level ability students to profit more from non-verbal cues present in AV listening tests than test takers with a higher proficiency level. This would support the results of Mueller (1980:340), who established that the presence of visual input significantly enhanced the performance of test takers with a lower proficiency level, while the performance of higher-level ability test takers stayed constant. However, further research would be necessary to unveil the effects of visual input on test takers with different ability levels.

Furthermore, the results of the analysis of the questionnaire revealed that the test takers with a lower ability level attended to the videos in the AV test to a slightly, though not significantly, higher degree than the other test takers and felt more strongly that the visual input supported comprehension<sup>33</sup>. In combination with their increased listening comprehension performance this results might further indicate a certain tendency of lower-level ability students to profit from visual cues, which they might use to compensate for comprehension difficulties. This would support Buck's (2001: 50-51) argument that the utilisation of non-verbal cues forms an integral aspect of the compensation skills of ESL and EFL learners. While average- and especially higher-level ability students probably have well-developed listening strategies and compensation skills, which work effectively in audio-only conditions, students at a lower ability level might still be in search for such strategies and turn to the visual input to find assistance in listening comprehension. This would need to be investigated in a further study focusing on the relationship of test taker performance and the perception of visual input as a supportive or hindering factor in FL listening comprehension.

In connection with the performance of bimodal visual and auditive learners and their perception of the visual stimuli, the study draws quite a different picture. These learners paid the closest attention to the videos, had the strongest impression that the visual input supported listening comprehension and felt significantly less distracted by seeing the visual context than the test takers with other learning styles. If visual stimuli can indeed have a positive impact on EFL test takers, it could be assumed that the bimodal visual and auditive learners would profit from the available visual cues which would lead to an enhanced listening comprehension performance. The analysis of the performance of these

---

<sup>33</sup> The results might have been insignificant because of the small number of participants.

test takers showed, however, that they actually performed worse in the presence of visual input in the AV listening test than in the baseline listening test<sup>34</sup>. However, this result has to be treated with caution because of the small number of students in every learning style group and because test taker performance can vary from test to test for a variety of reasons.

With regards to the effects of visual input on test taker performance, this study confirms Buck's (2001: 253) statement that these effects are still unclear. It remains doubtful whether the access to visual cues increases or decreases listening comprehension performance or simply has no effect on test takers. Further research, especially in the context of EFL listening comprehension, is needed to investigate these intricate, controversial and still barely researched effects.

Despite the inconclusive results concerning the effects of visuals on test takers in general and test takers with different learning styles and ability levels in particular, the study clearly revealed that the participants prefer audio-visual listening tests over traditional audio-only listening tests and supported the argument that the inclusion of videos in listening tests leads to increased test authenticity. A positive attitude towards the inclusion of visual materials in listening comprehension tests was also found in the studies conducted by Sueyoshi and Hardison (2005: 661), Progosh (1996: 34) and Wagner (2010: 494).

In connection with the increase in test authenticity, the analysis of the importance the students attach to visual cues in everyday communication and language learning showed that visual cues play an important role in communication and that the students usually engage in activities which involve auditive as well as visual information to improve their listening comprehension skills. As AV listening tests closely reconstruct listening situations which the students encounter outside the classroom, they can be said to constitute a more authentic form of listening test than traditional audio-only listening tests. This is supported by Kellerman's (1990: 279) argument that videos are an "authentic mode of presentation" because they include "the visual information inherent in all normal speech situations".

---

<sup>34</sup> The difference found between the performance of the bimodal auditive and visual learners on the baseline and the AV listening test was, however, insignificant.

Finally, the test takers' feedback also revealed that the large majority of test takers had a very positive opinion on the inclusion of visual materials in listening tests and that several test takers perceived the AV listening test as more authentic than AO listening tests. As authentic materials can serve as motivators in language learning (Bacon & Finnemann 1990: 459; Bachman & Palmer 1996: 32) and can have a positive effect on the students' level of satisfaction (Bacon & Finnemann 1990: 469), AV listening tests, which have a high level of authenticity, might lead to an increase in the students' level of motivation and satisfaction with language learning. Combined with the test takers preference for audio-visual over conventional audio-only listening tests, this argument would be reason enough to integrate videos also in listening testing and thereby align English language teaching and language testing.

## 9. Bibliography

- ACTFL American Council on the Teaching of Foreign Languages. 2015. *ACTFL performance descriptors for language learners*. (2<sup>nd</sup> edition). Alexandria, VA: ACTFL.
- Alcorn, Pete. 2009. "The World in 2200". In *TED Ideas worth spreading*. [https://www.ted.com/talks/pete\\_alcorn\\_s\\_vision\\_of\\_a\\_better\\_world#t-209334](https://www.ted.com/talks/pete_alcorn_s_vision_of_a_better_world#t-209334) (8 Apr. 2016).
- Alderson, J. Charles; Clapham, Caroline; Wall, Dianne. 1995. *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Antes, Theresa A. 1996. "Kinesics: the value of gesture in language and in the language classroom". *Foreign Language Annals* 29(3), 439-448.
- Bachman, Lyle F. 1990. *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, Lyle F.; Palmer, Adrian S. 1996. *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.
- Bacon, Susan M.; Finnemann, Michael D. 1990. "A study of the attitudes, motives, and strategies of university foreign language students and their disposition to authentic oral and written input". *Modern Language Journal* 74(4), 459-473.
- bifie – Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. 2011. *Haupttermin 2 2010/11 Englisch – Hören*. <https://www.bifie.at/node/800> (8 Apr. 2016).
- bifie – Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. 2012. *Haupttermin 2011/12 Englisch – Hören*. <https://www.bifie.at/node/1830> (8 Apr. 2016).
- bifie – Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. 2013. *Übersetzungen der Elterninformation zu den Bildungsstandards E8: Englisch E8*. <https://www.bifie.at/node/2009>.
- bifie – Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. 2015a. *Nebentermin 2 2013/14 – Englisch (AHS) – Hören*. <https://www.bifie.at/node/2800> (8 Apr. 2016).
- bifie – Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. 2015b. *Prüfungsaufgabensammlung Englisch (AHS) – März 2015*. <https://www.bifie.at/node/2837> (8 Apr. 2016).
- bifie – Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. 2016. <https://www.bifie.at/> (7 Apr. 2016).
- Bostrom, Robert N. 1997. "The testing of mother tongue listening skills". In Clapham, Caroline; Corson, David. *Encyclopedia of language and education: language testing and assessment*. Vol. 7. Dordrecht: Kluwer, 21-28.
- Brindley, Geoff. 1998. "Assessing listening abilities". *Annual Review of Applied Linguistics* 18, 171-191.
- Brown, H. Douglas; Abeywickrama, Priyanvada. 2010. *Language assessment: principles and classroom practices*. (2<sup>nd</sup> edition). New York, NY: Pearson Longman.
- Bruggmann, Patrick. *Young world 3 – unit 2: adjectives* <http://www.schule.rorschach.ch/potatoe/5EN002/5EN002.htm> (12 Apr. 2016).
- Buck, Gary. 1991. "The testing of listening comprehension: an introspective study." *Language Testing* 8, 67-91.
- Buck, Gary. 2001. *Assessing listening*. Cambridge: Cambridge University Press.

- Buck, Gary; Tatsuoka, Kikumi. 1998. "Application of the rule-space procedure to language testing: examining attributes of a free response listening test". *Language Testing* 15(2), 119-157.
- Bundesministerium für Bildung und Frauen. 31.08.2015. *Lebende Fremdsprache (Erste, Zweite)*. [https://www.bmbwf.gv.at/schulen/unterricht/lp/lp\\_ahs\\_os\\_lebende\\_fs\\_11854.pdf?4dzgm2](https://www.bmbwf.gv.at/schulen/unterricht/lp/lp_ahs_os_lebende_fs_11854.pdf?4dzgm2) (21 Sept. 2015).
- Burgoon, Judee K. 1985. "Nonverbal signals". In Knapp, Mark L.; Miller, Gerald R. (eds.). *Handbook of interpersonal communication*. Beverly Hills, CA: Sage Publications, 344-390.
- Burgoon, Judee K. 1994. "Nonverbal signals". In Knapp, Mark L.; Miller, Gerald R. (eds.). *Handbook of interpersonal communication*. (2<sup>nd</sup> edition). Thousand Oaks, CA: Sage Publications, 229-285.
- Calm down U2 fans*. 2015. <http://www.u2radio.com/2015/10/calm-down-u2-fans/> (12 Apr. 2016).
- Campbell, Ruth; Dodd, Barbara. 1987. "Hearing by eye". *Quarterly Journal of Experimental Psychology* 32(1), 85-99.
- Canale, Michael; Swain, Merrill. 1980. "Theoretical bases of communicative approaches to second language teaching and testing". *Applied Linguistics* 1(1), 1-47.
- Capper, Simon. 2000. "Nonverbal communication and the second language learner: some pedagogic considerations". *The Language Teacher* 24(5). [http://jalt-publications.org/old\\_tlt/articles/2000/05/capper](http://jalt-publications.org/old_tlt/articles/2000/05/capper) (19 Mar. 2016).
- Coniam, David. 2001. "The use of audio or video comprehension as an assessment instrument in the certificate of English language teachers: a case study". *System* 29, 1-14.
- Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. [http://www.coe.int/t/dg4/linguistic/cadre1\\_en.asp](http://www.coe.int/t/dg4/linguistic/cadre1_en.asp) (21 Sept. 2015).
- Dallow, Peter. 2010. "The visual complex: mapping some interdisciplinary dimensions of visual literacy". In Elkins, James (ed.). *Visual literacy*. New York, NY: Routledge, Taylor & Francis, 91-103.
- Dunkel, Patricia. 1991. "Listening in the native and second/foreign language: toward an integration of research and practice". *TESOL Quarterly* 25(3), 431-457.
- Dunkel, Patricia; Henning, Grant; Chaudron, Craig. 1993. "The assessment of an L2 listening comprehension construct: a tentative model for test specification and development". *Modern Language Journal* 77(2), 180-191.
- Education Planner. 2011. *What's your learning style?*. <http://www.educationplanner.org/students/self-assessments/learning-styles.shtml> (30 Sept. 2015).
- Ekman, Paul; Friesen, Wallace V. 1969. "The repertoire of nonverbal behaviour: categories, origins, usage, and coding". *Semiotica* 1(1), 49-98.
- Ginther, April. 2001. *Research reports: effects of the presence and absence of visuals on performance on TOEFL CBT listening-comprehension stimuli*. (Report 66). Princeton, NJ: Educational Testing Service.
- Ginther, April. 2002. "Context and content visuals and performance on listening comprehension stimuli". *Language Testing* 19(2), 133-167.
- Goldin-Meadow, Susan. 1999. "The role of gesture in communication and thinking". *Trends in Cognitive Sciences* 3(11), 419-429.
- Green, Rita. 2013. *Statistical analyses for language testers*. Houndmills: Palgrave Macmillan.
- Grimes, Tom. 1990. "Audio-video correspondence and its role in attention and memory". *Educational Technology Research and Development* 38, 15-25.

- Gruba, Paul. 1993. "A comparison study of audio and video in language testing". *JALT Journal* 15(1), 85-88.
- Gruba, Paul. 1994. "Design and development of a video-mediated test of communicative proficiency". *JALT Journal* 16(1), 25-40.
- Gruba, Paul. 1997. "The role of video media in listening assessment". *System* 25(3), 335-345.
- Gruba, Paul. 2014. "New media in language assessments". In Kunnan, Antony. *The Companion to language assessment. Vol. II*. Chichester: Wiley Blackwell, 997-1012.
- Gullberg, Marianne. 2006. "Some reasons for studying gesture and second language acquisition (hommage à Adam Kendon)". *International Review of Applied Linguistics* 44, 103-124.
- Hadar, Uri; Wenkert-Olenik, Dafna; Krauss, Robert; Soroker, Nachum. 1998. „Gesture and the processing of speech: neuropsychological evidence". *Brain and Language* 62, 107-126.
- Hardison, Debra M. 1999. "Bimodal speech perception by native and nonnative speakers of English: factors influencing the McGurk effect". *Language Learning* 49(1), 213-283.
- Hardison, Debra M. 2003. "Acquisition of second-language speech: effects of visual cues, context, and talker variability". *Applied Psycholinguistics* 24, 495-522.
- Hardison, Debra M. 2005. "Second-language spoken word identification: effects of perceptual training, visual cues, and phonetic environment". *Applied Psycholinguistics* 26, 579-596.
- Harris, Tony. 2003. "Listening with your eyes: the importance of speech-related gestures in the language classroom". *Foreign Language Annals* 36(2), 180-187.
- Henning, Grant. 1987. *A guide to language testing: developments, evaluation, research*. New York, NY: Newbury House.
- Herron, Carol; Morris, Matthew; Secules, Teresa; Curtis, Lisa. 1995. "A comparison study of the effects of video-based versus text-based instruction in the foreign language classroom". *The French Review* 68(5), 775-795.
- HSVMA Humane Society Veterinary Medical Association. 2008. *HSVMA leadership council member opposes handling of elephants with bullhooks*. [http://www.hsvma.org/leadership\\_council\\_member\\_opposes\\_handling#.VwymVPgaTIU](http://www.hsvma.org/leadership_council_member_opposes_handling#.VwymVPgaTIU) (12 Apr. 2016).
- IBM Corporation. 2012. *IBM SPSS Statistics*. (Version 21). [Computer Programme].
- Jobs, Steve. 2011. "Connecting the dots". In YouTube. <https://www.youtube.com/watch?v=sr07uR75Qk0> (8 Apr. 2016).
- Jones, Linda C. 2003. "Supporting listening comprehension and vocabulary acquisition with multimedia annotations: the students' voice". *CALICO Journal* 21(1), 41-65.
- Kellerman, Susan. 1990. "Lip service: the contribution of the visual modality to speech perception and its relevance to the teaching and testing of foreign language listening comprehension". *Applied Linguistics* 11(3), 272-280.
- Kellerman, Susan. 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.
- Kelly, Spencer D.; Barr, Dale J.; Church, R. Breckinridge; Lynch, Katheryn. 1999. "Offering a hand to pragmatic understanding: the role of speech and gesture in comprehension and memory". *Journal of Memory and Language* 40, 577-592.
- Kintsch, Walter. 1998. *Comprehension: a paradigm for cognition*. Cambridge, MA: Cambridge University Press.
- Madre Luna. 2011. <http://www.famousfix.com/topic/madre-luna/photos> (12 Apr. 2016).

- McGurk, Harry; MacDonald, John. 1976. "Hearing lips and seeing voices". *Nature* 264, 746-748.
- McNeill, David. 1985. "So you think gestures are nonverbal?" *Psychological Review* 92(3), 350-371.
- Messick, Samuel. 1996. "Validity and washback in language testing". *Research Report* 96(17), 1-18.
- Morley, Joan. 1991. "Listening comprehension in second/foreign language instruction". In Celce-Murcia, Marianne (ed.). *Teaching English as a second or foreign language*. (2<sup>nd</sup> edition). Boston, MA: Heinle & Heinle, 81-106.
- Mueller, Gunther A. 1980. "Visual contextual clues and listening comprehension: an experiment". *The Modern Language Journal* 64(3), 335-340.
- Ockey, Gary J. 2007. "Construct implications of including still image or video in computer-based listening tests". *Language Testing* 24(4), 517-537.
- Pallant, Julie. 2001. *SPSS survival manual: a step by step guide to data analysis using SPSS for Windows (Versions 10 and 11)*. Buckingham: Open University Press.
- Philpott, Jeffrey S. 1983. "The relative contribution to meaning of verbal and nonverbal channels of communication: a meta-analysis". MA thesis, University of Nebraska.
- Progosh, David. 1996. "Using video for listening assessment: opinions of test takers". *TESL Canada Journal/ La Revue TESL du Canada* 14(1), 34-44.
- Richards, Jack C. 1983. "Listening comprehension: approach, design, procedure". *TESOL Quarterly* 17(2), 219-240.
- Riseborough, Margaret. 1981. "Physiographic gestures as decoding facilitators: three experiments exploring a neglected facet of communication". *Journal of Nonverbal Behavior* 5(3), 172-183.
- Rost, Michael. 1990. *Listening in language learning*. London: Longman.
- Rubin, Joan. 1994. "A review of second language listening comprehension research". *The Modern Language Journal* 78(2), 199-221.
- Rubin, Joan; Mendelsohn, David J. 1995. *A guide for the teaching of second language listening*. San Diego, CA: Dominic Press.
- Sapir, Edward. 1999 [1928]. "The unconscious patterning of behaviour in society (1928)". In Darnell, Regna; Irvine, Judith T.; Handler, Richard (eds.). *The collected works of Edward Sapir: culture Vol. III*. Berlin: Mouton de Gruyter, 155-172.
- Secules, Teresa; Herron, Carol; Tomasello, Michael. 1992. "The effect of video context on foreign language learning". *The Modern Language Journal* 76(4), 480-490.
- Sherman, Jane. 2003. *Using authentic video in the language classroom*. Cambridge: Cambridge University Press.
- Shin, Dongil. 1998. "Using videotaped lectures for testing academic listening proficiency". *International Journal of Listening* 12(1), 57-80.
- Sueyoshi, Ayano; Hardison, Debra M. 2005. "The role of gestures and facial cues in second language listening comprehension". *Language Learning* 55(4), 661-699.
- Suvorov, Ruslan. 2011. "The effects of context visuals on L2 listening comprehension". *Cambridge ESOL: Research Notes* 45, 2-8.
- Universität Innsbruck. 2012. *SRP - Standardisierte Reifeprüfung: Testspezifikationen: B2 Listening Test Specifications*. [https://www.uibk.ac.at/srp/PDFs/Listening\\_Test\\_Specifications\\_B2\\_v10.pdf](https://www.uibk.ac.at/srp/PDFs/Listening_Test_Specifications_B2_v10.pdf) (7 Apr. 2016).
- Wagner, Elvis. 2002. "Video listening tests: a pilot study". *Working Papers in TESOL & Applied Linguistics* 2(1). <http://tesol-dev.journals.cdrs.columbia.edu/wp-content/uploads/sites/12/2015/05/4.-Wagner-2002.pdf> (17 Aug. 2015).
- Wagner, Elvis. 2007. "Are they watching? Test takers viewing behaviour during an L2 video listening test". *Language Learning & Technology* 11(1), 67-86.

- Wagner, Elvis. 2008. "Video listening tests: what are they measuring?". *Language Assessment Quarterly* 5(3), 218-243.
- Wagner, Elvis. 2010. "The effect of the use of video texts on ESL listening test-taker performance". *Language Testing* 27(4), 493-513.
- Weyers, Joseph R. 1999. "The effects of authentic video on communicative competence". *The Modern Language Journal* 83(3), 339-349.
- Widdowson, Henry G. 1978. *Teaching language as communication*. Oxford: Oxford University Press.
- YouTube. 2012. *Comic books go digital*. <https://www.youtube.com/watch?v=uSJW-MC7jDM> (8 Apr. 2016).

## **Appendix 1: Abstracts**

### **Abstract**

This study compares an audiotape-mediated and a video-mediated listening test in order to determine whether the presence of visual input, which provides access to non-verbal cues, leads to increased listening comprehension performance in Austrian students who are learning English as a foreign language. It also investigates the students' interaction with visual input in everyday communication and provides insights into the extent to which they engage with visual materials in a testing situation. Concerning their interaction with visual stimuli in a testing situation, differences are explored between higher- and lower-level ability students and students with different learning styles. Finally, the students' opinions are examined on the inclusion of videos in listening tests and the usefulness of visual input to support listening comprehension and differences are investigated in the opinion of students with different learning styles and ability levels. Prior research on the effects of visuals on foreign and second language listening comprehension has revealed rather conflicting results, suggesting that visual input can both assist and hinder listening comprehension. An analysis of the results of this study established that the presence of visual input did not affect the listening comprehension performance of students; however, it led to increased test authenticity and potentially also student motivation as the students clearly prefer video-mediated listening tests over traditional audio-only listening tests.

## **Zusammenfassung**

Diese Studie vergleicht eine herkömmliche Hörverständnisübung mit einer Video-Hörverständnisübung, bei welcher die Schülerinnen und Schüler die Sprecher nicht nur hören, sondern auf Videos auch sehen können, um festzustellen, ob sich Videos, welche Zugang zu nicht-verbaler Information anbieten, positiv auf das Hörverständnis von österreichischen Schülerinnen und Schülern, die Englisch als Fremdsprache lernen, auswirkt. Des Weiteren beschäftigt sich die Studie mit der Bedeutung der visuellen Wahrnehmung im alltäglichen Kommunikationsprozess und erforscht die Interaktion der Schülerinnen und Schüler mit Videos in audio-visuellen Hörverständnis-Tests. Bezüglich der Interaktion der Schülerinnen und Schüler mit den visuellen Stimuli in der Testsituation werden Unterschiede zwischen Schülerinnen und Schülern mit unterschiedlichen Hörverständnisfähigkeiten und unterschiedlichen Lerntypen erforscht. Schließlich wird auch die Meinung der Schülerinnen und Schüler zur Verwendung von Videos in Hörverständnis-Tests und zur Nützlichkeit solcher Videos zur Unterstützung des Hörverständnisses in der Fremdsprache Englisch erfragt und weitere Unterschiede in diesem Bereich zwischen Schülerinnen und Schülern mit unterschiedlichen Fähigkeiten und Lerntypen festgestellt. Bisherige Untersuchungen der Effekte von Videos auf das fremd- oder zweitsprachliche Hörverständnis brachten unterschiedliche, oft zwiespältige, Ergebnisse hervor und zeigten einerseits, dass visuelle Informationen das Hörverständnis unterstützen jedoch andererseits auch vermindern können. Die Auswertung der Ergebnisse dieser Studie zeigte, dass das Vorhandensein von visuellen Informationen keinen Einfluss auf das Hörverständnis der Schülerinnen und Schüler hat. Jedoch wurde auch festgestellt, dass die Verwendung von Videos in Hörverständnis-Tests die Authentizität der Tests steigern kann und sich möglicherweise auch positiv auf die Motivation der Schülerinnen und Schüler auswirkt, da diese eindeutig eine Präferenz für Video-Hörverständnisübungen im Vergleich zu herkömmlichen Hörverständnisübungen zeigen.

## Appendix 2: Taxonomy of listening comprehension skills

Table 27 Richards' (1983) taxonomy of listening comprehension skills: Conversational listening (Richards 1983: 228-229)

<b>Microskills: Conversational listening</b>	
1.	ability to retain chunks of language of different lengths for short periods
2.	ability to discriminate among the distinctive sounds of the target language
3.	ability to recognize the stress patterns of words
4.	ability to recognize the rhythmic structure of English
5.	ability to recognize the functions of stress and intonation to signal the information structure of utterances
6.	ability to identify words in stressed and unstressed positions
7.	ability to recognize reduced forms of words
8.	ability to distinguish word boundaries
9.	ability to recognize typical word order patterns in the target language
10.	ability to recognize vocabulary used in core conversational topics
11.	ability to detect keywords (i.e., those which identify topics and propositions)
12.	ability to guess the meanings of words from the contexts in which they occur
13.	ability to recognize grammatical word classes (parts of speech)
14.	ability to recognize major syntactic patterns and devices
15.	ability to recognize cohesive devices in spoken discourse
16.	ability to recognize elliptical forms of grammatical units and sentences
17.	ability to detect sentence constituents
18.	ability to distinguish between major and minor constituents
19.	ability to detect meanings expressed in differing grammatical forms/ sentence types (i.e., that a particular meaning maybe expressed in different ways)
20.	ability to recognize the communicative functions of utterances, according to situations, participants, goals
21.	ability to reconstruct or infer situations, goals, participants, procedures
22.	ability to use real world knowledge and experience to work out purposes, goals, settings, procedures
23.	ability to predict outcomes from events described
24.	ability to infer links and connections between events
25.	ability to deduce causes and effects from events
26.	ability to distinguish between literal and implied meanings
27.	ability to identify and reconstruct topics and coherent structure from ongoing discourse involving two or more speakers
28.	ability to recognize markers of coherence in discourse, and to detect such relations as main idea, supporting idea, given information, new information, generalization, exemplification
29.	ability to process speech at different rates
30.	ability to process speech containing pauses, errors, corrections
31.	ability to make use of facial, paralinguistic, and other clues to work out meanings
32.	ability to adjust listening strategies to different kinds of listener purposes or goals
33.	ability to signal comprehension or lack of comprehension, verbally and non-verbally

**Table 28 Richards' (1983) taxonomy of listening  
comprehension skills: Academic listening (Richards 1983: 229-230)**

**Microskills: Academic listening**

1. ability to identify purpose and scope of lecture
2. ability to identify topic of lecture and follow topic development
3. ability to identify relationships among units within discourse (e.g., major ideas, generalizations, hypotheses, supporting ideas, examples)
4. ability to identify role of discourse markers in signalling structure of a lecture (e.g., conjunctions, adverbs, gambits, routines)
5. ability to infer relationships (e.g., cause, effect, conclusion)
6. ability to recognize key lexical items related to subject/topic
7. ability to deduce meanings of words from context
8. ability to recognize markers of cohesion
9. ability to recognize function of intonation to signal information structure (e.g., pitch, volume, pace, key)
10. ability to detect attitude of speaker toward subject matter
11. ability to follow different modes of lecturing: spoken, audio, audio-visual
12. ability to follow lecture despite differences in accent and speed
13. familiarity with different styles of lecturing: formal, conversational, read, unplanned
14. familiarity with different registers: written versus colloquial
15. ability to recognize irrelevant matter: jokes, digressions, meanderings
16. ability to recognize functions of non-verbal cues as markers of emphasis and attitude
17. knowledge of classroom conventions (e.g., turn taking, clarification, requests)
18. ability to recognize instructional/learner tasks (e.g., warnings, suggestions, recommendations, advice, instructions)

## **Appendix 3: Letter to parents**

03. Jänner 2016

### **Elternbrief**

**Sehr geehrte Eltern und Erziehungsberechtigte!**

**Liebe Schülerinnen und Schüler der 7. Klassen!**

Mein Name ist Julia Zirhan. Ich bin eine ehemalige Schülerin des ORGs Vöcklabruck und studiere jetzt Englisch und Geschichte auf Lehramt an der Universität Wien. Ich schreibe gerade an meiner Diplomarbeit im Fach Englisch, die sich mit dem Vergleich von traditionellen englischen Hörverständnisübungen, bei welchen nur die Tonspur zu hören ist, und Video-Hörverständnisübungen, bei welchen die Schülerinnen und Schüler die Sprecher nicht nur hören, sondern auf Videos auch sehen können, beschäftigt.

Die Verwendung von Videos in Hörverständnisübungen wird in der Fachliteratur in den letzten Jahren immer wieder diskutiert. Es wird argumentiert, dass Video-Hörverständnisübungen viel realitätsnaher sind als gewöhnliche Hörverständnisübungen, da in den meisten Situationen, in denen Englisch gehört wird, auch die visuelle Wahrnehmung eine bedeutende Rolle spielt.

Ich möchte in meiner Diplomarbeit herausfinden, ob die Verwendung von Videos in englischen Hörverständnisübungen das Hörverständnis der Schülerinnen und Schüler verbessert. Des Weiteren möchte ich die Meinung der Schülerinnen und Schüler zur Verwendung von Videos in Hörverständnisübungen erfragen und herausfinden, ob es hierbei einen Unterschied zwischen Schülerinnen und Schülern mit unterschiedlichen Lerntypen (visuell, auditiv oder kinästhetisch) gibt.

Um dieses Thema zu untersuchen, habe ich zwei kurze Fragebögen entwickelt und zwei Hörverständnistests zusammengestellt. Ich möchte, in der Englischstunde am 02.02.2016, eine traditionelle Hörverständnisübung mit den Schülerinnen und Schüler der 7. Klassen durchführen, um das Hörverständnislevel der Schülerinnen und Schüler festzustellen und einen Vergleichswert zu haben. Dafür habe ich eine Matura-Hörverständnisübung des Bundesinstituts für Bildungsforschung, Innovation und Entwicklung des Österreichischen Schulwesens (Bifie) ausgesucht, die den Schülerinnen und Schülern gleichzeitig als Übung für die Englisch-Matura dient.

In der Englischstunde am 03.02.2016, möchte ich dann mit einer der drei Englischgruppen, die als Kontrollgruppe dient, eine weitere traditionelle Hörverständnisübung (nur mit Tonspur) durchführen, während die anderen zwei Gruppen die gleiche Hörverständnisübung jedoch mit den dazugehörigen Videos bekommen. Diese Hörverständnisübungen habe ich aus weiteren standardisierten Maturavorbereitungsmaterialien von Bifie, die Videos als Grundlage der Hörverständnisübungen hatten, zusammengestellt. Somit dienen auch diese Hörverständnisübungen der weiteren Vorbereitung der Schülerinnen und Schüler auf die Matura. Die zwei kurzen Fragebögen werden jeweils nach den Hörverständnisübungen am 02.02. und am 03.02.2016 von den Schülerinnen und Schülern ausgefüllt.

---

Ich wäre Ihnen sehr dankbar, wenn Sie mich in meiner Diplomarbeitsstudie unterstützen könnten und mir die Untersuchung einer neuen Art von Hörverständnisübung, die vielleicht eines Tages zum Schulalltag gehören wird, ermöglichen!

**Herzlichen Dank im Voraus!**

*Julia Zirhan*

### **Bestätigung**

Ich, \_\_\_\_\_, bestätige hiermit, dass ich darüber informiert worden bin, dass mein/e Sohn/Tochter, \_\_\_\_\_, an dieser Studie teilnimmt, die am 02.02. und am 03.02.2016 durchgeführt wird.

**Datum:** \_\_\_\_\_

**Unterschrift** der Eltern/Erziehungsberechtigten

oder der volljährigen Schülerinnen und Schüler: \_\_\_\_\_

\_\_\_\_\_

**Appendix 4: Baseline listening test (including first student questionnaire, correction key and test specifications)**

# **Baseline listening test**

Based on **HÖREN E HT2 2010/11**

## **Schriftliche Reifeprüfung aus English**

2. Haupttermin

10. Mai 2011

First name: \_\_\_\_\_

Second name: \_\_\_\_\_

English teacher: \_\_\_\_\_

Points: \_\_\_\_\_ out of **23**

### **Instructions**

1. This test contains 3 tasks and 23 questions.
2. Write all your answers in this test booklet.
3. Please fill in the questionnaire after the listening test.

Test source: <https://www.bifie.at/node/800>

## Task 1

\_\_\_ / 8 P.

You are going to listen to a recording about Miguel Sabido's soap operas. First you will have 45 seconds to study the task below, then you will hear the recording twice. While listening, choose the correct answer for questions 1-8. Put a cross in the correct box. The first one (0) has been done for you.

After the second listening, you will have 45 seconds to check your answers.

35



### The Sabido Method

- Q 0 **The interviewer Brook Gladstone suggests that soap operas are**
- A less influential than newspapers.
  - B mostly watched by uncritical viewers.
  - C unlikely to change people's minds.
  - D able to touch people emotionally.
- Q 1 **Inspired by *Simple Maria*, Miguel Sabido aimed to**
- A become more successful commercially.
  - B teach TV viewers through soap operas.
  - C focus more on entertainment.
  - D mirror everyday life in South America.
- Q 2 **Hannah Rosen says that a traditional telenovela focuses on the**
- A personal growth of the main character.
  - B main character's love and happiness.
  - C extraordinary lives of the rich and beautiful.
  - D sad events in the main character's life.

<sup>35</sup> *Madre Luna*. 2011. <http://www.famousfix.com/topic/madre-luna/photos> (12 Apr. 2016).

- Q 3 The main characteristic of a Sabido telenovela is the**
- A likeable character behaving unexpectedly.
  - B beautiful blond as a main character.
  - C description of the sad lives of poor people.
  - D focus on love, romance and relationships.
- Q 4 Screen writers for Sabido telenovelas should**
- A partly share their audience's values.
  - B do their job unspoiled by any training.
  - C try to create roles for local actors.
  - D avoid topics that hurt people's feelings.
- Q 5 PCI is an organization that first focused on**
- A economic development in poor countries.
  - B educating the native population.
  - C keeping population growth stable.
  - D raising money for poor countries.
- Q 6 The interviewer is amazed that Sabido's telenovelas**
- A are more popular in Africa and Asia.
  - B have almost replaced traditional ones.
  - C are so popular among poorer people.
  - D can really change how people act.
- Q 7 In an experiment one part of Tanzania's population was cut off from**
- A news delivery.
  - B TV telenovelas.
  - C radio broadcasts.
  - D the internet.
- Q 8 The experiment showed that telenovelas mainly change how we**
- A act in certain situations.
  - B think about something.
  - C plan our lives.
  - D treat other people.

**Task 2**

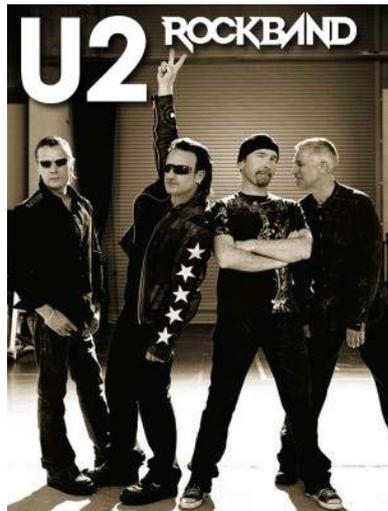
\_\_\_ / 8 P.

You are going to listen to a recording about the band U2. First you will have 45 seconds to study the task below, then you will hear the recording twice. While listening, match the beginnings of the sentences (9-16) with the sentence endings (A-K). There are two sentence endings that you should not use. The first one (0) has been done for you.

After the second listening, you will have 45 seconds to check your answers.

**U2**

Q 0	The speaker interviewed U2 in a _____.	<i>D</i>
Q 9	U2 first performed in a _____.	
Q 10	Bono changed his name because of _____.	
Q 11	When singing a song by Peter Frampton, Bono started to like _____.	
Q 12	A short time ago Bono worked with _____.	
Q 13	Relatives on his mother's side were _____.	
Q 14	During the economic crisis Bono worked with _____.	
Q 15	He hates the fact that the world doesn't do enough to help _____.	
Q 16	Politicians are scared of _____.	



A	his voice.
B	TV talent contest.
C	children in developing countries.
D	rehearsal studio in Dublin.
E	some famous Americans.
F	the large number of U2 fans.
G	a quarrel with his father.
H	popular punk rockers.
I	people fighting AIDS.
J	people who travelled around selling goods.
K	school sports hall.

<sup>36</sup> *Calm down U2 fans.* 2015. <http://www.u2radio.com/2015/10/calm-down-u2-fans/> (12 Apr. 2016)

### Task 3

\_\_\_ / 7 P.

You are going to listen to a recording about elephants in zoos. First you will have 45 seconds to study the task below, then you will hear the recording twice. While listening, answer the questions (17-23) using a maximum of 4 words. Write your answers in the spaces provided. The first one (0) has been done for you.

After the second listening, you will have 45 seconds to check your answers.

### Life Behind Bars



37

Q 0	How long do elephants in zoos usually live?	<i>very short</i>
Q 17	By how much are the lives of Asian elephants reduced in zoos?	
Q 18	Name one thing most zoo elephants suffer from.	
Q 19	What does being born in a zoo mean for an elephant?	
Q 20	Name one thing that might cause stress and health problems for zoo elephants.	
Q 21	What did Mrs Stevenson and her colleagues publish?	
Q 22	Why do wild elephants walk from place to place?	
Q 23	What are zoo elephants given the chance to do?	

<sup>37</sup> HSVMA Humane Society Veterinary Medical Association. 2008. *HSVMA Leadership Council Member Opposes Handling of Elephants with Bullhooks*. [http://www.hsvma.org/leadership\\_council\\_member\\_opposes\\_handling#.VwymVPgaTIU](http://www.hsvma.org/leadership_council_member_opposes_handling#.VwymVPgaTIU) (12 Apr. 2016).

# Questionnaire

## PERSONAL BACKGROUND INFORMATION

Please fill in the information below.

1. **Gender:**  male  female

2. **What is/are your native language/s** (several answers possible)? \_\_\_\_\_  
\_\_\_\_\_

3. **In which situations are you in contact with English?**

in school

reading books

with friends

reading the newspaper

with family

listening to the radio

internet surfing

watching TV

reading or writing e-mails,  
letters, blogs...

Other(s) \_\_\_\_\_  
\_\_\_\_\_

## GENERAL QUESTION

4. **Did you take this listening before?**

Yes  No  Partly

If partly, which listening task(s) did you know? \_\_\_\_\_

## LEARNING STYLES

Please answer the following 20 questions to find out what learning type or what combination of learning types you are. The results will be given to you in another lesson.

1. **What kind of book would you like to read for fun?**

A book with a lot of pictures in it

A book with a lot of words

A book with word searches or crossword puzzles

**2. When you are not sure how to spell a word, what are you mostly likely to do?**

- Write it down to see if it looks right
- Spell it out loud to see if it sounds right
- Write letters in the air

**3. You are out shopping for clothes and you're waiting in line to pay. What are you most likely to do while you are waiting?**

- Look around at other clothes on the racks
- Talk to the person next to you in line
- Move back and forth or play with something

**4. When you see the word "cat", what do you do first?**

- Picture a cat in your mind
- Say the word "cat" to yourself
- Think about being with a cat (and, for example, stroking its head)

**5. What's the best way for you to study for a test?**

- Read the book or your notes and review pictures or charts
- Have someone ask you questions that you can answer out loud
- Make up index cards that you can review

**6. What's the best way for you to learn about how something works?**

- Get someone to show you
- Read about it or listen to someone explain it
- Try it out on your own

**7. If you went to a disco, what would you be most likely to remember the next day?**

- The faces of the people who were there
- The music that was played
- The dance moves you did

**8. What do you find most annoying when you are trying to study?**

- People walking past you
- Loud noises
- An uncomfortable chair

**9. When you are angry, what are you most likely to do?**

- Put on your “mad” face
- Shout
- Slam doors

**10. When you are happy, what are you most likely to do?**

- Smile from ear to ear
- Talk a lot
- Behave hyperactively

**11. When in a new place, how do you find your way around?**

- Look for a map that shows you where everything is
- Ask someone for directions
- Just start walking around until you find what you are looking for

**12. Of these three classes, which is your favourite?**

- Art class
- Music class
- Sports class

**13. When you hear a song on the radio, what are you most likely to do?**

- Picture the video that goes along with it
- Sing or hum along to the music
- Start dancing or tapping your foot

**14. What do you find most distracting when in class?**

- Lights that are too bright or too dim
- Noises from the hallway or outside the building
- The temperature being too hot or too cold

**15. What do you like to do to relax?**

- Read
- Listen to music
- Exercise (for example, walk, run, play sport, etc.)

**16. What is the best way for you to remember a friend's phone number?**

- Picture the numbers on the phone display
- Say it out loud over and over
- Write it down or store it in your phone contact list

**17. If you won a game, which of these three prizes would you choose?**

- A poster for the wall
- A music CD or mp3 download
- A game of some kind (or a football or soccer ball, etc.)

**18. Where would you rather go to with a group of friends?**

- A movie
- A concert
- An entertainment park

**19. What are you most likely to remember about new people you meet?**

- Their face but not their name
- Their name but not their face
- What you talked about with them

**20. When you give someone directions to your house, what are you most likely to tell them?**

- A description of building and landmarks they will pass on the way
- The names of the roads or streets they will be on
- "Follow me – it will be easier if I just show you how you get there"

**THANK YOU!!**



38

<sup>38</sup> Bruggmann, Patrick. *Young world 3 – unit 2: adjectives* <http://www.schule.rorschach.ch/potatoe/5EN002/5EN002.htm> (12 Apr. 2016).

## Correction key: *Baseline listening test*<sup>39</sup>

### **Task 1: The Sabido Method**

#### **Key**

Q 0	Q 1	Q 2	Q 3	Q 4	Q 5	Q 6	Q 7	Q 8
D	B	D	A	A	C	D	C	B

### **Task 2: U2**

#### **Key**

Q 0	Q 9	Q 10	Q 11	Q 12	Q 13	Q 14	Q 15	Q 16
D	K	G	A	I	J	E	C	F

### **Task 3: Elephants behind bars**

#### **Key**

	<b>accepted</b>	<b>not accepted</b>
Q 0	<i>very short</i>	
Q 17	by half 50% half as long half of their lives a half a life	one third 60 percent 58 percent 50 more than a half
Q 18	foot problems overweight lameness they can't walk walking problems lame have no normal walk	foth problems food problems children die laming
Q 19	less chance of survival less chance of long life reduces chance of survival shorter life reduce chance for life no high life expectations reduce survival reduce surviving chance stress you die early	unnatural survival are with their mother reduce the change of survival unnatural
Q 20	husbandry small spaces unnatural environment hard floors unnatural conditions	transport visitors environment change of environment have no place they can't move

<sup>39</sup> The correction key and the corresponding justifications can be found on bifie's homepage under the following link: <https://www.bifie.at/node/800> (bifie 2010).

Q 21	guidelines (for elephants) guidelines from research husbandry guidelines	guides research problems of elephants
Q 22	to find food	
Q 23	go for walks take exercise forage for food look for food exercise them can be exercised	substrates provided food do tasks

## Test specifications: Baseline listening test<sup>40</sup>

**General statement of purpose:** The purpose of this listening test is to assess the listening proficiency of three groups of 7<sup>th</sup> grade students participating in this diploma thesis study who are learning English as a foreign language. The test is further designed to establish whether there is a difference in the level of listening ability between the participating student groups.

**Test category:** proficiency test

**Description of test takers:** The test takers are 50 Austrian pupils attending the 7<sup>th</sup> grade of an upper secondary school in Vöcklabruck in Upper Austria. They are 16 to 17 years old and have German as their mother tongue. Thirty six of them are female while 14 are male.

**Test level:** B2

**Target language use situation:** The test takers can comprehend authentic and unmodified native English if the speaker or speakers use a standard dialect of English.

**Mode of presentation:** audio only

### Test construct<sup>41</sup>

#### General purpose

- a) listening for specific information
- b) listening for important details
- c) search listening
- d) listening to infer (propositional) meaning
- e) listening to deduce the meaning of a word

#### Specific purpose

1. Can understand the main ideas of propositionally and linguistically complex speech on both concrete and abstract topics delivered in a standard dialect [...] (Council of Europe 2001: 66).
2. Can follow extended speech and complex lines of argument provided the topic is reasonably familiar and the direction of the talk is sign posted by explicit markers (Council of Europe 2001: 66).
3. Can follow the essentials of lectures, talks and reports and other forms of academic / professional presentation which are propositionally and linguistically complex (Council of Europe 2001: 67).
4. can understand most radio documentaries and most other recorded or broadcast material delivered in standard dialect and can identify the speaker's mood, tone etc. (Council of Europe 2001: 68).
5. Can use a variety of strategies to achieve comprehension including listening for main points and checking comprehension by using contextual clues (Council of Europe 2001: 72).
6. Can understand standard spoken language, live or broadcast, on both familiar and unfamiliar topics normally encountered in personal, social, academic or vocational life.

<sup>40</sup> These test specifications are modelled on the test specifications operationalised in Matura listening exams (Universität Innsbruck 2012) and the guidelines for devising test specifications by Alderson, Clapham and Wall (1995: 11-15) and Bachman and Palmer (1996: 48-49).

<sup>41</sup> This listening comprehension construct is a modified version of the construct definition of listening comprehension developed on behalf of bifie (Universität Innsbruck 2012) which used the descriptors for language learners contained in the CEFR (Council of Europe 2001).

only extreme background noise, inadequate discourse structure and/or idiomatic usage influences the ability to understand (Council of Europe 2001: 66).

### Structure of the test

- *Number of listening tasks:* 3
- *Number of test items altogether:* 20-25
- *Number of test items per task:* 6-9

### Time allotment

- *Time for the entire test:* approximately 30 minutes
- *Time for each task:* approximately 10 minutes (including the time for playing the recordings twice and the pauses before, after and between the tasks)
- Recordings are played twice. Students have 45 seconds time before the first recording to study the tasks, 20 seconds between the tasks to complete the preceding task and 45 seconds after hearing the last recording a second time to complete the tasks and check the answers.

### Test methods

- multiple choice
- matching
- limited response (maximum of 4 words)

### Input

- *text types:* spoken  
recordings of monologues, dialogues, conversations – e.g. interviews, documentaries, reports, speeches, talks, etc.
- *authenticity of input:* authentic, non-scripted input or partly semi-scripted input (for example, in interviews and documentaries)
- *discourse type:* descriptive, narrative, persuasive
- *nature of content:* concrete
- *speaker characteristics:* male and female speakers of different ages using standard dialects of English
- *source of listening texts and tasks:* bifie (2016)

### Test rubrics

- *instructions:* are provided in the target language in written and spoken form for each listening comprehension task.
- one example answer will be given

### Assessment criteria

- *objective testing*
- *weighing per item:*

multiple choice	- 1 point per test item
matching	- 1 point per test item
limited response	- 1 point if the answer is appropriate and grammatically accurate 0,5 points if the answer is appropriate, however, grammatically incorrect in a way that is not included in the provided correction key

spelling mistakes are not counted as  
mistakes

- correction key is available at the homepage of bife (2016)

**Appendix 5: Audio-only listening test (including AO student questionnaire, correction key and test specifications)**

## **Audio-only listening test**

First name: \_\_\_\_\_

Second name: \_\_\_\_\_

English teacher: \_\_\_\_\_

Points: \_\_\_\_\_ out of **22**

### **Instructions**

1. This test contains 3 tasks and 22 questions.
2. Write all your answers in this test booklet.
3. Please fill in the questionnaire after the listening test.

Test sources: <https://www.bifie.at/node/2800>

<https://www.bifie.at/node/2837>

<https://www.bifie.at/node/1830>

**Task 1**

\_\_\_ / 7 P.

You are going to listen to a recording about a new era in comic book design. First you will have 45 seconds to study the task below, then you will hear the recording twice. While listening, complete the sentences (1-7) using a maximum of four words. Write your answers in the spaces provided on the answer sheet. The first one (0) has been done for you.

After the second listening, you will have 45 seconds to check your answers.

**A new generation of comic books**

Q 0	For decades Golden Apple's comic books have been bought by both _____.
Q 1	One recent change has been the introduction of _____.
Q 2	The costs of digital and printed comics _____.
Q 3	Mark Wait works in a format which _____.
Q 4	With Operation Ajax, it is possible for the reader to _____.
Q 5	Comic book readers today want to see _____.
Q 6	Luckily for Ryan Liewbowitz, the new development _____.
Q 7	The new development is expected to appeal to _____.

## A new generation of comic books

Q 0	<i>adults and children</i>
Q 1	
Q 2	
Q 3	
Q 4	
Q 5	
Q 6	
Q 7	

von der  
Lehrperson  
auszufüllen

richtig falsch

## Task 2

\_\_\_ / 6 P.

You are going to listen to a recording about Pete Alcorn comparing his view of the world with that of Thomas Robert Malthus. First you will have 45 seconds to study the task below, then you will hear the recording twice. While listening, choose the correct answer (A, B, C or D) for each question (8-13). Put a  in the correct box. The first one (0) has been done for you.

After the second listening, you will have 45 seconds to check your answers.



### On the world's future

Q 0 **Malthus' idea was that**

- A we need more than one mental model of the world.
- B the rising number of people on Earth is a disaster.
- C the world needs a new Age of Enlightenment.
- D we have moved beyond the ideas of the past.

Q 8 **Alcorn expects that with a smaller world population**

- A economic problems will arise.
- B depressions will increase.
- C property will become cheaper.
- D there will be fewer social benefits.

---

<sup>42</sup> Alcorn, Pete. 2009. "The World in 2200". In *TED Ideas worth spreading*. [https://www.ted.com/talks/pete\\_alcorn\\_s\\_vision\\_of\\_a\\_better\\_world#t-209334](https://www.ted.com/talks/pete_alcorn_s_vision_of_a_better_world#t-209334) (8 Apr. 2016).

- Q 9 **Another effect will be that**
- A people will earn more money.
  - B the poor will need to work longer.
  - C there will be strikes for higher pay.
  - D the workers will have harder lives.
- Q 10 **In Europe a decline in population once led to**
- A higher birth rates.
  - B a loss of know-how and knowledge.
  - C the rise of a new social group.
  - D large-scale land sales.
- Q 11 **People have always tended to believe that**
- A life would be easier in the future.
  - B the past was better than the present.
  - C the Renaissance was a period of romance.
  - D in two generations everything would get worse.
- Q 12 **According to Alcorn, periods of change are**
- A optimistic times.
  - B fast-paced.
  - C interesting.
  - D unsafe.
- Q 13 **In the 22<sup>nd</sup> century, people will**
- A be heading towards another disaster.
  - B live up to 150 years.
  - C long for the good old times.
  - D have an optimistic outlook.

### Task 3

\_\_\_ / 9 P.

You are going to listen to a speech about understanding connections and relationships. First you will have 45 seconds to study the task below, then you will hear the recording twice. While listening, choose the correct answer (A, B, C or D) for each question (14-22). Put a  in the correct box. The first one (0) has been done for you.

After the second listening, you will have 45 seconds to check your answers.



## Connecting the Dots

Q 0 **After dropping out of college, Steve Jobs**

- A decided to stay on for another six months.
- B lost his interest in an academic career.
- C resolved to interrupt his studies for eighteen months.
- D stayed on the campus for eighteen months.

Q 14 **Steve's natural mother**

- A was too young to look after him.
- B was determined that he should be adopted.
- C couldn't afford to bring him up.
- D had been brought up by adoptive parents.

---

<sup>43</sup> Jobs, Steve. 2011. "Connecting the Dots". In *YouTube*. <https://www.youtube.com/watch?v=sr07uR75Qk0> (8 Apr. 2016).

- Q 15 **The couple that should have adopted Steve originally**
- A were both lawyers.
  - B got the chance to adopt a baby girl.
  - C rejected him because of his sex.
  - D lost interest in an adoption.
- Q 16 **Steve's biological mother only agreed to the adoption if**
- A he was guaranteed a college education.
  - B his adoptive parents were university graduates.
  - C his adoptive parents would send him to an expensive college.
  - D they promised to give him a proper start in life.
- Q 17 **After six months at college, Steve**
- A had wasted all his parents' savings.
  - B was dissatisfied with college life.
  - C couldn't afford the university fees anymore.
  - D doubted that he would benefit from college education.
- Q 18 **Dropping out provided a chance to**
- A regain self-confidence.
  - B rethink his goals in life.
  - C take up more attractive courses.
  - D pay back his debts to his parents.
- Q 19 **Steve's life was made difficult by the fact that he**
- A only got decent food once a week.
  - B had to do terrible jobs.
  - C had to let friends share his room.
  - D had to endure growing pressure from his family.
- Q 20 **Back then, Reed College was famous for offering**
- A special courses on how to design posters.
  - B the most excellent calligraphy tuition nationwide.
  - C free calligraphy instruction.
  - D outstanding courses on graphic design.

**Q 21 While learning these skills, Steve was certain they would**

- A never be of much use in his life.
- B be valuable for his professional life.
- C be of historical importance.
- D help him to build a career as a computer designer.

**Q 22 After some time Steve discovered that things in life**

- A only happen coincidentally.
- B only make sense if planned carefully.
- C always make an unexpected turn.
- D only make sense in retrospection.

# Questionnaire

## VISUAL CUES IN EVERYDAY COMMUNICATION

Please indicate on a scale of 1-4 (with 1 meaning that you **disagree strongly** and 4 meaning that you **agree strongly**) to what extent you agree with the following statements by circling the appropriate number.

<b>disagree strongly</b>	<b>disagree somewhat</b>	<b>agree somewhat</b>	<b>agree strongly</b>
1	2	3	4

1. I think that it is easier to comprehend spoken English if I am able to see the speaker's facial expressions.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
2. I think that it is easier to understand spoken English if I am able to see the speaker's lip movements.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
3. I think that it is easier to understand spoken English if I am able to see the speaker's gestures.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
4. I think it is easier to understand spoken English if I am able to see the context of the speaking situation.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
5. I think that it is easier to understand English on TV rather than English on the radio.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
6. In a face-to-face conversation, I observe my conversation partner's facial expressions.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
7. In a face-to-face conversation, I observe my conversation partner's lip movements.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
8. In a face-to-face conversation, I observe my conversation partner's gestures.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
9. The following activities are very helpful in improving my English listening skills.				
a. watching TV	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
b. watching videos on the internet	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
c. listening to the radio	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
d. listening to music	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>

e. having conversations in English	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
f. attending to my English lessons	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
g. The following activities are also helpful in improving my English listening skills: _____				

**YOUR OPINION ON THE INCLUSION OF VISUAL INPUT IN A LISTENING TEST**

10. I think that seeing the speaker would have helped me to understand the listening texts.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
11. I think that seeing the context of the speaking situation would have helped me to understand the listening texts.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
12. I would prefer a video-listening test where I could see the speaker(s) and the visual context rather than a conventional audio-listening test.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>

**PRIOR KNOWLEDGE AND ADDITIONAL COMMENTS**

**13. Did you take any of the listening tasks before?**

a. First listening task: **A new generation of comic books**

Yes       No

b. Second listening task: **On the world's future**

Yes       No

c. Third listening task: **Connecting the dots**

Yes       No

**14. Do you have any additional comments concerning this test?** (for example about sound quality, difficulty, time to answer the questions,...)

**THANK YOU!!**



---

<sup>44</sup> Bruggmann, Patrick. *Young world 3 – unit 2: adjectives* <http://www.schule.rorschach.ch/potatoe/5EN002/5EN002.htm> (12 Apr. 2016).

## Correction key: *Audio-only and audio-visual listening test*

### Task 1: A new generation of comic books<sup>45</sup>

#### Key

	accepted	not accepted
Q 0	<i>adults and children</i>	
Q 1	<p><b>digital comic books</b>                      digital comicbooks                      digital comics                      digital versions of comics</p> <p><i>both concepts “digital” + “comics” have to be there for the answer to be accepted</i></p>	catching up worldwide comic book business comic books digital digital devices digital comic books major transitions media merger transmission online stores something new something new in business the comic business the comics the iPad traditional comic books transition
Q 2	<p><b>are about the same</b>                      are approximately the same                      are equal                      are equally expensive                      are nearly the same                      are not different                      are the same                      are the same prices                      cost about the same                      have same price</p>	are different are exploding are lower are very expensive are very high have changed is booming is increasing rise up the same ( <i>verb is needed</i> )
Q 3	<p><b>fits the digital screen</b>                      fits all digital screens                      fits digital screens                      fits in digital screens                      fits to digital screens</p>	a digital screen a digital screen a special screen adjust on digital device can fit the screen create a digital site created the digital comic design digital screens designs digital comic books designs screens fit the screen fit the screen fits all screens fits digital stream fits the original stream focused on comics

<sup>45</sup> The correction key and the corresponding justifications for task 1 can be found on bifie’s homepage under the following link: <https://www.bifie.at/node/2800> (bifie 2015a).

		<p>has a digital screen  is called digital screens  is digital on screen  is digital screen  is experimenting  is on screen  is very popular  produces screen-fitting comics  provides comics on screens  sound emotions</p>
Q 4	<p><b>get extra historical information</b>  feature sound and motion  get additional information  get extra history  get extra information  get historical background  get historical background information  get historical information  get more information  get other information  have sound and motion  have sound information  hear extra sounds  hear sounds while reading  incorporate sound and motion  learn more about history  listen to comics  listen to sounds  provide sounds and motions  to have sound effects  use sound</p>	<p>add sound emotion  add sounds (<i>they are already there</i>)  feature sound and picture  fit sound emotion  have digital comics  hear the comics  learn  learn more  produce sound  reach readers worldwide  see sound and motion  to sound and motion  write fiction work</p>
Q 5	<p><b>something other than violence</b>  non-violent comics  other things than violence  other things without violence  something different from violence  something different than violence  something else than violence  something other than violents</p>	<p>a lot more  extra information  guns  guns and fights  his works  less violence (<i>they want something other than violence, not just less violence</i>)  more  more than just violence (<i>this suggests they do want violence, plus something else</i>)  more than violence (<i>this suggests they do want violence, plus something else</i>)  more violence  not only violence (<i>this suggests they do want violence, plus something else</i>)  no violence  non-violent things  other than violence  other than violent  others as violence  others than violence  outside the world  pictures</p>

		something after violence something extravagant something other with violence the virtual world violence
Q 6	<b>is helping his business</b> helps his business helps his business growing <i>various spellings of business were also accepted</i>	are too violent are video games has a big audience helps growing market helps his site is getting more popular is growing his business is holding his business is less violent is solving copies is successful of website reaches people worldwide went on works out well
Q 7	<b>new generations of readers</b> a new generation a new reader generation a new reader's generation a new readership attract a new generation new comic readers new generations new readers newcomers	all people all readers boost the business both of them both worlds buy both of them real worlds the digital format to be a new generation virtual and real world

### **Task 2: On the world's future**<sup>46</sup>

#### **Key**

Q 0	Q 8	Q 9	Q 10	Q 11	Q 12	Q 13
B	C	A	C	B	D	D

### **Task 3: Connecting the Dots**<sup>47</sup>

#### **Key**

Q 0	Q 14	Q 15	Q 16	Q 17	Q 18	Q 19	Q 20	Q 21	Q 22
D	B	C	A	D	C	A	B	A	D

<sup>46</sup> The correction key for task 2 can be found on bifie's homepage under the following link: <https://www.bifie.at/node/2837> (bifie 2015b).

<sup>47</sup> The correction key and the corresponding justifications for task 3 can be found bifie's homepage under the following link: <https://www.bifie.at/node/1830> (bifie 2012).

## Test specifications: Audio-only listening test<sup>48</sup>

**General statement of purpose:** This audio-only listening test is specifically designed for the purpose of this diploma thesis, which investigates the effects of the presence or absence of visual cues in listening comprehension tests by comparing an audio-only test with an audio-visual test that differ from each other only in the mode of presentation. This test has the purpose of measuring the listening comprehension skills of one of the three groups of 7<sup>th</sup> grade students, i.e. the control group, in an audio-only listening condition. The results of this test will serve as a basis for comparison with the results of the audio-visual listening comprehension test.

**Test category:** specific purpose test

**Description of test takers:** The test takers are 18 Austrian pupils attending the 7<sup>th</sup> grade of an upper secondary school in Vöcklabruck in Upper Austria. They are 16 to 17 years old and have German as their mother tongue. Fourteen of them are female and 4 are male.

**Test level:** B2

**Target language use situation:** The test takers can comprehend authentic and unmodified native English if the speaker or speakers use a standard dialect of English.

**Mode of presentation:** audio only

### Test construct<sup>49</sup>

#### General purpose

- a) listening for specific information
- b) listening for important details
- c) search listening
- d) listening to infer (propositional) meaning
- e) listening to deduce the meaning of a word

#### Specific purpose

1. Can understand the main ideas of propositionally and linguistically complex speech on both concrete and abstract topics delivered in a standard dialect [...] (Council of Europe 2001: 66).
2. Can follow extended speech and complex lines of argument provided the topic is reasonably familiar and the direction of the talk is sign posted by explicit markers (Council of Europe 2001: 66).
3. Can follow the essentials of lectures, talks and reports and other forms of academic / professional presentation which are propositionally and linguistically complex (Council of Europe 2001: 67).
4. Can understand most radio documentaries and most other recorded or broadcast material delivered in standard dialect and can identify the speaker's mood, tone etc. (Council of Europe 2001: 68).
5. Can use a variety of strategies to achieve comprehension including listening for main points and checking comprehension by using contextual clues (Council of Europe 2001: 72).
6. Can understand standard spoken language, live or broadcast, on both familiar and unfamiliar topics normally encountered in personal, social, academic or vocational life.

<sup>48</sup> These test specifications are modelled on the test specification operationalised in Matura listening exams (Universität Innsbruck 2012) and the guidelines for devising test specifications by Alderson, Clapham and Wall (1995: 11-15) and Bachman and Palmer (1996: 48-49).

<sup>49</sup> This listening comprehension construct is a modified version of the construct definition of listening comprehension developed on behalf of bifie (Universität Innsbruck 2012), which used the descriptors for language learners contained in the CEFR (Council of Europe 2001).

Only extreme background noise, inadequate discourse structure and/or idiomatic usage influences the ability to understand (Council of Europe 2001: 66).

### Structure of the test

- *Number of listening tasks:* 3
- *Number of test items altogether:* 20-25
- *Number of test items per task:* 6-9

### Time allotment

- *Time for the entire test:* approximately 30 minutes
- *Time for each task:* approximately 10 minutes (including the time for playing the recordings twice and the pauses before, after and between the tasks)
- Recordings are played twice. Students have 45 seconds time before the first recording to study the tasks, 20 seconds between the tasks to complete the preceding task and 45 seconds after hearing the last recording a second time to complete the tasks and check the answers.

### Test methods

- multiple choice
- limited response (maximum of 4 words)

### Input

- *text types:* spoken recordings of monologues, dialogues, conversations – e.g. interviews, documentaries, reports, speeches, talks, etc.
- *authenticity of input:* authentic, non-scripted input or partly semi-scripted input (for example, in a speech or a talk)
- *discourse type:* descriptive, narrative, persuasive
- *nature of content:* concrete
- *speaker characteristics:* male and female speakers of different ages using standard dialects of English
- *source of listening texts and tasks:* bifie (2016)

### Test rubrics

- *instructions:* are provided in the target-language in written and spoken form for each listening comprehension task.
- one example answer will be given

### Assessment criteria

- *objective testing*
- *weighing per item:* multiple choice - 1 point per test item  
limited response - 1 point if the answer is appropriate and grammatically accurate  
0,5 points if the answer is appropriate, however, grammatically incorrect in a way that is not included in the provided correction key  
spelling mistakes are not counted as mistakes
- correction keys are available at the homepage of bifie (2016)

**Appendix 6: Audio-visual listening test (including AV student questionnaire and test specifications)**

# **Audio-visual listening test**

First name: \_\_\_\_\_

Second name: \_\_\_\_\_

English teacher: \_\_\_\_\_

Points: \_\_\_\_\_ out of **22**

## **Instructions**

1. This test contains 3 tasks and 22 questions.
2. Write all your answers in this test booklet.
3. Please fill in the questionnaire after the listening test.

Test sources: <https://www.bifie.at/node/2800>, <https://www.bifie.at/node/2837>,  
<https://www.bifie.at/node/1830>

Video sources: <https://www.youtube.com/watch?v=uSJW-MC7jDM>,  
[https://www.ted.com/talks/pete\\_alcorn\\_s\\_vision\\_of\\_a\\_better\\_world#t-209334](https://www.ted.com/talks/pete_alcorn_s_vision_of_a_better_world#t-209334),  
<https://www.youtube.com/watch?v=sr07uR75Qk0>

**Task 1**

\_\_\_ / 7 P.

You are going to listen to a recording about a new era in comic book design. First you will have 45 seconds to study the task below, then you will hear the recording twice. While listening, complete the sentences (1-7) using a maximum of four words. Write your answers in the spaces provided on the answer sheet. The first one (0) has been done for you.

After the second listening, you will have 45 seconds to check your answers.

**A new generation of comic books**

Q 0	For decades Golden Apple's comic books have been bought by both _____.
Q 1	One recent change has been the introduction of _____.
Q 2	The costs of digital and printed comics _____.
Q 3	Mark Wait works in a format which _____.
Q 4	With Operation Ajax, it is possible for the reader to _____.
Q 5	Comic book readers today want to see _____.
Q 6	Luckily for Ryan Liewbowitz, the new development _____.
Q 7	The new development is expected to appeal to _____.

## A new generation of comic books

Q 0	<i>adults and children</i>
Q 1	
Q 2	
Q 3	
Q 4	
Q 5	
Q 6	
Q 7	

von der  
Lehrperson  
auszufüllen

richtig falsch

## Task 2

\_\_\_ / 6 P.

You are going to listen to a recording about Pete Alcorn comparing his view of the world with that of Thomas Robert Malthus. First you will have 45 seconds to study the task below, then you will hear the recording twice. While listening, choose the correct answer (A, B, C or D) for each question (8-13). Put a ☒ in the correct box. The first one (0) has been done for you.

After the second listening, you will have 45 seconds to check your answers.



### On the world's future

Q 0 **Malthus' idea was that**

- A we need more than one mental model of the world.
- B the rising number of people on Earth is a disaster.
- C the world needs a new Age of Enlightenment.
- D we have moved beyond the ideas of the past.

Q 8 **Alcorn expects that with a smaller world population**

- A economic problems will arise.
- B depressions will increase.
- C property will become cheaper.
- D there will be fewer social benefits.

---

<sup>50</sup> Alcorn, Pete. 2009. "The World in 2200". In *TED Ideas worth spreading*. [https://www.ted.com/talks/pete\\_alcorn\\_s\\_vision\\_of\\_a\\_better\\_world#t-209334](https://www.ted.com/talks/pete_alcorn_s_vision_of_a_better_world#t-209334) (8 Apr. 2016).

- Q 9 Another effect will be that**
- A people will earn more money.
  - B the poor will need to work longer.
  - C there will be strikes for higher pay.
  - D the workers will have harder lives.
- Q 10 In Europe a decline in population once led to**
- A higher birth rates.
  - B a loss of know-how and knowledge.
  - C the rise of a new social group.
  - D large-scale land sales.
- Q 11 People have always tended to believe that**
- A life would be easier in the future.
  - B the past was better than the present.
  - C the Renaissance was a period of romance.
  - D in two generations everything would get worse.
- Q 12 According to Alcorn, periods of change are**
- A optimistic times.
  - B fast-paced.
  - C interesting.
  - D unsafe.
- Q 13 In the 22<sup>nd</sup> century, people will**
- A be heading towards another disaster.
  - B live up to 150 years.
  - C long for the good old times.
  - D have an optimistic outlook.

### Task 3

\_\_\_ / 9 P.

You are going to listen to a speech about understanding connections and relationships. First you will have 45 seconds to study the task below, then you will hear the recording twice. While listening, choose the correct answer (A, B, C or D) for each question (14-22). Put a  in the correct box. The first one (0) has been done for you.

After the second listening, you will have 45 seconds to check your answers.



## Connecting the Dots

Q 0 **After dropping out of college, Steve Jobs**

- A decided to stay on for another six months.
- B lost his interest in an academic career.
- C resolved to interrupt his studies for eighteen months.
- D stayed on the campus for eighteen months.

Q 14 **Steve's natural mother**

- A was too young to look after him.
- B was determined that he should be adopted.
- C couldn't afford to bring him up.
- D had been brought up by adoptive parents.

---

<sup>51</sup> Jobs, Steve. 2011. "Connecting the Dots". In *YouTube*. <https://www.youtube.com/watch?v=sr07uR75Qk0> (8 Apr. 2016).

- Q 15 **The couple that should have adopted Steve originally**
- A were both lawyers.
  - B got the chance to adopt a baby girl.
  - C rejected him because of his sex.
  - D lost interest in an adoption.
- Q 16 **Steve's biological mother only agreed to the adoption if**
- A he was guaranteed a college education.
  - B his adoptive parents were university graduates.
  - C his adoptive parents would send him to an expensive college.
  - D they promised to give him a proper start in life.
- Q 17 **After six months at college, Steve**
- A had wasted all his parents' savings.
  - B was dissatisfied with college life.
  - C couldn't afford the university fees anymore.
  - D doubted that he would benefit from college education.
- Q 18 **Dropping out provided a chance to**
- A regain self-confidence.
  - B rethink his goals in life.
  - C take up more attractive courses.
  - D pay back his debts to his parents.
- Q 19 **Steve's life was made difficult by the fact that he**
- A only got decent food once a week.
  - B had to do terrible jobs.
  - C had to let friends share his room.
  - D had to endure growing pressure from his family.
- Q 20 **Back then, Reed College was famous for offering**
- A special courses on how to design posters.
  - B the most excellent calligraphy tuition nationwide.
  - C free calligraphy instruction.
  - D outstanding courses on graphic design.

**Q 21 While learning these skills, Steve was certain they would**

- A never be of much use in his life.
- B be valuable for his professional life.
- C be of historical importance.
- D help him to build a career as a computer designer.

**Q 22 After some time Steve discovered that things in life**

- A only happen coincidentally.
- B only make sense if planned carefully.
- C always make an unexpected turn.
- D only make sense in retrospection.

# Questionnaire

## VISUAL CUES IN EVERYDAY COMMUNICATION

Please indicate on a scale of 1-4 (with 1 meaning that you **disagree strongly** and 4 meaning that you **agree strongly**) to what extent you agree with the following statements by circling the appropriate number.

<b>disagree strongly</b>	<b>disagree somewhat</b>	<b>agree somewhat</b>	<b>agree strongly</b>
1	2	3	4

1. I think that it is easier to comprehend spoken English if I am able to see the speaker's facial expressions.	1	2	3	4
2. I think that it is easier to understand spoken English if I am able to see the speaker's lip movements.	1	2	3	4
3. I think that it is easier to understand spoken English if I am able to see the speaker's gestures.	1	2	3	4
4. I think it is easier to understand spoken English if I am able to see the context of the speaking situation.	1	2	3	4
5. I think that it is easier to understand English on TV rather than English on the radio.	1	2	3	4
6. In a face-to-face conversation, I observe my conversation partner's facial expressions.	1	2	3	4
7. In a face-to-face conversation, I observe my conversation partner's lip movements.	1	2	3	4
8. In a face-to-face conversation, I observe my conversation partner's gestures.	1	2	3	4
9. The following activities are very helpful in improving my English listening skills.				
a. watching TV	1	2	3	4
b. watching videos on the internet	1	2	3	4
c. listening to the radio	1	2	3	4
d. listening to music	1	2	3	4

e. having conversations in English	1	2	3	4
f. attending to my English lessons	1	2	3	4
g. The following activities are also helpful in improving my English listening skills: _____				

### ATTENTION PAID TO VISUAL CUES IN THE VIDEO-LISTENING TEST

#### 10. How often did you watch the video in the test?

- very often     hardly ever  
 often     never

#### 11. In the first video-listening (“*A new generation of comic books*”), how much attention did you pay to ...

	a lot of attention	some attention	hardly any attention	no atten- tion at all
a. the speakers’ facial expressions?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. the speakers’ lip movements?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c. the speakers’ gestures?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d. the visual contexts of the speaking situations (e.g. the comic book store, the digital comics showed, ...)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

#### 12. In the second video-listening (“*On the world’s future*”), how much attention did you pay to ...

a. the speaker’s facial expressions?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. the speaker’s lip movements?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c. the speaker’s gestures?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d. the visual contexts of the speaking situations (e.g. the slides for the talk,...)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**13. In the third video-listening (“Connecting the dots”), how much attention did you pay to ...**

- |   |                          |                          |                          |                          |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| a. the speaker’s facial expressions?  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| b. the speaker’s lip movements?   | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| c. the visual contexts of the speaking situations (e.g. the context of the speech,...)? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

**YOUR OPINION ON THE INCLUSION OF VISUAL INPUT IN A LISTENING TEST**

Please indicate on a scale of 1-4 (with 1 meaning that you **disagree strongly** and 4 meaning that you **agree strongly**) to which extent you agree with the following statements by circling the appropriate number.

<b>disagree strongly</b>	<b>disagree somewhat</b>	<b>agree somewhat</b>	<b>agree strongly</b>
1	2	3	4

14. In general, the visuals in this video-listening test supported my comprehension of the spoken text and helped me answer the test questions.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
15. I think that the following elements of non-verbal communication supported my comprehension of the spoken text and helped me answer the test questions:				
• facial expressions	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
• lip movements	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
• gestures	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
• visual contexts of the speaking situations	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
16. I think that seeing the speakers in this video-listening test distracted me and had a negative effect on my comprehension of the spoken text.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
17. I think that seeing the visual contexts of the speaking situations distracted me and had a negative effect on my comprehension of the spoken text.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>

18. I prefer a video-listening test over a conventional audio-listening test.	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
---	----------	----------	----------	----------

**PRIOR KNOWLEDGE AND ADDITIONAL COMMENTS**

**19. Did you take any of the listening tasks before?**

a. First listening task: **A new generation of comic books**

Yes       No

b. Second listening task: **On the world’s future**

Yes       No

c. Third listening task: **Connecting the dots**

Yes       No

**20. Do you have any additional comments concerning this test and the use of videos in listening tests?** (for example about sound quality, difficulty, time to answer the questions,...)

52

**THANK YOU!!**



<sup>52</sup> Bruggmann, Patrick. *Young world 3 – unit 2: adjectives* <http://www.schule.rorschach.ch/potatoe/5EN002/5EN002.htm> (12 Apr. 2016).

## Test specifications: Audio-visual listening test<sup>53</sup>

**General statement of purpose:** This audio-visual listening test is specifically designed for the purpose of this diploma thesis, which investigates the effects of the presence or absence of visual cues in listening comprehension tests by comparing an audio-only test with an audio-visual test that differ from each other only in the mode of presentation. This test has the purpose of measuring the listening comprehension skills of two of the three groups of 7<sup>th</sup> grade students, i.e. the experimental groups, in the presence of visual cues provided by videos accompanying the listening test. The results of this test will serve as a basis for comparison with the results of the audio-audio listening comprehension test.

**Test category:** specific purpose test

**Description of test takers:** The test takers are 32 Austrian pupils attending the 7<sup>th</sup> grade of an upper secondary school in Vöcklabruck in Upper Austria. They are 16 to 17 years old and have German as their mother tongue. Twenty two of them are female and 10 are male.

**Test level:** B2

**Target language use situation:** The test takers can comprehend authentic and unmodified native English if the speaker or speakers use a standard dialect of English.

**Mode of presentation:** audio-visual

### Test construct<sup>54</sup>

#### General purpose

- a) listening for specific information
- b) listening for important details
- c) search listening
- d) listening to infer (propositional) meaning
- e) listening to deduce the meaning of a word

#### Specific purpose

1. Can understand the main ideas of propositionally and linguistically complex speech on both concrete and abstract topics delivered in a standard dialect [...] (Council of Europe 2001: 66).
2. Can follow extended speech and complex lines of argument provided the topic is reasonably familiar and the direction of the talk is sign posted by explicit markers (Council of Europe 2001: 66).
3. Can follow the essentials of lectures, talks and reports and other forms of academic / professional presentation which are propositionally and linguistically complex (Council of Europe 2001: 67).
4. Can understand most radio documentaries and most other recorded or broadcast material delivered in standard dialect and can identify the speaker's mood, tone etc. (Council of Europe 2001: 68).
5. Can use a variety of strategies to achieve comprehension including listening for main points and checking comprehension by using contextual clues (Council of Europe 2001: 68).

<sup>53</sup> These test specifications are modelled on the test specification operationalised in Matura listening exams (Universität Innsbruck 2012) and the guidelines for devising test specifications by Alderson, Clapham and Wall (1995: 11-15) and Bachman and Palmer (1996: 48-49).

<sup>54</sup> This construct definition of listening comprehension is based on the 'modified B2-level listening comprehension construct for video-mediated listening test' (see Section 5.3). This modified construct was based on the construct definition of listening comprehension developed on behalf of bifie (Universität Innsbruck 2012), which used the descriptors for language learners contained in the CEFR (Council of Europe 2001).

72).

6. Can understand standard spoken language, live or broadcast, on both familiar and unfamiliar topics normally encountered in personal, social, academic or vocational life. Only extreme background noise, inadequate discourse structure and/or idiomatic usage influences the ability to understand (Council of Europe 2001: 66).

**Visual competencies involved in B2-level listening comprehension in audio-visual listening conditions**

- I. Can utilise information derived from the gestures of a speaker to support their comprehension of the spoken message.
- II. Can utilise information derived from the facial expressions of a speaker to support their comprehension of the spoken message.
- III. Can utilise information derived from the lip movements of a speaker to support their comprehension of the spoken message.
- IV. Can utilise information derived from the visual context of the speaking situation to support their comprehension of the spoken message.
- V. Can gain information about the intentions, emotions and attitudes of a speaker from his or her gestures and facial expressions.
- VI. Can utilise information derived from lip movements to assist speech perception.
- VII. Can gain information about the roles of different speakers in a conversation, their interpersonal relationship and the discourse type of the spoken text from the visual context of the speaking situation.
- VIII. Can simultaneously process and utilise auditory and visual information and answer test questions.

**Structure of the test**

- *Number of listening tasks:* 3
- *Number of test items altogether:* 20-25
- *Number of test items per task:* 6-9

**Time allotment**

- *Time for the entire test:* approximately 30 minutes
- *Time for each task:* approximately 10 minutes (including the time for playing the videos twice and the pauses before, after and between the tasks)
- Videos are played twice. Students have 45 seconds time before the first video to study the tasks, 20 seconds between the tasks to complete the preceding task and 45 seconds after viewing the last video a second time to complete the tasks and check the answers.

**Test methods**

- multiple choice
- limited response (maximum of 4 words)

**Input**

- *text types:* video texts  
videos showing monologues, dialogues, conversations – e.g. interviews, documentaries, reports, speeches, talks, etc.
- *authenticity of input:* authentic videos, non-scripted input or partly semi-scripted input (for example, in a speech or a talk)
- *discourse type:* descriptive, narrative, persuasive
- *nature of content:* concrete
- *speaker characteristics:* male and female speakers of different ages using standard dialects of English
- *source of video texts:* internet (YouTube, TED: Ideas worth spreading,...)
- *sources of listening tasks:* bifie (2016)

**Test rubrics**

- *instructions*: are provided in the target-language in written and spoken form for each listening comprehension task.
- one example answer will be given

**Assessment criteria**

- *objective testing*
- *weighing per item*: multiple choice - 1 point per test item  
limited response - 1 point if the answer is appropriate and grammatically accurate  
0,5 points if the answer is appropriate, however, grammatically incorrect in a way that is not included in the provided correction key  
spelling mistakes are not counted as mistakes
- correction keys are available at the homepage of bifie (2016)

## Appendix 7: Calculations

### 1. Frequency statistics

**Table 29 AV/AO Qu. item 9: “The following activities are very helpful in improving my English listening skills.”**

activities	disagree strongly	disagree somewhat	agree somewhat	agree strongly
watching TV	10.6 <sup>a</sup>	8.5	25.5	55.3
watching videos on the internet	2.1	8.5	38.3	51.1
listening to the radio	8.5	36.2	46.8	14.9
listening to music	10.6	27.7	46.8	14.9
having conversations in English	2.1	10.6	25.5	61.7
attending to your English lessons		23.4	53.2	23.4

<sup>a</sup> all numbers mean percentage of test takers

### 2. T-tests

**Table 30 AV/AO test: Comparison of performance of control vs. experimental groups on individual test tasks**

Independent samples t-test		Levene's test for equality of variances		t-test for equality of means	
		F	sig.	t	sig. (2-tailed)
task 1	equal variances assumed	0.225	0.638	-1.395	0.170
task 2	equal variances assumed	0.271	0.605	-0.793	0.432
task 3	equal variances assumed	0.286	0.596	-0.024	0.981

**Table 31 AV/AO test: Comparison of performance of control vs. experimental groups on individual test items**

Independent samples t-test		Levene's test for equality of variances		t-test for equality of means	
		F	sig.	t	sig. (2-tailed)
item 1	equal variances assumed	0.016	0.899	-0.064	0.949
	equal variances not assumed			-0.064	0.949
item 4	equal variances assumed	14.706	0.000	-1.297	0.201
	equal variances not assumed			-1.112	0.278
item 5	equal variances assumed	2.464	0.123	-2.707	0.010
	equal variances not assumed			-2.713	0.010
item 8	equal variances assumed	0.245	0.623	0.251	0.803
	equal variances not assumed			0.250	0.804
item 9	equal variances assumed	9.105	0.004	-1.339	0.187
	equal variances not assumed			-1.448	0.155
item 10	equal variances assumed	0.001	0.978	-0.014	0.989
	equal variances not assumed			-0.014	0.989
item 11	equal variances assumed	4.836	0.033	-1.102	0.276
	equal variances not assumed			-1.023	0.315
item 14	equal variances assumed	0.952	0.334	-0.470	0.640
	equal variances not assumed			-0.476	0.637
item 15	equal variances assumed	0.627	0.433	-0.411	0.683
	equal variances not assumed			-0.404	0.688
item 16	equal variances assumed	3.644	0.063	-1.065	0.293
	equal variances not assumed			-1.030	0.311
item 17	equal variances assumed	3.646	0.063	1.178	0.245
	equal variances not assumed			1.146	0.260
item 18	equal variances assumed	0.664	0.419	-0.414	0.681
	equal variances not assumed			-0.403	0.689
item 19	equal variances assumed	5.768	0.021	1.114	0.271
	equal variances not assumed			1.153	0.256
item 20	equal variances assumed	3.644	0.063	-1.065	0.293
	equal variances not assumed			-1.030	0.311
item 21	equal variances assumed	0.100	0.753	-0.161	0.873
	equal variances not assumed			-0.160	0.874
item 22	equal variances assumed	10.459	0.002	1.417	0.163
	equal variances not assumed			1.600	0.117

**Table 32 Paired samples t-test: Comparison of performance of experimental test group split by learning style on test A vs. AV listening test<sup>a</sup>**

Learning style	test pairs	mean	t	sig. (2-tailed)
mixed learning style	pair 1 results test A in % - results AV test in %	-11.41	-1.822	0.088
bimodal visual/auditive	pair 1 results test A in % - results AV test in %	8.68	1.241	0.303
predominantly visual	pair 1 results test A in % - results AV test in %	13.86	1.401	0.296
predominantly auditive	pair 1 results test A in % - results AV test in %	-7.37	-1.400	0.395
predominantly tactile	pair 1 results test A in % - results AV test in %	16.05	1.034	0.489

<sup>a</sup> no statistics calculated for learning style bimodal auditive and tactile as no matching pairs were available

**Table 33 Paired samples t-test: Comparison of performance of test takers who differed in their opinion on AV Qu. item 14 on test A vs. AV listening test<sup>a</sup>**

AV Qu. item 14: abbr. visuals supported	test pairs	mean	t	sig. (2-tailed)
disagree somewhat	pair 1 results test A in % - results AV test in %	-1.05263	-0.171	0.875
agree somewhat	pair 1 results test A in % - results AV test in %	-1.31579	-0.230	0.821
agree strongly	pair 1 results test A in % - results AV test in %	-10.15038	-0.866	0.420

<sup>a</sup> no statistics calculated for answer disagree strongly as no matching pairs were available

**Table 34 Paired samples t-test: Comparison of performance of test takers who differed in their opinion on AV Qu. item 16 on test A vs. AV listening test<sup>a</sup>**

AV Qu. item 16: abbr. seeing speakers distracted	test pairs	mean	t	sig. (2-tailed)
disagree somewhat	pair 1 results test A in % - results AV test in %	-2.40132	-0.360	0.724
agree somewhat	pair 1 results test A in % - results AV test in %	-6.51316	-0.840	0.429
agree strongly	pair 1 results test A in % - results AV test in %	-1.97368	-0.383	0.727

<sup>a</sup> no statistics calculated for answer disagree strongly as no matching pairs were available

**Table 35 Paired samples t-test: Comparison of performance of test takers who differed in their opinion on AV Qu. item 17 on test A vs. AV listening test<sup>a</sup>**

AV Qu. item 17: abbr. seeing visual context distracted	test pairs	mean	t	sig. (2-tailed)
disagree somewhat	pair 1 results test A in % - results AV test in %	-4.32749	-0.706	0.490
agree somewhat	pair 1 results test A in % - results AV test in %	-11,31579	-1.957	0.108
agree strongly	pair 1 results test A in % - results AV test in %	5.26316	1.192	0.355

<sup>a</sup> no statistics calculated for answer disagree strongly as no matching pairs were available

### 3. ANOVA

<b>ANOVA 1</b>
<i>Comparison of performance of test takers with different learning styles on the baseline listening test and the AV listening test</i>

**Table 36 ANOVA 1: Test of homogeneity of variances**

	Levene Statistic	df1	df2	sig.
Test A – results in %	1.858	4	25	0.149
AV Test – results in %	1.572	4	22	0.217

**Table 37 ANOVA 1: One-way ANOVA**

		sum of squares	df	mean square	F	sig.
Test A – results in %	between groups	3488.89	5	697.78	1.915	0.128
	within groups	9111.11	25	364.44		
	total	12600.00	30			
AV Test – results in %	between groups	2509.23	5	501.85	1.753	0.165
	within groups	6299.63	22	286.35		
	total	8808.86	27			

<b>ANOVA 2</b>
<i>Comparison of performance on the AV listening test of test takers who differed in their opinion on AV Qu. item 14 (abbr.: visuals supported listening comprehension and helped answer test questions)</i>

**Table 38 ANOVA 2: Test of homogeneity of variances**

	Levene Statistic	df1	df2	sig.
AV Test – results in %	0.468	2	25	0.632

**Table 39 ANOVA 2: One-way ANOVA**

		sum of squares	df	mean square	F	sig.
AV Test – results in %	between groups	778.09	3	259.36	0.797	0.507
	within groups	8137.76	25	325.51		
	total	8915.85	28			

### ANOVA 3

*Comparison of performance on the AV listening test of test takers who differed in their opinion AV Qu. item 16 (abbr.: seeing the speaker was distracting and had negative effect on listening comprehension)*

**Table 40 ANOVA 3: Test of homogeneity of variances**

	Levene Statistic	df1	df2	sig.
AV Test – results in %	0.792	2	26	0.463

**Table 41 ANOVA 3: One-way ANOVA**

		sum of squares	df	mean square	F	sig.
AV Test – results in %	between groups	231.43	2	115.72	0.346	0.710
	within groups	8684.41	26	334.02		
	total	8915.85	28			

### ANOVA 4

*Comparison of performance on the AV listening test of test takers who differed in their opinion AV Qu. item 17 (abbr.: seeing the visual context was distracting and had negative effect on listening comprehension)*

**Table 42 ANOVA 4: Test of homogeneity of variances**

	Levene Statistic	df1	df2	sig.
AV Test – results in %	0.179	2	25	0.837

**Table 43 ANOVA 4: One-way ANOVA**

		sum of squares	df	mean square	F	sig.
AV Test – results in %	between groups	1318.29	3	439.43	1.446	0.253
	within groups	7597.56	25	303.90		
	total	8915.85	28			

#### 4. Correlations

**Table 44 Correlation: Listening ability and AV Qu. item 10 (“How often did you watch the video in the test?”)**

		listening ability	AV Qu. item 10
Spearman’s rho	listening ability	1.000	-0.105
		.	0.589
		50	29
AV Qu. item 10	AV Qu. item 10	-0.105	1.000
		0.589	.
		29	29

**Table 45 Correlation: Learning style and AV Qu. item 10 (“How often did you watch the video in the test?”)**

		learning style	AV Qu. item 10
Spearman’s rho	learning style	1.000	-0.079
		.	0.691
		29	28
AV Qu. item 10	AV Qu. item 10	-0.079	1.000
		0.691	.
		28	31

**Table 46 Correlation: Listening ability and AV Qu. item 14 (abbr.:  
Visuals supported listening comprehension and helped answer test questions)**

		listening ability	AV Qu. item 14
Spearman's rho	listening ability	correlation coefficient	1.000
		sig. (2-tailed)	-.319
		N	.092
AV Qu. item 14		50	29
		correlation coefficient	-0.319
		sig. (2-tailed)	1.000
	N	0.092	.
		29	29

**Table 47 Correlation: Listening ability and AV Qu. item 16 (abbr.:  
Seeing the speakers was distracting and had negative effect on listening comprehension)**

		listening ability	AV Qu. item 16
Spearman's rho	listening ability	correlation coefficient	1.000
		sig. (2-tailed)	.107
		N	.582
AV Qu. item 16		50	29
		correlation coefficient	0.107
		sig. (2-tailed)	1.000
	N	0.582	.
		29	29

**Table 48 Correlation: Listening ability and AV Qu. item 17 (abbr.:  
Seeing the visual context was distracting and had negative effect on listening comprehension)**

		listening ability	AV Qu. item 17
Spearman's rho	listening ability	correlation coefficient	1.000
		sig. (2-tailed)	.304
		N	.109
AV Qu. item 17		50	29
		correlation coefficient	0.304
		sig. (2-tailed)	1.000
	N	0.109	.
		29	29

**Table 49 Correlation: Learning style and AV Qu. item 14 (abbr.:  
Visuals supported listening comprehension and helped answer test questions)**

		learning style	AV Qu. item 14
learning style	correlation coefficient	1.000	-0.118
	sig. (2-tailed)	.	0.549
	N	31	28
Spearman's rho	correlation coefficient	-0.118	1.000
	AV Qu. item 14	sig. (2-tailed)	0.549
	N	28	29

**Table 50 Correlation: Learning style and AV Qu. item 16 (abbr.:  
Seeing the speakers was distracting and had negative effect on listening comprehension)**

		learning style	AV Qu. item 16
learning style	correlation coefficient	1.000	0.127
	sig. (2-tailed)	.	0.518
	N	31	28
Spearman's rho	correlation coefficient	0.127	1.000
	AV Qu. item 16	sig. (2-tailed)	0.518
	N	28	29

**Table 51 Correlation: Learning style and AV Qu. item 17 (abbr.:  
Seeing the visual context was distracting and had negative effect on listening comprehension)**

		listening ability	AV Qu. item 17
learning style	correlation coefficient	1.000	0.409*
	sig. (2-tailed)	.	0.031
	N	31	28
Spearman's rho	correlation coefficient	0.409*	1.000
	AV Qu. item 17	sig. (2-tailed)	0.031
	N	28	29

\*Correlation is significant at the 0.05 level (2-tailed).